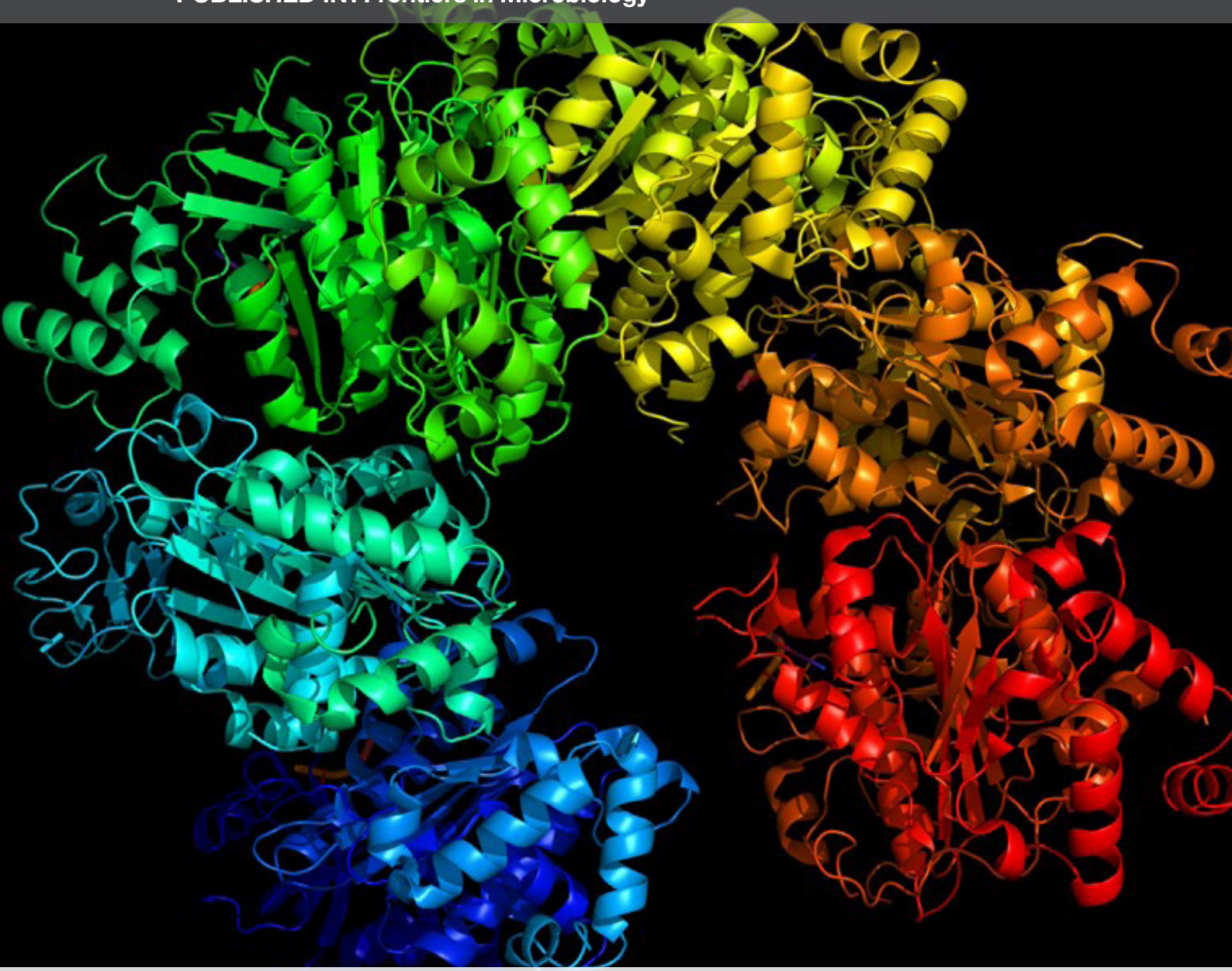


# PROTEIN SOLUBILITY AND AGGREGATION IN BACTERIA

EDITED BY : Salvador Ventura

PUBLISHED IN: Frontiers in Microbiology





# frontiers

## Frontiers Copyright Statement

© Copyright 2007-2016 Frontiers Media SA. All rights reserved.

All content included on this site, such as text, graphics, logos, button icons, images, video/audio clips, downloads, data compilations and software, is the property of or is licensed to Frontiers Media SA ("Frontiers") or its licensees and/or subcontractors. The copyright in the text of individual articles is the property of their respective authors, subject to a license granted to Frontiers.

The compilation of articles constituting this e-book, wherever published, as well as the compilation of all other content on this site, is the exclusive property of Frontiers. For the conditions for downloading and copying of e-books from Frontiers' website, please see the Terms for Website Use. If purchasing Frontiers e-books from other websites or sources, the conditions of the website concerned apply.

Images and graphics not forming part of user-contributed materials may not be downloaded or copied without permission.

Individual articles may be downloaded and reproduced in accordance with the principles of the CC-BY licence subject to any copyright or other notices. They may not be re-sold as an e-book.

As author or other contributor you grant a CC-BY licence to others to reproduce your articles, including any graphics and third-party materials supplied by you, in accordance with the Conditions for Website Use and subject to any copyright notices which you include in connection with your articles and materials.

All copyright, and all rights therein, are protected by national and international copyright laws.

The above represents a summary only. For the full conditions see the Conditions for Authors and the Conditions for Website Use.

ISSN 1664-8714

ISBN 978-2-88919-976-1

DOI 10.3389/978-2-88919-976-1

## About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

## Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

## Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.

Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view.

By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

## What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: [researchtopics@frontiersin.org](mailto:researchtopics@frontiersin.org)



# PROTEIN SOLUBILITY AND AGGREGATION IN BACTERIA

Topic Editor:

**Salvador Ventura**, Universitat Autònoma de Barcelona, Spain



Structure of Rho Termination Factor, a *Clostridium botulinum* protein with prion-like properties.

Image by Salvador Ventura

Proteins suffer many conformational changes and interactions through their life, from their synthesis at ribosomes to their controlled degradation. Only folded and soluble proteins are functional. Thus, protein folding and solubility are controlled genetically, transcriptionally, and at the protein sequence level. In addition, a well-conserved cellular machinery assists the folding of polypeptides to avoid misfolding and ensure the attainment of soluble and functional structures. When these redundant protective strategies are overcome, misfolded proteins are recruited into aggregates.

Recombinant protein production is an essential tool for the biotechnology industry and also supports expanding areas of basic and biomedical research, including structural genomics and proteomics. Although bacteria still represent a convenient production system, many recombinant polypeptides produced in prokaryotic hosts undergo irregular or incomplete folding processes that usually result in their accumulation as insoluble aggregates, narrowing thus the spectrum of

protein-based drugs that are available in the biotechnology market. In fact, the solubility of bacterially produced proteins is of major concern in production processes, and many orthogonal strategies have been exploited to try to increase soluble protein yields. Importantly, contrary to the usual assumption that the bacterial aggregates formed during protein production are totally inactive, the presence of a fraction of molecules in a native-like structure in these assemblies endorse them with a certain degree of biological activity, a property that is allowing the use of bacteria as factories to produce new functional materials and catalysts.

The protein embedded in intracellular bacterial deposits might display different conformations, but they are usually enriched in beta-sheet-rich assemblies resembling the amyloid fibrils

characteristic of several human neurodegenerative diseases. This makes bacterial cells simple, but biologically relevant model systems to address the mechanisms behind amyloid formation and the cellular impact of protein aggregates. Interestingly, bacteria also exploit the structural principles behind amyloid formation for functional purposes such as adhesion or cytotoxicity.

In the present research topic we collect papers addressing all the issues mentioned above from both the experimental and computational point of view.

**Citation:** Ventura, S., ed. (2016). Protein Solubility and Aggregation in Bacteria. Lausanne: Frontiers Media. doi: 10.3389/978-2-88919-976-1



# Table of Contents

- 06 Editorial: Protein Solubility and Aggregation in Bacteria**  
Salvador Ventura
- 09 Identification of multiple physicochemical and structural properties associated with soluble expression of eukaryotic proteins in cell-free bacterial extracts**  
Alexander A. Tokmakov
- 15 Large-scale analysis of macromolecular crowding effects on protein aggregation using a reconstituted cell-free translation system**  
Tatsuya Niwa, Ryota Sugimoto, Lisa Watanabe, Shugo Nakamura, Takuya Ueda and Hideki Taguchi
- 20 Localization of aggregating proteins in bacteria depends on the rate of addition**  
Karlton Scheu, Rakinder Gill, Saeed Saberi, Pablo Meyer and Eldon Emberly
- 25 Refolding and purification of recombinant L-asparaginase from inclusion bodies of E. coli into active tetrameric protein**  
Arun K. Upadhyay, Anupam Singh, K. J. Mukherjee and Amulya K. Panda
- 35 Multitasking SecB chaperones in bacteria**  
Ambre Sala, Patricia Bordes and Pierre Genevaux
- 47 Chaperonin GroEL/GroES Over-Expression Promotes Aminoglycoside Resistance and Reduces Drug Susceptibilities in Escherichia coli Following Exposure to Sublethal Aminoglycoside Doses**  
Lise Goltermann, Menachem V. Sarusie and Thomas Bentin
- 52 Fold modulating function: bacterial toxins to functional amyloids**  
Adnan K. Syed and Blaise R. Boles
- 61 Functional bacterial amyloid increases Pseudomonas biofilm hydrophobicity and stiffness**  
Guanghong Zeng, Brian S. Vad, Morten S. Dueholm, Gunna Christiansen, Martin Nilsson, Tim Tolker-Nielsen, Per H. Nielsen, Rikke L. Meyer and Daniel E. Otzen
- 75 Identification of Key Amino Acid Residues Modulating Intracellular and In vitro Microcin E492 Amyloid Formation**  
Paulina Aguilera, Andrés Marcoleta, Pablo Lobos-Ruiz, Rocío Arranz, José M. Valpuesta, Octavio Monasterio and Rosalba Lagos
- 91 Computational analysis of candidate prion-like proteins in bacteria and their role**  
Valentin Iglesias, Natalia S. de Groot and Salvador Ventura

**104    *The Rho Termination Factor of Clostridium botulinum Contains a Prion-Like Domain with a Highly Amyloidogenic Core***

Irantzu Pallarès, Valentin Iglesias and Salvador Ventura

**116    *Engineered bacterial hydrophobic oligopeptide repeats in a synthetic yeast prion, [REP-PSI<sup>+</sup>]***

Fátima Gasset-Rosa and Rafael Giraldo





# Editorial: Protein Solubility and Aggregation in Bacteria

Salvador Ventura \*

Departament de Bioquímica i Biologia Molecular, Institut de Biotecnologia i Biomedicina, Universitat Autònoma de Barcelona, Barcelona, Spain

**Keywords:** bacteria, protein folding, protein aggregation, protein expression, functional amyloid, bacterial chaperones, prion-like proteins

## The Editorial on the Research Topic

### Protein Solubility and Aggregation in Bacteria

For many years, the aggregation of proteins and polypeptides remained a neglected area of protein chemistry. It was only with the discovery that the insoluble deposits found in the organs and tissues of patients suffering from different diseases were enriched in a single, but different polypeptide, that the interest in understanding how and why these proteins aggregate arose (Fernández-Busquets et al., 2008). From that time on, the study of protein aggregation has evolved to become a key research topic, whose implications span disciplines like biochemistry, biomedicine, biotechnology, and nanotechnology.

The formation of insoluble protein aggregates is linked to more than 40 human diseases (Invernizzi et al., 2012). In all these disorders, the aggregated proteins assemble into a common  $\beta$ -sheet enriched supra-molecular structure, known as amyloid. A large number of higher eukaryotic biochemical pathways, from DNA replication to protein degradation, have been modeled first in prokaryotic organisms, providing important clues on the molecular basis of pathology (Dwyer et al., 2012). Scientists have long known that, very often, the heterologous expression of proteins in bacteria results in the accumulation of the target protein in insoluble deposits. However, only recently a number of laboratories have dared to exploit this well-characterized phenomenon to dissect the determinants of protein aggregation into amyloid assemblies (Villar-Piqué and Ventura, 2012). Bacteria possess an intracellular environment that, while differing from that of eukaryotic cells, is physiologically more relevant than a test tube. Accordingly, they constitute privileged model systems to understand the mechanisms behind amyloid assembly and the cellular fitness cost associated with the formation of these aggregates (Navarro et al., 2014), but also to screen for effective modulators of amyloid aggregation with potential therapeutic applications (Villar-Piqué et al., 2012).

Protein aggregation constitutes a major bottleneck in the biotechnological production of protein based therapeutics. Consequently, a large effort has been devoted to the development of orthogonal strategies to increase soluble protein yields, including *in vitro* protein synthesis (Ventura, 2005). Bacterial extracts have become a convenient means to produce recombinant proteins *in vitro*, because they contain all the machinery required for protein synthesis. This technology has rendered a significant volume of data on the behavior of proteins during their cell-free synthesis. As illustrated by Tokmakov, these datasets can be exploited to derive the physicochemical and structural properties associated with the solubility and aggregation of eukaryotic proteins

## OPEN ACCESS

### Edited by:

Marc Strous,  
University of Calgary, Canada

### Reviewed by:

Dong-Woo Lee,  
Kyungpook National University,  
South Korea

### \*Correspondence:

Salvador Ventura  
salvador.ventura@uab.es

### Specialty section:

This article was submitted to  
Microbial Physiology and Metabolism,  
a section of the journal  
Frontiers in Microbiology

**Received:** 27 May 2016

**Accepted:** 18 July 2016

**Published:** 29 July 2016

### Citation:

Ventura S (2016) Editorial: Protein  
Solubility and Aggregation in Bacteria.  
Front. Microbiol. 7:1178.  
doi: 10.3389/fmicb.2016.01178

(Tokmakov). However, the cell is a very crowded environment and during cell-free experiments the concentration of cellular components is significantly diluted, questioning whether these assays recapitulate intracellular conditions. To address this question, the group of Taguchi has made an effort to synthesize more than 100 proteins in a bacterial cell-free system, either in the absence or in the presence of specific crowding agents. Their results demonstrate that the impact of crowding on aggregation is not generic, but protein-dependent (Niwa et al.).

Protein aggregates are not distributed homogeneously in the cytoplasm of bacteria, instead, they are mainly located at one or at the two poles of the organism. This localization has important functional consequences, from regulating signaling to protect the rest of the cell from misfolded species. In an original study, Emberly and co-workers from Simon Fraser University demonstrated that the localization of aggregating proteins in bacteria depends on their expression rates. This suggests that the localization of protein aggregates is constrained by aggregation itself, but also by nucleoid occlusion (Scheu et al.). When the accumulation of these polar aggregates results from the recombinant protein expression, they are named inclusion bodies (IBs). Despite they contain certain impurities, IBs are usually highly enriched in the target polypeptide. This property, together with their insolubility, allows for an easy purification of the recombinant polypeptide. In many cases, the protein can be recovered afterwards in its biologically active form upon *in vitro* unfolding and subsequent refolding, as demonstrated by the group of Panda for L-asparaginase (Upadhyay et al.).

Molecular chaperones, like trigger factor, the Dna KJE system and GroEL/GroES survey the protein quality in the bacterial cytosol. However, a dedicated mechanism is needed when proteins should be targeted to the bacterial membrane. Genevaux and co-workers from Université Paul Sabatier review the crucial role played by multitasking Sec B chaperones in this complex process (Sala et al.). Chaperones regulate proteostasis, but they also may allow the apparition of novel beneficial phenotypes. In this way, the Thomas Bentin's group demonstrates how GroEL/GroES over-expression increases cellular fitness and expands the mutational space. These effects provide an opportunity for bacteria to acquire tolerance, and even resistance, to antibiotics (Goltermann et al.).

Accumulating evidence indicates that different organisms exploit the special architecture of amyloid protein aggregates for functional purposes (Otzen, 2010). Functional bacterial amyloids constitute amazing macromolecular systems, where shifts in the folding and solubility of the embedded proteins in response to

environmental factors critically affect activity, as reviewed by Boles and co-workers at University of Iowa (Syed and Boles). Two nice examples of the role played by these functional assemblies are provided in the works of Otzen's group and Lagos lab. In the first case, the authors described how amyloids in the *Pseudomonas* biofilm make a major contribution to the mechanical robustness of this extracellular matrix (Zeng et al.). In the second example, the authors identify the key residues accounting for the amyloid propensity of MccE492, a pore-forming bacteriocin whose antibacterial activity seems to be inactivated in the aggregated state (Aguilera et al.).

Prions are a special class of amyloids, in which the aggregated state becomes self-perpetuating. The prion phenomenon is best-known by its association with encephalopathies in mammals, but it also occurs in lower eukaryotic organisms, like yeast, where it is exploited for functional purposes. The self-assembly of yeast prions relies on the presence of long and intrinsically disordered glutamine/asparagine rich domains. These domains are both necessary and sufficient for self-templating protein aggregation. Giraldo and his group, showed that a fragment of these domains could be replaced by the protein sequence of RepA-WH1, a bacterial protein with amyloid-like properties, without losing the intracellular aggregation potential of the resulting chimera in yeast (Gasset-Rosa and Giraldo). This finding opens up the possibility that prion-like proteins would also exist in prokaryotes. Accordingly, the group of Ventura, using a previously developed computational approach (Espinosa Angarica et al., 2014), identified more than 2000 putative prion candidates in bacterial proteomes (Iglesias et al.). A significant number of these proteins are involved in DNA transcription and protein translation, therefore, playing a crucial role in the regulation of biochemical pathways. One outstanding example of this type of proteins is the Rho terminator factor. Ventura and co-workers demonstrate that in the pathogen *Clostridium Botulinum* this essential protein contains a prion-like domain, with the ability to self-assemble into amyloid structures similar to those found in yeast prions (Pallares et al.).

Overall, it is clear that the study of protein solubility and aggregation in bacteria is a highly dynamic field with the potential to provide very relevant insights and tools to understand and control deleterious and beneficial protein self-assembly.

## AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and approved it for publication.

## REFERENCES

- Dwyer, D. J., Camacho, D. M., Kohanski, M. A., Callura, J. M., and Collins, J. J. (2012). Antibiotic-induced bacterial cell death exhibits physiological and biochemical hallmarks of apoptosis. *Mol. Cell* 46, 561–572. doi: 10.1016/j.molcel.2012.04.027
- Espinosa Angarica, V., Angulo, A., Giner, A., Losilla, G., Ventura, S., and Sancho, J. (2014). PrionScan: an online database of predicted prion domains in complete proteomes. *BMC Genomics* 15:102. doi: 10.1186/1471-2164-15-102
- Fernández-Busquets, X., De Groot, N. S., Fernandez, D., and Ventura, S. (2008). Recent structural and computational insights into conformational diseases. *Curr. Med. Chem.* 15, 1336–1349. doi: 10.2174/092986708784534938
- Invernizzi, G., Papaleo, E., Sabate, R., and Ventura, S. (2012). Protein aggregation: mechanisms and functional consequences. *Int. J. Biochem. Cell Biol.* 44, 1541–1554. doi: 10.1016/j.biocel.2012.05.023
- Navarro, S., Villar-Piqué, A., and Ventura, S. (2014). Selection against toxic aggregation-prone protein sequences in bacteria.



- Biochim. Biophys. Acta* 1843, 866–874. doi: 10.1016/j.bbamcr.2014.01.020
- Otzen, D. (2010). Functional amyloid: turning swords into plowshares. *Prion* 4, 256–264. doi: 10.4161/pri.4.4.13676
- Ventura, S. (2005). Sequence determinants of protein aggregation: tools to increase protein solubility. *Microb. Cell Fact.* 4:11. doi: 10.1186/1475-2859-4-11
- Villar-Piqué, A., Espargaró, A., Sabaté, R., De Groot, N. S., and Ventura, S. (2012). Using bacterial inclusion bodies to screen for amyloid aggregation inhibitors. *Microb. Cell Fact.* 11:55. doi: 10.1186/1475-2859-11-55
- Villar-Pique, A., and Ventura, S. (2012). Modeling amyloids in bacteria. *Microb. Cell Fact.* 11:166. doi: 10.1186/1475-2859-11-166

**Conflict of Interest Statement:** The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Ventura. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Identification of multiple physicochemical and structural properties associated with soluble expression of eukaryotic proteins in cell-free bacterial extracts

Alexander A. Tokmakov\*

Research Center for Environmental Genomics, Kobe University, Kobe, Japan

## Edited by:

Salvador Ventura, *Universitat Autònoma de Barcelona, Spain*

## Reviewed by:

George-John Nychas, *Agricultural University of Athens, Greece*  
Kirill Alexandrov, *University of Queensland, Australia*

## \*Correspondence:

Alexander A. Tokmakov, *Research Center for Environmental Genomics, Kobe University, Rokko dai 1-1 Nada, Kobe, Hyogo 657-8501, Japan*  
e-mail: tokmak@phoenix.kobe-u.ac.jp

Bacterial extracts are widely used to synthesize recombinant proteins. Vast data volumes have been accumulated in cell-free expression databases, covering a whole range of existing proteins. It makes possible comprehensive bioinformatics analysis and identification of multiple features associated with protein solubility and aggregation. In the present paper, an approach to identify the multiple physicochemical and structural properties of amino acid sequences associated with soluble expression of eukaryotic proteins in cell-free bacterial extracts is presented. The method includes: (1) categorical assessment of expression data; (2) calculation and prediction of multiple properties of expressed sequences; (3) correlation of the individual properties with the expression scores; and (4) evaluation of statistical significance of the observed correlations. Using this method, a number of significant correlations between calculated and predicted properties of amino acid sequences and their propensity for soluble cell-free expression have been revealed.

**Keywords:** cell-free protein synthesis, protein solubility, physicochemical and structural protein properties, categorical data analysis, correlation analysis

## INTRODUCTION

Heterologous protein synthesis is widely used for production of recombinant proteins. Particularly, eukaryotic proteins and their domains are often expressed in bacterial hosts (Yokoyama, 2003; Sorensen and Mortensen, 2005; Sivashanmugam et al., 2009; Chen, 2012). However, only a minor fraction of all proteins can be successfully produced in bacterial host systems. Presently, the factors determining expression success in these systems are poorly understood. Various physicochemical features of an amino acid sequence have been implicated as determining factors of soluble protein expression in bacteria (Bertone et al., 2001; Dyson et al., 2004; Goh et al., 2004; Idicula-Thomas and Balaji, 2005).

Recently, cell-free systems of protein synthesis have been developed that offer numerous advantages over cell-based expression (reviewed in Spirin, 2004; Katzen et al., 2005; He, 2008). The cell-free systems allow genome-scale expression of various amino acid sequences under strictly controlled uniform conditions. The productivity of bacterial cell-free synthesis reaches several milligrams of protein per milliliter of reaction mixture (Kigawa et al., 1999). Most often, the purpose of heterologous cell-free synthesis is to produce properly folded and functionally active protein product in the amounts sufficient for structural and functional studies. However, the folding of eukaryotic proteins is greatly compromised in bacterial extracts due to intrinsic differences between the cytoplasmic environments of prokaryotic and eukaryotic cells. Moreover, many eukaryotic proteins require multiple post-translational modifications (PTMs) to attain a native, biologically active state. However, the bacterial expression systems have only a limited capacity for PTMs.

In the present paper, we describe an approach aimed at identification of numerous physicochemical, structural and functional properties of amino acid sequences, including the sites of multiple PTMs, associated with soluble expression of eukaryotic proteins in bacterial cell-free extracts, and highlight major correlations obtained using this approach.

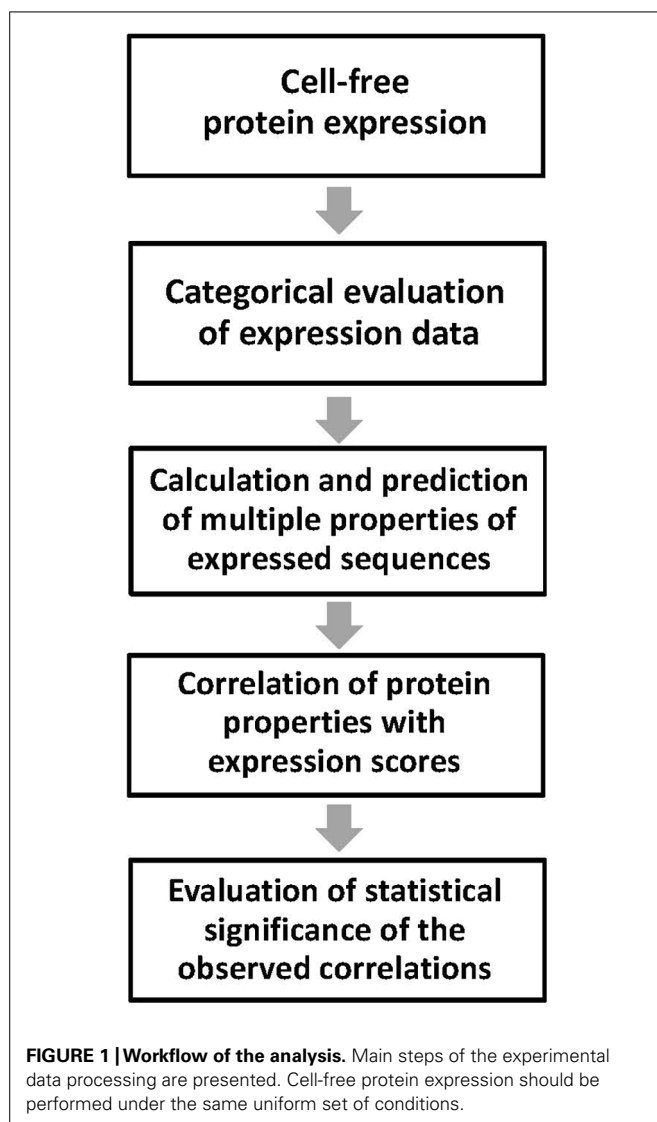
## METHOD

### METHOD OVERVIEW

The developed method is intended for analysis of output from an existing cell-free protein production pipeline. Thus, this paper does not cover the experimental workflow of protein production. It is described in detail in the previous publications (Yabuki et al., 2007; Kigawa et al., 2008; Kurotani et al., 2010; Tokmakov et al., 2012). Here, the focus is set on the processing of experimental data with the purpose of identification of multiple physicochemical and structural properties associated with soluble expression of eukaryotic proteins in cell-free bacterial extracts. Important for the developed approach is that all the proteins in the analyzed dataset are expressed under the same uniform set of conditions. This minimizes the influence of sequence-independent factors and makes possible adequate categorical assessment of expression data (see Categorical Assessment of Expression Data section). The affinity purification tags should be avoided in the expressed sequences because they hinder the analysis of expression correlations by decreasing the role of sequence-specific determinants.

The main steps of the proposed method are summarized in **Figure 1**. They include: (1) categorical assessment of the experimental results of protein expression; (2) determination of multiple physicochemical and structural properties of the expressed amino

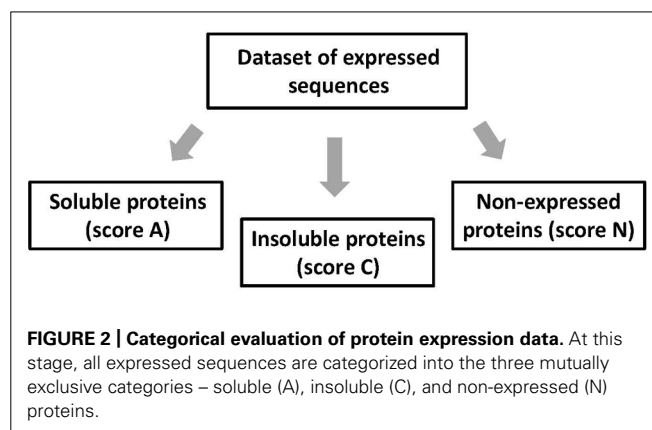




acid sequences using computational and predictive bioinformatics tools; (3) correlation of the individual protein properties with the experimental expression scores; and (4) evaluation of statistical significance of the observed correlations. The developed approach has been extensively used to analyze experimental expression of human proteins and their domains in *Escherichia coli* bacterial extracts (Kurotani et al., 2010; Tokmakov et al., 2012; see Results and Discussion section). However, it can be universally applied to any other cell-free system of heterologous protein synthesis. Each step of the above protocol is detailed below.

#### CATEGORICAL ASSESSMENT OF EXPRESSION DATA

At the stage of expression assessment, all studied proteins are classified into three mutually exclusive categories – soluble (A), insoluble (C), and non-expressed (N) proteins (Figure 2). Each sequence can only be placed into one expression category and not into another. Soluble and insoluble products of protein synthetic reaction can be separated by centrifugation at  $10,000 \times g$  for 10 min and visualized by Coomassie Blue staining after SDS

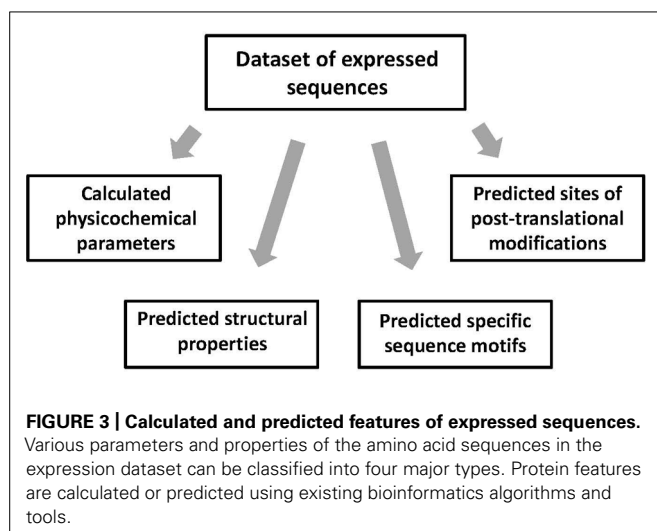


PAGE. The scores A, C, and N are assigned as follows: A, soluble proteins expressed at the level of more than 0.1 mg per ml of cell-free extract; C, expressed, but insoluble proteins; and N, non-expressed proteins with the expression level below 0.1 mg/ml. The protein products expressed at the level below 0.1 mg/ml are difficult to visualize on the Coomassie-stained gels, because the specific protein bands are masked by the endogenous proteins of the bacterial extract. Proteins that are expressed at a lower than expected molecular size should be classified into the category N, as they cannot attain proper structure and function. Notably, in this setting, the score A provides the upper estimation of soluble protein expression, because the procedure of centrifugation at  $10,000 \times g$  cannot discriminate between small protein aggregates and truly soluble proteins. Often, expressed proteins can be found in both soluble and insoluble fractions of the bacterial extract. Lane-to-lane comparison of total and supernatant fractions of the extract in PAGE gels is usually sufficient to establish the preferential pattern of protein expression.

#### CALCULATION AND PREDICTION OF MULTIPLE PROPERTIES OF EXPRESSED SEQUENCES

In this step, multiple features of the amino acid sequences in the expression dataset are calculated or predicted using existing bioinformatics tools. Various protein properties can be classified into the four major types, including physicochemical parameters, structural properties, the presence of specific sequence motifs, and the presence of PTM sites (Figure 3). Many of the physicochemical parameters, such as protein length, molecular weight, amino acid composition, number of charged residues, pI, hydrophobicity, etc., can be calculated using the free ProtParam tool available at the ExPASy server<sup>1</sup>. On the other hand, it is difficult to precisely calculate high-dimensional protein properties, because the 3D structures of expressed protein targets are usually unknown. Still, it is possible to deduce some structural features of the proteins in the expression dataset using existing prediction algorithms. Admittedly, some of these algorithms have quite low prediction accuracy, not exceeding 80%. The low accuracy of prediction thwarts the following correlation analysis, making impossible detection of weak correlations.

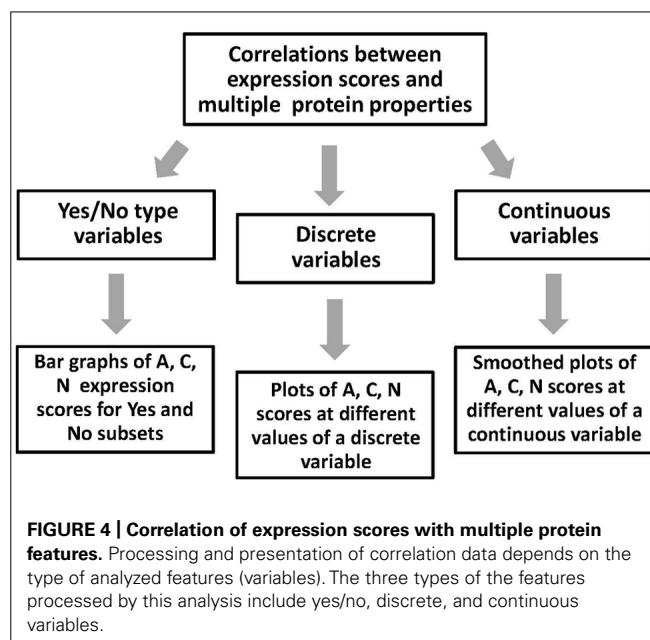
<sup>1</sup><http://www.expasy.org/tools/>



Solvent accessibility can be assessed with the ACCpro 4.0 software downloaded from the SCRATCH Protein Predictor server (Cheng et al., 2005<sup>2</sup>) and content of secondary structure is evaluated with the PREDATOR 2.1.2 tool (Frishman and Argos, 1997) provided online<sup>3</sup>. Coiled coil structures are predicted with the pepcoil tool provided online<sup>4</sup> (Lupas et al., 1991) and content of disordered structure is predicted with the RONN software (Yang et al., 2005<sup>5</sup>). The specific sequence motifs in proteins can also be predicted using available bioinformatics tools. PEST regions, signal sequences, and transmembrane domains are predicted with the tools provided online<sup>6,7,8</sup>. The sites of multiple PTMs, such as phosphorylation, glycosylation, amidation, Asx hydroxylation, sulfation, prenylation, etc., can be predicted using the PROSITE scanning tool PS\_SCAN available online at [http://www.hpa-bioinfotools.org.uk/cgi-bin/ps\\_scan/ps\\_scanCGI.pl](http://www.hpa-bioinfotools.org.uk/cgi-bin/ps_scan/ps_scanCGI.pl). The sites of ubiquitination and SUMOylation are predicted using the site-specific predictors UbPred (Radivojac et al., 2010) and SUMOsp 2.0 (Ren et al., 2009) freely downloadable for academic research from <http://ubpred.org/> and <http://sumosp.biocuckoo.org/>, respectively. The sites of S-palmitoylation are predicted with the CSS-Palm tool (Ren et al., 2008<sup>9</sup>) and S-S bonds can be predicted using the DIpro tool (Cheng et al., 2006) downloadable free from <http://download.igb.uci.edu/intro.html>.

### CORRELATION OF THE INDIVIDUAL PROPERTIES WITH EXPRESSION SCORES

The multiple protein properties calculated and predicted using the above bioinformatics tools can be categorized into the three types,



including yes/no, discrete, and continuous variables (Figure 4). Data processing and presentation differs for the three types of variables. The yes/no type variables, such as single-event PTMs, are the features that can be either present in or absent from proteins. To present the expression data associated with these variables, the bar graphs can be built, which show the ratio of proteins in the expression categories A, C, and N. The graphs should represent two subsets of proteins, excluding and including the analyzed feature. Total number of sequences in the two subsets should be defined. Using these graphs, it is easy to make a side-by-side comparison of the data for the two subsets and deduce the tendencies in protein expression amenability associated with the analyzed feature. To present the expression correlations associated with the discrete variables related to the protein features repeatedly observed in the analyzed sequences, such as abundant multi-site PTMs, another type of data presentation is more convenient. In this case, the percentage of proteins in the expression categories A, C, and N is plotted at different values of analyzed parameter, covering the entire parameter range in the dataset. In addition, the distribution of dataset proteins according to parameter values should be presented. The distribution graphs provide important information concerning the abundance of studied protein features in the analyzed dataset. The processing of data associated with continuous variables, such as sequence hydrophobicity, solvent accessibility, content of intrinsic disorder, etc., is similar to that described for discrete variables. The graphs of A, C, and N scores, as well as the distribution graphs should be provided in the full range of continuous feature values. Curve smoothing is recommended to straighten the graphs obtained with continuous variables. It can be performed using the Excel chart smoothing algorithm. The examples of data presentation for the three types of variables associated with different protein properties are provided in our recent publication (Tokmakov et al., 2014).

<sup>2</sup><http://scratch.proteomics.ics.uci.edu/explanation.html>

<sup>3</sup><http://mobyle.pasteur.fr/cgi-bin/portal.py?#forms::predator>

<sup>4</sup><http://emboss.sourceforge.net/apps/cvs/emboss/apps/pepcoil.html>

<sup>5</sup><http://www.strubi.ox.ac.uk/RONN>

<sup>6</sup><http://emboss.bioinformatics.nl/cgi-bin/emboss/pestfind>

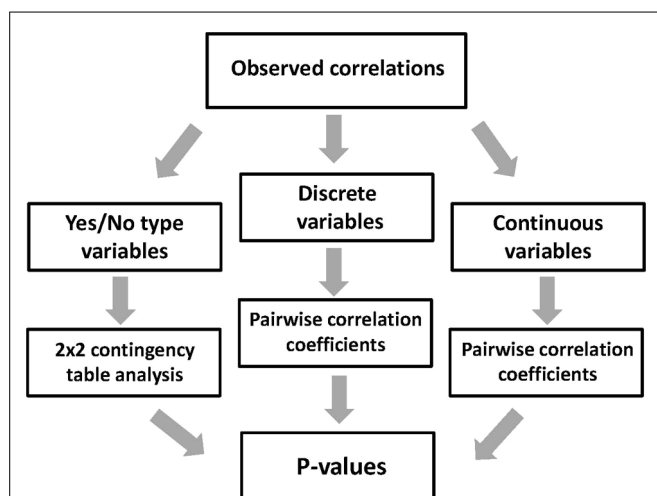
<sup>7</sup><http://www.cbs.dtu.dk/services/SignalP/>

<sup>8</sup><http://harrier.nagahama-i-bio.ac.jp/sosui/>

<sup>9</sup><http://csspalm.biocuckoo.org>

### STATISTICAL SIGNIFICANCE OF THE OBSERVED CORRELATIONS

The expression data processed by the proposed method represent categorical datasets, where all expressed sequences are classified into three categories – soluble (A), insoluble (C), and non-expressed targets (Figure 2). Thus, to evaluate the statistical significance of the observed correlations between the multiple protein features and protein amenability to cell-free expression, the categorical data analysis should be applied (Xu et al., 2010). The estimation of statistical significance should be provided for each expression category (A, C, and N). In addition, multiple protein properties are also categorized into the three types, such as yes/no, discrete, and continuous variables (Figure 4). Evaluation of statistical significance differs for the three types of variables. To deduce the statistical differences associated with yes/no type variables, the two-way contingency table test can be applied (Figure 5). The Fisher's exact *p*-values can be computed using the tool provided on line at <http://statpages.org/ctab2x2.html>. Usually, a confidence level of 95% is set up as the null hypothesis rejection threshold. To evaluate the statistical significance of expression correlations associated with the discrete variables, which have a finite number of possible values, as well as the continuous variables, Pearson's pairwise correlation coefficients should be calculated (Figure 5). The percentage of proteins in the expression categories A, C, and N should be paired with the values of the analyzed variable in the full range of variable values observed in the dataset. Statistical significance of the correlation coefficients is validated by calculating one-tailed probability values, given the value of correlation coefficient (*r*) and the sample size (*n*), with the significance level set to 0.05. Calculations of both correlation coefficients and *p*-values can be performed using the online statistics calculators available at <http://www.danielsoper.com/statcalc3/>. As a general comment, it should be noted that the confidence level of categorical data analysis increases greatly with the number of sequences in the expression datasets (Norman and Streiner, 2000).



**FIGURE 5 | Evaluation of statistical significance of the observed correlations.** Method for statistical evaluation of correlation data is chosen according to the type of analyzed protein features (variables). The three types of the features processed by this analysis include yes/no, discrete, and continuous variables.

### RESULTS AND DISCUSSION

Using the developed method, expression of 3066 human proteins and their domains in a cell-free bacterial system has been analyzed. It was found that the rate of soluble expression (score A) in the investigated dataset constituted 25.7% (Kurotani et al., 2010). This value should be considered as a benchmark, as the similar success rate has been reported for a different subset of human proteins expressed in *E. coli* (Ding et al., 2002). Furthermore, a number of statistically significant correlations between calculated and predicted properties of amino acid sequences and their amenability to bacterial cell-free expression have been identified using the developed approach. The most influential features that affect protein amenability to cell-free expression are listed in Table 1.

Notably, some of these features, such as protein *pI*, hydrophobicity, presence of localization signals, etc., are mostly related to protein solubility, whereas the others, such as protein length,

**Table 1 | Correlations of cell-free protein expression with calculated and predicted properties of amino acid sequences.**

Expression property	Soluble	Insoluble	Undetectable
Length	–	ND	+
<i>pI</i>	±	±	ND
Charge	+	±	–
Hydrophobicity	–	+	ND
Solvent accessibility	+	–	+
Secondary structure	+	±	–
Intrinsic disorder	+	–	+
Protein domains	–	–	+
S–S bonds	–	+	+
Coiled coil	+	–	–
Transmembrane seqs	–	–	+
Localization signals	–	+	ND
PEST regions	+	–	+
Prenylation	+	ND	ND
Phosphorylation	+	–	–
Asn glycosylation	–	+	ND
Palmitoylation	–	±	+
Ubiquitination	+	–	ND
SUMOylation	+	–	±
Amidation	ND	ND	ND
Asx hydroxylation	ND	ND	ND
Sulfation	ND	ND	ND

The signs (+) and (–) indicate positive and negative correlations, respectively; (±) refers to the opposite tendencies of expression estimates at different values of calculated parameters; and ND denotes the lack of correlation.

charge, solvent accessibility, presence of S–S bonds, transmembrane sequences, PEST regions, etc., also affect the overall expression propensity. The presence of some specific sequence motifs was found to be one of the most discriminative parameters for expression propensity. The correlations revealed can be of practical use for protein engineering with the aim of increasing expression success. The rationales for these correlations are discussed in detail in the published paper (Kurotani et al., 2010).

In addition, it was found that amenability of human polypeptide sequences to bacterial cell-free expression correlates with the presence of multiple PTM sites bioinformatically predicted in these sequences (Tokmakov et al., 2012; **Table 1**). Surprisingly, the presence of predicted sites for several PTMs, such as ubiquitination, SUMOylation, etc. (**Table 1**), was associated with increased production of properly folded soluble protein. However, no SUMOylation and ubiquitination machineries are known to exist in bacteria, suggesting that the presence of these PTM sites in amino acid sequences is related to intrinsically better protein solubility even in the absence of the modifications. It was hypothesized that physicochemical and/or structural characteristics of the modification sites themselves convey the better solubility (Tokmakov et al., 2012). Altogether, these findings indicate that identification of potential PTM sites in polypeptide sequences can be of practical use for predicting expression success and optimizing heterologous protein synthesis. Currently, a discriminant-based machine-learning algorithm that utilizes multiple features of amino acid sequences to predict the success rate of heterologous protein synthesis is being developed based on the reported findings. The algorithm will provide a basis for the internet-based tool for predicting amenability of eukaryotic proteins to cell-free expression in a prokaryotic system.

## ACKNOWLEDGMENTS

This work was supported by the research fund for Foreign Visiting Professor from Kobe University and the Grant-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology of Japan (no. 25440023).

## REFERENCES

- Bertone, P., Kluger, Y., Lan, N., Zheng, D., Christendat, D., Yee, A., et al. (2001). SPINE: an integrated tracking database and data mining approach for identifying feasible targets in high-throughput structural proteomics. *Nucleic Acids Res.* 29, 2884–2898. doi: 10.1093/nar/29.13.2884
- Chen, R. (2012). Bacterial expression systems for recombinant protein production: *E. coli* and beyond. *Biotechnol. Adv.* 30, 1102–1107. doi: 10.1016/j.biotechadv.2011.09.013
- Cheng, J., Randall, A. Z., Sweredoski, M. J., and Baldi, P. (2005). SCRATCH: a protein structure and structural feature prediction server. *Nucleic Acids Res.* 33, W72–W76. doi: 10.1093/nar/gki396
- Cheng, J., Saigo, H., and Baldi, P. (2006). Large-scale prediction of disulfide bridges using kernel methods, two-dimensional recursive neural networks, and weighted graph matching. *Proteins* 62, 617–629. doi: 10.1002/prot.20787
- Ding, H. T., Ren, H., Chen, Q., Fang, G., Li, L. F., Li, R., et al. (2002). Parallel cloning, expression, purification, and crystallization of human proteins for structural genomics. *Acta Crystallogr. D Biol. Crystallogr.* 58, 2102–2108. doi: 10.1107/S0907444902016359
- Dyson, M. R., Shadbolt, S. P., Vincent, K. J., Perera, R. L., and McCafferty J. (2004). Production of soluble mammalian proteins in *Escherichia coli*: identification of protein features that correlate with successful expression. *BMC Biotechnol.* 4:32. doi: 10.1186/1472-6750-4-32
- Frishman, D., and Argos, P. (1997). Seventy-five percent accuracy in protein secondary structure prediction. *Proteins* 27, 329–335. doi: 10.1002/(SICI)1097-0134(199703)27:3<329::AID-PROT1>3.0.CO;2-8
- Goh, C. S., Lan, N., Douglas, S. M., Wu, B., Echols, N., Smith, A., et al. (2004). Mining the structural genomics pipeline: identification of protein properties that affect high-throughput experimental analysis. *J. Mol. Biol.* 336, 115–130. doi: 10.1016/j.jmb.2003.11.053
- He, M. (2008). Cell-free protein synthesis: applications in proteomics and biotechnology. *Nat. Biotechnol.* 25, 126–132. doi: 10.1016/j.nbt.2008.08.004
- Idicula-Thomas, S., and Balaji, P. V. (2005). Understanding the relationship between the primary structure of proteins and its propensity to be soluble on over-expression in *Escherichia coli*. *Protein Sci.* 14, 582–592. doi: 10.1110/ps.041009005
- Katzen, F., Chang, G., and Kudlicki, W. (2005). The past, present and future of cell-free protein synthesis. *Trends Biotechnol.* 23, 150–156. doi: 10.1016/j.tibtech.2005.01.003
- Kigawa, T., Matsuda, T., Yabuki, T., and Yokoyama, S. (2008). “Bacterial cell-free system for highly efficient protein synthesis,” in *Cell-Free Protein Synthesis*, eds A. S. Spirin and J. R. Swartz (Wiley VCH: Weinheim), 83–97.
- Kigawa, T., Yabuki, T., Yoshida, Y., Tsutsui, M., Ito, Y., Shibata, T., et al. (1999). Cell-free production and stable-isotope labeling of milligram quantities of proteins. *FEBS Lett.* 442, 15–19. doi: 10.1016/S0014-5793(98)01620-2
- Kurotani, A., Takagi, T., Toyama, M., Shirouzu, M., Yokoyama, S., Fukami, Y., et al. (2010). Comprehensive bioinformatics analysis of cell-free protein synthesis: identification of multiple protein properties that correlate with successful expression. *FASEB J.* 24, 1095–1104. doi: 10.1096/fj.09-139527
- Lupas, A., Van Dyke, M., and Stock, J. (1991). Predicting coiled coils from protein sequences. *Science* 252, 1162–1164. doi: 10.1126/science.252.5009.1162
- Norman, G. R., and Streiner, D. L. (2000). *Biostatistics: The Bare Essentials*. B. C. Decker, Hamilton.
- Radivojac, P., Vacic, V., Haynes, C., Cocklin, R. R., Mohan, A., Heyen, J. W., et al. (2010). Identification, analysis, and prediction of protein ubiquitination sites. *Proteins* 78, 365–380. doi: 10.1002/prot.22555
- Ren, J., Gao, X., Jin, C., Zhu, M., Wang, X., Shaw, A., et al. (2009). Systematic study of protein sumoylation: development of a site-specific predictor of SUMOylation. *Proteomics* 9, 3409–3412. doi: 10.1002/pmic.200800646
- Ren, J., Wen, L., Gao, X., Jin, C., Xue, Y., and Yao, X. (2008). CSS-Palm 2.0: an updated software for palmitoylation sites prediction. *Protein Eng. Des. Sel.* 21, 639–644. doi: 10.1093/protein/gzn039
- Sivashanmugam, A., Murray, V., Cui, C., Zhang, Y., Wang, J., and Li, Q. (2009). Practical protocols for production of very high yields of recombinant proteins using *Escherichia coli*. *Protein Sci.* 18, 936–948. doi: 10.1002/pro.102
- Sorensen, H. P., and Mortensen, K. K. (2005). Advanced genetic strategies for recombinant protein expression in *Escherichia coli*. *J. Biotechnol.* 115, 113–128. doi: 10.1016/j.jbiotec.2004.08.004
- Spirin, A. S. (2004). High-throughput cell-free systems for synthesis of functionally active proteins. *Trends Biotechnol.* 22, 538–545. doi: 10.1016/j.tibtech.2004.08.012
- Tokmakov, A. A., Kurotani, A., Shirouzu, M., Fukami, Y., and Yokoyama, S. (2014). Bioinformatics analysis and optimization of cell-free protein synthesis. *Methods Mol. Biol.* 1118, 17–33. doi: 10.1007/978-1-62703-782-2\_2
- Tokmakov, A. A., Kurotani, A., Takagi, T., Toyama, M., Shirouzu, M., Fukami, Y., et al. (2012). Multiple post-translational modifications affect heterologous protein synthesis. *J. Biol. Chem.* 287, 27106–27116. doi: 10.1074/jbc.M112.366351
- Xu, B., Feng, X., and Burdine, R. D. (2010). Categorical data analysis in experimental biology. *Dev. Biol.* 348, 3–11. doi: 10.1016/j.ydbio.2010.08.018
- Yabuki, T., Motoda, Y., Hanada, K., Nunokawa, E., Saito, M., Seki, E., et al. (2007). A robust two-step PCR method of template DNA production for high-throughput cell-free protein synthesis. *J. Struct. Funct. Genomics* 8, 173–191. doi: 10.1007/s10969-007-9038-z
- Yang, Z. R., Thomson, R., McMeil, P., and Esnouf, R. M. (2005). RONN: the bio-basis function neural network technique applied to the detection of natively disordered regions in proteins. *Bioinformatics* 21, 3369–3376. doi: 10.1093/bioinformatics/bti534
- Yokoyama, S. (2003). Protein expression systems for structural genomics and proteomics. *Curr. Opin. Chem. Biol.* 7, 39–43. doi: 10.1016/S1367-5931(02)00019-4



**Conflict of Interest Statement:** The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 10 April 2014; accepted: 29 May 2014; published online: 20 June 2014.

Citation: Tokmakov AA (2014) Identification of multiple physicochemical and structural properties associated with soluble expression of eukaryotic proteins in cell-free bacterial extracts. *Front. Microbiol.* 5:295. doi: 10.3389/fmicb.2014.00295

*This article was submitted to Microbial Physiology and Metabolism, a section of the journal Frontiers in Microbiology.*

Copyright © 2014 Tokmakov. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Large-scale analysis of macromolecular crowding effects on protein aggregation using a reconstituted cell-free translation system

Tatsuya Niwa<sup>1†</sup>, Ryota Sugimoto<sup>1†</sup>, Lisa Watanabe<sup>1</sup>, Shugo Nakamura<sup>2</sup>, Takuya Ueda<sup>3</sup> and Hideki Taguchi<sup>1\*</sup>

<sup>1</sup> Department of Biomolecular Engineering, Graduate School of Bioscience and Biotechnology, Tokyo Institute of Technology, Yokohama, Japan, <sup>2</sup> Department of Biotechnology, Graduate School of Agricultural and Life Sciences, The University of Tokyo, Tokyo, Japan, <sup>3</sup> Department of Computational Biology and Medical Sciences, Graduate School of Frontier Sciences, The University of Tokyo, Chiba, Japan

## OPEN ACCESS

### Edited by:

Salvador Ventura,  
Universitat Autònoma de Barcelona,  
Spain

### Reviewed by:

Dong-Woo Lee,  
Kyungpook National University, South  
Korea

Pierre Genevoux,

Centre National de la Recherche  
Scientifique, France

### \*Correspondence:

Hideki Taguchi  
taguchi@bio.titech.ac.jp

<sup>†</sup>These authors have contributed  
equally to this work.

### Specialty section:

This article was submitted to  
Microbial Physiology and Metabolism,  
a section of the journal  
Frontiers in Microbiology

**Received:** 26 June 2015

**Accepted:** 25 September 2015

**Published:** 08 October 2015

### Citation:

Niwa T, Sugimoto R, Watanabe L,  
Nakamura S, Ueda T and Taguchi H  
(2015) Large-scale analysis  
of macromolecular crowding effects  
on protein aggregation using  
a reconstituted cell-free translation  
system. *Front. Microbiol.* 6:1113.  
doi: 10.3389/fmicb.2015.01113

Proteins must fold into their native structures in the crowded cellular environment, to perform their functions. Although such macromolecular crowding has been considered to affect the folding properties of proteins, large-scale experimental data have so far been lacking. Here, we individually translated 142 *Escherichia coli* cytoplasmic proteins using a reconstituted cell-free translation system in the presence of macromolecular crowding reagents (MCRs), Ficoll 70 or dextran 70, and evaluated the aggregation propensities of 142 proteins. The results showed that the MCR effects varied depending on the proteins, although the degree of these effects was modest. Statistical analyses suggested that structural parameters were involved in the effects of the MCRs. Our dataset provides a valuable resource to understand protein folding and aggregation inside cells.

**Keywords:** protein aggregation, protein folding, cell-free translation system, macromolecular crowding, large-scale analysis

Most proteins must properly fold into their native tertiary structures, defined by the primary amino acid sequences, to perform their functions (Anfinsen, 1973; Dobson, 2003). However, protein folding is a highly complicated physicochemical process, and many proteins require the aid of molecular chaperones to fold into their correct structures, both *in vitro* and *in vivo* (Tyedmers et al., 2010; Hartl et al., 2011). Misfolded proteins often form protein aggregates, which leads to the loss of protein function and sometimes cause toxic effects in the cells (Tyedmers et al., 2010).

To clarify the principles of the protein aggregation and the properties associated with it, we conducted a comprehensive analysis of protein aggregation under the completely chaperone-free condition by using a *Escherichia coli* reconstituted cell-free translation system (Niwa et al., 2009). In this analysis, thousands of bacterial proteins were expressed separately, and their aggregation propensities were evaluated by using a centrifugation-based method. Statistical analyses revealed significant insights concerning protein aggregation (Niwa et al., 2009).

In the previous analysis, the aggregation propensity was evaluated in a diluted solution, in which the protein concentration was at most 1–2 mg/mL (Shimizu et al., 2001, 2005; Niwa et al., 2009).

However, the intracellular environment is much more crowded with macromolecules such as proteins and nucleic acids, and such an environment has been thought to affect the protein folding properties and the aggregation propensity (Zhou et al., 2008; Zhou, 2013). The effect of macromolecular crowding on protein folding and aggregation has been studied extensively, from both theoretical and experimental viewpoints, for decades (Zhou et al., 2008; Elcock, 2010; Gershenson and Gierasch, 2011; Zhou, 2013), and some studies suggested that the macromolecular crowding effects increase the intermolecular interactions mainly by its excluded volume effect, and hence facilitate the aggregation of some proteins (van den Berg et al., 1999; Munishkina et al., 2004). In contrast, other studies predicted that the crowding effects increase the stability of the native state and tend to bias proteins toward the native structure, although the effect on the stability was suggested to be modest (Cheung et al., 2005; Christiansen et al., 2010; Hong and Gierasch, 2010; Mittal and Best, 2010; Wang et al., 2010; Gershenson and Gierasch, 2011). However, in either case, these studies were limited to the experiments with a small number of model substrates or theoretical approaches.

To gain insight into the effects of macromolecular crowding on protein folding and aggregation and confirm these theories, we performed a large-scale analysis of the macromolecular crowding effects with a variety of proteins, by attempting the “*in vitro* proteome” approach reported previously (Niwa et al., 2009, 2012). By using a reconstituted cell-free translation system (Shimizu et al., 2001, 2005), we can easily evaluate the macromolecular crowding effects during the translation reaction for various kinds of proteins. In this analysis, we chose two macromolecular crowding reagents (MCRs), Ficoll 70, and dextran 70, because both two MCRs are hydrophilic polysaccharide and expected to have low interaction in specific amino acid side chains. Hence, the effects of these two MCRs can be thought to be mainly attributed to its excluded volume effect without any significant inhibition of expression reactions by the cell-free translation system (Zhou et al., 2008). In fact, we tried to use polyethylene glycol (PEG) 3350 as another MCR, but we could not evaluate its effect because the presence of PEG 3350 almost entirely abolished the protein expression by the cell-free translation system.

The method for the evaluation followed the previous comprehensive analyses of protein aggregation (see Materials and Methods). The measurement error of the solubility in the presence of the MCRs was about  $\pm 10\%$ , which is nearly equal to that in the absence of MCRs, as reported previously (Niwa et al., 2009).

We performed this experiment for 150 *E. coli* proteins under three conditions: no addition of MCRs, Ficoll-added, and dextran-added conditions. These 150 proteins were chosen at random among the proteins that were annotated as cytoplasmic proteins and whose aggregation propensities were evaluated in the previous comprehensive analysis (Niwa et al., 2009). Among the tested proteins, 142 proteins were quantified under the three conditions. All obtained data are shown in Supplementary Table S1 in the dataset, which is available at

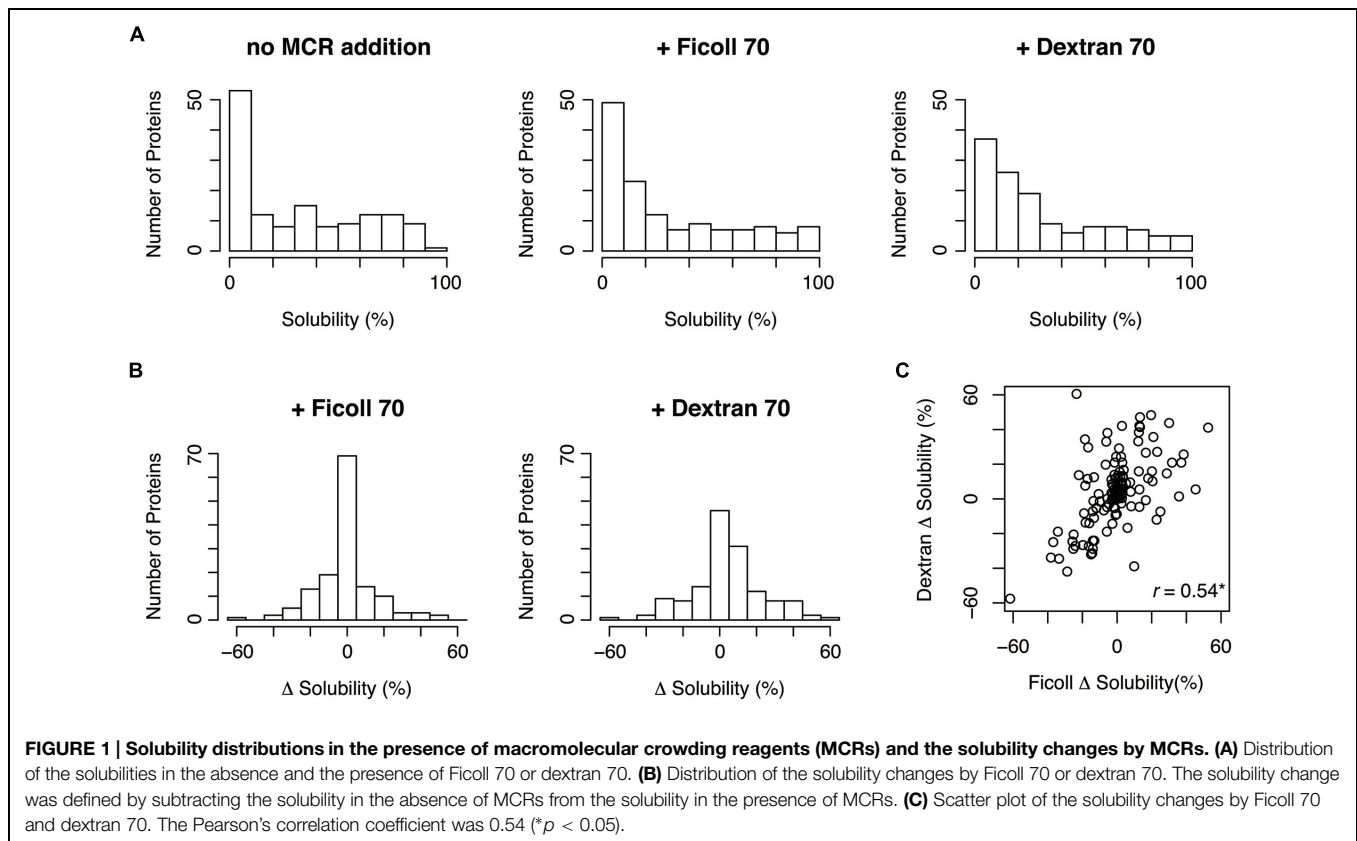
figshare repository<sup>1</sup>. The distributions of the solubilities under the Ficoll- and dextran-added conditions were similar to that in the absence of MCRs (Figure 1A). This result suggested that Ficoll and dextran do not exert strong effects on the overall aggregation propensity. However, the distribution of the solubility changes under the dextran-added conditions was slightly biased toward a higher level, suggesting that dextran tends to act to prevent protein aggregation (Figure 1B). Furthermore, the solubility changes by Ficoll or dextran were widely distributed between  $-50$  and  $+50\%$ , suggesting that the MCRs could act both positively and negatively on aggregate formation. In other words, the degree or direction of the effect of the MCRs on the aggregation propensity depends on the properties of the proteins. Moreover, the solubility changes by Ficoll correlated well with those by dextran (Figure 1C), indicating the similar effects of the two MCRs on the aggregation propensity.

As expected, the addition of both two MCRs did not cause drastic changes in the synthetic yield of the cell-free translation system. Furthermore, the change of the synthetic yield did not show a significant correlation with the solubility changes, suggesting that the effects of the changes in the synthetic yield on the solubility changes by the MCRs were small.

To determine which properties were related to the effects of the MCRs, we compared the solubility changes by MCRs and the physicochemical properties, such as molecular weight and isoelectric point. Although the molecular weight did not correlate with the solubility change, the solubility change by dextran positively correlated with the isoelectric point (Figure 2A). Moreover, the net charge, calculated by the number of charged amino acid residues, also correlated with the solubility change by dextran. These results suggested that dextran tends to act as an aggregation inhibitor for positively charged proteins. Concerning the properties derived from the primary sequence, we compared the solubility changes with the content ratios of the four amino acid groups, classified according to their properties. However, no obvious correlation was observed in the ratio of negatively charged (Asp and Glu), positively charged (Lys, Arg, and His), aromatic (Phe, Tyr, and Trp), or hydrophobic (Val, Leu, and Ile) amino acids.

Most proteins adopt a unique tertiary structure defined by the amino acid sequence, and a wide variety of structures exist. To compare the structural properties of proteins with the effects of the MCRs, we used the structural classification of proteins (SCOP) database (Murzin et al., 1995). As reported previously, the SCOP fold seems to have a formidable influence on the aggregation propensity, and proteins with specific folds have a strong tendency to form aggregates (Niwa et al., 2009). We then extracted the proteins with aggregation-prone folds, and investigated the distribution of their solubility changes by the MCRs. All annotations of the SCOP folds for tested proteins are listed in Supplementary Table S2 in the dataset<sup>1</sup>, and the aggregation-prone folds in the previous report (Niwa et al., 2009) were as follows; c37: P-loop containing nucleoside triphosphate hydrolases, a4: DNA/RNA-binding 3-helical bundle, c1: TIM

<sup>1</sup><http://dx.doi.org/10.6084/m9.figshare.1495333>



$\beta/\alpha$ -barrel, c3: FAD/NAD(P)-binding domain, c55: Ribonuclease H-like motif, and c94: Periplasmic binding protein-like II. The histograms of the solubility changes for the proteins with the aggregation-prone folds revealed strong biases toward lower solubility, indicating that the MCRs tended to enhance the aggregate formation of these proteins (Figure 2B).

To compare further structural features, we constructed structural models for 41 proteins, by using a template-based modeling method. The structural templates used for the modeling are listed in Supplementary Table S3 in the dataset<sup>2</sup>. Comparisons of the radius of gyration and the surface area of the amino acid main chains showed that both parameters positively correlated with the solubility change by dextran (Figures 2C,D). In addition, we compared the solubility change by the MCRs and the relative contact order, which is considered to be related to protein folding (Plaxco et al., 1998). Although the contact order negatively correlated with the solubility change by dextran, the correlation between them was not statistically significant.

The data obtained from this study suggested that the macromolecular crowding effect enhances aggregate formation for some proteins and prevents it for others. Previous experimental and theoretical studies suggested that the macromolecular crowding effects are often quite complicated, and particularly difficult to understand quantitatively (Zhou et al., 2008; Elcock, 2010; Zhou, 2013). The results obtained here

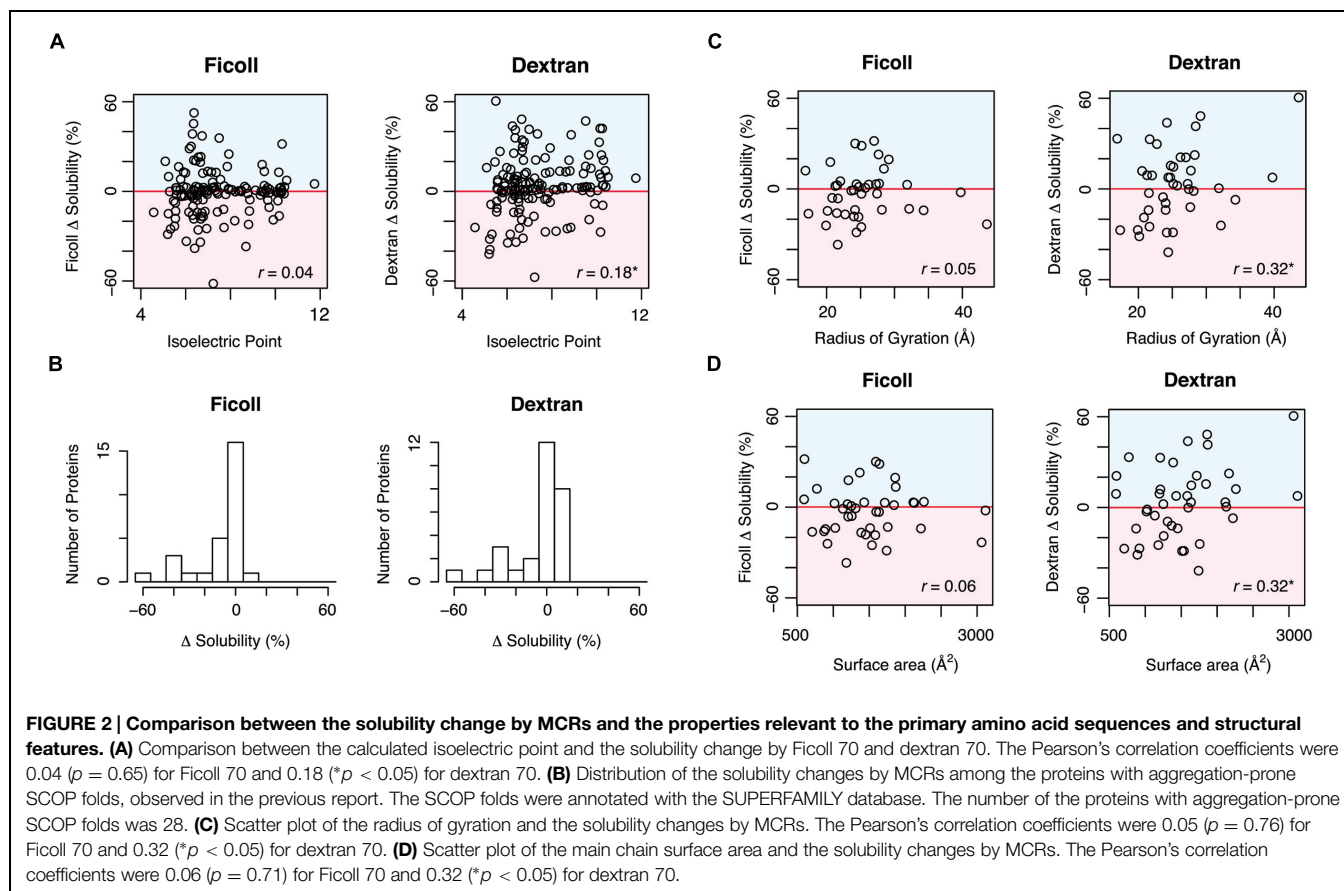
seem to reflect this complexity of the macromolecular crowding effects on protein folding and aggregation. In addition, some studies suggested that the effect of macromolecular crowding on protein folding is modest (Zhou et al., 2008; Mittal and Best, 2010; Zhou, 2013). Our data seem to be in agreement with these ideas, because the effects of macromolecular crowding were not strong, in comparison with the influences of molecular chaperones reported previously (Figure 1A; Niwa et al., 2012). Although our statistical analyses gave some insights for understanding the macromolecular crowding effects as described above, their influences on protein folding and aggregation are quite complicated and further detailed analysis are needed. Our dataset obtained from the “*in vitro* proteome” approach has great potential, as a valuable dataset that will contribute to further understanding of the effects of macromolecular crowding and protein folding inside cells.

## MATERIALS AND METHODS

### Method for the Evaluation of the Aggregation Propensity

The method for the evaluation of the aggregation propensity followed those used in previous comprehensive analysis (Niwa et al., 2009). The template DNA for expression by the cell-free translation system was amplified from an *E. coli* ORF library [ASKA library (Kitagawa et al., 2005; Riley et al., 2006)] by PCR,

<sup>2</sup><http://dx.doi.org/10.6084/m9.figshare.1495333>



as described previously (Niwa et al., 2009). The transcription-translation-coupled expression was conducted by a reconstituted cell-free translation system [PURE system (Shimizu et al., 2001, 2005)] at 37°C for 1 h. For detection, L-[<sup>35</sup>S]-methionine was added to the PURE system. Ficoll 70 (GE Healthcare) or dextran 70 (Sigma-Aldrich) was also included at the concentration of 80 mg/ml in the reaction, to evaluate the effects of MCRs. After the expression, an aliquot was withdrawn as the total fraction, and the remainder was centrifuged at 20,000 × *g* for 30 min. The total and supernatant fractions were separated by SDS-PAGE, and the band intensities were quantified by autoradiography (FLA7000 image analyzer and Multi Gauge software, Fujifilm). The ratio of the supernatant to the total protein was defined as the solubility, as referred to as the index of aggregation propensity.

## Data Analysis

The molecular weight, amino acid content, and net charge were calculated from the amino acid sequences obtained from GenoBase<sup>3</sup> (Kitagawa et al., 2005; Riley et al., 2006). Estimation of pI values was conducted with a web tool<sup>4</sup> (Sillero and Maldonado, 2006). The SCOP (Murzin et al., 1995) classification was obtained from the dataset distributed

by GenoBase. The SCOP fold annotation in GenoBase was based on the SUPERFAMILY database (Madera et al., 2004). The SCOP folds annotated as aggregation-prone folds were as follows; c37: P-loop containing nucleoside triphosphate hydrolases, a4: DNA/RNA-binding 3-helical bundle, c1: TIM β/α-barrel, c3: FAD/NAD(P)-binding domain, c55: Ribonuclease H-like motif, and c94: Periplasmic binding protein-like II (Niwa et al., 2009). The modeled structures were obtained from the database by Zhang's group<sup>5</sup> or modeled by the MODELER program<sup>6</sup> (Eswar et al., 2006). Among the 41 modeled structures, 28 were selected from Zhang's database with the following criteria: >80% template identity, >80% template coverage, and >0.7 TM-score to the template. The remaining 13 structures were modeled by MODELER with the template PDBs determined by a PSI-BLAST search with the following criteria: >80% template identity and >80% template coverage. The radius of gyration and the relative contact order were calculated by using in-house developed scripts. Surface area was calculated with the NACCESS software<sup>7</sup> (Hubbard and Thornton, 1993). All statistical tests were conducted with the R software<sup>8</sup>.

<sup>3</sup><http://zhanglab.cmb.med.umich.edu/QUARK/ecoli2/>

<sup>6</sup><https://salilab.org/modeller/>

<sup>7</sup><http://www.bioinf.manchester.ac.uk/naccess/>

<sup>8</sup><http://www.r-project.org/>

<sup>3</sup><http://ecoli.naist.jp/GB/>

<sup>4</sup><http://isoelectric.ovh.org/>



## ACKNOWLEDGMENTS

This work was supported in part by JSPS KAKENHI Grant Numbers 22870010 and 25840045 (to T. N.), MEXT KAKENHI Grant Numbers 19058002 and 26116002 (to H. T.), and the Platform Project for Supporting in Drug Discovery and Life Science Research (Platform for Drug Discovery, Informatics, and Structural Life Science) from the Ministry of Education, Culture, Sports, Science and Technology (MEXT) and the

Japan Agency for Medical Research and Development (AMED) (to S. N. and T. U.)

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fmicb.2015.01113>

## REFERENCES

- Anfinsen, C. B. (1973). Principles that govern the folding of protein chains. *Science* 181, 223–230. doi: 10.1126/science.181.4096.223
- Cheung, M. S., Klimov, D., and Thirumalai, D. (2005). Molecular crowding enhances native state stability and refolding rates of globular proteins. *Proc. Natl. Acad. Sci. U.S.A.* 102, 4753–4758. doi: 10.1073/pnas.0409630102
- Christiansen, A., Wang, Q., Samiotakis, A., Cheung, M. S., and Wittung-Stafshede, P. (2010). Factors defining effects of macromolecular crowding on protein stability: an in vitro/in silico case study using cytochrome c. *Biochemistry* 49, 6519–6530. doi: 10.1021/bi100578x
- Dobson, C. M. (2003). Protein folding and misfolding. *Nature* 426, 884–890. doi: 10.1038/nature02261
- Elcock, A. H. (2010). Models of macromolecular crowding effects and the need for quantitative comparisons with experiment. *Curr. Opin. Struct. Biol.* 20, 196–206. doi: 10.1016/j.sbi.2010.01.008
- Eswar, N., Webb, B., Marti-Renom, M. A., Madhusudhan, M. S., Eramian, D., Shen, M. Y., et al. (2006). Comparative protein structure modeling using Modeller. *Curr. Protoc. Bioinformatics* Chap. 5, Unit 5.6. doi: 10.1002/0471250953.bi0506s15
- Gershenson, A., and Gierasch, L. M. (2011). Protein folding in the cell: challenges and progress. *Curr. Opin. Struct. Biol.* 21, 32–41. doi: 10.1016/j.sbi.2010.11.001
- Hartl, F. U., Bracher, A., and Hayer-Hartl, M. (2011). Molecular chaperones in protein folding and proteostasis. *Nature* 475, 324–332. doi: 10.1038/nature10317
- Hong, J., and Gierasch, L. M. (2010). Macromolecular crowding remodels the energy landscape of a protein by favoring a more compact unfolded state. *J. Am. Chem. Soc.* 132, 10445–10452. doi: 10.1021/ja103166y
- Hubbard, S. J. T., and Thornton, J. M. (1993). “NACCESS; Computer Program”. London: Department of Biochemistry and Molecular Biology, University College London.
- Kitagawa, M., Ara, T., Arifuzzaman, M., Ioka-Nakamichi, T., Inamoto, E., Toyonaga, H., et al. (2005). Complete set of ORF clones of *Escherichia coli* ASKA library (a complete set of *E. coli* K-12 ORF archive): unique resources for biological research. *DNA Res.* 12, 291–299. doi: 10.1093/dnares/dsi012
- Madera, M., Vogel, C., Kummerfeld, S. K., Chothia, C., and Gough, J. (2004). The SUPERFAMILY database in 2004: additions and improvements. *Nucleic Acids Res.* 32, D235–D239. doi: 10.1093/nar/gkh117
- Mittal, J., and Best, R. B. (2010). Dependence of protein folding stability and dynamics on the density and composition of macromolecular crowders. *Biophys. J.* 98, 315–320. doi: 10.1016/j.bpj.2009.10.009
- Munishkina, L. A., Cooper, E. M., Uversky, V. N., and Fink, A. L. (2004). The effect of macromolecular crowding on protein aggregation and amyloid fibril formation. *J. Mol. Recognit.* 17, 456–464. doi: 10.1002/jmr.699
- Murzin, A. G., Brenner, S. E., Hubbard, T., and Chothia, C. (1995). SCOP: a structural classification of proteins database for the investigation of sequences and structures. *J. Mol. Biol.* 247, 536–540. doi: 10.1016/S0022-2836(05)80134-2
- Niwa, T., Kanamori, T., Ueda, T., and Taguchi, H. (2012). Global analysis of chaperone effects using a reconstituted cell-free translation system. *Proc. Natl. Acad. Sci. U.S.A.* 109, 8937–8942. doi: 10.1073/pnas.1201380109
- Niwa, T., Ying, B. W., Saito, K., Jin, W., Takada, S., Ueda, T., et al. (2009). Bimodal protein solubility distribution revealed by an aggregation analysis of the entire ensemble of *Escherichia coli* proteins. *Proc. Natl. Acad. Sci. U.S.A.* 106, 4201–4206. doi: 10.1073/pnas.0811922106
- Plaxco, K. W., Simons, K. T., and Baker, D. (1998). Contact order, transition state placement and the refolding rates of single domain proteins. *J. Mol. Biol.* 277, 985–994. doi: 10.1006/jmbi.1998.1645
- Riley, M., Abe, T., Arnaud, M. B., Berlyn, M. K., Blattner, F. R., Chaudhuri, R. R., et al. (2006). *Escherichia coli* K-12: a cooperatively developed annotation snapshot–2005. *Nucleic Acids Res.* 34, 1–9. doi: 10.1093/nar/gkj405
- Shimizu, Y., Inoue, A., Tomari, Y., Suzuki, T., Yokogawa, T., Nishikawa, K., et al. (2001). Cell-free translation reconstituted with purified components. *Nat. Biotechnol.* 19, 751–755. doi: 10.1038/90802
- Shimizu, Y., Kanamori, T., and Ueda, T. (2005). Protein synthesis by pure translation systems. *Methods* 36, 299–304. doi: 10.1016/j.ymeth.2005.04.006
- Sillero, A., and Maldonado, A. (2006). Isoelectric point determination of proteins and other macromolecules: oscillating method. *Comput. Biol. Med.* 36, 157–166. doi: 10.1016/j.compbiomed.2004.09.006
- Tyedmers, J., Mogk, A., and Bukau, B. (2010). Cellular strategies for controlling protein aggregation. *Nat. Rev. Mol. Cell Biol.* 11, 777–788. doi: 10.1038/nrm2993
- van den Berg, B., Ellis, R. J., and Dobson, C. M. (1999). Effects of macromolecular crowding on protein folding and aggregation. *EMBO J.* 18, 6927–6933. doi: 10.1093/emboj/18.24.6927
- Wang, Y., He, H., and Li, S. (2010). Effect of Ficoll 70 on thermal stability and structure of creatine kinase. *Biochemistry (Mosc.)* 75, 648–654. doi: 10.1134/S0006297910050160
- Zhou, H. X. (2013). Influence of crowded cellular environments on protein folding, binding, and oligomerization: biological consequences and potentials of atomistic modeling. *FEBS Lett.* 587, 1053–1061. doi: 10.1016/j.febslet.2013.01.064
- Zhou, H. X., Rivas, G., and Minton, A. P. (2008). Macromolecular crowding and confinement: biochemical, biophysical, and potential physiological consequences. *Annu. Rev. Biophys.* 37, 375–397. doi: 10.1146/annurev.biophys.37.032807.125817

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Niwa, Sugimoto, Watanabe, Nakamura, Ueda and Taguchi. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Localization of aggregating proteins in bacteria depends on the rate of addition

Karlton Scheu<sup>1</sup>, Rakinder Gill<sup>1</sup>, Saeed Saberi<sup>1</sup>, Pablo Meyer<sup>2</sup> and Eldon Emberly<sup>1\*</sup>

<sup>1</sup> Department of Physics, Simon Fraser University, Burnaby, BC, Canada

<sup>2</sup> IBM T.J. Watson Research Center, New York, NY, USA

## Edited by:

Salvador Ventura, Universitat Autònoma de Barcelona, Spain

## Reviewed by:

Jason Warren Cooley, University of Missouri, USA

Andrew David Rutenberg, Dalhousie University, Canada

## \*Correspondence:

Eldon Emberly, Department of Physics, Simon Fraser University, 8888 University Drive, Burnaby V5A 1S6, BC, Canada  
e-mail: eemberly@sfu.ca

Many proteins are observed to localize to specific subcellular regions within bacteria. Recent experiments have shown that proteins that have self-interactions that lead them to aggregate tend to localize to the poles. Theoretical modeling of the localization of aggregating protein within bacterial cell geometries shows that aggregates can spontaneously localize to the pole due to nucleoid occlusion. The resulting polar localization, whether it be to a single pole or to both was shown to depend on the rate of protein addition. Motivated by these predictions we selected a set of genes from *Escherichia coli*, whose protein products have been reported to localize when tagged with green fluorescent protein (GFP), and explored the dynamics of their localization. We induced protein expression from each gene at different rates and found that in all cases unipolar patterning is favored at low rates of expression whereas bipolar is favored at higher rates of expression. Our findings are consistent with the predictions of the model, suggesting that localization may be due to aggregation plus nucleoid occlusion. When we expressed GFP by itself under the same conditions, no localization was observed. These experiments highlight the potential importance of protein aggregation, nucleoid occlusion and rate of protein expression in driving polar localization of functional proteins in bacteria.

**Keywords:** polar localization, protein aggregate, protein induction, nucleoid occlusion, GFP labeled

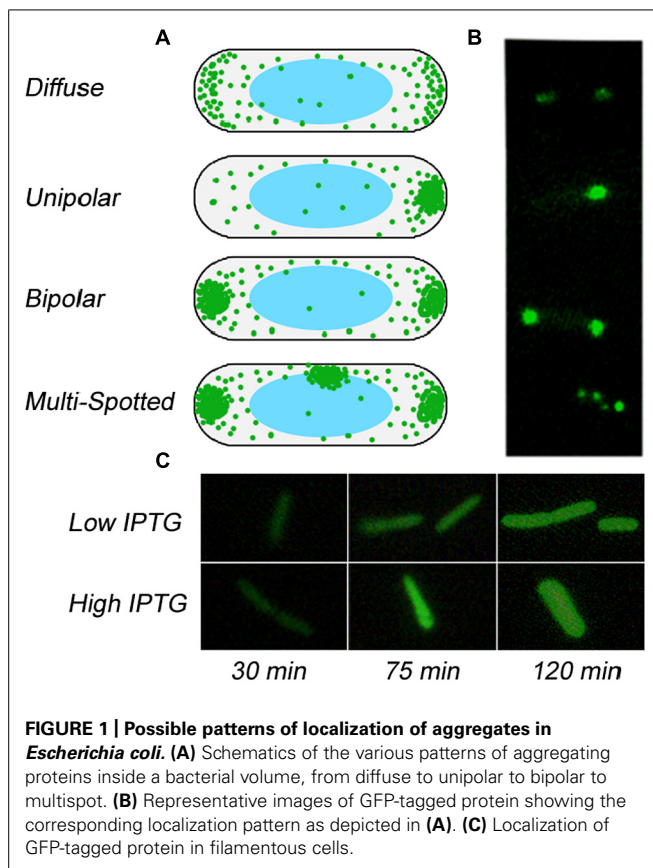
## INTRODUCTION

Many proteins are observed to localize within bacterial cells. Patterns range from forming ordered groupings of receptors on the cell membrane (Greenfield et al., 2009), to spatial waves (Hu and Lutkenhaus, 1999; Raskin and de Boer, 1999; Loose et al., 2008) to highly localized patterns at either one or both poles (Bowman et al., 2008; Ebersbach et al., 2008; Ramamurthi and Losick, 2009). Such localization has been shown in many cases to play an important role in function, whether in detecting extracellular signals (Alley et al., 1992), to guiding the dynamics of chromosome segregation (Bowman et al., 2008).

A number of these localization patterns have been shown to arise due to specific interactions of the protein with the bacterial membrane (Alley et al., 1992; Raskin and de Boer, 1999; Ramamurthi and Losick, 2009). Some of these membrane associated proteins are able to localize to regions of specific curvature through lipid mediated interactions (Ramamurthi and Losick, 2009). However some localization has been shown to arise purely within the cytoplasm, without any specific requirements for interactions with the membrane. Indeed some polar localized proteins have been shown to be driven to the poles through their aggregation and occlusion from the central portion of the cell by the nucleoid (Bowman et al., 2008; Ebersbach et al., 2008; Maisonneuve et al., 2008; Winkler et al., 2010). Some of these are functional proteins, such as PopZ that acts as a polar scaffold for many proteins in *Caulobacter crescentus* (Bowman et al., 2008; Ebersbach et al., 2008). But other aggregating proteins, such as those that are misfolded have also shown such localization through the formation of inclusion bodies (Maisonneuve et al., 2008; de

Groot et al., 2009; Winkler et al., 2010; Garcia-Fruitos et al., 2011). Recent theoretical (Saberi and Emberly, 2010, 2013; Winkler et al., 2010) and experimental (Bowman et al., 2008; Ebersbach et al., 2008; Winkler et al., 2010; Laloux and Jacobs-Wagner, 2013) work supports the hypothesis that the nucleoid can force aggregating structures within the cell to the poles. Additional levels of control, such as coupling to cell cycle associated spatial oscillations (Laloux and Jacobs-Wagner, 2013) or expression from genes that are spatially localized (Montero Llopis et al., 2010; Kuhlman and Cox, 2012) can further aid the nucleoid driven mechanism. Modeling efforts (Saberi and Emberly, 2013) highlight that the nature of the localization of the aggregate depends strongly on how fast protein is added to the cell, with slow rates leading just to a single polar aggregate whereas faster rates support multiple aggregates at both poles (see Figure 1).

A survey of the patterns of all fluorescently tagged enzymes in the *Escherichia coli* genome (Kitagawa et al., 2005) found that a number display either unipolar or bipolar patterning (222 polar localized). In this library (ASKA+), each strain has a high copy number plasmid containing an *E. coli* gene tagged with green fluorescent protein (GFP) whose expression is controlled by the *lac* promoter. The reported patterns were generated by expressing each strain with the same level of inducer, in this case IPTG. Could the polar localization of these proteins be the result of sequence specific localization cues, or potentially due to aggregation and nucleoid occlusion as well? The latter is predicted to lead to a strong rate of expression dependence on the resulting polar pattern. To explore whether aggregation plus nucleoid occlusion could be the mechanism behind a portion of the polar localization observed in



*E. coli*, we tested whether we could alter the observed patterns in a manner consistent with the predictions of theory by simply changing the rate of expression of protein (Saber and Emberly, 2013). To do this we started by selecting five, otherwise arbitrary, globular proteins that have no transmembrane domain and that showed some form of polar localization in the initial screen of the ASKA+ library. For each strain we then imaged the formation of localization patterns of the GFP-tagged protein in time at various levels of induction. The model predicts that the localization should depend strongly on how fast the protein is added. If, however, localization is specific to particular cellular locations due to sequence, then the rate of addition should not affect localization significantly; the localization should appear once the concentration crosses some threshold for the specific binding, which does not depend on rate. From our experimental analysis, we find that for each strain tested, the observed localization is indeed rate dependent that is at least consistent with the predictions of the model based on aggregation with nucleoid occlusion.

## MATERIALS AND METHODS

### SAMPLE PREPARATION

Five samples of *E. coli* K-12, strain AG1 [recA1 endA1 gyrA96 thi-1 hsdR17 (rK- mK+) supE44 relA1], that have different plasmids inserted in them were obtained from the NBRP-*E. coli* at NIG ASKA+ Library (Kitagawa et al., 2005). The plasmids contained a lac promoter and one of the genes found in *E. coli* tagged with GFP ( $\lambda_{em}$ : 509 nm,  $\lambda_{ex}$ : 395 nm). The different *E. coli* cells will be

referred to by the name of the gene on their plasmid (*glnQ*, *pykA*, *cysJ*, *aceA*, and *pfkA*). Cells were grown at 37°C for 17 h in Luria-Bertani (LB) medium [1% Bio Tryptone (Bioshop), 0.5% yeast extract (EMD), 1% NaCl] containing 14  $\mu$ g/ml Chloramphenicol (BDH Chemicals) for selection. The cells were then diluted 1:50 and grown for 3 h to be well into log phase. Isopropyl  $\beta$  D thiogalactoside (IPTG) was then used in varying concentrations (0.025, 0.05, 0.075, 0.1, and 0.25 mM) to induce expression of the cloned gene.

### FLUORESCENT MICROSCOPY AND IMAGE ANALYSIS

After induction, bright field and fluorescent microscopy were performed (Zeiss Axioskop with Photometrics CF) for 2 h in 15 min intervals. About 60 cells were photographed at each time step. Afterward the GFP intensities of imaged cells were quantified and the resulting pattern was classified by eye as either non-fluorescing, diffuse, unipolar, bipolar, or multi-spotted (see Figure 1B and Figures S1–S3). Bipolar patterns clearly show a circular GFP spot at each pole (see Figure S1B), whereas at low expression rates some diffuse patterns show some increased crescent shaped GFP fluorescence at both poles that is easily distinguishable from a bipolar pattern based on both spatial and significant intensity differences (see Figure S1B and Figure S2). Filamentous *E. coli* cells were not included in the count. Cells that appeared in the fluorescent image but not the bright field image were also not included in the classification of patterning. At each time point the fraction of each localization pattern within the total number of characterized cells was evaluated.

## RESULTS

Our model for polar localization argues that protein aggregation along with nucleoid occlusion can cause an aggregate to form at one or both poles depending on how fast the aggregating protein is added to the cell (Saber and Emberly, 2013). Motivated by this model, we explored whether it may be at work in the patterning of polar localizing tagged proteins in the ASKA+ *E. coli* library (Kitagawa et al., 2005). We selected five genes (*glnQ*, *pykA*, *cysJ*, *aceA*, and *pfkA*) from this library whose proteins showed polar localization to see if we could alter their localization patterns by changing the rate of addition. The proteins were selected because they were globular, had no transmembrane domain and showed polar localization in the reported images in the initial screen of the ASKA+ library.

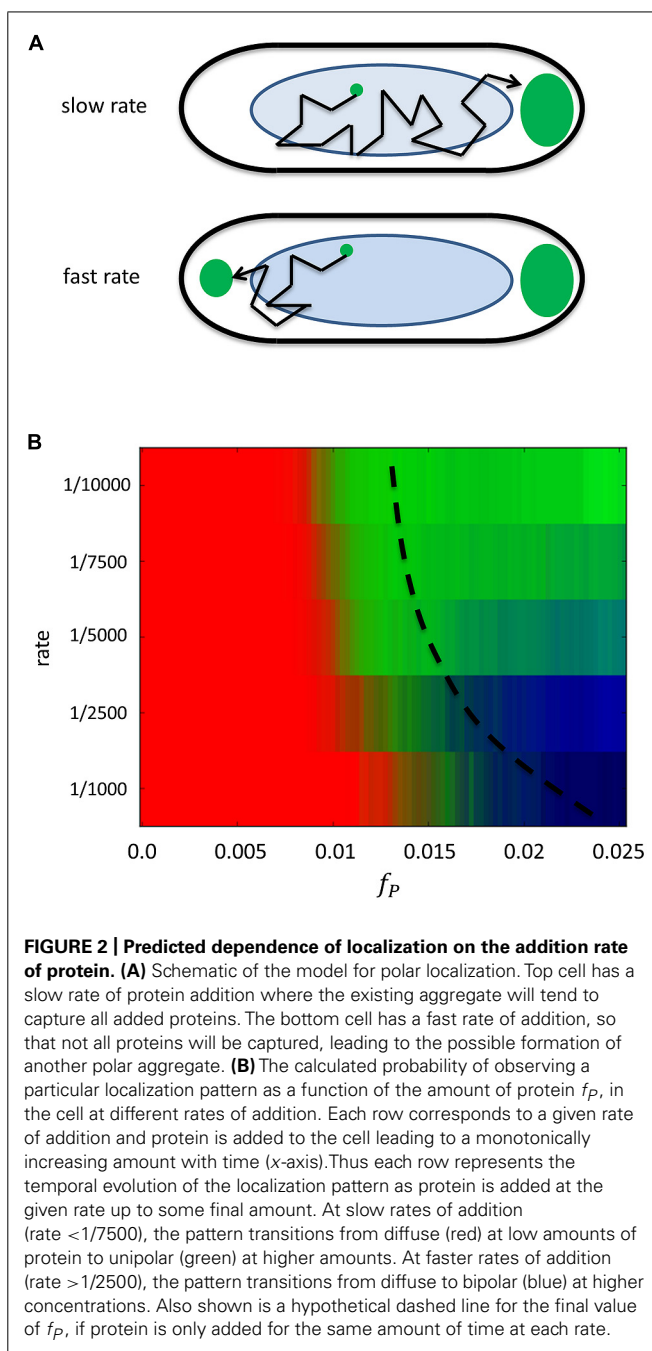
Since each strain could potentially possess different numbers of plasmids and therefore express protein at different rates under otherwise identical conditions, we chose to grow up each strain at several induction levels. (In Figure S3 we show the increase of GFP intensity with time at different levels of induction, showing how higher amounts of IPTG lead to higher expression rates). This would allow us to make correspondences between the strains in terms of expression rates. Each strain was grown up on six different levels of IPTG (0.025, 0.05, 0.075, 0.1, and 0.25 mM) leading to a different rates of expression for each GFP tagged protein. The cells were then imaged every 15 min, with the patterns of each imaged cell recorded into one of four patterns (see Figures S1, S2): diffuse, unipolar, bipolar, and multi-spot (see Figure 1B, Figure S1A). For each strain, GFP starts to come on 20 min after induction, and

so we have very few fluorescing cells at the earliest time points, making it a challenge to fully characterize the nature of the pattern at times less than 30 min. Only one of the five strains, *cysJ*, showed considerably lower levels of GFP induction, and hence less cells to characterize for patterning and was left out of the analysis that follows.

The model predicts that at slow rates of addition, only one polar aggregate should form, since all added protein will have time to diffuse and eventually get captured by the growing aggregate. This is shown in **Figure 2B**, which plots how the pattern changes as protein is added to the cell in time at a given rate of expression. At low rates, as time progresses and the amount of protein  $f_P$ , increases in the cell, the pattern transitions from being predominantly diffuse (red) to unipolar (green). At higher rates of expression, as more and more protein is added to the cell the pattern goes from being diffuse to a brief period of being unipolar and finally to bipolar (dark blue). The localization behavior at both low and high rates of expression can be understood in terms of a diffusion-to-capture model that has been used to explain receptor localization on the bacterial membrane (Wang et al., 2008; Greenfield et al., 2009), except here in the context of the cytoplasm (see **Figure 2A**). At low rates of addition added proteins have enough time to be captured by a lone aggregate before the next protein is added, causing it to continue to grow. At faster rates, diffusing proteins do not have time to cover the full length of the cell and eventually reach densities that are sufficient to start a new aggregate. It is entropically and energetically favorable for aggregates to migrate to the poles due to the entropic force exerted on them by the presence of the bacterial nucleoid and the ability to grow to larger sizes in the polar regions where there is less DNA.

For each strain we could find a concentration of IPTG that lead to a rate of expression of the GFP tagged protein such that unipolar patterning dominates the population at later times (see **Figure S4**). At early time points the pattern is diffuse and transitions to being predominantly unipolar  $\sim 80\%$  at the final time point. At these low induction concentrations, some bipolar patterning exists, but at very low levels in all strains. For the strain expressing CysJ, similar behavior was observed, though the statistics are based on only a few fluorescing cells (data not shown).

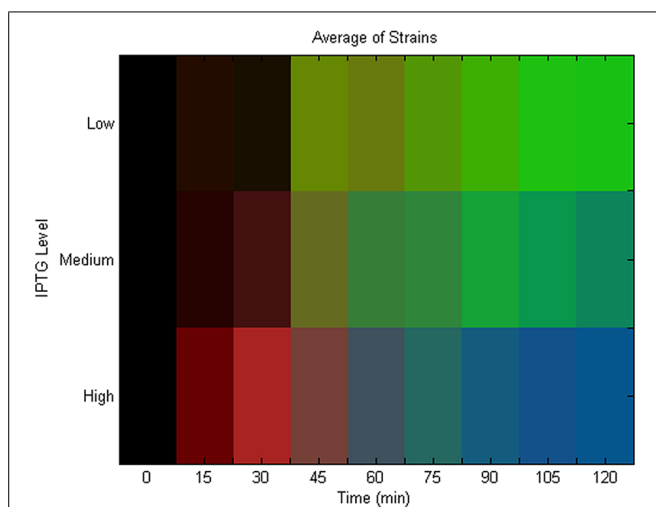
As the rate of addition of protein is increased by increasing the concentration of IPTG, the model predicts that patterning should move from being predominantly unipolar to that of bipolar (or possibly multi-spot) as (see **Figure 2B** at the higher rates). This should occur at a rate that goes as  $r \sim D/L^2$  where  $D$  is the diffusion coefficient of the soluble protein and  $L$  is the length of the bacterial cell. [For GFP in *E. coli*,  $D = 7.7 \mu\text{m}^2/\text{s}$  (Elowitz et al., 1999) and  $L = 2\text{--}3 \mu\text{m}$ , so  $r = 0.9\text{--}2 \text{ GFP/s}$  as an upper estimate]. At rates faster than this, the pattern transitions from unipolar to bipolar since another aggregate can form in the cell, moving to the other pole. To test whether this would happen in the chosen set of GFP-tagged proteins, we grew up each strain in increasingly higher levels of IPTG. For each strain we found that at a particular higher level of IPTG, and hence a faster rate of addition (see **Figure S3**), bipolar patterning dominates at longer times (see **Figure S5**). Thus adding protein at a faster rate leads to the formation of another aggregate that is then occluded by the nucleoid to the other pole.



**FIGURE 2 | Predicted dependence of localization on the addition rate of protein. (A)** Schematic of the model for polar localization. Top cell has a slow rate of protein addition where the existing aggregate will tend to capture all added proteins. The bottom cell has a fast rate of addition, so that not all proteins will be captured, leading to the possible formation of another polar aggregate. **(B)** The calculated probability of observing a particular localization pattern as a function of the amount of protein  $f_P$ , in the cell at different rates of addition. Each row corresponds to a given rate of addition and protein is added to the cell leading to a monotonically increasing amount with time (x-axis). Thus each row represents the temporal evolution of the localization pattern as protein is added at the given rate up to some final amount. At slow rates of addition (rate  $< 1/7500$ ), the pattern transitions from diffuse (red) at low amounts of protein to unipolar (green) at higher amounts. At faster rates of addition (rate  $> 1/2500$ ), the pattern transitions from diffuse to bipolar (blue) at higher concentrations. Also shown is a hypothetical dashed line for the final value of  $f_P$ , if protein is only added for the same amount of time at each rate.

We summarize our experimental findings in **Figure 3** that shows the average localization pattern over the population for all strains at different IPTG levels versus time. This shows that the localization behavior of these GFP tagged polar localizing proteins depends on the rate of addition. This finding is consistent with the predictions of the model (see **Figure 2B**), namely that unipolar patterning dominates at slow rates of addition, transitioning between a mix of unipolar and bipolar at intermediate rates to finally bipolar at the highest levels of induction. We note that we do not detect much of a diffuse phase at early time points at low rates of expression. As we previously mentioned, characterization of the





**FIGURE 3 | Average localization pattern versus time and induction rate.** Population and strain averaged localization pattern versus time at different induction levels. The RGB color represents the mix of patterns at each time point from diffuse (red), unipolar (green) to bipolar (blue). For each strain, an induction level was found that lead to unipolar patterning at late times ("Low"; see Figure S4) and also bipolar ("High"; see Figure S5). The IPTG concentration in between the "Low" and "High" concentration for each strain was selected for the midpoint expression level ("Medium"). The average localization pattern over all strains was then calculated for each of these induction levels. The population average for diffuse (red), unipolar (green) and bipolar (blue) is then plotted as an RGB level at each time point to make the heat map, and where black represents time points that had no data.

pattern at early time points is compromised due to the slow folding time of GFP. **Figure 3** is also an average over all strains, and looking at each strain individually (Figure S4), *pykA* shows the diffuse pattern being present at early time, and similarly for the other strains if the diffuse data is extrapolated to earlier times. As a control, we expressed GFP alone under the same conditions – here at 0.025 and 0.25 mM (see **Figure 1C**). Under both levels of expression, no localization could be detected. Thus for the five strains tested there is some form of specific interaction that exists between the tagged proteins leading to their aggregation that is either inherent to the native protein or arises due to the tagging. This aggregation in the presence of the bacterial nucleoid leads to their localization to the poles.

## DISCUSSION

Protein aggregation leading to polar localization within the cytoplasm of bacteria has many functional consequences, from regulating signaling, to tethering chromosomes, to segregating misfolded proteins to guard against potential deleterious consequences. Such localization could be targeted by spatially specific cues (Montero Llopis et al., 2010; Kuhlman and Cox, 2012) or aided by actively driven processes as recently shown for the polar localized scaffold protein PopZ (Laloux and Jacobs-Wagner, 2013). Modeling efforts have shown that nucleoid occlusion in addition to protein aggregation can be a sufficient mechanism to drive the spontaneous formation of polar localization. Our theoretical modeling has shown that such polar should have strong dependence on the

rate at which protein is added to the cell. To explore this prediction we selected five cytoplasmic proteins that showed polar localization when tagged with GFP, and expressed them at different rates within the cell. The mRNA is expressed off of a high-copy number plasmid and so the resulting GFP-tagged proteins should be produced uniformly within the cell. In all cases we found that there was a strong dependency on the rate of protein addition consistent with the predictions of the model of nucleoid occlusion plus protein aggregation.

It is possible that the interactions leading these polar localized proteins to aggregate is native to these proteins and is essential for them to target either one or both poles. Indeed, as our experiments show, the selection of unipolar versus bipolar localization can be selected for by tuning the rate at which the protein is expressed in the cell. In principle the bacteria could tune such expression levels to select for particular polar patterns. One prediction is that one might see rate dependent polar patterning for constitutively expressed genes in bacteria that functionally aggregate. Examining the localization data obtained in the recent library of chromosomal YFP tagged genes in *E. coli* would give a potential test of this prediction (Taniguchi et al., 2010). Another possibility that could explain some of the polar localization is that it results due to tagging with GFP as recently highlighted (Landgraf et al., 2012). The tagging by GFP for certain proteins leads to their aggregation and formation of inclusion bodies that then can be driven to the poles via the mechanisms put forward here. Our own findings for expressing GFP alone showed no preference for localizing to the poles. Nevertheless, further work using antibody staining or other fluorescent tags would help to further clarify whether the localization is the wild-type pattern or not. It should also be noted that the theoretical modeling revealed that initial conditions play a strong influence on the dynamics of the resulting pattern. Had there been a pre-existing pattern present due to the wild-type protein, once the tagged protein is introduced it would then get quickly incorporated to what is already there. Our experimental findings show an emergence of localization patterns for the GFP tagged protein arguing against there be any strong initial conditions. Categorization and functional testing of these polar localized proteins will give further clues into the biological significance of using nucleoid occlusion and protein aggregation as a method for localization.

## ACKNOWLEDGMENTS

We thank Nancy Forde and Andrew Wiczorek for helpful discussions in carrying out experiments. Eldon Emberly would like to acknowledge NSERC for supporting this research.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://www.frontiersin.org/journal/10.3389/fmicb.2014.00418/abstract>

## REFERENCES

- Alley, M. R., Maddock, J. R., and Shapiro, L. (1992). Polar localization of a bacterial chemoreceptor. *Genes Dev.* 6, 825–836. doi: 10.1101/gad.6.5.825
- Bowman, G. R., Comolli, L. R., Zhu, J., Eckart, M., Koenig, M., Downing, K. H., et al. (2008). A polymeric protein anchors the chromosomal origin/ParB complex at a bacterial cell pole. *Cell* 134, 945–955. doi: 10.1016/j.cell.2008.07.015



- de Groot, N. S., Sabate, R., and Ventura, S. (2009). Amyloids in bacterial inclusion bodies. *Trends Biochem. Sci.* 34, 408–416. doi: 10.1016/j.tibs.2009.03.009
- Ebersbach, G., Briegel, A., Jensen, G. J., and Jacobs-Wagner, C. (2008). A self-associating protein critical for chromosome attachment, division, and polar organization in *caulobacter*. *Cell* 134, 956–968. doi: 10.1016/j.cell.2008.07.016
- Elowitz, M. B., Surette, M. G., Wolf, P. E., Stock, J. B., and Leibler, S. (1999). Protein mobility in the cytoplasm of *Escherichia coli*. *J. Bacteriol.* 181, 197–203.
- Garcia-Fruitos, E., Sabate, R., de Groot, N. S., Villaverde, A., and Ventura, S. (2011). Biological role of bacterial inclusion bodies: a model for amyloid aggregation. *FEBS J.* 278, 2419–2427. doi: 10.1111/j.1742-4658.2011.08165.x
- Greenfield, D., McEvoy, A. L., Shroff, H., Crooks, G. E., Wingreen, N. S., Betzig, E., et al. (2009). Self-organization of the *Escherichia coli* chemotaxis network imaged with super-resolution light microscopy. *PLoS Biol.* 7:e1000137. doi: 10.1371/journal.pbio.1000137
- Hu, Z., and Lutkenhaus, J. (1999). Topological regulation of cell division in *Escherichia coli* involves rapid pole to pole oscillation of the division inhibitor MinC under the control of MinD and MinE. *Mol. Microbiol.* 34, 82–90. doi: 10.1046/j.1365-2958.1999.01575.x
- Kitagawa, M., Ara, T., Arifuzzaman, M., Ioka-Nakamichi, T., Inamoto, E., Toyonaga, H., et al. (2005). Complete set of ORF clones of *Escherichia coli* ASKA library (a complete set of *E. coli* K-12 ORF archive): unique resources for biological research. *DNA Res.* 12, 291–299. doi: 10.1093/dnares/dsi012
- Kuhlman, T. E., and Cox, E. C. (2012). Gene location and DNA density determine transcription factor distributions in *Escherichia coli*. *Mol. Syst. Biol.* 8, 610. doi: 10.1038/msb.2012.42
- Laloux, G., and Jacobs-Wagner, C. (2013). Spatiotemporal control of PopZ localization through cell cycle-coupled multimerization. *J. Cell Biol.* 201, 827–841. doi: 10.1083/jcb.201303036
- Landgraf, D., Okumus, B., Chien, P., Baker, T. A., and Paulsson, J. (2012). Segregation of molecules at cell division reveals native protein localization. *Nat. Methods* 9, 480–482. doi: 10.1038/nmeth.1955
- Loose, M., Fischer-Friedrich, E., Ries, J., Kruse, K., and Schwill, P. (2008). Spatial regulators for bacterial cell division self-organize into surface waves in vitro. *Science* 320, 789–792. doi: 10.1126/science.1154413
- Maisonneuve, E., Frayse, L., Moinier, D., and Dukan, S. (2008). Existence of abnormal protein aggregates in healthy *E. coli* cells. *J. Bacteriol.* 190, 887–893. doi: 10.1128/JB.01603-07
- Montero Llopis, P., Jackson, A. F., Sliusarenko, O., Surovtsev, I., Heinritz, J., Emonet, T., et al. (2010). Spatial organization of the flow of genetic information in bacteria. *Nature* 466, 77–81. doi: 10.1038/nature09152
- Ramamurthi, K. S., and Losick, R. (2009). Negative membrane curvature as a cue for subcellular localization of a bacterial protein. *Proc. Natl. Acad. Sci. U.S.A.* 106, 13541–13545. doi: 10.1073/pnas.0906851106
- Raskin, D. M., and de Boer, P. A. (1999). Rapid pole-to-pole oscillation of a protein required for directing division to the middle of *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* 96, 4971–4976. doi: 10.1073/pnas.96.9.4971
- Saberi, S., and Emberly, E. (2010). Chromosome driven spatial patterning of proteins in bacteria. *PLoS Comput. Biol.* 6:e1000986. doi: 10.1371/journal.pcbi.1000986
- Saberi, S., and Emberly, E. (2013). Non-equilibrium polar localization of proteins in bacterial cells. *PLoS ONE* 8:e64075. doi: 10.1371/journal.pone.0064075
- Taniguchi, Y., Choi, P. J., Li, G. W., Chen, H., Babu, M., Hearn, J., et al. (2010). Quantifying *E. coli* proteome and transcriptome with single-molecule sensitivity in single cells. *Science* 329, 533–538. doi: 10.1126/science.1188308
- Wang, H., Wingreen, N. S., and Mukhopadhyay, R. (2008). Self-organized periodicity of protein clusters in growing bacteria. *Phys. Rev. Lett.* 101, 218101. doi: 10.1103/PhysRevLett.101.218101
- Winkler, J., Seybert, A., König, L., Pruggnaller, S., Haselmann, U., Sourjik, V., et al. (2010). Quantitative and spatio-temporal features of protein aggregation in *E. coli* and consequences on protein quality control and cellular ageing. *EMBO J.* 29, 910–923. doi: 10.1038/emboj.2009.412

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 01 May 2014; accepted: 22 July 2014; published online: 06 August 2014.

Citation: Scheu K, Gill R, Saberi S, Meyer P and Emberly E (2014) Localization of aggregating proteins in bacteria depends on the rate of addition. *Front. Microbiol.* 5:418. doi: 10.3389/fmicb.2014.00418

This article was submitted to *Microbial Physiology and Metabolism*, a section of the journal *Frontiers in Microbiology*.

Copyright © 2014 Scheu, Gill, Saberi, Meyer and Emberly. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Refolding and purification of recombinant L-asparaginase from inclusion bodies of *E. coli* into active tetrameric protein

Arun K. Upadhyay<sup>1</sup>, Anupam Singh<sup>1</sup>, K. J. Mukherjee<sup>2</sup> and Amulya K. Panda<sup>1\*</sup>

<sup>1</sup> Product Development Cell, National Institute of Immunology, New Delhi, India

<sup>2</sup> School for Biotechnology, Jawaharlal Nehru University, New Delhi, India

## Edited by:

Salvador Ventura, Universitat Autònoma de Barcelona, Spain

## Reviewed by:

Dirk Linke, Max Planck Society, Germany

Elena García-Fruitós, Centro de Investigación Biomédica en Red, Spain

## \*Correspondence:

Amulya K. Panda, Product Development Cell, National Institute of Immunology, Aruna Asaf Ali Marg, New Delhi-110067, India  
e-mail: amulya@nii.res.in

A tetrameric protein of therapeutic importance, *Escherichia coli* L-asparaginase-II was expressed in *Escherichia coli* as inclusion bodies (IBs). Asparaginase IBs were solubilized using low concentration of urea and refolded into active tetrameric protein using pulsatile dilution method. Refolded asparaginase was purified in two steps by ion-exchange and gel filtration chromatographic techniques. The recovery of bioactive asparaginase from IBs was around 50%. The melting temperature ( $T_m$ ) of the purified asparaginase was found to be 64°C. The specific activity of refolded, purified asparaginase was found to be comparable to the commercial asparaginase (190 IU/mg). Enzymatic activity of the refolded asparaginase was high even at four molar urea solutions, where the IB aggregates are completely solubilized. From the comparison of chemical denaturation data and activity at different concentrations of guanidine hydrochloride, it was observed that dissociation of monomeric units precedes the complete loss of helical secondary structures. Protection of the existing native-like protein structure during solubilization of IB aggregates with 4 M urea improved the propensity of monomer units to form oligomeric structure. Our mild solubilization technique retaining native-like structures, improved recovery of asparaginase in bioactive tetrameric form.

**Keywords:** L-asparaginase-II, *Escherichia coli*, IBs, mild solubilization, refolding

## INTRODUCTION

Most of the times, expression of recombinant proteins in *Escherichia coli* leads to the formation of insoluble aggregates known as IBs (Hartley and Kane, 1988; Fahnert et al., 2004). Recovery of active protein from IB aggregates remains to be a cumbersome task and requires standardization of solubilization and refolding methods (De Bernardez et al., 1999; Burgess, 2009). The major hurdle associated with purification of proteins from IBs is the sub-optimal refolding of recombinant proteins into native conformation (Rudolph and Lilie, 1996; Panda, 2003; Vallejo and Rinas, 2004). Poor refolding is often associated to high concentrations of urea or guanidine hydrochloride (GdmCl) used to solubilize the IB proteins. At higher concentrations, chaotropes such as urea and GdmCl completely denature the proteins and increases its propensity to aggregate during refolding resulting in low recovery of bioactive protein from IBs. IB proteins are reported to have structure and functional activities (Umetsu et al., 2004; Ventura and Villaverde, 2006; Peternel and Komel, 2011). These active IBs can be isolated from bacterial cells using different methods like homogenization, enzymatic lysis, and sonication where homogenization was observed to be most appropriate (Peternel and Komel, 2010). It will be ideal to protect these secondary protein structures during IB solubilization process. Mild solubilization of IB aggregates protects the existing native-like protein structure of IB proteins and helps in its improved recovery

into bioactive form Singh and Panda (2005). In a few cases, solubilization with mild denaturing conditions has been proved to be more efficient for the recovery of bioactive protein from the IBs (Panda, 2003; Singh et al., 2012; Upadhyay et al., 2012).

Refolding yield of active oligomeric proteins from IBs is even lower (Scrofani et al., 2000; Karumuri et al., 2007; Garrido et al., 2011). Formation of active monomer and its association is a prerequisite for refolding into fully active oligomeric proteins. Often it is hindered due to complete unfolding of proteins in IBs into random coil structure while using high concentration of chaotropes. Solubilized protein molecules have propensity to form intermolecular aggregates leading to massive aggregation during refolding. It is thus essential to protect the secondary helical structure of IB proteins so that it reduces intermolecular aggregation between monomers. It can be achieved by adopting mild solubilization process for solubilization of IBs. Chemical denaturation studies provide information about the solubility profile of IB aggregates. Based on this information, IBs can be solubilized at low denaturation concentration while protecting the existing native-like secondary protein structure. Refolding of protein from monomers having secondary structural element will promote monomer association leading to the formation of active oligomeric protein and thus will improve the overall recovery of bioactive protein from IBs. Even though mild solubilization processes have been used to

recover bioactive protein from IBs, there is very little information available on the refolding of oligomeric proteins into bioactive form.

Bacterial asparaginases from *Escherichia coli* and *Erwinia chrysanthemi* have been extensively used as drugs for the treatment of acute lymphoblastic leukemia (Muller and Boos, 1998; Graham, 2003; Verma et al., 2007). L-asparaginases (EC 3.5.1.1) catalyze the hydrolysis of L-asparagine to L-aspartic acid and ammonia. All asparaginases consist of four identical subunits A, B, C, and D and exist in homo-tetrameric form (Kozak et al., 2002) having masses in the range of 140–150 kDa (Aung et al., 2000). One subunit consists of two  $\alpha/\beta$  domains that are connected by linking sequence. Interaction between N and C-terminal domain of adjacent monomers forms each active site. Therefore, the asparaginase tetramer can be treated as dimer of dimers because active site is either created by subunits A and C or B and D. The active form of the enzyme is a tetramer, and the dimers lack enzyme activity (Swain et al., 1993). L-asparaginase contains one tryptophan molecule at 66 positions in each monomer. It consists of four active sites formed at the interfaces of N and C-terminal domains of two interacting monomers (Swain et al., 1993). L-asparaginase has been reported to be produced using recombinant *E. coli* (Khushoo et al., 2004; Oza et al., 2011), *P. pastoris* (Ferrara et al., 2010) and from other microorganisms (Mahajan et al., 2012). Most of the times, the enzymes is produced in soluble form and purified with 40–60% recovery. There are no reports till date on solubilization and refolding of L-asparaginase from IBs into bioactive tetrameric form.

In this work, *E. coli* L-asparaginase II was expressed in *E. coli* cells as IBs and used as a model system to refold into bioactive tetrameric form. Attempts were made to solubilize the IBs of asparaginase in mild denaturing conditions. Solubilized proteins were then refolded and purified into active oligomeric form. Role of structural integrity between monomers and their importance in regulation of enzyme activity and stability was monitored by chemical, pH, temperature and organic solvent based denaturation methods. The results are of indication that mild solubilization followed by pulsatile refolding leads to improved recovery of bioactive multimeric L-asparaginase from IBs of *E. coli*.

## MATERIALS AND METHODS

### CHEMICALS

Culture media ingredients, tryptone, and yeast extract were from Difco Laboratories, India. Tris buffer, glycine, IPTG, sodium dodecyl sulfate, PMSF, and deoxycholic acid were from Amresco, USA. Ammonium persulfate, acrylamide, bis-acrylamide, *E. coli* asparaginase, L-arginine and Urea were from Sigma Chemicals, USA. DEAE-sepharose and S-200 gel filtration matrix was from GE Healthcare, Sweden. TEMED, EDTA, bromophenol blue from Biorad, USA. Coomassie brilliant blue R-250 and ampicillin from USB Corporation, Cleveland, Ohio. Glucose, NaCl, Nessler's reagent, and other chemicals were from Qualigen, India.

### ESCHERICHIA COLI STRAINS AND CLONING OF L-ASPARAGINASE

*Escherichia coli* BL21 (DE3) strain was from Novagen, USA. *E. coli* DH5 $\alpha$  strain was obtained from Amersham Bioscience, USA. Plasmid pET14b was from Novagen, USA.

L-asparaginase II (ansB) gene was amplified from the genomic DNA of *E. coli* K-12 strain (JM109) using primers (forward) 5'-GTGCAGCACATATGTTACCCAATA TCACCA 3' and (reverse) 5'-GGCGGGATCCTTAGTACTGATTGAAGA 3'. NdeI and BamHI restriction sites were incorporated in the primers to facilitate cloning of the structural asparaginase gene (without its native signal sequence) in the *E. coli* expression vector pET14b. *E. coli* BL21 (DE3) cells were transformed with recombinant pET14b plasmid vector and used for expression of L-asparaginase II.

### EXPRESSION AND ISOLATION OF ASPARAGINASE IBs

Transformed *E. coli* BL21 (DE3) cells were grown in modified LB media (1% tryptone, 0.5% yeast extract, 0.5% glucose and 1% NaCl) with 100  $\mu$ g/ml ampicillin at 37°C in incubator shaker at 200 rpm. Culture at an OD<sub>600</sub> of 0.6 was induced with 1 mM IPTG and was further grown for 3.5 h. Cells (1 L culture) were pelleted at 6000 rpm (rotor SLA-3000, Sorvall evolution RC) for 10 min at 4°C. Supernatant was discarded and culture pellet was resuspended in 20 ml buffer I (50 mM Tris-HCl, 1 mM EDTA, 1 mM PMSF, pH 8.5) containing egg white lysozyme (1 mg/ml) and incubated at room temperature for 2 h. Re-suspended cells were sonicated for 10 cycle (1 min cycle with 1 min gap) using Branson sonifier 450, Germany (probe diameter 13 mm, output voltage 60 and 50% duty cycle). Sonicated cells were centrifuged at 12,000 rpm (SA-300 rotor, Sorvall evolution RC) for 20 min at 4°C and the pellet containing recombinant asparaginase as IBs was separated. These pellets were resuspended in 20 ml buffer I containing 0.1% (w/v) deoxycholic acid, sonicated and centrifuged as described earlier (Patra et al., 2000). Pellet was further washed two times with buffer I and once with de-ionized water. Washed pellet (purified IBs) was resuspended in 1 ml 50 mM Tris-HCl buffer, pH 8.5 and used for solubilization and refolding.

### SOLUBILIZATION AND REFOLDING OF ASPARAGINASE FROM IBs

Isolated asparaginase IBs (1 ml) were solubilized in 9 ml buffer II (4 M urea, 50 mM Tris-HCl, 1 mM PMSF, 20 mM  $\beta$ -mercaptoethanol, 10 mM NaCl, pH 8.5) and kept for one h at room temperature. Solubilized asparaginase was centrifuged at 15,000 rpm for 30 min and supernatant was filtered through 0.2  $\mu$ m filter (Millipore, USA). Solubilized protein was refolded into 90 ml buffer III (0.5 M urea, 50 mM Tris-HCl, 0.1 M arginine, 10 mM NaCl) by pulsatile dilution method at flow rate of 0.1 ml/min at 4°C. Refolded sample was centrifuged at 15,000 rpm (rotor SA-300, Sorvall evolution RC) for 30 min at 4°C. Supernatant was collected and dialyzed against buffer (50 mM Tris-HCl, 0.5 M urea, pH 8.5) using dialysis tubing (10 kDa, molecular weight cut-off) for 4 h at 4°C. Buffer was exchanged three times with 50 mM Tris-HCl, pH 8.5, at 4 h interval. Dialyzed protein was pooled and used for purification.

### PURIFICATION OF REFOLDED RECOMBINANT L-ASPARAGINASE

Refolded L-asparaginase was purified using DEAE-sepharose anion exchange matrix packed in XK 16 column (GE Healthcare, Sweden). Refolded asparaginase was loaded on to the column

at flow rate of 1 ml/min and column was washed with 3 bed volumes of 50 mM Tris-HCl buffer at pH 8.5. Recombinant asparaginase was eluted from column using 0 to 0.5 M continuous gradient of NaCl. Fractions (2 ml each) were collected and analyzed by 12% SDS polyacrylamide gel electrophoresis. Ion exchange fractions containing asparaginase were pooled and concentrated by ultra-filtration (10 kDa molecular weight cut-off) to the final volume 5 ml. Concentrated asparaginase was loaded on S-200 Sephacryl column (XK16/60, GE Healthcare, Sweden) pre-equilibrated with 50 mM Tris-HCl, 10 mM NaCl, pH 7.5 buffer at flow rate 0.5 ml/min. Eluted tetramer fractions were pooled, concentrated and used for activity assay and characterization. All chromatography experiments were carried out using AKTA protein purifier (GE Healthcare, Sweden). SDS polyacrylamide gel electrophoresis was performed according to method of Laemmli on a slab gel containing 12% running gel and 5% stacking gel (Laemmli, 1970).

### PROTEIN ESTIMATION AND MASS SPECTROMETRY

Protein concentration was determined by micro BCA assay kit (Pierce, Germany) using BSA as a standard. For enzyme assay and spectral analysis purified protein concentration was calculated by spectroscopic method using  $\lambda_{280}$  nm ( $1 \text{ mg/ml} = 0.77$  for purified asparaginase). Purified recombinant asparaginase was passed through desalting column to remove salts and other small molecules. Recombinant asparaginase ( $10 \mu\text{g/ml}$ ) dissolved in 1% formic acid was used for mass spectroscopy analysis. Mass spectrum was collected using LC MS MS system (Waters, USA).

### ANALYTICAL GEL FILTRATION CHROMATOGRAPHY USING HPLC SYSTEM

Purified recombinant asparaginase was loaded on protein Bio-Sep S-2000 column (Phenomenex, Torrance, USA) attached to HPLC system (Shimadzu, Japan) to analyze oligomeric form of the recombinant protein. For this, a Bio-Sep S-2000 column was equilibrated with 50 mM Tris-HCl, 10 mM NaCl, pH 7.5 and  $20 \mu\text{l}$  of  $1.5 \text{ mg/ml}$  purified recombinant asparaginase solution was injected in the column. Equilibration and elution was carried out at a flow rate of 1 ml/min. Gel filtration marker proteins (GE Healthcare, Sweden) were applied on the column for estimation of molecular mass of recombinant L-asparaginase.

### ASPARAGINASE ACTIVITY ASSAY

Asparaginase activity was assayed using method described by Wriston (1985). Briefly, reaction mixture consisted of 50 mM Tris-HCl (pH 8.6) and 8.6 mM L-asparagine incubated at  $37^\circ\text{C}$  for 10 min. L-asparaginase enzyme solution ( $10 \mu\text{g/ml}$ ) was added in reaction mixture and incubated at  $37^\circ\text{C}$ . Reaction was stopped by adding 1.5 M trichloro acetic acid at different time points and samples were centrifuged and used for estimation of released ammonia by Nessler's reagent using ammonium sulfate as standard. Time dependent release of ammonia was determined by taking OD measurements at 432 nm and linear range of ammonia release was found to be up to 30 min. An international unit (IU) of L-asparaginase is defined as the amount of enzyme required to release one micromole of ammonia per min under

the condition of the assay at saturating substrate concentration. Assays were run in triplicate.

### FLUORESCENCE AND FAR-UV CD SPECTROSCOPY OF PURIFIED ASPARAGINASE

Fluorescence emission spectra of purified recombinant asparaginase were recorded using the Cary Eclipse spectrophotometer (Varian, Australia) attached with Peltier temperature controller. Spectra were recorded on 1 cm path length cuvettes with excitation and emission slit width 5 nm.  $25 \mu\text{g/ml}$  concentrations were used for acquisition of fluorescence emission spectra. Samples were excited at 280 nm and emission spectra were collected from 290 to 400 nm. Far-UV CD spectra were recorded using Jasco-spectropolarimeter equipped with a Peltier temperature controller. Spectra were acquired with a bandwidth of 1 nm, a step size of 1 nm, and an accumulation of 100 nm per min for three scans. Protein concentration used for CD spectroscopy was  $0.4 \text{ mg/ml}$ . Asparaginase was denatured in different concentrations of GdmCl. Melting curves were recorded at 222 nm during sample heating and consecutive cooling from 20 to  $90^\circ\text{C}$ , with  $5^\circ\text{C}$  increment in base region and at  $1^\circ\text{C}$  in transition region of thermal denaturation.

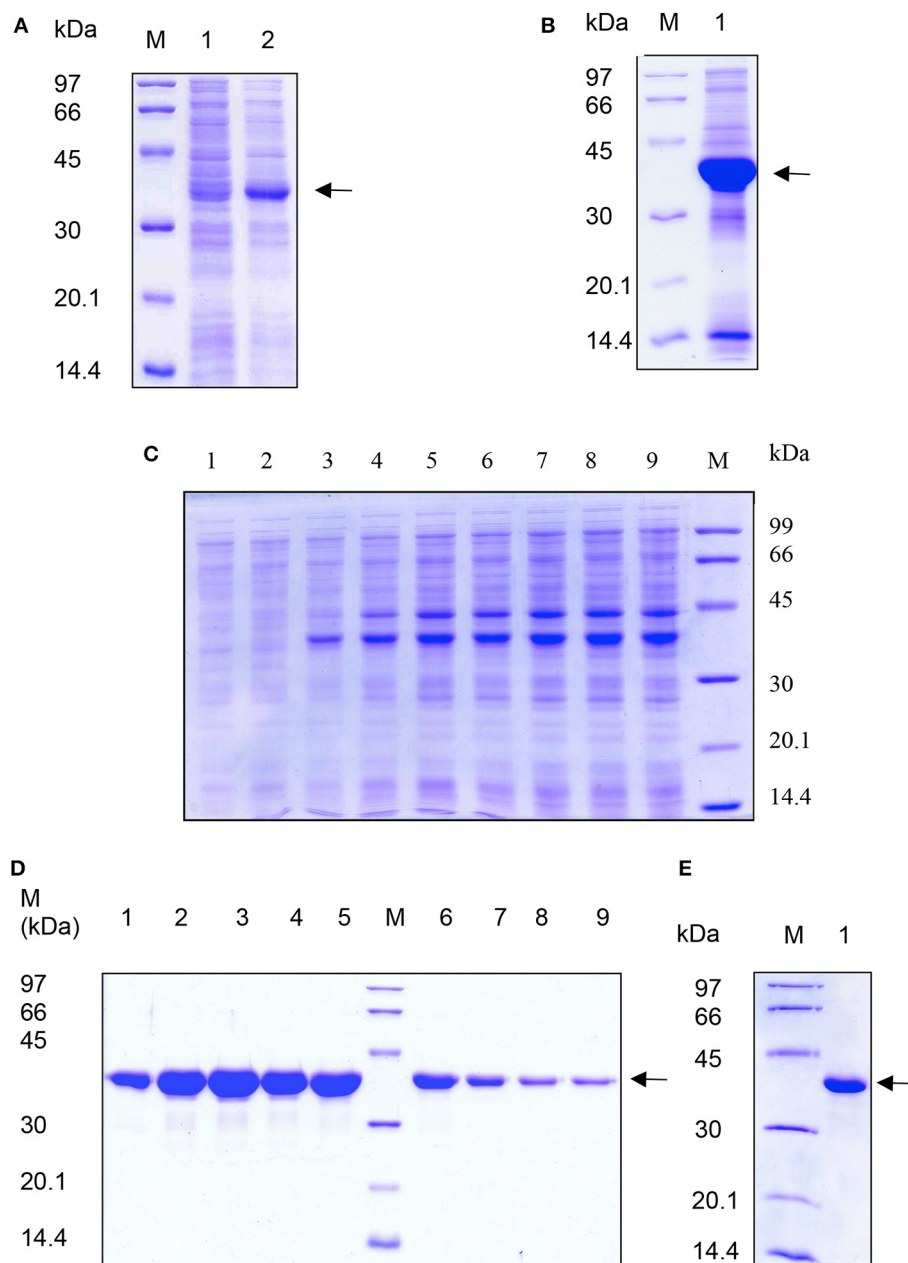
### DETERMINATION OF CONFORMATIONAL STABILITY OF L-ASPARAGINASE

Stability of the active asparaginase tetramer was monitored by measurements of its tryptophan fluorescence at various concentrations of denaturants. Purified asparaginase was denatured in different concentrations of urea and GdmCl. In all the denaturation studies, final concentration of asparaginase in denaturing buffer was  $10 \mu\text{g/ml}$ . Fluorescence emission spectra were recorded in a Cary Eclipse spectrophotometer (Varian, Australia). Samples were excited at 280 nm and emission spectra were recorded between 300 and 400 nm. Analysis of denaturation process was carried out by plotting the ratios of fluorescence intensities (319/355 nm) as function of denaturant concentration.

## RESULTS

### EXPRESSION OF RECOMBINANT L-ASPARAGINASE IN *E. COLI*

Transformed *E. coli* cells were induced at optical density of 0.6 at 600 nm with 1 mM IPTG and expression of recombinant asparaginase was checked on 12% SDS polyacrylamide gel. Recombinant L-asparaginase was expressed as a  $\sim 37 \text{ kDa}$  protein and most of it accumulated as intracellular aggregates (Figure 1A). In shaker-flask culture at  $\text{OD}_{600}$  of 1.5, around 120 mg of L-asparaginase was produced as IB aggregates. We lysed the cells by sonication in the presence of lysozyme and IB pellet was extensively washed with 0.1% Na deoxycholate. Use of lysozyme during cell lysis helped in better recovery of IB aggregates from cells. Washing the IB pellet with 0.1% deoxycholic acid helped in the removal of contaminating membrane proteins. Purified IB pellet was observed to consist of two major bands on SDS-PAGE gel, one corresponds to recombinant L-asparaginase (85%) and another one to the lysozyme used during IBs isolation (Figure 1B). It was observed that small fraction of lysozyme may stick to IBs during isolation step. This was also observed by others in their study (Peternel and Komel, 2010).



**FIGURE 1 | SDS-PAGE analysis of expression and purification of recombinant L-asparaginase II from *E. coli* IBs.** Samples were analyzed on 12% polyacrylamide gel. Lane M; low molecular weight marker proteins (Pharmacia, Sweden) in (A–E). Arrow indicates recombinant asparaginase. (A) Lane 1: uninduced cell lysate; lane 2: induced cell lysate. (B) Lane 1: purified IBs. (C) Lane 1–9: supernatants

of solubilized asparaginase from IBs in 0 to 8 molar urea respectively in 50 mM Tris pH 8.5 buffer. (D) Eluted fractions of asparaginase from DEAE ion exchange chromatography. Lanes 1–9 are eluted fractions 1, 3, 5, 7, 9, 11, 13, 15, and 17 from DEAE ion exchange chromatography. (E) Lane 1: purified recombinant asparaginase after S-200 gel chromatography.

#### REFOLDING AND PURIFICATION OF ASPARAGINASE FROM IBs

Isolated and purified L-asparaginase IBs were solubilized in different concentrations of urea in 50 mM Tris-HCl buffer at pH 8.5. 4 M urea completely solubilized the recombinant L-asparaginase IBs (Figure 1C). An extra major protein band was observed above the recombinant L-asparaginase band in Figure 1C. This protein band probably represents the host chaperone protein

(DnaJ ~40 kDa) expressed during recombinant expression of L-asparaginase. The solubility of asparaginase IBs in 4 M urea containing Tris-HCl was found to be ~10 mg/ml. Solubilized recombinant asparaginase was refolded by pulsatile dilution method. Presence of 0.1 M arginine in refolding buffer reduced protein aggregation, promoted tetramer formation and improved the protein recovery during refolding. Refolded asparaginase was

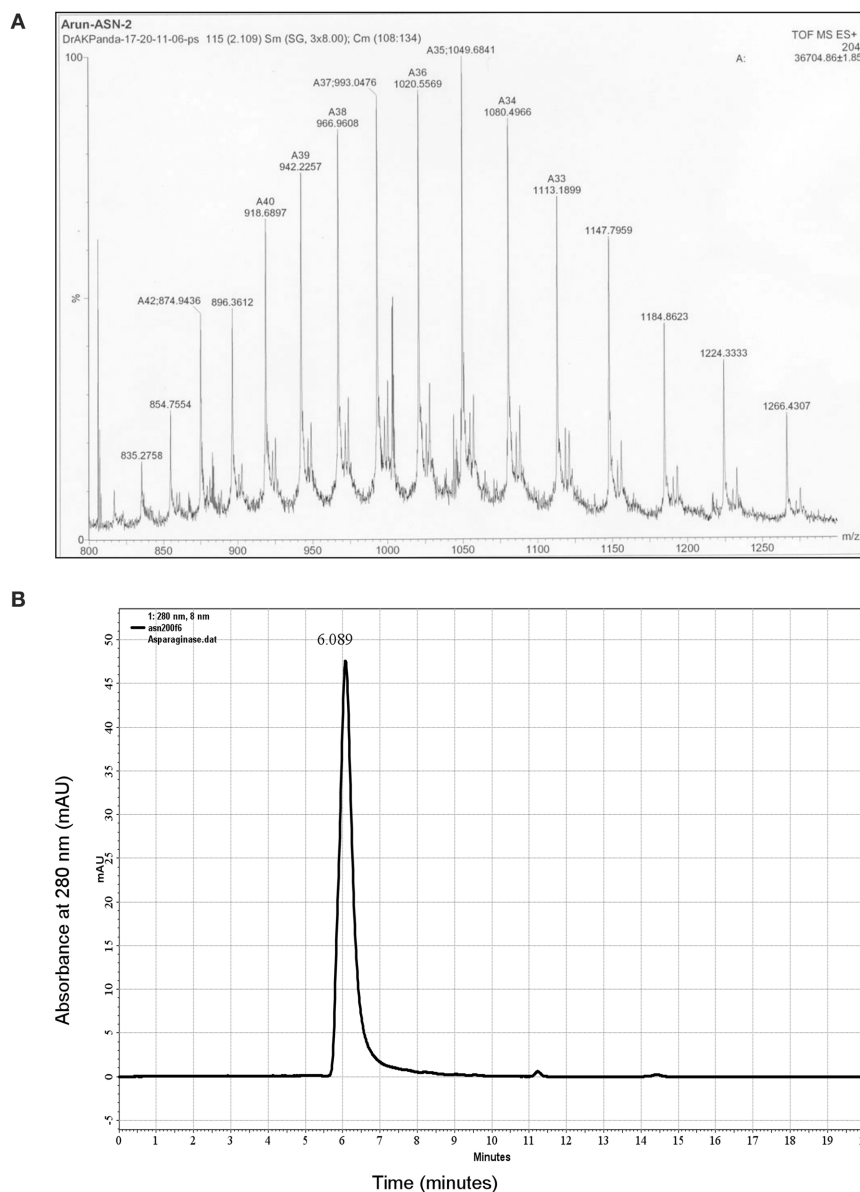


**Table 1 | Purification of recombinant *E. coli* L-asparaginase II from IBs (culture volume 1 liter).**

Fractions	Total protein (mg)	Specific activity (IU/mg) <sup>a</sup>	Fold purification <sup>b</sup>	% Yield
Total cell protein	580	–	–	–
IBs	118	–	–	100
Solubilized protein	110	–	–	93.2
Refolded protein	97	125	1	82.2
DEAE ion-exchange chromatography	78	160	1.3	66.1
Gel filtration chromatography	60	190	1.5	50.8

<sup>a</sup> One international unit is defined as amount of enzyme required to produce one micromole of ammonia per min under standard condition of reaction.

<sup>b</sup> Fold purification was calculated after refolding as the ratio of specific activity at given step of purification to specific activity of refolded sample.



**FIGURE 2 | Characterization of purified asparaginase. (A)** Mass spectrogram of purified asparaginase. Protein was dissolved in 0.1% formic acid and spectrum was obtained from LCMS instrument, Waters, USA. **(B)** HPLC size exclusion chromatogram of purified

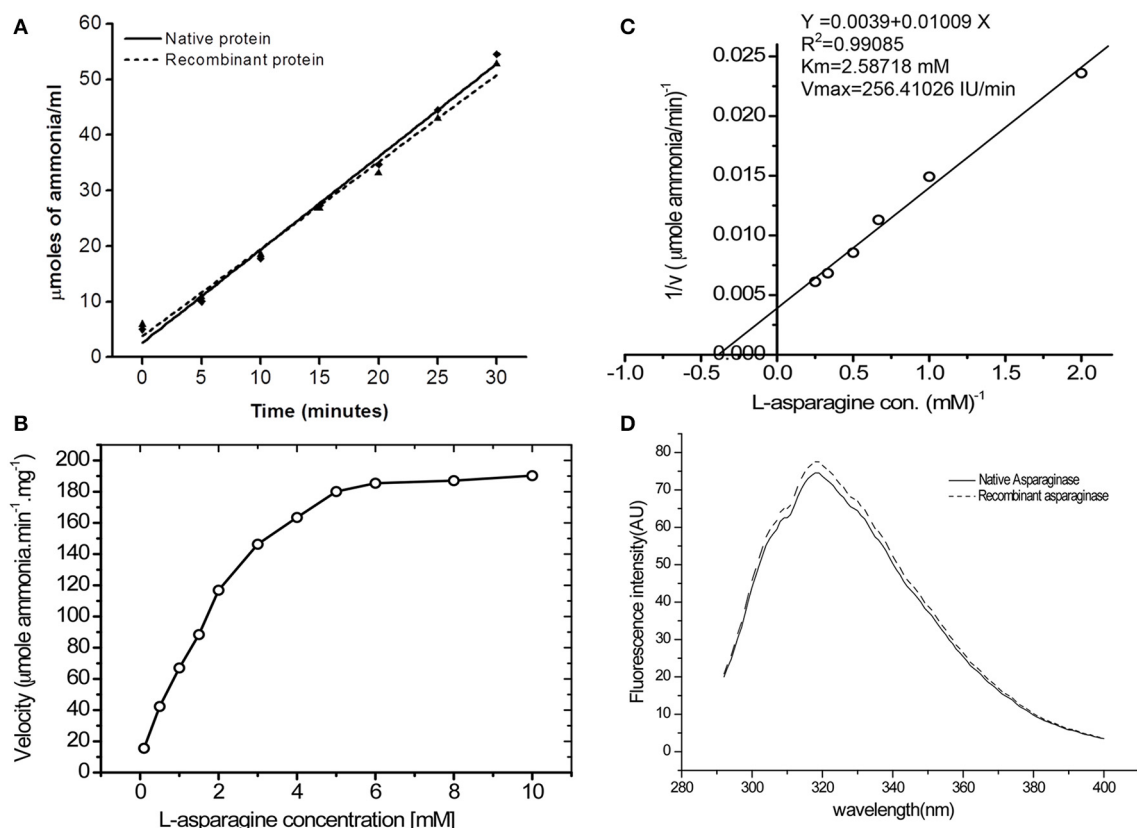
recombinant asparaginase loaded on Bio-Sep S-2000 column (Phenomenex, USA). Elution profile shows the presence of single peak corresponding to the asparaginase tetramer of ~150 kDa and retention time 6.08 min.

dialyzed and loaded on DEAE ion-exchange chromatographic column. Refolded asparaginase was eluted at a conductance of 8–20 mS/cm. Eluates from ion exchange chromatography were analyzed by SDS-PAGE and found to be 95% pure (**Figure 1D**). Ion exchange fractions containing asparaginase were pooled and concentrated by ultra-filtration (10 kDa molecular weight cut-off) unit and loaded on S-200 column to separate the monomer and other higher soluble aggregates species from tetramer. The eluted asparaginase was pure (**Figure 1E**). Maximum step recovery of protein was achieved during solubilization and refolding steps (**Table 1**). Approximately 60 mg of purely refolded L-asparaginase was recovered from 1 L shaker -flask culture. The overall recovery of the bioactive enzyme from the IBs was ~50%. Mass spectroscopy data analysis showed the molecular weight of purified asparaginase single subunit to be equal to ~36.7 kDa that is equal to calculated molecular weight of recombinant asparaginase with six histidine tags (**Figure 2A**). To check whether purified recombinant asparaginase consists of a mixture of large molecular aggregates and monomer form other than the native tetramer, purified protein was loaded on BioSep S-2000 column. Purified recombinant asparaginase was

eluted at retention time 6.08 min at 1 ml/min flow rate with 100% peak intensity (**Figure 2B**). Molecular weight of eluted asparaginase was calculated using standard calibration curve plotted using different marker proteins and found to be close to 150 kDa. Molecular weight of purified asparaginase was found to be 8–10 kDa more due to the addition ~2 kDa N-terminal His-tag in each subunit of active tetramer. This result showed that the purified recombinant asparaginase was in tetrameric form.

#### ENZYME ACTIVITY AND DETERMINATION OF KINETIC PARAMETERS OF PURIFIED ASPARAGINASE

Steady-state kinetic analysis of recombinant L-asparaginase was carried out at saturating substrate concentration of L-asparagine to determine its specific activity. Time duration of the reaction was determined by calculating linear range of ammonia release from both recombinant and native asparaginase (**Figure 3A**). It was observed that up to 30 min, the amount of ammonia release was linear ( $R^2$  value for native protein on linear fit was 0.991 and for recombinant asparaginase it was 0.990). As the protein was refolded and purified, the specific activity increased from 125 to



**FIGURE 3 | Bioactivity of purified asparaginase. (A)** Time dependent ammonia release per ml of reaction mixture (Y axis) for native (—◆—) and recombinant (—▲—) asparaginase was calculated by Nessler's reagent method. Amount of asparaginase enzyme in final reaction buffer was 10 μg/ml. **(B)** Effect of L-asparagine concentration on specific activity of recombinant asparaginase. **(C)** Double reciprocal

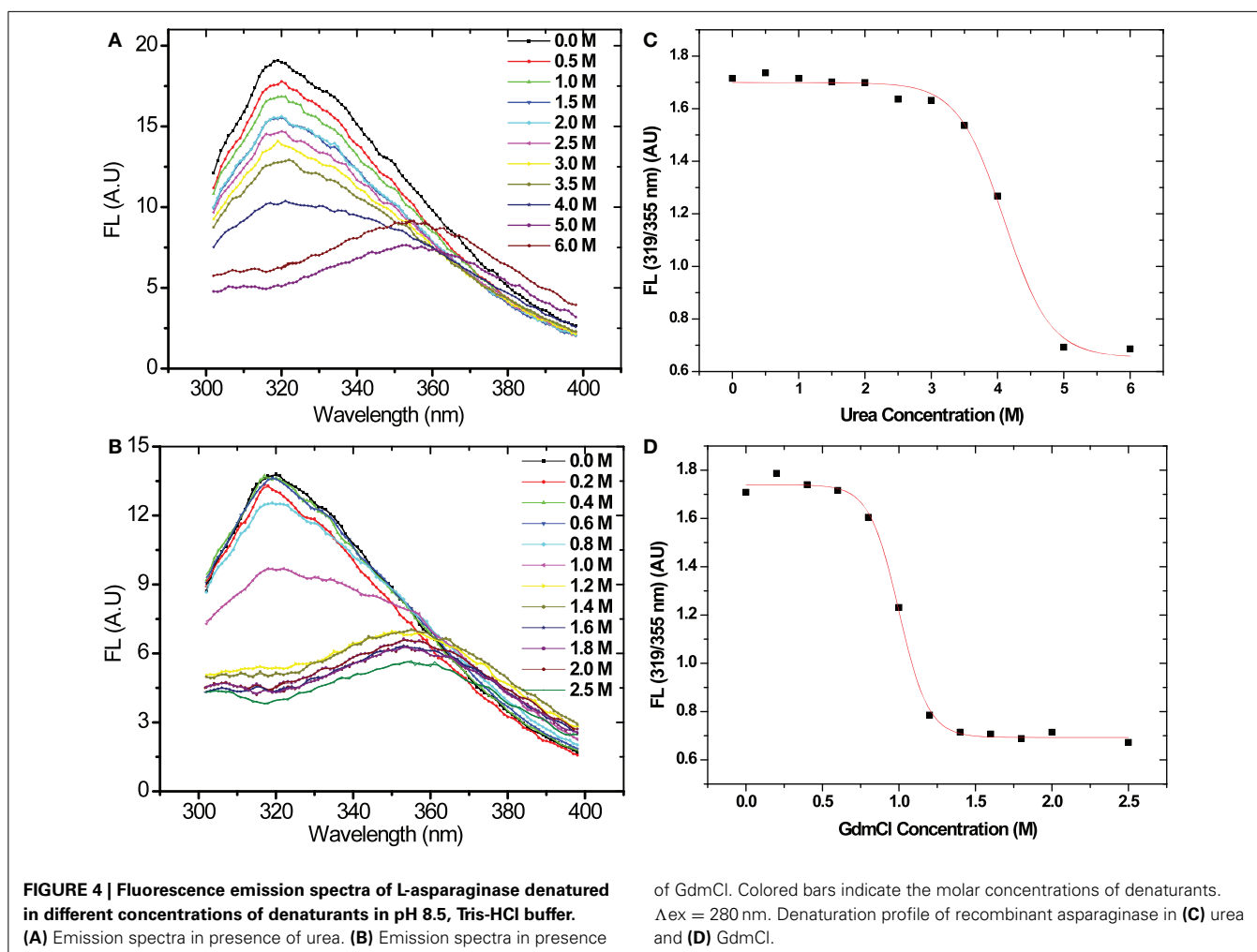
plot between velocity and Lasparagine concentration for determination of  $K_m$  and  $V_{max}$  values. **(D)** Fluorescence emission spectra of the native (—) and purified recombinant (---) Lasparaginase II at 15 μg/ml concentration in 50 mM Tris-HCl buffer, pH 7.5. Emission spectra show maxima at 320 nm with two shoulders at 307 and 325 nm at 280 nm excitation wavelengths.

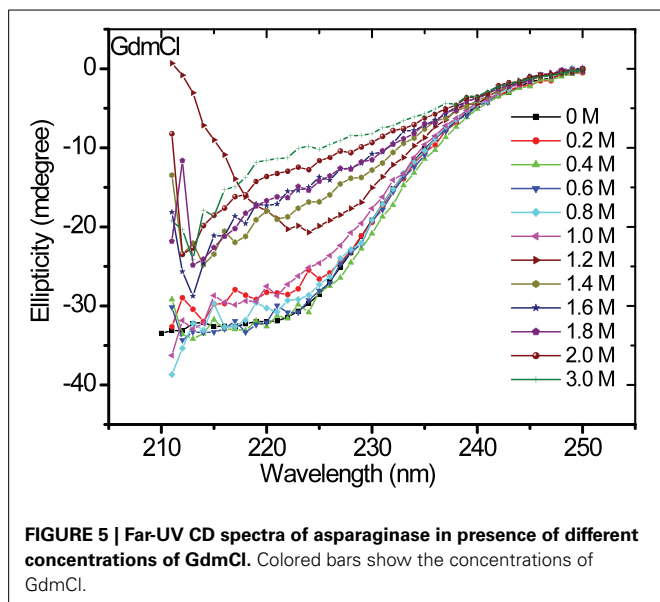
190 IU/mg (Table 1). Specific activity of final purified protein was found close to native *E. coli* asparaginase (200 IU) and was around  $190 \pm 5$  IU. The kinetic parameters for purified asparaginase were calculated by determining the specific activity of asparaginase at different concentrations of L-asparagine (0.1–10 mM) (Figure 3B). The  $K_m$  and  $V_{max}$  value were calculated from the linear fit of double reciprocal Lineweaver-Burk plot (Figure 3C). The  $K_m$  and  $V_{max}$  values for purified protein were found to be 2.58 mM L-asparagine and 256 IU/mg respectively. The fluorescence spectrum of the refolded asparaginase was comparable to that of the native enzyme (Figure 3D). The similarity of specific activities and fluorescence spectrum of the refolded enzyme with the native asparaginase indicated that the recombinant protein has been refolded into native-like conformation.

#### CHEMICAL DENATURATION OF ASPARAGINASE BY UREA AND GDMCL

Chemical denaturation of asparaginase was probed by fluorescence spectroscopy in the presence of different concentrations of urea and GdmCl. The protein was excited at 280 nm, and emission spectra were recorded between 300 and 400 nm. Fluorescence spectra of asparaginase in the presence of different concentrations of urea and GdmCl are presented in Figures 4A,B. Native

asparaginase showed emission maximum at 319 nm. In the presence of 6 M urea or 2 M GdmCl, complete denaturation of the protein was observed with emission maximum at 356 nm that is a characteristic feature of solvent-exposed tryptophan residues in polar environment. Fluorescence intensity was reduced with the increase in urea concentration (up to 4 M) without any significant shift in emission maxima. However, there was a 36 nm red shift in emission maximum of asparaginase at higher than 4 M urea concentrations. Similar denaturation behavior was observed in the presence of GdmCl. However, being a stronger denaturant, loss of fluorescence intensity and red shift in emission maximum occurred at 1.2 M concentration of GdmCl. As asparaginase has characteristic emission maxima wavelength in native and denatured states, denaturation curves of asparaginase were analyzed as ratios of fluorescence intensities at 319/356 nm at different concentrations of urea and GdmCl. Asparaginase showed cooperative denaturation profile in urea and GdmCl (Figures 4C,D). It can be concluded that asparaginase is more susceptible to GdmCl denaturation in comparison to urea. This conclusion was further supported by the denaturation profiles of asparaginase in presence of GdmCl observed with circular dichroism (CD) spectroscopy (Figure 5).

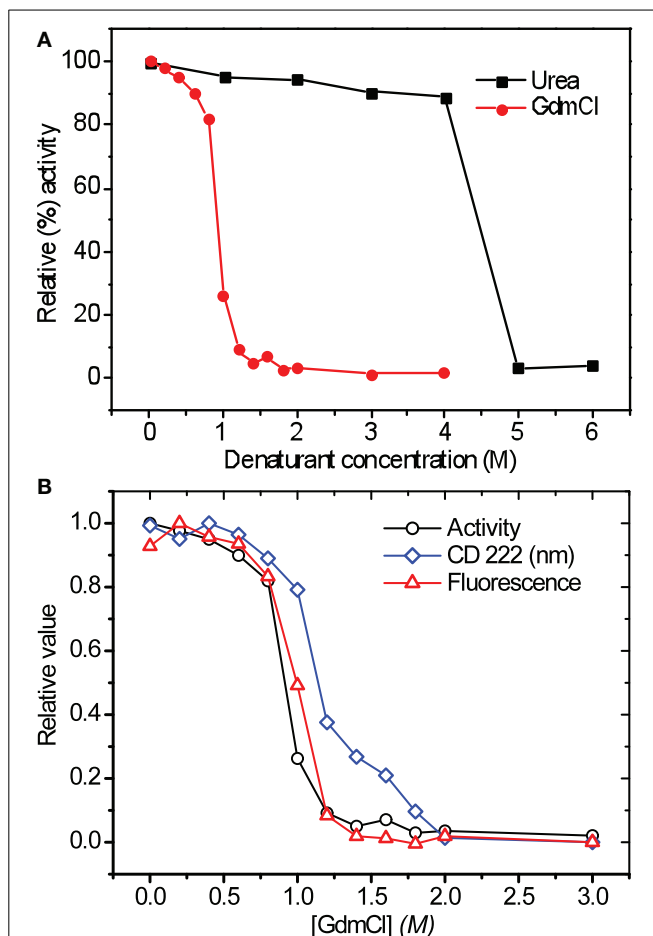




We determined the asparaginase enzymatic activity in the presence of denaturants such as urea and GdmCl at different concentrations to gain information on the stability of quaternary structure. The denaturant concentration-dependent activity profile of asparaginase is shown in **Figure 6A**. No significant loss in activity of asparaginase till 4 M urea was observed. Similarly, asparaginase retained 80% activity in 0.8 M GdmCl. The GdmCl concentration-dependence of fluorescence spectra, CD signal at 222 nm and enzyme activity of asparaginase was compared (**Figure 6B**). Fluorescence signal and activity were found to be in agreement. However, the CD signal at 222 nm showed a biphasic denaturation behavior in which first transition state was similar to the fluorescence and activity profile. The second transition state was observed between 1.4 and 1.8 M concentrations of GdmCl where the protein retained some proportion of secondary structures.

#### THERMAL DENATURATION AND DETERMINATION OF MELTING TEMPERATURE

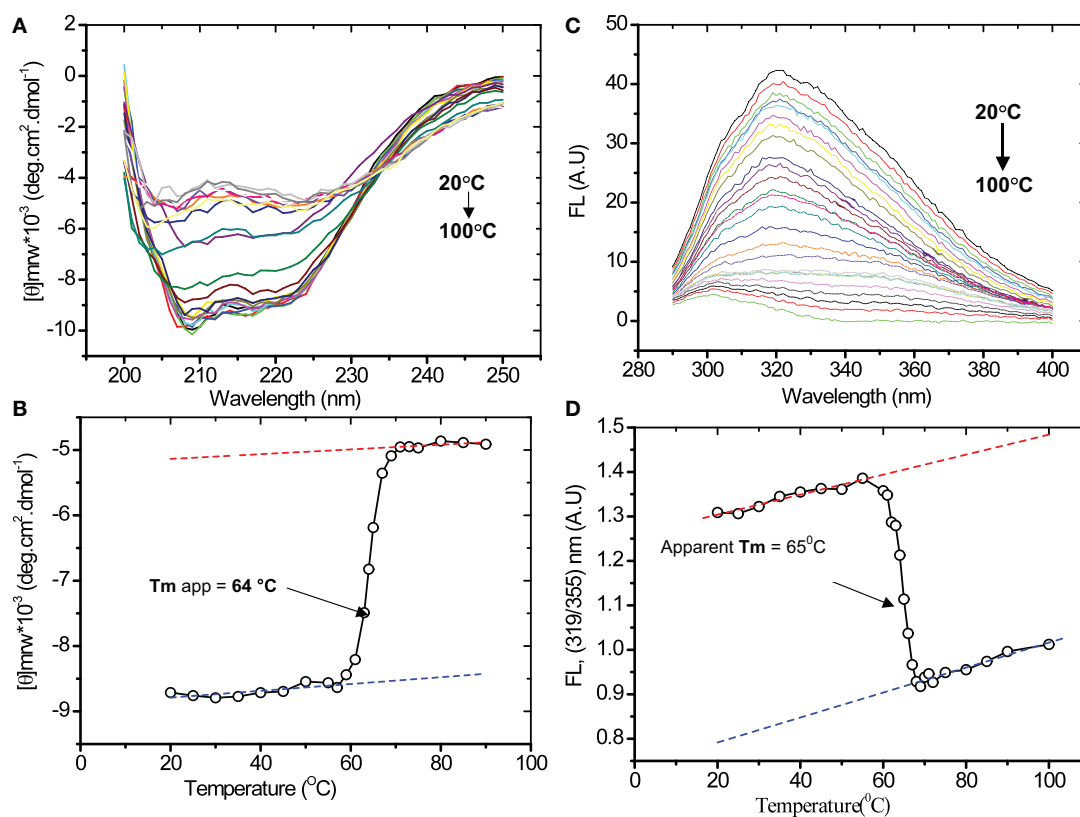
Thermal stability of purified recombinant asparaginase was probed by CD and fluorescence spectroscopy. The CD spectra of asparaginase were recorded at different temperatures (**Figure 7A**). Samples were left for 5 min for equilibration at each point, and three scans were recorded. In pre- and post-transition range of thermal denaturation, scans were performed at 5°C intervals and 1°C intervals within the transition range. CD 222 nm signals were plotted against temperature (**Figure 7B**). Thermal denaturation curves showed the loss of CD signal at start of 57°C and gradually decreased to the lowest value at 70°C. Further increase in temperature beyond 70°C had no significant effect on CD 222 nm signal. The melting temperature ( $T_m$ ) of asparaginase was found to be 64°C at pH 8.0. Similarly, the thermal denaturation of asparaginase was probed by acquiring the fluorescence spectra at different temperatures (**Figure 7C**). Ratio of fluorescence intensities at 319 and 355 nm (FL at 319/355 nm), was plotted against temperature (**Figure 7D**). The denaturation curve showed the increase in



the FL at 319/355 nm) value up to 57°C. After that, there was a sharp decrease in this value. In post-transition range, a similar trend was observed as in pre-transition range. The curve fitting showed that the apparent  $T_m$  to be 65°C. It was close to the  $T_m$  value calculated from thermal denaturation curve plotted from CD signals.

#### DISCUSSION

Refolding yield of bioactive proteins from IBs of *E. coli* are in general very low and contribute toward a major cost for production of recombinant proteins (Datar et al., 1993). However, expression of protein as IBs has the advantages of very high level of protein expression and resistance to protease attack. Thus, IB expression helps in purifying the denatured protein with a lower number of operating steps. If suitable high throughput refolding procedure can be developed for a particular protein, IB formation will be very helpful in the production of recombinant protein. Such an approach will be more beneficial for oligomeric proteins where the refolding yields are even lower. Strategy used to



**FIGURE 7 | Thermal denaturation of recombinant asparaginase.**

(A) Far-UV CD spectra of asparaginase at different temperatures at pH 8.0. (B) Thermal denaturation curve of asparaginase plotted between CD 222 nm signal and temperature. (C) Fluorescence

spectra of asparaginase at different temperatures at pH 8.0.

(D) Thermal denaturation curve of asparaginase was plotted between FL (319/355) nm and temperature.  $T_m$  values are indicated by arrow.

recover bioactive protein from IBs mainly involves four steps: isolation of IBs from *E. coli* cells; solubilization of the IB aggregates; refolding and purification of the solubilized protein. Among these steps, solubilization of the protein aggregates and refolding of the solubilized protein are the most crucial steps and need careful attention for high recovery of protein. Aggregation leading to low recovery of the recombinant protein occurs due to the use of sub-optimal refolding procedure. It is expected that mild solubilization of the IB aggregates using low concentration of denaturants followed by pulsatile refolding of the solubilized protein would be an ideal approach for maximal recovery of bioactive proteins from the IBs.

In this study, *E. coli* L-asparaginase II was expressed as IB aggregates in *E. coli*, purified by detergent washing and solubilized using only 4 molar urea solution. Maximum of 10–11 mg/ml of the protein could be solubilized at 4 M urea solution. The solubilized protein was subsequently refolded into native tetrameric form using pulsatile dilution method. Around 50% of IB protein could be refolded into bioactive tetrameric protein using the mild solubilization procedure. The production data are comparable to that reported for soluble expression using LB medium in shaker-flask culture (Khushoo et al., 2004). It is expected that if such expression system are optimized using high cell density fermentation followed by the present IB refolding process, the

volumetric productivities will be even better. Crystal structure of *E. coli* asparaginase shows that the active site is present between subunits of intimate dimer formed by interaction of N-terminal domain of one subunit with C-terminal domain of the other subunit. Each dimer has two active sites, but only tetramer shows the activity. Apart from this, 0.1 M arginine in refolding buffer also promoted the formation of tetramer and reduced the extent of protein aggregation. It showed that the arginine not only helped in preventing aggregation during refolding but also promoted the formation of correctly folded subunits for proper association into tetrameric form. Fluorescence spectrum of the refolded asparaginase was similar to that of native asparaginase indicating the formation of proper quaternary structure of the recombinant asparaginase after refolding.

According to the crystal structure of asparaginase,  $\alpha$ -helices are predominantly present on the surface of tetramer. Active site of asparaginase is present on the interface of subunits with tryptophan residues at the core of N-terminal domains of monomers near the active site. The cooperative fluorescence and activity profile and biphasic nature of CD spectra in the presence of different concentrations of GdmCl could be possible if, during the denaturation, dissociation of subunits precedes the complete loss of secondary structures into monomers. Protecting the existing secondary structure element of the protein during solubilization



thus favors the association monomer into oligomeric form and results in the formation of tetrameric bioactive protein.

## CONCLUSIONS

Understanding the solubilization profile of IB aggregates at different denaturation concentration helped in designing a mild solubilization procedure. Such a mild process protected the existing native-like protein structure and thus helped in improved recovery of bioactive protein during refolding. Solubilized asparaginase from the IBs was refolded into tetrameric form having high specific enzymatic activity by optimization of solubilization and refolding conditions. Purified asparaginase was found to be active. Kinetic parameters of the enzyme were determined. Purified asparaginase was characterized for its denaturation profiles in the presence of urea and GdmCl. Thermal denaturation of the purified asparaginase was also studied, and the  $T_m$  was determined to be about 64°C. Mild solubilization of IB protein is thus the key requirements in high throughput recovery of IB proteins into bioactive form. This concept thus can be used for recovery of bioactive oligomeric proteins from the IBs of *E. coli*.

## ACKNOWLEDGMENTS

This work is supported by the core funding of the National Institute of Immunology, New Delhi, to Dr. Amulya K. Panda. Arun K. Upadhyay and Anupam Singh are supported by Council for Scientific and Industrial Research (CSIR), New Delhi, India.

## REFERENCES

- Aung, H. P., Bocola, M., Schleper, S., and Rohm, K. H. (2000). Dynamics of a mobile loop at the active site of *Escherichia coli* asparaginase. *Biochim. Biophys. Acta* 1481, 349–359. doi: 10.1016/S0167-4838(00)00179-5
- Burgess, R. R. (2009). Refolding solubilized IB proteins. *Methods Enzymol.* 463, 259–282. doi: 10.1016/S0076-6879(09)63017-2
- Datar, R. V., Cartwright, T., and Rosen, C. G. (1993). Process economics of animal cell and bacterial fermentations: a case study analysis of tissue plasminogen activator. *Biotechnology (N.Y.)* 11, 349–357.
- De Bernardez, C. E., Schwarz, E., and Rudolph, R. (1999). Inhibition of aggregation side reactions during *in vitro* protein folding. *Methods Enzymol.* 309, 217–236.
- Fahnert, B., Lilie, H., and Neubauer, P. (2004). Inclusion bodies: formation and utilisation. *Adv. Biochem. Eng. Biotechnol.* 89, 93–142. doi: 10.1007/b93995
- Ferrara, M. A., Severino, N. M. B., Valente, R. H., Perales, J., and Bon, E. P. S. (2010). High yield extraction of periplasmic asparaginase produced by recombinant *P. pastoris* harbouring the *S. cerevisiae* ASP3 gene. *Enzyme Microb. Technol.* 47, 71–76. doi: 10.1016/j.enzmictec.2010.05.001
- Garrido, F., Estrela, S., Alves, C., Sanchez-Perez, G. F., Sillero, A., and Pajares, M. A. (2011). Refolding and characterization of methionine Adenosyltransferase from *Euglena gracilis*. *Protein Expr. Purif.* 79, 128–136. doi: 10.1016/j.pep.2011.05.004
- Graham, M. L. (2003). Pegaspargase: a review of clinical studies. *Adv. Drug Deliv. Rev.* 55, 1293–1302. doi: 10.1016/S0169-409X(03)00110-8
- Hartley, D. L., and Kane, J. F. (1988). Properties of inclusion bodies from recombinant *Escherichia coli*. *Biochem. Soc. Trans.* 16, 101–102.
- Karumuri, N. N., Gangireddy, S. R., Narala, V. R., Majee, S. S., Gunwar, S., and Reddy, R. C. (2007). Simple, rapid, high-purity preparation of recombinant human platelet-derived growth factor-BB. *Biotechnol. Lett.* 29, 1333–1339. doi: 10.1007/s10529-007-9411-9
- Khushoo, A., Pal, Y., Singh, B. N., and Mukherjee, K. J. (2004). Extracellular expression and single step purification of recombinant *E. coli* L asparaginase II. *Protein Expr. Purif.* 38, 29–36. doi: 10.1016/j.pep.2004.07.009
- Kozak, M., Borek, D., Janowski, R., and Jaskolski, M. (2002). Crystallization and preliminary crystallographic studies of five crystal forms of *Escherichia coli* L-asparaginase II (Asp90Glu mutant). *Acta Crystallogr. D Biol. Crystallogr.* 58, 130–132. doi: 10.1107/S0907444901016663
- Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227, 680–685. doi: 10.1038/227680a0
- Mahajan, R. V., Saran, S., Kameswaran, K., Kumar, V., and Saxena, R. K. (2012). Efficient production of L-asparaginase from *Bacillus licheniformis* with low glutaminase activity: optimization of scale-up and acrylamide degradation studies. *Bioresour. Technol.* 125, 11–16. doi: 10.1016/j.biortech.2012.08.086
- Muller, H. J., and Boos, J. (1998). Use of L-asparaginase in childhood ALL. *Crit. Rev. Oncol. Hematol.* 28, 97–113. doi: 10.1016/S1040-8428(98)00015-8
- Oza, V. P., Parmer, P. P., Patel, D. H., and Subramanian, R. B. (2011). Cloning, expression and characterization of L-asparaginase from *Withania somnifera* L. for large scale production. *3 Biotech* 1, 21–26. doi: 10.1007/s13205-011-0003-y
- Panda, A. K. (2003). Bioprocessing of therapeutic proteins from the inclusion bodies of *Escherichia coli*. *Adv. Biochem. Eng. Biotechnol.* 85, 43–93. doi: 10.1007/3-540-36466-8\_3
- Patra, A. K., Mukhopadhyay, R., Mukhija, R., Krishnan, A., Garg, L. C., and Panda, A. K. (2000). Optimization of IB solubilization and renaturation of recombinant human growth hormone from *Escherichia coli*. *Protein Expr. Purif.* 18, 182–192. doi: 10.1006/prep.1999.1179
- Peternel, S., and Komel, R. (2010). Isolation of biologically active nanomaterial (inclusion bodies) from bacterial cells. *Microb Cell Fact.* 9:66. doi: 10.1186/1475-2859-9-66
- Peternel, S., and Komel, R. (2011). Active protein aggregates produced in *Escherichia coli*. *Int. J. Mol. Sci.* 12, 8275–8287. doi: 10.3390/ijms12118275
- Rudolph, R., and Lilie, H. (1996). *In vitro* folding of IB proteins. *FASEB J.* 10, 49–56.
- Scrofani, S. D., Fabri, L. J., Xu, P., Maccarone, P., and Nash, A. D. (2000). Purification and refolding of vascular endothelial growth factor-B. *Protein Sci.* 9, 2018–2025. doi: 10.1110/ps.9.10.2018
- Singh, S. M., and Panda, A. K. (2005). Solubilization and refolding of bacterial IB proteins. *J. Biosci. Bioeng.* 99, 303–310. doi: 10.1263/jbb.99.303
- Singh, S. M., Sharma, A., Upadhyay, A. K., Singh, A., Garg, L. C., and Panda, A. K. (2012). Solubilization of IB proteins using n-propanol and its refolding into bioactive form. *Protein Expr. Purif.* 81, 75–82. doi: 10.1016/j.pep.2011.09.004
- Swain, A. L., Jaskolski, M., Housset, D., Rao, J. K., and Wlodawer, A. (1993). Crystal structure of *Escherichia coli* L-asparaginase, an enzyme used in cancer therapy. *Proc. Natl. Acad. Sci. U.S.A.* 90, 1474–1478.
- Umetsu, M., Tsumoto, K., Ashish, K., Nitta, S., Tanaka, Y., Adschiri, T., et al. (2004). Structural characteristics and refolding of *in vivo* aggregated hyperthermophilic archaeon proteins. *FEBS Lett.* 557, 49–56. doi: 10.1016/S0014-5793(03)01441-8
- Upadhyay, A. K., Murmu, A., Singh, A., and Panda, A. K. (2012). Kinetics of IB formation and its correlation with the characteristics of protein aggregates in *Escherichia coli*. *PLoS ONE* 7:e33951. doi: 10.1371/journal.pone.0033951
- Vallejo, L. F., and Rinas, U. (2004). Strategies for the recovery of active proteins through refolding of bacterial IB proteins. *Microb. Cell Fact.* 3:11. doi: 10.1186/1475-2859-3-11
- Ventura, S., and Villaverde, A. (2006). Protein quality in bacterial inclusion bodies. *Trends Biotechnol.* 24, 179–185. doi: 10.1016/j.tibtech.2006.02.007
- Verma, N., Kumar, K., Kaur, G., and Anand, S. (2007). L-asparaginase: a promising chemotherapeutic agent. *Crit. Rev. Biotechnol.* 27, 45–62. doi: 10.1080/07388550601173926
- Wriston, J. C. Jr. (1985). Asparaginase. *Methods Enzymol.* 113, 608–618.

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 05 May 2014; accepted: 28 August 2014; published online: 15 September 2014.

Citation: Upadhyay AK, Singh A, Mukherjee KJ and Panda AK (2014) Refolding and purification of recombinant L-asparaginase from inclusion bodies of *E. coli* into active tetrameric protein. *Front. Microbiol.* 5:486. doi: 10.3389/fmicb.2014.00486

This article was submitted to Microbial Physiology and Metabolism, a section of the journal *Frontiers in Microbiology*.

Copyright © 2014 Upadhyay, Singh, Mukherjee and Panda. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Multitasking SecB chaperones in bacteria

Ambre Sala, Patricia Bordes and Pierre Genevoux\*

Laboratoire de Microbiologie et Génétique Moléculaire, Centre National de la Recherche Scientifique, Université Paul Sabatier, Toulouse, France

## Edited by:

Salvador Ventura, Universitat Autònoma de Barcelona, Spain

## Reviewed by:

Claes von Wachenfeldt, Lund University, Sweden

Julien Brillard, Institut National de la Recherche Agronomique, France

## \*Correspondence:

Pierre Genevoux, Laboratoire de Microbiologie et Génétique Moléculaire, Centre National de la Recherche Scientifique, Université Paul Sabatier, 118 route de Narbonne, 31062 Toulouse Cedex 9, France  
e-mail: pierre.genevoux@ibcg.biotoul.fr

Protein export in bacteria is facilitated by the canonical SecB chaperone, which binds to unfolded precursor proteins, maintains them in a translocation competent state and specifically cooperates with the translocase motor SecA to ensure their proper targeting to the Sec translocon at the cytoplasmic membrane. Besides its key contribution to the Sec pathway, SecB chaperone tasking is critical for the secretion of the Sec-independent heme-binding protein HasA and actively contributes to the cellular network of chaperones that control general proteostasis in *Escherichia coli*, as judged by the significant interplay found between SecB and the trigger factor, DnaK and GroEL chaperones. Although SecB is mainly a proteobacterial chaperone associated with the presence of an outer membrane and outer membrane proteins, *secB*-like genes are also found in Gram-positive bacteria as well as in certain phages and plasmids, thus suggesting alternative functions. In addition, a SecB-like protein is also present in the major human pathogen *Mycobacterium tuberculosis* where it specifically controls a stress-responsive toxin–antitoxin system. This review focuses on such very diverse chaperone functions of SecB, both in *E. coli* and in other unrelated bacteria.

**Keywords:** protein folding and targeting, SecA, DnaK, trigger factor, proteases, toxin–antitoxins

## PROTEIN FOLDING AND TARGETING IN BACTERIA

A major challenge for the cells is to ensure the proper folding and targeting of newly synthesized proteins to the different cellular compartments. Indeed, ongoing protein synthesis in the crowded cellular environment offers a window of opportunities for non-native interactions, which may eventually lead to proteostasis breakdown (Kramer et al., 2009). Therefore, to cope with noxious off pathways in protein biogenesis, cells have evolved universally conserved molecular chaperones and targeting factors, which act co- and/or post-translationally to guide the precise partitioning, localization and folding of newly synthesized proteins (Kramer et al., 2009; Kim et al., 2013).

In bacteria, the folding of newly synthesized proteins is mainly assisted by three highly conserved cytosolic chaperones, namely trigger factor (TF), DnaK/DnaJ/GrpE (DnaKJE), and GroEL/GroES (GroESL; Deuerling et al., 1999; Agashe et al., 2004; Kerner et al., 2005). The ribosome-bound TF is the first chaperone to interact co-translationally with most newly synthesized proteins (Valent et al., 1995). Although the majority of the cytosolic proteins can reach their native state following interaction with TF, a substantial amount of proteins (about 30%) need further co- and/or post-translational assistance by the downstream DnaKJE and GroESL chaperones (Bukau et al., 2000). Forceful genetic and biochemical analyses have demonstrated significant overlap and cooperation between these three major chaperones, revealing a dynamic network of chaperones to control intracellular proteostasis (Teter et al., 1999; Genevoux et al., 2004; Calloni et al., 2012).

Targeting of newly synthesized proteins to the bacterial cytoplasmic membrane can occur either co- or post-translationally. While certain small membrane proteins are

targeted post-translationally to the YidC insertase at the inner membrane (Dalbey et al., 2011), most integral membrane proteins as well as some presecretory proteins are targeted co-translationally by the ribosome-associated RNA-protein complex SRP (Saraogi and Shan, 2014). SRP binds to hydrophobic signal-anchor or signal sequence in nascent chains and targets them to the Sec translocon via interaction with its membrane receptor FtsY (Luirink and Sinning, 2004). The majority of presecretory proteins are translocated post-translationally either folded via the twin-arginine translocation (Tat) pathway or in a non-native state via the Sec pathway. The Tat system is known to translocate folded proteins or assembled protein complexes (up to 70 Å in diameter) through the cytoplasmic membrane. Tat substrate proteins possess an amino-terminal signal sequence with a conserved twin-arginine motif, which mediates post-translational targeting to the Tat translocon (Palmer and Berks, 2012; Patel et al., 2014). They are often assisted by specific cytosolic chaperones called redox enzyme maturation proteins (REMPs) and by the generic chaperones DnaK and GroEL, which likely prevent their degradation and premature export, and facilitate their assembly and functional interaction with the translocon (Castanie-Cornet et al., 2014).

The Sec translocon is conserved in all three domains of life. Its core is composed of a heterotrimeric membrane complex SecYEG in bacteria and Sec61 $\alpha\beta\gamma$  in eukaryotes (du Plessis et al., 2011). While translocation in the endoplasmic reticulum via the Sec translocon is mainly mediated co-translationally and thus energized by polypeptide chain elongation, Sec translocation across the bacterial plasma membrane preferentially occurs post-translationally and energy is provided by the SecA ATPase motor component (Chatzi et al., 2013). In this case, SecA binds to presecretory proteins with mildly hydrophobic signal sequences,

targets them to the Sec translocon at the inner membrane via a direct interaction with SecY, and subsequently drives the translocation process by successive cycles of ATP binding and hydrolysis.

Productive interaction with Sec is influenced by the folding rate of the substrate and facilitated by cytosolic chaperones capable of preventing premature folding, aggregation or degradation of precursor proteins (Randall and Hardy, 2002). Accordingly, all three main generic chaperone machines involved in *de novo* protein folding, namely TF, DnaKJE and GroEL, have been shown to participate at different levels in this process (Castanie-Cornet et al., 2014). Remarkably, most proteobacteria also possess the chaperone SecB, which in addition to its generic chaperone function has the ability to specifically interact with SecA to facilitate post-translational delivery of presecretory proteins to the Sec translocon (Bechtluft et al., 2010). This review describes the Sec-dependent and Sec-independent cellular functions of SecB, its interplay with other molecular chaperones as well as the distribution of SecB homologs in very diverse bacteria. The role of the recently identified SecB-like proteins in the control of intracellular stress-responsive toxin–antitoxin (TA) systems is also discussed.

## SecB AND THE Sec PATHWAY

SecB is a homotetrameric chaperone of 69 kDa with a cellular concentration estimated to be between 4 and 20  $\mu$ M in *Escherichia coli*. SecB binds co- and/or post-translationally to non-native precursor proteins, maintaining them in competent state for delivery to the Sec translocon via a well-described interaction with its SecA partner (Randall and Hardy, 1995, 2002; Chatzi et al., 2013).

The *E. coli secB* gene was initially identified genetically by selecting for mutants that were defective in the export of a fusion protein composed of the N-terminal part of maltose-binding protein (MBP) preMBP (containing the signal sequence) and  $\beta$ -galactosidase (Kumamoto and Beckwith, 1983). Further experiments showed that *secB* mutations delayed or blocked the processing of a subset of preproteins and exhibited a synergistic effect with temperature-sensitive alleles of *secA*, thus revealing a role for SecB in export (Kumamoto and Beckwith, 1983, 1985). *E. coli secB* mutant strains were initially shown to be defective for growth on rich Luria broth media agar plates, but it later appeared that this phenotype was due to a polar effect on the downstream *gpsA* gene encoding a glycerol-3-phosphate dehydrogenase involved in phospholipid biosynthesis (Kumamoto and Beckwith, 1985; Shimizu et al., 1997). Deletion of *secB* without apparent polarity on *gpsA* results in a strong cold-sensitive phenotype below 23°C, a moderate temperature-sensitive phenotype at temperatures above 45°C and a hypersensitivity to several antibiotics (Ullers et al., 2007; **Table 1**). Most of the relevant phenotypes associated with *secB* mutations or SecB overexpression are shown in **Table 1**. Genetic interactions between *secB* and the various protein localization locus (*prl*) mutations known to suppress the export defect of sequence signal deficient precursors are also presented (**Table 1**).

The crystal structures of SecB from *Haemophilus influenza* (Xu et al., 2000) and *E. coli* (Dekker et al., 2003) revealed that it forms a tetramer that assembles as a dimer of dimers (**Figure 1A**).

SecB monomer is composed of four stranded antiparallel  $\beta$ -sheets (the first two strands being at opposite sides and connected by a cross over loop) and two  $\alpha$ -helices separated by a helix connecting loop (**Figure 1A**). SecB dimer is formed via interactions between strands  $\beta$ 1 and helices  $\alpha$ 1 of two monomers. The tetramer forms by packing the helices  $\alpha$ 1 of four monomers in between the eight stranded antiparallel  $\beta$ -sheets formed by each dimer, mainly via polar interactions. A peptide binding groove was suggested from these structures, lying between the end of the cross over loop and strand  $\beta$ 2 on one side, and the helix connecting loop and the helix  $\alpha$ 2 on the other side. This proposed substrate binding region likely contains two subsites: the aromatic, deep subsite 1, and the shallower hydrophobic subsite 2, as presented in **Figure 1**. Two peptide binding grooves are present on each side of the SecB tetramer, each potentially allowing the binding of  $\sim$ 20 amino acids-long extended segments. The fact that SecB can bind long fragments of approximately 150 residues in pre-protein substrates (Khisty et al., 1995) suggests that these might wrap around the chaperone using several possible routes. Accordingly, electron paramagnetic resonance spectrometry analysis of spin labeled SecB variants in the presence of the physiologic SecB substrate galactose binding protein revealed that in addition to the proposed peptide binding groove, a much larger area of SecB appears to make contact with the substrate (Crane et al., 2006; **Figure 1B**).

SecB binds to non-native protein substrates with low specificity and high affinity ( $K_d$  in the nanomolar range), generally in a one to one ratio of tetrameric chaperone to substrate (Randall and Hardy, 1995). SecB binds to regions within the mature part of preprotein substrates and does not specifically recognize signal sequences (Gannon et al., 1989; Liu et al., 1989). Substrate selectivity by SecB is thought to occur via a kinetic partitioning between binding to the chaperone and folding, which is modulated by the affinity and the folding rate of the substrate protein (Hardy and Randall, 1991). Seminal work performed on the SecB substrate preMBP revealed the appearance of proteolysis resistant conformation of preMBP in the absence of SecB, thus suggesting that binding to SecB prevents precursor proteins from acquiring a stable tertiary structure incompatible with Sec-dependent translocation (Collier et al., 1988). A single molecule study recently confirmed that binding to SecB maintains preMBP in a molten globule-like state, preventing the formation of stable tertiary interactions (Bechtluft et al., 2007). SecB binding motif was identified by peptide scan of protein substrates as a nine amino acids-long segment enriched in aromatic and basic residues, with acidic residues strongly disfavored. Such motifs statistically occur every 20–30 amino acid residues in both exported and cytosolic proteins, thus suggesting low substrate specificity (Knoblauch et al., 1999).

Several SecB dependent presecretory substrates have been identified in *E. coli* by pulse chase experiments, sequence prediction or following analysis of protein aggregates that accumulate in the absence of the chaperone. This includes 25 presecretory proteins, namely CsgF, DegP, FhuA, FkpA, GBP, LamB, MBP, OmpA, OmpC, OmpF, OmpT, OmpX, OppA, PhoE, TolB, TolC, YagZ, YaiO, YbgF, YcgK, YfaZ, YgiW, YftM, YliI, and YncE (Hayashi and Wu, 1985; Kumamoto and Beckwith, 1985; Kusters et al., 1989; Laminet et al.,

**Table 1 | Most relevant phenotypes associated with mutations or overexpression of the *E. coli* SecB chaperone.**

SecB	Phenotypes <sup>a</sup>
Single mutation	Cs below 23°C and Ts at 46°C on LB agar plates <sup>(1)</sup> ; sensitive to copper, ethanol, cholate, low pH, dibucaine, triclosan, verapamil, and to several antibiotics, including bacitracin, novobiocin, amoxicillin, carbenicillin, tetracycline, cefaclor, glufosfomycin, ceftazidime, tunicamycin <sup>(2)</sup> ; partially resistant to phage U3 <sup>(3)</sup> ; produces slightly bigger cells <sup>(4)</sup> ; and induces synthesis of heat-shock proteins <sup>(4,5)</sup> . Mutation in <i>secB</i> with polar effect on the downstream <i>gpsA</i> gene inhibits growth on LB agar plates <sup>(6)</sup> .
Genetic interactions	Mutation in <i>secB</i> suppresses erythromycin and rifampin sensitivity of <i>lptE</i> mutants with increased outer-membrane permeability <sup>(7)</sup> ; enhances growth and export defects of <i>secA51</i> mutation <sup>(4)</sup> ; exacerbates Lon, DnaJ <sup>(8)</sup> and TF <sup>(1)</sup> toxicity. Ts, Cs and export defect of <i>secB</i> mutation are suppressed by <i>tig</i> mutation <sup>(1)</sup> ; Cs is suppressed by <i>lon</i> mutation <sup>(8)</sup> and by overexpression of $\sigma^{32}$ <sup>(9)</sup> , DnaK/DnaJ <sup>(1,10)</sup> , GroEL/GroES <sup>(10,11)</sup> , Rv1957 <sup>(12)</sup> , SmegB <sup>(13)</sup> , and less efficiently by DnaJ259 <sup>(8)</sup> and SecA <sup>(1)</sup> . Export defect of <i>secB</i> mutations is partially suppressed by <i>secA853-128</i> mutation <sup>(14)</sup> . Synthetic lethal with <i>dnaKdnaJ</i> mutation <sup>(1)</sup> and possibly with mutations in forty-one additional genes, including <i>groEL</i> , the <i>dsbC</i> , <i>lolA</i> , and <i>cpxP</i> genes encoding periplasmic stress proteins and/or chaperones, as well as <i>rpIW</i> encoding for L23, the main chaperone docking site at the ribosomal peptide exit <sup>(15)</sup> . Likely presents positive or negative epistasis with eighty-nine additional mutations in cell envelope biogenesis genes <sup>(15)</sup> .
Protein localization loci	Mutation in <i>secB</i> blocks the phenotypic effects of the <i>prlC8</i> (mutation in <i>opdA</i> encoding the cytoplasmic Oligopeptidase A) suppressor of <i>lamB</i> signal sequence mutation <sup>(16)</sup> ; inhibits <i>prlA4</i> ( <i>secY</i> [F286Y, I408N]) mediated suppression of maltose-binding protein (MBP) signal sequence mutations <sup>(17,18)</sup> and <i>prlA4</i> and <i>prlZ1</i> mediated suppression of LamB signal sequence mutations <sup>(19,20)</sup> . The <i>prlA1001</i> ( <i>secY</i> [I90N]) and <i>prlA1024</i> ( <i>secY</i> [I408N]) mutations suppress export deficient maltose-binding protein in the absence of SecB <sup>(17)</sup> ; the <i>prlF1</i> mutation in the antitoxin gene <i>sohA</i> of the SohA-YhaV toxin-antitoxin system suppresses SecB-dependent accumulation of LamB precursors <sup>(21)</sup> .
Overexpression	Partially suppresses the Ts of the double <i>dnaK tig</i> mutant <sup>(22)</sup> ; affects expression of the cytoplasmic response regulator OmpR <sup>(23)</sup> ; prevents activation of the mycobacterial HigBA1 toxin-antitoxin system expressed in <i>E. coli</i> <sup>(12)</sup> .

<sup>a</sup>Phenotypes associated with mutation or overexpression of SecB, and genetic interactions between *secB* and other mutations. Cs and Ts stand for cold- and temperature-sensitive phenotype, respectively.

<sup>1</sup>Ullers et al. (2007); <sup>2</sup>Nichols et al. (2011); <sup>3</sup>Kumamoto and Beckwith (1985); <sup>4</sup>Baars et al. (2006); <sup>5</sup>Wild et al. (1993); <sup>6</sup>Shimizu et al. (1997); <sup>7</sup>Grabowicz et al. (2013); <sup>8</sup>Sakr et al. (2010); <sup>9</sup>Altman et al. (1991); <sup>10</sup>Castanie-Cornet et al. (2014); <sup>11</sup>Danese et al. (1995); <sup>12</sup>Bordes et al. (2011); <sup>13</sup>Sala et al. (2013b); <sup>14</sup>Mcfarland et al. (1993); <sup>15</sup>Babu et al. (2011); <sup>16</sup>Trun and Silhavy (1989); <sup>17</sup>Francetic et al. (1993); <sup>18</sup>Derman et al. (1993); <sup>19</sup>Wei and Stader (1994); <sup>20</sup>Trun et al. (1988); <sup>21</sup>Snyder and Silhavy (1992); <sup>22</sup>Ullers et al. (2004); <sup>23</sup>Jin and Inouye (1995).

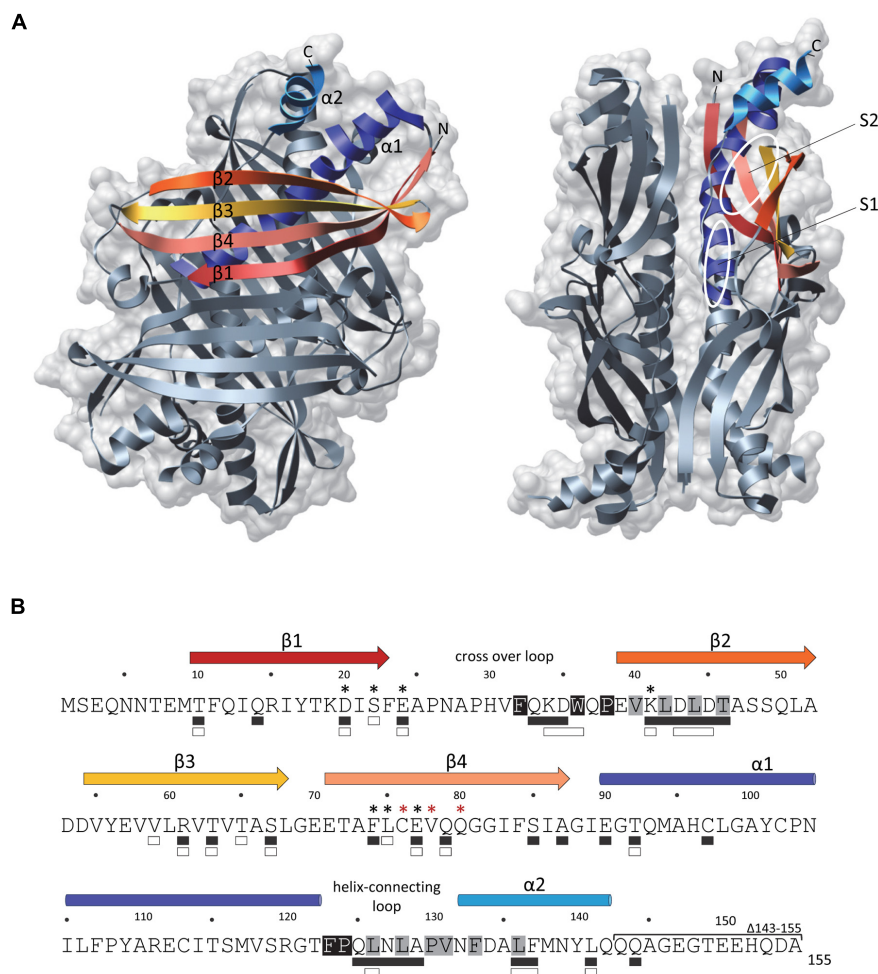
1991; Powers and Randall, 1995; Baars et al., 2006; Marani et al., 2006). Proteomic analyzes of protein aggregates that accumulate in a *secB* mutant also revealed the presence of a small number of aggregated cytosolic proteins (Baars et al., 2006; Sakr et al., 2010; see SecB Networking).

As stated above, SecB directly targets presecretory proteins to the Sec pathway via its specific interaction with the peripheral ATPase SecA: the motor component of the Sec translocon (Hartl et al., 1990; **Figure 2A**). SecA is an essential cytosolic protein of 102 kDa with an estimated cellular concentration of  $\sim 7 \mu\text{M}$  in *E. coli* (Kusters et al., 2011; Chatzi et al., 2013). SecA forms a homodimer in solution, is found either soluble or membrane-bound, and can specifically interact with translating ribosomes mainly via its N-terminal helix (Singh et al., 2014). SecB can interact with both membrane-bound and soluble SecA, albeit with a significantly lower affinity for soluble SecA ( $1.5 \mu\text{M}$  versus 30 nM Kds, respectively; den Blaauwen et al., 1997). Interaction between SecB and membrane-bound SecA is further increased in the presence of precursor proteins (Kd of  $\sim 10 \text{ nM}$ ), in order to facilitate

the targeting of precursor proteins to the translocon (Fekkes et al., 1997).

Contact regions between SecB and SecA have been studied as well. SecB mutants with amino acid substitutions at positions D20, E24, L75, and E77 that were originally selected on the basis of their export defect (Gannon and Kumamoto, 1993; Kimsey et al., 1995), were later shown to be defective for binding to SecA (Woodbury et al., 2000). Accordingly, the crystal structures of SecB revealed that all these residues localize within the negatively charged surface formed by the  $\beta$ -sheets on both sides of the tetramer (**Figure 1**; Xu et al., 2000; Dekker et al., 2003). The main SecB binding site of SecA, which encompasses the last 22 C-terminal amino acid of SecA is highly enriched in basic residues and possesses a zinc binding site required for a functional interaction with SecB (Fekkes et al., 1998, 1999). The structure of *H. influenzae* SecB in complex with the last 24 amino acids of SecA further established such a specific binding occurring mainly through electrostatic interactions, with one SecA C-terminal peptide being bound to each  $\beta$ -sheet surface on both sides of a SecB tetramer (Zhou and Xu, 2003).





**FIGURE 1 | The SecB chaperone. (A)** Three-dimensional structure of an *Escherichia coli* SecB tetramer (1QYN) in front (left) and side (right) views displayed in ribbon and molecular surface using Chimera. Secondary structural elements are colored in both representations as follows:  $\beta$ -sheet 1 in red,  $\beta$ -sheet 2 in orange,  $\beta$ -sheet 3 in yellow,  $\beta$ -sheet 4 in salmon,  $\alpha$ -helix 1 in dark blue and  $\alpha$ -helix 2 in light blue. On the side view, the proposed subsites 1 (S1) and 2 (S2) of interaction with the substrate are indicated. **(B)** The primary amino acid sequence of SecB is annotated with the secondary

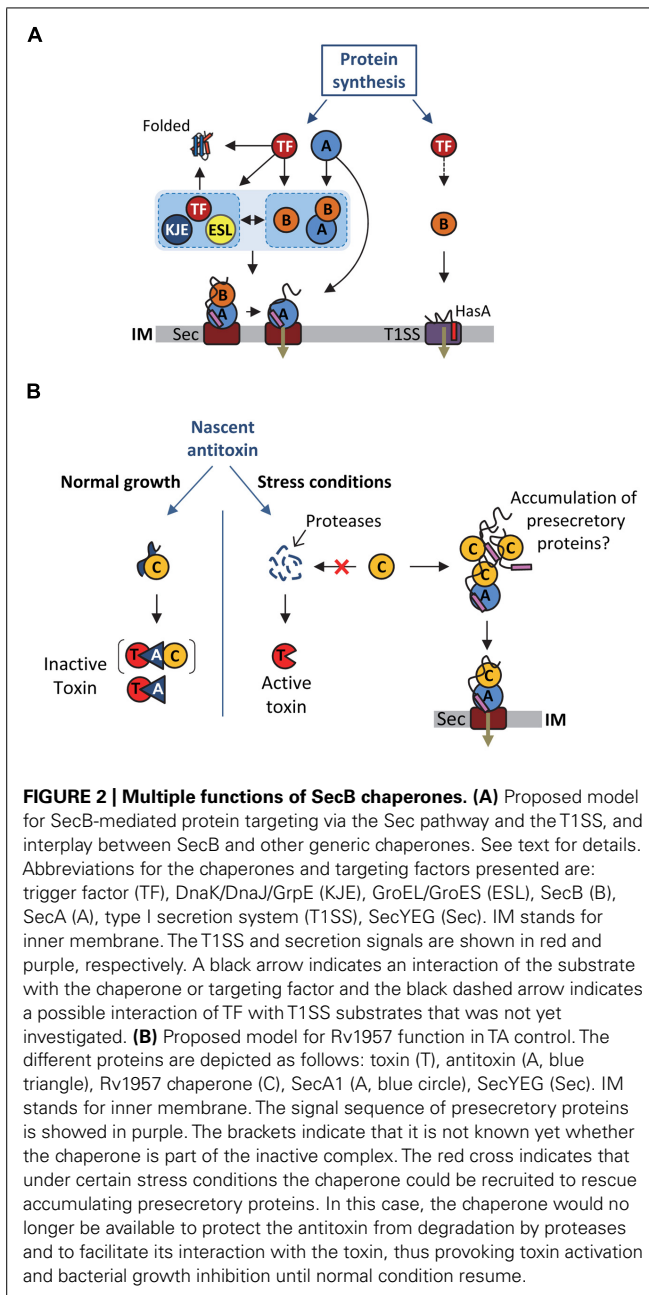
structural elements on the top with colors corresponding to **(A)**. Residues involved in the interaction with SecA or with the substrate are indicated below the sequence with white or black rectangles, respectively. Black asterisks indicate positions known to alter SecB function in export and red asterisks positions known to trigger aggregation of the protein. The C-terminal deletion 143–155 which alters the interaction with SecA is indicated. The residues predicted as being part of the subsites 1 and 2 of interaction with the substrate are highlighted in black and gray, respectively.

This is consistent with a model in which one SecA dimer binds to one SecB tetramer (Randall et al., 2005). An additional contact site between the two proteins has been described, which consists of the C-terminal  $\alpha$ -helices of SecB and the N-terminal part of SecA involved in dimerization and ribosome binding (Randall et al., 2004, 2005; Singh et al., 2014). Such interaction was proposed to trigger dissociation of the SecA dimer, thus allowing the opening of a peptide binding groove that would favor substrate transfer from SecB to SecA (Randall et al., 2005). These two surfaces of contact were confirmed by spin-labeling analyses of SecB upon SecA binding (Crane et al., 2005). Interestingly, this work also showed that the surfaces of SecB that interact with the precursor and with SecA significantly overlap, thus likely facilitating substrate transfer for translocation (**Figure 1B**; Crane et al., 2005). Efficient transfer of the precursor protein from SecB to SecA requires both a correct

interaction with SecB and the binding of the functional signal sequence to SecA, which has a strong affinity for signal sequences (Fekkes et al., 1998). To date, the precise mechanism of substrate transfer remains unknown. Once the ATP-dependent translocation process initiates, SecB is released from the Sec translocon and is now free to initiate a new cycle of binding to precursor proteins (Fekkes et al., 1997). The fact that SecB has the ability to stimulate SecA ATPase activity suggests that it could contribute to the translocation initiation process as well (Miller et al., 2002).

In addition to the post-translational targeting of the SecB-precursor complex to the SecYEG-bound SecA, recent studies suggest that SecB might be directly recruited to the preformed cytosolic SecA-precursor complex prior to SecA interaction with the protein conducting channel SecYEG (Chatzi et al., 2013; **Figure 2A**). This model is supported by the recently described





interaction of SecA with the L23 ribosomal protein platform for ribosome interacting factors at the ribosome exit tunnel (Huber et al., 2011; Singh et al., 2014). Together these data further highlight the multifaceted interaction between SecA and SecB, and its key contribution to the selective post-translational targeting of precursor proteins in *E. coli*.

### SecB AND TYPE 1 SECRETION SYSTEMS

Besides its chaperone function during protein export via the Sec pathway SecB is a key player in the secretion of the small Sec-independent HasA hemoprotein (19.3 kDa), which is part of the heme acquisition system of *Serratia marcescens* (Letoffe et al., 1994). So far, HasA is the only known type 1 secretion system

(T1SS) substrate that is strictly dependent on SecB. The T1SS, which is widespread among Gram-negative bacteria directs the one step translocation of polypeptides across both the inner and outer membranes, directly to the extracellular medium (Delepelaire, 2004). It allows secretion of proteins of diverse sizes (19–800 kDa) and functions (toxins, lipases, heme-binding, or S-layer proteins), which are presumably transported in an unfolded state via a C-terminal uncleaved secretion signal (Delepelaire, 2004; Holland et al., 2005). HasA of *Serratia marcescens* is secreted by an archetypal T1SS comprising an inner membrane ABC (ATP binding cassette) protein HasD, a periplasmic adaptor HasE, and an outer membrane channel-forming protein of the TolC family, named HasF (Letoffe et al., 1994).

SecB interacts with nascent HasA early during synthesis and holds it in an unfolded conformation competent for productive interaction with the ABC transporter HasD at the inner membrane (Figure 2A; Delepelaire and Wandersman, 1998; Debarbieux and Wandersman, 2001). In support of such antifolding activity of SecB, it has been shown that slow folding mutants of HasA are secreted independently of SecB (Wolff et al., 2003). Despite the fact that SecB allows a functional interaction between the N-terminal region of HasA and HasD, no direct interaction could be detected between SecB and the transporter (Sapriel et al., 2002, 2003; Wolff et al., 2003). Remarkably, point mutations in SecB that are known to affect its interaction with SecA (i.e., mutations D20A, E24A, L75R, and E77V; Figure 1B) exhibited very little or no effect on HasA secretion, thus indicating that SecB functions independently of SecA in this process. In contrast, SecB mutations affecting its oligomeric state and thus its chaperone function (mutations C76Y and Q80R; Kimsey et al., 1995; Muren et al., 1999) have a very strong effect on HasA secretion (Sapriel et al., 2003), suggesting that substrate binding by SecB is sufficient in this case. To date, the use of SecB generic chaperone function by SecA-independent secretion systems has only been shown for HasDEF and it remains to be determined whether other systems require similar assistance by SecB, and to what extent such chaperone redeployment could affect proper functioning of the SecA/SecB cascade *in vivo*.

### SecB NETWORKING

Significant interplay between SecB and other major cytosolic chaperones has been described (Castanie-Cornet et al., 2014; Figure 2A). The functional cooperation and/or overlap, as well as the strong genetic interactions observed between SecB, TF and DnaKJE suggest a key role for SecB as part of the chaperone network that orchestrates proper protein folding and targeting in *E. coli*. Albeit significantly less studied, a discrete link between SecB and the chaperonin GroEL has also been shown in some cases. In this part, we describe the intricate relationship between SecB and these main chaperones and discuss how SecB chaperone tasking contributes to such proteostasis network.

### SecB AND THE RIBOSOME-BOUND TRIGGER FACTOR CHAPERONE

The TF chaperone interacts with most newly synthesized polypeptides in *E. coli* (Valent et al., 1995). It is believed that about 70% of the *E. coli* cytosolic proteins interacting with TF reach their native state without further assistance (Deuerling et al., 1999; Teter

et al., 1999). TF specifically binds to the ribosomal protein L23 in the vicinity of the polypeptide exit tunnel and cycles on and off the ribosome in an ATP-independent manner (Kramer et al., 2002; Ferbitz et al., 2004; Genevaux et al., 2004). Following release from the ribosome, TF can stay bound to elongating polypeptides and facilitate substrate transfer to downstream chaperones or possibly to the Sec translocon (Crooke et al., 1988b; Kaiser et al., 2006; Raine et al., 2006; Hoffmann et al., 2010; Saio et al., 2014). TF can delay the folding of large proteins and exhibits unfolding activity (Agashe et al., 2004; Hoffmann et al., 2012; O'Brien et al., 2012), which may facilitate targeting of presecretory proteins to the Sec translocon, as observed for SecB. TF interacts with outer membrane proteins (OMPs) and several OMPs and periplasmic proteins are significantly decreased in the absence of TF (Oh et al., 2011). Remarkably, a substantial number of these exported substrates is shared between SecB and TF: this includes precursors of OmpA, OmpC, OmpF, LamB, PhoE, TolC, DegP, FkpA, OppA, Bla, and MBP (Castanie-Cornet et al., 2014). Yet, in contrast with SecB, a direct role for TF in stabilizing translocation competent precursors has only been shown for proOmpA and in this case, deletion of the *tig* gene encoding TF exhibited no significant defect on proOmpA processing (Crooke and Wickner, 1987; Crooke et al., 1988a,b). Instead, *tig* mutation was shown to accelerate translocation of several known SecB substrates, namely OmpA, OmpC, and OmpF (Lill et al., 1988; Guthrie and Wickner, 1990; Lee and Bernstein, 2002; Genevaux et al., 2004; Ullers et al., 2007) and to fully suppress both cold-sensitive and temperature-sensitive phenotypes of a *secB* null strain (Table 1; Guthrie and Wickner, 1990; Lee and Bernstein, 2002; Genevaux et al., 2004; Ullers et al., 2007). These data suggest that ribosome-bound TF could facilitate post-translational targeting of precursors by maintaining them competent either for binding to membrane-bound SecA (Gouridis et al., 2009) or for transfer to SecB, DnaKJE, or GroESL (Figure 2A; see subsections below). The fact that both TF and SecA bind to L23 at the ribosomal polypeptide exit suggests that TF could either cooperate with SecA or prevent unproductive SecA binding to precursors that first need to transit via SecB (Figure 2; Karamyshev and Johnson, 2005; Huber et al., 2011; Singh et al., 2014). Although more work is needed to shed light on such possible interplays between SecA, TF and SecB, it is important to note that both *secB* and *rpLW* (the gene encoding L23) mutations likely synergize *in vivo*, further supporting an important role for SecB in this process (Table 1).

### SecB AND THE DnaKJE CHAPERONE MACHINE

The ATP-dependent chaperone DnaK of *E. coli* is a well-characterized member of the Hsp70 chaperone family. It is an abundant cytosolic chaperone expressed constitutively and induced in response to different stresses (Genevaux et al., 2007). The DnaK chaperone cycle is tightly regulated by essential co-chaperones: (i) the DnaJ (Hsp40) co-chaperone family members that stimulate DnaK's weak ATP activity and facilitate substrate delivery to DnaK, and (ii) the nucleotide exchange factor GrpE, which mediates the dissociation of ADP and the subsequent binding of a new ATP that triggers substrate release from DnaK (Liberek et al., 1991; Harrison et al., 1997; Brehmer et al., 2001). DnaK preferentially interacts with short extended hydrophobic

polypeptide sequences accessible during *de novo* protein folding, translocation through biological membranes, during stress or within native protein complexes (Rudiger et al., 1997). In agreement with such a variety of potential interactors, the recently described *in vivo* interactome of DnaK obtained in the presence of SecB revealed that DnaK interacts with more than six hundred *E. coli* proteins at 37°C, including cytosolic (~80%), inner membrane (~11%), outer membrane (~3%) and periplasmic proteins (~3%; Calloni et al., 2012).

Most of our current knowledge concerning DnaKJE's contribution to the Sec pathway originates from studies concerning *secB* mutants and/or SecB substrates (Castanie-Cornet et al., 2014). Indeed, it has been shown that export of the SecB substrates OmpA, OmpC, and OmpF strongly relies on DnaK when protein translocation is compromised (Qi et al., 2002), and that overexpression of DnaK suppresses both the cold-sensitive phenotype of a *secB* null strain and the export defect of the SecB-dependent substrates LamB and MBP (Wild et al., 1992; Ullers et al., 2007; Castanie-Cornet et al., 2014). Although, export of both LamB and MBP is not affected by a *dnaK* mutation (Wild et al., 1992), depletion of DnaKJ in the absence of SecB showed a further decrease in the processing of these proteins and a robust accumulation of protein aggregates in the *E. coli* cytoplasm (Wild et al., 1992; Ullers et al., 2007). These aggregated proteins include known DnaK substrates and several OMPs (i.e., OmpA, OmpC, OmpX, and PhoE) previously known as SecB substrates (Ullers et al., 2007). Such a major overlap between these two chaperones is further supported by the fact that SecB substrates were recently identified as *bona fide* DnaK interactors *in vivo* (Calloni et al., 2012). This includes the OMPs OmpA, OmpC, OmpF, OmpT and OmpX, and the periplasmic proteins OppA and DegP. Accordingly, peptide binding scans revealed that SecB and DnaK share many potential binding sites in polypeptide substrates and could interact with similar regions within protein (Knoblauch et al., 1999). These data are in complete agreement with the fact that mutations in *secB* and *dnaK* (or *dnaJ*) exhibit synthetic lethality (Table 1), and that expression of DnaK is upregulated in the absence of SecB, and reciprocally (Muller, 1996; Ullers et al., 2007). These data also suggest that both chaperones could work in concert to assist the post-translational translocation of certain Sec substrates (Sakr et al., 2010). The physical interaction recently found between SecB and DnaK *in vivo* is in agreement with such hypothesis (Calloni et al., 2012).

### SecB AND THE TF/DnaK PATHWAY FOR CYTOSOLIC PROTEIN FOLDING

In addition to protein export, a role for SecB in rescuing cytosolic protein folding has been proposed. Such a SecB function has emerged from studies generally focusing on both TF and DnaK chaperones. Indeed, it has been shown that SecB overexpression efficiently rescues the severe growth defect of a chaperone-deficient strain carrying both *dnaK* and *tig* mutations, and suppresses the DnaK/TF-dependent accumulation of aggregated cytosolic proteins (Ullers et al., 2004). *In vitro* cross-linking experiments further showed that SecB is indeed capable of interacting co- and/or post-translationally with nascent RpoB in the absence of both chaperones. Such a possible SecB function is further supported by the fact that (i) SecB has preference for

unstructured stretches of polypeptides that are not specifically found in exported proteins (Knoblauch et al., 1999), (ii) SecB prevents luciferase aggregation and cooperates with DnaKJE in the refolding of luciferase *in vitro* (Knoblauch et al., 1999), and (iii) cytosolic proteins can be isolated from aggregated protein fractions in both *secB* and *secB lon* mutant strains (Baars et al., 2006; Sakr et al., 2010). More work is warranted to elucidate whether SecB indeed has cytosolic protein substrates *in vivo*. Of note, the SecB-like chaperone Rv1957 in *Mycobacterium tuberculosis* was shown to directly assist the folding of a cytosolic antitoxin, arguing for such possible SecB function in other bacteria (see part below).

### SecB AND THE CHAPERONIN GroESL

The third main molecular chaperone potentially linked to SecB in *E. coli* is the chaperonin GroESL. The ATP-dependent chaperonin GroEL is a well-characterized member of the Hsp60 chaperone family. Together with its co-chaperone GroES (Hsp10), it provides both a protected environment and a functional assistance to polypeptides generally up to 60 kDa. GroEL forms a barrel-shaped complex composed of two heptameric rings assembled back-to-back (Saibil et al., 2013). The GroEL folding cavity can be closed by a seven GroES co-chaperone lid, which allows confinement of the polypeptide. It is believed that GroESL interacts with more than 10% of the *E. coli* cytosolic proteins, including aggregation-prone proteins that are strictly chaperonin-dependent for their folding *in vivo* (Kerner et al., 2005).

Although poorly investigated, a direct involvement of GroESL in the Sec pathway has been observed and several SecB substrates have been shown to interact with or to be processed by GroEL (Kusukawa et al., 1989; Lecker et al., 1989; Phillips and Silhavy, 1990). Remarkably, five known SecB substrates were recently identified as GroEL interactors *in vivo*. These include three OMPs, namely OmpA, OmpC and OmpF, and two periplasmic proteins OppA and YncE (Watanabe et al., 1988; Kerner et al., 2005; Baars et al., 2006). In addition, GroEL was previously shown to interact with prePhoE and proOmpA *in vitro*, and to stabilize proOmpA for translocation (Kusukawa et al., 1989; Lecker et al., 1989). Although *groESL* mutations exhibit no apparent effect on proOmpA and proOmpF processing, overexpression of GroESL efficiently rescues the cold-sensitive phenotype of a *secB* null strain (unpublished data). The fact that endogenous SecB level also increases in strains with impaired GroESL is in agreement with such findings (Muller, 1996). Together these data suggest that GroESL may actively contribute to the Sec-dependent export process, perhaps rescuing SecB substrates under certain stresses or even cooperate with SecB to facilitate their transfer to SecA, as proposed for TF and DnaK.

### SecB-LIKE CHAPERONES AND TOXIN-ANTITOXIN SYSTEMS

As stated above, SecB is usually found in proteobacteria. Yet, some SecB-like sequences are also found in other taxonomic groups, including Gram-positive bacteria (Sala et al., 2013b). The major human pathogen *M. tuberculosis* also encodes a SecB-like protein, namely Rv1957, which shares 19% amino acid sequence identity with the *E. coli* SecB. The fact that mycobacteria have a well-defined and characteristic outer membrane, named the mycomembrane, with a significant number of predicted OMPs suggests that these

bacteria could make use of such SecB chaperone function for the targeting of their OMPs to the Sec translocon (Zuber et al., 2008; Mah et al., 2010). Previous work showed that Rv1957 can replace SecB export function in *E. coli*, partially restoring the processing of both proOmpA and preMBP, and complement the cold-sensitive phenotype of a *secB* mutant strain (Table 1). *In vitro*, Rv1957 also forms a tetramer in solution and efficiently prevents aggregation of the known *E. coli* SecB substrate proOmpC at a level comparable to that of SecB (Bordes et al., 2011). These results strongly suggest that Rv1957 could act as a *bona fide* SecB chaperone to assist protein export in *M. tuberculosis*.

In contrast with *E. coli* SecB, the Rv1957 encoding gene is clustered together with genes that are part of a stress-responsive type II TA system related to the HigBA family (Host Inhibition of Growth; Gupta, 2009; Ramage et al., 2009; Bordes et al., 2011; Sala et al., 2013a,b; Schuessler et al., 2013). Type II TA systems are genetic modules composed of a stable toxin and a less stable antitoxin, which interact together to form a complex in which the toxin is inactive (Gerdes et al., 2005; Goeders and Van Melderen, 2014). Under specific stress condition, the antitoxin is degraded by activated proteases, provoking the release of the active toxin, which will then act on its intracellular targets. Toxins from type II TA generally target essential cellular functions, such as translation or replication, resulting in growth inhibition. Modulation of bacterial growth by TA systems in response to environmental insults likely favors adaptation to stress (Lewis, 2010; Yamaguchi and Inouye, 2011). Remarkably, the SecB-like chaperone Rv1957 from *M. tuberculosis* specifically controls the inhibition of the HigBA TA system (Figure 2B). Indeed, Rv1957 interacts directly with the HigA antitoxin and protects it from both aggregation and degradation by proteases, thus facilitating its folding and subsequent interaction with the toxin. This chaperone function is necessary for the efficient inhibition of the toxin by the antagonistic antitoxin (Bordes et al., 2011). Such a tripartite system, named TAC for toxin-antitoxin-chaperone, is the first example of a TA system controlled by a molecular chaperone.

The hypothetical dual role of Rv1957 both as a generic chaperone potentially assisting protein export and as a specialized chaperone controlling a bacterial growth tuning system raises the question of a possible link between these two functions under certain conditions (Figure 2B). An attractive hypothesis is that in case of a compromised translocon accumulation of preproteins could compete with the antitoxin for Rv1957 binding, resulting in antitoxin degradation and subsequent toxin activation. In this model, the SecB-like chaperone would thus function as a molecular sentinel to watch over protein export. The fact that the presence of a *secB* open reading frame associated with TA modules is not unique to *M. tuberculosis* or to mycobacteria (Sala et al., 2013b) indicates that such a mechanism might be conserved (see part below).

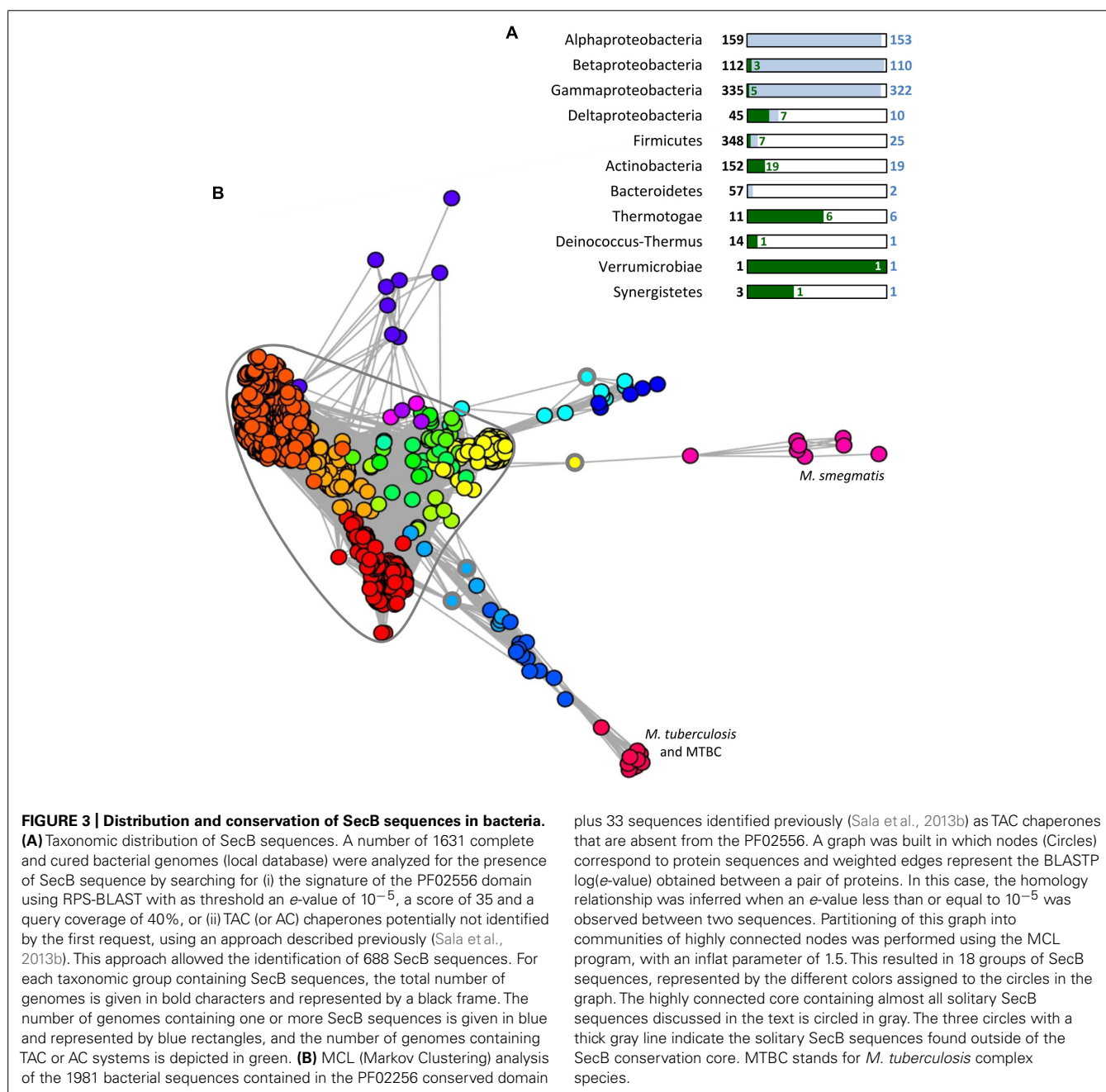
### TAXONOMIC DISTRIBUTION OF SOLITARY AND TA-ASSOCIATED SecB

It has been proposed that SecB appeared in the last common ancestor of  $\alpha$ -,  $\beta$ -, and  $\gamma$ -proteobacteria and that its conservation is linked to the presence of an outer membrane, and



thus an increased need in protein export (van der Sluis and Driessen, 2006). Nevertheless, analysis of the taxonomic repartition of PF02556, the Pfam domain characterizing SecB sequences (<http://pfam.sanger.ac.uk/>) in a set of 1631 complete and cured bacterial genomes revealed the presence of this domain in seven groups outside proteobacteria. Noticeably, these groups are mainly composed of diderm bacteria, except from the Firmicutes phylum (**Figure 3A**) and in most cases, SecB sequences occur at low frequency when compared to the total number of genomes. This is in sharp contrast with  $\alpha$ -,  $\beta$ -, and  $\gamma$ -proteobacteria, where most of the genomes contain at least one SecB sequence (Sala et al., 2013b; **Figure 3A**).

A subset of SecB genes, representing approximately 7.5% of the total number of SecB sequences (52/688), are associated with genes encoding TA systems (or in some cases antitoxin genes alone), as observed for the TAC system of *M. tuberculosis*. This suggests that these putative SecB chaperones might function in the specific control of their cognate TA systems in a manner comparable to that of Rv1957 (Bordes et al., 2011; Sala et al., 2013b). When SecB sequences are present in groups outside of  $\alpha$ -,  $\beta$ -, and  $\gamma$ -proteobacteria, they seem to preferentially associate with a TA system (63%, 44/70). In this case, the vast majority of the genomes (>90%) do not possess an additional copy of solitary SecB. Interestingly, TA systems associated with SecB sequences often belong



to different families of toxins and/or antitoxins, strongly suggesting that the event of association of a SecB encoding gene with a TA module is a widespread mechanism that occurred several times during evolution (Sala et al., 2013b). The possible involvement of these SecB chaperones in Sec-dependent protein export remains to be determined.

Further analysis on all the bacterial sequences available on the Pfam server for the PF02556 conserved domain (i.e., from both complete and in progress genomes) was performed to study the homology links between SecB sequences using a graph partitioning approach. This revealed that solitary SecB sequences are grouped together in a highly connected core, reflecting a high level of conservation (Figure 3B). In this core, several SecB communities (corresponding to the different colors) are well-defined and generally correspond to the taxonomy: the red family contains mainly  $\alpha$ -proteobacterial sequences, the dark orange mainly  $\gamma$ -proteobacterial sequences and the light orange mainly  $\beta$ -proteobacterial sequences. Another clearly defined group that emerges from this core, in yellow, contains mainly sequences from *Streptococcus pneumoniae* strains (158/196). The other communities within the core are poorly defined and mainly correspond to other Firmicutes sequences. Most of the TAC (or AC) chaperones are grouped in eight different communities, which seem to have diverged from the solitary SecB core from distinct origins (Figure 3B). The light orange group within the core also contains four TA-associated SecB sequences from  $\delta$ -proteobacteria, thus strongly suggesting a common evolutionary history between TAC chaperones and canonical solitary SecB (Sala et al., 2013b). Yet, in sharp contrast with solitary SecB, the groups of TAC chaperones do not follow the taxonomy and most of them are comprised in regions containing horizontal gene transfer signatures, as it is the case for classical TA systems (Makarova et al., 2009; Sala et al., 2013b).

Paralogs of SecA, SecE, SecY, and SecG are found in Actinobacteria and Firmicutes, either forming a parallel pathway with a dedicated translocon or exploiting the generic Sec translocon to export a specific set of substrates, as it is the case for SecA2 in mycobacteria (Rigel et al., 2009; Sullivan et al., 2012). Interestingly, among the 1631 complete bacterial genomes analyzed 26 of them are predicted to have more than one solitary SecB sequence (up to three in *Acetobacter pasteurianus* IFO 3283-01-42C), eight genomes contain both a solitary and a TA associated SecB, one genome contains two solitary and one TA associated SecB, and two genomes contain two TAC or AC. These additional SecB sequences could function as specialized chaperones for the control of TA systems or for the export of specific substrates, or as generic chaperone induced in response to certain stress conditions. Interestingly, single deletion of *secB1* or *secB2* from *Francisella tularensis* subsp. *novicida* exhibits a reduced biofilm formation, suggesting that both chaperone paralogs participate in the secretion of specific factors important for the attachment to abiotic surfaces (Margolis et al., 2010). Remarkably, the double *secB* mutant was not viable, suggesting that in the case of *F. tularensis*, SecB chaperones have overlapping functions essential for bacterial survival (Margolis et al., 2010).

## CONCLUDING REMARKS

Extensive genetic and biochemical analyses of SecB chaperone tasking have undoubtedly revealed its key cellular roles as part of the network of generic chaperones that orchestrate proteostasis in *E. coli*. The fact that SecB binds its substrates in a non-native state and prevents their unproductive folding is in agreement with its major role in delivering translocation competent proteins to the inner membrane, as observed for a large number of Sec-dependent presecretory proteins, for the ABC transporter substrate HasA, and perhaps for other proteins whose secretion relies on specific secretion systems that lack dedicated chaperones.

In contrast with the specific and well-described cooperative cascade between SecB and SecA during post-translational targeting of Sec-dependent precursors, the interplay between SecB, TF and DnaK remains poorly understood, and there is a clear lack of knowledge about the substrates that are shared between the three chaperones *in vivo*. In addition, it is not known whether these chaperones actively cooperate to facilitate export of certain proteins, both under normal and stress conditions, and to what extent such cooperation influences early partitioning of newly synthesized proteins. Similarly, it remains to be determined whether some cytosolic proteins or protein complexes do require SecB for their folding and/or assembly, as it is the case for the SecB-like protein Rv1957 and its TA system in *M. tuberculosis*. These are truly open questions that need to be addressed.

The relatively frequent association of SecB proteins with different TA families is intriguing and may reveal interesting new SecB functions, perhaps reflecting a link between toxin activation and membrane jamming. In this respect, an important mechanistic issue will be to determine how TA systems have acquired such a unique addiction for SecB chaperones. Finally, the sporadic presence of solitary SecB-like proteins in monoderm bacteria also suggests novel SecB functions to be discovered.

## ACKNOWLEDGMENTS

We thank Petra Langendijk and Olivera Francetic for insightful discussion and Gwennaele Fichant's group for sharing their in-house bacterial genome database. This work was supported by a French MENRT fellowship and an FRM grant (FDT20140930836) to Ambre Sala and a mycoTAC ANR grant (ANR-13-BSV8-0010-01) to Pierre Genevaux.

## REFERENCES

- Agashe, V. R., Guha, S., Chang, H. C., Genevaux, P., Hayer-Hartl, M., Stemp, M., et al. (2004). Function of trigger factor and DnaK in multidomain protein folding: increase in yield at the expense of folding speed. *Cell* 117, 199–209. doi: 10.1016/S0092-8674(04)00299-5
- Altman, E., Kumamoto, C. A., and Emr, S. D. (1991). Heat-shock proteins can substitute for SecB function during protein export in *Escherichia coli*. *EMBO J.* 10, 239–245.
- Baars, L., Ytterberg, A. J., Drew, D., Wagner, S., Thilo, C., Van Wijk, K. J., et al. (2006). Defining the role of the *Escherichia coli* chaperone SecB using comparative proteomics. *J. Biol. Chem.* 281, 10024–10034. doi: 10.1074/jbc.M509929200
- Babu, M., Diaz-Mejia, J. J., Vlasblom, J., Gagarinova, A., Phanse, S., Graham, C., et al. (2011). Genetic interaction maps in *Escherichia coli* reveal functional crosstalk among cell envelope biogenesis pathways. *PLoS Genet.* 7:e1002377. doi: 10.1371/journal.pgen.1002377
- Bechtluft, P., Nouwen, N., Tans, S. J., and Driessen, A. J. M. (2010). SecB-a chaperone dedicated to protein translocation. *Mol. Biosyst.* 6, 620–627. doi: 10.1039/b915435c



- Bechtluft, P., Van Leeuwen, R. G. H., Tyreman, M., Tomkiewicz, D., Nouwen, N., Tepper, H. L., et al. (2007). Direct observation of chaperone-induced changes in a protein folding pathway. *Science* 318, 1458–1461. doi: 10.1126/science.1144972
- Bordes, P., Cirinesi, A. M., Ummels, R., Sala, A., Sakr, S., Bitter, W., et al. (2011). SecB-like chaperone controls a toxin-antitoxin stress-responsive system in *Mycobacterium tuberculosis*. *Proc. Natl. Acad. Sci. U.S.A.* 108, 8438–8443. doi: 10.1073/pnas.1101189108
- Brehmer, D., Rudiger, S., Gassler, C. S., Klostermeier, D., Packschies, L., Reinstein, J., et al. (2001). Tuning of chaperone activity of Hsp70 proteins by modulation of nucleotide exchange. *Nat. Struct. Biol.* 8, 427–432. doi: 10.1038/87588
- Bukau, B., Deuerling, E., Pfund, C., and Craig, E. A. (2000). Getting newly synthesized proteins into shape. *Cell* 101, 119–122. doi: 10.1016/S0092-8674(00)80806-5
- Calloni, G., Chen, T., Schermann, S. M., Chang, H. C., Genevaux, P., Agostini, F., et al. (2012). DnaK functions as a central hub in the *E. coli* chaperone network. *Cell Rep.* 1, 251–264. doi: 10.1016/j.celrep.2011.12.007
- Castanie-Cornet, M. P., Bruel, N., and Genevaux, P. (2014). Chaperone networking facilitates protein targeting to the bacterial cytoplasmic membrane. *Biochim. Biophys. Acta* 1843, 1442–1456. doi: 10.1016/j.bbamcr.2013.11.007
- Chatzi, K. E., Sardis, M. F., Karamanou, S., and Economou, A. (2013). Breaking on through to the other side: protein export through the bacterial Sec system. *Biochem. J.* 449, 25–37. doi: 10.1042/BJ20121227
- Collier, D. N., Bankaitis, V. A., Weiss, J. B., and Bassford, P. J. Jr. (1988). The antifolding activity of SecB promotes the export of the *E. coli* maltose-binding protein. *Cell* 53, 273–283. doi: 10.1016/0092-8674(88)90389-3
- Crane, J. M., Mao, C. F., Lilly, A. A., Smith, V. F., Suo, Y. Y., Hubbell, W. L., et al. (2005). Mapping of the docking of SecA onto the chaperone SecB by site-directed spin labeling: insight into the mechanism of ligand transfer during protein export. *J. Mol. Biol.* 353, 295–307. doi: 10.1016/j.jmb.2005.08.022
- Crane, J. M., Suo, Y. Y., Lilly, A. A., Mao, C. F., Hubbell, W. L., and Randall, L. L. (2006). Sites of interaction of a precursor polypeptide on the export chaperone SecB mapped by site-directed spin labeling. *J. Mol. Biol.* 363, 63–74. doi: 10.1016/j.jmb.2006.07.021
- Crooke, E., Brundage, L., Rice, M., and Wickner, W. (1988a). ProOmpA spontaneously folds in a membrane assembly competent state which trigger factor stabilizes. *EMBO J.* 7, 1831–1835.
- Crooke, E., Guthrie, B., Lecker, S., Lill, R., and Wickner, W. (1988b). ProOmpA is stabilized for membrane translocation by either purified *E. coli* trigger factor or canine signal recognition particle. *Cell* 54, 1003–1011. doi: 10.1016/0092-8674(88)90115-8
- Crooke, E., and Wickner, W. (1987). Trigger factor: a soluble protein that folds pro-OmpA into a membrane-assembly-competent form. *Proc. Natl. Acad. Sci. U.S.A.* 84, 5216–5220. doi: 10.1073/pnas.84.15.5216
- Dalbey, R. E., Wang, P., and Kuhn, A. (2011). Assembly of bacterial inner membrane proteins. *Annu. Rev. Biochem.* 80, 161–187. doi: 10.1146/annurev-biochem-060409-092524
- Danese, P. N., Murphy, C. K., and Silhavy, T. J. (1995). Multicopy suppression of cold-sensitive sec mutations in *Escherichia coli*. *J. Bacteriol.* 177, 4969–4973.
- Debarbieux, L., and Wandersman, C. (2001). Folded HasA inhibits its own secretion through its ABC exporter. *EMBO J.* 20, 4657–4663. doi: 10.1093/emboj/20.17.4657
- Dekker, C., De Kruijff, B., and Gros, P. (2003). Crystal structure of SecB from *Escherichia coli*. *J. Struct. Biol.* 144, 313–319. doi: 10.1016/j.jsb.2003.09.012
- Delepelaire, P. (2004). Type I secretion in gram-negative bacteria. *Biochim. Biophys. Acta* 1694, 149–161. doi: 10.1016/j.bbamcr.2004.05.001
- Delepelaire, P., and Wandersman, C. (1998). The SecB chaperone is involved in the secretion of the *Serratia marcescens* HasA protein through an ABC transporter. *EMBO J.* 17, 936–944. doi: 10.1093/emboj/17.4.936
- den Blaauwen, T., Terpetschnig, E., Lakowicz, J. R., and Driessen, A. J. (1997). Interaction of SecB with soluble SecA. *FEBS Lett.* 416, 35–38. doi: 10.1016/S0014-5793(97)01142-3
- Derman, A. I., Puziss, J. W., Bassford, P. J. Jr., and Beckwith, J. (1993). A signal sequence is not required for protein export in prlA mutants of *Escherichia coli*. *EMBO J.* 12, 879–888.
- Deuerling, E., Schulze-Specking, A., Tomoyasu, T., Mogk, A., and Bukau, B. (1999). Trigger factor and DnaK cooperate in folding of newly synthesized proteins. *Nature* 400, 693–696. doi: 10.1038/23301
- du Plessis, D. J., Nouwen, N., and Driessen, A. J. (2011). The Sec translocase. *Biochim. Biophys. Acta* 1808, 851–865. doi: 10.1016/j.bbamem.2010.08.016
- Fekkes, P., De Wit, J. G., Boorsma, A., Friesen, R. H., and Driessen, A. J. (1999). Zinc stabilizes the SecB binding site of SecA. *Biochemistry* 38, 5111–5116. doi: 10.1021/bi982818r
- Fekkes, P., De Wit, J. G., Van Der Wolk, J. P., Kimsey, H. H., Kumamoto, C. A., and Driessen, A. J. (1998). Preprotein transfer to the *Escherichia coli* translocase requires the co-operative binding of SecB and the signal sequence to SecA. *Mol. Microbiol.* 29, 1179–1190. doi: 10.1046/j.1365-2958.1998.00997.x
- Fekkes, P., Van Der Does, C., and Driessen, A. J. (1997). The molecular chaperone SecB is released from the carboxy-terminus of SecA during initiation of precursor protein translocation. *EMBO J.* 16, 6105–6113. doi: 10.1093/emboj/16.20.6105
- Ferbitz, L., Maier, T., Patzelt, H., Bukau, B., Deuerling, E., and Ban, N. (2004). Trigger factor in complex with the ribosome forms a molecular cradle for nascent proteins. *Nature* 431, 590–596. doi: 10.1038/nature02899
- Francetic, O., Hanson, M. P., and Kumamoto, C. A. (1993). PrlA suppression of defective export of maltose-binding protein in SecB mutants of *Escherichia coli*. *J. Bacteriol.* 175, 4036–4044.
- Gannon, P. M., and Kumamoto, C. A. (1993). Mutations of the molecular chaperone protein SecB which alter the interaction between SecB and maltose-binding protein. *J. Biol. Chem.* 268, 1590–1595.
- Gannon, P. M., Li, P., and Kumamoto, C. A. (1989). The mature portion of *Escherichia coli* maltose-binding protein (MBP) determines the dependence of MBP on SecB for export. *J. Bacteriol.* 171, 813–818.
- Genevaux, P., Georgopoulos, C., and Kelley, W. L. (2007). The Hsp70 chaperone machines of *Escherichia coli*: a paradigm for the repartition of chaperone functions. *Mol. Microbiol.* 66, 840–857. doi: 10.1111/j.1365-2958.2007.05961.x
- Genevaux, P., Keppel, F., Schwager, F., Langendijk-Genevaux, P. S., Hartl, F. U., and Georgopoulos, C. (2004). In vivo analysis of the overlapping functions of DnaK and trigger factor. *EMBO Rep.* 5, 195–200. doi: 10.1038/sj.embor.7400067
- Gerdas, K., Christensen, S. K., and Lobner-Olesen, A. (2005). Prokaryotic toxin-antitoxin stress response loci. *Nat. Rev. Microbiol.* 3, 371–382. doi: 10.1038/nrmicro1147
- Goeders, N., and Van Melderen, L. (2014). Toxin-antitoxin systems as multilevel interaction systems. *Toxins (Basel)* 6, 304–324. doi: 10.3390/toxins6010304
- Gouridis, G., Karamanou, S., Gelis, I., Kalodimos, C. G., and Economou, A. (2009). Signal peptides are allosteric activators of the protein translocase. *Nature* 462, 363–367. doi: 10.1038/nature08559
- Grabowicz, M., Yeh, J., and Silhavy, T. J. (2013). Dominant negative LptE mutation that supports a role for LptE as a plug in the LptD barrel. *J. Bacteriol.* 195, 1327–1334. doi: 10.1128/JB.02142-12
- Gupta, A. (2009). Killing activity and rescue function of genome-wide toxin-antitoxin loci of *Mycobacterium tuberculosis*. *FEMS Microbiol. Lett.* 290, 45–53. doi: 10.1111/j.1574-6968.2008.01400.x
- Guthrie, B., and Wickner, W. (1990). Trigger factor depletion or overproduction causes defective cell division but does not block protein export. *J. Bacteriol.* 172, 5555–5562.
- Hardy, S. J., and Randall, L. L. (1991). A kinetic partitioning model of selective binding of nonnative proteins by the bacterial chaperone SecB. *Science* 251, 439–443. doi: 10.1126/science.1989077
- Harrison, C. J., Hayer-Hartl, M., Di Liberto, M., Hartl, F., and Kuriyan, J. (1997). Crystal structure of the nucleotide exchange factor GrpE bound to the ATPase domain of the molecular chaperone DnaK. *Science* 276, 431–435. doi: 10.1126/science.276.5311.431
- Hartl, F. U., Lecker, S., Schiebel, E., Hendrick, J. P., and Wickner, W. (1990). The binding cascade of SecB to SecA to SecY/E mediates preprotein targeting to the *E. coli* plasma membrane. *Cell* 63, 269–279. doi: 10.1016/0092-8674(90)90160-G
- Hayashi, S., and Wu, H. C. (1985). Accumulation of prolipoprotein in *Escherichia coli* mutants defective in protein secretion. *J. Bacteriol.* 161, 949–954.
- Hoffmann, A., Becker, A. H., Zachmann-Brand, B., Deuerling, E., Bukau, B., and Kramer, G. (2012). Concerted action of the ribosome and the associated chaperone trigger factor confines nascent polypeptide folding. *Mol. Cell* 48, 63–74. doi: 10.1016/j.molcel.2012.07.018
- Hoffmann, A., Bukau, B., and Kramer, G. (2010). Structure and function of the molecular chaperone Trigger Factor. *Biochim. Biophys. Acta* 1803, 650–661. doi: 10.1016/j.bbamcr.2010.01.017

- Holland, I. B., Schmitt, L., and Young, J. (2005). Type 1 protein secretion in bacteria, the ABC-transporter dependent pathway (review). *Mol. Membr. Biol.* 22, 29–39. doi: 10.1080/09687860500042013
- Huber, D., Rajagopalan, N., Preissler, S., Rocco, M. A., Merz, F., Kramer, G., et al. (2011). SecA interacts with ribosomes in order to facilitate posttranslational translocation in bacteria. *Mol. Cell* 41, 343–353. doi: 10.1016/j.molcel.2010.12.028
- Jin, T., and Inouye, M. (1995). Identification of the genes in multicopy plasmids affecting ompC and ompF expression in *Escherichia coli*. *FEMS Microbiol. Lett.* 133, 225–231. doi: 10.1111/j.1574-6968.1995.tb07889.x
- Kaiser, C. M., Chang, H. C., Agashe, V. R., Lakshmipathy, S. K., Etchells, S. A., Hayer-Hartl, M., et al. (2006). Real-time observation of trigger factor function on translating ribosomes. *Nature* 444, 455–460. doi: 10.1038/nature05225
- Karamyshev, A. L., and Johnson, A. E. (2005). Selective SecA association with signal sequences in ribosome-bound nascent chains: a potential role for SecA in ribosome targeting to the bacterial membrane. *J. Biol. Chem.* 280, 37930–37940. doi: 10.1074/jbc.M509100200
- Kerner, M. J., Naylor, D. J., Ishihama, Y., Maier, T., Chang, H. C., Stines, A. P., et al. (2005). Proteome-wide analysis of chaperonin-dependent protein folding in *Escherichia coli*. *Cell* 122, 209–220. doi: 10.1016/j.cell.2005.05.028
- Khisty, V. J., Munske, G. R., and Randall, L. L. (1995). Mapping of the binding frame for the chaperone SecB within a natural ligand, galactose-binding protein. *J. Biol. Chem.* 270, 25920–25927. doi: 10.1074/jbc.270.43.25920
- Kim, Y. E., Hipp, M. S., Bracher, A., Hayer-Hartl, M., and Hartl, F. U. (2013). Molecular chaperone functions in protein folding and proteostasis. *Annu. Rev. Biochem.* 82, 323–355. doi: 10.1146/annurev-biochem-060208-092442
- Kimsey, H. H., Dagarag, M. D., and Kumamoto, C. A. (1995). Diverse effects of mutation on the activity of the *Escherichia coli* export chaperone SecB. *J. Biol. Chem.* 270, 22831–22835. doi: 10.1074/jbc.270.39.22831
- Knoblauch, N. T., Rudiger, S., Schonfeld, H. J., Driessen, A. J., Schneider-Mergener, J., and Bukau, B. (1999). Substrate specificity of the SecB chaperone. *J. Biol. Chem.* 274, 34219–34225. doi: 10.1074/jbc.274.48.34219
- Kramer, G., Boehringer, D., Ban, N., and Bukau, B. (2009). The ribosome as a platform for co-translational processing, folding and targeting of newly synthesized proteins. *Nat. Struct. Mol. Biol.* 16, 589–597. doi: 10.1038/nsmb.1614
- Kramer, G., Rauch, T., Rist, W., Vorderwulbecke, S., Patzelt, H., Schulze-Specking, A., et al. (2002). L23 protein functions as a chaperone docking site on the ribosome. *Nature* 419, 171–174. doi: 10.1038/nature01047
- Kumamoto, C. A., and Beckwith, J. (1983). Mutations in a new gene, secB, cause defective protein localization in *Escherichia coli*. *J. Bacteriol.* 154, 253–260.
- Kumamoto, C. A., and Beckwith, J. (1985). Evidence for specificity at an early step in protein export in *Escherichia coli*. *J. Bacteriol.* 163, 267–274.
- Kusters, I., Van Den Bogaart, G., Kedrov, A., Krasnikov, V., Fulyani, F., Poolman, B., et al. (2011). Quaternary structure of SecA in solution and bound to SecYEG probed at the single molecule level. *Structure* 19, 430–439. doi: 10.1016/j.str.2010.12.016
- Kusters, R., De Vrije, T., Breukink, E., and De Kruijff, B. (1989). SecB protein stabilizes a translocation-competent state of purified prePhoE protein. *J. Biol. Chem.* 264, 20827–20830.
- Kusukawa, N., Yura, T., Ueguchi, C., Akiyama, Y., and Ito, K. (1989). Effects of mutations in heat-shock genes groES and groEL on protein export in *Escherichia coli*. *EMBO J.* 8, 3517–3521.
- Lamiet, A. A., Kumamoto, C. A., and Pluckthun, A. (1991). Folding in vitro and transport in vivo of pre-Beta-Lactamase are SecB independent. *Mol. Microbiol.* 5, 117–122. doi: 10.1111/j.1365-2958.1991.tb01832.x
- Lecker, S., Lill, R., Ziegelhoffer, T., Georgopoulos, C., Bassford, P. J., Kumamoto, C. A., et al. (1989). Three pure chaperone proteins of *Escherichia coli* SecB, trigger factor and groEL-form soluble complexes with precursor proteins in vitro. *EMBO J.* 8, 2703–2709.
- Lee, H. C., and Bernstein, H. D. (2002). Trigger factor retards protein export in *Escherichia coli*. *J. Biol. Chem.* 277, 43527–43535. doi: 10.1074/jbc.M205950200
- Letoffe, S., Ghigo, J. M., and Wandersman, C. (1994). Secretion of the *Serratia marcescens* HasA protein by an ABC transporter. *J. Bacteriol.* 176, 5372–5377.
- Lewis, K. (2010). Persister cells. *Annu. Rev. Microbiol.* 64, 357–372. doi: 10.1146/annurev.micro.112408.134306
- Liberek, K., Marszałek, J., Ang, D., Georgopoulos, C., and Zylicz, M. (1991). *Escherichia coli* DnaJ and GrpE heat shock proteins jointly stimulate ATPase activity of DnaK. *Proc. Natl. Acad. Sci. U.S.A.* 88, 2874–2878. doi: 10.1073/pnas.88.7.2874
- Lill, R., Crooke, E., Guthrie, B., and Wickner, W. (1988). The “trigger factor cycle” includes ribosomes, presecretory proteins, and the plasma membrane. *Cell* 54, 1013–1018. doi: 10.1016/0092-8674(88)90116-X
- Liu, G., Topping, T. B., and Randall, L. L. (1989). Physiological role during export for the retardation of folding by the leader peptide of maltose-binding protein. *Proc. Natl. Acad. Sci. U.S.A.* 86, 9213–9217. doi: 10.1073/pnas.86.23.9213
- Luirink, J., and Sinning, I. (2004). SRP-mediated protein targeting: structure and function revisited. *Biochim. Biophys. Acta* 1694, 17–35. doi: 10.1016/j.bbamcr.2004.03.013
- Mah, N., Perez-Iratxeta, C., and Andrade-Navarro, M. A. (2010). Outer membrane pore protein prediction in mycobacteria using genomic comparison. *Microbiology* 156, 2506–2515. doi: 10.1099/mic.0.040089-0
- Makarova, K. S., Wolf, Y. I., and Koonin, E. V. (2009). Comprehensive comparative-genomic analysis of type 2 toxin-antitoxin systems and related mobile stress response systems in prokaryotes. *Biol. Direct* 4, 19. doi: 10.1186/1745-6150-4-19
- Marani, P., Wagner, S., Baars, L., Genevaux, P., De Gier, J. W., Nilsson, I., et al. (2006). New *Escherichia coli* outer membrane proteins identified through prediction and experimental verification. *Protein Sci.* 15, 884–889. doi: 10.1110/ps.051889506
- Margolis, J. J., El-Etr, S., Joubert, L. M., Moore, E., Robison, R., Rasley, A., et al. (2010). Contributions of *Francisella tularensis* subsp. novicida chitinases and Sec secretion system to biofilm formation on chitin. *Appl. Environ. Microbiol.* 76, 596–608. doi: 10.1128/AEM.02037-09
- McFarland, L., Francetic, O., and Kumamoto, C. A. (1993). A mutation of *Escherichia coli* SecA protein that partially compensates for the absence of SecB. *J. Bacteriol.* 175, 2255–2262.
- Miller, A., Wang, L., and Kendall, D. A. (2002). SecB modulates the nucleotide-bound state of SecA and stimulates ATPase activity. *Biochemistry* 41, 5325–5332. doi: 10.1021/bi025639p
- Muller, J. P. (1996). Influence of impaired chaperone or secretion function on SecB production in *Escherichia coli*. *J. Bacteriol.* 178, 6097–6104.
- Muren, E. M., Suci, D., Topping, T. B., Kumamoto, C. A., and Randall, L. L. (1999). Mutational alterations in the homotetrameric chaperone SecB that implicate the structure as dimer of dimers. *J. Biol. Chem.* 274, 19397–19402. doi: 10.1074/jbc.274.27.19397
- Nichols, R. J., Sen, S., Choo, Y. J., Beltrao, P., Zietek, M., Chaba, R., et al. (2011). Phenotypic landscape of a bacterial cell. *Cell* 144, 143–156. doi: 10.1016/j.cell.2010.11.052
- O'Brien, E. P., Christodoulou, J., Vendruscolo, M., and Dobson, C. M. (2012). Trigger factor slows co-translational folding through kinetic trapping while sterically protecting the nascent chain from aberrant cytosolic interactions. *J. Am. Chem. Soc.* 134, 10920–10932. doi: 10.1021/ja302305u
- Oh, E., Becker, A. H., Sandikci, A., Huber, D., Chaba, R., Gloge, F., et al. (2011). Selective ribosome profiling reveals the cotranslational chaperone action of trigger factor in vivo. *Cell* 147, 1295–1308. doi: 10.1016/j.cell.2011.10.044
- Palmer, T., and Berks, B. C. (2012). The twin-arginine translocation (Tat) protein export pathway. *Nat. Rev. Microbiol.* 10, 483–496. doi: 10.1038/nrmicro2814
- Patel, R., Smith, S. M., and Robinson, C. (2014). Protein transport by the bacterial Tat pathway. *Biochim. Biophys. Acta* 1843, 1620–1628. doi: 10.1016/j.bbamcr.2014.02.013
- Phillips, G. J., and Silhavy, T. J. (1990). Heat-shock proteins DnaK and GroEL facilitate export of LacZ hybrid proteins in *E. coli*. *Nature* 344, 882–884. doi: 10.1038/344882a0
- Powers, E. L., and Randall, L. L. (1995). Export of periplasmic galactose-binding protein in *Escherichia coli* depends on the chaperone SecB. *J. Bacteriol.* 177, 1906–1907.
- Qi, H. Y., Hyndman, J. B., and Bernstein, H. D. (2002). DnaK promotes the selective export of outer membrane protein precursors in SecA-deficient *Escherichia coli*. *J. Biol. Chem.* 277, 51077–51083. doi: 10.1074/jbc.M209238200
- Raine, A., Lovmar, M., Wikberg, J., and Ehrenberg, M. (2006). Trigger factor binding to ribosomes with nascent peptide chains of varying lengths and sequences. *J. Biol. Chem.* 281, 28033–28038. doi: 10.1074/jbc.M605753200
- Ramage, H. R., Connolly, L. E., and Cox, J. S. (2009). Comprehensive functional analysis of *Mycobacterium tuberculosis* toxin-antitoxin systems: implications for pathogenesis, stress responses, and evolution. *PLoS Genet.* 5:e1000767. doi: 10.1371/journal.pgen.1000767

- Randall, L. L., Crane, J. M., Lilly, A. A., Liu, G. P., Mao, C. F., Patel, C. N., et al. (2005). Asymmetric binding between SecA and SecB two symmetric proteins: implications for function in export. *J. Mol. Biol.* 348, 479–489. doi: 10.1016/j.jmb.2005.02.036
- Randall, L. L., Crane, J. M., Liu, G., and Hardy, S. J. (2004). Sites of interaction between SecA and the chaperone SecB, two proteins involved in export. *Protein Sci.* 13, 1124–1133. doi: 10.1110/ps.03410104
- Randall, L. L., and Hardy, S. J. (1995). High selectivity with low specificity: how SecB has solved the paradox of chaperone binding. *Trends Biochem. Sci.* 20, 65–69. doi: 10.1016/S0968-0004(00)88959-8
- Randall, L. L., and Hardy, S. J. (2002). SecB, one small chaperone in the complex milieu of the cell. *Cell. Mol. Life Sci.* 59, 1617–1623. doi: 10.1007/PL00012488
- Rigel, N. W., Gibbons, H. S., Mccann, J. R., McDonough, J. A., Kurtz, S., and Braunstein, M. (2009). The accessory SecA2 system of mycobacteria requires ATP binding and the canonical SecA1. *J. Biol. Chem.* 284, 9927–9936. doi: 10.1074/jbc.M900325200
- Rudiger, S., Germeroth, L., Schneider-Mergener, J., and Bukau, B. (1997). Substrate specificity of the DnaK chaperone determined by screening cellulose-bound peptide libraries. *EMBO J.* 16, 1501–1507. doi: 10.1093/emboj/16.7.1501
- Saibil, H. R., Fenton, W. A., Clare, D. K., and Horwich, A. L. (2013). Structure and allostery of the chaperonin GroEL. *J. Mol. Biol.* 425, 1476–1487. doi: 10.1016/j.jmb.2012.11.028
- Saio, T., Guan, X., Rossi, P., Economou, A., and Kalodimos, C. G. (2014). Structural basis for protein antiaggregation activity of the trigger factor chaperone. *Science* 344:1250494. doi: 10.1126/science.1250494
- Sakr, S., Cirinesi, A. M., Ullers, R. S., Schwager, F., Georgopoulos, C., and Genevau, P. (2010). Lon protease quality control of presecretory proteins in *Escherichia coli* and its dependence on the SecB and DnaJ (Hsp40) chaperones. *J. Biol. Chem.* 285, 23504–23512. doi: 10.1074/jbc.M110.133058
- Sala, A., Bordes, P., Fichant, G., and Genevau, P. (2013a). “Toxin-antitoxin loci in *Mycobacterium tuberculosis*,” in *Prokaryotic Toxin-Antitoxins*, ed. K. Gerdes (Berlin: Springer), 295–314.
- Sala, A., Calderon, V., Bordes, P., and Genevau, P. (2013b). TAC from *Mycobacterium tuberculosis*: a paradigm for stress-responsive toxin-antitoxin systems controlled by SecB-like chaperones. *Cell Stress Chaperones* 18, 129–135. doi: 10.1007/s12192-012-0396-5
- Sapriel, G., Wandersman, C., and Delepelaire, P. (2002). The N terminus of the HasA protein and the SecB chaperone cooperate in the efficient targeting and secretion of HasA via the ATP-binding cassette transporter. *J. Biol. Chem.* 277, 6726–6732. doi: 10.1074/jbc.M108632200
- Sapriel, G., Wandersman, C., and Delepelaire, P. (2003). The SecB chaperone is bifunctional in *Serratia marcescens*: SecB is involved in the Sec pathway and required for HasA secretion by the ABC transporter. *J. Bacteriol.* 185, 80–88. doi: 10.1128/JB.185.1.80-88.2003
- Saraogi, I., and Shan, S. O. (2014). Co-translational protein targeting to the bacterial membrane. *Biochim. Biophys. Acta* 1843, 1433–1441. doi: 10.1016/j.bbamcr.2013.10.013
- Schuessler, D. L., Cortes, T., Fivian-Hughes, A. S., Lougheed, K. E., Harvey, E., Buxton, R. S., et al. (2013). Induced ectopic expression of HigB toxin in *Mycobacterium tuberculosis* results in growth inhibition, reduced abundance of a subset of mRNAs and cleavage of tmRNA. *Mol. Microbiol.* 90, 195–207. doi: 10.1111/mmi.12358
- Shimizu, H., Nishiyama, K., and Tokuda, H. (1997). Expression of *gpsA* encoding biosynthetic sn-glycerol 3-phosphate dehydrogenase suppresses both the LB-phenotype of a *secB* null mutant and the cold-sensitive phenotype of a *secG* null mutant. *Mol. Microbiol.* 26, 1013–1021. doi: 10.1046/j.1365-2958.1997.6392003.x
- Singh, R., Kraft, C., Jaiswal, R., Sejwal, K., Kasaragod, V. B., Kuper, J., et al. (2014). Cryo-electron microscopic structure of SecA protein bound to the 70S ribosome. *J. Biol. Chem.* 289, 7190–7199.
- Snyder, W. B., and Silhavy, T. J. (1992). Enhanced export of beta-galactosidase fusion proteins in *prlF* mutants is Lon dependent. *J. Bacteriol.* 174, 5661–5668.
- Sullivan, J. T., Young, E. E., Mccann, J. R., and Braunstein, M. (2012). The *Mycobacterium tuberculosis* SecA2 system subverts phagosome maturation to promote growth in macrophages. *Infect. Immun.* 80, 996–1006. doi: 10.1128/IAI.05987-11
- Teter, S. A., Houry, W. A., Ang, D., Tradler, T., Rockabrand, D., Fischer, G., et al. (1999). Polypeptide flux through bacterial Hsp70: DnaK cooperates with trigger factor in chaperoning nascent chains. *Cell* 97, 755–765. doi: 10.1016/S0092-8674(00)80787-4
- Trun, N. J., and Silhavy, T. J. (1989). *PrlC*, a suppressor of signal sequence mutations in *Escherichia coli*, can direct the insertion of the signal sequence into the membrane. *J. Mol. Biol.* 205, 665–676. doi: 10.1016/0022-2836(89)90312-4
- Trun, N. J., Stader, J., Lupas, A., Kumamoto, C., and Silhavy, T. J. (1988). Two cellular components, *PrlA* and SecB, that recognize different sequence determinants are required for efficient protein export. *J. Bacteriol.* 170, 5928–5930.
- Ullers, R. S., Ang, D., Schwager, F., Georgopoulos, C., and Genevau, P. (2007). Trigger factor can antagonize both SecB and DnaK/DnaJ chaperone functions in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* 104, 3101–3106. doi: 10.1073/pnas.0608232104
- Ullers, R. S., Luirink, J., Harms, N., Schwager, F., Georgopoulos, C., and Genevau, P. (2004). SecB is a bona fide generalized chaperone in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* 101, 7583–7588. doi: 10.1073/pnas.0402398101
- Valent, Q. A., Kendall, D. A., High, S., Kusters, R., Oudega, B., and Luirink, J. (1995). Early events in preprotein recognition in *E. coli*: interaction of SRP and trigger factor with nascent polypeptides. *EMBO J.* 14, 5494–5505.
- van der Sluis, E. O., and Driessen, A. J. (2006). Stepwise evolution of the Sec machinery in Proteobacteria. *Trends Microbiol.* 14, 105–108. doi: 10.1016/j.tim.2006.01.009
- Watanabe, T., Hayashi, S., and Wu, H. C. (1988). Synthesis and export of the outer membrane lipoprotein in *Escherichia coli* mutants defective in generalized protein export. *J. Bacteriol.* 170, 4001–4007.
- Wei, S. Q., and Stader, J. (1994). A new suppressor of a *lamB* signal sequence mutation, *prlZ1*, maps to 69 minutes on the *Escherichia coli* chromosome. *J. Bacteriol.* 176, 5704–5710.
- Wild, J., Altman, E., Yura, T., and Gross, C. A. (1992). DnaK and DnaJ heat shock proteins participate in protein export in *Escherichia coli*. *Genes Dev.* 6, 1165–1172. doi: 10.1101/gad.6.7.1165
- Wild, J., Walter, W. A., Gross, C. A., and Altman, E. (1993). Accumulation of secretory protein precursors in *Escherichia coli* induces the heat shock response. *J. Bacteriol.* 175, 3992–3997.
- Wolff, N., Sapriel, G., Bodenreider, C., Chaffotte, A., and Delepelaire, P. (2003). Antifolding activity of the SecB chaperone is essential for secretion of HasA, a quickly folding ABC pathway substrate. *J. Biol. Chem.* 278, 38247–38253. doi: 10.1074/jbc.M302322200
- Woodbury, R. L., Topping, T. B., Diamond, D. L., Suciu, D., Kumamoto, C. A., Hardy, S. J., et al. (2000). Complexes between protein export chaperone SecB and SecA – Evidence for separate sites on SecA providing binding energy and regulatory interactions. *J. Biol. Chem.* 275, 24191–24198. doi: 10.1074/jbc.M002885200
- Xu, Z., Knafels, J. D., and Yoshino, K. (2000). Crystal structure of the bacterial protein export chaperone SecB. *Nat. Struct. Biol.* 7, 1172–1177. doi: 10.1038/82040
- Yamaguchi, Y., and Inouye, M. (2011). Regulation of growth and death in *Escherichia coli* by toxin-antitoxin systems. *Nat. Rev. Microbiol.* 9, 779–790. doi: 10.1038/nrmicro2651
- Zhou, J., and Xu, Z. (2003). Structural determinants of SecB recognition by SecA in bacterial protein translocation. *Nat. Struct. Biol.* 10, 942–947. doi: 10.1038/nsb980
- Zuber, B., Chami, M., Houssin, C., Dubochet, J., Griffiths, G., and Daffe, M. (2008). Direct visualization of the outer membrane of mycobacteria and corynebacteria in their native state. *J. Bacteriol.* 190, 5672–5680. doi: 10.1128/JB.01919-07

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 25 September 2014; accepted: 17 November 2014; published online: 05 December 2014.

Citation: Sala A, Bordes P and Genevau P (2014) Multitasking SecB chaperones in bacteria. *Front. Microbiol.* 5:666. doi: 10.3389/fmicb.2014.00666

This article was submitted to *Microbial Physiology and Metabolism*, a section of the journal *Frontiers in Microbiology*.

Copyright © 2014 Sala, Bordes and Genevau. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



OPEN ACCESS

**Edited by:**

Salvador Ventura,  
Universitat Autònoma de Barcelona,  
Spain

**Reviewed by:**

Helen Zgurskaya,  
The University of Oklahoma, USA  
Blanca Barquera,  
Rensselaer Polytechnic Institute, USA  
Frank T. Robb,  
University of Maryland School of  
Medicine, USA

**\*Correspondence:**

Thomas Bentin  
bentin@sund.ku.dk

**† Present Address:**

Lise Goltermann,  
Department of Immunology and  
Microbiology, Costerton Biofilm  
Center, Copenhagen University,  
Copenhagen, Denmark;  
Menachem V. Sarusie,  
Department of Biological Sciences,  
National University of Singapore,  
Singapore, Singapore

‡ Joint first authors.

**Specialty section:**

This article was submitted to  
Microbial Physiology and Metabolism,  
a section of the journal  
Frontiers in Microbiology

**Received:** 28 June 2015

**Accepted:** 28 December 2015

**Published:** 26 January 2016

**Citation:**

Goltermann L, Sarusie MV and  
Bentin T (2016) Chaperonin  
GroEL/GroES Over-Expression  
Promotes Aminoglycoside Resistance  
and Reduces Drug Susceptibilities in  
*Escherichia coli* Following Exposure to  
Sublethal Aminoglycoside Doses.  
Front. Microbiol. 6:1572.  
doi: 10.3389/fmicb.2015.01572

# Chaperonin GroEL/GroES Over-Expression Promotes Aminoglycoside Resistance and Reduces Drug Susceptibilities in *Escherichia coli* Following Exposure to Sublethal Aminoglycoside Doses

Lise Goltermann <sup>†‡</sup>, Menachem V. Sarusie <sup>†‡</sup> and Thomas Bentin <sup>\*</sup>

Department of Cellular and Molecular Medicine, University of Copenhagen, Copenhagen, Denmark

Antibiotic resistance is an increasing challenge to modern healthcare. Aminoglycoside antibiotics cause translation corruption and protein misfolding and aggregation in *Escherichia coli*. We previously showed that chaperonin GroEL/GroES depletion and over-expression sensitize and promote short-term tolerance, respectively, to this drug class. Here, we show that chaperonin GroEL/GroES over-expression accelerates acquisition of streptomycin resistance and reduces susceptibility to several other antibiotics following sub-lethal streptomycin antibiotic exposure. Chaperonin buffering could provide a novel mechanism for emergence of antibiotic resistance.

**Keywords:** chaperonin, aminoglycosides, antibiotic resistance, protein misfolding, mRNA translation

## INTRODUCTION

Chaperones comprise an integral part of the protein folding machinery of bacterial cells and help maintain cellular homeostasis (Houry, 2001; Lin and Rye, 2006). Chaperones may also mask deleterious effects of mutations (Rutherford and Lindquist, 1998). Chaperonin GroEL/GroES promotes evolution of recombinant protein (Tokuriki and Tawfik, 2009) and a comparison of 446 bacterial genomes revealed that protein evolutionary rates in nature correlate positively with their dependency on GroEL/GroES (Bogumil and Dagan, 2010). Aminoglycoside (AG) antibiotics are known to promote translational misreading (Davies et al., 1964; Gorini and Kataja, 1964). Our lab (Goltermann et al., 2013) and others (Ling et al., 2012) have demonstrated that AG action promotes misfolding of newly synthesized protein. We have further demonstrated that GroEL/GroES over-expression countered nascent protein misfolding and promoted bacterial survival and growth in exponential cultures whereas chaperonin depletion sensitized cells to AG antibiotics (Goltermann et al., 2013). Consistent with chaperones playing key roles in the response to AG antibiotics, chaperone gene expression has been reported to be up-regulated in response to AG exposure across different bacterial species (Lin et al., 2005; Kohanski et al., 2008; Cardoso et al., 2010). Although not a mutagen in itself, the AG antibiotic streptomycin has been reported to induce mutations *via* translational misreading in streptomycin sensitive (*rpsL*<sup>+</sup>) wild type *E. coli* (Boe, 1992) and facilitate a mutator phenotype (Ren et al., 1999). Given that GroEL/GroES can promote short-term tolerance to AG antibiotics (Goltermann et al., 2013) and that streptomycin has indirect mutagenic properties (Boe, 1992; Ren et al., 1999), we speculated that chaperonin over-expression could buffer

deleterious protein misfolding resulting from translational misreading during AG exposure and hence promote adaptation and resistance to AG antibiotics.

## MATERIALS AND METHODS

### Strains and Plasmids

We used *E. coli* strain MG1655 (a gift from Stanley Brown, University of Copenhagen, strain #548 in our inventory) transformed with pACYC184 derived plasmids with protein expression driven by the arabinose inducible *Pbad* promoter. The following plasmids were used: pGroEL/GroES (p544 in our inventory) expressing GroEL/GroES, pΔGroEL/GroES (p730-c1 in our inventory) deleted for the *groS-groL* genes (Goltermann et al., 2013), and pGFP (p488-c1 in our inventory) expressing GFPuv (Crameri et al., 1996) containing the F64L, S65T (Heim et al., 1995; Cormack et al., 1996) mutations (Goltermann et al., 2013). pGroEL/GroES is commercially available from Takara Biosciences as pGro7 (Nishihara et al., 1998). The p730-c1 and p488-c1 plasmids were previously described (Goltermann et al., 2013). The construction of p730-c1 was erroneously described and we therefore describe this plasmid in detail: Plasmid p730-c1 (ΔGroEL/GroES) was generated by inverse PCR using pGroEL/GroES as a template and 5'-phosphorylated primers otb674 (5'-P-TTGAGAAAGTCCGTATCTGTTATGGGTG) and otb695 (5'-P-GCCCTGCACCTCGCAGAAATAAAC). The PCR product/template mix was *DpnI* treated to remove template. The PCR amplicon was circularized with DNA ligase and transformed into *E. coli*. The resulting construct was characterized by restriction mapping using *EcoRI* and *HindIII* yielding bands compatible with the expected fragment sizes (1947 and 1523 bp) and by sequencing across the *groS-groL* deletion using otb710 (CAAAAGCGTACAGTTCAGGCG).

### Susceptibility Assay and MIC Value Determinations

Over-night MG1655/plasmid cultures in LB containing 40 μg/ml chloramphenicol (LBC; for plasmid maintenance) and 0.02% L-arabinose (for GroEL/GroES induction) were diluted into fresh LBC containing arabinose and sub-inhibitory concentrations of selection antibiotic (12–14 μg/ml streptomycin, 15 μg/ml spectinomycin, or 18 μg/ml ampicillin (see Figure S1) followed by continued growth over-night. Cultures were diluted to OD<sub>595 nm</sub> = 0.003 in fresh medium, and the resulting cultures were grown in Erlenmeyer flasks at 37°C, 180 rpm ON. The cultures were passaged every 24 h over the course of 3 days. At each passage, aliquots of out-grown cultures were removed normalized to 1 OD<sub>595 nm</sub> per ml and 100 μl of undiluted culture as well as ten-fold dilutions were plated on LBC-agar containing inhibitory amounts of selection antibiotic (40 μg/ml streptomycin, 80 μg/ml spectinomycin, or 60 μg/ml ampicillin) to determine the number of resistant colony forming units (cfu), and on LBC-agar to determine the total cfu count. Drug tolerant cfu and total cfu were thus determined on plates without L-arabinose. Growth inhibitory concentrations of selection antibiotic were determined as the concentration that did not

allow visible growth following 24 h incubation at 37°C of an unexposed culture. Following over-night incubation, cfu were counted manually. Examples of colonies growing on inhibitory selection antibiotic (streptomycin) were re-streaked on fresh plates to determine whether the reduced drug susceptibility was stably inherited. Drug susceptibility to streptomycin and other antibiotics was investigated further by growth in LBC liquid culture using microtiter plates and using the indicated antibiotic concentrations (Table 1). MIC determinations were done by replica plating from a fresh over-night culture using a 48-pin tool (frogger) into 200 μl of LBC and the desired test antibiotic (Table 2). Plates were double taped to avoid evaporation and scored for ±growth after over-night incubation at 37°C at 900 rpm in a Heidolph microplate incubator. Sequencing of the chromosomal *rpsL* gene in streptomycin resistant isolates was done using primer otb722 following PCR amplification of the *rpsL* locus using forward primer otb721 (TCTGCGTAATGCCCCCATTAAGG) and reverse primer otb722 (AACTTCGGATCCGGCAGAATTTTAC).

## RESULTS AND DISCUSSION

In order to determine whether GroEL/GroES over-expression promotes resistance to AG antibiotics, we examined drug susceptibilities of MG1655 over-expressing the chaperonin complex following exposure to sub-lethal concentrations

**TABLE 1 | Susceptibility to other antibiotics of streptomycin selected isolates.**

	pGroEL/GroES	pΔGroEL/GroES
Streptomycin	100% (95/95)	100% (127/127)
Ampicillin	15% (14/95)	2% (3/127)
Tetracycline	0% (0/95)	0% (0/127)
Spectinomycin	19% (18/95)	3% (4/127)
Kanamycin	9% (9/95)	1% (1/127)
Amp/Spc/Str	5% (5/95)	0 (0/127)
Spc/Kan/Str	4% (4/95)	0 (0/127)

Sensitivity to other antibiotics of isolates growing on inhibitory streptomycin plates from pGroEL/GroES and pΔGroEL/GroES transformed MG1655 cultures after 24 h of sub-inhibitory streptomycin selection. Streptomycin, Str (100 μg/ml), Ampicillin, Amp (100 μg/ml), Tetracycline (10 μg/ml), Spectinomycin, Spc (50 μg/ml, Spc), Kanamycin, Kan (50 μg/ml, Kan).

**TABLE 2 | MIC values of streptomycin selected pGroEL/GroES isolates.**

	Control	Isolates
Streptomycin (μg/ml)	4	≥128
Kanamycin (μg/ml)	8	128
Spectinomycin (μg/ml)	16	64
Ampicillin (μg/ml)	16	64
Tetracycline (μg/ml)	2	2

"Isolates" and "Control" designate pGroEL/GroES MG1655 cultures that had or had not been exposed to streptomycin selection prior to MIC determination, respectively. Numbers are based on the experiment shown in Table S2.



of various “selection antibiotics”: (i) Streptomycin, an aminoglycoside causing ribosomal misreading and protein misfolding, (ii) spectinomycin, a bacteriostatic antibiotic that causes translational blockage without misreading or protein misfolding, and (iii) ampicillin, which targets the peptidoglycan layer and therefore is not expected to impact protein folding. The antibiotic concentrations used were based on a titration, which showed similar growth phenotypes among the differently used antibiotics (Figure S1). We used the *E. coli* strain MG1655 transformed with one of three arabinose inducible plasmids. One transformant contained the plasmid pGroEL/GroES (Nishihara et al., 1998) expressing chaperonin GroEL/GroES. As controls, we used MG1655 transformed with pΔGroEL/GroES in which the entire *groS-groL* coding region was deleted (empty control) and MG1655 transformed with pGFP, a plasmid that expresses GFP (over-expression control). Cultures were passaged with arabinose inducer for 3 days in total (Figure S2) and drug susceptibility was then determined without chaperonin induction.

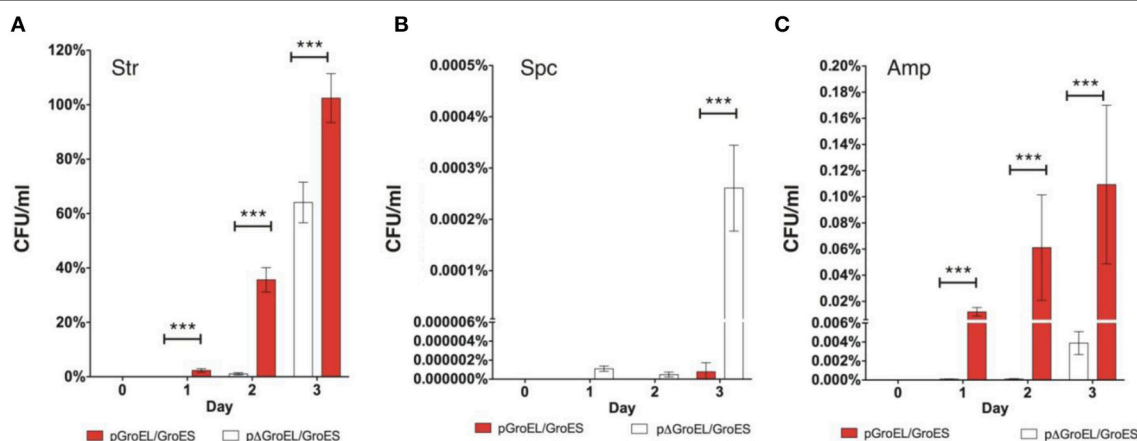
The pGroEL/GroES transformed strain showed a rapid reduction of drug susceptibility following streptomycin selection as determined by the percentage of cfu growing on plates containing an inhibitory streptomycin concentration (Figure 1A). In fact, the pGroEL/GroES transformed culture was virtually purged for streptomycin sensitive bacteria in only 3 days. MG1655 transformed with pΔGroEL/GroES also showed reduced drug susceptibility following growth in sub-lethal streptomycin but this was much less pronounced requiring extended exposure (Figure 1A). This tendency also holds true in absolute numbers (Figure S3B), even though the pGroEL/GroES transformed strain grew to a lower cell density on day 1 of selection (see legend to Figure S2). Similar to the pΔGroEL/GroES transformed strain, streptomycin tolerance

in MG1655/pGFP also developed slowly (Figure S3), indicating that protein over-expression in itself did not enhance antibiotic adaptation. We conclude that acceleration of adaptation toward streptomycin was a result of chaperonin action.

Colonies derived following sub-lethal streptomycin selection and growth on inhibitory streptomycin plates, were incubated in liquid culture containing 100 μg/ml streptomycin and all showed growth (Table 1). This suggests that these colonies were not only temporarily tolerant but heritably resistant toward streptomycin.

Among the 95 chaperonin over-expressing isolates, 17 randomly selected colonies from the first day of selection were sequenced for changes in the *rpsL* gene encoding ribosomal protein S12. Mutations in the S12 protein that confer resistance to streptomycin are known. Two independent point mutations giving rise to an arginine to serine mutation at codon 85 (R85S) and a lysine to arginine mutation at codon 87 (K87R) were found (Table S1). A streptomycin sensitive control, which had not been exposed to streptomycin selection, was sequenced and revealed the wild type *rpsL* gene sequence (Table S1). The K87R mutation confers resistance to streptomycin (Timms et al., 1992). We have not found a prior description of the specific R85S mutation. These observations show that GroEL/GroES over-expression can enable accelerated resistance acquisition during sub-lethal streptomycin selection and that this resistance is genetically based.

A number of resistant isolates identified on day 1 after streptomycin or ampicillin selection were tested for susceptibility to other antibiotics at or above the MIC value of the parent strain (Table 1). Interestingly, several pGroEL/GroES isolates resistant to streptomycin were found to show reduced susceptibility to one (ampicillin, spectinomycin, or kanamycin) or two (spectinomycin and kanamycin or spectinomycin and



**FIGURE 1 | Chaperonin dependent antibiotic resistance development following sub-inhibitory selection.** Percentage of colonies growing on inhibitory concentrations of the indicated antibiotics used for selection with or without over-expression of GroEL/GroES (pGroEL/GroES), or an empty control plasmid (pΔGroEL/GroES). The number of cfu growing on inhibitory antibiotic plates was normalized to total cfu count (Each data point consists of 2-4 experimental repeats each composed of 2-4 technical repeats. Error bars show SD). Day 1: pGroEL/GroES vs. pΔGroEL/GroES  $p < 0.0001$  (\*\*\*), day 2: pGroEL/GroES vs. pΔGroEL/GroES  $p < 0.0001$  (\*\*\*), day 3: pGroEL/GroES vs. pΔGroEL/GroES  $p < 0.0001$  (\*\*\*). Note the split y-axis used to enable visualization of large differences within the same graph. (A) Str, streptomycin; (B) Spc, spectinomycin; (C) Amp, ampicillin.

ampicillin) additional antibiotics. Susceptibility to tetracycline was unaltered.

To explore drug susceptibilities further, we repeated the selection experiment in sub-lethal streptomycin concentrations and picked fresh pGroEL/GroES isolates growing on inhibitory streptomycin plates. These isolates were grown ON in a 96 well format. We then determined MIC values by replica plating and growth ON in media (Table 2; Table S2). Compared to the control (included in duplicate), which had not been subjected to selection, isolates were observed to show reduced drug susceptibilities to streptomycin (as expected), to kanamycin, and to a lesser extent to spectinomycin and ampicillin. As above, tetracycline sensitivity was unaltered. MIC values are known to be sensitive to the inoculum density and since these isolates were inoculated using a pin tool (Table S2) whereas the data presented in Table 1 was obtained following inoculation of cfu, we did not expect identical numbers. Nevertheless, we observe a clear trend that isolates over-expressing GroEL/GroES during growth with sub-lethal streptomycin exposure rapidly acquired reduced susceptibility to the selection drug and also to additional drugs.

Selection with spectinomycin resulted in a very low fraction of colonies showing reduced susceptibility to the drug even after 3 days of selection, regardless of chaperonin status (Figure 1B). Cultures grown without antibiotic selection did not yield any colonies under any of the tested conditions indicating that spontaneous formation of resistant isolates w/o antibiotic selection was under our detection limit (estimated to be  $\sim 1$  cfu in  $10^8$  cells).

Selection with ampicillin produced only few colonies showing reduced ampicillin susceptibility (Figure 1C), and only 1–4% of these isolates showed reduced susceptibility to an additional antibiotic (Table S3). Unexpectedly, the pGroEL/GroES transformed strain showed a small but significant reduction of susceptibility to ampicillin as compared to MG1655 harboring p $\Delta$ GroEL/GroES (Figure 1C) even though ampicillin action targets the peptidoglycan layer and does not involve protein misfolding. We do not know the mechanism(s) connecting chaperonin over-expression to reduced ampicillin susceptibility. One speculative mechanism could involve GroEL/GroES dependent accelerated protein evolution, as described for GroES/GroEL substrates *in vitro* (Tokuriki and Tawfik, 2009) and *in vivo* (Bogumil and Dagan, 2010). To this end, GroEL/GroES is a known cytosolic chaperone, but it nevertheless does interact with a few membrane proteins (Kerner et al., 2005) and could chaperone a wider collection of proteins than those identified by interaction analyses (Chapman et al., 2006).

We repeated the selection experiment using two other AG antibiotics: Kanamycin and gentamicin (Figure S4). Resistance toward these antibiotics occurs infrequently as compared with streptomycin. Nevertheless, like in the case of streptomycin, GroEL/GroES over-expressing bacteria grown under sub-inhibitory kanamycin (14  $\mu$ g/ml) or gentamicin

(1.5  $\mu$ g/ml) concentrations showed more colonies on inhibitory plates as compared with the  $\Delta$ GroEL/GroES strain following plating on plates containing 60  $\mu$ g/ml kanamycin or 10  $\mu$ g/ml gentamicin, respectively. Compared to the streptomycin selection experiment, colonies from kanamycin or gentamicin selection were smaller indicating slower growth. The observation that more cfu grew following selection with GroEL/GroES over-expression, however, suggests that chaperonin over-expression also promotes reduced drug susceptibilities to these members of the AG class of antibiotics.

The present results reveal that chaperonin action can accelerate resistance development to streptomycin following exposure to sub-lethal streptomycin concentrations and simultaneously reduce drug susceptibilities to other antibiotics to which the bacteria had not previously been exposed. These observations are compatible with chaperonin over-expression conferring an increased fitness due to reduced protein misfolding during exposure to AG. In turn, this enables selection for reduced drug susceptibility, which is accelerated due to the mutator phenotype resulting from streptomycin exposure (Ren et al., 1999). Such selected strains would also have acquired numerous additional mutations. Mutations reducing influx or increasing efflux (Pagès et al., 2008; Nikaido and Pagès, 2012) could explain reduced susceptibility to multiple antibiotics.

Together, the results suggest that chaperonins could enable formation of complex antibiotic susceptibility phenotypes from a single AG exposure regime. We propose a model, where GroEL/GroES over-expression reduce protein misfolding, increase fitness, expand the “mutational space,” and provide a window of opportunity for bacteria to acquire resistance or tolerance and hence evade drug mediated killing.

## AUTHOR CONTRIBUTIONS

LG and TB conceived the project. LG, MVS, and TB designed experiments. MVS and LG performed experiments. MVS, LG, and TB analyzed the data. LG and TB wrote the manuscript.

## ACKNOWLEDGMENTS

This work was supported by grant 2010-01-0801 from the Carlsberg Foundation (to TB) and by grant 1333-00113B from the Danish Medical Research Council (To LG). We thank Liam Good for comments on the manuscript.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fmicb.2015.01572>

## REFERENCES

- Boe, L. (1992). Translational errors as the cause of mutations in *Escherichia coli*. *Mol. Gen. Genet.* 231, 469–471. doi: 10.1007/BF00292717
- Bogumil, D., and Dagan, T. (2010). Chaperonin-dependent accelerated substitution rates in prokaryotes. *Genome Biol. Evol.* 2, 602–608. doi: 10.1093/gbe/evq044
- Cardoso, K., Gandra, R. F., Wisniewski, E. S., Osaku, C. A., Kadowaki, M. K., Felipach-Neto, V., et al. (2010). DnaK and GroEL are induced in response to antibiotic and heat shock in *Acinetobacter baumannii*. *J. Med. Microbiol.* 59, 1061–1068. doi: 10.1099/jmm.0.020339-0
- Chapman, E., Farr, G. W., Usaite, R., Furtak, K., Fenton, W. A., Chaudhuri, T. K., et al. (2006). Global aggregation of newly translated proteins in an *Escherichia coli* strain deficient of the chaperonin GroEL. *Proc. Natl. Acad. Sci. U.S.A.* 103, 15800–15805. doi: 10.1073/pnas.0607534103
- Cormack, B. P., Valdivia, R. H., and Falkow, S. (1996). FACS-optimized mutants of the green fluorescent protein (GFP). *Gene* 173, 33–38. doi: 10.1016/0378-1119(95)00685-0
- Cramer, A., Whitehorn, E. A., Tate, E., and Stemmer, W. P. (1996). Improved green fluorescent protein by molecular evolution using DNA shuffling. *Nat. Biotechnol.* 14, 315–319. doi: 10.1038/nbt0396-315
- Davies, J., Gilbert, W., and Gorini, L. (1964). Streptomycin, suppression, and the code. *Proc. Natl. Acad. Sci. U.S.A.* 51, 883–890. doi: 10.1073/pnas.51.5.883
- Goltermann, L., Good, L., and Bentin, T. (2013). Chaperonins fight aminoglycoside-induced protein misfolding and promote short-term tolerance in *Escherichia coli*. *J. Biol. Chem.* 288, 10483–10489. doi: 10.1074/jbc.M112.420380
- Gorini, L., and Kataja, E. (1964). Phenotypic repair by streptomycin of defective genotypes in *E. coli*. *Proc. Natl. Acad. Sci. U.S.A.* 51, 487–493. doi: 10.1073/pnas.51.3.487
- Heim, R., Cubitt, A. B., and Tsien, R. Y. (1995). Improved green fluorescence. *Nature* 373, 663–664. doi: 10.1038/373663b0
- Houry, W. A. (2001). Chaperone-assisted protein folding in the cell cytoplasm. *Curr. Protein Pept. Sci.* 2, 227–244. doi: 10.2174/1389203013381134
- Kerner, M. J., Naylor, D. J., Ishihama, Y., Maier, T., Chang, H. C., Stines, A. P., et al. (2005). Proteome-wide analysis of chaperonin-dependent protein folding in *Escherichia coli*. *Cell* 122, 209–220. doi: 10.1016/j.cell.2005.05.028
- Kohanski, M. A., Dwyer, D. J., Wierzbowski, J., Cottarel, G., and Collins, J. J. (2008). Mistranslation of membrane proteins and two-component system activation trigger antibiotic-mediated cell death. *Cell* 135, 679–690. doi: 10.1016/j.cell.2008.09.038
- Lin, J. T., Connelly, M. B., Amolo, C., Otani, S., and Yaver, D. S. (2005). Global transcriptional response of *Bacillus subtilis* to treatment with subinhibitory concentrations of antibiotics that inhibit protein synthesis. *Antimicrob. Agents Chemother.* 49, 1915–1926. doi: 10.1128/AAC.49.5.1915-1926.2005
- Lin, Z., and Rye, H. S. (2006). GroEL-mediated protein folding: making the impossible, possible. *Crit. Rev. Biochem. Mol. Biol.* 41, 211–239. doi: 10.1080/10409230600760382
- Ling, J., Cho, C., Guo, L. T., Aerni, H. R., Rinehart, J., and Söll, D. (2012). Protein aggregation caused by aminoglycoside action is prevented by a hydrogen peroxide scavenger. *Mol. Cell* 48, 713–722. doi: 10.1016/j.molcel.2012.10.001
- Nikaido, H., and Pagès, J. M. (2012). Broad-specificity efflux pumps and their role in multidrug resistance of Gram-negative bacteria. *FEMS Microbiol. Rev.* 36, 340–363. doi: 10.1111/j.1574-6976.2011.00290.x
- Nishihara, K., Kanemori, M., Kitagawa, M., Yanagi, H., and Yura, T. (1998). Chaperone coexpression plasmids: differential and synergistic roles of DnaK-DnaJ-GrpE and GroEL-GroES in assisting folding of an allergen of Japanese cedar pollen, Cryj2, in *Escherichia coli*. *Appl. Environ. Microbiol.* 64, 1694–1699.
- Pagès, J. M., James, C. E., and Winterhalter, M. (2008). The porin and the permeating antibiotic: a selective diffusion barrier in Gram-negative bacteria. *Nat. Rev. Microbiol.* 6, 893–903. doi: 10.1038/nrmicro1994
- Ren, L., Rahman, M. S., and Humayun, M. Z. (1999). *Escherichia coli* cells exposed to streptomycin display a mutator phenotype. *J. Bacteriol.* 181, 1043–1044.
- Rutherford, S. L., and Lindquist, S. (1998). Hsp90 as a capacitor for morphological evolution. *Nature* 396, 336–342. doi: 10.1038/24550
- Timms, A. R., Steingrimsdottir, H., Lehmann, A. R., and Bridges, B. A. (1992). Mutant sequences in the rpsL gene of *Escherichia coli* B/r: mechanistic implications for spontaneous and ultraviolet light mutagenesis. *Mol. Gen. Genet.* 232, 89–96. doi: 10.1007/BF00299141
- Tokuriki, N., and Tawfik, D. S. (2009). Chaperonin overexpression promotes genetic variation and enzyme evolution. *Nature* 459, 668–673. doi: 10.1038/nature08009

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Goltermann, Sarusie and Bentin. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Fold modulating function: bacterial toxins to functional amyloids

Adnan K. Syed<sup>1</sup> and Blaise R. Boles<sup>2</sup> \*

<sup>1</sup> Department of Molecular Cellular and Developmental Biology, University of Michigan, Ann Arbor, MI, USA

<sup>2</sup> Department of Microbiology, Roy J. and Lucille A. Carver College of Medicine, University of Iowa, Iowa City, IA, USA

## Edited by:

Salvador Ventura, Universitat Autònoma de Barcelona, Spain

## Reviewed by:

Paras Jain, Albert Einstein College of Medicine, USA

Claudio Soto, University of Texas, USA

## \*Correspondence:

Blaise R. Boles, Department of Microbiology, Roy J. and Lucille A. Carver College of Medicine, University of Iowa, 51 Newton Road, Iowa City, IA, USA  
e-mail: blaise-boles@uiowa.edu

Many bacteria produce cytolytic toxins that target host cells or other competing microbes. It is well known that environmental factors control toxin expression, however, recent work suggests that some bacteria manipulate the fold of these protein toxins to control their function. The  $\beta$ -sheet rich amyloid fold is a highly stable ordered aggregate that many toxins form in response to specific environmental conditions. When in the amyloid state, toxins become inert, losing the cytolytic activity they display in the soluble form. Emerging evidence suggest that some amyloids function as toxin storage systems until they are again needed, while other bacteria utilize amyloids as a structural matrix component of biofilms. This amyloid matrix component facilitates resistance to biofilm disruptive challenges. The bacterial amyloids discussed in this review reveal an elegant system where changes in protein fold and solubility dictate the function of proteins in response to the environment.

**Keywords:** functional amyloid, bacterial toxin, biofilm, bifunctional protein, aggregation

## AMYLOIDS

The secondary structure of the amyloid fold is one that is seen throughout life. Amyloids have long been studied because of their importance in human neurodegenerative diseases such as Alzheimer's and Huntington's diseases. There have been many reviews published on disease associated amyloids and they will not be discussed in this review (Chiti and Dobson, 2006; Eisenberg and Jucker, 2012). Amyloids were initially described in the 1850s as deposits in human tissues that stained with iodine, which was a characteristic of starch. A few years later though, it was shown that there were no carbohydrates in amyloid deposits but they consisted of proteins. Since then, amyloids have been found to be produced in many organisms and the biophysical and chemical properties of amyloids have been significantly investigated.

Amyloids are composed of fibrous oligomers of proteins that are characterized by a cross  $\beta$ -sheet structure running perpendicular to the fiber axis. The core of the amyloid fiber is made of protein backbones that form many hydrogen bonds between them leading to strong molecular forces. Because the strength of the amyloid is mainly independent of the side chains, proteins that can fold into amyloids do not have any sequence motifs, making it difficult to predict amyloid forming proteins. Even though they do not contain sequence homology, amyloids can be identified using biophysical properties including SDS-insolubility, protease resistance, and binding to the amyloid specific dyes Thioflavin T and Congo red (CR).

Amyloids have long been thought to be a result of protein misfolding, but over the past decade, this view has evolved to the understanding that some organisms utilize the amyloid fold for various functions aptly named Functional Amyloids (Table 1; Chapman et al., 2002). The most well studied functional amyloids

are those made by bacteria that help form microbial communities called biofilms. These biofilms contain bacterial cells as well as matrix containing carbohydrates and proteins that hold them together affixed to a surface. Increasingly, it is being found that bacteria utilize the strength of the amyloid fold to make the strong biofilm matrix that resists disruption from stressors. As will be discussed below, many bacteria have developed toxin systems that are able to attack niche competitors or the host that can be abrogated by sequestering these toxins as amyloids where some have a second function in biofilm stability.

## AMYLOIDS AS STRUCTURAL MOLECULES

### CURLI

Curli are the most well studied bacterial functional amyloid. Through a dedicated pathway, curli form amyloids on the surface of *Enterobacteria*, such as *Escherichia coli*, and *Salmonella*, that aid bacteria in attaching to surfaces as well as defending the population from stress (Saldaña et al., 2009; Goulter-Thorsen et al., 2011; Zhou et al., 2012). Curli are made through a highly controlled master regulator CsgD, which induces the transcription of other curli specific genes (*csg*) to produce these amyloids (Figure 1; Brombacher et al., 2003). The major functional subunit of curli, CsgA, is secreted from the cell in a soluble form, leaving the outer-membrane through the pore formed by a hexamer of CsgG (Figure 1; Chapman et al., 2002; Robinson et al., 2006; Epstein et al., 2009). The minor fiber subunit CsgB is linked with the membrane and facilitates the nucleation of CsgA into amyloid fibers (Figure 1; Bian and Normark, 1997; Hammer et al., 2007). Proper assembly, localization, and regulation of curli fibers are modulated by CsgC, CsgE, and CsgF (Gibson et al., 2007; Nenninger et al., 2009, 2011; Evans and Chapman, 2014). Not only are the curli genes under strict genetic regulation, but it



**Table 1 | Bacterial functional amyloids.**

Organism	Protein	Functions	Reference
<i>Escherichia coli</i> , other <i>Enterobacteriaceae</i>	CsgA (Curli)	Biofilm matrix protein, surface attachment	Chapman et al. (2002)
<i>Pseudomonads</i>	FapC	Biofilm matrix protein	Dueholm et al. (2010, 2013)
<i>Mycobacterium tuberculosis</i>	MTP	Binding to human proteins and biofilm formation	Alteri et al. (2007), Ramsugit et al. (2013)
<i>Streptococcus mutans</i>	Adhesin P1 (antigen I/II, PAc)	Role in dental carries	Oli et al. (2012)
<i>Streptomyces coelicolor</i>	Chaplin	Hydrophobic coat for aerial hyphae	de Jong et al. (2009), Claessen et al. (2003)
<i>Staphylococcus aureus</i> and other <i>Staphylococci</i>	Phenol soluble modulins (PSMs)	Soluble virulence factor, biofilm dispersal/amyloid stabilizes biofilms	Wang et al. (2007), Schwartz et al. (2012)
<i>Klebsiella pneumoniae</i>	Microcin E492	Bacteriocin that is inactivated by amyloidogenesis	Biéler et al. (2005)
<i>Bacillus subtilis</i>	TasA	Soluble toxin/amyloid necessary for biofilm architecture	Stover and Driks (1999c), Romero et al. (2010)
<i>Listeria monocytogenes</i>	Listeriolysin O (LLO)	Forms pores in phagolysosome/inactive by pH shift in cytosol	Bavdek et al. (2012)
<i>Xanthomonas axonopodis</i>	HpaG (Harpins)	Triggers hypersensitive response in plants	Oh et al. (2007)

has been shown that cellular chaperones can modulate the fold of CsgA to prevent improper folding in the cells (Evans et al., 2011).

Curli fibers are important for *E. coli* surface colonization and biofilm formation (Chapman et al., 2002; Saldaña et al., 2009; Crémet et al., 2013; DePas et al., 2013; Giaouris et al., 2013). The expression of curli is a tightly regulated process in regards to the environment around the bacteria as well as within a biofilms community. Recently, it has been shown that there is spatial regulation within an *E. coli* rugose biofilms where curli producing cells are localized to the exterior of the biofilms, whereas cells on the interior of the community were not producing curli fibers (DePas et al., 2013; Serra et al., 2013). This bimodal growth allows for a protective shell of matrix-encased cells that contain a population of cells that ready to disperse and disseminate when conditions become favorable.

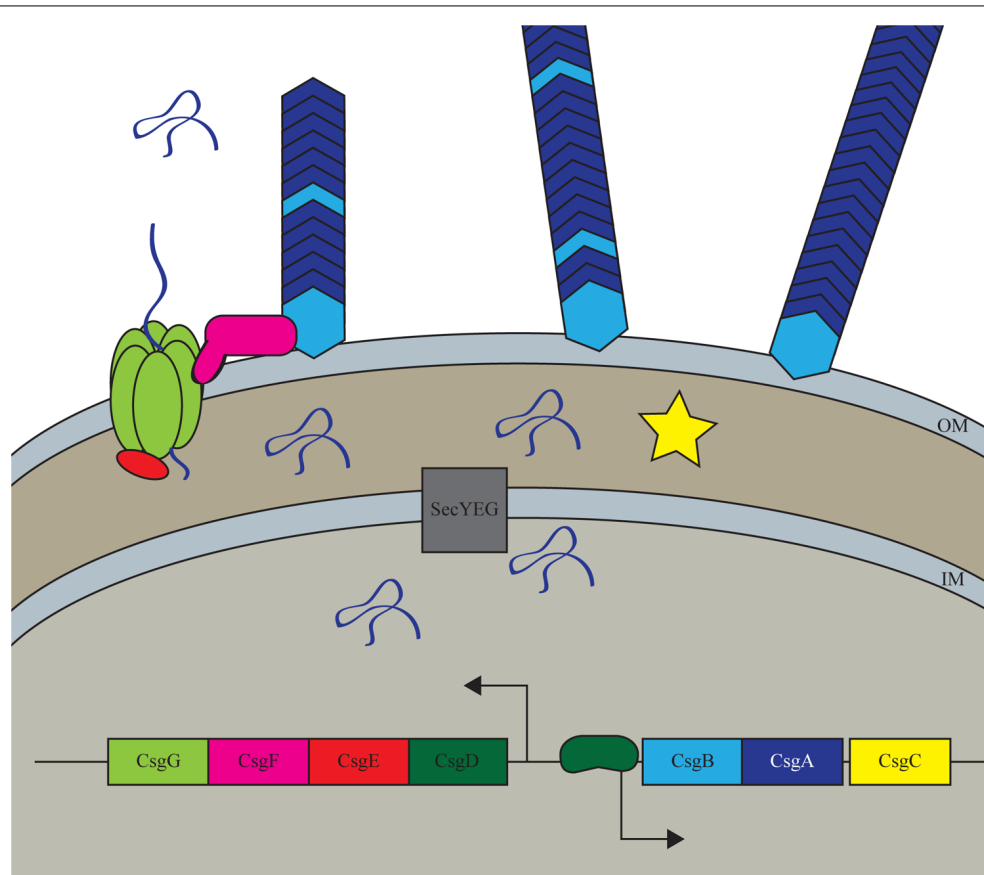
#### OTHER FUNCTIONAL AMYLOIDS PRODUCED BY BACTERIA

Emerging evidence suggest that amyloids likely play a structural role in some naturally occurring environmental biofilms. Recent work utilizing conformational antibodies that specifically bind to the amyloid fold, and the amyloid-specific dye thioflavin-T, provide evidence of amyloids being present in biofilm samples for fresh water lakes, drinking water, and activated sludge from a water treatment facility (Larsen et al., 2007). The bacteria present in these biofilms include representatives from Actinobacteria, Bacteroidetes, Chloroflexi, and Proteobacteria. Further studies revealed one member of this community, *Pseudomonas fluorescens*, was able to produce an amyloid found in the biofilm matrix (Dueholm et al., 2010). Proteomic analysis revealed the major subunit of the amyloid to consist of a protein named FapC (Dueholm et al., 2010). The genes necessary for formation of this amyloid were traced to the *fapA-F* operon, which is conserved in many

*Pseudomonas* species. FapC contains repeat motifs and conserved Asn/Gln consensus residues similar to curli and the prion and spider silk amyloid proteins (Dueholm et al., 2010). Further studies have demonstrated that other *Pseudomonads* also form Fap fibrils that result in biofilm formation (Dueholm et al., 2013). These finding suggest functional amyloids are likely abundant in naturally occurring biofilms consisting of diverse microbial members.

The pathogens *Mycobacterium tuberculosis* and *Streptococcus mutans* have also been found to produce functional amyloids. In the case of *M. tuberculosis*, thin, aggregative flexible pili, named MTP, were observed during human infection (Alteri et al., 2007). These pili possess biophysical and morphological characteristics of amyloids and bind to the human extracellular matrix component, laminin. Proteomic analysis suggests the structural subunit of MTP is a proteolytically processed version of a 10.5 kDa protein encoded by the open reading frame Rv3312A (*mtp*) in *M. tuberculosis* strain H37Rv (Alteri et al., 2007). In addition, serum from tuberculosis patients contained antibodies that specifically recognized MTP, suggesting a role for MTP during infection (Alteri et al., 2007). MTP was also found to be important in the formation of biofilms by *M. tuberculosis* (Ramsugit et al., 2013). *S. mutans* is a member of the oral microbiome and is linked to the disease dental caries because of it's ability to produce acid from the utilization of dietary sugars. Recent work suggests that the *S. mutans* adhesin P1 (antigen I/II, PAc) is an amyloid-forming protein (Oli et al., 2012). During biofilm growth *S. mutans* displayed amyloid fibers as evidenced by transmission electron microscopy, bound the amyloidophilic dyes CR and Thioflavin T (ThT), and possessed green birefringent properties of CR-stained protein aggregates when viewed under cross-polarized light (Oli et al., 2012). Importantly, human dental plaques contain microbial amyloids, suggesting a role for this protein fold in dental carries (Oli et al., 2012).





**FIGURE 1 | Curli biogenesis model.** The curli system in *Escherichia coli* is a highly controlled process that only expressed the curli amyloid under conditions that promote biofilm formation. The system is transcriptionally controlled by the master regulator CsgD which increases the transcription of the major and minor subunits CsgA and CsgB. All Csg proteins other than

CsgD are secreted through the Sec secretion pathway into the periplasm where CsgA, CsgB, and CsgF are then translocated outside of the cell through the CsgG pore complex. CsgE and CsgF aid in proper export and localization of the structural components while CsgC has a less well understood role in the periplasm.

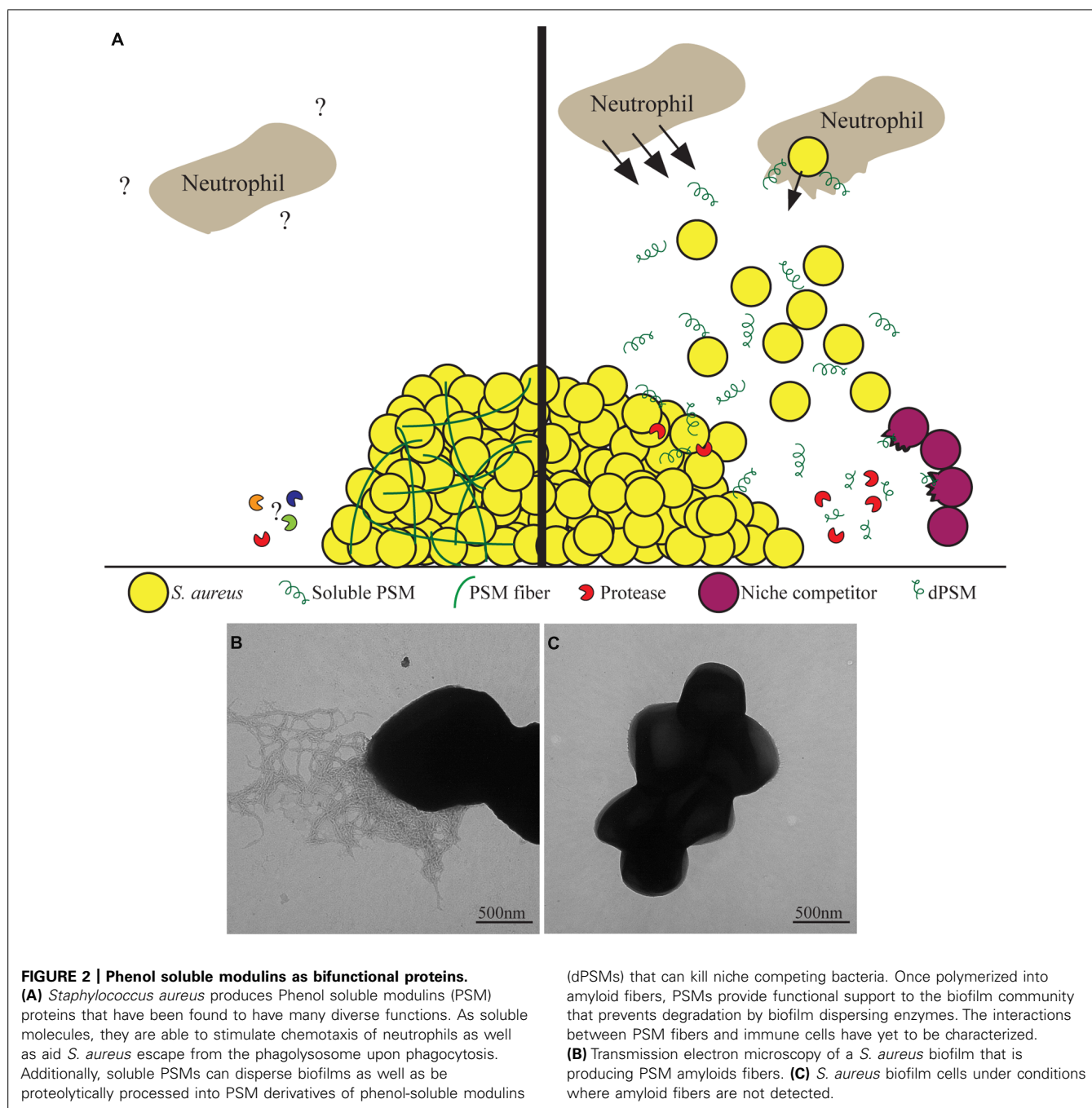
Chaplins are a class of hydrophobic proteins that spontaneously self-assemble into amyloid fibrils (Claessen et al., 2003). The spore-forming filamentous bacterium *S. coelicolor* uses chaplin amyloids to complete its lifecycle progression (Claessen et al., 2003). Under starvation conditions *S. coelicolor* produces aerial hyphae that extend upward out of the soil. Spores are produced in these hyphae and they are released once the soil surface has been breached. Vegetative *S. coelicolor* cell surfaces are hydrophilic, so to break the soil/air interface, the cells must first develop a hydrophobic coat. To this end, *S. coelicolor* secretes monomeric chaplin proteins (encoded by *chpA-H*; Claessen et al., 2003; Elliot et al., 2003). These hydrophobic proteins have been shown to form  $\beta$ -sheet rich amyloid fibers on contact with air, therefore chaplin amyloids are essential for *S. coelicolor* to complete its lifecycle from vegetative cells to spore containing hyphae.

## BIFUNCTIONAL PROTEINS

### PHENOL SOLUBLE MODULINS

Phenol soluble modulins (PSMs) are a family of proteins that are found in Staphylococci, most notably the significant human

pathogen *Staphylococcus aureus* and the human commensal *Staphylococcus epidermidis* (Mehlin et al., 1999; Wang et al., 2007). *S. aureus* has nine characterized PSM peptides that are all regulated by the accessory gene regulator (AGR) quorum sensing system (Janzon et al., 1989; Wang et al., 2007). There are four PSM $\alpha$ , two PSM $\beta$ , and  $\delta$ -toxin that are present in three separate regions of the chromosome. The newest member to this family is the N-terminal signal sequence of the AgrD molecule N-AgrD (Schwartz et al., 2014). This sequence is critical for localization of the propeptide to the membrane and once cleaved from the rest of the AgrD molecule has many structural and functional similarities to the other PSMs (Schwartz et al., 2014). In addition, some strains of *S. aureus* contain a pathogenicity island that harbors a ninth PSM called PSM-mec (Queck et al., 2009). The PSMs are secreted from the cells by a dedicated, essential secretion system called phenol-soluble modulin transporter (PMT; Chatterjee et al., 2013). These PSM peptides are amphipathic  $\alpha$ -helices, meaning that one face of the helix is hydrophobic while the other is hydrophilic (Wang et al., 2007). This shared property is thought to allow for them to form pores in the membranes of competing microbes and host cells to invade tissues and evade immune cells.



Phenol soluble modulins have been shown to be critical determinants of *S. aureus* to cause skin abscesses and wounds in murine models as well as aiding in the survival of *S. aureus* in murine bacteremia models (Wang et al., 2007). PSMs stimulate neutrophil chemotaxis through the human formyl peptide receptor 2 (FPR2), at nanomolar concentrations, independent of the formylation state of the peptides (Figure 2; Kretschmer et al., 2010). Once the neutrophils are in close proximity, PSMs are able to infiltrate cells and cause cell death (Figure 2; Wang et al., 2007). Recently though, the field has shifted toward the hypothesis that in the host, PSMs may be important in virulence once *S. aureus* is phagocytosed by

neutrophils (Surewaard et al., 2012). This hypothesis is supported by data demonstrating the serum lipoproteins are able to bind to and inactivate PSMs, meaning that they would be unable to function in the presence of serum in the host (Surewaard et al., 2012). Secondly, once phagocytosed by neutrophils, *S. aureus* cells highly upregulate the production of PSM peptides which aid in escaping from the phagolysosome (Figure 2; Surewaard et al., 2012).

Phenol soluble modulins are not only reported to be important for *S. aureus* pathogenesis and virulence against the host. PSMs have also been shown to be antimicrobial against potential

competitors. PSMs were first determined to have antimicrobial effects from *S. epidermidis* (Cogen et al., 2010). PSMs share structural similarity to mammalian antimicrobial peptides such as LL-37, thus it was tested to determine if *S. epidermidis* PSMs were able to kill mammalian pathogens (Cogen et al., 2010). Two PSMs from *S. epidermidis* were found to have antimicrobial effects against *S. aureus* as well as group A *Streptococcus* (GAS) and worked in conjunction with LL-37 to have synergy (Cogen et al., 2010). Focus then turned to determine if and how *S. aureus* PSMs may act against niche competing bacteria. It was found that full length PSM peptides possessed little antimicrobial activity but derivatives of PSMs (dPSMs), PSMs that have been proteolytically processed to be missing the first few amino acids, have strong antimicrobial properties against *S. pyrogenes*, *S. epidermidis*, and GAS (Figure 2; Joo et al., 2011; Gonzalez et al., 2012). Furthermore, when a colony of *S. aureus* is grown in close proximity to *S. epidermidis* or GAS, dPSMs are localized to the zone of inhibition of the competing bacteria (Gonzalez et al., 2012).

Along with their role as toxins, the biophysical properties of PSMs give them several unique properties in modulating *S. aureus* communities. First these are the ability of *S. aureus* PSM $\alpha$  and PSM $\beta$  to facilitate dissemination and spreading of a colony over soft agar plates (Tsompanidou et al., 2011). This suggests that some PSMs are able to act with surfactant-like properties lowering hydrophathy and allowing for *S. aureus* to spread (Tsompanidou et al., 2013). Additionally, PSMs have been shown to be important for the formation of biofilms. The PSMs are important biosurfactants that aid in the characteristic waves of dissemination of parts of the biofilms to colonize other areas (Figure 2; Periasamy et al., 2012).

Along with these properties attributed to the soluble PSM peptides, they are able to form amyloid fibers which stabilize *S. aureus* biofilms (Schwartz et al., 2012). This switch changes the soluble  $\alpha$ -helical peptides into  $\beta$ -sheet rich protein aggregates (Schwartz et al., 2012). This aggregation, like other amyloids, is through a self-templating mechanism that facilitates the transformation of other nearby proteins to adopt this amyloid fold. The PSMs formed amyloid fibers in biofilms that were grown in a non-standard rich media containing peptone, glucose, and NaCl. These biofilms were completely resistant to known biofilm dispersing enzymes Proteinase K, DNase, and Dispersin B, suggesting that the PSM amyloids are able to structurally stabilize the biofilm against enzymatic targeting of the previously characterized matrix components (Figure 2; Schwartz et al., 2012). Importantly, PSMs were demonstrated to have bifunctional abilities to either strengthen biofilms or disperse them dependent on their secondary structure (Figure 2). If monomeric PSMs were added to an established biofilm they exhibits surfactant like properties, dispersing the biofilms in a concentration dependent manner, whereas PSM fiber addition does not disperse biofilms (Schwartz et al., 2012).

Further studies are needed to investigate the flexibility of the PSM peptides to switch from soluble peptides to amyloid fibers and if this change is irreversible or only temporary. Interestingly, where PSM $\alpha$  and PSM $\beta$  peptides were shown to be essential for *S. aureus* colony spreading, it has been shown that colony spreading can be

inhibited by  $\delta$ -toxin (Omae et al., 2012). This may suggest a role in amyloid nucleation by  $\delta$ -toxin on the other PSMs that inhibit their ability to act as surfactants. It is tempting to speculate that these fibers may be reservoirs of toxins that *S. aureus* can utilize to both defend itself while also causing the population to disseminate and escape. It would also be interesting to see if the aggregation of these peptides into amyloids fibers is able to abrogate neutrophil chemotaxis thus acting as a way to hide from the immune system when forming biofilms in the host. Much more work is needed to fully understand how *S. aureus* and other Staphylococcal species utilize the fold of these PSM peptides to modulate their function.

## MICROCIN E492

Microcin E492 (MccE492) is part of a large family of bacteriocins that are antimicrobials secreted by bacteria to kill niche competitors. In general, bacteriocins are pore-forming proteins that kill competitor microbes by forming pores in their membranes, decreasing membrane potential (de Lorenzo and Pugsley, 1985). MccE492 is produced by *Klebsiella pneumoniae* that is able to target many *Enterobacteria* species such as *E. coli*, and *Salmonella* (de Lorenzo and Pugsley, 1985; Destoumieux-Garzón et al., 2003). MccE492 is found as both an unmodified peptide as well as a posttranslationally modified molecule with a catechol-type siderophore molecule (Thomas et al., 2004). This post translational modification allows for microcin to be recognized by siderophore catecholate receptors of target organisms that cause an uptake of the mature MccE492 molecule into the periplasm (Destoumieux-Garzón et al., 2003). The exact mechanism of cell death is unknown due to the fact that there needs to be much more MccE492 present for antimicrobial activity compared to the amount needed for membrane permeabilization (Destoumieux-Garzón et al., 2003).

Microcin E492 is unique to known microcins in that it is produced through exponential and stationary phase, whereas other microcins are only produced in stationary phase. Interestingly, MccE492 loses its antimicrobial activity when the population enters stationary phase even though the protein is still present at high levels (Corsini et al., 2002). This observation led to the discovery that in stationary phase, MccE492 aggregates to form amyloid fibers (Biéler et al., 2005). Aggregation into an amyloid abolishes the toxic effects of this peptide (Biéler et al., 2005). The aggregation of the peptide is modulated by many environmental factors as well as the state of posttranslational modification of the peptides.

MccE492 is produced in both the unmodified and modified forms in culture. The modified, antimicrobial MccE492 is the predominant form while the bacteria are growing in exponential phase of the culture (Marcoleta et al., 2013). When the bacteria begin to enter stationary phase, they decrease the production of the modified form making the unmodified MccE492 more prevalent in the population (Marcoleta et al., 2013). Unmodified MccE492 polymerizes faster than the modified peptide in forming amyloid fibers leading them to hypothesis that the bacteria may begin to produce more unmodified peptide in stationary phase is to begin to detoxify the environment by sequestering these peptides in inert amyloids (Marcoleta et al., 2013). Even

though the unmodified form is more efficient in polymerizing, the modified MccE492 is found in fibers with the unmodified form (Marcoleta et al., 2013). Moreover, polymerization of the modified MccE492 is accelerated in the presence of unmodified seeds, small amyloids that can nucleate amyloid elongation (Marcoleta et al., 2013).

Apart from the influence of posttranslational modification on MccE492 polymerization, the environment has a profound effect of the rate of polymerization and can even cause disassembly of MccE492 amyloids (Shah Nawaz and Soto, 2012). Basic pH, low ionic strength, and dilution of the fibers all led to fiber disassembly in two hours (Shah Nawaz and Soto, 2012). These disassembled fibers regained their antimicrobial activity that was absent while in amyloid fibers (Shah Nawaz and Soto, 2012). Even more striking was the ability of these disassembled toxins to reform amyloids when the environment was again changed (Shah Nawaz and Soto, 2012). These data demonstrate that in the case of MccE492, there are many factors that influence the toxicity of the peptides suggesting that the *Klebsiella pneumoniae* has evolved a mechanism of efficiently modulating the toxicity of MccE492. This data also leads to the exciting hypothesis that other amyloids may have conditions that lead to the disassembly of the amyloid fiber.

### TasA

TasA was originally described as a protein in *Bacillus subtilis* that was involved in sporulation but was later found to be expressed in stationary phase cells (Stover and Driks, 1999b,c). It has been demonstrated to have widespread antimicrobial activities against both plant and animal pathogens and commensal bacteria (Stover and Driks, 1999c). Recently though, TasA has been shown to produce amyloid fibers in *B. subtilis* biofilms that contributes heavily to the formation of complex community architecture similar to curli in *E. coli* (Romero et al., 2010). TasA is part of an operon that also encodes for TapA, the minor amyloid subunit and fiber anchor, and SipW, the signal peptidase that processes TapA and TasA (Stover and Driks, 1999a,c). It has been proposed that the antimicrobial effects of TasA may be due to the formation of toxic oligomers that many amyloidogenic proteins generate, but to our knowledge, this has not yet been investigated (Romero et al., 2010).

### LISTERIOLYSIN O

Listeriolysin O (LLO) of *Listeria monocytogenes* is a cholesterol-dependent cytolysin whose activity is dependent on pH. LLO formed pores in the membrane of phagolysosome allowing for *L. monocytogenes* to escape and carry out its replicative phase in the host cell cytosol (Portnoy et al., 1988; Cossart et al., 1989). The pH dependent activity of LLO was shown to be due to an irreversible structural change in the protein that lead to a decreased ability in hemolytic activity (Schuerch et al., 2005). Later, it was appreciated that this structural change was actually the result of LLO forming an amyloid (Bavdek et al., 2012). This pH change of the protein into an amyloid suggests that this proteins has evolved to form pores while the bacteria is trapped in a phagolysosome, but once it escapes into the cytosol, the higher pH inactivates the proteins by triggering amyloidogenesis (Bavdek et al., 2012). This amyloidogenesis may prevent the LLO toxin from lysing

the infected cell while *L. monocytogenes* replicates in the host cell. It is unknown if LLO amyloids demonstrate any activity intracellularly.

### HARPINS

Harpins are a class of proteins that are produced by gram-negative plant pathogens. They are characterized by being glycine rich, heat stable proteins that are secreted by a type III secretion system that can trigger a hypersensitive response (HR) in plants and are predicted to have  $\alpha$ -helical regions (Wei et al., 1992). Harpins trigger this HR when they are present in the intercellular space. The plant cells detect these proteins and respond using the early defense response through an apoptosis-like cell death. Pathogens lacking harpins, such as HpaG of *Xanthomonas axonopodis*, have decreased virulence (Kim et al., 2003, 2004). The mechanism by which harpins trigger HR in plants is not fully understood, but there is some data supporting harpins interacting with and disrupting cell membranes leading to depolarization (Lee et al., 2001).

In 2007, a harpin from *Xanthomonas*, HpaG, was characterized biochemically (Oh et al., 2007). This group found that under conditions that mimic plant apoplasts, HpaG formed amyloid fibers (Oh et al., 2007). A mutant of the proteins (L50P) that did not trigger HR in plants and was also unable to form amyloid fibers (Oh et al., 2007). From this, the authors suggest that the transition to amyloid fibers is an important step in triggering HR. Furthermore, harpins from other plant pathogens, *Ercinia amylovora*, and *Pseudomonas syringae* also formed amyloid fibers in plant apoplast-like conditions (Oh et al., 2007). More detailed studies on harpins are needed to determine exactly what properties can be attributed to soluble and amyloid forms of these proteins.

### AMYLOID INHIBITORS

An exciting field is emerging that is trying to speed up or slow down the formation of amyloid fibers using small molecules. The idea is that for many amyloidogenic proteins, the toxicity is due to the formation of intermediate oligomers that can disrupt membranes. By accelerating the polymerization of amyloid subunits, we may be able to bypass the toxic, degenerative affect of amyloids slowing the progression of neurodegenerative disease. Along similar lines, since many bacterial amyloids have been shown to aid in adherence to surfaces, if we use small molecules to interfere with the polymerization, bacteria may not be able to anchor themselves to the host in infections leading to faster clearance of pathogens. Conversely, in cases like the PSMs of *S. aureus*, since the soluble form is a toxin that is able to disrupt the immune response, by causing polymerization of the monomers could prevent their toxic function to cells allowing the immune system to clear the infection.

Curlicides and pilicides are ring-fused 2-pyridone molecules that are designed to look like peptide backbones. They have been designed to mimic a proteins backbone and interact with proteins that form amyloids by either disrupting their ability to polymerize, or nucleating and accelerating the amyloid maturation (Andersson et al., 2013). They have so far been characterized with their interaction with *E. coli* curli and type 1 pili (Andersson et al., 2013). These molecules are not only able to influence the



*in vitro* polymerization of CsgA, but they are able to affect the biogenesis of curli and pili in *E. coli* biofilms (Andersson et al., 2013). Additionally, uropathogenic *E. coli* treated with curlicides were attenuated in a murine model of a urinary tract infection (Cegelski et al., 2009). Other groups have taken the approach of designing small, non-natural peptides that can disrupt amyloid formation (Sievers et al., 2011). This has been shown to be successful *in vitro* with disease associated amyloids (Sievers et al., 2011). Recently, TasA has been proposed to be model amyloid for screening molecules with widespread activity against amyloids necessary for biofilm formation (Romero et al., 2013). Ongoing research is increasing the efficacy of these molecules as well as characterizing their ability to modulate the biogenesis of other amyloids.

### FUNCTIONAL AMYLOIDS IN OTHER KINGDOMS OF LIFE

Eukaryotic amyloids have been the subject of many recent reviews (Liebman and Chernoff, 2012; Watt et al., 2013; Wickner et al., 2013). Importantly, functional amyloids are not exclusive to bacteria. They have been found to be important in diverse eukaryotes from yeast to humans. It has recently become appreciated that humans have several functional amyloids. The protein Pmel17, a protein made in melanosomes for mammalian pigmentation, forms amyloids fibers (Fowler et al., 2005). The production of this amyloids is highly regulated with cells utilizing proteolytic processing to mature the proteins into a form that can form amyloids (Berson et al., 2001; Kummer et al., 2009). These processing steps prevent the proteins from forming amyloids in other cellular compartments preventing the toxicity that is associated with disease-associated amyloids. Additionally, it has been found that peptide hormones that are stored in the secretory granules of the mammalian pituitary form amyloids (Maji et al., 2009). It was found that *in vitro*, 31 of 42 studied hormones were able to form amyloids fibers (Maji et al., 2009). Mouse pituitary's also contained amyloid fibers using various approaches to identify amyloids (Maji et al., 2009). It is thought that these peptide hormones are stored in secretory granules as amyloids until they are needed. The cells can then secrete the granules and dilution of the amyloids causes disassociation and activation of the peptide hormones (Maji et al., 2009).

Many yeast form prions, self-propagating amyloids that are heritable elements. This method of non-mendelian inheritance was first proposed for [URE3] (Wickner, 1994). It was shown that the cytoplasmically inherited [URE3] element has the opposite effect on ureidosuccinate metabolism as the Ure2 protein and that cells cured of [URE3] were able to regain the element when Ure2 is overexpressed (Aigle and Lacroute, 1975; Wickner, 1994). This lead Wickner to the hypothesis that cytoplasmically inherited elements in yeast were prions (Wickner, 1994). Since then, several yeast proteins have been shown to form prions that, in most cases, abrogate the function of proteins.

### CONCLUDING REMARKS

It is becoming ever more appreciated that the amyloid fold is not just a product of protein misfolding, but it is a ubiquitously used protein fold throughout the kingdoms of life. Amyloids provide structure or control availability of proteins such as toxins

or signaling molecules. Even more exciting is the discovery that some of these proteins have been found to have different functions when they are in their soluble or insoluble forms. The production of functional amyloids is a highly controlled and regulated process that is controlled on several levels including transcriptional, translational, and posttranslational. The difficulty associated with breaking up these proteins is the property that has made them so valuable for many organisms. In the case of many bacteria, these amyloids provide a structural component that keep the community protected against mechanical and enzymatic disruption (Schwartz et al., 2012). In others, they are used as reservoirs of toxins that are ready to become active once the environment changes (Shah Nawaz and Soto, 2012). This field is only beginning to look at the effect that these functional amyloids play in the dynamic relationships between bacterial species as well as how these proteins may be involved in bacterial interactions with the host at a commensal or pathogenic level. It will be exciting to see where this field of bifunctional bacterial proteins goes as well as targeting these proteins to disperse biofilms or to sequester toxins.

### ACKNOWLEDGMENTS

The authors would like to thank the members of the laboratories of Blaise Boles and Matt Chapman at the University of Michigan for insightful conversation. This work was funded by a NIH grant NIAID AI081748 to Blaise R. Boles and American Heart Association Fellowship 13PRE13810001 to Adnan K. Syed.

### REFERENCES

- Aigle, M., and Lacroute, F. (1975). Genetical aspects of [URE3], a non-mitochondrial, cytoplasmically inherited mutation in yeast. *Mol. Gen. Genet.* 136, 327–335. doi: 10.1007/BF00341717
- Alteri, C. J., Xicohtencatl-Cortes, J., Hess, S., Caballero-Olin, G., Giron, J. A., and Friedman, R. L. (2007). *Mycobacterium tuberculosis* produces pili during human infection. *Proc. Natl. Acad. Sci. U.S.A.* 104, 5145–5150. doi: 10.1073/pnas.0602304104
- Andersson, E. K., Bengtsson, C., Evans, M. L., Chorell, E., Sellstedt, M., Lindgren, A. E. G., et al. (2013). Modulation of curli assembly and pellicle biofilm formation by chemical and protein chaperones. *Chem. Biol.* 20, 1245–1254. doi: 10.1016/j.chembiol.2013.07.017
- Bavdek, A., Kostanjšek, R., Antonini, V., Lakey, J. H., Dalla Serra, M., Gilbert, R. J. C., et al. (2012). pH dependence of listeriolysin O aggregation and pore-forming ability. *FEBS J.* 279, 126–141. doi: 10.1111/j.1742-4658.2011.08405.x
- Berson, J. F., Harper, D. C., Tenza, D., Raposo, G., and Marks, M. S. (2001). Pmel17 initiates premelanosome morphogenesis within multivesicular bodies. *Mol. Biol. Cell* 12, 3451–3464. doi: 10.1091/mbc.12.11.3451
- Bian, Z., and Normark, S. (1997). Nucleator function of CsgB for the assembly of adhesive surface organelles in *Escherichia coli*. *EMBO J.* 16, 5827–5836. doi: 10.1093/emboj/16.19.5827
- Biéler, S., Estrada, L., Lagos, R., Baeza, M., Castilla, J., and Soto, C. (2005). Amyloid formation modulates the biological activity of a bacterial protein. *J. Biol. Chem.* 280, 26880–26885. doi: 10.1074/jbc.M502031200
- Brombacher, E., Dorel, C., Zehnder, A. J., and Landini, P. (2003). The curli biosynthesis regulator CsgD co-ordinates the expression of both positive and negative determinants for biofilm formation in *Escherichia coli*. *Microbiology* 149, 2847–2857. doi: 10.1099/mic.0.26306-0
- Cegelski, L., Pinkner, J. S., Hammer, N. D., Cusumano, C. K., Hung, C. S., Chorell, E., et al. (2009). Small-molecule inhibitors target *Escherichia coli* amyloid biogenesis and biofilm formation. *Nat. Chem. Biol.* 5, 913–919. doi: 10.1038/nchembio.242
- Chapman, M. R., Robinson, L. S., Pinkner, J. S., Roth, R., Heuser, J., Hammar, M., et al. (2002). Role of *Escherichia coli* curli operons in directing amyloid fiber formation. *Science* 295, 851–855. doi: 10.1126/science.1067484



- Chatterjee, S. S., Joo, H.-S., Duong, A. C., Dieringer, T. D., Tan, V. Y., Song, Y., et al. (2013). Essential *Staphylococcus aureus* toxin export system. *Nat. Med.* 19, 364–367. doi: 10.1038/nm.3047
- Chiti, F., and Dobson, C. M. (2006). Protein misfolding, functional amyloid, and human disease. *Annu. Rev. Biochem.* 75, 333–366. doi: 10.1146/annurev.biochem.75.101304.123901
- Claessen, D., Rink, R., de Jong, W., Siebring, J., de Vreugd, P., Boersma, F. G. H., et al. (2003). A novel class of secreted hydrophobic proteins is involved in aerial hyphae formation in *Streptomyces coelicolor* by forming amyloid-like fibrils. *Genes Dev.* 17, 1714–1726. doi: 10.1101/gad.264303
- Cogen, A. L., Yamasaki, K., Sanchez, K. M., Dorschner, R. A., Lai, Y., MacLeod, D. T., et al. (2010). Selective antimicrobial action is provided by phenol-soluble modulins derived from *Staphylococcus epidermidis*, a normal resident of the skin. *J. Invest. Dermatol.* 130, 192–200. doi: 10.1038/jid.2009.243
- Corsini, G., Baeza, M., Monasterio, O., and Lagos, R. (2002). The expression of genes involved in microcin maturation regulates the production of active microcin E492. *Biochimie* 84, 539–544. doi: 10.1016/S0300-9084(02)01415-3
- Cossart, P., Vicente, M. F., Mengaud, J., Baquero, F., Perez-Diaz, J. C., and Berche, P. (1989). Listeriolysin O is essential for virulence of *Listeria monocytogenes*: direct evidence obtained by gene complementation. *Infect. Immun.* 57, 3629–3636.
- Crémet, L., Corvec, S., Batard, E., Auger, M., Lopez, I., Pagniez, F., et al. (2013). Comparison of three methods to study biofilm formation by clinical strains of *Escherichia coli*. *Diagn. Microbiol. Infect. Dis.* 75, 252–255. doi: 10.1016/j.diagmicrobio.2012.11.019
- de Jong, W., Wösten, H. A. B., Dijkhuizen, L., and Claessen, D. (2009). Attachment of *Streptomyces coelicolor* is mediated by amyloid fimbriae that are anchored to the cell surface via cellulose. *Mol. Microbiol.* 73, 1128–1140. doi: 10.1111/j.1365-2958.2009.06838.x
- de Lorenzo, V., and Pugsley, A. P. (1985). Microcin E492, a low-molecular-weight peptide antibiotic which causes depolarization of the *Escherichia coli* cytoplasmic membrane. *Antimicrob. Agents Chemother.* 27, 666–669. doi: 10.1128/AAC.27.4.666
- DePas, W. H., Hufnagel, D. A., Lee, J. S., Blanco, L. P., Bernstein, H. C., Fisher, S. T., et al. (2013). Iron induces bimodal population development by *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* 110, 2629–2634. doi: 10.1073/pnas.1218703110
- Destoumieux-Garzon, D., Thomas, X., Santamaria, M., Goulard, C., Barthélémy, M., Boscher, B., et al. (2003). Microcin E492 antibacterial activity: evidence for a TonB-dependent inner membrane permeabilization on *Escherichia coli*. *Mol. Microbiol.* 49, 1031–1041. doi: 10.1046/j.1365-2958.2003.03610.x
- Dueholm, M. S., Petersen, S. V., Sønderkaer, M., Larsen, P., Christiansen, G., Hein, K. L., et al. (2010). Functional amyloid in *Pseudomonas*. *Mol. Microbiol.* 77, 1009–1020. doi: 10.1111/j.1365-2958.2010.07269.x
- Dueholm, M. S., Søndergaard, M. T., Nilsson, M., Christiansen, G., Stensballe, A., Overgaard, M. T., et al. (2013). Expression of Pap amyloids in *Pseudomonas aeruginosa*, *P. fluorescens*, and *P. putida* results in aggregation and increased biofilm formation. *Microbiologyopen* 2, 365–382. doi: 10.1002/mbo3.81
- Eisenberg, D., and Jucker, M. (2012). The amyloid state of proteins in human diseases. *Cell* 148, 1188–1203. doi: 10.1016/j.cell.2012.02.022
- Elliot, M. A., Karoonuthaisiri, N., Huang, J., Bibb, M. J., Cohen, S. N., Kao, C. M., et al. (2003). The chaplins: a family of hydrophobic cell-surface proteins involved in aerial mycelium formation in *Streptomyces coelicolor*. *Genes Dev.* 17, 1727–1740. doi: 10.1101/gad.264403
- Epstein, E. A., Reizian, M. A., and Chapman, M. R. (2009). Spatial clustering of the curli secretion lipoprotein requires curli fiber assembly. *J. Bacteriol.* 191, 608–615. doi: 10.1128/JB.01244-08
- Evans, M. L., and Chapman, M. R. (2014). Curli biogenesis: order out of disorder. *Biochim. Biophys. Acta* 1844, 1551–1558. doi: 10.1016/j.bbamcr.2013.09.010
- Evans, M. L., Schmidt, J. C., Ilbert, M., Doyle, S. M., Quan, S., Bardwell, J. C. A., et al. (2011). *E. coli* chaperones DnaK, Hsp33 and Spy inhibit bacterial functional amyloid assembly. *Prion* 5, 323–334. doi: 10.4161/pri.18555
- Fowler, D. M., Koulou, A. V., Alory-Jost, C., Marks, M. S., Balch, W. E., and Kelly, J. W. (2005). Functional amyloid formation within mammalian tissue. *PLoS Biol.* 4:e6. doi: 10.1371/journal.pbio.0040006.sg001
- Giaouris, E., Samoilis, G., Chorianopoulos, N., Ercolini, D., and Nychas, G.-J. (2013). Differential protein expression patterns between planktonic and biofilm cells of *Salmonella enterica* serovar Enteritidis PT4 on stainless steel surface. *Int. J. Food Microbiol.* 162, 105–113. doi: 10.1016/j.ijfoodmicro.2012.12.023
- Gibson, D. L., White, A. P., Rajotte, C. M., and Kay, W. W. (2007). AgfC and AgfE facilitate extracellular thin aggregative fimbriae synthesis in *Salmonella enteritidis*. *Microbiology* 153, 1131–1140. doi: 10.1099/mic.0.2006/000935-0
- Gonzalez, D. J., Okumura, C. Y., Hollands, A., Kersten, R., Akong-Moore, K., Pence, M. A., et al. (2012). Novel phenol-soluble modulins derivatives in community-associated methicillin-resistant *Staphylococcus aureus* identified through imaging mass spectrometry. *J. Biol. Chem.* 287, 13889–13898. doi: 10.1074/jbc.M112.349860
- Goulter-Thorsen, R. M., Taran, E., Gentle, I. R., Gobius, K. S., and Dykes, G. A. (2011). CsgA production by *Escherichia coli* O157:H7 alters attachment to abiotic surfaces in some growth environments. *Appl. Environ. Microbiol.* 77, 7339–7344. doi: 10.1128/AEM.00277-11
- Hammer, N. D., Schmidt, J. C., and Chapman, M. R. (2007). The curli nucleator protein, CsgB, contains an amyloidogenic domain that directs CsgA polymerization. *Proc. Natl. Acad. Sci. U.S.A.* 104, 12494–12499. doi: 10.1073/pnas.0703310104
- Janzon, L., Löfdahl, S., and Arvidson, S. (1989). Identification and nucleotide sequence of the delta-lysin gene, hld, adjacent to the accessory gene regulator (agr) of *Staphylococcus aureus*. *Mol. Gen. Genet.* 219, 480–485. doi: 10.1007/BF00259623
- Joo, H.-S., Cheung, G. Y. C., and Otto, M. (2011). Antimicrobial activity of community-associated methicillin-resistant *Staphylococcus aureus* is caused by phenol-soluble modulins derivatives. *J. Biol. Chem.* 286, 8933–8940. doi: 10.1074/jbc.M111.221382
- Kim, J.-G., Jeon, E., Oh, J., Moon, J. S., and Hwang, I. (2004). Mutational analysis of *Xanthomonas* harpin HpaG identifies a key functional region that elicits the hypersensitive response in nonhost plants. *J. Bacteriol.* 186, 6239–6247. doi: 10.1128/JB.186.18.6239-6247.2004
- Kim, J.-G., Park, B. K., Yoo, C.-H., Jeon, E., Oh, J., and Hwang, I. (2003). Characterization of the *Xanthomonas axonopodis* pv. glycines Hrp pathogenicity island. *J. Bacteriol.* 185, 3155–3166. doi: 10.1128/JB.185.10.3155-3166.2003
- Kretschmer, D., Gleske, A.-K., Rautenberg, M., Wang, R., Köberle, M., Bohn, E., et al. (2010). Human formyl peptide receptor 2 senses highly pathogenic *Staphylococcus aureus*. *Cell Host Microbe* 7, 463–473. doi: 10.1016/j.chom.2010.05.012
- Kummer, M. P., Maruyama, H., Huelsmann, C., Baches, S., Weggen, S., and Koo, E. H. (2009). Formation of Pmel17 amyloid is regulated by juxtamembrane metalloproteinase cleavage, and the resulting C-terminal fragment is a substrate for  $\gamma$ -secretase. *J. Biol. Chem.* 284, 2296–2306. doi: 10.1074/jbc.M808904200
- Larsen, P., Nielsen, J. L., Dueholm, M. S., Wetzel, R., Otzen, D., and Nielsen, P. H. (2007). Amyloid adhesins are abundant in natural biofilms. *Environ. Microbiol.* 9, 3077–3090. doi: 10.1111/j.1462-2920.2007.01418.x
- Lee, J., Klessig, D. F., and Nürnberger, T. (2001). A harpin binding site in tobacco plasma membranes mediates activation of the pathogenesis-related gene HIN1 independent of extracellular calcium but dependent on mitogen-activated protein kinase activity. *Plant Cell* 13, 1079–1093. doi: 10.1105/tpc.13.5.1079
- Liebman, S. W., and Chernoff, Y. O. (2012). Prions in yeast. *Genetics* 191, 1041–1072. doi: 10.1534/genetics.111.137760
- Maji, S. K., Perrin, M. H., Sawaya, M. R., Jessberger, S., Vadodaria, K., Rissman, R. A., et al. (2009). Functional amyloids as natural storage of peptide hormones in pituitary secretory granules. *Science* 325, 328–332. doi: 10.1126/science.1173155
- Marcoleta, A., Marin, M., Mercado, G., Valpuesta, J. M., Monasterio, O., and Lagos, R. (2013). Microcin e492 amyloid formation is retarded by posttranslational modification. *J. Bacteriol.* 195, 3995–4004. doi: 10.1128/JB.00564-13
- Mehlin, C., Headley, C. M., and Klebanoff, S. J. (1999). An inflammatory polypeptide complex from *Staphylococcus epidermidis*: isolation and characterization. *J. Exp. Med.* 189, 907–918. doi: 10.1084/jem.189.6.907
- Neeninger, A. A., Robinson, L. S., Hammer, N. D., Epstein, E. A., Badtke, M. P., Hultgren, S. J., et al. (2011). CsgE is a curli secretion specificity factor that prevents amyloid fibre aggregation. *Mol. Microbiol.* 81, 486–499. doi: 10.1111/j.1365-2958.2011.07706.x
- Neeninger, A. A., Robinson, L. S., and Hultgren, S. J. (2009). Localized and efficient curli nucleation requires the chaperone-like amyloid assembly protein CsgF. *Proc. Natl. Acad. Sci. U.S.A.* 106, 900–905. doi: 10.1073/pnas.0812143106
- Oh, J., Kim, J.-G., Jeon, E., Yoo, C.-H., Moon, J. S., Rhee, S., et al. (2007). Amyloidogenesis of type III-dependent harpins from plant pathogenic bacteria. *J. Biol. Chem.* 282, 13601–13609. doi: 10.1074/jbc.M602576200
- Oli, M. W., Otoo, H. N., Crowley, P. J., Heim, K. P., Nascimento, M. M., Ramsook, C. B., et al. (2012). Functional amyloid formation by *Streptococcus mutans*. *Microbiology* 158, 2903–2916. doi: 10.1099/mic.0.060855-0

- Omae, Y., Sekimizu, K., and Kaito, C. (2012). Inhibition of colony-spreading activity of *Staphylococcus aureus* by secretion of  $\delta$ -hemolysin. *J. Biol. Chem.* 287, 15570–15579. doi: 10.1074/jbc.M112.357848
- Periasamy, S., Joo, H.-S., Duong, A. C., Bach, T.-H. L., Tan, V. Y., Chatterjee, S. S., et al. (2012). How *Staphylococcus aureus* biofilms develop their characteristic structure. *Proc. Natl. Acad. Sci. U.S.A.* 109, 1281–1286. doi: 10.1073/pnas.1115006109
- Portnoy, D. A., Jacks, P. S., and Hinrichs, D. J. (1988). Role of hemolysin for the intracellular growth of *Listeria monocytogenes*. *J. Exp. Med.* 167, 1459–1471. doi: 10.1084/jem.167.4.1459
- Queck, S. Y., Khan, B. A., Wang, R., Bach, T.-H. L., Kretschmer, D., Chen, L., et al. (2009). Mobile genetic element-encoded cytolysin connects virulence to methicillin resistance in MRSA. *PLoS Pathog.* 5:e1000533. doi: 10.1371/journal.ppat.1000533.g010
- Ramsugit, S., Guma, S., Pillay, B., Jain, P., Larsen, M. H., Danaviah, S., et al. (2013). Pili contribute to biofilm formation *in vitro* in *Mycobacterium tuberculosis*. *Antonie Van Leeuwenhoek* 104, 725–735. doi: 10.1007/s10482-013-9981-6
- Robinson, L. S., Ashman, E. M., Hultgren, S. J., and Chapman, M. R. (2006). Secretion of curli fibre subunits is mediated by the outer membrane-localized CsgG protein. *Mol. Microbiol.* 59, 870–881. doi: 10.1111/j.1365-2958.2005.04997.x
- Romero, D., Aguilar, C., Losick, R., and Kolter, R. (2010). Amyloid fibers provide structural integrity to *Bacillus subtilis* biofilms. *Proc. Natl. Acad. Sci. U.S.A.* 107, 2230–2234. doi: 10.1073/pnas.0910560107
- Romero, D., Sanabria-Valentín, E., Vlamakis, H., and Kolter, R. (2013). Biofilm inhibitors that target amyloid proteins. *Chem. Biol.* 20, 102–110. doi: 10.1016/j.chembiol.2012.10.021
- Saldaña, Z., Xicohtencatl-Cortes, J., Avelino, F., Phillips, A. D., Kaper, J. B., Puente, J. L., et al. (2009). Synergistic role of curli and cellulose in cell adherence and biofilm formation of attaching and effacing *Escherichia coli* and identification of Fis as a negative regulator of curli. *Environ. Microbiol.* 11, 992–1006. doi: 10.1111/j.1462-2920.2008.01824.x
- Schuerch, D. W., Wilson-Kubalek, E. M., and Tweten, R. K. (2005). Molecular basis of listeriolysin O pH dependence. *Proc. Natl. Acad. Sci. U.S.A.* 102, 12537–12542. doi: 10.1073/pnas.0500558102
- Schwartz, K., Sekedat, M. D., Syed, A. K., O'Hara, B., Payne, D. E., Lamb, A., et al. (2014). The AgrD N-terminal leader peptide of *Staphylococcus aureus* has cytolytic and amyloidogenic properties. *Infect. Immun.* doi: 10.1128/IAI.02111-14
- Schwartz, K., Syed, A. K., Stephenson, R. E., Rickard, A. H., and Boles, B. R. (2012). Functional amyloids composed of phenol soluble modulins stabilize *Staphylococcus aureus* biofilms. *PLoS Pathog.* 8:e1002744. doi: 10.1371/journal.ppat.1002744.t001
- Serra, D. O., Richter, A. M., Klauck, G., Mika, F., and Hengge, R. (2013). Microanatomy at cellular resolution and spatial order of physiological differentiation in a bacterial biofilm. *MBio* 4:e00103–13. doi: 10.1128/mBio.00103-13
- Shahnawaz, M., and Soto, C. (2012). Microcin amyloid fibrils A are reservoir of toxic oligomeric species. *J. Biol. Chem.* 287, 11665–11676. doi: 10.1074/jbc.M111.282533
- Sievers, S. A., Karanickolas, J., Chang, H. W., Zhao, A., Jiang, L., Zirafi, O., et al. (2011). Structure-based design of non-natural amino-acid inhibitors of amyloid fibril formation. *Nature* 475, 96–100. doi: 10.1038/nature10154
- Stover, A. G., and Driks, A. (1999a). Control of synthesis and secretion of the *Bacillus subtilis* protein YqxM. *J. Bacteriol.* 181, 7065–7069.
- Stover, A. G., and Driks, A. (1999b). Regulation of synthesis of the *Bacillus subtilis* transition-phase, spore-associated antibacterial protein TasA. *J. Bacteriol.* 181, 5476–5481.
- Stover, A. G., and Driks, A. (1999c). Secretion, localization, and antibacterial activity of TasA, a *Bacillus subtilis* spore-associated protein. *J. Bacteriol.* 181, 1664–1672.
- Surewaard, B. G., Nijland, R., Spaan, A. N., Kruijtz, J. A., de Haas, C. J., and van Strijp, J. A. (2012). Inactivation of staphylococcal phenol soluble modulins by serum lipoprotein particles. *PLoS Pathog.* 8:e1002606. doi: 10.1371/journal.ppat.1002606.g007
- Thomas, X., Destoumieux-Garzón, D., Peduzzi, J., Afonso, C., Blond, A., Birlirakis, N., et al. (2004). Siderophore peptide, a new type of post-translationally modified antibacterial peptide with potent activity. *J. Biol. Chem.* 279, 28233–28242. doi: 10.1074/jbc.M400228200
- Tsompanidou, E., Denham, E. L., Becher, D., de Jong, A., Buist, G., van Oosten, M., et al. (2013). Distinct roles of phenol-soluble modulins in spreading of *Staphylococcus aureus* on wet surfaces. *Appl. Environ. Microbiol.* 79, 886–895. doi: 10.1128/AEM.03157-12
- Tsompanidou, E., Sibbald, M. J. J. B., Chlebowicz, M. A., Dreisbach, A., Back, J. W., van Dijk, J. M., et al. (2011). Requirement of the agr locus for colony spreading of *Staphylococcus aureus*. *J. Bacteriol.* 193, 1267–1272. doi: 10.1128/JB.01276-10
- Wang, R., Braughton, K. R., Kretschmer, D., Bach, T.-H. L., Queck, S. Y., Li, M., et al. (2007). Identification of novel cytolytic peptides as key virulence determinants for community-associated MRSA. *Nat. Med.* 13, 1510–1514. doi: 10.1038/nm1656
- Watt, B., van Niel, G., Raposo, G., and Marks, M. S. (2013). PMEL: a pigment cell-specific model for functional amyloid formation. *Pigment Cell Melanoma Res.* 26, 300–315. doi: 10.1111/pcmr.12067
- Wei, Z. M., Laby, R. J., Zumoff, C. H., Bauer, D. W., He, S. Y., Collmer, A., et al. (1992). Harpin, elicitor of the hypersensitive response produced by the plant pathogen *Erwinia amylovora*. *Science* 257, 85–88. doi: 10.1126/science.1621099
- Wickner, R. B. (1994). [URE3] as an altered URE2 protein: evidence for a prion analog in *Saccharomyces cerevisiae*. *Science* 264, 566–569. doi: 10.1126/science.7909170
- Wickner, R. B., Edskes, H. K., Bateman, D. A., Kelly, A. C., Gorkovskiy, A., Dayani, Y., et al. (2013). Amyloids and yeast prion biology. *Biochemistry* 52, 1514–1527. doi: 10.1021/bi301686a
- Zhou, Y., Smith, D., Leong, B. J., Brännström, K., Almqvist, F., and Chapman, M. R. (2012). Promiscuous cross-seeding between bacterial amyloids promotes interspecies biofilms. *J. Biol. Chem.* 287, 35092–35103. doi: 10.1074/jbc.M112.383737

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 05 May 2014; accepted: 16 July 2014; published online: 01 August 2014.

Citation: Syed AK and Boles BR (2014) Fold modulating function: bacterial toxins to functional amyloids. *Front. Microbiol.* 5:401. doi: 10.3389/fmicb.2014.00401

This article was submitted to *Microbial Physiology and Metabolism*, a section of the journal *Frontiers in Microbiology*.

Copyright © 2014 Syed and Boles. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Functional bacterial amyloid increases *Pseudomonas* biofilm hydrophobicity and stiffness

Guanghong Zeng<sup>1</sup>, Brian S. Vad<sup>1</sup>, Morten S. Dueholm<sup>2</sup>, Gunna Christiansen<sup>3</sup>, Martin Nilsson<sup>4</sup>, Tim Tolker-Nielsen<sup>4</sup>, Per H. Nielsen<sup>2</sup>, Rikke L. Meyer<sup>1\*</sup> and Daniel E. Otzen<sup>1\*</sup>

<sup>1</sup> Interdisciplinary Nanoscience Centre, Aarhus University, Aarhus, Denmark, <sup>2</sup> Center for Microbial Communities, Aalborg University, Aalborg, Denmark, <sup>3</sup> Department of Biomedicine-Medical Microbiology and Immunology, Aarhus University, Aarhus, Denmark, <sup>4</sup> Department of Immunology and Microbiology, Costerton Biofilm Center, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark

## OPEN ACCESS

### Edited by:

Salvador Ventura,  
Universitat Autònoma  
de Barcelona, Spain

### Reviewed by:

Akos T. Kovacs,  
Friedrich Schiller University Jena,  
Germany  
Matthew Richard Chapman,  
University of Michigan, USA

### \*Correspondence:

Rikke L. Meyer and Daniel E. Otzen,  
Interdisciplinary Nanoscience Centre,  
Aarhus University, Gustav Wieds Vej  
14, DK – 8000 Aarhus, Denmark  
rikke.meyer@inano.au.dk;  
dao@inano.au.dk

### Specialty section:

This article was submitted to  
Microbial Physiology and Metabolism,  
a section of the journal  
Frontiers in Microbiology

**Received:** 29 July 2015

**Accepted:** 22 September 2015

**Published:** 07 October 2015

### Citation:

Zeng G, Vad BS, Dueholm MS,  
Christiansen G, Nilsson M,  
Tolker-Nielsen T, Nielsen PH,  
Meyer RL and Otzen DE (2015)  
Functional bacterial amyloid increases  
*Pseudomonas* biofilm hydrophobicity  
and stiffness. *Front. Microbiol.* 6:1099.  
doi: 10.3389/fmicb.2015.01099

The success of *Pseudomonas* species as opportunistic pathogens derives in great part from their ability to form stable biofilms that offer protection against chemical and mechanical attack. The extracellular matrix of biofilms contains numerous biomolecules, and it has recently been discovered that in *Pseudomonas* one of the components includes  $\beta$ -sheet rich amyloid fibrils (functional amyloid) produced by the *fap* operon. However, the role of the functional amyloid within the biofilm has not yet been investigated in detail. Here we investigate how the *fap*-based amyloid produced by *Pseudomonas* affects biofilm hydrophobicity and mechanical properties. Using atomic force microscopy imaging and force spectroscopy, we show that the amyloid renders individual cells more resistant to drying and alters their interactions with hydrophobic probes. Importantly, amyloid makes *Pseudomonas* more hydrophobic and increases biofilm stiffness 20-fold. Deletion of any one of the individual members of in the *fap* operon (except the putative chaperone FapA) abolishes this ability to increase biofilm stiffness and correlates with the loss of amyloid. We conclude that amyloid makes major contributions to biofilm mechanical robustness.

**Keywords:** amyloid, *Pseudomonas*, biofilm, AFM, force spectroscopy, Young's modulus, contact angle

## Introduction

Most bacteria are able to form biofilms or “microbial cities,” which renders the bacteria resistant to conventional antibiotics as well as mechanical and chemical attack. *Pseudomonas* strains in particular show a remarkable ability to form biofilm in a wide range of environments, allowing, e.g., *P. aeruginosa* to colonize lungs as an opportunistic pathogen in cystic fibrosis (Alhede et al., 2014). Biofilm protects bacterial communities by encasing them within a matrix of extracellular polymeric substances (EPS) (Rasamiravaka et al., 2015). This matrix stabilizes the establishment of microbial cells on a surface, and is likely to be a major factor facilitating most microbial infections (Costerton et al., 1999). EPS consists of different biomolecules, including polysaccharides, extracellular DNA

**Abbreviations:** AFM, Atomic Force Microscopy; EPS, extracellular polymeric substances; Fap, Functional amyloid in *Pseudomonas*; Dfap, *Pseudomonas* strain lacking the *fap* operon; pFap, *Pseudomonas* strain transformed with a plasmid overexpressing *fap*; WT, wildtype.

(eDNA) and proteins, leading to a highly hydrated and polar structural scaffold (Flemming and Wingender, 2010). There has been much focus on the role of polysaccharides and eDNA, which are both involved in early stages of biofilm formation (Whitchurch et al., 2002; Vasseur et al., 2005). Polysaccharides contribute to structural stability and protection in addition to helping bind and retain water and nutrients (Sutherland, 2001), while eDNA maintains coherent cell alignments (Gloag et al., 2013) and also serves as a source of nutrients during starvation (Mulcahy et al., 2010). In contrast, the role(s) of the protein components remain less well studied. Nevertheless, there is increasing evidence that proteins also play a major role in the build-up of biofilm.

A few years ago, we discovered an operon in a *Pseudomonas* sp. which we dubbed *fap* (Functional Amyloid in *Pseudomonas*) (Dueholm et al., 2010) and which upon overexpression in *E. coli* led to the formation of significant amounts of biofilm. This operon leads to the formation of amyloid fibrils, i.e., long and thin structures extending from the surface of the bacterial outer membrane which are arranged in the classical amyloid pattern with  $\beta$ -strands stacked orthogonal to the main fibril axis. The *fap* operon contains 6 genes (*fapA-F*) that encode the respective Fap proteins. Of these, FapC is the main fibril monomer, while FapB and FapE only make up a small percentage of purified fibrils (Dueholm et al., 2013a). FapB can take over as the main amyloid component *in vivo* if FapA is knocked out (Dueholm et al., 2013a), implying that FapA has a more regulatory role. Sequence analysis suggests that FapF is a  $\beta$ -barrel membrane protein and FapD a peptidase (Dueholm et al., 2013a). While such amyloid structures are typically associated with neurodegenerative diseases such as Alzheimer's and Parkinson's (Otzen, 2012), it is increasingly clear that they also have a vital and beneficial role to play in bacteria in general (Chapman et al., 2002; Dueholm et al., 2012). Amyloid is found widespread in the bacterial kingdom (Larsen et al., 2007, 2008) and we have very recently reported their existence in *Archaea* (Dueholm et al., 2015). Though not as widespread as the operon which in *E. coli* produces amyloid called curli, the *fap* operon (whose amyloid product we refer to as Fap amyloid) is found not just in *Pseudomonas* species (belonging to the Gammaproteobacteria) but also among Beta- and Deltaproteobacteria (Dueholm et al., 2013b). Bacterial amyloids serve a variety of functions, and so far the structural role appears to dominate, either in the form of self-associating fibrils or possibly even as direct components of the cell wall (Jordal et al., 2009; Dueholm et al., 2015). In addition to being able to associate laterally and entangle with each other, fibrils are also able to bind small metabolites such as quorum-sensing molecules which are instrumental in bacterial intercellular communication and biofilm build-up (Seviour et al., 2015). This makes them potential reservoirs to retain metabolites and nutrients within the biofilm, which is a very useful feature under conditions of e.g., turbulent flow in aquatic environments like rivers or wastewater treatment plants.

Hitherto it has not been addressed to what extent amyloid affects the mechanical properties of the biofilm. We would expect amyloid to play a significant role, given that amyloid can show

tensile strength comparable to steel on a weight-to-weight basis, making them high-performance biomaterials (Knowles et al., 2007). Biofilms themselves need to be sufficiently robust to withstand mechanical insults (typically shearing forces Tiirola et al., 2009) while being soft enough to accommodate bacterial movement and proliferation (Tolker-Nielsen et al., 2000). Biofilm is generally robust toward chemical treatment; in fact,  $\text{Fe}^{3+}$  can increase biofilm elasticity by a factor of  $\sim 500$  (from ca. 1 kPa to 0.5 MPa), probably through interactions with negatively charged groups in the EPS, though this is countered by the addition of the Fe-chelator citric acid (Lieleg et al., 2011). Here we use atomic force microscopy (AFM) to probe the effect of functional amyloid on the robustness and stiffness of the amyloid-producing *Pseudomonas* sp. UK4 (UK4) found within the *P. fluorescens* group (Dueholm et al., 2010). We show that amyloid makes the cells much more resistant toward drying out, increases the hydrophobicity of the biofilm surface and increases biofilm stiffness 20-fold. Deletion of individual genes in the *fap* operon highlights different contributions of individual components. We conclude that bacterial amyloid makes a major contribution to the build-up of biofilm's mechanical robustness.

## Materials and Methods

### Materials

Unless otherwise stated, all chemicals were from Sigma-Aldrich. The UK4 *fap* operon overexpression vector pMMB190Tc-UK4fap (pFap) was obtained from an earlier study (Dueholm et al., 2013a).

### Cloning of the UK4 $\Delta$ *fap* Mutant

To remove the entire *fap* operon (*fapA-fapF*) from UK4, a knockout fragment containing a gentamicin (Gm) resistance cassette was generated by PCR overlap extension essentially as described by Choi and Schweizer (2005) and Dueholm et al. (2013a). Briefly, primers (see **Table 1**) were used to amplify chromosomal regions upstream (UKfapAUpF-GWL/UKfapAUpR-GM) and downstream (UKfapFDnF-GM/UKfapFDnR-GWL) of *fapA-fapF*, and to amplify a gentamicin resistance cassette (Gm-F/Gm-R) from plasmid pPS856 (Hoang et al., 1998). The PCR fragments were fused together and amplified with primers GW-attB1 and GW-attB2, incorporating the *attB1* and *attB2* recombination sites at either end of the knockout cassette. Using the Gateway cloning system (Invitrogen), the resulting knockout fragment was first transferred by the BP reaction into pDONR221, generating entry plasmid pDONR221- $\Delta$ UK4fap, and subsequently transferred by the LR reaction into pEX18ApGW generating the knockout plasmid pEX18ApGW- $\Delta$ UK4fap. This plasmid was then transferred into UK4 by two-parental mating using the donor strain *E. coli* S17-1 with selection on *Pseudomonas* isolation agar plates supplemented with 30  $\mu\text{g}/\text{mL}$  gentamicin. Resolution of single crossover events was achieved by streaking on 5% sucrose plates via the counter-selectable *sacB* marker on the knockout plasmid. The mutant construction was confirmed by PCR analysis.

TABLE 1 | Bacteria, plasmids and primers used in this study.

Species and strain	Characteristics/Sequence	References
<b>Escherichia coli</b>		
Mach1	Used for routine subcloning	Invitrogen
S17-1	<i>pro thi recA hsdR (r- m+) Tp<sup>r</sup> Sm<sup>r</sup> Km<sup>r</sup> Δ RP4-2-Tc::Mu-Km::Tn7</i>	Simon et al., 1983
<b>Pseudomonas sp. UK4</b>		
WT	Wildtype	Dueholm et al., 2010
Δ <i>fap</i>	<i>fapA-fapF</i> inactivated in PAO1 by allelic displacement with a gentamicin-resistance cassette using pEX18ApGW-UK4 <i>fap</i> ; Gm <sup>r</sup>	This study
pVC	Δ <i>fap</i> with pMMB190Tc	This study
pFap	Δ <i>fap</i> with pMMB190Tc-UK4 <i>fap</i>	This study
pFapΔA	Δ <i>fap</i> with pMMB190Tc-UK4 <i>fap</i> Δ <i>fapA</i>	This study
pFapΔB	Δ <i>fap</i> with pMMB190Tc-UK4 <i>fap</i> Δ <i>fapB</i>	This study
pFapΔC	Δ <i>fap</i> with pMMB190Tc-UK4 <i>fap</i> Δ <i>fapC</i>	This study
pFapΔD	Δ <i>fap</i> with pMMB190Tc-UK4 <i>fap</i> Δ <i>fapD</i>	This study
pFapΔE	Δ <i>fap</i> with pMMB190Tc-UK4 <i>fap</i> Δ <i>fapE</i>	This study
pFapΔF	Δ <i>fap</i> with pMMB190Tc-UK4 <i>fap</i> Δ <i>fapF</i>	This study
<b>PLASMIDS</b>		
pMMB190Tc	pMMB190 Δ <i>bla</i> <i>tet</i> (Tc <sup>R</sup> ) P <sub>lacUV5</sub> controlled expression	Dueholm et al., 2013a
pMMB190Tc-UK4 <i>fap</i>	pMMB190Tc containing the complete UK4 <i>fap</i> operon	Dueholm et al., 2013a
pMMB190Tc-UK4 <i>fap</i> ΔA	pMMB190Tc-UK4 <i>fap</i> with a deletion of <i>fapA</i>	This study
pMMB190Tc-UK4 <i>fap</i> ΔB	pMMB190Tc-UK4 <i>fap</i> with a deletion of <i>fapB</i>	This study
pMMB190Tc-UK4 <i>fap</i> ΔC	pMMB190Tc-UK4 <i>fap</i> with a deletion of <i>fapC</i>	This study
pMMB190Tc-UK4 <i>fap</i> ΔD	pMMB190Tc-UK4 <i>fap</i> with a deletion of <i>fapD</i>	This study
pMMB190Tc-UK4 <i>fap</i> ΔE	pMMB190Tc-UK4 <i>fap</i> with a deletion of <i>fapE</i>	This study
pMMB190Tc-UK4 <i>fap</i> ΔF	pMMB190Tc-UK4 <i>fap</i> with a deletion of <i>fapF</i>	This study
pDONR221	Gateway donor vector; Km <sup>r</sup>	Invitrogen
pEX18ApGW	Gateway compatible gene replacement vector; Suc <sup>S</sup> , Amp <sup>r</sup>	Choi and Schweizer, 2005
pPS856	0.83 kb blunt-ended SacI fragment from pUCGM ligated into the EcoRV site of pPS854; Amp <sup>r</sup> , Gm <sup>r</sup>	Hoang et al., 1998
pDONR221-ΔUK4 <i>fap</i>	<i>fapA-fapF</i> entry clone; Km <sup>r</sup> , Gm <sup>r</sup>	This study
pEX18ApGW-ΔUK4 <i>fap</i>	UK4 <i>fap</i> knockout vector; Suc <sup>S</sup> , Amp <sup>r</sup> , Gm <sup>r</sup>	This study
<b>PRIMERS</b>		
UK <i>fap</i> AUpF-GWL	5'-TACAAAAAGCAGGCTAAGATCA GCCATGACCGCG	This study
UK <i>fap</i> AUpR-GM	5'- TCAGAGCGCTTTTGAAGCTAA TTCGTGAGAGTCGCCACGGCAG	This study

(Continued)

TABLE 1 | Continued

Species and strain	Characteristics/Sequence	References
UK <i>fap</i> FDnF-GM	5'-AGGAACCTCAAGATCCCCAATT CGCATTACCTACGCGCGCTATGA	This study
UK <i>fap</i> FDnR-GWL	5'- TACAAGAAAGCTGGGTAGGCT GAAGGTGAAGTCGGG	This study
Gm-F	5'- CGAATTAGCTTCAAAAGCGCT CTGA	Choi and Schweizer, 2005
Gm-R	5'- CGAATTGGGGATCTTGAAGTT CCT	Choi and Schweizer, 2005
GW-attB1	5'- GGGGACAAGTTTGTACAAAA AGCAGGCT	Choi and Schweizer, 2005
GW-attB2	5'- GGGGACCACCTTTGTACAAGAA AGCTGGGT	Choi and Schweizer, 2005
pMMB190-UpF	5'-ACACAGGAACTAGGCAC	This study
pMMB190-DnR	5'- AAATCTTCTCTCATCCGCC	This study
LmarkR-UK4 <i>fapA</i> -UpR	5'-AGGAACCTCAAGATCCCCAATT CGAGCCTTGTTGAGAGTCG	This study
LmarkR-UK4 <i>fapB</i> -UpR	5'-AGGAACCTCAAGATCCCCAATT CGAGCAGCCAAGAACGGTGA	This study
LmarkR-UK4 <i>fapC</i> -UpR	5'-AGGAACCTCAAGATCCCCAATT CGGGTTTGAGAGCCATTGTAG	This study
LmarkR-UK4 <i>fapD</i> -UpR	5'-AGGAACCTCAAGATCCCCAATT CGCTGGAAATAAAAGGCCT	This study
LmarkR-UK4 <i>fapE</i> -UpR	5'-AGGAACCTCAAGATCCCCAATT CGAACGGGAAGTGTTTCATCAT	This study
LmarkR-UK4 <i>fapF</i> -UpR	5'-AGGAACCTCAAGATCCCCAATT CGTTAAACAGACAACGGCACG	This study
LmarkF-UK4 <i>fapA</i> -DnF	5'-CGAATTGGGGATCTTGAAGTT CCTTGGCACCGTTGAATAACA	This study
LmarkF-UK4 <i>fapB</i> -DnF	5'-CGAATTGGGGATCTTGAAGTT CCTGAACCGCATGGCTAACAC	This study
LmarkF-UK4 <i>fapC</i> -DnF	5'-CGAATTGGGGATCTTGAAGTT CCTGCAAAGCAACACCCTCAC	This study
LmarkF-UK4 <i>fapD</i> -DnF	5'-CGAATTGGGGATCTTGAAGTT CCTCGCCAAGACCAAATGAAC	This study
LmarkF-UK4 <i>fapE</i> -DnF	5'-CGAATTGGGGATCTTGAAGTT CCTACTGCAATCTGGATCAACT	This study
LmarkF-UK4 <i>fapF</i> -DnF	5'-CGAATTGGGGATCTTGAAGTT CCTACCTGACAATTGTGCCCA	This study
EcoRI-UK4 <i>fapF</i>	5'-CACTGAATTCGCTCTTGCT CTATTCTCAC	Dueholm et al., 2013a
HindIII-UK4 <i>fapR</i>	5'-CACTAAGCTTGCGCAGCGGT TTTAGAAGT	Dueholm et al., 2013a

Construction of pFap Single Gene Knockout Derivatives

Single gene knockout derivatives of pFap were obtained by PCR overlap extension (Horton et al., 2013). Briefly, primers were used to amplify the regions spanning the upstream region of the *fap* operon to the 5'-terminal of the target gene (pMMB190-UpF/LmarkR-UK4*fapX*-UpR) and the 3'-terminal of the target gene to the downstream region of the *fap* operon (LmarkF-UK4*fapX*-DnF and pMMB190-DnR) in pMMB190Tc-UK4*fap*. The PCR fragments were fused together using an incorporated reverse complemented sequence and amplified with the EcoRI-UK4*fapF* and HindIII-UK4*fapR* primers (Dueholm et al., 2013a).



The final PCR fragments, which contained the *fap* operon with the individual genes disrupted, were subcloned into pCR4-TOPO vectors (Life technologies). Inserts for cloning were obtained by digestion of the subcloned vectors with FastDigest EcoRI/HindIII (Thermo Scientific) followed by gel purification using the UltraClean 15 DNA kit (MO-BIO, Carlsbad, CA). The expression vector pMMB190Tc was prepared for ligation using the same restriction enzyme combinations and purified as above. Ligations were done using T4 DNA ligase (Life technologies) according to the manufacturers' recommendation and the resulting vectors were validated by shotgun sequencing of the whole plasmids (Macrogen).

### Transformation of UK4Δ*fap*

Electrocompetent UK4Δ*fap* cells were prepared as previously described (Choi et al., 2006). Fifty microliter of electrocompetent cells were mixed with 2 μL plasmid (400 ng/μL) and 40 μL of the suspension was transferred to a room temperature 1 mm gap electroporation cuvette. A pulse of 1.40 kV was applied using a MicroPulser electroporator (BioRad). One milliliter room temperature LB medium was immediately applied and the samples transferred to a 15 mL tube. The transformation was incubated (28°C, 200 rpm, 2 h) and 100 μL were plated out on a LB agar plate containing 50 μg/mL tetracycline. Transformed colonies were picked after 3 days.

### Transmission Electron Microscopy

UK4Δ*fap* transformed pFap and single gene knockout derivatives were grown overnight (26°C, 200 rpm) in colony factor antigen (CFA) medium (10 g/L casein hydrolysate (Fluka), 1.5 g/L yeast extract (Sigma), 50 mg/L MgSO<sub>4</sub>, and 5 mg/mL MnCl<sub>2</sub>, adjusted to pH 7.4 with NaOH) containing 50 μg/mL tetracycline. These starter cultures were used to inoculate 10 mL of CFA with 50 μg/mL tetracycline and 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) to OD<sub>600</sub> 0.05. The samples were grown at 26°C with shaking at 200 rpm. Samples for TEM were collected at the stationary phase. Grids were washed in two drops of MilliQ water and stained with 1% phosphotungstic acid (pH 6.8) and blotted dry on filter paper. Samples were viewed with a JEOL 1010 transmission electron microscope.

### Congo Red Binding Assay and Thioflavin T (ThT) Staining

The same strains used for TEM analysis were grown on Congo Red indicator plates composed of CFA medium supplemented with 20 μg/mL Congo red and 10 μL/mL Coomassie brilliant blue G-250 solidified with 2% agar (AppliChem). One micrometer IPTG was used to induce production of Fap amyloid in overexpressing mutants. Growth was at 26°C. ThT was dissolved in 10 mM TRIS buffer at pH 8.0 at concentration of 7.5 μM. Bacterial colonies from agar plates were stained with ThT for 15 min before observation under epi-fluorescence microscope (Axiovert 200 M, Zeiss, Germany) with Zeiss filterset 06 (excitation 431–441, emission 470 LP).

### Growth of UK4 for AFM Studies

UK4 and derivative were cultured in CFA medium (26°C, 180 rpm) or on CFA agar. Forty microgram per milliliter

tetracycline was used to select for mutants carrying the pMMB190Tc plasmids. For this assay, leaky expression of cloned *fap* operons was sufficient and no IPTG was used (Dueholm et al., 2013a).

### AFM Imaging

Bacteria were grown in CFA medium (40 h, 26°C, 180 rpm), harvested and washed three times by centrifugation (5000 rpm, 3 min) followed by resuspension in MilliQ water. This was sufficient to detach individual cells for AFM imaging. A 5 μL droplet of bacterial suspension was placed on freshly cleaved mica and dried in air for 15–30 min and immediately imaged by AFM in air. In this way, morphological changes due to drying over time were minimized. For imaging of WT cell aggregates, extra precautions were taken: A cell aggregate was identified and tracked under optical microscope during the drying process and then imaged afterwards. This is necessary because single cells tend to be forced to come close and form fake aggregates after drying out.

A NanoWizard II (JPK Instruments, Germany) combined with an inverted optical microscope (Axiovert 200 M, Zeiss, Germany) was used for all AFM measurements. Tapping mode imaging was conducted with OMCL-AC160TS (Olympus) cantilevers. For each bacterial strain, at least three images at random locations were obtained for each sample, and measurements were repeated with independently grown cultures to confirm that similar morphologies were observed.

### AFM Force Spectroscopy on Single Cells

AFM force spectroscopy on single cells was performed in the following way: a colloid glued on an AFM cantilever (see below) was approached on single cells immobilized on glass slides, kept at force set point for a specific amount of time, and then retracted. The deflection of the cantilever and displacement of the cantilever as a result of the piezo movement was recorded. The deflection-displacement curves were converted to force-distance curves (force curves hereafter) after calibration of cantilever sensitivity and spring constant. Sensitivity was determined by recording force curves on glass slides and spring constant was calibrated using the thermal tuning method (Hutter and Bechhoefer, 1993).

Tipless cantilevers with a nominal spring constant of 0.03 N/m (HQ:CSC38/TIPLESS/NO AL, MikroMasch Europe) were used for force spectroscopy. To attach a colloid, a tipless cantilever was approached on a small drop of UV curable adhesive (LOCTITE, part no. 17944) on glass slide and retracted. One to two additional approaches on clean glass slide were conducted to remove excessive glue, and the cantilever was then approached on a 5.6 μm silica colloid (Microparticles GmbH) under optical microscope and retracted after 1 min. One minutes UV irradiation with the built-in mercury lamp was used to cure the glue, leading to formation of a colloid probe. The colloid probe was cleaned by UV/Ozone for 20 min to yield a hydrophilic surface. To make the colloid hydrophobic, the cleaned colloid probe was incubated in anhydrous toluene with 1% v/v (3,3,3-Trifluoropropyl)trimethoxysilane overnight, and washed with toluene and trichloromethane.

For successful force spectroscopy on single cells, cells must be firmly attached on the substrate. To achieve this, glass slides were coated with wet adhesive polydopamine. Four milligram per milliliter dopamine in 10 mM Tris buffer pH 8.5 were incubated on glass slides for 1 h, after which the glass slides were washed with MilliQ water.

Bacterial samples were prepared almost the same way as for AFM imaging, except that cells were washed with PBS instead of water to avoid stressing the cells. This was sufficient to obtain individual cells of the WT strain. As pFap cells are mostly severely flocculated, 30 s sonication (Branson B5510, 135W) was used to break aggregates and obtain a reasonable amount of single cells for measurements. A 5  $\mu\text{m}$  cell suspension in PBS was placed on a coated glass slide for 5 min and washed extensively with PBS to remove unattached cells. A 100  $\mu\text{L}$  drop of PBS was added for liquid AFM force spectroscopy. A colloid probe was calibrated on a glass slide, approached at 4  $\mu\text{m/s}$  on an immobilized single cell for 5 s at 1 nN setpoint, and then retracted. The positioning of the colloid probe was done with optical microscope under 40X long working distance objective lens. The process can be facilitated by using DirectOverlay on NanoWizard II, which allows identification of the tip of the colloid and automation of the measurement on multiple cells. Successful approaches on single cells were seen as a gradual increase of repulsive force after contact point, whereas approaches on coated glass gave a sharp linear increase after contact. More than 10 force curves were recorded on each cell, and at least 10 cells were measured.

### AFM Force Spectroscopy on Cell Aggregates

The same procedure for making samples for AFM imaging was used. Cell aggregates were first imaged by AFM imaging in air to aid the selection of positions of measurement, and then added with PBS to perform force spectroscopy with colloid probe. The drying and rehydration process ensured the firm attachment of cell aggregates in liquid and therefore no coating was needed. More than 10 force curves were recorded at each position, and at least 5 positions were measured.

### AFM Nanoindentation on Biofilm

Biofilm of WT and the 8 different mutants was grown on 20 mm square glass coverslips placed in 12-well plates with CFA medium inoculated with colonies from agar plates. The plates were incubated at 26°C for 40 h with no shaking. A glass microbead (nominal size 59.2  $\mu\text{m}$ , G4649-10G, Sigma-Aldrich) was glued to a tipless AFM cantilever using the same procedure as for the colloid probe. The actual size of the glass microbead was determined from optical images. To avoid contamination of the microbead probe during measurement, an antifouling coating was applied to the UV/ozone cleaned microbead probe by incubating in 100  $\mu\text{g/mL}$  PLL (20 kDa)-g(4.0)-PEG (5 kDa) (Susos AG, Dubendorf, Switzerland) in 10 mM HEPES for 2 h and subsequent washing with MilliQ water. Ten force curves were recorded on each of the 5 locations, and measurements were repeated once on different samples. Measurement locations were selected in the area where the biofilm is estimated to be more than 20  $\mu\text{m}$  thick by

reading the movement of the optical stage from the bottom to the top.

To calculate the Young's modulus ( $E$ ) of biofilm from the indentation force curves, Sneddon's modification of Hertz model of contact mechanics was employed (Hertz, 1882; Sneddon, 1965). When a stiff sphere indents a soft planar sample, the loading force  $F$  is related to indentation depth  $\delta$  by:

$$F = \frac{4}{3} \frac{E}{(1 - \nu^2)} R \delta^{3/2}, \quad (1)$$

where  $R$  is the radius of the sphere and  $\nu$  is the Poisson ratio (assumed to be 0.5) of the sample. The Hertz model requires several assumptions to be met: the adhesion force between sphere and sample is negligible; the sample is homogeneous and semifinite (i.e., the indentation is small compared to the thickness of the film) and the sample undergoes small strains within the elastic limit. While these assumptions are almost impossible to be met at the same time for most biological samples, the model was proven to be useful for soft biological samples when indentation was limited to a small range (<10%) compared to the sample thickness and the loading speed is low enough to minimize contribution from plastic deformation and hydrodynamic forces. The indentation depth in the current study was limited to 2  $\mu\text{m}$  and loading speed was set at 1  $\mu\text{m}$ . The indentation curves were fit by Hertz model using JPK Data Processing software, giving the contact points and Young's moduli ( $E$ ) of the samples.

The plastic properties of the samples were characterized by the plasticity index  $\psi_P$  in the form:

$$\psi_P = \frac{A_1}{A_1 + A_2}, \quad (2)$$

where  $A_1$  and  $A_2$  are the energy of plastic deformation and elastic recovery, respectively, calculated from the area in force curves.

### Hydrophobicity of Bacteria

Bacterial cells were harvested and washed in water as described above. A bacterial lawn was coated on a glass slide by drying a 100  $\mu\text{L}$  droplet of bacterial suspension overnight. Hydrophobicity of bacteria was characterized by dynamic contact angle measurement using a Drop Shape Analyzer DSA100 (Krüss GmbH, Germany). A water droplet was formed on the bacterial lawn and manually expanded by syringe injection. The change of the contact angle over time was monitored by video recording and submitted to frame by frame analyzation, and contact angles were plotted against time. Due to the high surface heterogeneity, contact angles rose, and fell in repetitive patterns over time. The highest peak of contact angles, which reflected the state the liquid was about to wet new surface, was determined as advancing contact angles. Three samples were tested for each strain.

### Statistics

Data were presented as mean  $\pm$  SD. Student's  $t$ -test was used for comparing two groups of data, and significant difference was claimed when  $P < 0.05$ .

## Results

### Staining Confirms the Presence of Amyloid in the *Pseudomonas* Strain Overexpressing pFap

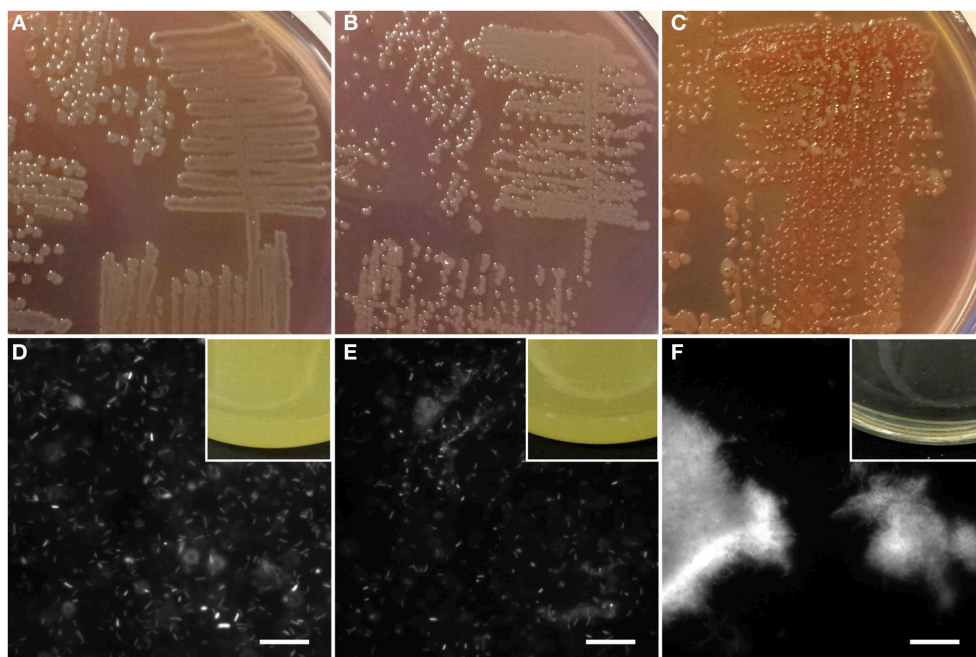
We started out by creating appropriate derivatives of *Pseudomonas* sp. UK4 to allow for proper comparison. We constructed a whole *fap* operon deletion mutant ( $\Delta fap$ ). This involves deletion of the open reading frames for the major amyloid component FapC as well as the ancillary proteins FapA-B and FapD-F (Dueholm et al., 2010, 2013b). Since UK4 wildtype (WT) only produces amyloid to a small degree, we also cloned the *fap* operon overexpression vector pMMB190Tc-UK4*fap*, where expression was under the control of the lacUV5 promoter, into the  $\Delta fap$  strain. This led to the pFap strain.

To confirm the differences in amyloid production, all three strains were grown on agar plates containing the amyloid-binding dye Congo Red. On such agar plates, amyloid-producing strains can be identified based on their red-brown coloration caused by the binding of Congo Red as well as their opaque or agglutinated appearance resulting from cell aggregation. Gratifyingly, only pFap showed the characteristic red-brown color indicative of amyloid formation, in contrast to WT and  $\Delta fap$ , which showed the green color characteristic of amyloid-free *Pseudomonas fluorescens* (Figures 1A–C). In addition, pFap yielded smaller colonies as a result of reduced mobility caused by the production of the aggregative amyloid and to a minor extent the increased growth cost due to the overexpression of the amyloids. The phenotype difference was further confirmed by growth in liquid culture, where

WT and  $\Delta fap$  led to a homogeneous cloudy suspension of bacteria while pFap flocculated and appeared as a clear solution with biofilm accumulating on the sides of the growth vessel (Figures 1D–F).

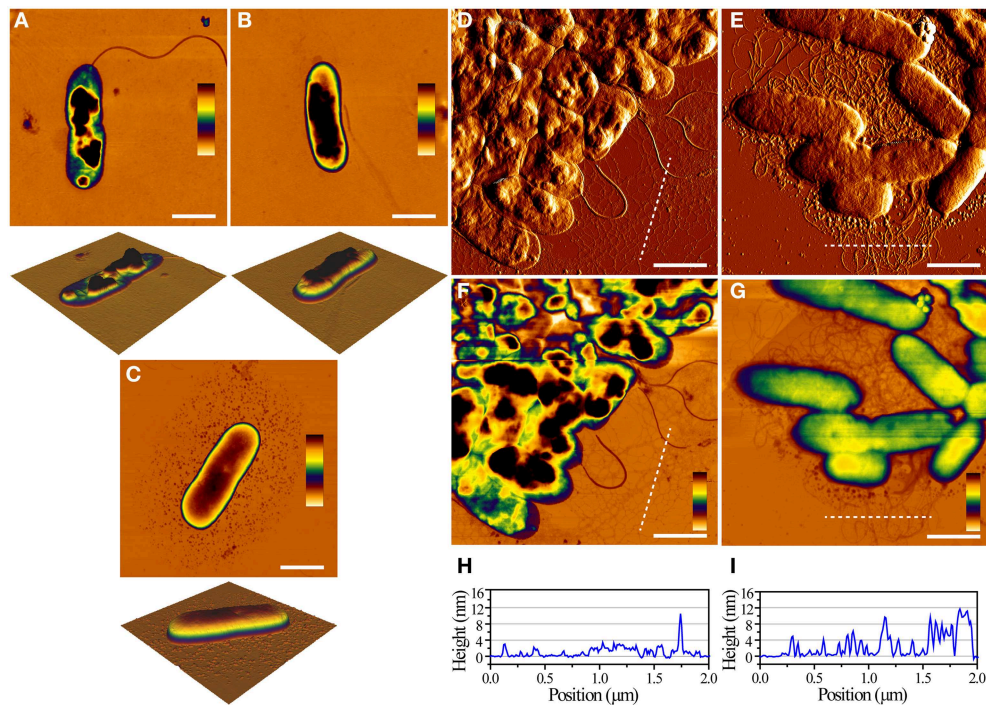
### Fap Amyloid Confers Resistance Against Desiccation According to AFM

We next imaged the individual cells using Atomic Force Microscopy. It was very straightforward to obtain and visualize individual WT and  $\Delta fap$  cells due to their low degree of intercellular contacts, while pFap cells had to be sonicated briefly to dissociate the cellular clumps. We consistently observed that both the WT and  $\Delta fap$  cells collapsed after brief drying in air, while pFap cells were able to maintain the rod-like shapes seen for bacteria in solution (Figures 2A–C). This difference in cell integrity was also observed when analyzing clumps of cells of WT and pFap (Figures 2D–I). Flagella are easily visible for both strains, while extracellular fibrils are clearly much more prominent for pFap. The height profiles show that pFap forms fibrils that are much higher and wider than those of WT and tend to aggregate to form larger co-existing bundles, consistent with the cells' pronounced tendency to flocculate. Overall there was little morphological difference between WT and  $\Delta fap$  in these studies, in good agreement with their shared inability to make detectable amounts of amyloid and flocculate. Consequently we concentrate on WT rather than  $\Delta fap$  in the following sections. pFap's ability to resist collapse upon drying, on the other hand, indicates a high level of structural robustness that is presumably conferred by the amyloid.

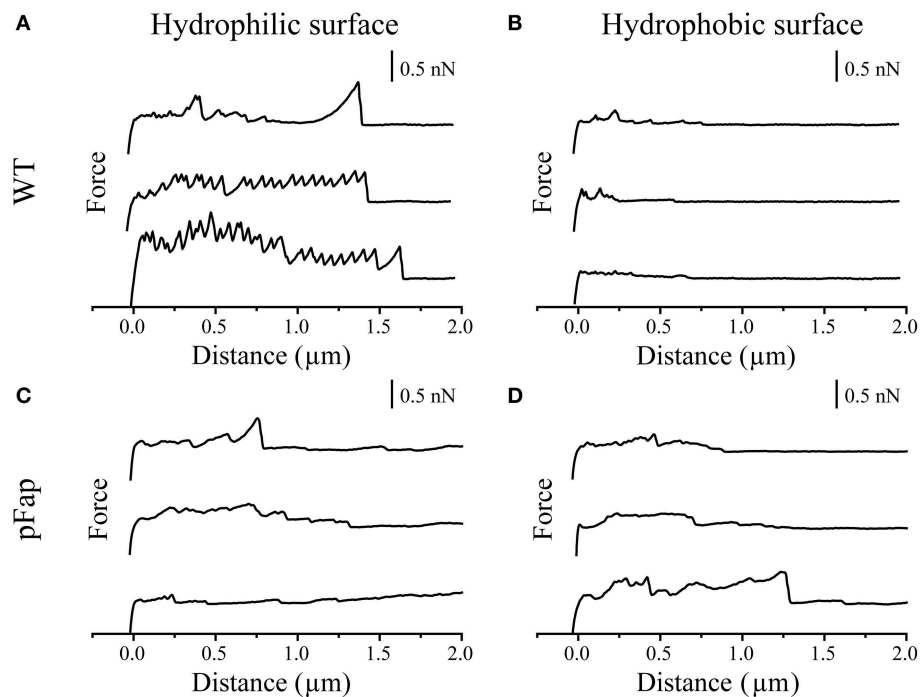


**FIGURE 1 | Identification of amyloid production by staining.** Bacterial colonies of WT (A),  $\Delta fap$  (B), and pFap (C) grown on Congo red-CFA agar plates. Fluorescent images of liquid cultures of WT (D),  $\Delta fap$  (E), and pFap (F) stained with ThT, scale bar = 10  $\mu$ m. Insets: side-views of liquid cultures grown in flasks.





**FIGURE 2 | AFM images of single cells and cell aggregates.** Height images of WT (A, height scale = 104 nm),  $\Delta fap$  (B, height scale = 122 nm), and pFap (C, height scale = 213 nm) cells, scale bar = 1  $\mu m$ . Error images of WT (D) and pFap (E) aggregates. Corresponding height images of WT (F) and pFap (G) aggregates, scale bar = 1  $\mu m$ , and line profiles of WT (H) and pFap (I) height images marked by dashed lines.



**FIGURE 3 | AFM force curves of colloids with different surface properties on single cells.** (A) Hydrophilic surface on WT; (B) Hydrophobic surface on WT; (C) Hydrophilic surface on pFap; (D) Hydrophobic surface on pFap. Only retraction parts of the force curves are shown.

## AFM Force Spectroscopy on Single Cells Reveals that Fap Amyloid Induces Different Cellular Adhesion Patterns

To better understand the adhesion forces conferred by amyloid on the surface of *Pseudomonas* sp. UK4, we investigated to what extent cells with and without amyloid adhered to hydrophilic and hydrophobic probes. These were constructed by gluing a 5- $\mu\text{m}$  silica colloid onto the very tip of the cantilever, curing the glue with UV and functionalizing the probe surface to make it hydrophilic (UV-oxidation) or hydrophobic (silanization). Individual cells were then immobilized in an evenly distributed pattern onto a polydopamine surface and were exposed to contact with the colloid probe. The probe first approaches each cell and is then retracted while the associated cantilever deflection and movement is recorded. These data are then converted to force and probe-sample distance and plotted into force curves. The retraction force curves are shown in **Figure 3** for both WT and pFap.

The sawtooth-like adhesion peaks of WT on hydrophilic surface can probably be assigned to the unfolding of the protein adhesin LapA, which is important for adhesion and surface attachment of many *Pseudomonas* species, in particular those belonging to the *P. fluorescens* and *P. putida* group (Duque et al., 2013). Accordingly, sequence analysis suggests that *Pseudomonas* sp. UK4 has a *lapA* homolog. LapA is a giant protein containing tandem repeats whose individual unfolding during retraction leads to the characteristic saw-tooth pattern, as previously reported (El-Kirat-Chatel et al., 2014). For WT, these peaks are absent on a hydrophobic surface due to different interaction modes of LapA with hydrophobic surfaces, as reported earlier (El-Kirat-Chatel et al., 2014). The absence of these saw-tooth-like peaks from force curves of pFap on hydrophilic surfaces suggests that the pFap cells are covered by other surface molecules (i.e., amyloid), thus shielding the interactions from LapA, similar to the effect of capsular polysaccharides on the short-range adhesion Ag43 in *E. coli* (Schembri et al., 2004).

The force curves directly provide the maximal adhesion force (the highest point in the retraction curve) and the final rupture length (the longest separation length at which the cell ruptures from the probe). As summarized in **Table 2**, the average values suggest that both the maximal adhesion force and rupture length of pFap are larger than those of WT on the hydrophobic surface. However, likely because LapA contributes to the adhesion, WT adheres more strongly than pFap to hydrophilic surfaces. Note that the errors on these measurements are considerable despite measurements on a considerable number of samples, and the differences are not statistically significant. This is remedied when looking at collections of cells in biofilm rather than the individual cells (see below).

## Fap Amyloid Enhances Adhesion to Spherical Probes, Especially those with Hydrophobic Surfaces, and Increases the Contact Angle of Water Droplets

The situation became considerably clearer when we instead turned to force spectroscopy measurements on cell aggregates

**TABLE 2 | Maximal adhesion force and final rupture length of WT and pFap cells on hydrophilic and hydrophobic surfaces<sup>a</sup>.**

Surface type	WT		pFap	
	Max. adhesion force (nN)	Final rupture length ( $\mu\text{m}$ )	Max. adhesion force (nN)	Final rupture length ( $\mu\text{m}$ )
Hydrophilic	0.61 $\pm$ 0.32	1.07 $\pm$ 0.45	0.30 $\pm$ 0.17	1.38 $\pm$ 0.76
Hydrophobic	0.32 $\pm$ 0.17	0.70 $\pm$ 0.47	0.50 $\pm$ 0.26	1.40 $\pm$ 0.82

<sup>a</sup>Based on data shown in **Figure 3**.

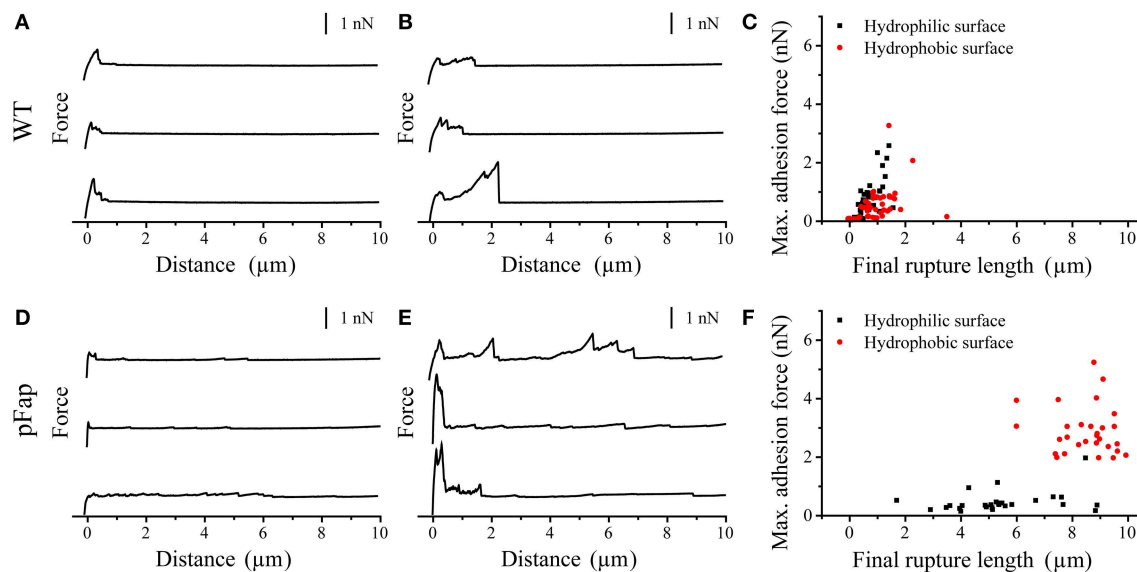
(**Figure 4**). We did not observe any saw-tooth patterns in the force curves on biofilm, indicating that all cells in biofilm are completely covered by EPS. EPS are therefore the major components which contribute to the adhesive properties of biofilm. For WT, rupture lengths and maximal adhesion forces were uniformly small and not statistically different for the two types of surfaces (**Table 3**). In contrast, pFap shows much longer rupture lengths on both surfaces than WT does, and the maximal adhesion force is 5-fold increased compared to WT on hydrophobic surfaces (but not on hydrophilic surfaces). We see a large variation in the final rupture length particularly for pFap on hydrophilic surfaces. We attribute this to the fact that EPS in biofilm are composed of polymers of various lengths. At each measurement random molecules or molecular assemblies are picked up by the probe at random positions along the molecules, leading to varying final rupture length. While EPS is present in biofilm for both WT and pFap, it is possible that interactions between amyloid, other EPS components and the hydrophilic surface lead to a particularly large variation.

The high adhesion forces and increased rupture lengths for pFap on hydrophobic surfaces suggested that the biofilm might be more hydrophobic itself. We confirmed this by measuring the contact angle formed when water drops are deposited on cell lawns (**Figure 5**). Water will try to minimize contact with a hydrophobic surface, so the larger the contact angle, the more hydrophobic the surface. Due to the instability of static water drops on porous and absorptive cell lawns, we had to carry out dynamic measurements where contact angles were recorded over time. Contact angles at the critical points where water drops spreads to make contact with new surfaces were defined as advancing contact angles. The advancing contact angles of WT and pFap were  $95.2 \pm 1.1$  and  $128.1 \pm 0.6^\circ$ , respectively. The measured contact angle may overestimate the actual contact angle of the bacterial surface, because the micro scale porosity of the bacterial lawn could lead to an increase in the contact angle. However, the data clearly suggest that pFap is more hydrophobic, in good agreement with our AFM measurements.

## AFM Nanoindentation Measurements Reveal that Fap Amyloid Increases Biofilm Stiffness 20-Fold

To gain insight into the mechanical properties of the biofilm, we carried out nanoindentation experiments, in which a large microbead was pressed into the bacterial biofilm using the AFM cantilever (**Figure 6**). During the approach of the probe, a repulsive force is generated immediately after the bead establishes contact with the biofilm, and the approach continues until a





**FIGURE 4 | AFM force curves of colloids with different surface properties on cell aggregates. (A)** Hydrophilic surface on WT; **(B)** Hydrophobic surface on WT; **(C)** Max. adhesion force vs. final rupture length of the force curves on WT; **(D)** Hydrophilic surface on pFap; **(E)** Hydrophobic surface on pFap; **(F)** Max. adhesion force vs. final rupture length of the force curves on pFap. Only retraction parts of the force curves are shown.

**TABLE 3 | Maximal adhesion force and final rupture length of WT and pFap cell aggregates on hydrophilic and hydrophobic surfaces<sup>a</sup>.**

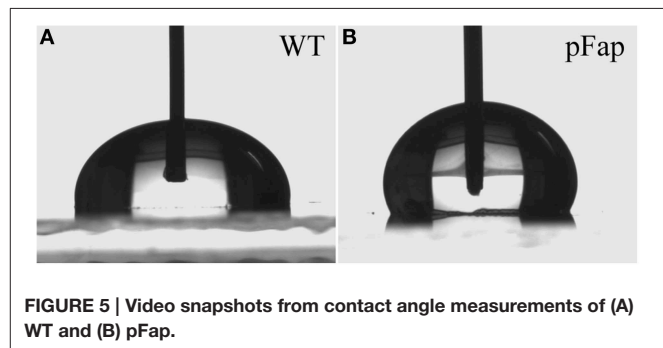
Surface type	WT		pFap	
	Max. adhesion force (nN)	Final rupture length (μm)	Max. adhesion force (nN)	Final rupture length (μm)
Hydrophilic	0.707 ± 0.616	0.653 ± 0.368	0.473 ± 0.365	5.430 ± 1.762
Hydrophobic	0.524 ± 0.566	0.822 ± 0.863	2.891 ± 0.819	8.486 ± 1.007

<sup>a</sup>Based on data shown in Figure 4.

repulsive force of ~2 nN is reached, after which the the probe is retracted. The approach and retraction curves do not overlap completely, because deformation of biofilm is not completely elastic, i.e., there is a certain degree of plastic deformation that occurs. This can be quantified by the plasticity index  $\psi_P$  (Equation 2). Plasticity indices  $\psi_P$  of WT and pFap were  $0.31 \pm 0.03$  and  $0.33 \pm 0.02$ , respectively ( $P > 0.05$ ,  $df = 6$ ), indicating that they have similar plastic deformation. However, the indentation depth under the same load (2 nN) was much larger for WT ( $1.540 \pm 0.047 \mu\text{m}$ ) than for pFap ( $0.238 \pm 0.025 \mu\text{m}$ ), indicating that the pFap film is much stiffer. The indentation curves can be further fitted to the Hertz model of contact mechanics to provide Young's modulus, which is a measure of biofilm stiffness. Indeed, the fitted Young's modulus of pFap ( $2.01 \pm 0.08 \text{ kPa}$ ) is 20 times as high as for WT ( $0.10 \pm 0.01 \text{ kPa}$ ).

### All Fap Genes Except FapA Contribute to Biofilm Stiffness

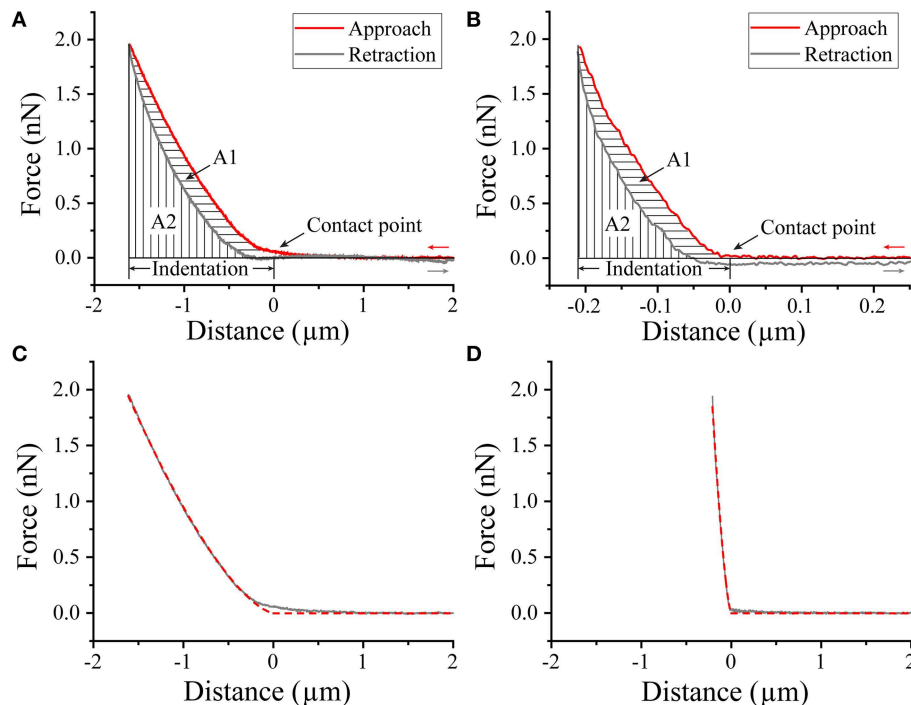
To evaluate the contribution of individual protein components to the biofilm, we constructed plasmids containing *fap* operons



**FIGURE 5 | Video snapshots from contact angle measurements of (A) WT and (B) pFap.**

missing each one of the six genes in turn. We then evaluated how these individual deletions affected Congo Red binding, fibril formation and biofilm stiffness. The only deletion mutant which produced the classical amyloid positive colony morphology on the Congo Red agar plates was pFapΔA (Figure 7A). Since Congo Red is not able to penetrate the intact bacterial outer membrane (Sleytr et al., 1988), binding implies successful export of amyloid for pFapΔA. pFapΔB and pFapΔF produced transparent colonies that bound Congo Red weakly. The weak binding of Congo Red by pFapΔB and pFapΔF are likely caused by non-amyloid components, probably polysaccharides. It has previously been shown that overexpression of the *fap* operon in *P. aeruginosa* can induce alginate production (Herbst et al., 2015) and this effect may be independent of FapB and FapF. The remaining deletion strains displayed morphologies similar to Δ*fap*.

The consequences of the deletion of these genes become clearer when individual cells are visualized by TEM. Deletion of all individual genes except *fapA* significantly reduces the amount



**FIGURE 6 | AFM Nanoindentation of biofilm.** Approach-retraction force curves on WT (A) and pFap (B) biofilm illustrating contact point, indentation depth, and plastic deformation (A1) and elastic recovery (A2). Indentation curves (approach part of the force curves, solid lines) and fit to Hertz model (dashed lines) for WT (C) and pFap (D).

of *bona fide* curly fibrillary outgrowths from *Pseudomonas* (Figure 7B). Only pFap $\Delta$ A shows curly structures comparable to (though less dense than) those of pFap. pFap $\Delta$ C has a halo of fuzzy structures which cannot be amyloid (cfr. Figure 7A), and pFap $\Delta$ D has a mesh-like cover that likely represents excreted vesicles. These differences are strikingly confirmed by biofilm stiffness measurements (Figure 7C), which reveal that only the pFap $\Delta$ A mutant leads to biofilm with a stiffness comparable to that of the intact *fap* operon; all other constructs show stiffness levels comparable to that of the empty vector. Note also that the pFap $\Delta$ A mutant leads to overproduction of FapB at the expense of FapC (Dueholm et al., 2013a), explaining the change in appearance of the cells in Figure 7B.

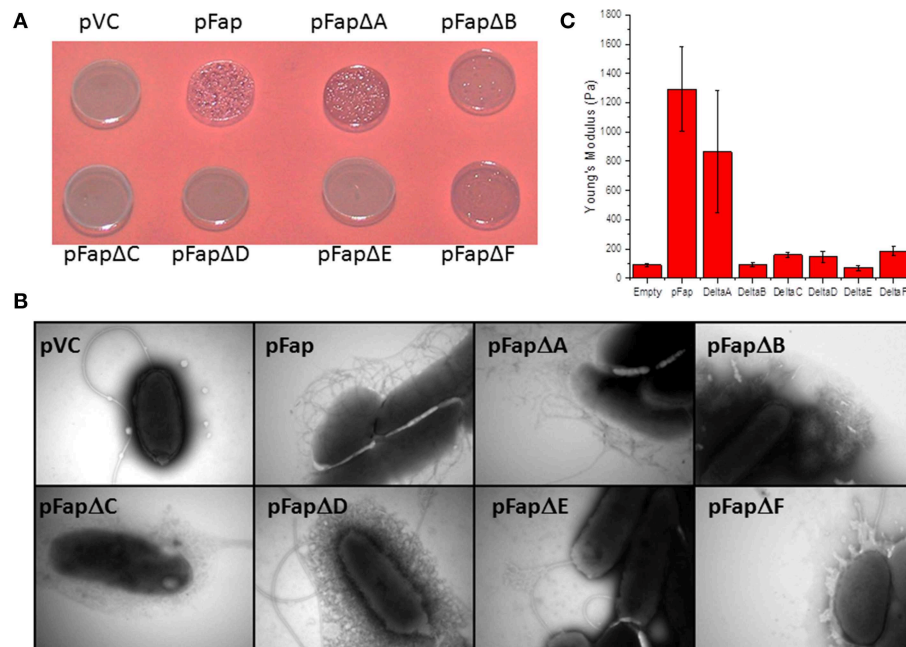
## Discussion

### Role of Amyloid in the Mechanical Robustness of Biofilms

We have established that functional amyloid expression in *Pseudomonas* significantly affects cell properties, such as ability to withstand drying, hydrophobicity and biofilm rigidity. It cannot be ruled out that these effects are linked to secondary changes induced by overexpression of amyloid, such as altered expression patterns of other components in the EPS. A recent proteomic study (Herbst et al., 2015) showed that *fap* expression led to upregulation of alginate biosynthesis, turning the *P. aeruginosa* PAO1 into a mucoid phenotype. While mucoid biofilms are thicker, rougher and more antibiotic-resistant than

the non-mucoid versions which are flat and dense (Hentzer et al., 2001), their impact on biofilm stiffness is unclear and might intuitively be expected to lead to a more expanded structures. Our data make it clear that amyloid expression and formation on the surface is linked to major changes in cell robustness. Furthermore, deletion of individual members of the *fap* operon (which is not expected to alter the cellular response to *fap* expression significantly) completely removes this increased biofilm stiffness. We have previously reported that *Pseudomonas* amyloid is able to bind small metabolites (Seviour et al., 2015), and this binding ability correlates with metabolite hydrophobicity. This strongly suggests that amyloid in itself is hydrophobic and provides a straightforward driving force for intercellular contacts. It also highlights the role of amyloid in providing high elastic strength and thus mechanical robustness to the biofilm. We speculate that this property is conferred because the relatively stiff amyloid fibrils are embedded within a more flexible matrix provided by eDNA and polysaccharides, analogous to the amyloid-amorphous phase combination seen in spider dragline silk (Hagn, 2012) or even reinforced concrete (steel wires within the concrete matrix). Increased protein content in the EPS has also been associated with increased biofilm strength (Pellicer-Nácher and Smets, 2014) in a study of nitrifying membrane-aerated biofilms, though in this case the proteins rendered the biofilm more hydrophilic.

It is clear that increased biofilm strength requires production of a substantial amount of amyloid. Although the mutants lacking



**FIGURE 7 | Analysis of the effect of removing individual genes from the *fap* operon.**  $\Delta fap$  was transformed with the empty plasmid (pVC), a plasmid containing the complete *fap* operon (pFap) or plasmids containing the *fap* operon with the individual *fap* genes deleted. **(A)** Congo Red binding. Cells were grown on CFA agar plates containing Congo Red. A dark red color indicates amyloid formation. **(B)** Transmission electron microscopy images of cells taken from cell cultures at stationary phase. **(C)** Indentation experiments showing the value of Young's modulus for biofilms grown with the different mutants.

FapB and FapF showed some level of Congo Red binding, the overall production of extracellular amyloid had clearly been reduced, and correspondingly the biofilm stiffness was just as low as that of the  $\Delta fap$  mutant. Based on its sequence homology to FapC, we have proposed that FapB is a nucleator for amyloid formation on the outer membrane surface (Dueholm et al., 2010), analogous to CsgB (Hammer et al., 2007). Sequence comparisons also suggest that FapF provides a conduit for transport of FapC (and FapB) through the outer membrane. Thus, removal of these two components should be definitely reduce FapC export and formation of the amyloid state on the bacterial cell surface. It remains more mysterious why deletion of FapE (a minor component of amyloid) and FapD (a putative peptidase) should completely abolish amyloid production, and further studies are clearly needed to delineate their specific contributions to the production of amyloid. In contrast, the removal of FapA does not compromise the overall function of the *fap* operon, but largely seems to affect the balance between FapC and the other amyloid component FapB which is predicted to be a fibril nucleator (Dueholm et al., 2013a) but can also fibrillate well on its own *in vitro* (B.S.V. and D.E.O., unpublished).

We note that the major component in *Pseudomonas* amyloid, the protein FapC, consists of 3 imperfect repeats connected by linkers of variable length (Dueholm et al., 2010). The repeats likely constitute the amyloid core since their removal completely abolishes FapC's ability to fibrillate (B.S.V. and D.E.O., unpublished), while the linkers show variable length but largely contain hydrophilic residues. It is possible that

the hydrophobicity of the biofilm, as well as the interactions with other EPS components, may be modulated by the composition and length of the linkers. This in turn may also affect the mechanical properties of the ensuing biofilm. Our nanoindentation assay provides a very convenient tool to assess this. Given the important role of mechanical forces between bacterial cells in the early build-up of biofilm (Grant et al., 2014), amyloid can also play a role in promoting rapid establishment of biofilm. We have already observed this in a number of *Pseudomonas* strains overexpressing different variants of the *fap* operon, which led to more rapid surface colonization and alterations in biofilm morphology (Dueholm et al., 2013a).

### Measurement of Hydrophobicity by Contact Angle

Static contact angle measurement has been used by researchers to characterize the hydrophobicity of bacterial cells (Busscher et al., 1984; Fernández et al., 2007; Seale et al., 2008; Gallardo-Moreno et al., 2011). We found it difficult to obtain consistent results across different replicates using this measurement, probably due to the fast penetration of water droplets into the bacteria lawns. We instead used advancing contact angle measurement for determining hydrophobicity (Liu et al., 2008), which yielded more reproducible contact angles. However, caution must be taken to use contact angles on bacterial lawns as a direct measurement of hydrophobicity. As pointed out by a recent report investigating this complex topic (Gallardo-Moreno et al., 2011), contact angles on dried bacterial lawns are so strongly

dependent on measuring time and environmental conditions that an absolute measurement of hydrophobicity and Gibbs energy of bacterial cell surfaces is essentially impossible. Nevertheless, the method is still useful as an indication of relative hydrophobicity, especially when all measurements are done using the same procedure.

## Colloidal Probes as a Tool to Determine Biofilm Mechanical Properties

Since its invention in the early 1990s (Butt, 1991; Ducker et al., 1991), AFM force spectroscopy with colloidal probes has been widely used to measure the mechanical properties of soft biological materials (Rosenbluth et al., 2006; Sokolov et al., 2007; Plodinec et al., 2012). Compared to the conventional sharp AFM tips, colloidal probes have more well-defined contact areas, and they are therefore recommended for measurement when spatial resolution is not of concern (Dimitriadis et al., 2002; Loparic et al., 2010). Penetration of soft materials can also be avoided by colloidal probes. In the current study, relatively large microbeads were selected for nanoindentation measurements primarily because large contact area provides averaging of the mechanical properties, that is, the global instead of local mechanical properties are measured (Loparic et al., 2010). This eliminates large variation of the results as the biofilm is usually highly heterogeneous. This is confirmed by large data variation from indentation with 5  $\mu\text{m}$  colloid probe (data not shown). Besides, the Hertz model for spherical indenter used to calculate the mechanical properties of the sample, assumes that the contact radius is less than 10% of the indenter radius. This assumption cannot be met when using a small sphere on a soft sample which usually results in a large indentation depth for biofilm samples, even with small indentation forces.

Biofilm elasticity measured by colloidal probe indentation is in the current study within the range of 0.017–170 kPa reported previously using different techniques (Stoodley et al., 1999, 2001; Körstgens et al., 2001; Lau et al., 2009). The wide range of biofilm elasticities is due to the inherent differences among biofilm matrix composition in various bacterial strains as well as the growth conditions, sample preparation, and different methodologies used. One needs to be careful in comparing results using different techniques. A most relevant case to our study is that Casey and co-workers recently used AFM colloidal probe to measure the elastic modulus of *P. fluorescens* PCL1701 biofilm, and reported mean elastic moduli of  $33 \pm 22$  kPa and  $2.7 \pm 1.3$  kPa for calcium-free and calcium-supplemented biofilm,

respectively (Safari et al., 2014). Moreover, they also detected a softer surface layer of the calcium-supplemented biofilm with an elastic modulus of  $0.39 \pm 0.24$  kPa up to an indentation depth of  $1.27 \pm 0.33$   $\mu\text{m}$ . This data is comparable to our measurement on WT, which should not be surprising because our CFA growth medium contains divalent ions of magnesium and manganese.

The roles of amyloids as biofilm structural components have been demonstrated in several species. Curli amyloids are essential for *Escherichia coli* (Vidal et al., 1998) and *Salmonella enteritidis* (Austin et al., 1998) to attach to the surface and form thick biofilm. Amyloid fibers formed from TasA protein in *Bacillus subtilis* have also been shown to be required for the formation of robust pellicle biofilms, while aggregation of phenol-soluble modulins protect *S. aureus* biofilm against mechanical and enzymatic attack (Schwartz et al., 2012). We have previously shown that Fap amyloids facilitated *Pseudomonas* biofilm formation (Dueholm et al., 2013a). All these studies, however, are limited to morphological investigation. The study we present here is, as far as we know, the first to investigate the effect of amyloids on biofilm mechanical properties in a quantitative way.

## Robust Biofilm as a Key to Successful Colonization?

The extraordinary mechanical robustness that functional amyloids provide to bacterial biofilms raises the question of their role in biofilm survival in different contexts. *Pseudomonas fluorescens* is non-pathogenic and ubiquitous in the environment where it successfully colonizes inorganic or biological (e.g., plants) surfaces in, e.g., soil and aquatic environments. Other *Pseudomonas* species are also wide spread. One might speculate that the functional amyloid of *P. fluorescens* and other *Pseudomonas* species is one of the keys to the robustness required to successfully colonize such diverse habitats. Future studies will no doubt reveal new insights into the role of functional amyloids in the ecology of bacteria, and how amyloid production in biofilm affect the biofilm's resistance to environmental stressors, grazing predators, or production of antimicrobial compounds from competing microorganisms.

## Acknowledgments

RM and GZ gratefully acknowledge The Danish Council for Independent Research (DFF) Sapere Aude Starting Grant (0602-02130B).

## References

- Alhede, M., Bjarnsholt, T., and Givskov, M. (2014). *Pseudomonas aeruginosa* biofilms: mechanisms of immune evasion. *Adv. Appl. Microbiol.* 86, 1–40. doi: 10.1016/B978-0-12-800262-9.00001-9
- Austin, J. W., Sanders, G., Kay, W. W., and Collinson, S. K. (1998). Thin aggregative fimbriae enhance *Salmonella enteritidis* biofilm formation. *FEMS Microbiol. Lett.* 162, 295–301. doi: 10.1111/j.1574-6968.1998.tb13012.x
- Busscher, H. J., Weerkamp, A. H., Vandermei, H. C., Vanpelt, A. W. J., Dejong, H. P., and Arends, J. (1984). Measurement of the surface free-energy of bacterial-cell surfaces and its relevance for adhesion. *Appl. Environ. Microbiol.* 48, 980–983.
- Butt, H. J. (1991). Measuring electrostatic, van der Waals, and hydration forces in electrolyte solutions with an atomic force microscope. *Biophys. J.* 60, 1438–1444. doi: 10.1016/S0006-3495(91)82180-4
- Chapman, M. R., Robinson, L. S., Pinkner, J. S., Roth, R., Heuser, J., Hammar, M., et al. (2002). Role of *Escherichia coli* curli operons in directing



- amyloid fiber formation. *Science* 295, 851–855. doi: 10.1126/science.1067484
- Choi, K. H., Kumar, A., and Schweizer, H. P. (2006). A 10-min method for preparation of highly electrocompetent *Pseudomonas aeruginosa* cells: application for DNA fragment transfer between chromosomes and plasmid transformation. *J. Microbiol. Methods* 64, 391–397. doi: 10.1016/j.mimet.2005.06.001
- Choi, K. H., and Schweizer, H. P. (2005). An improved method for rapid generation of unmarked *Pseudomonas aeruginosa* deletion mutants. *BMC Microbiol.* 5:30. doi: 10.1186/1471-2180-5-30
- Costerton, J. W., Stewart, P. S., and Greenberg, E. P. (1999). Bacterial biofilms: a common cause of persistent infections. *Science* 284, 1318–1322. doi: 10.1126/science.284.5418.1318
- Dimitriadis, E. K., Horkay, F., Maresca, J., Kachar, B., and Chadwick, R. S. (2002). Determination of elastic moduli of thin layers of soft material using the atomic force microscope. *Biophys. J.* 82, 2798–2810. doi: 10.1016/S0006-3495(02)75620-8
- Ducker, W. A., Senden, T. J., and Pashley, R. M. (1991). Direct measurement of colloidal forces using an atomic force microscope. *Nature* 353, 239–241. doi: 10.1038/353239a0
- Dueholm, M., Halkjær, P. H., Chapman, M. R., and Otzen, D. E. (2012). “Functional amyloids in Bacteria,” in *Amyloid Fibrils and Prefibrillar Aggregates*, ed D. E. Otzen (Wiley-VCH Verlag GmbH), 411–438.
- Dueholm, M., Otzen, D. E., and Nielsen, P. H. (2013b). Evolutionary insight into the functional amyloids of the Pseudomonads. *PLoS ONE* 8:e76630. doi: 10.1371/journal.pone.0076630
- Dueholm, M., Søndergaard, M. T., Nilsson, M., Christiansen, G., Stensballe, A., Overgaard, M. T., et al. (2013a). Expression of Fap amyloids in *Pseudomonas aeruginosa*, *P. fluorescens*, and *P. putida* results in aggregation and increased biofilm formation. *Microbiol. Open* 2, 365–382. doi: 10.1002/mbo3.81
- Dueholm, M. S., Larsen, P., Finster, K., Stenvang, M. R., Christiansen, G., Vad, B. S., et al. (2015). The tubular sheaths encasing methanosaeta thermophila are functional amyloids. *J. Biol. Chem.* doi: 10.1074/jbc.M115.654780
- Dueholm, M. S., Petersen, S. V., Sønderkær, M., Larsen, P., Christiansen, G., Hein, K. L., et al. (2010). Functional Amyloid in *Pseudomonas*. *Mol. Microbiol.* 77, 1009–1020. doi: 10.1111/j.1365-2958.2010.07269.x
- Duque, E., de la Torre, J., Bernal, P., Molina-Henares, M. A., Alaminos, M., Espinosa-Urgel, M., et al. (2013). Identification of reciprocal adhesion genes in pathogenic and non-pathogenic *Pseudomonas*. *Environ. Microbiol.* 15, 36–48. doi: 10.1111/j.1462-2920.2012.02732.x
- El-Kirat-Chatel, S., Beaussart, A., Boyd, C. D., O'Toole, G. A., and Dufrene, Y. F. (2014). Single-cell and single-molecule analysis deciphers the localization, adhesion, and mechanics of the biofilm adhesin LapA. *ACS Chem. Biol.* 9, 485–494. doi: 10.1021/cb400794e
- Fernández, I. C. S., van der Mei, H. C., Lochhead, M. J., Grainger, D. W., and Busscher, H. J. (2007). The inhibition of the adhesion of clinically isolated bacterial strains on multi-component cross-linked poly(ethylene glycol)-based polymer coatings. *Biomaterials* 28, 4105–4112. doi: 10.1016/j.biomaterials.2007.05.023
- Flemming, H. C., and Wingender, J. (2010). The biofilm matrix. *Nat. Rev. Microbiol.* 8, 623–633. doi: 10.1038/nrmicro2415
- Gallardo-Moreno, A. M., Navarro-Pérez, M. L., Vadillo-Rodriguez, V., Bruque, J. M., and González-Martin, M. L. (2011). Insights into bacterial contact angles: difficulties in defining hydrophobicity and surface Gibbs energy. *Colloid Surface B* 88, 373–380. doi: 10.1016/j.colsurfb.2011.07.016
- Gloag, E. S., Turnbull, L., Huang, A., Vallotton, P., Wang, H., Nolan, L. M., et al. (2013). Self-organization of bacterial biofilms is facilitated by extracellular DNA. *Proc. Natl. Acad. Sci. U.S.A.* 110, 11541–11546. doi: 10.1073/pnas.1218898110
- Grant, M. A., Waclaw, B., Allen, R. J., and Cicuta, P. (2014). The role of mechanical forces in the planar-to-bulk transition in growing *Escherichia coli* microcolonies. *J. R. Soc. Interface* 11:400. doi: 10.1098/rsif.2014.0400
- Hagn, F. (2012). A structural view on spider silk proteins and their role in fiber assembly. *J. Pept. Sci.* 18, 357–365. doi: 10.1002/psc.2417
- Hammer, N. D., Schmidt, J. C., and Chapman, M. R. (2007). The curli nucleator protein, CsgB, contains an amyloidogenic domain that directs CsgA polymerization. *Proc. Natl. Acad. Sci. U.S.A.* 104, 12494–12499. doi: 10.1073/pnas.0703310104
- Hentzer, M., Teitzel, G. M., Balzer, G. J., Heydorn, A., Molin, S., Givskov, M., et al. (2001). Alginate overproduction affects *Pseudomonas aeruginosa* biofilm structure and function. *J. Bacteriol.* 183, 5395–5401. doi: 10.1128/JB.183.18.5395-5401.2001
- Herbst, F. A., Søndergaard, M. T., Kjeldal, H., Stensballe, A., Nielsen, P. H., and Dueholm, M. S. (2015). Major proteomic changes associated with amyloid-induced biofilm formation in *Pseudomonas aeruginosa* PAO1. *J. Proteome Res.* 14, 72–81. doi: 10.1021/pr500938x
- Hertz, H. (1882). Ueber die Berührung fester elastischer Körper. *Reine Ang. Mathem.* 92, 156–171.
- Hoang, T. T., Karkhoff-Schweizer, R. R., Kutchma, A. J., and Schweizer, H. P. (1998). A broad-host-range Flp-FRT recombination system for site-specific excision of chromosomally-located DNA sequences: application for isolation of unmarked *Pseudomonas aeruginosa* mutants. *Gene* 212, 77–86. doi: 10.1016/S0378-1119(98)00130-9
- Horton, R. M., Cai, Z., Ho, S. M., and Pease, L. R. (2013). Gene splicing by overlap extension: tailor-made genes using the polymerase chain reaction. *Biotechniques* 54, 129–133.
- Hutter, J. L., and Bechhoefer, J. (1993). Calibration of atomic-force microscope tips. *Rev. Sci. Instrum.* 64, 1868–1873. doi: 10.1063/1.1143970
- Jordal, P. B., Dueholm, M., Larsen, P., Pedersen, S. V., Enghild, J. J., Christiansen, G., et al. (2009). Widespread abundance of Functional Bacterial Amyloid in Mycolata and other Gram positive Bacteria. *Appl. Env. Microbiol.* 75 4101–4110. doi: 10.1128/aem.02107-08
- Knowles, T. P. J., Fitzpatrick, A. W., Meehan, S., Mott, H. R., Vendruscolo, M., Dobson, C. M., et al. (2007). Role of intermolecular forces in defining material properties of protein nanofibrils. *Science* 318, 1900–1903. doi: 10.1126/science.1150057
- Körstgens, V., Flemming, H. C., Wingender, J., and Borchard, W. (2001). Uniaxial compression measurement device for investigation of the mechanical stability of biofilms. *J. Microbiol. Methods* 46, 9–17. doi: 10.1016/S0167-7012(01)00248-2
- Larsen, P., Dueholm, M., Christiansen, G., Nielsen, J. L., Otzen, D. E., and Nielsen, P. H. (2007). Amyloid adhesins are abundant in natural biofilms. *Env. Microbiol.* 9, 3077–3090. doi: 10.1111/j.1462-2920.2007.01418.x
- Larsen, P., Nielsen, J. L., Otzen, D. E., and Nielsen, P. H. (2008). Amyloid-like adhesins in floc-forming and filamentous bacteria in activated sludge. *Appl. Env. Microbiol.* 74, 1517–1526. doi: 10.1128/AEM.02274-07
- Lau, P. C. Y., Dutcher, J. R., Beveridge, T. J., and Lam, J. S. (2009). absolute quantitation of bacterial biofilm adhesion and viscoelasticity by microbead force spectroscopy. *Biophys. J.* 96, 2935–2948. doi: 10.1016/j.bpj.2008.12.3943
- Lieleg, O., Caldara, M., Baumgartel, R., and Ribbeck, K. (2011). Mechanical robustness of *Pseudomonas aeruginosa* biofilms. *Soft Matter* 7, 3307–3314. doi: 10.1039/c0sm01467b
- Liu, X. M., Sheng, G. P., Wang, J., and Yu, H. Q. (2008). Quantifying the surface characteristics and flocculability of *Ralstonia eutropha*. *Appl. Microbiol. Biotechnol.* 79, 187–194. doi: 10.1007/s00253-008-1426-1
- Loparic, M., Wirz, D., Daniels, A. U., Raiteri, R., VanLandingham, M. R., Guex, G., et al. (2010). Micro- and nanomechanical analysis of articular cartilage by indentation-type atomic force microscopy: validation with a gel-microfiber composite. *Biophys. J.* 98, 2731–2740. doi: 10.1016/j.bpj.2010.02.013
- Mulcahy, H., Charron-Mazenod, L., and Lewenza, S. (2010). *Pseudomonas aeruginosa* produces an extracellular deoxyribonuclease that is required for utilization of DNA as a nutrient source. *Environ. Microbiol.* 12, 1621–1629. doi: 10.1111/j.1462-2920.2010.02208.x
- Otzen, D. E. (ed.). (2012). *Amyloid Fibrils and Prefibrillar Aggregates*. Weinheim: Wiley-VCH Verlag GmbH.
- Pellicer-Nácher, C., and Smets, B. F. (2014). Structure, composition, and strength of nitrifying membrane-aerated biofilms. *Water Res.* 57, 151–161. doi: 10.1016/j.watres.2014.03.026
- Plodinec, M., Loparic, M., Monnier, C. A., Obermann, E. C., Zanetti-Dallenbach, R., Oertle, P., et al. (2012). The nanomechanical signature of breast cancer. *Nat. Nanotechnol.* 7, 757–765. doi: 10.1038/nnano.2012.167

- Rasamiravaka, T., Labtani, Q., Duez, P., and El Jaziri, M. (2015). The formation of biofilms by *Pseudomonas aeruginosa*: a review of the natural and synthetic compounds interfering with control mechanisms. *Biomed Res. Int.* 2015:759348. doi: 10.1155/2015/759348
- Rosenbluth, M. J., Lam, W. A., and Fletcher, D. A. (2006). Force microscopy of nonadherent cells: a comparison of leukemia cell deformability. *Biophys. J.* 90, 2994–3003. doi: 10.1529/biophysj.105.067496
- Safari, A., Habimana, O., Allen, A., and Casey, E. (2014). The significance of calcium ions on *Pseudomonas fluorescens* biofilms - a structural and mechanical study. *Biofouling* 30, 859–869. doi: 10.1080/08927014.2014.938648
- Schembri, M. A., Dalsgaard, D., and Klemm, P. (2004). Capsule shields the function of short bacterial adhesins. *J. Bact.* 186, 1249–1257. doi: 10.1128/JB.186.5.1249-1257.2004
- Schwartz, K., Syed, A. K., Stephenson, R. E., Rickard, A. H., and Boles, B. R. (2012). Functional amyloids composed of phenol soluble modulins stabilize *staphylococcus aureus* biofilms. *PLoS Pathog.* 8:e1002744. doi: 10.1371/journal.ppat.1002744
- Seale, R. B., Flint, S. H., McQuillan, A. J., and Bremer, P. J. (2008). Recovery of spores from thermophilic dairy bacilli and effects of their surface characteristics on attachment to different surfaces. *Appl. Environ. Microbiol.* 74, 731–737. doi: 10.1128/AEM.01725-07
- Seviour, T., Hansen, S. H., Yang, L., Yau, Y. H., Wang, V. B., Stenvang, M., et al. (2015). Functional amyloids keep quorum sensing molecules in check. *J. Biol. Chem.* 290, 6457–6469. doi: 10.1074/jbc.M114.613810
- Simon, R., Priefer, U., and Pühler, A. N. B. (1983). A broad host range mobilization system for *in vivo* genetic engineering: transposon mutagenesis in gram negative bacteria. *Nat. Biotech.* 1, 784–791. doi: 10.1038/nbt1183-784
- Sleytr, U. B., Messner, P., Pum, D., and Sara, M. (eds.). (1988). *Crystalline Bacterial Cell Surface Layers*. Heidelberg: Springer.
- Sneddon, I. N. (1965). The relation between load and penetration in the axisymmetric boussinesq problem for a punch of arbitrary profile. *Int. J. Eng. Sci.* 3, 47–57. doi: 10.1016/0020-7225(65)90019-4
- Sokolov, I., Iyer, S., Subba-Rao, V., Gaikwad, R. M., and Woodworth, C. D. (2007). Detection of surface brush on biological cells *in vitro* with atomic force microscopy. *Appl. Phys. Lett.* 91:023902. doi: 10.1063/1.2757104
- Stoodley, P., Jacobsen, A., Dunsmore, B. C., Purevdorj, B., Wilson, S., Lappin-Scott, H. M., et al. (2001). The influence of fluid shear and AIC3 on the material properties of *Pseudomonas aeruginosa* PAO1 and *Desulfovibrio* sp. EX265 biofilms. *Water Sci. Technol.* 43, 113–120.
- Stoodley, P., Lewandowski, Z., Boyle, J. D., and Lappin-Scott, H. M. (1999). Structural deformation of bacterial biofilms caused by short-term fluctuations in fluid shear: an *in situ* investigation of biofilm rheology. *Biotechnol. Bioeng.* 65, 83–92.
- Sutherland, I. (2001). Biofilm exopolysaccharides: a strong and sticky framework. *Microbiology* 147, 3–9. doi: 10.1099/00221287-147-1-3
- Tirola, M., Lahtinen, T., Vuento, M., and Oker-Blom, C. (2009). Early succession of bacterial biofilms in paper machines. *J. Ind. Microbiol. Biotechnol.* 36, 929–937. doi: 10.1007/s10295-009-0571-6
- Tolker-Nielsen, T., Brinch, U. C., Ragas, P. C., Andersen, J. B., Jacobsen, C. S., and Molin, S. (2000). Development and dynamics of *Pseudomonas* sp. biofilms. *J. Bacteriol.* 182, 6482–6489. doi: 10.1128/JB.182.22.6482-6489.2000
- Vasseur, P., Vallet-Gely, I., Soscia, C., Genin, S., and Filloux, A. (2005). The pel genes of the *Pseudomonas aeruginosa* PAK strain are involved at early and late stages of biofilm formation. *Microbiology* 151, 985–997. doi: 10.1099/mic.0.27410-0
- Vidal, O., Longin, R., Prigent-Combaret, C., Dorel, C., Hooreman, M., and Lejeune, P. (1998). Isolation of an *Escherichia coli* K-12 mutant strain able to form biofilms on inert surfaces: involvement of a new ompR allele that increases curli expression. *J. Bacteriol.* 180, 2442–2449.
- Whitchurch, C. B., Tolker-Nielsen, T., Ragas, P. C., and Mattick, J. S. (2002). Extracellular DNA required for bacterial biofilm formation. *Science* 295, 1487. doi: 10.1126/science.295.5559.1487

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Zeng, Vad, Dueholm, Christiansen, Nilsson, Tolker-Nielsen, Nielsen, Meyer and Otzen. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Identification of Key Amino Acid Residues Modulating Intracellular and *In vitro* Microcin E492 Amyloid Formation

Paulina Aguilera<sup>1</sup>, Andrés Marcoleta<sup>1\*</sup>, Pablo Lobos-Ruiz<sup>1</sup>, Rocío Arranz<sup>2</sup>, José M. Valpuesta<sup>2</sup>, Octavio Monasterio<sup>1</sup> and Rosalba Lagos<sup>1\*</sup>

<sup>1</sup> Laboratorio de Biología Estructural y Molecular, Departamento de Biología, Facultad de Ciencias, Universidad de Chile, Santiago, Chile, <sup>2</sup> Centro Nacional de Biotecnología, Consejo Superior de Investigaciones Científicas, Madrid, Spain

## OPEN ACCESS

### Edited by:

Salvador Ventura,  
Universitat Autònoma de Barcelona,  
Spain

### Reviewed by:

Xuefeng Lu,  
Qingdao Institute of Bioenergy  
and Bioprocess Technology –  
Chinese Academy of Sciences, China  
Kürşad Turgay,  
Leibniz Universität Hannover,  
Germany  
Filip Meersman,  
University College London, UK

### \*Correspondence:

Rosalba Lagos  
rolagos@uchile.cl;  
Andrés Marcoleta  
amarcoleta@uchile.cl

### Specialty section:

This article was submitted to  
Microbial Physiology and Metabolism,  
a section of the journal  
Frontiers in Microbiology

**Received:** 09 October 2015

**Accepted:** 11 January 2016

**Published:** 28 January 2016

### Citation:

Aguilera P, Marcoleta A, Lobos-Ruiz P,  
Arranz R, Valpuesta JM,  
Monasterio O and Lagos R (2016)  
Identification of Key Amino Acid  
Residues Modulating Intracellular  
and *In vitro* Microcin E492 Amyloid  
Formation. *Front. Microbiol.* 7:35.  
doi: 10.3389/fmicb.2016.00035

Microcin E492 (MccE492) is a pore-forming bacteriocin produced and exported by *Klebsiella pneumoniae* RYC492. Besides its antibacterial activity, excreted MccE492 can form amyloid fibrils *in vivo* as well as *in vitro*. It has been proposed that bacterial amyloids can be functional playing a biological role, and in the particular case of MccE492 it would control the antibacterial activity. MccE492 amyloid fibril's morphology and formation kinetics *in vitro* have been well-characterized, however, it is not known which amino acid residues determine its amyloidogenic propensity, nor if it forms intracellular amyloid inclusions as has been reported for other bacterial amyloids. In this work we found the conditions in which MccE492 forms intracellular amyloids in *Escherichia coli* cells, that were visualized as round-shaped inclusion bodies recognized by two amyloidophilic probes, 2-4'-methylaminophenyl benzothiazole and thioflavin-S. We used this property to perform a flow cytometry-based assay to evaluate the aggregation propensity of MccE492 mutants, that were designed using an *in silico* prediction of putative aggregation hotspots. We established that the predicted amino acid residues 54–63, effectively act as a pro-amyloidogenic stretch. As in the case of other amyloidogenic proteins, this region presented two gatekeeper residues (P57 and P59), which disfavor both intracellular and *in vitro* MccE492 amyloid formation, preventing an uncontrolled aggregation. Mutants in each of these gatekeeper residues showed faster *in vitro* aggregation and bactericidal inactivation kinetics, and the two mutants were accumulated as dense amyloid inclusions in more than 80% of *E. coli* cells expressing these variants. In contrast, the MccE492 mutant lacking residues 54–63 showed a significantly lower intracellular aggregation propensity and slower *in vitro* polymerization kinetics. Electron microscopy analysis of the amyloids formed *in vitro* by these mutants revealed that, although with different efficiency, all formed fibrils morphologically similar to wild-type MccE492. The physiological implication of MccE492 intracellular amyloid formation is probably similar to the inactivation process observed for extracellular amyloids, and could be used as a mean of sequestering potentially toxic species inside the cell when this bacteriocin is produced in large amounts.

**Keywords:** microcin E492, intracellular amyloids, gatekeeper residues, protein aggregation, inclusion bodies

**Abbreviations:** BTA-1, 2-4'-methylaminophenyl benzothiazole; MccE492, microcin E492; PBS, phosphate-buffered saline; PFA, paraformaldehyde; TCA, trichloroacetic acid; ThS, thioflavin S.

## INTRODUCTION

Amyloid fibrils are highly organized protein aggregates with a unique quaternary structure named cross- $\beta$ , consisting of protein monomers assembled into intermolecular hydrogen bonded  $\beta$ -strands placed perpendicularly to the fibril axis (Chiti and Dobson, 2006; Fowler et al., 2007). The unit of the cross- $\beta$  spine is a  $\beta$ -sheet bilayer with side chains within the bilayer forming a tight “steric zipper” (Sawaya et al., 2007; Fitzpatrick et al., 2013). Amyloids share chemical properties such as specific binding of probes (Nilsson, 2004) as well as denaturation and proteolysis resistance. Traditionally, amyloid-fibrils formation has been related to neurodegenerative pathologies such as Alzheimer’s, Parkinson’s, and Huntington’s disease. However, several examples of amyloids playing a biological role have appeared during the last years. These “functional amyloids” have been described in many organisms: from bacteria and fungi, to insects, fish, and mammals (Fowler et al., 2007). The first reported example of a bacterial functional amyloid was curli, a well-studied type of extracellular amyloid fibrils produced by *Escherichia coli* and *Salmonella* that participate in biofilm formation, host cell adhesion and invasion (Chapman et al., 2002; Wang et al., 2006). Other examples include TasA produced by the gram-positive bacteria *Bacillus subtilis*, where the amyloid fibrils also participate in biofilm development (Romero et al., 2010), and the filamentous bacteria *Streptomyces coelicolor*, in which the aerial hyphae growth is mediated by amyloid structures formed by the chaplin proteins (Claessen et al., 2003).

Most examples of bacterial amyloids are extracellular. One exceptional and interesting case is RepA, the replication initiator protein of *Pseudomonas* plasmid pPS10. In the presence of short dsDNA oligonucleotides, the RepA-WH1 domain forms amyloid fibrils *in vitro* (Giraldo, 2007). The expression in *E. coli* of the hyper-amyloidogenic domain variant A31V fused to a red fluorescent protein, led to the accumulation of amyloid inclusions into the cytosol. Remarkably, these amyloid inclusions were transmitted vertically to the progeny and bacteria carrying them showed decreased cell fitness, constituting a bacterial proteinopathy (Fernández-Tresguerres et al., 2010). Regarding intracellular amyloid formation, it has been found that mammalian amyloid proteins expressed in *E. coli* form intracellular amyloid aggregates as well (Dasari et al., 2011; Espargaró et al., 2012). Moreover, when analyzing the ultrastructure of inclusion bodies formed by a growing number of proteins (even those not formally defined as amyloidogenic), it was noticed that they have an amyloid-like cross- $\beta$  structure (Carrió et al., 2005; Wang et al., 2008a). Additionally, it was demonstrated that the co-expression of the yeast amyloidogenic proteins Sup35 and New1 in *E. coli* cells leads to the formation of cytoplasmic inclusions with amyloid properties, and its formation is correlated with the propagation of the prionic/amyloid form of Sup35, termed [PSI<sup>+</sup>] (Yuan et al., 2014).

It is now evident that proteins capable of forming amyloid structures are quite diverse, and that the amyloidogenesis phenomenon is more ubiquitous than originally thought. Thus, it

has been suggested that amyloid formation is a generic property of the polypeptide chain and not a feature of a small number of proteins (Stefani and Dobson, 2003). Moreover, it has been shown that the propensity to form amyloid fibrils depends on specific regions or residues within the polypeptide chain. The “pro-amyloidogenic regions” or “aggregation hotspots” initiate or favor the conformational transition into the cross- $\beta$  assembly and often consist of 5–15 adjacent hydrophobic residues of low net charge and a high tendency to form  $\beta$ -strands (Beerten et al., 2012). In addition, aggregation propensity can also be modulated by a group of residues called *gatekeepers*, located near or within the pro-amyloidogenic region. These residues prevent an uncontrolled aggregation disfavoring  $\beta$ -structure formation, and are usually charged residues or prolines (Rousseau et al., 2006; Beerten et al., 2012).

Microcin E492 is a low-molecular-weight channel-forming bacteriocin produced by *Klebsiella pneumoniae* RYC492 (de Lorenzo, 1984; Lagos et al., 1993, 2009). It is found in two forms: unmodified (7,887 Da) and post-translationally modified in its C-terminal end by the covalent linkage of glycosylated salmochelin derivatives of different molecular masses (Thomas et al., 2004). This modification is required for antibacterial activity, since toxin uptake by the target cells depends on the recognition of the salmochelin-like moiety by the outer membrane catecholate siderophore receptors FepA, Fiu, and Cir (Strahsburger et al., 2005). Once in the periplasm, MccE492 exerts its toxic activity through the formation of pores in the cytoplasmic membrane and the consequent membrane potential dissipation (de Lorenzo and Pugsley, 1985; Lagos et al., 1993). One salient feature of MccE492 is its ability to form amyloid fibrils, which was observed *in vitro* and *in vivo* in the extracellular space, and was associated with the loss of antibacterial activity (Bieler et al., 2005; Marcoleta et al., 2013). Hence, amyloid formation was proposed as a mechanism of antibacterial inactivation. Even though MccE492 amyloid fibril’s morphology and formation kinetics *in vitro* have been well-characterized (Bieler et al., 2005; Arranz et al., 2012; Marcoleta et al., 2013), it is not known if MccE492, as in the case of RepA, forms amyloid *in vivo* inside the cell. Besides the influence of the post-translational modification on retarding the kinetics of amyloid formation (Marcoleta et al., 2013), there are no studies on how the primary structure, specifically which amino acid residues, determines its amyloidogenic propensity. In this work we report the conditions in which MccE492 forms intracellular amyloid inclusions. We used this phenomenon to perform a flow cytometry-based screen for MccE492 mutants with altered aggregation propensity, establishing which regions and residues act as pro-amyloidogenic regions or as gatekeepers, promoting or disfavoring both intracellular and *in vitro* amyloid formation. Additionally, our results suggest that there is a factor encoded in the MccE492 genetic cluster that may act promoting intracellular MccE492 amyloidogenesis. The promotion of cytoplasmic MccE492 amyloid formation suggests that this phenomenon may have a physiological implication as a mechanism of capturing and inactivating an excess of toxic species inside the bacteria.



## MATERIALS AND METHODS

### Bacterial Strain and Plasmids

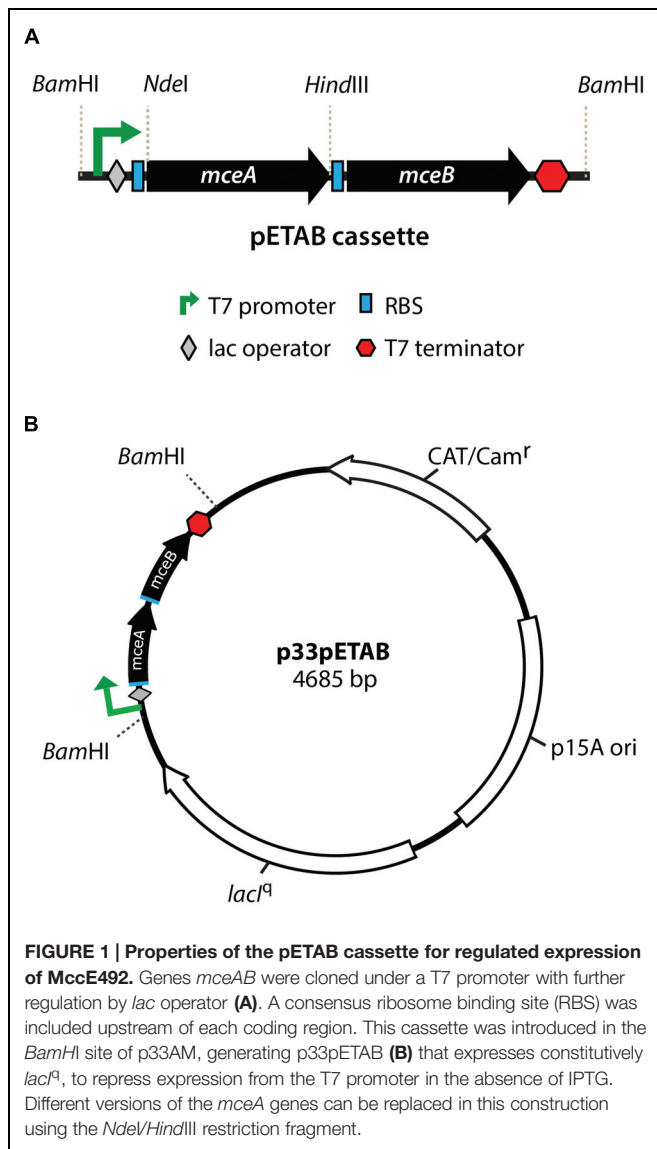
The bacterial strain and plasmids used in this work are shown in **Table 1**. A vector for tight expression of MccE492 and mutagenized derivatives was designed. This construct harbors both *mceA* (MccE492) and *mceB* (immunity protein) genes under the control of a T7 promoter, a *lac* operator and a T7 terminator (pETAB cassette, **Figure 1A**). The expression of the immunity is required since MccE492 expression is toxic in the absence of this protein (Bieler et al., 2006). It is important to note that naturally, *mceB* and *mceA* (in this order) form a single transcriptional unit where the last 23 nucleotides of *mceB* and the first 23 nucleotides of *mceA* overlap. Results from our laboratory indicate that there is an internal promoter inside the *mceB* coding region that contributes to the expression of *mceA*. For this reason we decided to invert the order of the genes in the cassette, and to insert a consensus ribosome binding site to each coding region in order to achieve a comparable expression of both genes, and to avoid leaky transcription of *mceA* from the *mceB* internal promoter. Additionally, to allow the easy replacement of the *mceA* gene by distinct variants and the cloning of the pETAB cassette, *NdeI/HindIII* and *BamHI* sites were conveniently included. pETAB was cloned in the p33AM plasmid backbone generating p33pETAB, which harbors a copy of the *lacI<sup>q</sup>* repressor gene and a p15A origin (**Figure 1B**).

Two different plasmids compatible with the p33pETAB system, carrying the whole or part of the MccE492 production cluster were used in this work: pMccE492 and np220. pMccE492 comprises all the necessary components for active microcin production, i.e., the structural and immunity genes and the genes involved in post-translational modification, export, regulation, and others of unknown function, in the same disposition as in *K. pneumoniae* RYC492 chromosome, as it is depicted in Lagos et al. (2009). The expression of this plasmid results in the production of active MccE492. Meanwhile, np220 is a plasmid used in our lab as a background to produce and export post-translationally modified MccE492 (Mercado et al., 2008) expressed from a compatible plasmid, such as p33pETAB and pBAML. np220 encodes all the components necessary for MccE492 activity, such as immunity, export and post-translational modification, with a gene disposition that it is not exactly the same than in *K. pneumoniae* RYC492. In this plasmid the structural gene *mceA* is interrupted by a Tn5 insertion, thus np220 by itself is unable to produce MccE492 (Lagos et al., 2001). Another plasmid used in this work is pBAML, which only harbors the structural *mceA* and the immunity *mceB* genes with their natural promoter and in the configuration described above for *K. pneumoniae* RYC492. The expression of pBAML as well as p33pETAB in the absence of np220 results in the production of unmodified and non-exported MccE492.

**TABLE 1 | Bacterial strain and plasmids used in this work.**

Bacterial strain or plasmid	Relevant genotype or features	Source
<b><i>Escherichia coli</i> strain</b>		
BL21-AI	F- <i>ompT hsdSB (rB-mB-) gal dcm araB::T7RNAP tetA</i>	Invitrogen
<b>Plasmids</b>		
pHC79	ColE1 cosmid derivative. Amp <sup>r</sup> Tet <sup>r</sup>	Hohn and Collins, 1980
pMccE492	Whole MccE492 production system, pHC79 derived. Amp <sup>r</sup>	González, 2011
np220	Whole MccE492 production system, but <i>mceA::Tn5</i> . Amp <sup>r</sup> Kan <sup>r</sup>	Mercado et al., 2008
pBAML	MccE492 structural ( <i>mceA</i> ) and immunity ( <i>mceB</i> ) genes. Cam <sup>r</sup>	Lagos R. laboratory collection
p33AM	<i>lacI<sup>q</sup></i> . Allows the regulation of the pET system. P15A origin, Cam <sup>r</sup>	Lagos R. laboratory collection
p33pETAB	p33AM derivative. Contains the pETAB cassette ( <i>mceA</i> and <i>mceB</i> under the control of the T7 promoter) and <i>lacI<sup>q</sup></i> . Cam <sup>r</sup>	This work
N16A	p33pETAB derivative. <i>mceA</i> was mutagenized replacing asparagine 16 by alanine. Cam <sup>r</sup>	This work
P26A	p33pETAB derivative. <i>mceA</i> was mutagenized replacing proline 26 by alanine. Cam <sup>r</sup>	This work
Δ18–35	p33pETAB derivative. <i>mceA</i> was mutagenized deleting amino acids 18–35. Cam <sup>r</sup>	This work
P53A	p33pETAB derivative. <i>mceA</i> was mutagenized replacing proline 53 by alanine. Cam <sup>r</sup>	This work
N55A	p33pETAB derivative. <i>mceA</i> was mutagenized replacing asparagine 55 by alanine. Cam <sup>r</sup>	This work
P57A	p33pETAB derivative. <i>mceA</i> was mutagenized replacing proline 57 by alanine. Cam <sup>r</sup>	This work
P59A	p33pETAB derivative. <i>mceA</i> was mutagenized replacing proline 59 by alanine. Cam <sup>r</sup>	This work
P64A	p33pETAB derivative. <i>mceA</i> was mutagenized replacing proline 64 by alanine. Cam <sup>r</sup>	This work
Δ54–63	p33pETAB derivative. <i>mceA</i> was mutagenized deleting amino acids 54–63. Cam <sup>r</sup>	This work





## Growth Conditions

Bacterial growth was performed incubating with shaking (180–220 rpm) at 37°C. For confocal microscopy analysis cells carrying MccE492 producing systems were grown in M9 minimal medium supplemented with citrate 0.2% and glucose 0.2% for 48 h (late stationary phase). For intracellular aggregation propensity determinations, cells carrying p33pETAB variants with or without np220 (this plasmid allows the modification and exportation of MccE492) were grown in LB medium. Antibiotics were used in the following concentrations: ampicillin (Amp) 100 µg/ml, kanamycin (Kan) 50 µg/ml, and chloramphenicol (Cam) 50 µg/ml.

## MccE492 Induction Assay

MccE492 induction assay in the p33pETAB/BL21-AI system was performed as follows: 15 ml of LB (with the corresponding antibiotic) were inoculated with 1:20 dilution of fresh overnight

*E. coli* BL21-AI p33pETAB culture. The cells were grown at 37°C with shaking at 180 rpm until early exponential phase was reached. At this point, 10 ml of culture were split in two (5 ml each): one aliquot was induced with 0.2% arabinose and 1 mM IPTG, and the other was left untreated. After 6 h, aliquots of induced and non-induced cells were collected to continue with the protein extraction protocol or with the fixation and staining procedure.

## Total Protein Extraction

For total protein extraction, cells were centrifuged at  $12,900 \times g$  for 10 min. The pellet was re-suspended in PBS buffer, incubated at 95°C for 20 min and centrifuged at  $12,900 \times g$  for 15 min at 4°C. 400 µl of the supernatant were collected and 100 µl of cold TCA were added. The mixture was incubated 10 min at 4°C and then centrifuged at  $12,900 \times g$  for 15 min at 4°C. The supernatant was discarded and the pellet was washed twice with 200 µl of chilled acetone. Finally, acetone was evaporated and the pellet was stored at –80°C.

## SDS-PAGE and Immunoblotting

Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) was performed as described by Schägger and von Jagow (1987). Nitrocellulose membranes (Millipore) were used for immunoblot transfer (1 h, 350 mA, using chilled 25 mM Tris-HCl, 190 mM glycine, 20% methanol, as the transfer buffer). MccE492 was detected with a polyclonal antibody prepared in rabbit against the last 20 amino acids of the protein (antisera dilution, 1:1,000) and with a goat anti-rabbit alkaline phosphatase-conjugated secondary antibody (dilution 1:5,000). The alkaline phosphatase colorimetric reaction was performed as described by Marcoleta et al. (2013). The membrane was washed with FAL buffer (100 mM Tris-HCl [pH 9.5], 100 mM NaCl, 5 mM MgCl<sub>2</sub>) and incubated in 10 ml of a mixture of BCIP (5-bromo-4-chloro-3'-indolylphosphate *p*-toluidine salt) and NBT (nitroblue tetrazolium chloride; 0.3 and 0.15 mg/ml in FAL buffer, respectively), until an optimal signal was observed.

## Cell Fixation and Staining for Confocal Microscopy

Five hundred µl of bacterial culture were washed twice with the same volume of PBS buffer by centrifuging at  $1,100 \times g$  for 4 min at room temperature, fixed by suspending in 250 µl of 4% PFA in PBS and incubating for 30 min at room temperature. BTA-1 staining was performed as described (Fernández-Tresguerres et al., 2010). After fixation, cells were washed twice with 250 µl of PBS buffer by centrifuging at  $1,100 \times g$  for 4 min at room temperature, re-suspended in 1 mM BTA-1 prepared in 100% ethanol, and incubated for 30 min at room temperature. Finally, cells were washed twice with PBS. ThS staining was carried out based on the protocol of Espargaró et al. (2012) with some modifications: 500 µl of bacterial culture were centrifuged at  $1,100 \times g$  for 4 min at room temperature, the supernatant was discarded, and the bacterial pellet was washed twice with 500 µl of PBS buffer. Cells were suspended in 250 µl of 0.05% (w/v) ThS

in 12.5% ethanol and incubated for 1 h at room temperature. We found that ThS staining in the presence of 12.5% ethanol, instead of PBS as originally described, leads to a smaller number of false negative cells. After staining, cells were washed three times with 250  $\mu$ l of PBS buffer. Stained cells were mounted over a glass slide covered by layer of 1% agarose in PBS.

## Confocal Microscopy

Microscopic observations were performed using an LSM 710 (Zeiss) confocal microscope, with a 63x/NA 1.40 oil immersion objective. ThS fluorescence was excited using a 488 nm argon laser and the emission was registered in a range from 410 to 520 nm. BTA-1 fluorescence was excited with a 405 nm laser diode and the emission was registered between 493 and 552 nm. Images were digitally captured with ZEN 2012 and analyzed with ImageJ software (Schneider et al., 2012).

## Flow Cytometry

Thioflavin-S staining for flow cytometry analysis was carried out in the same way as described for confocal microscopy, but starting with unfixed cells. Flow cytometry measurements were performed using a BD FACSCanto™ II flow cytometer. Cells stained with ThS were first gated by forward scatter (FSC) and side scatter (SSC) signals, and then analyzed for ThS fluorescence by exciting at 405 nm and registering the emission at 510/550 nm.

## MccE492 Purification

MccE492 (wild-type or mutant) was purified from culture supernatants of *E. coli* BL21-AI cells carrying each p33pETAB variant along with np220, to allow MccE492 processing and export. Briefly, 4 L of M9 medium supplemented with 0.2% citrate and 0.1% glucose were inoculated with 1:1000 dilution of fresh overnight culture and grown at 37°C with shaking at 180 rpm for 6 h (until exponential phase). At this point, 0.2% arabinose and 1 mM IPTG were added. After 12–14 h, the supernatant was collected by centrifugation and filtered through a 0.22  $\mu$ m polyethersulfone membrane. The cell-free medium was incubated with 25 g of previously ACN-activated Bondapak C18 resin (Waters) at 4°C for 2 h. The resin was filtered by negative pressure through a Buchner funnel, washed with 200 ml of 40% methanol, then with 200 ml of 25% ACN, and finally eluted with a 30–100% ACN stepwise gradient. MccE492-enriched fractions were dialyzed twice for 2 h against 40 volumes of 5 mM Tris (pH 8.5) and then lyophilized and stored at –20°C.

## In vitro MccE492 Amyloid Formation followed by Congo Red Binding Assay

An appropriate amount of lyophilized MccE492 powder was dissolved in 5 mM Tris (pH 8.5), and centrifuged for 30 min at 16,000  $\times$  g to eliminate preformed aggregates. The supernatant was diluted to 0.4 mg/ml in the aggregation buffer (100 mM PIPES-NaOH, 0.5 M NaCl [pH 6.5]) and incubated with agitation (800 rpm) during the entire assay. Amyloid formation was quantified by the diminution of free Congo red using the following procedure: at the indicated times, MccE492 samples were incubated at 37°C for 15 min with 33  $\mu$ M Congo red and

centrifuged at 16,000  $\times$  g for 40 min at 4°C. The supernatant absorbance was registered at 490 nm, and the free Congo red fraction was determined as the ratio between the absorbance registered at each time and the absorbance at time zero.

## Determination of Soluble MccE492 During the Aggregation Curve

To visualize soluble MccE492 during the aggregation assays, aliquots of the samples taken at different times were collected and centrifuged at 16,000  $\times$  g for 40 min at 4°C. The supernatant was recovered and the remaining soluble protein was detected by SDS-PAGE and immunoblotting.

## MccE492 Antibacterial Activity Determination

MccE492 activity was determined by the critical dilution method (Mayr-Harting et al., 1972). At the indicated times, aliquots from the aggregation assays were collected and serially diluted in sterile nanopure water. Three  $\mu$ l of each dilution were seeded onto a lawn of a sensitive *E. coli* strain, prepared by mixing 0.3 ml of the *E. coli* culture with 3 ml soft agar and overlaying the resulting mixture onto LB plates. MccE492 antibacterial activity was detected by the formation of growth inhibition halos, and the activity was expressed in arbitrary units based on the highest dilution in which a halo was observed.

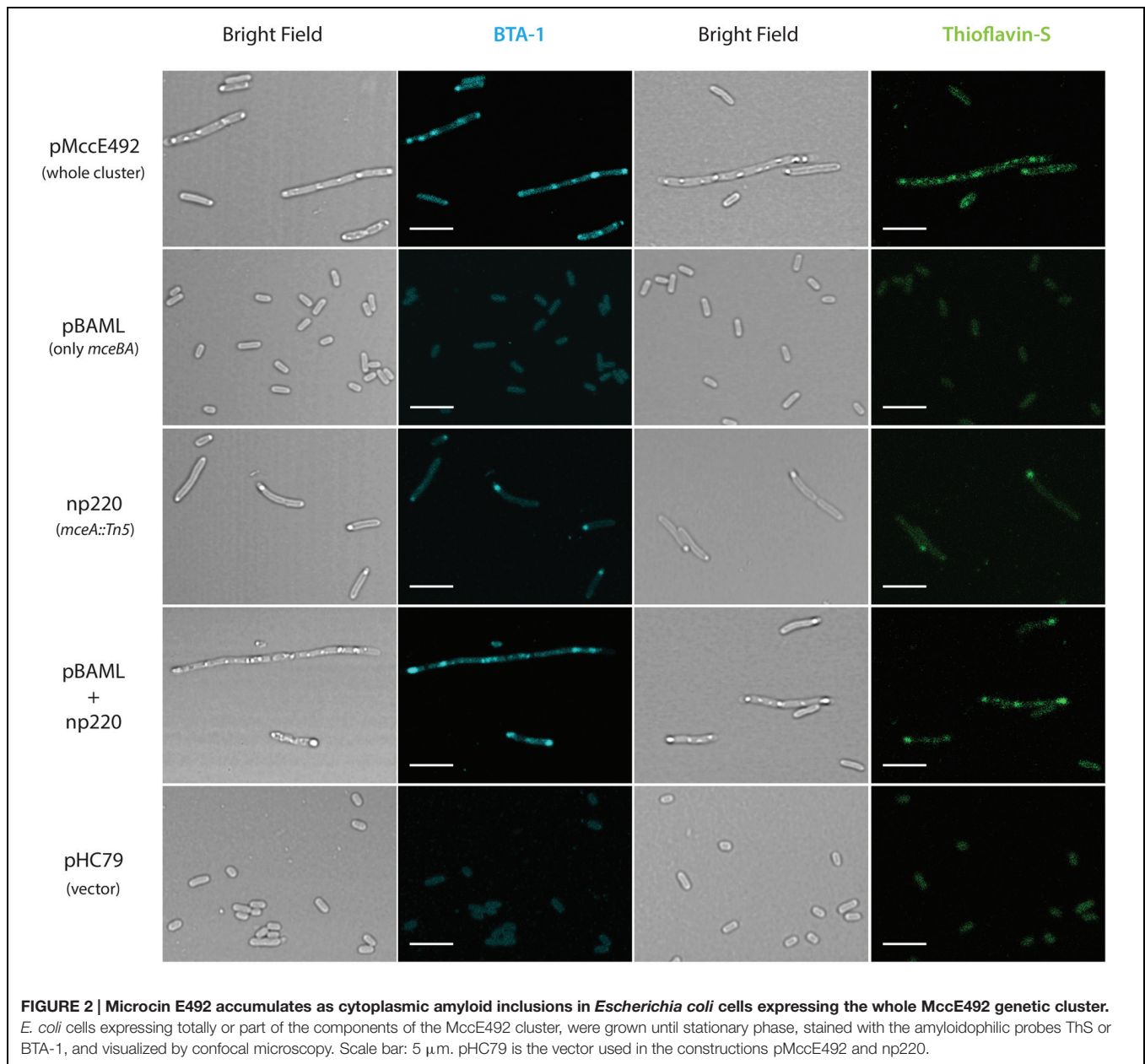
## Electron Microscopy

Samples from the MccE492 aggregation assays were placed onto 300-square-mesh copper-rhodium grids coated with carbon and negatively stained with 2% uranyl acetate. Micrographs were taken in a JEOL 1200EX microscope with a tungsten filament operated at 80 kV and with a 50,000X magnification.

## RESULTS

### Microcin E492 forms Intracellular Amyloid Aggregates in *E. coli* Cells

It has been shown that amyloidogenic proteins from different origins, when expressed in *E. coli*, are accumulated as cytoplasmic inclusions of amyloid nature (Fernández-Tresguerres et al., 2010; Dasari et al., 2011; Espargaró et al., 2012). Extracellular MccE492 amyloid formation has been observed under several conditions, but it is not known if there is amyloid aggregation in the cytoplasm. To investigate this possibility, *E. coli* cells carrying three different constructions expressing the whole, or part of the genetic cluster encoding for the production of active MccE492 were grown in M9 minimal medium at 37°C until late-stationary phase. After harvested, cells were fixed with 4% PFA, stained with the probes BTA-1 and ThS that specifically bind to amyloid aggregates, and visualized by confocal microscopy (Figure 2). Cells carrying pMccE492, a plasmid that encodes for the whole MccE492 gene cluster that includes not only the structural and immunity genes but all those necessary for export and maturation, contained a variable number of cytoplasmic inclusions that were visible under bright-field imaging, and that



were recognized by both amyloidophilic fluorescent probes as round-shaped foci. These inclusions of amyloid nature were heterogeneous in form and size, and were located at distinct positions of the major axis of the cells (not only at the poles), as reported previously for the hyper-amyloidogenic variant of the WH1 domain of RepA protein from *P. aeruginosa* (Fernández-Tresguerres et al., 2010). Unexpectedly, cells carrying pBAML (comprising only the MccE492 and its immunity genes) that do not export MccE492, did not accumulate inclusions, and behaved as *E. coli* cells carrying the pHC79 control vector. Cells carrying np220, a plasmid expressing the whole gene cluster with the exception of the structural gene of MccE492 that is interrupted by Tn5, showed a few cells with single polar inclusions that were recognized by both amyloidophilic probes, indicating that most

of the inclusions observed in the wild type construct are produced by the expression of the *mceA* gene. This was corroborated upon complementation of the mutant np220 with the *mceA* gene from pBAML. Cells carrying both plasmids restored the multiple amyloid inclusions phenotype caused by pMccE492 (Figure 2).

Interestingly, the accumulation of MccE492 amyloid inclusions is accompanied with a polymorphism in the cell length. There is a minority of cells that form filaments with a length equivalent to the sum of more than 15 normal cells (Supplementary Figure S1) that can harbor up to 20 amyloid foci. This suggests that massive amyloid accumulation could somehow interfere with the septation process, although it remains unclear why this occurs only in a minor fraction of the bacterial population.



## Identification of Amino Acid Residues and Regions Potentially Involved in MccE492 Amyloid Formation

One central aspect on the study of amyloid formation is how the primary structure of a protein influences the fibril formation process, and the identification of regions and residues that may act promoting or disfavoring its occurrence. Based on the evaluation of distinct physicochemical properties of its residues, several algorithms have been developed to predict and identify aggregation hotspots inside the primary structure of amyloidogenic proteins (Chiti et al., 2003; Fernández-Escamilla et al., 2004; López de la Paz and Serrano, 2004; Maurer-Stroh et al., 2010; Tsolis et al., 2013). In order to identify putative aggregation hotspots of MccE492, we analyzed its sequence with the AMYLPRED2 web tool, which employs a consensus of 11 different methods and algorithms that predict features related to the formation of amyloid fibrils (Tsolis et al., 2013). Analysis of the MccE492 protein sequence showed that the region comprising residues 57–63 of the processed peptide, especially residues 60–62, has the highest consensus being recognized as pro-amyloidogenic by six different algorithms (Figure 3). This region contains several hydrophobic residues (IPVLIG), some of which (VLI) are very commonly found in pro-amyloidogenic regions of several proteins (López de la Paz and Serrano, 2004). Also, inside and near to this region there are several proline residues (P53, P57, P59, P64), which are  $\beta$ -strands disruptors that normally act as “amyloid gatekeepers” diminishing the overall aggregation propensity of the protein (López de la Paz and Serrano, 2004; Beerten et al., 2012). An asparagine residue (N55) was found near to the predicted hydrophobic aggregation hotspot, that due to its polar nature could also act as a gatekeeper. The authors of AMYLPRED2 suggest at least a consensus of five for a reliable prediction. However, the region between residues 20–24 that is recognized only by four algorithms (Figure 3) was considered as a secondary amyloidogenic region for the following reasons: first, MccE492 region encompassing residues 16–38 presents 57% of identity and 74% of similarity with the prion PrP protein regions 111–133 (Baeza, 2003); second, this region is also hydrophobic and it is flanked by the putative gatekeeper residues asparagine (N16) and proline (P26). To test these predictions, MccE492 mutants in the residues mentioned above and deletions of the corresponding regions were constructed and its amyloidogenic properties were studied.

## Quantitation of the Proportion of Cells Carrying MccE492 Intracellular Amyloids, Evaluation of Variants with Altered Amyloidogenic Properties, and Identification of Amino Acid Residues Regulating Intracellular Amyloid Formation

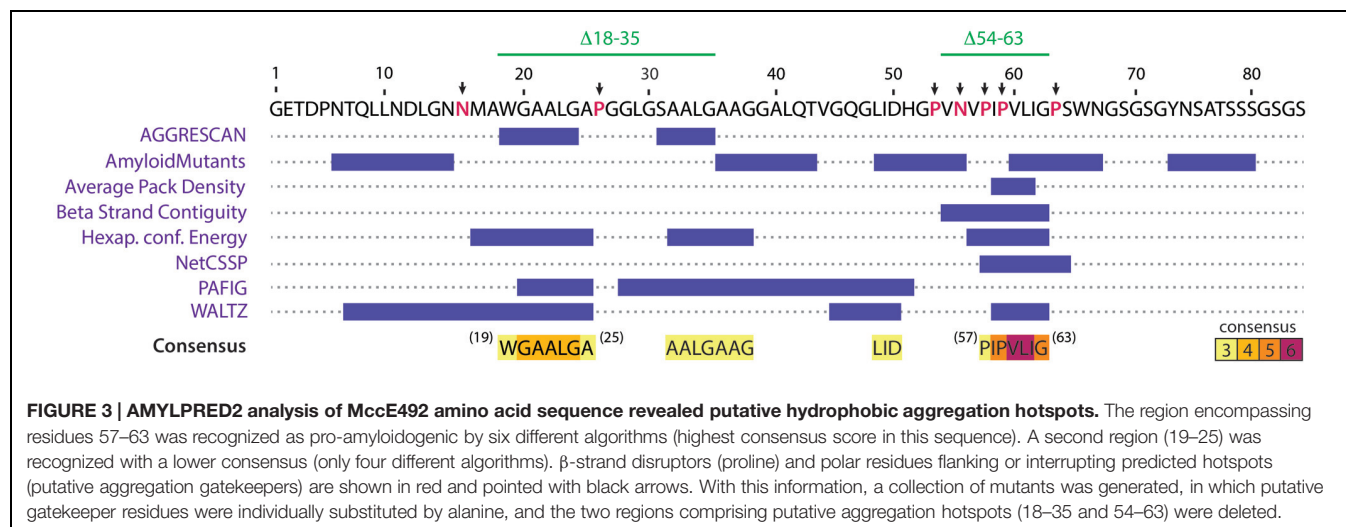
Intracellular MccE492 amyloid formation is a phenomenon that can be exploited to implement a fast and simple method for the screening of MccE492 mutants with altered amyloid-formation properties. A comprehensive analysis of these mutants

should allow the identification of protein regions controlling the aggregation propensity, providing new insights of how MccE492 amyloid formation occurs. Based on an assay described by Espargaró et al. (2012), we used ThS-staining and flow cytometry analysis to quantitate the proportion of cells harboring cytoplasmic amyloid inclusions upon expression of distinct MccE492 variants. In this way, we were able to measure and compare the amyloid formation propensity of each variant and consequently, to evaluate the relevance of the mutated residue or region in the aggregation process. As a first step, the assay setup required the design and construction of a tightly regulated expression system, which allowed a comparable expression of each MccE492 variant when induced for a defined time lapse. For this purpose, we synthesized the pETAB expression cassette (Figure 1, see Materials and Methods) that was ligated to the p33AM plasmid generating p33pETAB. From here, *mceA* gene can be easily replaced by the variants. These constructions were transformed into *E. coli* BL21-AI cells, which expresses the T7 polymerase only after induction with arabinose, therefore MccE492 expression from this system requires the concomitant action of IPTG and arabinose as inducers.

To ascertain that this regulated expression system was working as expected, we tested six random clones of the wild-type form that were induced by incubating them with arabinose and IPTG during 4 h. SDS-PAGE-Immunoblot of total protein extracts showed a prominent band corresponding to MccE492 only in the induced samples (Supplementary Figure S2). In addition, the antibacterial activity of the expressed MccE492 was assessed by co-transformation of the *E. coli* BL21-AI p33pETAB strain with the compatible plasmid np220 (mutated in *mceA*), that provides all the elements required for MccE492 maturation and export. The production of functional active MccE492 of six clones was detected as the presence of growth-inhibition halos over a layer of sensitive bacteria (Supplementary Figure S2). The toxic activity was notably higher in presence of the inducers.

Based on the information provided by the *in silico* prediction, we designed a collection of MccE492 variants with substitutions or deletions of specific amino acids. We hypothesized that if the predicted regions act as pro-amyloidogenic stretches, deletion of these regions should diminish aggregation propensity, while substitution of the corresponding putative gatekeeper residues should increase it. Based on this, seven putative gatekeeper residues were individually substituted by alanine (Figure 3, black arrows) and two deletion mutants,  $\Delta$ 18–35 and  $\Delta$ 54–63, were generated. All these variants were cloned in p33pETAB and transformed in *E. coli* BL21-AI cells. To evaluate and compare the intracellular aggregation propensity of each variant, their expression was induced for 6 h, and then the cells were stained with ThS and analyzed by flow cytometry. Thus, we determined the proportion of cells carrying amyloid inclusions that were recognized by the dye (Figure 4). Expression of wild-type MccE492 led to a very low proportion of cells carrying amyloid inclusions ( $\sim$ 3%), indicating that this protein by itself does not produce a significant amount of amyloid aggregation in the cytoplasm (Figure 4, pETAB). This is in agreement with the absence of inclusions observed



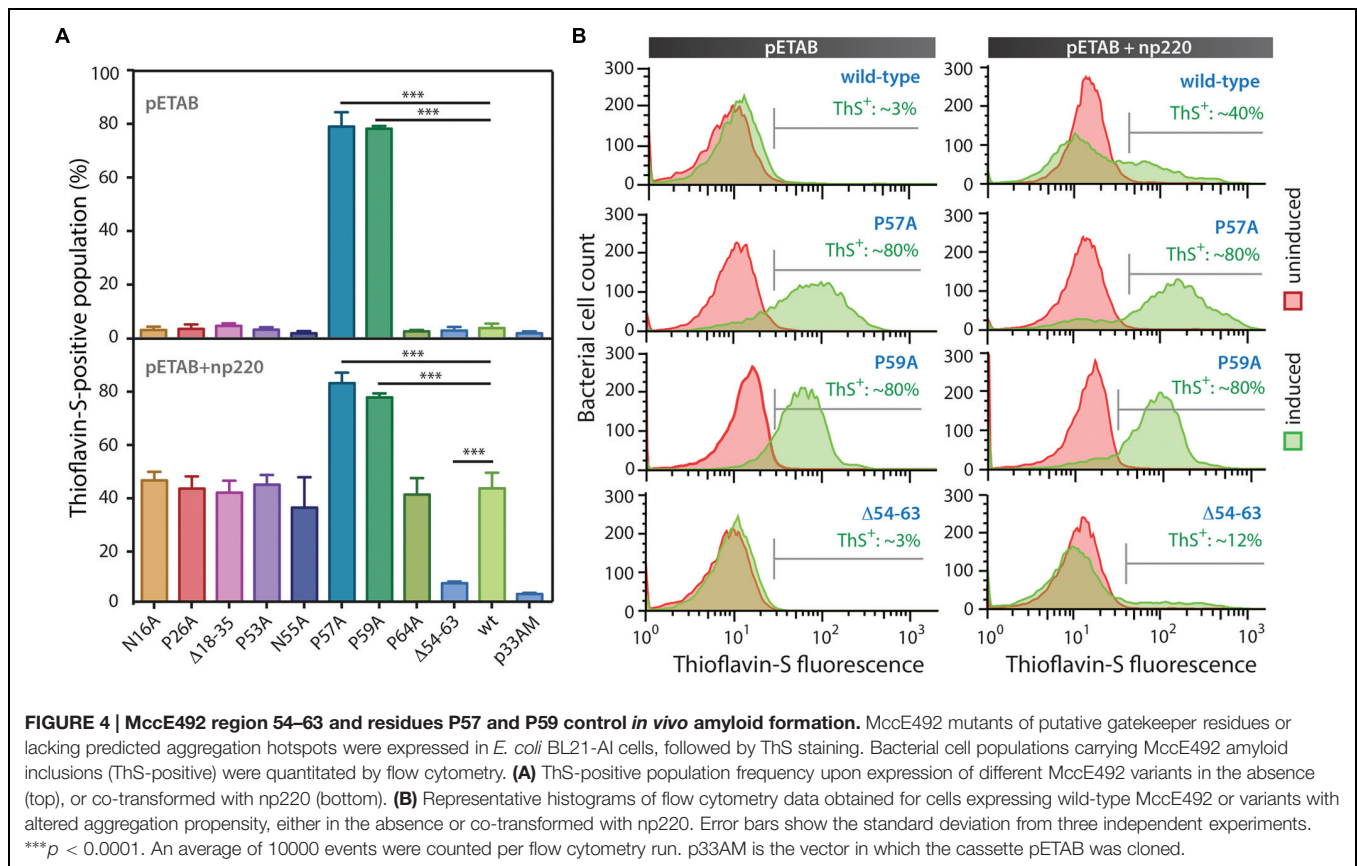


in cells carrying pBAML, expressing only the MccE492 and its immunity protein (Figure 2). Strikingly, variants P57A and P59A formed intracellular amyloids in a very high proportion of cells (~80%), indicating that substitution of those residues leads to a significant increase in the MccE492 aggregation propensity. This observation shows that both proline residues act as amyloid gatekeepers disfavoring MccE492 amyloid formation. The substitution by alanine of N16, P26, P53, N55, and P65 did not have a significant effect on intracellular amyloid formation, suggesting that these amino acids have a minor or no role controlling MccE492 amyloidogenesis.

Since wild-type protein expression led to a very small amount of ThS-positive cells, at this stage we were not able to evaluate if the deletions of the predicted aggregation hotspots affect MccE492 intracellular aggregation propensity. Considering that more inclusions were observed in cells expressing MccE492 in presence of the other components encoded in the MccE492 genetic cluster (Figure 2), the expression of the wild-type MccE492 and its variants was induced in cells carrying np220 (Figure 4, pETAB+np220). In this condition, a significantly higher proportion of cells accumulating inclusions upon expression was observed in most of the variants tested. Wild-type MccE492 intracellular amyloid was detected in nearly 40% of the cells, and an indistinguishable behavior was found for all the mutants except for P57A, P59A, and  $\Delta 54$ –63. A significant reduction of aggregation propensity was observed in the variant with the deletion of residues 54–63 ( $p < 0.0001$ ), while the deletion of residues 18–35 had no effect. On the other hand, P57A and P59A mutants maintained the high tendency to form aggregates, although co-transformation with np220 did not increase the proportion of cells accumulating amyloids. This proportion seems to be near the maximum detectable by this method, since no higher frequency was observed in any of the experiments performed, not even with longer induction times (data not shown). The intracellular amyloid formation in these assays was always dependent on the expression of MccE492, since a very low proportion of ThS-positive cells were observed in all the uninduced samples and in the induced cells carrying

just the vector p33AM, either alone or co-expressed with np220 (Figure 4A and Supplementary Figure S3). This observation suggests that the increase in the frequency of ThS-positive cells caused by np220 co-expression could be due to a synergistic effect between MccE492 and a cluster-encoded factor, and not to an independent aggregation of this putative factor.

To further investigate the effect on cell morphology and distribution of the intracellular amyloid detected by ThS-staining and flow cytometry, cells expressing wild type MccE492 from p33pETAB and the identified hypo- and hyper-amyloidogenic variants were induced for 6 h, stained with either BTA-1 or ThS, and visualized by confocal microscopy (Figure 5). In agreement with the flow cytometry measurements, cells expressing only wild-type MccE492 from p33pETAB showed none or one very weakly stained intracellular inclusion. A similar situation was observed with  $\Delta 54$ –63, where a faint homogeneous fluorescence was detected upon staining. In contrast, a variable number of inclusions were observed in cells expressing variants P57A and P59A, which showed up under bright-field imaging and were recognized by both amyloidophilic dyes as intense fluorescent foci. These foci were located in different regions/areas of the cells and had different shapes and sizes. Moreover, accumulation of MccE492 inclusions came together with some degree of cell length polymorphism, with a small proportion of cells experimenting a dramatic filamentation and harboring more than 30 amyloid foci (Supplementary Figure S4), as also seen in cells carrying the whole MccE492 production cluster (Figure 2 and Supplementary Figure S1). As expected, the hyper-amyloidogenic variants as well as the wild-type MccE492 formed amyloid inclusions in cells carrying np220 (Supplementary Figure S5). In contrast, none or one faint polar inclusion per bacteria was observed in cells carrying np220 and expressing the  $\Delta 54$ –63 mutant. Taken together, these results indicate that the region 54–63 of MccE492 has a major role controlling intracellular amyloid formation, encompassing hydrophobic residues that likely form the amyloid core, and two proline  $\beta$ -strand disruptors that act as aggregation gatekeepers. Additionally, at least in the conditions tested, a cluster-encoded



factor is likely to act promoting MccE492 intracellular amyloid formation.

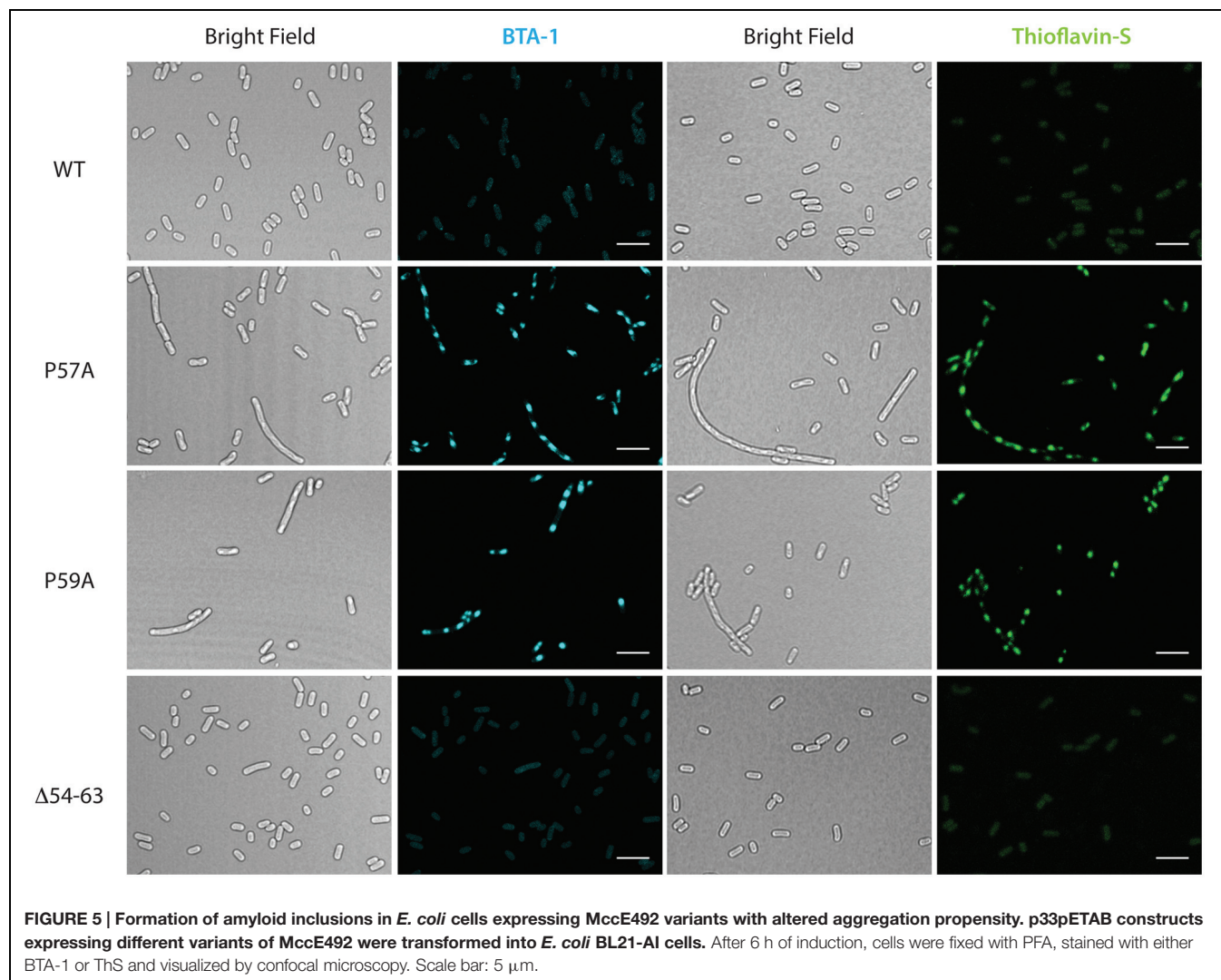
## Residues P57 and P59 of MccE492 Control *In vitro* the Kinetics of Amyloid Formation and Loss of Antibacterial Activity

To evaluate if the MccE492 aggregation hotspot and the gatekeeper residues identified in the *in vivo* assays are also relevant for the *in vitro* fibrils formation, we purified wild-type MccE492 and the variants P57A, P59A, and  $\Delta$ 54–63 from the supernatants of induced *E. coli* BL21-AI cells carrying np220 and p33pETAB. Lyophilized MccE492 samples of each variant were dissolved in aggregation buffer to a final concentration of 200  $\mu$ g/ml and incubated at 37°C with constant shaking for up to 72 h. Amyloid formation kinetics were followed for each variant, monitoring Congo red binding (**Figure 6A**) and determining soluble MccE492 by SDS-PAGE and immunoblotting (**Figure 6B**) at different times. Both Congo red binding and immunoblotting showed that most of the wild-type MccE492 remained soluble until 8 h and was completely aggregated at 24 h. In contrast, P57A and P59A variants aggregated significantly faster. Part of P57A was aggregated at the beginning of the incubation, and was practically completely aggregated at 2 h. The P59A variant also presented a reduced lag phase. It began to aggregate at 2 h, and was almost completely aggregated at

8 h, showing a more gradual aggregation kinetics than P57A. On the other hand, the mutant lacking the 54–63 pro-amyloidogenic region aggregated slower than the wild-type MccE492, and soluble protein was still detected even after 72 h of incubation.

Previous studies demonstrated that MccE492 amyloid formation causes loss of antibacterial activity (Bieler et al., 2005; Marcoleta et al., 2013). However, it is not known which MccE492 residues conform the toxic domain, nor if these residues are also important for amyloid formation. To gain further information about this, we tested the antibacterial activity of all the MccE492 mutants generated in this study (Supplementary Figure S6), stabbing colonies of *E. coli* expressing np220 and p33pETAB variants over a layer of sensitive bacteria grown in plates with or without the inducers. Growth inhibition halos were observed in wild-type MccE492, and mutants N16A, P26A, P53A, N55A, P57A, and P59A. On the other hand,  $\Delta$ 54–63,  $\Delta$ 18–23, and P64A lacked antibacterial activity, since no growth inhibition halos were observed. These results indicate that the regions encompassing residues 18–23 and 54–64 are somehow required for the antibacterial activity, while residues N16, P26, P53, N55, P57, and P59 seem to be dispensable.

To evaluate if the increased polymerization propensity of the hyper-amyloidogenic variants correlate with a faster toxin inactivation, we titrated the antibacterial activity during the aggregation assay described above using the critical dilution method (**Figure 6C**). We observed that although both P57A and



P59A mutants had antibacterial activity, the initial titer was  $\sim 20$ -fold lower than wild-type MccE492. As expected, inactivation and polymerization kinetics showed a close correlation, being the loss of the antibacterial activity of the hyperamyloidogenic variants significantly faster.

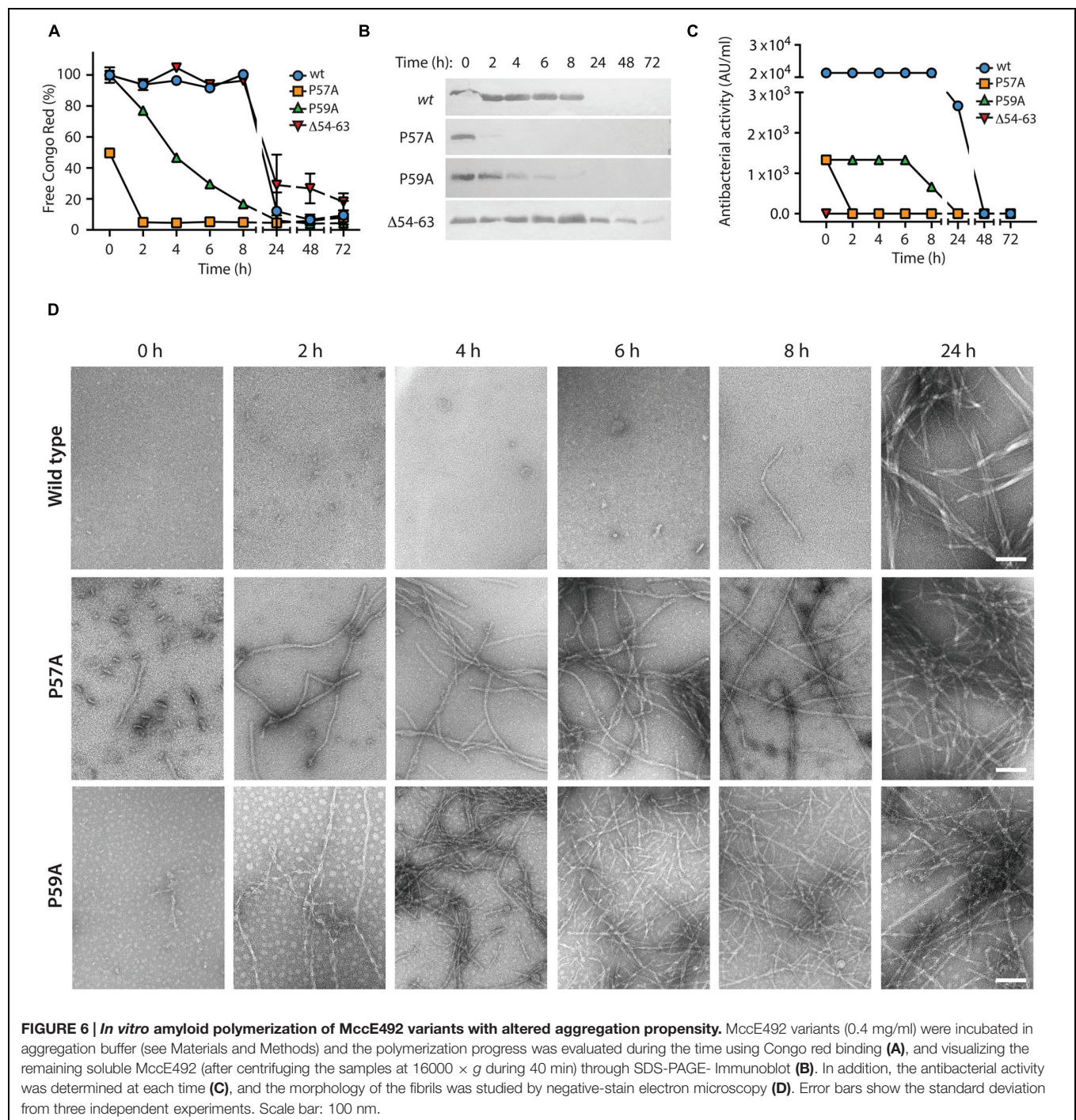
### The Fibrils formed by Wild-Type MccE492 and the Variants P57A and P59A have a Similar Morphology

*In vivo* and *in vitro* assays showed that P57A and P59A variants present a higher propensity to form fibrils with faster polymerization/inactivation kinetics. However, it is indispensable to know the morphology of the aggregates formed by the variants, to ascertain that they form amyloid fibrils. To this end, samples of the *in vitro* aggregation assay were analyzed by negative-stain TEM (Figure 6D). The results were consistent with those observed using Congo red. A few short fibrils and globular aggregates of the wild-type MccE492 were observed at 6 and 8 h, while longer mature fibrils with a width of about 10–12 nm

were observed at 24 h, as reported previously (Bieler et al., 2005; Arranz et al., 2012; Marcoleta et al., 2013). On the other hand, the amyloid fibrils formed by both proline mutants presented the same morphology than the wild-type protein, although with an accelerated kinetics of formation. For P57A, short fibrils and globular aggregates were observed even without incubation (time 0), whereas longer mature fibrils were observed at 2 h and in the remaining time points analyzed. In the case of P59A, no fibrils were detected at time 0, but long fibrils were observed after 2 h of incubation (Figure 6D).

Regarding the  $\Delta 54$ –63 mutant, TEM analysis of samples collected at the end of the aggregation assay showed that, despite its decreased aggregation propensity and slower polymerization kinetics, this mutant retains its ability to form fibrils with a similar morphology to the wild-type MccE492 (Figure 7A). This suggests that this region is not essential for amyloid formation, and probably there is a secondary aggregation hotspot that could act nucleating amyloid polymerization, although less efficiently. The ability to form amyloid fibrils was also assessed for the mutant  $\Delta 18$ –35, that *in vivo* did not present alterations in amyloid





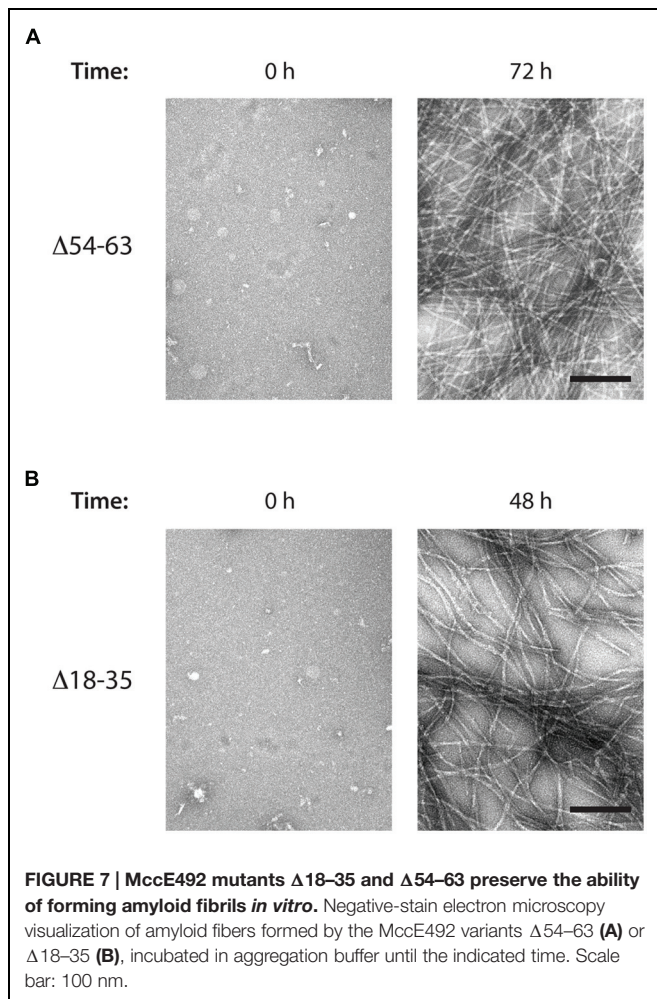
formation. As expected, typical amyloid fibrils although a little more relaxed were formed (Figure 7B).

## DISCUSSION

Currently, understanding of amyloid formation is experiencing a change of paradigm. Traditionally, amyloids have been associated with pathological manifestations of neurodegenerative diseases

in mammals, but today it is being recognized as a widespread phenomenon with examples from all domains of life (Fowler et al., 2007; Otzen, 2010; Invernizzi et al., 2012). Remarkably, microbes ended up being very successful in controlling and adapting amyloid formation to perform different functions (Hufnagel et al., 2013; Schwartz and Boles, 2013; Romero and Kolter, 2014), although they still remain vulnerable to the deleterious effects of uncontrolled aggregation (Fernández-Tresguerres et al., 2010; Beerten et al., 2012; Gasset-Rosa et al.,





2014). MccE492 represents a particular case of what has been defined as a functional amyloid, where the extracellular fibrils formation acts as a mechanism to control the antibacterial activity. In this study, we addressed two major aspects previously unexplored: first, if MccE492 could undergo amyloid aggregation in the cytoplasm of producing cells; and second, which regions or residues are involved in controlling MccE492 amyloid formation.

Our results showed that wild-type MccE492 has a low tendency to form intracellular amyloid, being barely detectable in cells over-expressing this gene. However, its occurrence was highly increased when a plasmid carrying the rest of the MccE492 cluster components was co-transformed, or when MccE492 hyperamyloidogenic variants were expressed. As seen for other amyloidogenic proteins like RepA, cytoplasmic MccE492 amyloid was accumulated as dense inclusion bodies of variable shape, size, and number per cell (Fernández-Tresguerres et al., 2010). This observation was exploited to perform a screening for MccE492 mutants with altered aggregation propensity, establishing that the region comprising residues 57–63, recognized as pro-amyloidogenic by the AMYLPRED2 prediction tool, effectively acts as an aggregation hotspot. Deletion of this region caused a significant decrease in MccE492

intracellular aggregation propensity and the slow down of *in vitro* fibril formation kinetics. The identified aggregation hotspot harbors VLI residues that likely form the  $\beta$ -strand core for amyloid assembly, as demonstrated for other amyloidogenic proteins (López de la Paz and Serrano, 2004). In addition, P57 and P59 residues that are inside this region, probably act as  $\beta$ -strand disruptors, disfavoring MccE492 amyloid formation. Indeed, alanine-substitution of each of these residues dramatically increased MccE492 aggregation propensity *in vivo*, and fastened *in vitro* fibril formation kinetics. The control of amyloid formation by gatekeeper residues seems to be a common strategy for different amyloids, since analysis of *Homo sapiens* and *E. coli* proteomes showed that the positions flanking aggregating stretches are enriched with residues such as lysine, arginine, glutamate, aspartate and proline. Strikingly, 90% of the 26,000 pro-amyloidogenic segments found in the *E. coli* proteome have at least one of these five residues at the first position on either side of the segment (Rousseau et al., 2006). Moreover, gatekeeper residues can be redundant, i.e., they are present several times in the same protein. Thus, aggregating regions from key human proteins such as p53 or huntingtin protein are among the most extensively gate-kept sequences, with a strong enrichment of mutations that disrupt gatekeeper motifs in a set of disease-associated mutations listed in the UniProt database (Reumers et al., 2009). The first four of the abovementioned gatekeeper amino acids are at the bottom of different aggregation propensity scales, mainly as a consequence of their very low hydrophobicity, charge and low  $\beta$ -sheet propensity (Monsellier and Chiti, 2007). Proline, as mentioned previously, is considered a  $\beta$ -strand breaker since its conformational rigidity imposes spatial constraints to the cross- $\beta$  folding. A more specific role of this residue in regulating amyloid formation has been proposed, where *cis* to *trans* isomerization can act as an intrinsic molecular switch modulating aggregation propensity. It has been shown that the peptidyl prolyl *cis/trans* isomerase cyclophilin A causes a prolongation of the lag phase and an increase in the yield and length of fibrils formed by the human amyloidogenic protein stefin B (Smajlović et al., 2009). Additionally, structural and biochemical characterization of early intermediates of  $\beta$ 2-microglobulin folding revealed that a *cis* to *trans* isomerization of proline 32 is determinant in the onset of amyloid formation of this protein (Eichner and Radford, 2009), while a similar role of proline isomerization was observed for Ribonuclease A amyloidogenesis (Miller et al., 2010). Understanding the nature of the proline-gatekeeping of MccE492 amyloid formation will provide further information about how this process is regulated and in which circumstances it is favored. The relatively high amount of proline residues found in MccE492 (six), four of them located near to the identified aggregation hotspot, supports the relevance of this amino acid in the modulation of MccE492 amyloidogenesis. It is important to note, however, that the position and sequence context of these residues seems to be determinant, since mutation of only P57 and P59 but not the rest of the proline residues caused a detectable alteration in the aggregation propensity.

Electron microscopy analysis showed that although with a different efficiency, MccE492 mutants P57A, P59A, and even that

lacking residues 54–63, are able to form fibrils morphologically similar to those formed by the native protein. This latter observation suggests that an alternative group of residues could act as a pro-amyloidogenic region. One possibly, is the region comprising residues 19–25, which was detected with a lower consensus by the AMYLPRED2 tool. This region also harbors hydrophobic residues flanked by a polar residue (N) and a proline, which qualify as putative aggregation gatekeepers. Additionally, this group of residues is located inside a region (16–37) that has 74% similarity with a portion of the prion-forming domain of the human prion protein (PrP), and also comprises an imperfect repeat of amino acid residues (21-AALGAPGG-28 and 32-AALGAAGG-39). Curiously, the only different amino acid (in bold) is a proline/alanine, surrounded by non-polar residues. It has been shown that amyloid formation of the CsgA protein (coding major curli subunit) is determined by five imperfect repeats (R1–R5), where only R1 and R5 promote responsiveness to CsgB nucleation and self-seeding by CsgA fibers (Wang and Chapman, 2008; Wang et al., 2008b). Repeats R2–R4 comprise specific aspartic and glycine residues that reduce the aggregation propensity, and thus modulate polymerization efficiency and potential toxicity (Wang et al., 2010). CsgA mutants lacking those gatekeeper residues polymerized *in vitro* significantly faster than wild-type protein, and remarkably, polymerized *in vivo* even in absence of its nucleator CsgB. This points out the possible relevance of the region 19–37 for MccE492 amyloidogenesis. Nevertheless, deletion of residues 18–35 did not affect the intracellular amyloid formation propensity (**Figure 4**), and the *in vitro* aggregation products were typical amyloid fibrils (**Figure 7**). Also, a mutant carrying substitution P26A (rendering a perfect repeat) showed the same intracellular aggregation behavior than the wild-type protein. A similar situation was observed when the alanine 37 was substituted by a proline (data not shown), arguing against the significance of the repeats. In spite of the evidence abovementioned, it cannot be excluded that this region may act leading MccE492 amyloid formation in the absence of the primary aggregation hotspot. Taken together, these results point out the very robust amyloidogenic capacity of MccE492, which is kept even after deleting 18 out of 42 residues of its N-terminal half, or at least 10 out of 42 residues of its C-terminal half.

In agreement with previous reports, MccE492 amyloid formation correlated with the loss of antibacterial activity (Bieler et al., 2005; Marcoleta et al., 2013). The hyperamyloidogenic variants P57A and P59A retained the antibacterial activity but displayed significantly faster inactivation kinetics (**Figure 6** and Supplementary Figure S6). This observation suggests that these residues are dispensable for the pore-forming activity, and that through these mutations it is possible to modulate how long the toxin remains active. From this, it is plausible that alternative substitutions of the gatekeeper residues could generate a battery of toxin variants with different inactivation kinetics.

We found that MccE492 intracellular amyloid formation was observed in three circumstances: when the whole MccE492 genetic system was expressed; when expressing the hyperamyloidogenic mutants P57A and P59A from the pETAB cassette; and when the wild-type form was expressed

from the abovementioned cassette in the presence of np220, a plasmid expressing the genetic determinants encoded in the MccE492 cluster with the exception of the structural gene. These observations suggest that a cluster-encoded factor could act nucleating or promoting MccE492 amyloid formation *in vivo*. Although at first sight this result may appear as unexpected, the general behavior of MccE492 amyloid formation is similar to the well-characterized curli amyloid system, for the following reasons: first, extracellular *in vivo* amyloid formation by CsgA requires an amyloid minor component, the nucleator CsgB (Chapman et al., 2002; Hammer et al., 2007), so it is perfectly plausible that MccE492 as well may require the presence of a nucleator for *in vivo* amyloid formation; second, both purified CsgA and MccE492 can form amyloids *in vitro* with the typical kinetics that has a lag, growth and stationary phase (Bieler et al., 2005; Wang et al., 2007). The requirements for amyloid formation *in vitro* seems to be less restrictive because there is no need of another component, probably because the use of high protein concentrations overcomes the requisite of a nucleator; and third, the duration of the lag phase of both proteins can be significantly shortened *in vitro* by the addition of sonicated preformed fibrils (Wang et al., 2007; Marcoleta et al., 2013). In the same line, the origin of the few single polar inclusions observed in cells carrying np220 could be explained by the aggregation of a factor such as CsgB, that also has amyloid properties (Hammer et al., 2007). Currently, we are working on the identification of the putative nucleator factor.

Whether intracellular MccE492 aggregation has a physiological role, and how its occurrence affects the cellular metabolism, are issues that have to be investigated. Regarding the latter, there are examples indicating that the accumulation of amyloid inclusion bodies are toxic for *E. coli* cells, as seen after overexpression of hyperamyloidogenic variants of RepA and CsgA (Fernández-Tresguerres et al., 2010; Wang et al., 2010). Moreover, it was shown that the identity of residues flanking an aggregation prone region ( $\sigma 32\beta$ ) fused to GFP had a significant effect on bacterial growth, where fusions harboring  $\sigma 32\beta$  flanked by its natural gatekeepers displayed the greatest competitive fitness (Beerten et al., 2012). The impact of intracellular amyloid accumulation over cellular functions makes necessary the existence of cellular mechanisms controlling its occurrence. In this respect, chaperone machineries seem to play an important role. In the case of the RepA-WH1(A31V) prionoid, it was shown that DnaK but not ClpB participates in the remodeling of amyloid inclusions, controlling the transition between mild toxic comet-shaped aggregates and highly toxic globular particles (Gasset-Rosa et al., 2014). In contrast, the disaggregase activity of the ClpB chaperone is required to propagate the yeast prion [PSI<sup>+</sup>] in *E. coli* cells, probably by means of fragmenting higher order intracellular amyloid aggregates to generate smaller seed particles (propagons) that can be inherited to daughter cells during cell division (Yuan et al., 2014). This apparently controversial role of ClpB chaperone in modulating amyloid formation in *E. coli* cells points out that amyloids formed of proteins from different origins not necessarily are subject of the same control mechanisms, even

when they are expressed in the same host. Additionally, it has been demonstrated for curli that an efficient secretion system and chaperone network ensures that CsgA does not form intracellular amyloid aggregates. Moreover, amyloid formation in the periplasm is prevented by the CsgC protein, which selectively inhibited aggregation of CsgA and also  $\alpha$ -synuclein (Evans and Chapman, 2014; Evans et al., 2015). We are currently working in establishing the impact of MccE492 intracellular amyloid formation on the cellular metabolism, and the potential role of chaperones or other factors in controlling this phenomenon. In this regard, preliminary observations show that MccE492 intracellular amyloid accumulation does not seem to have the highly toxic effects observed for other amyloids such as the RepA-WH1(A31V) prionoid. This suggests that MccE492 intracellular amyloid formation may have a role sequestering potentially harmful soluble or oligomeric forms, operating in a similar way as the extracellular toxin inactivation process. In addition, the fact that wild-type amyloid inclusions are observed only when the whole genetic cluster of MccE492 is expressed further support the notion that this phenomenon may have a physiological role. It is important to point out that formation of intracellular amyloid inclusions seems to be a dynamic process, because part of the expressed protein is exported, even in the case of the hyperamyloidogenic mutants. Thus, the formation of amyloid inclusions could be the consequence of MccE492 accumulation because of a limited exporting capacity. Also, although MccE492 expressed from the whole cluster context have a high tendency to form intracellular amyloid inclusions, it could be possible that other factors like chaperones ensure solubility for exporting a proportion of the produced bacteriocin.

One further aspect to be considered is if the immunity protein MceB plays any role in MccE492 intracellular amyloidogenesis. Since the expression of MccE492 in the absence of its immunity protein is lethal to the cell (Bieler et al., 2006), both proteins had to be co-expressed in all the experiments, hampering the possibility to directly compare MccE492 amyloid formation in presence or absence of MceB. However, we believe that it is very unlikely that this protein participates in the amyloidogenic process, because MceB is an integral membrane protein with three transmembrane helices and not detected in the cytoplasm (Lagos et al., 1999), therefore the neutralization of MccE492 by MceB does not occurs in this compartment. Although the exact mechanism by which the immunity protein neutralizes the MccE492 pore-forming activity is unknown, the interaction would occur at the moment that MccE492 is inserted into the inner membrane preventing the correct insertion to form

the pore. On the other hand, this protein does not form inclusions by itself, because when co-expressed with the wild type or hypo-amyloidogenic mutants no inclusions were observed.

MccE492 has advantageous properties for its use as a model to understand amyloid formation and its consequences. First, it is a small protein (84 amino acids), which allows the study of its amyloid behavior using variants of the full peptide, and not only regions or domains of the protein related with amyloid formation. Second, it polymerizes in a shorter time scale than other amyloidogenic proteins, and finally, it can be easily purified from culture supernatants. These properties, and the use of the intracellular amyloid formation phenomenon to screen for defects in amyloid formation, constitute a model useful to study the effect of extrinsic and cellular factors involved in the amyloidogenesis of this and possibly other proteins that can be expressed in *E. coli*.

## AUTHOR CONTRIBUTIONS

AM, OM, RL conceived the work. PA, AM, PL-R, RA, JMV, OM, RL designed the experiments, analyzed the data and interpreted the results. PA, AM, PL-R, RA, JMV conducted the experiments. PA, AM, RL wrote the manuscript. OM and JMV critically revised the manuscript. All the authors approved the final version of the manuscript.

## ACKNOWLEDGMENTS

This work was supported by grants 1140430 from FONDECYT to RL, VID UI 12 905/2-Universidad de Chile and FONDECYT 3140496 to AM, and grant BFU2013-44202 from the Spanish Ministry of Economy to JMV. We thank Felipe Hurtado for his help in determining the antibacterial activity of MccE492 mutants. PA received a CONICYT fellowship to carry out her MSc. Program, and her stay at Dr. Valpuesta's lab was funded by a Short Research Stay Fellowship from DPP- Universidad de Chile.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fmicb.2016.00035>

## REFERENCES

- Arranz, R., Mercado, G., Martín-Benito, J., Giraldo, R., Monasterio, O., Lagos, R., et al. (2012). Structural characterization of microcin E492 amyloid formation: identification of the precursors. *J. Struct. Biol.* 178, 54–60. doi: 10.1016/j.jsb.2012.02.015
- Baeza, M. (2003). *Mechanism of Immunity, Translocation, and Prion-Like Conversion of Microcin E492*. Ph.D. thesis, University of Chile, Faculty of Sciences, Santiago.
- Beerten, J., Jonckheere, W., Rudyak, S., Xu, J., Wilkinson, H., De Smet, F., et al. (2012). Aggregation gatekeepers modulate protein homeostasis of aggregating sequences and affect bacterial fitness. *Protein Eng. Des. Sel.* 25, 357–366. doi: 10.1093/protein/gzs031
- Bieler, S., Estrada, L., Lagos, R., Baeza, M., Castilla, J., and Soto, C. (2005). Amyloid formation modulates the biological activity of a bacterial protein. *J. Biol. Chem.* 280, 26880–26885. doi: 10.1074/jbc.M502031200
- Bieler, S., Silva, F., Soto, C., and Belin, D. (2006). Bactericidal activity of both secreted and nonsecreted microcin E492 requires the



- mannose permease. *J. Bacteriol.* 188, 7049–7061. doi: 10.1128/JB.00688-06
- Carrió, M., González-Montalbán, N., Vera, A., Villaverde, A., and Ventura, S. (2005). Amyloid-like properties of bacterial inclusion bodies. *J. Mol. Biol.* 347, 1025–1037. doi: 10.1016/j.jmb.2005.02.030
- Chapman, M., Robinson, L., Pinkner, J., Roth, R., Heuser, J., Hammar, M., et al. (2002). Role of *Escherichia coli* curli operons in directing amyloid fiber formation. *Science* 295, 851–855. doi: 10.1126/science.1067484
- Chiti, F., and Dobson, C. M. (2006). Protein misfolding, functional amyloid, and human disease. *Annu. Rev. Biochem.* 75, 333–366. doi: 10.1146/annurev.biochem.75.101304.123901
- Chiti, F., Stefani, M., Taddei, N., Ramponi, G., and Dobson, C. M. (2003). Rationalization of the effects of mutations on peptide and protein aggregation rates. *Nature* 424, 805–808. doi: 10.1038/nature01891
- Claessen, D., Rink, R., de Jong, W., Siebring, J., de Vreugd, P., Boersma, F. G. H., et al. (2003). A novel class of secreted hydrophobic proteins is involved in aerial hyphae formation in *Streptomyces coelicolor* by forming amyloid-like fibrils. *Genes Dev.* 17, 1714–1726. doi: 10.1101/gad.264303
- Dasari, M., Espargaró, A., Sabate, R., López del Almo, J. M., Fink, U., Grelle, G., et al. (2011). Bacterial inclusion bodies of Alzheimer's Disease  $\beta$ -amyloid peptides can be employed to study native-like aggregation intermediate states. *Chembiochem* 12, 407–423. doi: 10.1002/cbic.201000602
- de Lorenzo, V. (1984). Isolation and characterization of microcin E492 from *Klebsiella pneumoniae*. *Arch. Microbiol.* 139, 72–75. doi: 10.1007/BF00692715
- de Lorenzo, V., and Pugsley, A. P. (1985). Microcin E492, a low molecular weight peptide antibiotic which causes depolarization of *Escherichia coli* cytoplasmic membrane. *Antimicrob. Agents Chemother.* 27, 666–669. doi: 10.1128/AAC.27.4.666
- Eichner, T., and Radford, S. E. (2009). A generic mechanism of beta2-microglobulin amyloid assembly at neutral pH involving a specific proline switch. *J. Mol. Biol.* 386, 1312–1326. doi: 10.1016/j.jmb.2009.01.013
- Espargaró, A., Sabate, R., and Ventura, S. (2012). Thioflavin-S staining coupled to flow cytometry. A screening tool to detect in vivo protein aggregation. *Mol. Biosyst.* 8, 2839–2844. doi: 10.1039/c2mb25214g
- Evans, M. L., and Chapman, M. R. (2014). Curli Biogenesis: order out of disorder. *Biochim. Biophys. Acta* 1843, 1551–1558. doi: 10.1016/j.bbamcr.2013.09.010
- Evans, M. L., Chorell, E., Taylor, J. D., Aden, J., Göthesson, A., Li, F., et al. (2015). The bacterial curli system possesses a potent and selective inhibitor of amyloid formation. *Mol. Cell* 57, 1–11. doi: 10.1016/j.molcel.2014.12.025
- Fernández-Escamilla, A. M., Rousseau, F., Schymkowitz, J., and Serrano, L. (2004). Prediction of sequence-dependent and mutational effects on the aggregation of peptides and proteins. *Nat. Biotechnol.* 22, 1302–1306. doi: 10.1038/nbt1012
- Fernández-Tresguerres, M. E., Moreno-Díaz de la Espina, S., Gasset-Rosa, F., and Giraldo, R. (2010). A DNA-promoted amyloid proteinopathy in *Escherichia coli*. *Mol. Microbiol.* 77, 1456–1469. doi: 10.1111/j.1365-2958.2010.07299.x
- Fitzpatrick, A. W., Debelouchina, G. T., Bayro, M. J., Clare, D. K., Caporini, M. A., Bajaj, V. S., et al. (2013). Atomic structure and hierarchical assembly of a cross- $\beta$  amyloid fibril. *Proc. Natl. Acad. Sci. U.S.A.* 110, 5468–5473. doi: 10.1073/pnas.1219476110
- Fowler, D. M., Koulou, A. V., Balch, W. E., and Kelly, J. W. (2007). Functional amyloid: from bacteria to humans. *Trends Biochem. Sci.* 32, 217–224. doi: 10.1016/j.tibs.2007.03.003
- Gasset-Rosa, F., Coquel, A. S., Moreno-del Álamo, M., Serrano, A. M., Fernández-Tresguerres, M. E., Moreno-Díaz de la Espina, S., et al. (2014). Direct assessment in bacteria of prionoid propagation and phenotype selection by Hsp70 chaperone. *Mol. Microbiol.* 91, 1070–1087. doi: 10.1111/mmi.12518
- Giraldo, R. (2007). Defined DNA sequences promote the assembly of a bacterial protein into distinct amyloid nanostructures. *Proc. Natl. Acad. Sci. U.S.A.* 104, 17388–17393. doi: 10.1073/pnas.0702006104
- González, B. (2011). *Effect of mceD, mceF, and orfK Gene Expression on Microcin E492 Producing Strains*. M.Sc. thesis, University of Chile, Faculty of Sciences, Santiago, Chile.
- Hammer, N. D., Schmidt, J. D., and Chapman, M. R. (2007). The curli nucleator protein, CsgB contains an amyloidogenic domain that directs CsgA polymerization. *Proc. Natl. Acad. Sci. U.S.A.* 104, 12494–12499. doi: 10.1073/pnas.0703310104
- Hohn, B., and Collins, J. (1980). A small cosmid for efficient cloning of large DNA fragments. *Gene* 11, 291–298. doi: 10.1016/0378-1119(80)90069-4
- Hufnagel, D., Tükel, Ç., and Chapman, M. R. (2013). Disease to dirt: the biology of microbial amyloids. *PLoS Pathog.* 9:e1003740. doi: 10.1371/journal.ppat.1003740
- Invernizzi, G., Papaleo, E., Sabate, R., and Ventura, S. (2012). Protein aggregation: mechanisms and functional consequences. *Int. J. Biochem. Cell Biol.* 44, 1541–1554. doi: 10.1016/j.biocel.2012.05.023
- Lagos, R., Baeza, M., Corsini, G., Hetz, C., Strahsburguer, E., Castillo, J. A., et al. (2001). Structure, organization and characterization of the gene cluster involved in the production of microcin E492, a channel-forming bacteriocin. *Mol. Microbiol.* 42, 229–243. doi: 10.1046/j.1365-2958.2001.02630.x
- Lagos, R., Tello, M., Mercado, G., García, V., and Monasterio, O. (2009). Antibacterial and antitumorigenic properties of microcin E492, a pore-forming bacteriocin. *Curr. Pharm. Biotechnol.* 10, 74–85. doi: 10.2174/138920109787048643
- Lagos, R., Villanueva, J. E., and Monasterio, O. (1999). Identification and properties of the genes encoding microcin E492 and its immunity protein. *J. Bacteriol.* 181, 212–217.
- Lagos, R., Wilkens, M., Vergara, C., Cecchi, X., and Monasterio, O. (1993). Microcin E492 forms ion channels in phospholipid bilayer membranes. *FEBS Lett.* 321, 145–148. doi: 10.1016/0014-5793(93)80096-D
- López de la Paz, M., and Serrano, L. (2004). Sequence determinants of amyloid fibril formation. *Proc. Natl. Acad. Sci. U.S.A.* 10, 87–92. doi: 10.1073/pnas.2634884100
- Marcoleta, A., Marin, M., Mercado, G., Valpuesta, J. M., Monasterio, O., and Lagos, R. (2013). Microcin E492 amyloid formation is retarded by posttranslational modification. *J. Bacteriol.* 195, 1–10. doi: 10.1128/JB.00564-13
- Maurer-Stroh, S., Debulpaep, M., Kuemmerer, N., Lopez de la Paz, M., Martins, I. C., Reumers, J., et al. (2010). Exploring the sequence determinants of amyloid structure using position-specific scoring matrices. *Nat. Methods* 7, 237–245. doi: 10.1038/nmeth.1432
- Mayr-Harting, A., Hedges, A. J., and Berkeley, R. C. W. (1972). "Methods for studying bacteriocins," in *Methods in Microbiology*, eds J. R. Norris and D. W. Ribbons (New York, NY: Academic Press), 315–342.
- Mercado, G., Tello, M., Marín, M., Monasterio, O., and Lagos, R. (2008). The production in vivo of microcin E492 with antibacterial activity depends on salmochelin and EntF. *J. Bacteriol.* 190, 5464–5471. doi: 10.1128/JB.00351-08
- Miller, K. H., Karr, J. R., and Marqusee, S. (2010). A hinge Region Cis-proline in Ribonuclease A acts as a conformational gatekeeper for C-terminal domain Swapping. *J. Mol. Biol.* 400, 567–578. doi: 10.1016/j.jmb.2010.05.017
- Monsellier, E., and Chiti, F. (2007). Prevention of amyloid-like aggregation as a driving force of protein evolution. *EMBO Rep.* 8, 737–742. doi: 10.1038/sj.embor.7401034
- Nilsson, M. R. (2004). Techniques to study amyloid fibril formation in vitro. *Methods* 34, 151–160. doi: 10.1016/j.jymeth.2004.03.012
- Otzen, D. (2010). Functional amyloid. Turning swords into plowshares. *Prion* 4, 256–264. doi: 10.4161/pri.4.4.13676
- Reumers, J., Maurer-Stroh, S., Schymkowitz, J., and Rousseau, F. (2009). Protein sequences encode safeguards against aggregation. *Hum. Mutat.* 30, 431–437. doi: 10.1002/humu.20905
- Romero, D., Aguilar, C., Losick, R., and Kolter, R. (2010). Amyloid fibers provide structural integrity to *Bacillus subtilis* biofilms. *Proc. Natl. Acad. Sci. U.S.A.* 107, 2230–2234. doi: 10.1073/pnas.0910560107
- Romero, D., and Kolter, R. (2014). Functional amyloids in bacteria. *Int. Microbiol.* 17, 65–73. doi: 10.2436/20.1501.01.208
- Rousseau, F., Serrano, L., and Schymkowitz, J. W. (2006). How evolutionary pressure against protein aggregation shaped chaperone specificity. *J. Mol. Biol.* 55, 1037–1047. doi: 10.1016/j.jmb.2005.11.035
- Sawaya, M. R., Sambashivan, S., Nelson, R., Ianova, M. I., Sievers, S. A., Apostol, M. I., et al. (2007). Atomic structures of amyloid cross- $\beta$  spines reveal varied steric zippers. *Nature* 447, 453–457. doi: 10.1038/nature05695
- Schägger, H., and von Jagow, G. (1987). Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal. Biochem.* 166, 368–379. doi: 10.1016/0003-2697(87)90587-2
- Schneider, C. A., Rasband, W. S., and Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* 9, 671–675. doi: 10.1038/nmeth.2089



- Schwartz, K., and Boles, B. (2013). Microbial Amyloids-functions and interactions within the host. *Curr. Opin. Microbiol.* 16, 93–99. doi: 10.1016/j.mib.2012.12.001
- Smajlović, A., Berbić, S., Schiene-Fischer, C., Tušek-Žnidarič, M., Taler, A., and Jenko-Kokalj, S. (2009). Essential role of Pro 74 in stefin B amyloid-fibril formation: dual action of cyclophilin A on the process. *FEBS Lett.* 583, 1114–1120. doi: 10.1016/j.febslet.2009.02.037
- Stefani, M., and Dobson, C. M. (2003). Protein aggregation and aggregate toxicity: new insights into protein folding, misfolding diseases and biological evolution. *J. Mol. Med. (Berl.)* 81, 678–699. doi: 10.1007/s00109-003-0464-5
- Strahsburger, E., Baeza, M., Monasterio, O., and Lagos, R. (2005). Cooperative uptake of microcin E492 by receptors FepA, Fiu, and Cir and inhibition by the siderophore enterochelin and its dimeric and trimeric hydrolysis products. *Antimicrob. Agents Chemother.* 49, 3083–3086. doi: 10.1128/AAC.49.7.3083-3086.2005
- Thomas, X., Destoumieux-Garzón, D., Peduzzi, J., Alfonso, C., Blond, A., Birlirakis, N., et al. (2004). Siderophore peptide, a new type of post-translationally modified antibacterial peptide with potent activity. *J. Biol. Chem.* 279, 28233–28242. doi: 10.1074/jbc.M400228200
- Tsolis, A. C., Papandreou, N. C., Iconomidou, V. I., and Hamodrakas, S. J. (2013). A consensus method for the prediction of 'Aggregation-Prone' peptides in globular proteins. *PLoS ONE* 8:e54175. doi: 10.1371/journal.pone.0054175
- Wang, L., Maji, S. M., Sawaya, M. R., Eisenberg, D., and Riek, R. (2008a). Bacterial inclusion bodies contain amyloid-like structure. *PLoS Biol.* 6:e195. doi: 10.1371/journal.pbio.0060195
- Wang, X., Hammer, N. D., and Chapman, M. R. (2008b). The molecular basis of functional bacterial amyloid polymerization and nucleation. *J. Biol. Chem.* 283, 21530–21539. doi: 10.1074/jbc.M800466200
- Wang, X., and Chapman, M. R. (2008). Sequence determinants of bacterial amyloid formation. *J. Mol. Biol.* 380, 570–580. doi: 10.1016/j.jmb.2008.05.019
- Wang, X., Rochon, M., Lamprokostopoulou, A., Lünsdorf, H., Nimtz, M., and Römmling, U. (2006). Impact of biofilm matrix components on interaction of commensal *Escherichia coli* with the gastrointestinal cell line HT-29. *Cell. Mol. Life Sci.* 63, 2352–2363. doi: 10.1007/s00018-006-6222-4
- Wang, X., Smith, D. R., Jones, J. W., and Chapman, M. R. (2007). In vitro polymerization of a functional amyloid protein. *J. Biol. Chem.* 282, 3713–3719. doi: 10.1074/jbc.M609228200
- Wang, X., Zhou, Y., Ren, J. J., Hammer, N. D., and Chapman, M. R. (2010). Gatekeeper residues in the major curling subunit modulate bacterial amyloid fiber biogenesis. *Proc. Natl. Acad. Sci.* 107, 163–168. doi: 10.1073/pnas.0908714107
- Yuan, A. H., Garrity, S. J., Nako, E., and Hochschild, A. (2014). Prion propagation can occur in a prokaryote and requires the ClpB chaperone. *Elife* 3, e02949. doi: 10.7554/eLife.02949

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Aguilera, Marcoleta, Lobos-Ruiz, Arranz, Valpuesta, Monasterio and Lagos. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Computational analysis of candidate prion-like proteins in bacteria and their role

Valentin Iglesias, Natalia S. de Groot\* and Salvador Ventura\*

Departament de Bioquímica i Biologia Molecular, Institut de Biotecnologia i Biomedicina, Universitat Autònoma de Barcelona, Barcelona, Spain

## OPEN ACCESS

### Edited by:

Catherine Ayn Brissette,  
University of North Dakota, USA

### Reviewed by:

Blaise Boles,  
The University of Iowa, USA  
L. Jeannine Brady,  
University of Florida, USA

### \*Correspondence:

Natalia S. de Groot  
nsgroot@gmail.com;  
Salvador Ventura  
salvador.ventura@uab.cat

### Specialty section:

This article was submitted to  
Microbial Physiology and Metabolism,  
a section of the journal  
Frontiers in Microbiology

**Received:** 24 July 2015

**Accepted:** 28 September 2015

**Published:** 15 October 2015

### Citation:

Iglesias V, de Groot NS and Ventura S  
(2015) Computational analysis of  
candidate prion-like proteins in  
bacteria and their role.  
Front. Microbiol. 6:1123.  
doi: 10.3389/fmicb.2015.01123

Prion proteins were initially associated with diseases such as Creutzfeldt Jakob and transmissible spongiform encephalopathies. However, deeper research revealed them as versatile tools, exploited by the cells to execute fascinating functions, acting as epigenetic elements or building membrane free compartments in eukaryotes. One of the most intriguing properties of prion proteins is their ability to propagate a conformational assembly, even across species. In this context, it has been observed that bacterial amyloids can trigger the formation of protein aggregates by interacting with host proteins. As our life is closely linked to bacteria, either through a parasitic or symbiotic relationship, prion-like proteins produced by bacterial cells might play a role in this association. Bioinformatics is helping us to understand the factors that determine conformational conversion and infectivity in prion-like proteins. We have used PrionScan to detect prion domains in 839 different bacteria proteomes, detecting 2200 putative prions in these organisms. We studied this set of proteins in order to try to understand their functional role and structural properties. Our results suggest that these bacterial polypeptides are associated to peripheral rearrangement, macromolecular assembly, cell adaptability, and invasion. Overall, these data could reveal new threats and therapeutic targets associated to infectious diseases.

**Keywords:** prion, bacteria, protein aggregation, pathogenesis, amyloid

## INTRODUCTION

An increasing number of human diseases are being associated with amyloid forming proteins. Despite these polypeptides are diverse in function, sequence and origin, all share the propensity to form  $\beta$ -sheet aggregates (Karran et al., 2011). Amyloid fibril forming proteins appear to be highly conserved and have been detected in all kingdoms of life, suggesting that, despite they are usually thought to be involved in pathogenic processes, they might indeed provide selective advantages (Sanchez de Groot et al., 2012, 2015; Espinosa Angarica et al., 2013; Malinowska et al., 2013). In fact, cells exploit the formation of amyloid fibrils for diverse purposes (Coustou et al., 1997; Iconomidou et al., 2000; Podrabsky et al., 2001; Chapman et al., 2002; Graether et al., 2003; Fowler et al., 2006; Maji et al., 2009), from structure scaffolding, such as the melanin at the skin, to heritable information transmission, such as the yeast prions (Chien and Weissman, 2001; Shorter and Lindquist, 2005; Liebman and Chernoff, 2012; Staniforth and Tuite, 2012). Because amyloid fibers and their unstable intermediates can be highly cytotoxic (e.g., by disrupting the membrane integrity), the assembly of functional amyloids is a process tightly regulated by the cell, which

involves the assistance of chaperones and a spatiotemporal control (Blanco et al., 2012; Gsponer and Babu, 2012; Evans et al., 2015; Taylor and Matthews, 2015).

Among amyloids, prions are a singular subset of proteins able to change from one conformational state to another, often an amyloid aggregate, and transmit it to other homologous polypeptide sequences. Importantly, recent results suggest that amyloid proteins involved in Alzheimer's and Parkinson's diseases could be infectious and act as prion-like proteins in the brain (Chiti and Dobson, 2006; Stöhr et al., 2012). With the exception of the mammalian prion protein (PrP), prion-like proteins constitute a subset of aggregation-prone proteins with special sequential composition. Whereas, classical amyloid proteins contain specific regions rich in hydrophobic residues that lead the protein self-assembly, prion-like proteins exhibit domains that are commonly enriched in asparagine and glutamine (Q/N) (Dorsman et al., 2002; Fändrich and Dobson, 2002; Halfmann et al., 2011) but also in glycine, serine and tyrosine residues (Kato et al., 2012). This pattern has been found in human proteins associated to neurodegenerative diseases, such as FUS (dementia) or TDP43 (amyotrophic lateral sclerosis; Kato et al., 2012). This special residue content results in low complexity sequences displaying disordered structures, a crucial property that ensures conformational flexibility, permits self-assembly without a requirement for conformational unfolding and allows conversion between species (Tompa and Fuxreiter, 2008; Fuxreiter, 2012; Fuxreiter and Tompa, 2012; Malinowska et al., 2013). In fact, one of the main evolutionary strategies to control protein aggregation is to ensure a stable globular structure preventing, in this way, the exposition of aggregation prone stretches (Lim and Sauer, 1991; de Groot and Ventura, 2005; Ventura, 2005; Monsellier et al., 2007). However, a polypeptide sequence requires more than just low complexity to behave as a prion (Espinosa Angarica et al., 2013; Malinowska et al., 2013). Hence, it has been found that the propagation of amyloid aggregation depends on characteristics such as the degree of over/under representation of specific residues and the length of the considered low complexity region (Ross et al., 2004, 2005; Toombs et al., 2010).

The scientific community is getting closer to elucidate the characteristics that differentiate prion and non-prion amyloid proteins (Kushnirov et al., 2007; Newby and Lindquist, 2013; Sabate et al., 2015). In this way, the knowledge acquired in the last 5 years has allowed the design of prediction approaches to identify putative prion proteins. The first predictive algorithms were based on the properties of the primary sequence responsible for the formation of the classical amyloid aggregates (e.g., high hydrophobicity and intrinsic  $\beta$ -sheet propensity). However, they failed to detect Q/N-rich stretches since these are polar residues that do not fulfill the typical requirements associated with classical  $\beta$ -sheet-amyloid aggregation (Pawar et al., 2005). Then, the algorithms focused on localizing Q/N rich segments in the primary sequence (Michelitsch and Weissman, 2000; Harrison and Gerstein, 2003), without paying much attention to the contribution of the rest of residues (Ross et al., 2005), being unable to score the proteins in terms of their relative prionogenicity. A big improvement was achieved by combining

computational approaches with the experimental validation of new proteins displaying *in vitro* prionic properties. This strategy enlarged the set of prionic sequences and permitted the refinement of the available theoretical models. Alberti and co-workers employed a hidden Markov model (HMM), based on the four *bona fide* yeast prions identified to that moment, obtaining 200 yeast protein candidates carrying putative prion domains (PrDs; Alberti et al., 2009). The *in vivo* and *in vitro* analysis of the top 100 candidates rendered 29 proteins that proved heritable switch and significant *in vivo* amyloid formation. We have recently exploited this experimentally curated dataset to develop a probabilistic model of PrDs able to discover prionogenic proteins in complete proteomes (Espinosa Angarica et al., 2013). We have implemented this model in a web-based algorithm called PrionScan able to handle with large sequence databases and predict prion-like sequence stretches in the proteomes annotated in UniprotKB (Espinosa Angarica et al., 2014). In a previous work, we employed this predictor to analyze all the proteomes reported until that moment (1536 organisms; Espinosa Angarica et al., 2014). We discovered 20540 new putative prions present in 10 different taxonomic divisions, supporting prions universality. We also observed that in most cases the ratio of proteins with prion-forming domains is less than 1% of the whole proteome. Thus, in Archaea and Viruses the number is less than 10 per proteome, while in Bacteria, Fungi, Plantae, and Animalia the range is from few tens to few hundreds, depending on the organisms. Interestingly, we observed that, in different organisms, the predicted PrDs are associated with different cellular components and biological processes supporting prionic properties being employed for diverse biological purposes.

Bacteria are ubiquitous in the world, adapted to multiple environments and able to growth in the most extreme conditions. Moreover, bacterial infection remains a leading cause of death in both Western and developing world (WorldHealthOrganisation, WHO)<sup>1</sup>. Understanding which bacteria proteins display prionic properties could help to understand bacterial biology and pathogenesis. Indeed, despite no genuine prion has been characterized so far for prokaryotes, it is clear that at least *E. coli* can generate infectious conformations of heterologous fungal prions (Sabaté et al., 2009; Garrity et al., 2010; Espargaro et al., 2012; Yuan et al., 2014). In an analogous manner, the formation of amyloids was initially thought to be restricted to eukaryotic cells, but after the first report demonstrating that the curli fibers that emerge from the surfaces of *E. coli* cells had the same physical properties as human amyloids (Chapman et al., 2002), the number of discovered bacterial proteins displaying this ability is steadily increasing (Otzen and Nielsen, 2008; Blanco et al., 2012; Schwartz and Boles, 2013). Moreover, it has been observed that bacterial amyloids can initiate the formation of amyloid aggregates upon interaction with diverse host proteins (Otzen and Nielsen, 2008; Hufnagel et al., 2013; Friedland, 2015; Hill and Lukiw, 2015). With the aim to understand better the potential relevance of bacterial PrDs, here we focus on study the 2200 putative prion proteins predicted by PrionScan within

<sup>1</sup><http://www.who.int/mediacentre/factsheets/fs194/en/>.

the taxon domain bacteria, as derived from the study of 839 bacterial proteomes. Specifically, we analyze the functions and structures associated to these proteins and discuss the possible advantages that they could provide, ensuring their evolutionary conservation.

## MATERIAL AND METHODS

### Sequence Dataset

Our database was comprised of Uniprot Knowledgebase (UniProt, 2015) entries included both in Swissprot and TrEMBL (update 2012\_03) under the taxon domain bacteria in order to track the prion like domains present in bacterial proteomes.

### Discovering Putative Prion-like Domains

PrionScan, an algorithm developed by our group and described previously (Espinosa Angarica et al., 2014), was used in order to predict prion-like domains. Employing a cutoff of 50 bits, we identified 2200 PrD (Table S1). Further analysis was made *a posteriori* in order to identify common traits including the Gene Ontology GO terms for the molecular functions, biological processes, and cellular components and relevant domains according to Pfam database. Pfam domains and GO terms were manually annotated and counted in the 2200 positive PrD containing bacterial proteins according to the Uniprot annotations (UniProt, 2015). Due to the large amount of individual Pfam domains, only those ones represented more than five times were considered in the analysis. Then, we grouped the domains by similarity in their cellular function or process. The list of 18 selected pathogenic bacteria was manually annotated by looking for evidences of a human pathogenic association at the NCBI (Table S2). Then we calculated the enriched characteristics of the PrD containing proteins associated to these bacteria. The enrichment of the different GO terms and domains was calculated as explained below. The enrichment values obtained with the proteins detected in the subset of pathogenic bacteria were compared with those obtained for the complete 2200 PrD containing proteins dataset.

### Statistics

The enrichment analysis was performed with GStat (Beissbarth and Speed, 2004) against the goa\_uniprot database (UniProt, 2015). Out of 2200 initial proteins, 244 (11.09%) were annotated. A *p*-value of 0.1 was set as a cut-off and a false discovery rate (Benjamini) test was performed to obtain it. The initial clustering was performed by classifying the obtained Gene Ontologies according to their category: biological process, cellular component or molecular functions. We calculated the enrichment factors (*EF*) for every GO term to show how much higher is the proportion of hits in relation to the background sample (the total number of proteins). Accordingly, the *EF* is the number of hits among PrDs ( $n^l$ ) divided by the number of annotated proteins in our list ( $p^l$ ) and subsequently divided by the ratio between the hits of that GO term in goa\_annotation ( $n^b$ )

and the total number of proteins ( $p^b$ ) in this specific GO term:

$$EF = \frac{\frac{n^l}{p^l}}{\frac{n^b}{p^b}} = \frac{n^l p^b}{n^b p^l}$$

Only those GO terms with a log2-fold enrichment >0.5 were considered to be significant for their subsequent analysis.

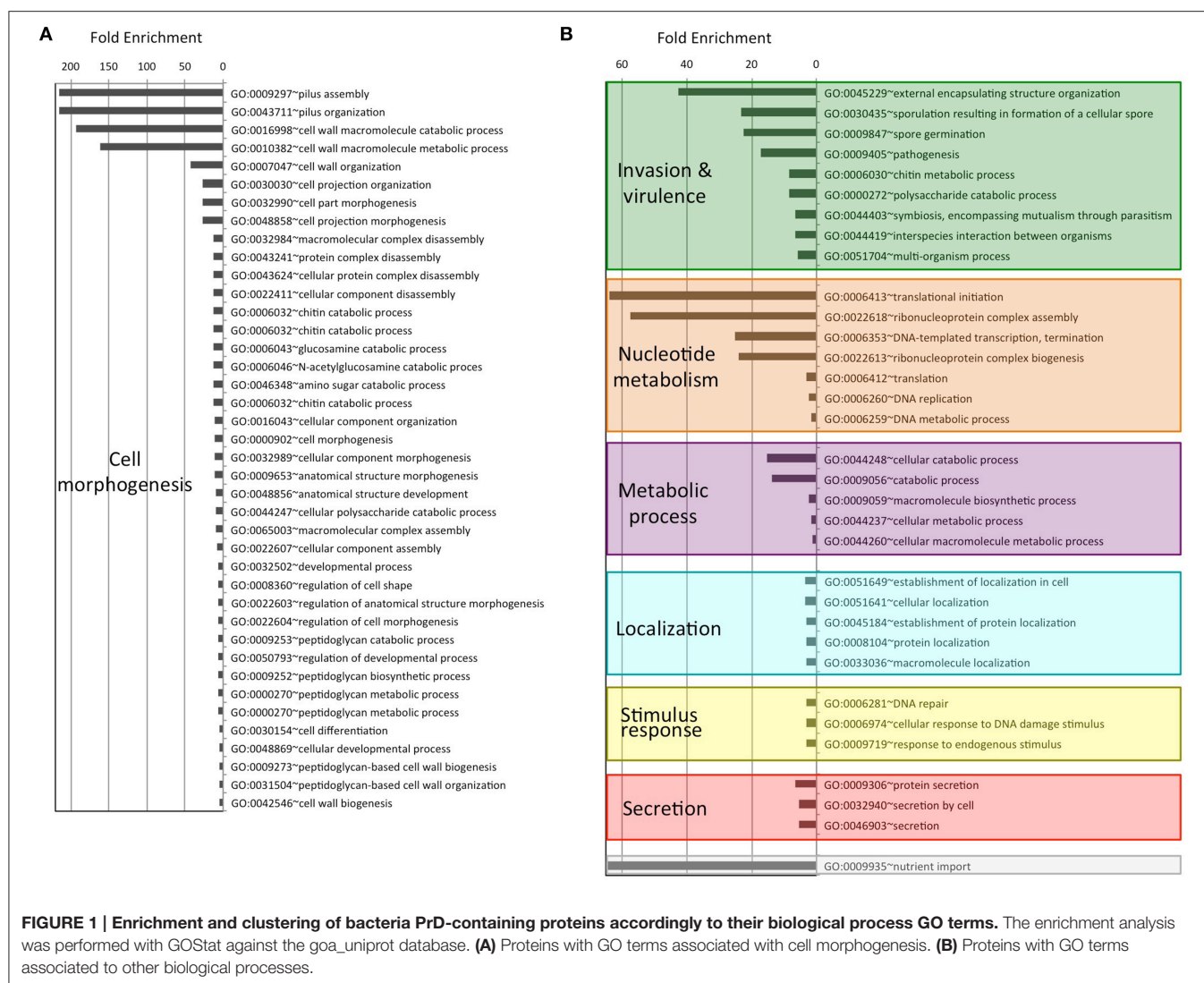
## RESULTS

### Identifying PrD in Bacteria Proteomes

We have analyzed 839 Bacteria proteomes containing a total of 860337 proteins with PrionScan, from which we detected 2200 putative prion proteins scoring higher than 50 bits in the algorithm scale (Espinosa Angarica et al., 2013) accounting for a 0.3% of the complete protein dataset. Interestingly, in the 18 selected pathogenic bacteria (Table S2) proteins containing PrDs are significantly more abundant (2.4%) and indeed they constitute 40% of all the detected PrDs (891 PrDs). Moreover, some specific pathogenic organisms appear to be specially enriched in PrDs: *Staphylococcus aureus* (18%), *Enterococcus faecalis* (10%), *Enterococcus faecium* (5%), or *Staphylococcus epidermidis* (3%). These data show the diversity of putative PrDs distribution and suggest certain associated functionality.

As an attempt to understand the biological purpose of these PrDs we analyzed the Gene Ontology of the corresponding proteins. Additionally, to facilitate its interpretation we have grouped the enriched GO terms by similar cellular function or process. After this classification, the biggest cluster of GO terms collects Biological Processes involved in cell morphogenesis, such as cell projection or cell wall dynamics. This group contains 40 different terms, some of them with fold enrichments above 200 (pilus assembly; **Figure 1A**). We also found several enriched Biological Processes involved in secretion, nutrient import, invasion and virulence; all of them processes involved in interaction with the surrounding environment. Interestingly, in invasion and virulence we find processes associated to encapsulation, sporulation, and interaction with other organisms. Between the Biological Processes, the metabolic ones are particularly involved in the assembly of macromolecules such as polysaccharides and peptidoglycan (**Figure 1B**). The other three Biological Processes clusters are nucleotide metabolism, stimulus to response and localization, which are associated to cellular adaptation and the formation of contacts between molecules. When we analyse the Molecular Functions (**Figure 2A**), the GO terms enriched can be grouped as: nucleic acid binding, metabolic processes, drug binding, and transport. All of them activities associated with the formation of functional interactions. Additionally, the clusters of metabolic process and drug binding perform functions related to cell wall such as peptidoglycan synthesis or chitin production. Moreover, nucleic acid binding functions could be associated to mechanisms of cellular adaptation. The proteins in this cluster are strongly associated to two essential functions such as translation initiation and DNA templated transcription. Surprisingly, the GO terms of the Cell Component





do not include any inside part of the cell, just terms associated to the external part: outer membrane, peptidoglycan based cell wall, plasma membrane, cell wall, and proton transport in flagella (**Figure 2B**; Namba, 2001). It is clear that many of the detected proteins, and specifically those involved in nucleotide binding, are located at the cytosol; however, because the large majority of bacterial proteins are categorized as cytosolic, this results in a poor enrichment factor for this compartment. Overall, the most remarkable characteristics of the bacteria proteins containing PrDs are their role in contact formation (e.g., macromolecular assembly), their relationship with the cell periphery and their involvement in nucleic acid mediated processes.

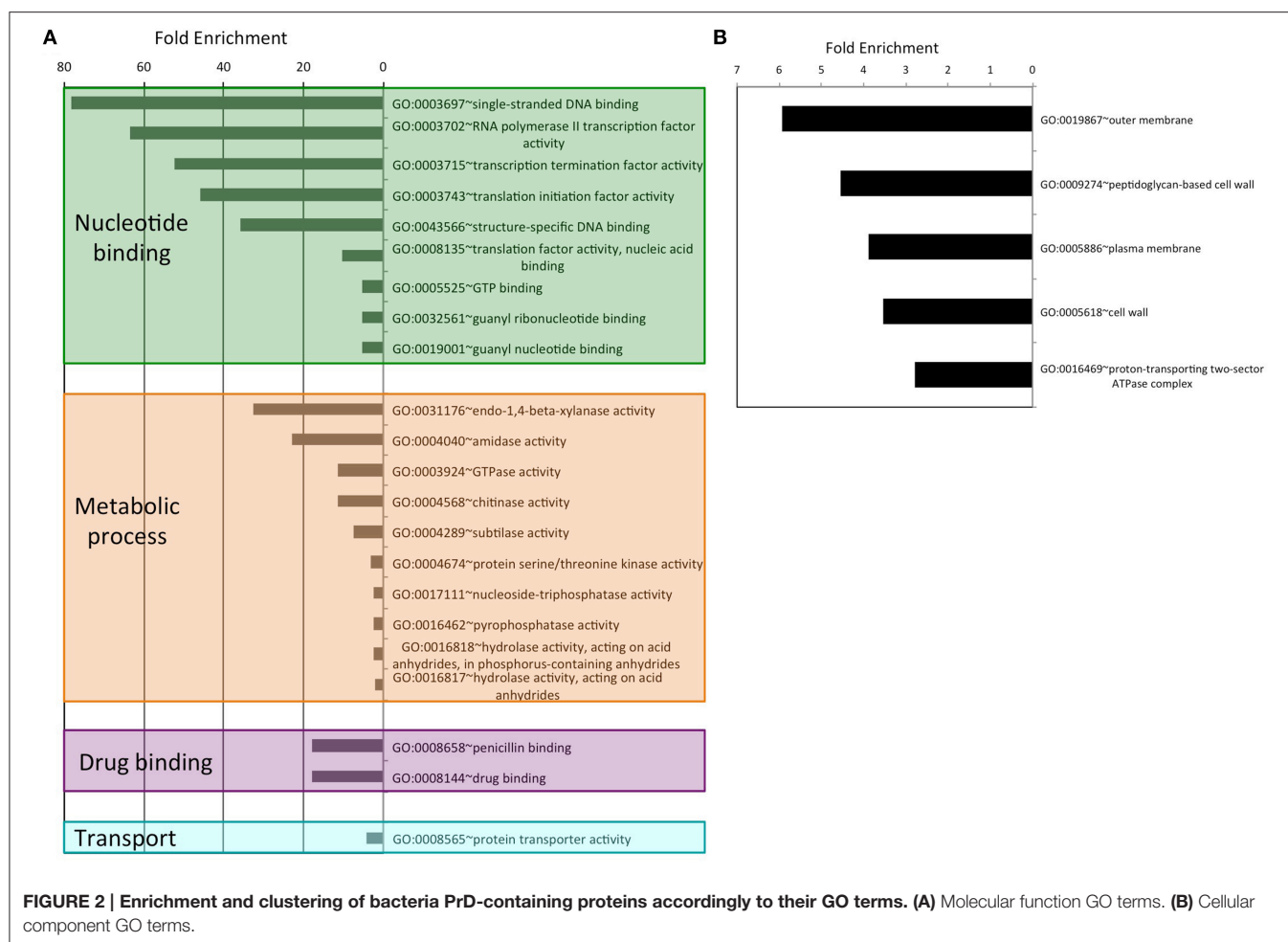
## Structural Domains Linked to Bacteria PrD Proteins

To learn more about the bacterial proteins that possess putative PrDs we examined their constituent functional domains (Finn et al., 2014; **Figure 3**). After clustering the Pfam domains we obtained eight functional groups: nucleotide binding, cell wall

dynamics, invasion and virulence, protein-protein interaction, iron transport, heat-shock, and unknown.

The most abundant group of Pfam families is the one involved in nucleotide binding (1183 domains). There are included domains associated to translation such as GTP-binding elongation factors (GTP\_EFTU), Rho termination factors (Rho\_RNA\_bind and Rho\_N) and translation initiation factors (IF2 and IF2-N). Canonical nucleotide binding domains are also be found such as the single stranded binding protein (SSB), the single zinc ribbon domain (zinc\_ribbon\_2), the major structural motif helix-turn-helix (HTHth-25), and the S1 RNA binding domain. Finally, in this group we can also find an ATP synthase domain, associated with Rho termination factors (ATP-synt\_ab), and the Ribonuclease B OB domain (Finn et al., 2014).

The second most abundant group of Pfam families is, once again, associated to cell wall dynamics (978 domains). This group clusters domains involved in cell wall metabolism (including biosynthesis and degradation) and proteins that bind

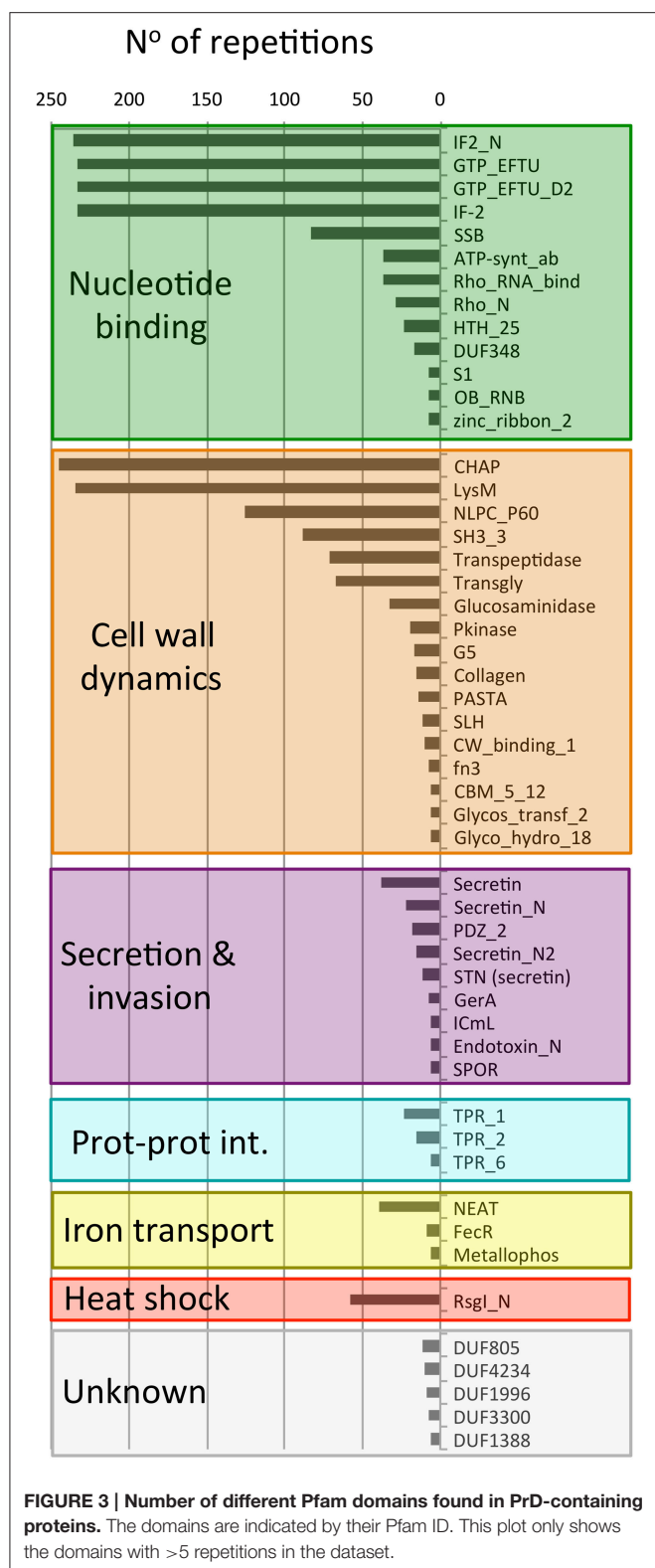


the wall to build functional structures. For example, the lysine motif (Lysm) is involved in bacterial cell wall degradation and may also have peptidoglycan binding function (Bateman and Bycroft, 2000). The Glucosaminidase, Glycosyl transferase family 2 (Glycos\_transf\_2) and Transpeptidase are three domains associated with the biosynthesis of polysaccharides and peptidoglycan (Finn et al., 2014). We also found 67 proteins with a transglycosylase domain (Transgly) that catalyze the polymerization of murein glycan chains as well as 12 proteins with a SLH domain that is associated with the assembly of (glyco)proteins that coat the bacteria surface. The PASTA domain is involved in cell wall biosynthesis and can bind the beta-lactam rings enclosed in antibiotics. The most abundant domain from this group is the CHAP domain (245 proteins) with an amidase activity implicated in cell wall metabolism. Other domains also linked to cell wall are: the collagen domain (connective structures), the NlpC/P60 family (Anantharaman and Aravind, 2003; peptidases associated to lipoproteins), the G5 domain (adhesion), the fibronectin type III (fn3, adhesion), the cell wall binding motif 1 (CW\_binding\_1, a repeat similar to some clostridia toxins) and the carbohydrate-binding module (CBM\_5\_12, enriched in chitinases and associated to cellulose

scaffolding). Additionally, the unknown domain DUF1388 has also been associated with surface lipoproteins.

The third group contains 130 proteins with domains associated to secretion and invasion. Here we have several domains associated to sporulation (SPOR) and spore germination (GerA). The secretin domains are involved in protein export via pore formation in a signal sequence-dependent manner (Van der Meeren et al., 2013; Tosi et al., 2014). The PDZ domains maintain together and organize signaling complexes located throughout the cellular membranes. Finally, the macrophage killing protein domain (ICmL) and the Endotoxin\_N are domains involved in the formation of pores at the host cell membrane (Finn et al., 2014).

Between the PrD containing proteins we have also found three different tetratricopeptide repeat domains (46 repetitions), which scaffold protein-protein interactions and mediate the assembly of multi-protein complexes. In addition, we also obtained 54 domains linked to iron binding and transport (Metallophos, NEAT and FecR) and 58 proteins involved in heat shock response (Anti-sigma factor N-terminus), both types of domains aimed to interact with or to transduce signals coming from the cell external microenvironment.



Overall, the functional families of the PrD containing proteins (Figure 3) match very well with their GO classifications (Figures 1, 2) and confirm that these proteins are associated to

the external part of the cell (e.g., cell wall) and interactions with other molecules (e.g., nucleotide binding).

## Structure Composition of Bacteria PrD Containing Proteins

As expected, the detected PrDs are located inside low complexity regions (e.g., disordered, coiled coil, etc; Figures 4, 5). Moreover, these regions are abundant in the PrD containing proteins and connect different domains (Figure 4) and elements with secondary structure (Figure 5).

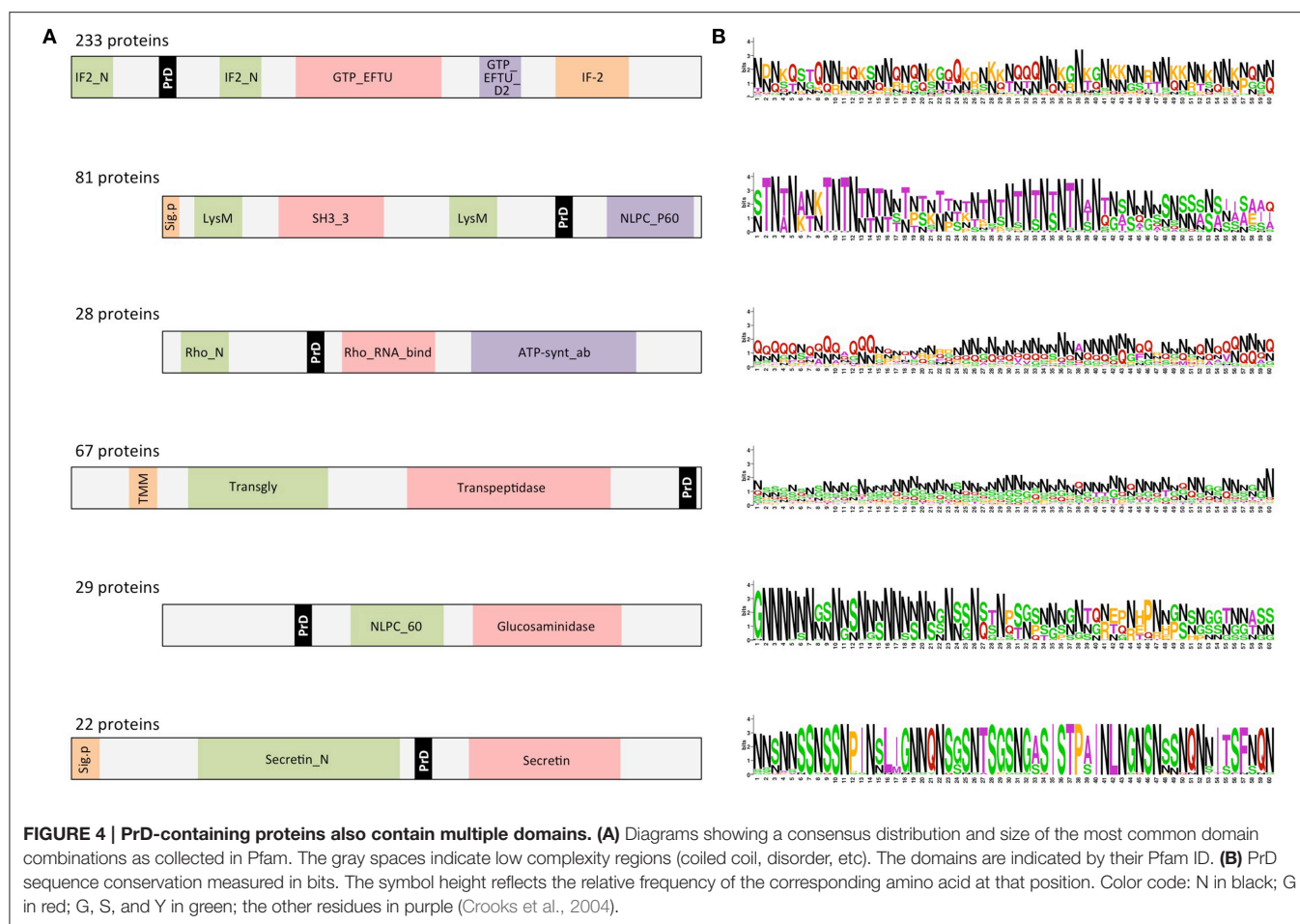
From 2200 PrD containing proteins, 1514 have at least one defined Pfam domain (69%). Additionally, 612 of these sequences (40%) have more than one structural domain (Ekman et al., 2005). When we focus on the PrD containing proteins from pathogenic bacteria (Figure 6), we observe that they have a lower number of designated Pfam domains (just 301 proteins) suggesting they could be less structured proteins or, more likely, carry still unknown domains and functions. Despite this, the proteins from pathogenic bacteria with reported Pfam domains tend to contain more than one structural domain family (Table S2). The percentage of proteins with multiple domains appears to be higher in these proteomes (60%) than in the complete protein dataset (Ekman et al., 2005).

When the proteins have multiple structural domains, the PrD regions can be located either close to an end or between structures (Figure 4A). Interestingly, the amino acid composition of the PrD regions is similar between proteins sharing similar domain arrangement but different between proteins with distinct domains composition (Figure 4B). In agreement with the data reported for yeast prions, we observe that the detected regions are abundant in N (30%), Q (21%), S (11%), and G (11%).

The domain combinations tend to be functionally associated. For example, we found 233 protein sequences containing two GTP-binding elongation factor domains and two translation initiation factor domains that are related with nucleotide binding and translation (Figure 4). During protein synthesis the initiation factors (IF2) form a ternary complex with GTP and the initiator Met-tRNA (Wienk et al., 2005). This complex binds the ribosome to interact with the AUG-codon of the starting methionine, once the codon is found IF2 has to hydrolyze its GTP to be released (Figures 5A,B).

P60 domain is a cell-wall-associated peptidase domain essential for adherence and invasion in some *Listeria* species. In agreement with previous studies (Ponting et al., 1999; Anantharaman and Aravind, 2003), we observed the P60 domain associated with SH3 and LysM domains (Figures 4, 5C,D). It has been hypothesized that this team facilitates the domains interaction with peptides, carbohydrates and lipids from the bacterial cell wall and thus their functionality (Ponting et al., 1999; Anantharaman and Aravind, 2003).

Rho factor proteins tend to be accompanied with an RNA-binding domain and an ATP-hydrolysis domain (Figure 4). The Rho termination factor disengages newly transcribed RNA from its DNA template. Rho catalyzes the 3' endpoint formation and the release of mRNA molecules from DNA templates (Skordalakes and Berger, 2003). The hydrolysis of ATP provides



the energy required to get the RNA-DNA region and break the hybrid structure.

Another example of functional domain combination that contains PrDs are the penicillin-binding proteins. They are bifunctional proteins involved in the final stages of the peptidoglycan synthesis (Figures 4, 5E). At the N-terminus there is a transglycosylase domain involved in the formation of linear glycan strands. And at the C-terminus there is a transpeptidase domain involved in the cross-linking of peptide subunits and drug binding, which is also responsible of the penicillin-sensitivity (Macheboeuf et al., 2005; Sauvage et al., 2008; Contreras-Martel et al., 2011).

NLPC/P60 and Glucosaminidase are two cell wall endopeptidase domains, which emerged together and that we have found accompanied with a PrD (Figure 4). These two domains are commonly employed to cleave the septa connecting the daughter cells during cell separation (Anantharaman and Aravind, 2003; Ruggiero et al., 2010).

The secretins are another example of domain combination found in our set of PrD bacteria containing proteins (Figures 4, 5F). Particularly it is the most abundant combination of two domains (67 times) found in the PrD containing proteins. The secretin domains detected take part in protein secretion systems type II and III. They build multimeric pores to transport

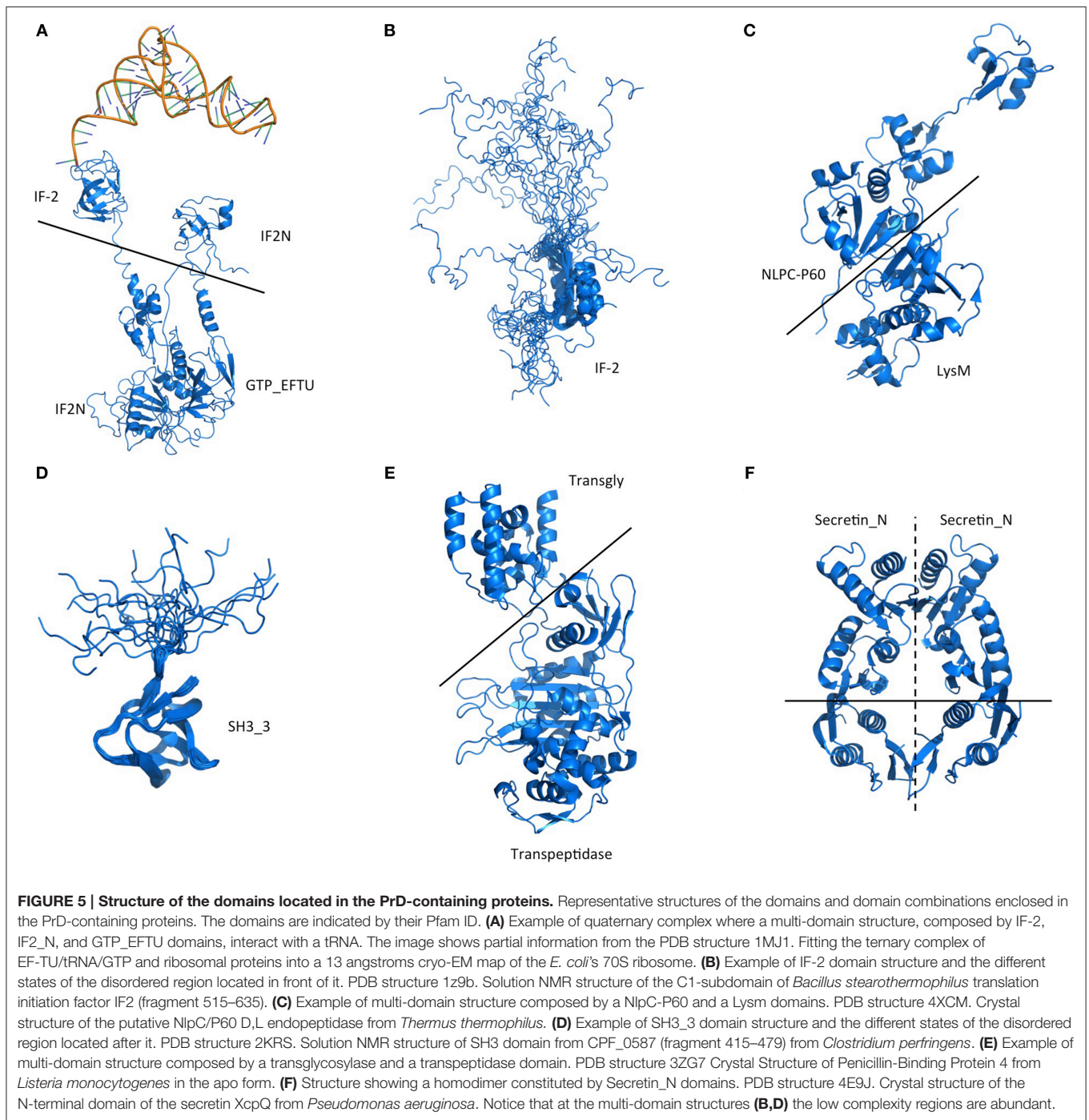
macromolecules either to the periplasm or to inject them into eukaryotic cells (Tosi et al., 2014). In general, secretin proteins consist of two domains: an N-terminal periplasmic domain responsible of the pore formation and a C-terminal domain responsible of the attachment to the outer membrane (Van der Meer et al., 2013; Tosi et al., 2014). Interestingly, the PrD domain detected is located between these two secretin domains (Figure 4).

## DISCUSSION

### Bacterial PrDs are Associated to Cellular Adaptability

We observed that a significant fraction of the bacteria PrD containing proteins are located at the cell periphery and are involved in cell wall metabolism, especially peptidoglycan biogenesis. Peptidoglycan is the major component of bacterial cell walls; it is essential for growth, cell division, and maintenance of the cellular shape, enabling the bacteria to resist intracellular pressures of several atmospheres. In some particular cases, the proteins present in the peptidoglycan can be anchored to the biofilm amyloid network and, more interesting, assist its assembly. This is the case of the TapA protein from *B. subtilis*,

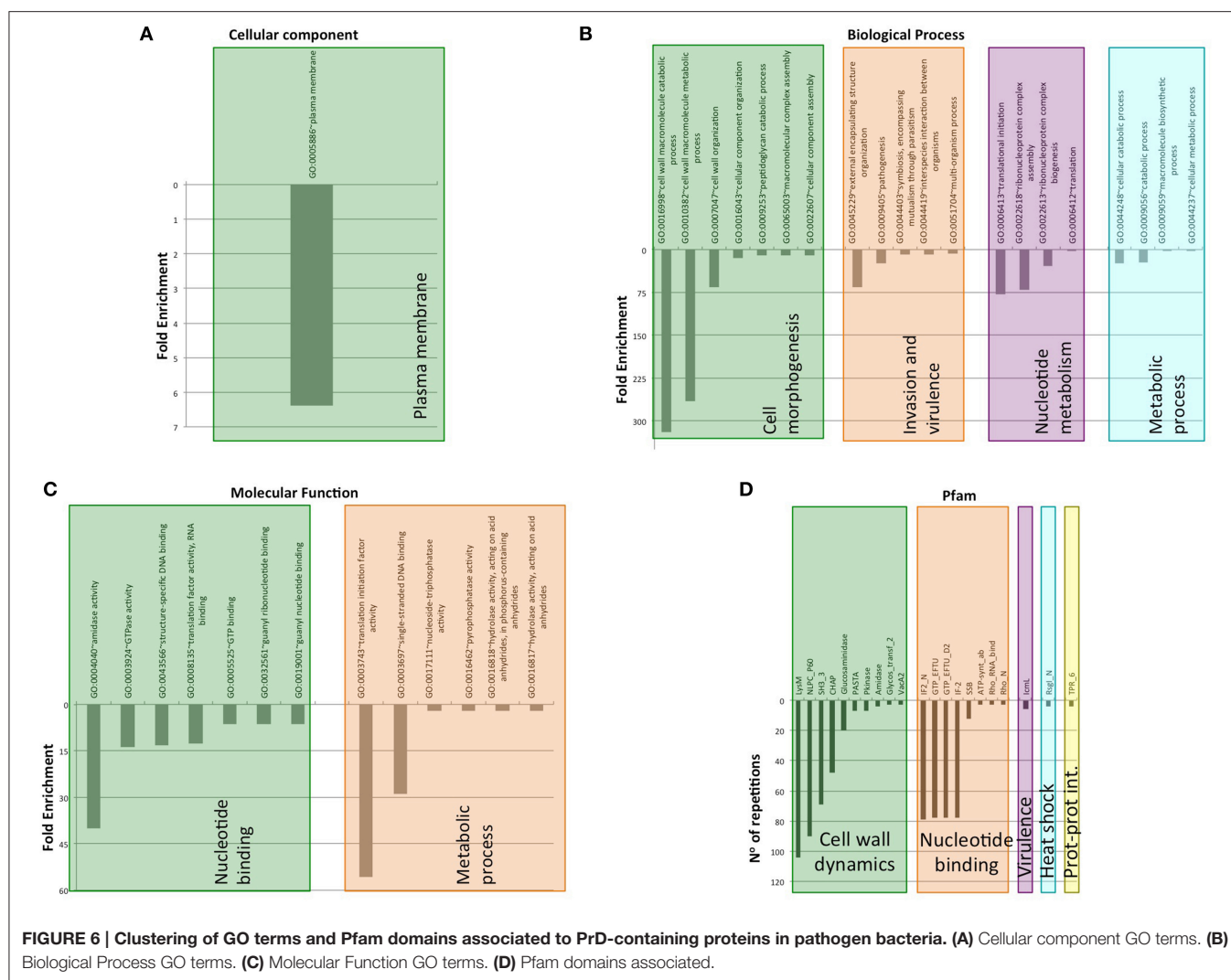




which is present in the peptidoglycan, where it functions as an anchor point for TasA fibers. (Sauvage et al., 2008; Romero et al., 2011; Friedland, 2015). The formation of biofilms is a powerful strategy that protects a bacterial community from chemicals and antibiotics and facilitates the attachment to different surfaces even host cells. Interestingly, *S. aureus*, a biofilm forming pathogen, is the bacteria specie with the highest content in PrDs. In this organism we found PrD-containing proteins linked to cell wall, proteins involved in secretion and proteins associated

to virulence. These data point to a possible relationship between the identified proteins and the biofilm formation. In fact, the *S. aureus* PrD-containing protein staphylococcal secretory antigen ssaA2 (Uniprot code Q2G2J2) is able to form amyloid fibrils *in vitro* (S.V. unpublished results). Thus, a more exhaustive analysis of these proteins might confirm their association to biofilms formation and their possible role as a drug targets.

The other processes enriched in the PrD containing proteins can also provide versatility and adaptability to different



**FIGURE 6 | Clustering of GO terms and Pfam domains associated to PrD-containing proteins in pathogen bacteria. (A)** Cellular component GO terms. **(B)** Biological Process GO terms. **(C)** Molecular Function GO terms. **(D)** Pfam domains associated.

environments. For instance, the proteins involved in stimulus response and invasion and in virulence have a clear role in supporting the bacteria development under variable conditions. From inside the cell the nucleotide binding proteins can be involved in functions that support cell adjustment such as transcription and translation (i.e., change the expression levels) or DNA repair that can enhance cell survival in stress conditions. Interestingly, most of the novel prion-like proteins discovered recently in humans play a role in RNA/DNA binding (King et al., 2012). In bacteria, we also found proteins involved in cellular localization that can rearrange different compounds adapting the cell to new requirements. Overall, as previously proposed for yeast prions, bacterial prions might serve as bet-hedging devices for diversifying microbial phenotypes.

## Bacterial PrDs are Associated to Functional and Interacting Proteins

The 69% of PrDs containing proteins have defined Pfam domains and 40% of them carry multiple domains. Since domains come

together to increase proteins functionality (Anantharaman and Aravind, 2003; Alberti et al., 2009), our data suggest that the proteins with PrDs tend to be functional. Moreover, in pathogenic bacteria PrD are associated to higher percentage of proteins with multiple domains, more than the average of the proteomes from this taxon (Ekman et al., 2005). This data suggests that, in pathogenic bacteria, PrD containing proteins might have a versatile character.

The detected PrDs are located in proteins rich in low complexity regions. These regions are important to provide the structural flexibility required to form interactions between proteins. This flexibility also allows the formation of reversible interactions, which are essential to build dynamic macromolecular assemblies. In fact, the GO terms associated to the PrDs detected by PrionScan comprise functions and processes linked to interaction and assembly. Many of these GO terms involve binding proteins, nucleotides or other cellular compounds. Human RNA/DNA binding proteins use their PrDs to attain functional macromolecular assemblies that regulate transcription and translation. In many cases these functions are

exerted in the so called ribonucleoprotein granules (Malinovska et al., 2013). Many of the proteins containing DNA/RNA binding domains identified in the present also work by forming large complexes and indeed are implied in ribonucleoprotein complex biogenesis and assembly suggesting that this property can be conserved across species. In addition, the association to cell wall dynamics suggests that certain proteins can be implied in the assembly and disassembly of peptidoglycans and polysaccharides. Overall, our data supports that, as previously suggested for eukaryotic PrDs, bacteria PrDs could play an important role in the arrangement of macromolecular structures (Malinovska et al., 2013).

## Prions in other Proteomes

*Saccharomyces cerevisiae* is the organism from which more information about its prion proteins has been so far collected (Alberti et al., 2009; Malinovska et al., 2013). These works showed for the first time that proteins could be employed for amazing functions such as epigenetic elements essential to adapt the cellular metabolism and increase the cell survival in front of environmental changes (Alberti et al., 2009; Newby and Lindquist, 2013). In *S. cerevisiae* the prion proteins are associated to functions that involve the formation of contacts such as RNA-binding, membrane-interacting, DNA binding and protein interaction domains (Malinovska et al., 2013). These proteins are located at the cytoskeleton, nucleus, ribonucleoprotein complexes, and chromatin. Comparing *S. cerevisiae* with other eukaryotic proteomes shows PrD-containing proteins with similar function and location. For example, in human and fruit fly these proteins are also involved in transcription, chromatin remodeling, ribonucleoprotein complex formation, and cytoskeleton (Malinovska et al., 2013). In animals, PrDs tend to be involved in the regulation of central biological processes and organism development, which in vertebrates includes the development of the neural crest. Hence, many human PrD are found in RNA-binding proteins, which deregulation has previously been associated with several neurodegenerative diseases (King et al., 2012).

Eukaryote PrD-containing proteins show less functional diversity than bacteria. In fact, here we have collected all the enriched eukaryote functions (i.e., transcription, RNA binding, and DNA binding) in just one cluster (nucleotide binding). Despite this difference, it appears that, independently of the considered taxon, PrD-containing proteins appear to be involved in a similar regulatory purpose: adapting the cell to a variable environment. This purpose is basically achieved through the control of the expression in eukaryotes, but in prokaryotes this is also reached by interacting with the environment, since microorganisms face the constant challenge of fluctuating conditions in their natural environments. These strategies may have facilitated the invasion of new environments (e.g., water, air) and the coexistence or exploitation of diverse life forms (e.g., host cells).

## Bacteria PrDs and Human Diseases

Our life is closely linked to bacteria, either through a parasitic or symbiotic relationship. On one hand, human microbiota is

required to assist many processes and ensure a healthy body. On the other hand, many common pathogenic bacteria are acquiring antibiotic resistance in all regions of the world (e.g., urinary tract infections, pneumonia, bloodstream infections; WorldHealthOrganisation, WHO). These bacteria cause many hospital-acquired infections, such as the methicillin-resistant *S. aureus*, with an associated high mortality rate (Contreras-Martel et al., 2011; WorldHealthOrganisation, WHO).

To the already intricate scenario where bacteria and host interact, the risk of their amyloid proteins concurring and altering their conformational states adds an extra level of complexity (Otzen and Nielsen, 2008). Additionally, the long periods that bacteria stay in the body, due to chronic infection or microbiota coexistence, enhances the chances of this event. In fact, recent studies have demonstrated that bacterial amyloids can initiate the formation of amyloid aggregates upon interaction with host proteins (Otzen and Nielsen, 2008; Zhou et al., 2012; Hufnagel et al., 2013; Hill and Lukiw, 2015). Moreover, it has been reported that the injection of bacteria amyloids in mice causes the development of amyloidosis (Lundmark et al., 2005). Overall, these data reminds the conformational template process associated to prion transmission and suggest that bacterial infection could be linked to neurodegenerative diseases (Friedland, 2015).

## General Conclusions

Despite PrD-containing proteins seem to be ubiquitous (Espinosa Angarica et al., 2013; Malinovska et al., 2013) they play distinct functional roles in different species. In this background, the mechanisms underlying host-bacteria relationship are just starting to be elucidated and, as a result, also the interplay between their amyloid proteins (Zhou et al., 2012; Schwartz and Boles, 2013; Seviour et al., 2015). The studies on bacteria amyloids are showing us that amyloid aggregates can be exploited to execute wide range of amazing functions (Blanco et al., 2012; DePas and Chapman, 2012; Gsponer and Babu, 2012; Zhou et al., 2012; Schwartz and Boles, 2013; Evans et al., 2015; Seviour et al., 2015; Taylor and Matthews, 2015). Because the formation of amyloids comes at expenses of the formation of transient toxic species cells tightly control the assembly of these macromolecular structures and how they can interact with proteins from other species (Zhou et al., 2012; Schwartz and Boles, 2013; Evans et al., 2015; Taylor and Matthews, 2015). Most of the bacterial amyloids described so far play a structural role and work extracellularly. Indeed, some of the PrD containing proteins with potential amyloidogenic properties could be linked to biofilms, structures that favor chronic human infections and, consequently, increase the chances of a potential bacterial prion to alter the conformation of host proteins. However, despite their *in vitro* amyloid potential and *in vivo* prionic behavior should be validated, the data in the present work suggest that, as it happens in yeast and humans, also in bacteria amyloid-like assemblies might play a regulatory role, since some of the detected candidates are linked to fundamental cellular functions such as transcription, translation or DNA repair. Intriguingly, linking the fact that we found at the same time association with extracellular environment and nucleic acid binding function,

it has been reported recently that extracellular DNA is bound tightly by bacterial amyloid fibrils during biofilm formation and that amyloid/DNA composites are immune stimulators when injected into mice, leading to autoimmunity (Gallo et al., 2015; Spaulding et al., 2015). Overall it becomes clear that a more exhaustive analysis of the putative bacterial prion proteins identified here is required in order to attain a better understanding of their functional role and their relationship with human diseases. This data could help to identify new drug targets and develop new therapies.

## REFERENCES

- Alberti, S., Halfmann, R., King, O., Kapila, A., and Lindquist, S. (2009). A systematic survey identifies prions and illuminates sequence features of prionogenic proteins. *Cell* 137, 146–158. doi: 10.1016/j.cell.2009.02.044
- Anantharaman, V., and Aravind, L. (2003). Evolutionary history, structural features and biochemical diversity of the NlpC/P60 superfamily of enzymes. *Genome Biol.* 4:R11. doi: 10.1186/gb-2003-4-2-r11
- Bateman, A., and Bycroft, M. (2000). The structure of a LysM domain from *E. coli* membrane-bound lytic murein transglycosylase D (MltD). *J. Mol. Biol.* 299, 1113–1119. doi: 10.1006/jmbi.2000.3778
- Beissbarth, T., and Speed, T. P. (2004). GStat: find statistically overrepresented Gene Ontologies within a group of genes. *Bioinformatics* 20, 1464–1465. doi: 10.1093/bioinformatics/bth088
- Blanco, L. P., Evans, M. L., Smith, D. R., Badtke, M. P., and Chapman, M. R. (2012). Diversity, biogenesis and function of microbial amyloids. *Trends Microbiol.* 20, 66–73. doi: 10.1016/j.tim.2011.11.005
- Chapman, M. R., Robinson, L. S., Pinkner, J. S., Roth, R., Heuser, J., Hammar, M., et al. (2002). Role of *Escherichia coli* curli operons in directing amyloid fiber formation. *Science* 295, 851–855. doi: 10.1126/science.1067484
- Chien, P., and Weissman, J. S. (2001). Conformational diversity in a yeast prion dictates its seeding specificity. *Nature* 410, 223–227. doi: 10.1038/35065632
- Chiti, F., and Dobson, C. M. (2006). Protein misfolding, functional amyloid, and human disease. *Annu. Rev. Biochem.* 75, 333–366. doi: 10.1146/annurev.biochem.75.101304.123901
- Contreras-Martel, C., Amoroso, A., Woon, E. C., Zervosen, A., Inglis, S., Martins, A., et al. (2011). Structure-guided design of cell wall biosynthesis inhibitors that overcome beta-lactam resistance in *Staphylococcus aureus* (MRSA). *ACS Chem. Biol.* 6, 943–951. doi: 10.1021/ch2001846
- Coustou, V., Deleu, C., Saupé, S., and Begueret, J. (1997). The protein product of the *het-s* heterokaryon incompatibility gene of the fungus *Podospora anserina* behaves as a prion analog. *Proc. Natl. Acad. Sci. U.S.A.* 94, 9773–9778. doi: 10.1073/pnas.94.18.9773
- Crooks, G. E., Hon, G., Chandonia, J. M., and Brenner, S. E. (2004). WebLogo: a sequence logo generator. *Genome Res.* 14, 1188–1190. doi: 10.1101/gr.849004
- de Groot, N. S., and Ventura, S. (2005). Amyloid fibril formation by bovine cytochrome *c*. *Spectroscopy* 19, 199–205. doi: 10.1155/2005/104348
- DePas, W. H., and Chapman, M. R. (2012). Microbial manipulation of the amyloid fold. *Res. Microbiol.* 163, 592–606. doi: 10.1016/j.resmic.2012.10.009
- Dorsman, J. C., Peppers, B., Langenberg, D., Kerkdijk, H., Ijszenga, M., den Dunnen, J. T., et al. (2002). Strong aggregation and increased toxicity of polyglutamine over polyglutamine stretches in mammalian cells. *Hum. Mol. Genet.* 11, 1487–1496. doi: 10.1093/hmg/11.13.1487
- Ekman, D., Björklund, A. K., Frey-Skött, J., and Elofsson, A. (2005). Multi-domain proteins in the three kingdoms of life: orphan domains and other unassigned regions. *J. Mol. Biol.* 348, 231–243. doi: 10.1016/j.jmb.2005.02.007
- Espargaro, A., Villar-Piqué, A., Sabaté, R., and Ventura, S. (2012). Yeast prions form infectious amyloid inclusion bodies in bacteria. *Microb. Cell Fact.* 11:89. doi: 10.1186/1475-2859-11-89
- ## ACKNOWLEDGMENTS
- This work was funded by the Spanish Ministry of Economy and Competitiveness BFU2013-44763-P to SV.
- ## SUPPLEMENTARY MATERIAL
- The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fmicb.2015.01123>
- Espinosa Angarica, V., Angulo, A., Giner, A., Losilla, G., Ventura, S., and Sancho, J. (2014). PrionScan: an online database of predicted prion domains in complete proteomes. *BMC Genomics* 15:102. doi: 10.1186/1471-2164-15-102
- Espinosa Angarica, V., Ventura, S., and Sancho, J. (2013). Discovering putative prion sequences in complete proteomes using probabilistic representations of Q/N-rich domains. *BMC Genomics* 14:316. doi: 10.1186/1471-2164-14-316
- Evans, M. L., Chorell, E., Taylor, J. D., Åden, J., Göthesson, A., Li, F., et al. (2015). The bacterial curli system possesses a potent and selective inhibitor of amyloid formation. *Mol. Cell* 57, 445–455. doi: 10.1016/j.molcel.2014.12.025
- Fändrich, M., and Dobson, C. M. (2002). The behaviour of polyamino acids reveals an inverse side chain effect in amyloid structure formation. *EMBO J.* 21, 5682–5690. doi: 10.1093/emboj/cdf573
- Finn, R. D., Bateman, A., Clements, J., Coggill, P., Eberhardt, R. Y., Eddy, S. R., et al. (2014). Pfam: the protein families database. *Nucleic Acids Res.* 42, D222–D230. doi: 10.1093/nar/gkt1223
- Fowler, D. M., Koulou, A. V., Alory-Jost, C., Marks, M. S., Balch, W. E., and Kelly, J. W. (2006). Functional amyloid formation within mammalian tissue. *PLoS Biol.* 4:e6. doi: 10.1371/journal.pbio.0040006
- Friedland, R. P. (2015). Mechanisms of molecular mimicry involving the microbiota in neurodegeneration. *J. Alzheimers. Dis.* 45, 349–362. doi: 10.3233/JAD-142841
- Fuxreiter, M., and Tompa, P. (2012). Fuzzy complexes: a more stochastic view of protein function. *Adv. Exp. Med. Biol.* 725, 1–14. doi: 10.1007/978-1-4614-0659-4\_1
- Fuxreiter, M. (2012). Fuzziness: linking regulation to protein dynamics. *Mol. Biosyst.* 8, 168–177. doi: 10.1039/C1MB05234A
- Gallo, P. M., Rapsinski, G. J., Wilson, R. P., Oppong, G. O., Sriram, U., Goulian, M., et al. (2015). Amyloid-DNA composites of bacterial biofilms stimulate autoimmunity. *Immunity* 42, 1171–1184. doi: 10.1016/j.immuni.2015.06.002
- Garrity, S. J., Sivanathan, V., Dong, J., Lindquist, S., and Hochschild, A. (2010). Conversion of a yeast prion protein to an infectious form in bacteria. *Proc. Natl. Acad. Sci. U.S.A.* 107, 10596–10601. doi: 10.1073/pnas.0913280107
- Graether, S. P., Slupsky, C. M., and Sykes, B. D. (2003). Freezing of a fish antifreeze protein results in amyloid fibril formation. *Biophys. J.* 84, 552–557. doi: 10.1016/S0006-3495(03)74874-7
- Gsponer, J., and Babu, M. M. (2012). Cellular strategies for regulating functional and nonfunctional protein aggregation. *Cell Rep.* 2, 1425–1437. doi: 10.1016/j.celrep.2012.09.036
- Halfmann, R., Alberti, S., Krishnan, R., Lyle, N., O'Donnell, C. W., King, O. D., et al. (2011). Opposing effects of glutamine and asparagine govern prion formation by intrinsically disordered proteins. *Mol. Cell* 43, 72–84. doi: 10.1016/j.molcel.2011.05.013
- Harrison, P. M., and Gerstein, M. (2003). A method to assess compositional bias in biological sequences and its application to prion-like glutamine/asparagine-rich domains in eukaryotic proteomes. *Genome Biol.* 4:R40. doi: 10.1186/gb-2003-4-6-r40
- Hill, J. M., and Lukiw, W. J. (2015). Microbial-generated amyloids and Alzheimer's disease (AD). *Front. Aging Neurosci.* 7:9. doi: 10.3389/fnagi.2015.00009



- Hufnagel, D. A., Tükel, C., and Chapman, M. R. (2013). Disease to dirt: the biology of microbial amyloids. *PLoS Pathog.* 9:e1003740. doi: 10.1371/journal.ppat.1003740
- Iconomidou, V. A., Vriend, G., and Hamodrakas, S. J. (2000). Amyloids protect the silkworm oocyte and embryo. *FEBS Lett.* 479, 141–145. doi: 10.1016/S0014-5793(00)01888-3
- Karran, E., Mercken, M., and De Strooper, B. (2011). The amyloid cascade hypothesis for Alzheimer's disease: an appraisal for the development of therapeutics. *Nat. Rev. Drug Discov.* 10, 698–712. doi: 10.1038/nrd3505
- Kato, M., Han, T. W., Xie, S., Shi, K., Du, X., Wu, L. C., et al. (2012). Cell-free formation of RNA granules: low complexity sequence domains form dynamic fibers within hydrogels. *Cell* 149, 753–767. doi: 10.1016/j.cell.2012.04.017
- King, O. D., Gitler, A. D., and Shorter, J. (2012). The tip of the iceberg: RNA-binding proteins with prion-like domains in neurodegenerative disease. *Brain Res.* 1462, 61–80. doi: 10.1016/j.brainres.2012.01.016
- Kushnirov, V. V., Vishnevskaya, A. B., Alexandrov, I. M., and Ter-Avanesyan, M. D. (2007). Prion and nonprion amyloids: a comparison inspired by the yeast Sup35 protein. *Prion* 1, 179–184. doi: 10.4161/pri.1.3.4840
- Liebmman, S. W., and Chernoff, Y. O. (2012). Prions in yeast. *Genetics* 191, 1041–1072. doi: 10.1534/genetics.111.137760
- Lim, W. A., and Sauer, R. T. (1991). The role of internal packing interactions in determining the structure and stability of a protein. *J. Mol. Biol.* 219, 359–376. doi: 10.1016/0022-2836(91)90570-V
- Lundmark, K., Westermark, G. T., Olsén, A., and Westermark, P. (2005). Protein fibrils in nature can enhance amyloid protein A amyloidosis in mice: cross-seeding as a disease mechanism. *Proc. Natl. Acad. Sci. U.S.A.* 102, 6098–6102. doi: 10.1073/pnas.0501814102
- Macheboeuf, P., Di Guilmi, A. M., Job, V., Vernet, T., Dideberg, O., and Dessen, A. (2005). Active site restructuring regulates ligand recognition in class A penicillin-binding proteins. *Proc. Natl. Acad. Sci. U.S.A.* 102, 577–582. doi: 10.1073/pnas.0407186102
- Maji, S. K., Perrin, M. H., Sawaya, M. R., Jessberger, S., Vadodaria, K., Rissman, R. A., et al. (2009). Functional amyloids as natural storage of peptide hormones in pituitary secretory granules. *Science* 325, 328–332. doi: 10.1126/science.1173155
- Malinowska, L., Kroschwald, S., and Alberti, S. (2013). Protein disorder, prion propensities, and self-organizing macromolecular collectives. *Biochim. Biophys. Acta* 1834, 918–931. doi: 10.1016/j.bbapap.2013.01.003
- Michelitsch, M. D., and Weissman, J. S. (2000). A census of glutamine/asparagine-rich regions: implications for their conserved function and the prediction of novel prions. *Proc. Natl. Acad. Sci. U.S.A.* 97, 11910–11915. doi: 10.1073/pnas.97.22.11910
- Monsellier, E., Ramazzotti, M., de Laureto, P. P., Tartaglia, G. G., Taddei, N., Fontana, A., et al. (2007). The distribution of residues in a polypeptide sequence is a determinant of aggregation optimized by evolution. *Biophys. J.* 93, 4382–4391. doi: 10.1529/biophysj.107.111336
- Namba, K. (2001). Roles of partly unfolded conformations in macromolecular self-assembly. *Genes Cells* 6, 1–12. doi: 10.1046/j.1365-2443.2001.00384.x
- Newby, G. A., and Lindquist, S. (2013). Blessings in disguise: biological benefits of prion-like mechanisms. *Trends Cell Biol.* 23, 251–259. doi: 10.1016/j.tcb.2013.01.007
- Otzen, D., and Nielsen, P. H. (2008). We find them here, we find them there: functional bacterial amyloid. *Cell Mol. Life Sci.* 65, 910–927. doi: 10.1007/s00018-007-7404-4
- Pawar, A. P., Dubay, K. F., Zurdo, J., Chiti, F., Vendruscolo, M., and Dobson, C. M. (2005). Prediction of “aggregation-prone” and “aggregation-susceptible” regions in proteins associated with neurodegenerative diseases. *J. Mol. Biol.* 350, 379–392. doi: 10.1016/j.jmb.2005.04.016
- Podrabsky, J. E., Carpenter, J. F., and Hand, S. C. (2001). Survival of water stress in annual fish embryos: dehydration avoidance and egg envelope amyloid fibers. *Am. J. Physiol. Regul. Integr. Comp. Physiol.* 280, R123–R131.
- Ponting, C. P., Aravind, L., Schultz, J., Bork, P., and Koonin, E. V. (1999). Eukaryotic signalling domain homologues in archaea and bacteria. Ancient ancestry and horizontal gene transfer. *J. Mol. Biol.* 289, 729–745.
- Romero, D., Vlamakis, H., Losick, R., and Kolter, R. (2011). An accessory protein required for anchoring and assembly of amyloid fibres in *B. subtilis* biofilms. *Mol. Microbiol.* 80, 1155–1168. doi: 10.1111/j.1365-2958.2011.07653.x
- Ross, E. D., Baxa, U., and Wickner, R. B. (2004). Scrambled prion domains form prions and amyloid. *Mol. Cell Biol.* 24, 7206–7213. doi: 10.1128/MCB.24.16.7206-7213.2004
- Ross, E. D., Edsikes, H. K., Terry, M. J., and Wickner, R. B. (2005). Primary sequence independence for prion formation. *Proc. Natl. Acad. Sci. U.S.A.* 102, 12825–12830. doi: 10.1073/pnas.0506136102
- Ruggiero, A., Marasco, D., Squeglia, F., Soldini, S., Pedone, E., Pedone, C., et al. (2010). Structure and functional regulation of RipA, a mycobacterial enzyme essential for daughter cell separation. *Structure* 18, 1184–1190. doi: 10.1016/j.str.2010.06.007
- Sabaté, R., Espargaró, A., Saupe, S. J., and Ventura, S. (2009). Characterization of the amyloid bacterial inclusion bodies of the HET-s fungal prion. *Microb. Cell Fact.* 8:56. doi: 10.1186/1475-2859-8-56
- Sabaté, R., Rousseau, F., Schymkowitz, J., and Ventura, S. (2015). What makes a protein sequence a prion? *PLoS Comput. Biol.* 11:e1004013. doi: 10.1371/journal.pcbi.1004013
- Sanchez de Groot, N., Gomes, R. A., Villar-Pique, A., Babu, M. M., Coelho, A. V., and Ventura, S. (2015). Proteome response at the edge of protein aggregation. *Open Biol.* 5:140221. doi: 10.1098/rsob.140221
- Sanchez de Groot, N., Torrent, M., Villar-Pique, A., Lang, B., Ventura, S., Gsponer, J., et al. (2012). Evolutionary selection for protein aggregation. *Biochem. Soc. Trans.* 40, 1032–1037. doi: 10.1042/BST20120160
- Sauvage, E., Kerff, F., Terrak, M., Ayala, J. A., and Charlier, P. (2008). The penicillin-binding proteins: structure and role in peptidoglycan biosynthesis. *FEMS Microbiol. Rev.* 32, 234–258. doi: 10.1111/j.1574-6976.2008.00105.x
- Schwartz, K., and Boles, B. R. (2013). Microbial amyloids—functions and interactions within the host. *Curr. Opin. Microbiol.* 16, 93–99. doi: 10.1016/j.mib.2012.12.001
- Seviour, T., Hansen, S. H., Yang, L., Yau, Y. H., Wang, V. B., Stenvang, M. R., et al. (2015). Functional amyloids keep quorum-sensing molecules in check. *J. Biol. Chem.* 290, 6457–6469. doi: 10.1074/jbc.M114.613810
- Shorter, J., and Lindquist, S. (2005). Prions as adaptive conduits of memory and inheritance. *Nat. Rev. Genet.* 6, 435–450. doi: 10.1038/nrg1616
- Skordalakes, E., and Berger, J. M. (2003). Structure of the Rho transcription terminator: mechanism of mRNA recognition and helicase loading. *Cell* 114, 135–146. doi: 10.1016/S0092-8674(03)00512-9
- Spaulding, C. N., Dodson, K. W., Chapman, M. R., and Hultgren, S. J. (2015). Fueling the fire with fibers: bacterial amyloids promote inflammatory disorders. *Cell Host Microbe* 18, 1–2. doi: 10.1016/j.chom.2015.06.013
- Staniforth, G. L., and Tuite, M. F. (2012). Fungal prions. *Prog. Mol. Biol. Transl. Sci.* 107, 417–456. doi: 10.1016/B978-0-12-385883-2.00007-2
- Stöhr, J., Watts, J. C., Mensinger, Z. L., Oehler, A., Grillo, S. K., DeArmond, S. J., et al. (2012). Purified and synthetic Alzheimer's amyloid beta (Aβeta) prions. *Proc. Natl. Acad. Sci. U.S.A.* 109, 11025–11030. doi: 10.1073/pnas.1206555109
- Taylor, J. D., and Matthews, S. J. (2015). New insight into the molecular control of bacterial amyloids. *Front. Cell. Infect. Microbiol.* 5:33. doi: 10.3389/fcimb.2015.00033
- Tomba, P., and Fuxreiter, M. (2008). Fuzzy complexes: polymorphism and structural disorder in protein-protein interactions. *Trends Biochem. Sci.* 33, 2–8. doi: 10.1016/j.tics.2007.10.003
- Toombs, J. A., McCarty, B. R., and Ross, E. D. (2010). Compositional determinants of prion formation in yeast. *Mol. Cell Biol.* 30, 319–332. doi: 10.1128/MCB.01140-09
- Tosi, T., Estrozi, L. F., Job, V., Guilvout, I., Pugsley, A. P., Schoehn, G., et al. (2014). Structural similarity of secretins from type II and type III secretion systems. *Structure* 22, 1348–1355. doi: 10.1016/j.str.2014.07.005
- UniProt, C. (2015). UniProt: a hub for protein information. *Nucleic Acids Res.* 43, D204–D212. doi: 10.1093/nar/gku989
- Van der Meeren, R., Wen, Y., Van Gelder, P., Tommassen, J., Devreese, B., Savvides, S. N., et al. (2013). New insights into the assembly of bacterial secretins: structural studies of the periplasmic domain of XcpQ from *Pseudomonas aeruginosa*. *J. Biol. Chem.* 288, 1214–1225. doi: 10.1074/jbc.M112.432096
- Ventura, N. S. D. G. A. S. (2005). Amyloid fibril formation by bovine cytochrome c. *Spectroscopy* 19:6. doi: 10.1155/2005/104348
- Wienk, H., Tomaselli, S., Bernard, C., Spurio, R., Picone, D., Gualerzi, C. O., et al. (2005). Solution structure of the

- C1-subdomain of *Bacillus stearothermophilus* translation initiation factor IF2. *Protein Sci.* 14, 2461–2468. doi: 10.1110/ps.051531305
- Yuan, A. H., Garrity, S. J., Nako, E., and Hochschild, A. (2014). Prion propagation can occur in a prokaryote and requires the ClpB chaperone. *Elife* 3:e02949. doi: 10.7554/eLife.02949
- Zhou, Y., Smith, D., Leong, B. J., Brännström, K., Almqvist, F., and Chapman, M. R. (2012). Promiscuous cross-seeding between bacterial amyloids promotes interspecies biofilms. *J. Biol. Chem.* 287, 35092–35103. doi: 10.1074/jbc.M112.383737

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Iglesias, de Groot and Ventura. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# The Rho Termination Factor of *Clostridium botulinum* Contains a Prion-Like Domain with a Highly Amyloidogenic Core

Irantzu Pallarès\*, Valentin Iglesias and Salvador Ventura\*

Institut de Biotecnologia i Biomedicina and Departament de Bioquímica i Biologia Molecular, Universitat Autònoma de Barcelona, Barcelona, Spain

## OPEN ACCESS

### Edited by:

Marc Bramkamp,  
Ludwig-Maximilians-Universität  
München, Germany

### Reviewed by:

Dennis Claessen,  
Leiden University, Netherlands  
Marina Lotti,  
University of Milano-Bicocca, Italy

### \*Correspondence:

Salvador Ventura  
salvador.ventura@uab.cat;  
Irantzu Pallarès  
irantzu.pallares@uab.cat

### Specialty section:

This article was submitted to  
Microbial Physiology and Metabolism,  
a section of the journal  
Frontiers in Microbiology

**Received:** 13 November 2015

**Accepted:** 16 December 2015

**Published:** 07 January 2016

### Citation:

Pallarès I, Iglesias V and Ventura S  
(2016) The Rho Termination Factor  
of *Clostridium botulinum* Contains  
a Prion-Like Domain with a Highly  
Amyloidogenic Core.  
Front. Microbiol. 6:1516.  
doi: 10.3389/fmicb.2015.01516

Prion-like proteins can switch between a soluble intrinsically disordered conformation and a highly ordered amyloid assembly. This conformational promiscuity is encoded in specific sequence regions, known as prion domains (PrDs). Prions are best known as the causative factors of neurological diseases in mammals. However, bioinformatics analyses reveal that proteins bearing PrDs are present in all kingdoms of life, including bacteria, thus supporting the idea that they serve conserved beneficial cellular functions. Despite the proportion of predicted prion-like proteins in bacterial proteomes is generally low, pathogenic species seem to have a higher prionic load, suggesting that these malleable proteins may favor pathogenic traits. In the present work, we performed a stringent computational analysis of the *Clostridium botulinum* pathogen proteome in the search for prion-like proteins. A total of 54 candidates were predicted for this anaerobic bacterium, including the transcription termination Rho factor. This RNA-binding protein has been shown to play a crucial role in bacterial adaptation to changing environments. We show here that the predicted disordered PrD domain of this RNA-binding protein contains an inner, highly polar, asparagine-rich short sequence able to spontaneously self-assemble into amyloid-like structures, bearing thus the potential to induce a Rho factor conformational switch that might rewire gene expression in response to environmental conditions.

**Keywords:** prion, bacteria, *Clostridium*, protein aggregation, amyloid

## INTRODUCTION

Amyloid forming proteins are found in all kingdoms of life, from Bacteria to Animalia (Fowler et al., 2007; Eichner and Radford, 2011; Sanchez de Groot et al., 2012). Although amyloid formation is associated with the onset of debilitating human disorders such as Alzheimer's, or Parkinson's (Maries et al., 2003; Stohr et al., 2012), the amyloid fold is also exploited for evolutionary selected biological functions by diverse species, including humans (Chiti and Dobson, 2006; Furukawa and Nukina, 2013). Prions are a particular type of amyloids that can switch between soluble and self-templating aggregated states. In the so-called functional prions, this property is used to perform important functions, acting as epigenetic elements and supporting beneficial roles in cell physiology (Newby and Lindquist, 2013).

The conformational duality of prion-like proteins resides in structurally independent, low complexity, prion-forming domains (PrDs), usually enriched in asparagine (N) and glutamine (Q) residues (Dorsman et al., 2002; Fandrich and Dobson, 2002; Halfmann et al., 2011). This composition endorses the domains with intrinsic structural disorder, which enables self-assembly without a requirement for conformational unfolding (Fuxreiter, 2012; Malinowska et al., 2013). Much research has gone in the recent years into uncovering how prion propensities are encoded in protein sequences (Alberti et al., 2009; Toombs et al., 2010; MacLea et al., 2015; Sabate et al., 2015a) and several algorithms exploit this knowledge to identify new putative prion proteins (Toombs et al., 2012; Espinosa Angarica et al., 2013, 2014; Lancaster et al., 2014; Sabate et al., 2015b; Zambrano et al., 2015). The high-throughput analysis of proteomes using these programs has led to the identification of thousands of new potential prion-like proteins in organisms belonging to all taxonomic subdivisions (Espinosa Angarica et al., 2013). The results show that, in general, the number of prions per genome is low, less than 1% of the complete proteome (Michelitsch and Weissman, 2000; Harrison and Gerstein, 2003; Espinosa Angarica et al., 2013). Ontology analysis indicates that PrD-containing proteins are associated with a great variety of physiological functions, supporting prion-like proteins acting as beneficial elements for organisms.

In a previous work, we have used our algorithm PrionScan to analyze 839 different bacteria proteomes, detecting 2200 putative prions in these organisms (Espinosa Angarica et al., 2013, 2014). Interestingly, we found a special enrichment in proteins containing PrDs in pathogenic bacteria (Espinosa Angarica et al., 2013). A significant number of these proteins are DNA or RNA binding proteins (Iglesias et al., 2015), which might be involved in host induced bacteria gene expression plasticity, recapitulating the response of yeast transcription factors with prion-like properties in front of environmental fluctuations (Alberti et al., 2009; Malinowska et al., 2013; Newby and Lindquist, 2013).

PrionScan identifies PrDs on the basis of their amino acid compositional similitude to *bona fide* yeast prions, which results in a very fast algorithm useful to scan very large databases, as those corresponding to a complete taxon (Espinosa Angarica et al., 2014). However, this speed comes at the cost of a lower specificity in the predictions, when compared with competing algorithms like PAPA (Toombs et al., 2012) and pWALTZ (Sabate et al., 2015b). PAPA exploits the compositional bias of PrDs to identify these domains in protein sequences using a experimentally derived amino acid prion propensity scale (Toombs et al., 2010), whereas pWALTZ implements a totally different concept, since it assumes that it is the presence and potency of specific short amyloid-prone sequences that occur within intrinsically disordered Q/N-rich regions that account for prion induction (Sabate et al., 2015b).

Here, we combined PAPA and pWALTZ algorithms to get highly specific PrDs predictions in the proteome of *Clostridium botulinum* (*C. botulinum*). This bacterium is widely spread in the environment, with reservoirs both in soil and water sediments and is a well-known pathogen that affects animals and humans worldwide (Espelund and Klaveness, 2014). This approach led

us to the identification of 54 putative prion proteins. Among them, it outstands the transcription termination factor Rho (Rho) (Richardson, 1990, 1996; Boudvillain et al., 2013). We show here that its predicted PrD contains a highly polar, N-rich, short sequence stretch able to form amyloid-like fibrils, which might endorse this RNA-binding protein with the ability to shift from soluble to aggregated states in order to modulate its functionality.

## MATERIALS AND METHODS

### Prion Forming Domains identification in Bacteria

The *C. botulinum* E1 str. 'BoNT E Beluga' proteome dataset was downloaded from Uniprot (release 2015\_05) and scanned for PrDs using PAPA (Toombs et al., 2012) with the default parameters, which includes the disorder prediction algorithm FoldIndex (Prilusky et al., 2005). From the initial 3678 proteins in the proteome, 63 prion-like candidates were identified. Their putative prion forming domains were further evaluated with pWALTZ (Sabate et al., 2015b) using the default parameters to identify those domains containing a putative amyloid core, which resulted in 54 final positive predictions.

### *Clostridium botulinum* PrD Peptide Preparation

A peptide with the sequence NNNNSNFNNNSNNSSFNNSN, corresponding to the predicted amyloid core in the PrD of *C. botulinum* Rho factor, was purchased from CASLO ApS. Stock solutions were prepared at 5 mM in DMSO and stored at  $-80^{\circ}\text{C}$ . For analysis, the peptide was diluted to 25, 50, and 100  $\mu\text{M}$  in PBS buffer.

### Aggregation Assays

Aggregation of initial soluble species was monitored by following the transition from non-aggregated to aggregated states by measuring light scattering at 360 nm in 25, 50, and 100  $\mu\text{M}$  peptide samples at  $25^{\circ}\text{C}$ . Light scattering changes were evaluated for samples incubated during 4, 48, and 120 h.

### Binding to Amyloid Dyes

The binding of 25  $\mu\text{M}$  of Thioflavin-T (Th-T) to Rho peptide was recorded using a Cary Eclipse Spectrofluorometer (Varian, Palo Alto, CA, USA) with an excitation wavelength of 440 nm and emission range from 460 to 600 nm at  $25^{\circ}\text{C}$  in PBS buffer. Spectra were recorded after 2 min of equilibration, and solutions without peptide were used as negative controls. Excitation and emission slit widths of 10 nm were used. For the staining assays with Thioflavin-S (Th-S), Rho peptide aggregates were incubated for 1 h in the presence of 125  $\mu\text{M}$  of dye. After centrifugation ( $14000 \times g$  for 5 min), the precipitated fraction was washed twice with PBS and placed on a microscope slide and sealed. Images of Rho peptide fibrils bound to Th-S were obtained at 40-fold magnification under UV light or using phase contrast in Leica fluorescence microscope (Leica DMRB, Heidelberg, Germany).



Congo red (CR) interaction with Rho peptide aggregates was tested using a Cary100 UV/Vis spectrophotometer (Varian, Palo Alto, CA, USA) by recording the absorbance spectra from 400 to 675 nm using a matched pair of quartz cuvettes of 1 cm optical length placed in a thermostated cell holder at 25°C. Final CR and peptide concentrations were 5  $\mu$ M in PBS buffer. In order to detect the typical amyloid band at  $\sim$ 541 nm, differential CR spectra in the presence and absence of peptide were recorded.

## Bis-ANS Binding

Binding of 4,4'-bis (1-anilino-naphthalene 8-sulphonat) (bis-ANS) to Rho peptide was evaluated by registering bis-ANS fluorescence between 400 and 700 nm after excitation at 370 nm on a Cary Eclipse Spectrofluorometer (Varian, Palo Alto, CA, USA). Spectra were recorded at 25°C in PBS buffer, final peptide and dye concentrations were 10 and 1  $\mu$ M, respectively. Excitation and emission slit widths of 10 nm were used.

## Aggregation Kinetics and Seeding Assays

Rho peptide aggregation was monitored by quantification of the changes in relative Th-T fluorescence at 475 nm when exciting at 440 nm along time. In the seeding assay, a solution of 0.1% (w/w) preformed fibrils was added at the beginning of the reaction. All experiments were carried out in PBS buffer under agitation ( $\sim$ 750 rpm with micro-stir bars) at 25°C with an initial soluble peptide concentration of 100  $\mu$ M.

## Secondary Structure Determination

ATR FT-IR spectroscopy analysis of Rho peptide aggregates was performed using a Bruker Tensor FT-IR Spectrometer (Bruker Optics, Berlin, Germany) with a Golden Gate MKII ATR accessory. Each spectrum consists of 16 independent scans, measured at spectral resolution of 1  $\text{cm}^{-1}$ . Infrared spectra between 1725 and 1575  $\text{cm}^{-1}$  were fitted through overlapping Gaussian curves, and the amplitude, and area for each Gaussian function were calculated employing the non-linear peak-fitting program (PeakFit package, Systat Software, San Jose, CA, USA).

## Transmission Electron Microscopy (TEM)

For negative staining, samples of Rho peptide incubated at 25°C for 4, 48, and 120 h were placed onto carbon-coated copper grids and left to stand for 5 min. The grids were washed with distilled water and stained with 2% (w/v) uranyl acetate for 2 min. Micrographs were recorded in a JEM-1400 (JEOL, Japan) transmission electron microscope (TEM) operated at 80-kV accelerating voltage.

# RESULTS

## Identifying Prion-Like Domains on the Pathogenic Bacteria *Clostridium botulinum*

Recent bioinformatics screenings revealed multiple prion candidates in bacteria, especially in pathogenic species (Espinosa

Angarica et al., 2013; Iglesias et al., 2015). In light of these data, we focused here on the Gram-positive, anaerobic bacterium *C. botulinum*, given its involvement in a number of pathological processes (Swaminathan and Eswaramoorthy, 2000; Kumaran et al., 2009; Rossetto et al., 2014). The analysis of the 3678 protein sequences in *C. botulinum* proteome was initially performed with PAPA (Toombs et al., 2012) and further refined with pWALTZ (Sabate et al., 2015b). Both PAPA and pWALTZ algorithms were trained on top of yeast prions; however, they are based on radically different concepts, a suitable composition of the PrD and the presence of an amyloid core embedded in it, respectively. This ensures that sequences that pass the two thresholds should have properties resembling previously verified yeast prions. According to their respective scores, 54 proteins, corresponding to 1.5% of the proteome, were identified as containing PrDs in *C. botulinum* (Supplementary Table S1). Ontology analysis indicates that the putative prion-like dataset is enriched in biological processes related to the cell wall dynamics. However, we also found proteins relevant in bacterial processes such as invasion, virulence and nucleotide metabolism (Supplementary Table S1).

We analyzed the role of the structural Pfam domains linked to the detected *C. botulinum* PrD-containing proteins. As expected, the biggest cluster of Pfam families is associated with cell wall dynamics, with 19 out of the 41 annotated putative prions having a cell wall binding repetition domain. Among the proteins in that cluster we can find a glycosyl transferase (C5UUW9), which is a glycan synthesis effector and a clear example of proteins involved in cell wall rearrangement, with a structure combining two different functional domains, one glucoamylase domain and two glycotransferase domains. The cell shape protein MreC (C5UR99), is another relevant protein in that cluster, which is thought to couple the internal bacterial cytoskeleton to the extracellular cell wall synthesizing complexes; interestingly, it is a protein that associates with penicillin-binding proteins and guides the insertion of newly synthesized cell wall precursors (Divakaruni et al., 2005; Tavares et al., 2015). Yet another protein in this subset is Brachyurin (C5UXB1), a cell-wall associated protein that contains two N-cadherin domains in its structure, suggesting a role in cell-cell contact, adhesion and biofilm formation (Anantharaman and Aravind, 2010). The second most abundant group of Pfam domain families is associated to invasion and virulence processes. This group includes proteins associated with encapsulation, sporulation and toxins. CotH (C5UUU1) and the spore cortex-lytic enzyme (C5U536) are proteins required either for spore coat formation (Zilhao et al., 1999) or for spore germination, thus facilitating *C. botulinum* aerial growth, surface attachment and pathogenesis. We also find a L,D-transpeptidase (C5UVDO), which cross-links peptidoglycan in presence of antibiotic drugs that block regular effectors (Biarrotte-Sorin et al., 2006; Magnet et al., 2007) allowing the bacteria to overcome classical  $\beta$ -lactams antibiotic blockage. We highlight in this cluster the presence of the Botulinum neurotoxin non-toxic-non-hemagglutinin component (NTNH). The neurotoxin complex is composed of NTNH, the toxin BoNT, hemagglutinin (HA) and associated subcomponent proteins and RNAs (Wren, 1991). It has been proposed that NTNH confers

protection against the harsh conditions the toxin faces in the digestive tract (Sugawara et al., 2014). The third group contains proteins with domains involved in nucleotide binding, such as the Transcription termination factor Rho (C5URV5) involved in transcription regulation and the Ribonucleoside-diphosphate reductase (C5UTH8) that is implicated in DNA replication. Other relevant putative prion-like proteins that cannot be clustered in the former groups but merit attention are StbA (C5UUD6), a putative Hsp70 family chaperone which has been seen to stabilize plasmids and control their number in *Escherichia coli* (*E. coli*) (Bork et al., 1992; Guynet et al., 2011) and a putative ggdef domain protein (C5UR68), with two relevant functional domains, a tetratricopeptide domain, involved in scaffold formation to mediate protein interactions and the assembly of multiprotein complexes and a GGDEF domain related with the synthesis of cyclic di-GMP and involved in the regulation of processes such as biofilm formation, motility and cell differentiation.

## Rho Factor Exhibits a Predicted PrD Containing a Putative N-rich Amyloid Core

Because many of the prion-like polypeptides identified in eukaryotes are RNA binding proteins (King et al., 2012; Kim et al., 2013), we focused our attention in the transcription termination factor Rho (Rho). Rho is required for the factor-dependent transcription termination by an RNA polymerase in prokaryotes and is essential for the viability of the cell (Richardson, 1996; Cardinale et al., 2008; Washburn and Gottesman, 2011; Krishna Leela et al., 2013). Recent studies indicate that besides being a housekeeping gene, Rho can function as a gene regulator and participates in the control of prophage maintenance in bacterial genomes (Boudvillain et al., 2013; Menouni et al., 2013). Accordingly, it plays a critical role in determining what proteins are present in the cell, in what amounts and thus modulating the organism's phenotype.

PAPA predicts an 80 residues long PrD close to the Rho factor N-terminus, which resides in a longer intrinsically disordered region, as predicted with FoldIndex (Prilusky et al., 2005) (Figure 1). pWALTZ predicts the presence of three overlapping 21 residues long amyloid stretches comprising residues 90–110, 92–112, and 93–113 inside the identified Rho PrD (Figure 1). When we analyzed the location of structured, unstructured and PrD regions in Rho factor, we found that, overall, its topology resembles that observed in certain *bona fide* yeast prions, like Ure2p (Figure 1). Globular domains in prion-like proteins are responsible for their biological function. The Rho factor consists of six identical subunits, each containing three functional domains. The RNA binding site has been localized to the N-terminal portion of the protein, the ATP binding site is located in the central portion of the primary sequence, and subunit interaction sites have been proposed to reside in the C-terminal region (Geiselman et al., 1993; Bogden et al., 1999). The interaction of Rho with RNA is critical to all the activities of the protein. Thus RNA binding is required to activate the RNA-dependent ATPase activity of Rho. The predicted PrD and the

RNA binding domain are contiguous in Rho, a topology that is also found in many eukaryotic prion-like proteins (King et al., 2012; Espinosa Angarica et al., 2013; Malinowska et al., 2013; Navarro et al., 2015).

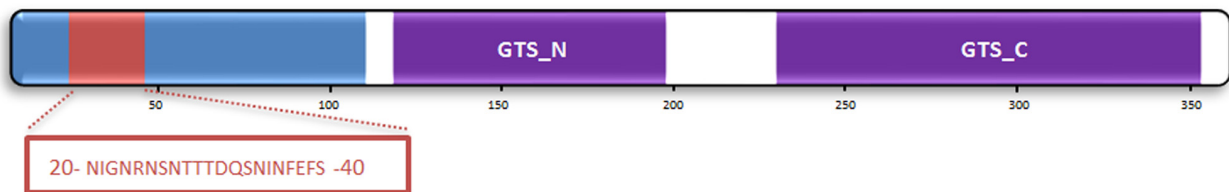
The widely accepted “amyloid-stretch” hypothesis proposes that the amyloid potential of amyloidogenic proteins resides in short, highly amyloidogenic regions that act by nucleating the aggregation reaction (Ventura et al., 2004; Esteras-Chopo et al., 2005). We have recently proposed that this view also applies for prion-like proteins, explaining why all known prions adopt amyloid conformations in their propagative state (Sabate et al., 2015a). In order to assess if this is the case of Rho factor, we experimentally characterized the predicted central amyloid core of the prion domain (cPrD) using a synthetic peptide corresponding to sequence 92–NNNNSNFNNNSNNSSFNNSN-112, with a 67% N content. Despite pWALTZ, which is specially intended to analyze PrDs, predicts that this N-rich sequence would endorse the surrounding PrD with significant amyloidogenic potential, well-contrasted aggregation predictors like AGGRESCAN (Conchillo-Solé et al., 2007), TANGO (Fernandez-Escamilla et al., 2004) or FoldAmyloid (Garbuzynskiy et al., 2010) fail to predict any aggregation-prone region in this peptide and, indeed, they predict it to be soluble.

## Rho cPrD Forms $\beta$ -sheet Enriched Aggregates

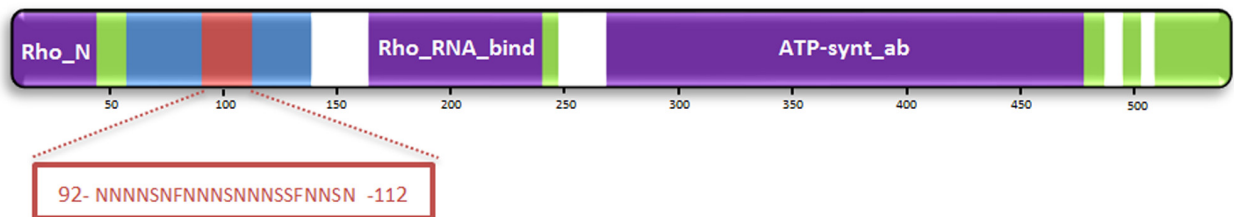
As a first step to experimentally characterize the selected cPrD we analyzed its *in vitro* aggregation properties. Rho cPrD was incubated at 25, 50, and 100  $\mu$ M at 25°C for 4, 48, and 120 h and aggregation from its initially soluble state was evaluated using synchronous light scattering (Figure 2). A concentration dependent scattering signal is observed after 4 h. However, the signal corresponding to the 25 and 50  $\mu$ M solutions does not evolve significantly with time, whereas the scattering signal of the 100  $\mu$ M peptide solution steadily increases to attain a maximum after 120 h (Figure 2C). Accordingly, unless otherwise indicated, all subsequent experiments were performed with the peptide at a concentration of 100  $\mu$ M.

For most amyloids, the self-assembly reaction depends on the formation of intra-chain hydrophobic clusters (Hills and Brooks, 2007). However, Rho cPrD is a highly polar peptide, with less than 10% of its residues being hydrophobic. We explored the presence of exposed hydrophobic clusters in the aggregates formed by Rho cPrD at different times by measuring their binding to bis-ANS (Figure 3), a dye that increases its fluorescence emission upon interaction with these regions (Gohlke, 1972; de Groot et al., 2007; Zhou et al., 2012). The bis-ANS fluorescence emission maximum blue-shifts from 530 nm, in the absence of peptide, to 509 nm in the presence of the peptide after 4 h. This spectral change is even more pronounced after 48 h, even if the global intensity decreases. Bis-ANS fluorescence emission attains a maximum at 120 h, with its spectral maximum blue-shifted to 490 nm. These data clearly indicate that the two phenylalanine (F) residues in Rho cPrD play an important role on its aggregation reaction, leading to the formation of strong hydrophobic patches in the final aggregates.

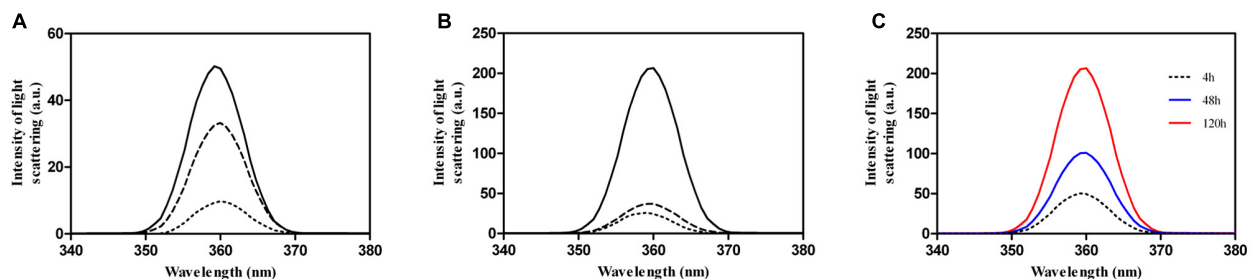
### Transcription regulator URE2



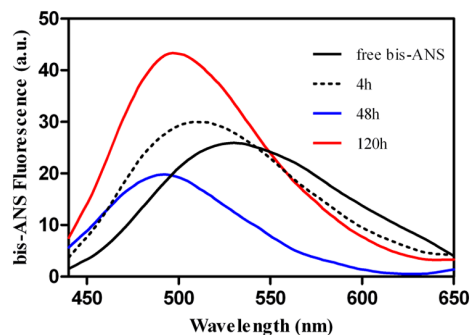
### Transcription termination factor Rho



**FIGURE 1 |** Prion domains and amyloid cores as predicted in the sequences of Rho from *Clostridium botulinum* and the prion protein URE2 from *Saccharomyces cerevisiae*. Identified Pfam domains, disordered regions, predicted PrDs and amyloid cores are shown in violet, green, blue, and red, respectively.



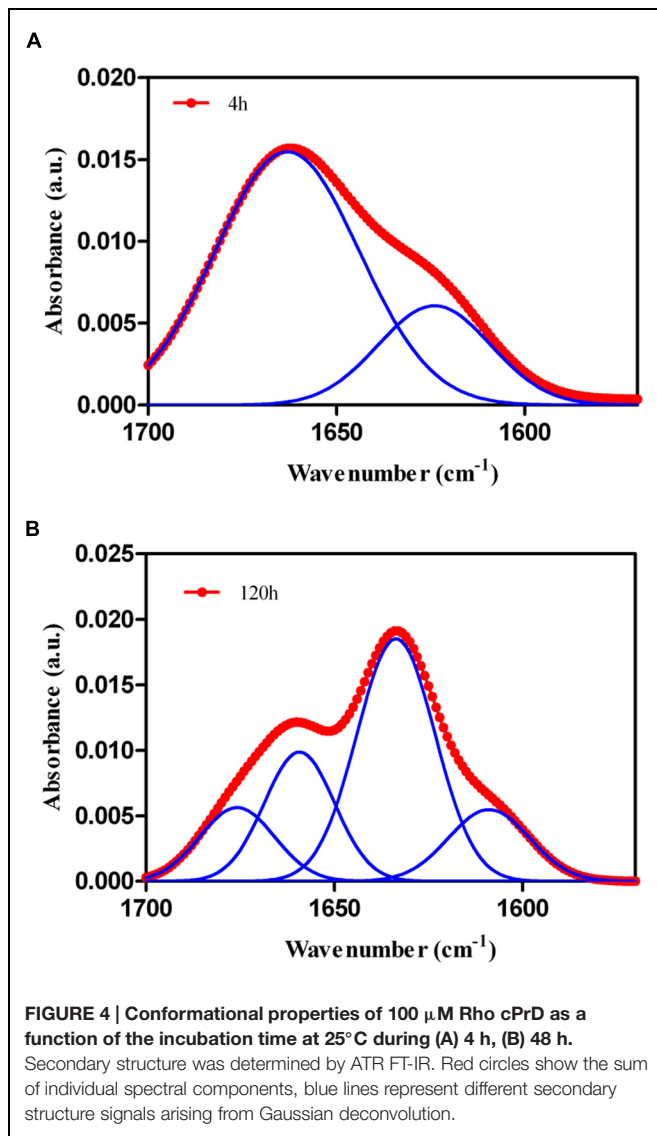
**FIGURE 2 |** Aggregation of Rho cPrD as a function of the concentration and the incubation time. Aggregation changes were monitored by light scattering at different Rho cPrD concentrations 25 (dotted line), 50 (dashed line), and 100  $\mu$ M (solid line) incubated at 25°C during (A) 4 h, (B) 120 h. The scattering at 4, 48, and 120 h for the 100  $\mu$ M peptide solution is shown in (C).



**FIGURE 3 |** Binding of bis-ANS to 100  $\mu$ M Rho cPrD as a function of the incubation time at 25°C.

The aggregation of proteins into amyloid fibrils results in the formation of intermolecular  $\beta$ -sheets (Nelson et al., 2005). To get insights into the secondary structure content of the assemblies

formed by Rho cPrD, we analyzed the amide I region of the FTIR spectrum (1700–1600  $\text{cm}^{-1}$ ) (Figure 4). This region corresponds to the absorption of the carbonyl peptide bond group of the protein main chain and is a sensitive marker of the protein secondary structure. Examination of the secondary structure of Rho cPrD peptide by deconvoluted FTIR spectra allow us to assign the individual secondary structure elements and their relative contribution to the main absorbance signal at the beginning (4 h) and ending (120 h) of the aggregation reaction (Figure 4; Table 1). After 4 h of incubation the spectrum of Rho cPrD is dominated by a band at 1663  $\text{cm}^{-1}$ , corresponding to disordered structures, accounting for 74% of the total area. However, the presence of an inter-molecular  $\beta$ -sheet component at 1624  $\text{cm}^{-1}$  is already observable at this time point. At the end of the reaction (120 h), the FTIR spectrum of Rho cPrD is dominated by a band at 1633  $\text{cm}^{-1}$  attributable to  $\beta$ -sheet conformations. At this stage, the low frequency  $\beta$ -sheet components at 1607 and 1633  $\text{cm}^{-1}$  together with the high frequency  $\beta$ -sheet component at 1676  $\text{cm}^{-1}$  account for 77% of the total area, with disordered



conformations contributing only 23% of the signal. These spectral properties are compatible with the assembly of Rho cPrD into a highly  $\beta$ -sheet enriched amyloid-like structure.

## Rho cPrD Self-Assembles into Amyloid Fibrils

We used the amyloid-specific dyes CR, Th-T and Th-S to confirm that the detected  $\beta$ -sheet enriched aggregates were organized into amyloid-like suprastructures. The absorbance of CR increases and the spectrum maximum red shifts to 505–510 nm in the presence of peptide aggregates formed at 100  $\mu$ M after 120 h of incubation at 25°C (Figure 5A). This spectral change corresponds to that promoted by different amyloid proteins in the aggregated state (Klunk et al., 1989). Moreover, the difference spectrum between the dye in the presence and absence of aggregated peptide allows detecting the characteristic amyloid band at  $\sim$ 541 nm (Figure 5B). The binding of Rho cPrD to CR at early time points is significantly lower.

Thioflavin-T fluorescence emission is enhanced in the presence of amyloid fibrils (LeVine, 1993; Sabate et al., 2013). The same behavior is observed upon incubation of Th-T with Rho cPrD (Figure 5C). In good agreement with light scattering signals, Th-T binding to peptide solutions increases with incubation time, the Th-T fluorescence at the 480 nm spectral maximum increasing 80-fold at 120 h. Furthermore, binding of Th-S to 120 h aggregates could be visualized by fluorescence microscopy (Figure 5D). Areas rich in fibrous material were stained with Th-S to yield green–yellow fluorescence against a dark background.

The dye binding results indicate that incubated Rho cPrD solutions contain detectable amounts of amyloid-like structure. To confirm this extent, the morphological features of the peptide assemblies in these samples were analyzed using TEM. As shown in Figure 6, we detect the presence of protein aggregates in all cases. Nevertheless, in good agreement with spectroscopic data, the size and morphology of the aggregates are significantly different. The peptide incubated for 4 h forms short, poorly ordered protofibrillar assemblies. These assemblies coexist with fibrillar structures at the 48 h, whereas only mature fibrils with a typical amyloid-like morphology are observed at the 120 h.

Seeded protein aggregation is a well-established mechanism for *in vivo* amyloid fibril formation and underlies prion propagation (Caughey, 2001; Wickner et al., 2001). The nucleation step of the amyloid assembly is shortened in the presence of preformed amyloid fibrils of the same protein that can act as nuclei for the subsequent polymerization reaction (Jarrett and Lansbury, 1992). Specific and short aggregation-prone regions have been shown to play a crucial role in this process (Pastor et al., 2007; Sabate et al., 2012). To test whether preformed Rho cPrD fibrils can seed the aggregation of the correspondent soluble peptide, we followed the aggregation kinetics of the peptide at 100  $\mu$ M in the presence and absence of 0.1% (w/w) preformed fibrils. As shown in Figure 7, the presence of fibrils strongly accelerated the formation of Th-T positive assemblies, raising the possibility that such specific amyloid-promoting interactions could also occur in the context of the complete Rho factor protein.

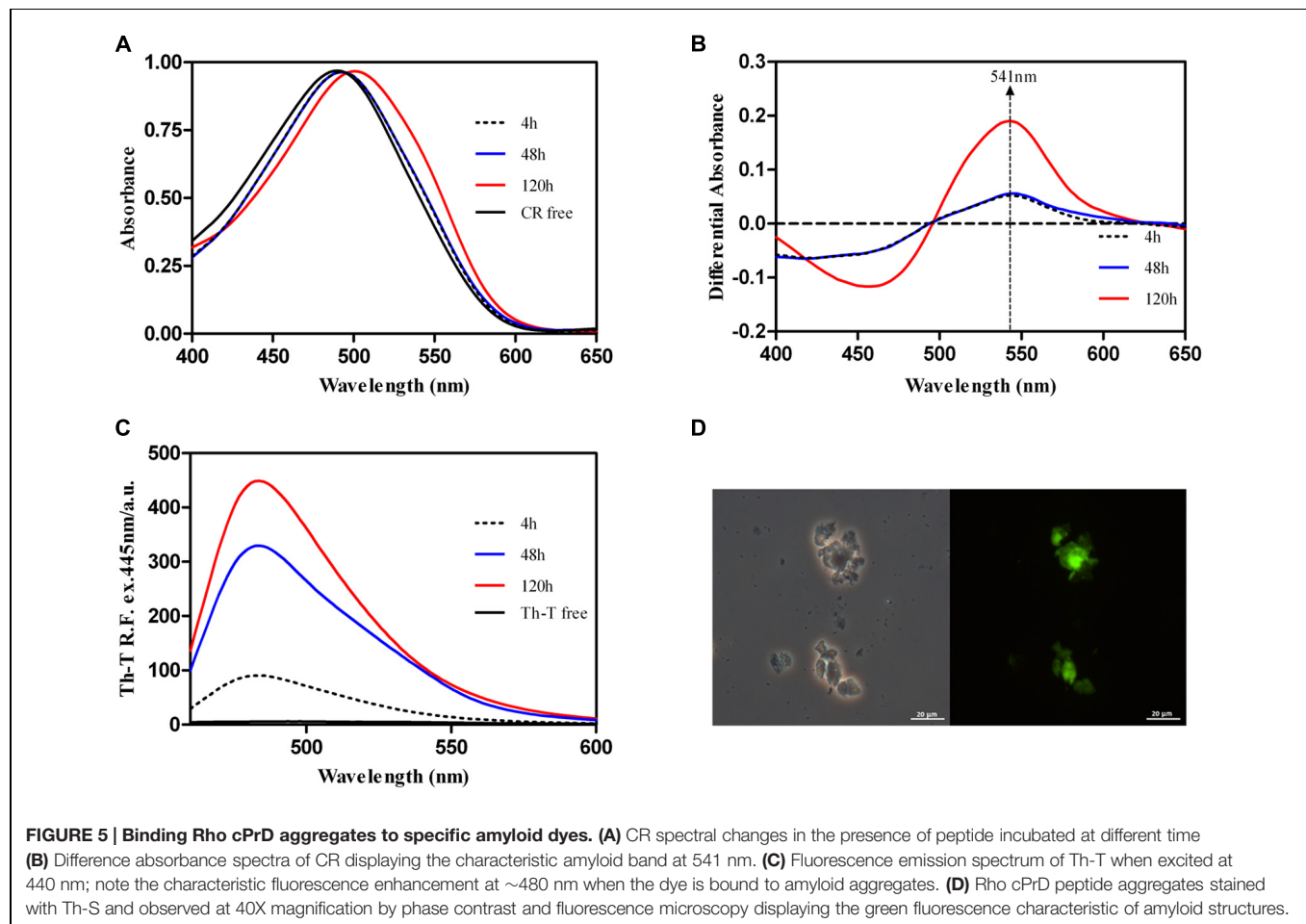
## DISCUSSION

Prion-like proteins were initially thought to be restricted to mammals, resulting in transmissible pathologies (Aguzzi and Weissmann, 1998). Later on, the discovery of yeast prions (Wickner, 1994; Du et al., 2008; Patel et al., 2009; Rogoza et al., 2010) and more recently of prion-like proteins in multicellular eukaryotes, from snail to human (Maji et al., 2009; Heinrich and Lindquist, 2011; Majumdar et al., 2012; Tariq et al., 2013; Cai and Chen, 2014), suggest that prion-like mechanisms would sustain evolutionary conserved functions in eukaryotic kingdoms. Despite no bacterial prion-like protein has been characterized so far, computational predictions support the existence of a significant number of proteins with potential prion-like properties in bacterial proteomes (Espinosa Angarica et al., 2014; Iglesias et al., 2015). This is not surprising, since bacterial



**TABLE 1 | Assignment of secondary structure components of Rho cPrD peptide in the amide I region of the FTIR spectra.**

Rho PrD 4 h			Rho PrD 120 h		
Band (cm <sup>-1</sup> )	Area (%)	Secondary structure	Band (cm <sup>-1</sup> )	Area (%)	Secondary structure
1624	26	β-sheet	1607	15	β-sheet
1663	74	Loops/turns	1633	48	β-sheet
			1661	23	Loops/turns
			1676	14	β-sheet

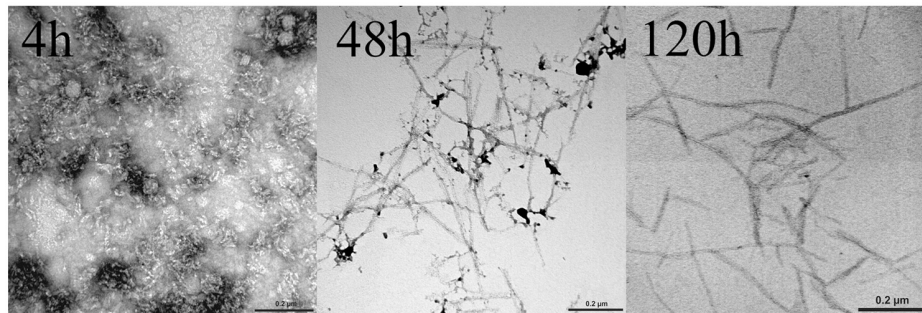


cells have been shown to support the formation of prion-like conformations of yeast prions (Sabaté et al., 2009; Garrity et al., 2010; Espargaró et al., 2012) and, more importantly, to propagate them for over a hundred generations, even when the cells can no longer make the protein that serves as the trigger for the initial conversion (Yuan et al., 2014), which suggests that functional prion-like mechanisms might be more ancient than previously thought (Desantis et al., 2012).

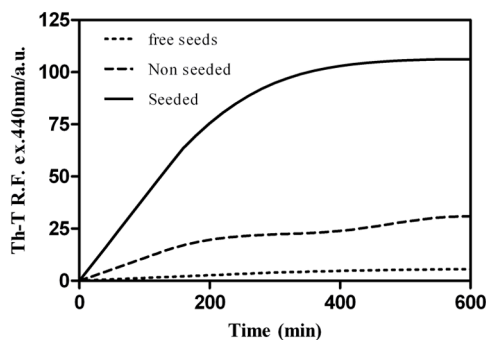
As a trend, prion-like sequences are predicted to be less abundant in bacteria than in eukaryotes (Espinosa Angarica et al., 2013), but, interestingly, pathogenic species seem to have a higher prion load than non-pathogenic ones. An exciting possibility is that these sequences represent a bet-hedging mechanism for pathogens, as suggested recently for yeast prions

(Newby and Lindquist, 2013). These mechanisms are used to diversify microbial phenotypes. In fluctuating environments this allows a fraction of the population to survive in conditions when most would perish. This mechanism would permit certain cells to persist in strenuous environments like in the presence of antibiotics or to escape the immunogenic response, saving the population from extinction. Shuffling the states of multiple prion-like proteins would allow rapid phenotypic diversification.

Here, we addressed the presence of potential prion-like proteins in the proteome of the pathogen *C. botulinum* using a stringent approach in which both a long region displaying amino acid compositional similitude to *bona fide* prions (Toombs et al., 2010, 2012) and the presence of a specific nucleating sequence



**FIGURE 6 |** Transmission electron micrographs of Rho cPrD after 4, 48, and 120 h of incubation at 25°C.



**FIGURE 7 |** Aggregation kinetics of Rho cPrD. The aggregation reactions of 100  $\mu$ M Rho cPrD were carried out under agitation at 25°C. 0.1  $\mu$ M of *in vitro* formed fibrils (representing 0.1% of the final peptide concentration) were used for the seeding assay. The formation of Rho cPrD amyloid fibrils was accelerated in the presence of pre-aggregated peptide.

inside it (Sabate et al., 2015b) should be present for a protein to be considered prion-like. This approach rendered a total of 54 candidates. Interestingly, the set of candidates is enriched in proteins that play a structural role and are linked to essential processes as cell wall metabolism or cellular shape maintenance. Although a more exhaustive analysis of these proteins is necessary, the data point to a possible relationship between the identified proteins and biofilm formation, which would confer a protecting strategy and facilitate the attachment of the bacteria to different surfaces. Indeed, the biofilms of a number of bacterial species have been shown to contain proteins in an amyloid conformation (Romero and Kolter, 2014). PrDs associated with proteins involved in survival and virulence were also found in *C. botulinum*. Sporulation and toxin production are powerful strategies that facilitate the invasion of new environments and bacterial survival in adverse conditions. In this context, proteins involved in spore formation and degradation, the degradation-protector NTNH in Botulinum toxin and the cell-wall cross-linker L,D-transpeptidase develop non-essential functions, but facilitate the bacteria to remove toxic agents and evade the action of antibiotics or from harsh natural environmental conditions and toxic compounds (Biarrotte-Sorin et al., 2006; Magnet et al., 2007).

A significant number of the prion-like sequences predicted in the human proteome correspond to RNA binding proteins (King et al., 2012; Espinosa Angarica et al., 2013; Malinowska et al., 2013), which fits well with the fact that several experimentally determined *genuine* prion-like proteins, including Ure2, Swi1, Spf1, Cyc8, and Mot3 in yeast (Wickner, 1994; Du et al., 2008; Alberti et al., 2009; Patel et al., 2009; Rogoza et al., 2010) and *Drosophila melanogaster*'s GAGAFactor (Tariq et al., 2013) act as transcriptional regulators. This is also the function of the Rho factor in *C. botulinum*, for which we predict the existence of a highly scoring putative PrD at the N-terminus, adjacent to the RNA binding domain. It has been suggested that, in the prion-like state, transcriptional regulators may alter gene expression by creating diffusion barriers that restrict protein movement toward specific subcellular locations, by decreasing the effective concentration of the freely available pool of protein, or, on the contrary, by increasing the effective concentration in a certain location; this might result in enough functional diversity to create phenotypic divergence (Si, 2015). Interestingly enough, recent works have shown that Rho inhibition allows prophage maintenance, as a strategy to keep beneficial prophage genes, while silencing those likely to be deleterious (Cardinale et al., 2008; Menouni et al., 2013). Importantly, the pathogenic trait in *C. botulinum*, the botulinum neurotoxin, is mainly linked to a large plasmidome consisting of plasmids and circular prophages (Skarin and Segerman, 2014). Indeed, it has been recently shown that, in *E. coli*, mutations promoting adaptive properties, such as adaptation to thermal stress, converge to cluster either in the RNA polymerase complex or the termination factor Rho (Tenaillon et al., 2012; Rodriguez-Verdugo et al., 2014; Hug and Gaut, 2015). When we analyzed the mutations reported to occur specifically in Rho factor with our aggregation prediction algorithm AGGRESCAN (Conchillo-Solé et al., 2007), we found out that 72% of them endorse the terminator factor with increased aggregation-propensity, thus suggesting a link between the self-assembly of Rho and adaptation to changing environments.

We provide here strong evidence that detected PrD in Rho factor contains a short amyloid-like segment with the ability to potentially nucleate the Rho factor PrD assembly; however, it remains to be demonstrated if, in the case it occurs *in vivo*, the

reaction would exhibit the reversibility required for considering this protein a *bona fide* prion.

In contrast to pWALTZ, conventional aggregation prediction algorithms do not capture the amyloidogenic potential of Rho cPrD. Because these latter algorithms usually display good accuracy when predicting the core of disease-linked amyloids (Sabate et al., 2015b), this suggests that the principles underlying their aggregation and that of Rho cPrD are somehow different. Indeed, the amyloid core of pathogenic proteins is usually very hydrophobic, whereas 90% of Rho cPrD sequence is made of N and S, and therefore polar, with only two hydrophobic residues. On the one hand, while a certain amyloid nucleation capacity favoring a sufficiently high aggregation rate is absolutely necessary, the final amyloid aggregate in a prion-like protein should at the same time display brittleness, a property that facilitates propagation. On the other hand, the protein should remain in a soluble state under physiological conditions, while keeping a cryptic amyloid capacity that allows it to self-assemble only in selected conditions. Both requirements imply that, in contrast to most amyloids, in PrDs, the aggregation reaction should not be nucleated by an extremely strong, and highly hydrophobic, amyloid core. We have proposed that the role of N residues in PrDs and their amyloid cores is to endorse these sequences with a basal aggregation propensity, while allowing them at the same time to remain soluble and disordered in normal cellular conditions (Sabate et al., 2015a; Zambrano et al., 2015). In contrast, the few hydrophobic residues found in these cores, especially aromatic ones, would play a key role in the initial amyloid oligomerization steps. This seems to be true for Rho cPrD since its assembly into amyloid-like fibrils is accompanied by an increase in the presence of hydrophobic clusters, as monitored by bis-ANS binding. It is very likely that, as described for amyloid peptides from the Sup35 prion (Balbirnie et al., 2001; Diaz-Avalos et al., 2003; van Der Wel et al., 2006; Zheng et al., 2006), complete hydrogen bonding of its N and S residues would also contribute to sustain the mature amyloid structure.

Aggregation constraints the evolution of proteins and accordingly nature have evolved different strategies to minimize protein aggregation in sequences and structures. Essentially, mutations that result in an increase in aggregation propensity tend to be purged out from the population, especially when they occur in a disordered context, since they are exposed

to solvent, being this the reason that intrinsically disordered protein segments are in general, very soluble (Santner et al., 2012; Uversky, 2013, 2015; Graña-Montes et al., 2014). In this context, the inherent amyloid potential of Rho cPrD strongly suggests that this protein segment, and the surrounding predicted PrD, are conserved because they serve functional purposes in *C. botulinum*, in agreement with the general view that PrDs are important for protein–protein interactions and provide the flexibility required to self-organizing macromolecular assemblies in living cells (Malinowska et al., 2013; Iglesias et al., 2015).

## CONCLUSION

Overall, despite the reversibility and the functionality of *C. botulinum* Rho factor self-assembly should still be validated, this study provides a first proof for the existence of amyloidogenic sequences embedded in the recurrent putative PrD identified in transcription regulators of pathogenic bacteria, a property that is compatible with them being biological capacitors that might respond to environmental conditions rewiring gene expression.

## AUTHOR CONTRIBUTIONS

Conception/design of the work: SV, IP; performed the experiments: IP, VI; generated and analyzed the data: IP, VI, SV; drafting the work: IP, VI, SV; final approval of the manuscript to be published: SV, IP.

## ACKNOWLEDGMENT

This work was funded by the Spanish Ministry of Economy and Competitiveness BFU2013-44763-P to SV.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fmicb.2015.01516>

## REFERENCES

- Aguzz, A., and Weissmann, C. (1998). Spongiform encephalopathies. The prion's perplexing persistence. *Nature* 392, 763–764. doi: 10.1038/33812
- Alberti, S., Halfmann, R., King, O., Kapila, A., and Lindquist, S. (2009). A systematic survey identifies prions and illuminates sequence features of prionogenic proteins. *Cell* 137, 146–158. doi: 10.1016/j.cell.2009.02.044
- Anantharaman, V., and Aravind, L. (2010). Novel eukaryotic enzymes modifying cell-surface biopolymers. *Biol. Direct* 5:1. doi: 10.1186/1745-6150-5-1
- Balbirnie, M., Grothe, R., and Eisenberg, D. S. (2001). An amyloid-forming peptide from the yeast prion Sup35 reveals a dehydrated beta-sheet structure for amyloid. *Proc. Natl. Acad. Sci. U.S.A.* 98, 2375–2380. doi: 10.1073/pnas.041617698
- Biarrotte-Sorin, S., Hugonnet, J. E., Delfosse, V., Mainardi, J. L., Gutmann, L., Arthur, M., et al. (2006). Crystal structure of a novel  $\beta$ -lactam-insensitive peptidoglycan transpeptidase. *J. Mol. Biol.* 359, 533–538. doi: 10.1016/j.jmb.2006.03.014
- Bogden, C. E., Fass, D., Bergman, N., Nichols, M. D., and Berger, J. M. (1999). The structural basis for terminator recognition by the rho transcription termination factor. *Mol. Cell* 3, 487–493. doi: 10.1016/S1097-2765(00)80476-1
- Bork, P., Sander, C., and Valencia, A. (1992). An ATPase domain common to prokaryotic cell cycle proteins, sugar kinases, actin, and hsp70 heat shock proteins. *Proc. Natl. Acad. Sci. U.S.A.* 89, 7290–7294. doi: 10.1073/pnas.89.16.7290
- Boudvillain, M., Figueroa-Bossi, N., and Bossi, L. (2013). Terminator still moving forward: expanding roles for Rho factor. *Curr. Opin. Microbiol.* 16, 118–124. doi: 10.1016/j.mib.2012.12.003

- Cai, X., and Chen, Z. J. (2014). Prion-like polymerization as a signaling mechanism. *Trends Immunol.* 35, 622–630. doi: 10.1016/j.it.2014.10.003
- Cardinale, C. J., Washburn, R. S., Tadiogola, V. R., Brown, L. M., Gottesman, M. E., and Nudler, E. (2008). Termination factor rho and its cofactors NusA and NusG silence foreign DNA in *E. coli*. *Science* 320, 935–938. doi: 10.1126/science.1152763
- Caughey, B. (2001). Prion protein interconversions. *Philos. Trans. R. Soc. B Biol. Sci.* 356, 197–202. doi: 10.1098/rstb.2000.0765
- Chiti, F., and Dobson, C. M. (2006). Protein misfolding, functional amyloid, and human disease. *Annu. Rev. Biochem.* 75, 333–366. doi: 10.1146/annurev.biochem.75.101304.123901
- Conchillo-Solé, O., de Groot, N. S., Avilés, F. X., Vendrell, J., Daura, X., and Ventura, S. (2007). AGGRESCAN: a server for the prediction and evaluation of “hot spots” of aggregation in polypeptides. *BMC Bioinformatics* 8:65. doi: 10.1186/1471-2105-8-65
- de Groot, N. S., Parella, T., Aviles, F. X., Vendrell, J., and Ventura, S. (2007). Ile-phe dipeptide self-assembly: clues to amyloid formation. *Biophys. J.* 92, 1732–1741. doi: 10.1529/biophysj.106.096677
- Desantis, M. E., Leung, E. H., Sweeny, E. A., Jackrel, M. E., Cushman-Nick, M., Neuhaus-Follini, A., et al. (2012). Operational plasticity enables Hsp104 to disaggregate diverse amyloid and nonamyloid clients. *Cell* 151, 778–793. doi: 10.1016/j.cell.2012.09.038
- Diaz-Avalos, R., Long, C., Fontano, E., Balbirnie, M., Grothe, R., Eisenberg, D., et al. (2003). Cross-beta order and diversity in nanocrystals of an amyloid-forming peptide. *J. Mol. Biol.* 330, 1165–1175. doi: 10.1016/S0022-2836(03)00659-4
- Divakaruni, A. V., Loo, R. R., Xie, Y., Loo, J. A., and Gober, J. W. (2005). The cell-shape protein MreC interacts with extracytoplasmic proteins including cell wall assembly complexes in *Caulobacter crescentus*. *Proc. Natl. Acad. Sci. U.S.A.* 102, 18602–18607. doi: 10.1073/pnas.0507937102
- Dorsman, J. C., Pepers, B., Langenberg, D., Kerkdijk, H., Ijszenga, M., den Dunnen, J. T., et al. (2002). Strong aggregation and increased toxicity of polyisoleucine over polyglutamine stretches in mammalian cells. *Hum. Mol. Genet.* 11, 1487–1496. doi: 10.1093/hmg/11.13.1487
- Du, Z., Park, K. K.-W., Yu, H., Fan, Q., and Li, L. (2008). Newly identified prion linked to the chromatin-remodeling factor Swi1 in *Saccharomyces cerevisiae*. *Nat. Genet.* 40, 460–465. doi: 10.1038/ng.112
- Eichner, T., and Radford, S. E. (2011). A diversity of assembly mechanisms of a generic amyloid fold. *Mol. Cell.* 43, 8–18. doi: 10.1016/j.molcel.2011.05.012
- Espargaró, A., Villar-Piqué, A., Sabaté, R., and Ventura, S. (2012). Yeast prions form infectious amyloid inclusion bodies in bacteria. *Microb. Cell Fact.* 11, 89–89. doi: 10.1186/1475-2859-11-89
- Espelund, M., and Klaveness, D. (2014). Botulism outbreaks in natural environments - an update. *Front. Microbiol.* 5:287. doi: 10.3389/fmicb.2014.00287
- Espinosa Angarica, V., Angulo, A., Giner, A., Losilla, G., Ventura, S., and Sancho, J. (2014). PrionScan: an online database of predicted prion domains in complete proteomes. *BMC Genomics* 15:102. doi: 10.1186/1471-2164-15-102
- Espinosa Angarica, V., Ventura, S., and Sancho, J. (2013). Discovering putative prion sequences in complete proteomes using probabilistic representations of Q/N-rich domains. *BMC Genomics* 14:316. doi: 10.1186/1471-2164-14-316
- Esteras-Chopo, A., Serrano, L., and López de la Paz, M. (2005). The amyloid stretch hypothesis: recruiting proteins toward the dark side. *Proc. Natl. Acad. Sci. U.S.A.* 102, 16672–16677. doi: 10.1073/pnas.0505905102
- Fandrich, M., and Dobson, C. M. (2002). The behaviour of polyamino acids reveals an inverse side chain effect in amyloid structure formation. *EMBO J.* 21, 5682–5690. doi: 10.1093/emboj/cdf573
- Fernandez-Escamilla, A.-M., Rousseau, F., Schymkowitz, J., and Serrano, L. (2004). Prediction of sequence-dependent and mutational effects on the aggregation of peptides and proteins. *Nat. Biotechnol.* 22, 1302–1306. doi: 10.1038/nbt1012
- Fowler, D. M., Koulov, A. V., Balch, W. E., and Kelly, J. W. (2007). Functional amyloid – from bacteria to humans. *Trends Biochem. Sci.* 32, 217–224. doi: 10.1016/j.tibs.2007.03.003
- Furukawa, Y., and Nukina, N. (2013). Functional diversity of protein fibrillar aggregates from physiology to RNA granules to neurodegenerative diseases. *Biochim. Biophys. Acta* 1832, 1271–1278. doi: 10.1016/j.bbdis.2013.04.011
- Fuxreiter, M. (2012). Fuzziness: linking regulation to protein dynamics. *Mol. BioSyst.* 8, 168–177. doi: 10.1039/c1mb05234a
- Garbuzynskiy, S. O., Lobanov, M. Y., and Galzitskaya, O. V. (2010). FoldAmyloid: a method of prediction of amyloidogenic regions from protein sequence. *Bioinformatics* 26, 326–332. doi: 10.1093/bioinformatics/btp691
- Garrity, S. J., Sivanathan, V., Dong, J., Lindquist, S., and Hochschild, A. (2010). Conversion of a yeast prion protein to an infectious form in bacteria. *Proc. Natl. Acad. Sci. U.S.A.* 107, 10596–10601. doi: 10.1073/pnas.0913280107
- Geiselmann, J., Wang, Y., Seifried, S. E., and von Hippel, P. H. (1993). A physical model for the translocation and helicase activities of *Escherichia coli* transcription termination protein Rho. *Proc. Natl. Acad. Sci. U.S.A.* 90, 7754–7758. doi: 10.1073/pnas.90.16.7754
- Gohlke, J. R. (1972). *The Most Popular Fluorescence Probes in use at This Time*. Palatine, IL: Weber.
- Graña-Montes, R., Marinelli, P., Reverter, D., and Ventura, S. (2014). N-terminal protein tails act as aggregation protective entropic bristles: the SUMO case. *Biomacromolecules* 15, 1194–1203. doi: 10.1021/bm401776z
- Guynet, C., Cuevas, A., Moncalián, G., and de la Cruz, F. (2011). The stb operon balances the requirements for vegetative stability and conjugative transfer of plasmid R388. *PLoS Genet.* 7:e1002073. doi: 10.1371/journal.pgen.102073
- Halfmann, R., Alberti, S., Krishnan, R., Lyle, N., O'Donnell, C. W., King, O. D., et al. (2011). Opposing effects of glutamine and asparagine govern prion formation by intrinsically disordered proteins. *Mol. Cell* 43, 72–84. doi: 10.1016/j.molcel.2011.05.013
- Harrison, P. M., and Gerstein, M. (2003). A method to assess compositional bias in biological sequences and its application to prion-like glutamine/asparagine-rich domains in eukaryotic proteomes. *Genome Biol.* 4, R40–R40. doi: 10.1186/gb-2003-4-6-r40
- Heinrich, S. U., and Lindquist, S. (2011). Protein-only mechanism induces self-perpetuating changes in the activity of neuronal Aplysia cytoplasmic polyadenylation element binding protein (CPEB). *Proc. Natl. Acad. Sci. U.S.A.* 108, 2999–3004. doi: 10.1073/pnas.1019368108
- Hills, R. D., and Brooks, C. L. (2007). Hydrophobic cooperativity as a mechanism for amyloid nucleation. *J. Mol. Biol.* 368, 894–901. doi: 10.1016/j.jmb.2007.02.043
- Hug, S. M., and Gaut, B. S. (2015). The phenotypic signature of adaptation to thermal stress in *Escherichia coli*. *BMC Evol. Biol.* 15:177. doi: 10.1186/s12862-015-0457-3
- Iglesias, V., de Groot, N. S., and Ventura, S. (2015). Computational analysis of candidate prion-like proteins in bacteria and their role. *Front. Microbiol.* 6:1123. doi: 10.3389/fmicb.2015.01123
- Jarrett, J. T., and Lansbury, P. T. (1992). Amyloid fibril formation requires a chemically discriminating nucleation event: studies of an amyloidogenic sequence from the bacterial protein OsmB. *Biochemistry* 31, 12345–12352. doi: 10.1021/bi00164a008
- Kim, H. J., Kim, N. C., Wang, Y. D., Scarborough, E. A., Moore, J., Diaz, Z., et al. (2013). Mutations in prion-like domains in hnRNPA2B1 and hnRNPA1 cause multisystem proteinopathy and ALS. *Nature* 495, 467–473. doi: 10.1038/nature11922
- King, O. D., Gitler, A. D., and Shorter, J. (2012). The tip of the iceberg: RNA-binding proteins with prion-like domains in neurodegenerative disease. *Brain Res.* 1462, 61–80. doi: 10.1016/j.brainres.2012.01.016
- Klunk, W. E., Pettegrew, J. W., and Abraham, D. J. (1989). Quantitative evaluation of congo red binding to amyloid-like proteins with a beta-pleated sheet conformation. *J. Histochem. Cytochem.* 37, 1273–1281. doi: 10.1177/37.8.2666510
- Krishna Leela, J., Syeda, A. H., Anupama, K., and Gowrishankar, J. (2013). Rho-dependent transcription termination is essential to prevent excessive genome-wide R-loops in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* 110, 258–263. doi: 10.1073/pnas.1213123110
- Kumaran, D., Eswaramoorthy, S., Furey, W., Navaza, J., Sax, M., and Swaminathan, S. (2009). Domain organization in *Clostridium botulinum* neurotoxin type E is unique: its implication in faster translocation. *J. Mol. Biol.* 386, 233–245. doi: 10.1016/j.jmb.2008.12.027
- Lancaster, A. K., Nutter-Upham, A., Lindquist, S., and King, O. D. (2014). PLAAC: a web and command-line application to identify proteins with Prion-Like Amino Acid Composition. *Bioinformatics* 30, 2–3. doi: 10.1093/bioinformatics/btu310



- LeVine, H. (1993). Thioflavine T interaction with synthetic Alzheimer's disease beta-amyloid peptides: detection of amyloid aggregation in solution. *Protein Sci.* 2, 404–410. doi: 10.1002/pro.5560020312
- MacLea, K. S., Paul, K. R., Ben-Musa, Z., Waechter, A., Shattuck, J. E., Gruca, M., et al. (2015). Distinct amino acid compositional requirements for formation and maintenance of the  $[\Psi\text{PSI}^{\text{N}} + \Psi\text{N}]$  Prion in Yeast. *Mol. Cell. Biol.* 35, 899–911. doi: 10.1128/MCB.01020-14
- Magnet, S., Arbeloa, A., Mainardi, J. L., Hugonnet, J. E., Fourgeaud, M., Dubost, L., et al. (2007). Specificity of L,D-transpeptidases from gram-positive bacteria producing different peptidoglycan chemotypes. *J. Biol. Chem.* 282, 13151–13159. doi: 10.1074/jbc.M610911200
- Maji, S. K., Perrin, M. H., Sawaya, M. R., Jessberger, S., Vadodaria, K., Rissman, R. A., et al. (2009). Functional amyloids as natural storage of peptide hormones in pituitary secretory granules. *Science* 325, 328–332. doi: 10.1126/science.1173155
- Majumdar, A., Cesario, W. C., White-Grindley, E., Jiang, H., Ren, F., Khan, M. R., et al. (2012). Critical role of amyloid-like oligomers of *Drosophila* Orb2 in the persistence of memory. *Cell* 148, 515–529. doi: 10.1016/j.cell.2012.01.004
- Malinowska, L., Kroschwald, S., and Alberti, S. (2013). Protein disorder, prion propensities, and self-organizing macromolecular collectives. *Biochim. Biophys. Acta* 1834, 918–931. doi: 10.1016/j.bbapap.2013.01.003
- Maries, E., Dass, B., Collier, T. J., Kordower, J. H., and Steece-Collier, K. (2003). The role of alpha-synuclein in Parkinson's disease: insights from animal models. *Nat. Rev. Neurosci.* 4, 727–738. doi: 10.1038/nrn1199
- Menouni, R., Champ, S., Espinosa, L., Boudvillain, M., and Ansaldi, M. (2013). Transcription termination controls prophage maintenance in *Escherichia coli* genomes. *Proc. Natl. Acad. Sci. U.S.A.* 110, 14414–14419. doi: 10.1073/pnas.1303400110
- Michelitsch, M. D., and Weissman, J. S. (2000). A census of glutamine/asparagine-rich regions: implications for their conserved function and the prediction of novel prions. *Proc. Natl. Acad. Sci. U.S.A.* 97, 11910–11915. doi: 10.1073/pnas.97.22.11910
- Navarro, S., Marinelli, P., Diaz-Caballero, M., and Ventura, S. (2015). The prion-like RNA-processing protein HNRPDL forms inherently toxic amyloid-like inclusion bodies in bacteria. *Microb. Cell Fact.* 14, 102–102. doi: 10.1186/s12934-015-0284-7
- Nelson, R., Sawaya, M. R., Balbirnie, M., Madsen, A. Ø., Riek, C., Grothe, R., et al. (2005). Structure of the cross-beta spine of amyloid-like fibrils. *Nature* 435, 773–778. doi: 10.1038/nature03680
- Newby, G. A., and Lindquist, S. (2013). Blessings in disguise: biological benefits of prion-like mechanisms. *Trends Cell Biol.* 23, 251–259. doi: 10.1016/j.tcb.2013.01.007
- Pastor, M. T., Esteras-Chopo, A., and Serrano, L. (2007). Hacking the code of amyloid formation: the amyloid stretch hypothesis. *Prion* 1, 9–14. doi: 10.4161/pri.1.1.14100
- Patel, B. K., Gavin-Smyth, J., and Lieberman, S. W. (2009). The yeast global transcriptional co-repressor protein Cyc8 can propagate as a prion. *Nat. Cell Biol.* 11, 344–349. doi: 10.1038/ncb1843
- Prilusky, J., Felder, C. E., Zeev-Ben-Mordehai, T., Rydberg, E. H., Man, O., and Beckmann, J. S. (2005). FoldIndex: a simple tool to predict whether a given protein sequence is intrinsically unfolded. *Bioinformatics* 21, 3435–3438. doi: 10.1093/bioinformatics/bti537
- Richardson, J. P. (1990). Rho-dependent transcription termination. *Biochim. Biophys. Acta* 1048, 127–138. doi: 10.1016/0167-4781(90)90048-7
- Richardson, J. P. (1996). Structural organization of transcription termination factor Rho. *J. Biol. Chem.* 271, 1251–1254. doi: 10.1074/jbc.271.3.1251
- Rodríguez-Verdugo, A., Carrillo-Cisneros, D., Gonzalez-Gonzalez, A., Gaut, B. S., and Bennett, A. F. (2014). Different tradeoffs result from alternate genetic adaptations to a common environment. *Proc. Natl. Acad. Sci. U.S.A.* 111, 12121–12126. doi: 10.1073/pnas.1406886111
- Rogozha, T., Goginashvili, A., Rodionova, S., Ivanov, M., Viktorovskaya, O., Rubel, A., et al. (2010). Non-Mendelian determinant [ISP+] in yeast is a nuclear-residing prion form of the global transcriptional regulator Sfp1. *Proc. Natl. Acad. Sci. U.S.A.* 107, 10573–10577. doi: 10.1073/pnas.1005949107
- Romero, D., and Kolter, R. (2014). Functional amyloids in bacteria. *Int. Microbiol.* 17, 65–73. doi: 10.2436/20.1501.01.208
- Rossetto, O., Pirazzini, M., and Montecucco, C. (2014). Botulinum neurotoxins: genetic, structural and mechanistic insights. *Nat. Rev. Microbiol.* 12, 535–549. doi: 10.1038/nrmicro3295
- Sabate, R., Espargaro, A., Graña-Montes, R., Reverter, D., and Ventura, S. (2012). Native structure protects SUMO proteins from aggregation into amyloid fibrils. *Biomacromolecules* 13, 1916–1926. doi: 10.1021/bm3004385
- Sabaté, R., Espargaró, A., Saupe, S. J., and Ventura, S. (2009). Characterization of the amyloid bacterial inclusion bodies of the HET-s fungal prion. *Microb. Cell Fact.* 8, 56–56. doi: 10.1186/1475-2859-8-56
- Sabate, R., Rodríguez-Santiago, L., Sodupe, M., Saupe, S. J., and Ventura, S. (2013). Thioflavin-T excimer formation upon interaction with amyloid fibers. *Chem. Commun.* 49, 5745–5745. doi: 10.1039/c3cc42040j
- Sabate, R., Rousseau, F., Schymkowitz, J., Batlle, C., and Ventura, S. (2015a). Amyloids or prions? That is the question. *Prion* 9, 200–206. doi: 10.1080/19336896.2015.1053685
- Sabate, R., Rousseau, F., Schymkowitz, J., and Ventura, S. (2015b). What makes a protein sequence a prion? *PLoS Comput. Biol.* 11:e1004013. doi: 10.1371/journal.pcbi.1004013
- Sanchez de Groot, N., Torrent, M., Villar-Piqué, A., Lang, B., Ventura, S., Gsponer, J., et al. (2012). Evolutionary selection for protein aggregation. *Biochem. Soc. Trans.* 40, 1032–1037. doi: 10.1042/BST20120160
- Santner, A. A., Croy, C. H., Vasanwala, F. H., Uversky, V. N., Van, Y. Y., and Dunker, A. K. (2012). Sweeping away protein aggregation with entropic bristles: intrinsically disordered protein fusions enhance soluble expression. *Biochemistry* 51, 7250–7262. doi: 10.1021/bi300653m
- Si, K. (2015). Prions: what are they good for? *Annu. Rev. Cell Dev. Biol.* 31, 149–169. doi: 10.1146/annurev-cellbio-100913-013409
- Skarin, H., and Segerman, B. (2014). Plasmidome interchange between *Clostridium botulinum*, *Clostridium novyi* and *Clostridium haemolyticum* converts strains of independent lineages into distinctly different pathogens. *PLoS ONE* 9:e107777. doi: 10.1371/journal.pone.0107777
- Stohr, J., Watts, J. C., Mensinger, Z. L., Oehler, A., Grillo, S. K., DeArmond, S. J., et al. (2012). Purified and synthetic Alzheimer's amyloid beta (A $\beta$ ) prions. *Proc. Natl. Acad. Sci. U.S.A.* 109, 11025–11030. doi: 10.1073/pnas.1206555109
- Sugawara, Y., Yutani, M., Amatsu, S., Matsumura, T., and Fujinaga, Y. (2014). Functional dissection of the *Clostridium botulinum* type B hemagglutinin complex: identification of the carbohydrate and E-cadherin binding sites. *PLoS ONE* 9:e111170. doi: 10.1371/journal.pone.0111170
- Swaminathan, S., and Eswaramoorthy, S. (2000). Structural analysis of the catalytic and binding sites of *Clostridium botulinum* neurotoxin B. *Nat. Struct. Biol.* 7, 693–699. doi: 10.1186/1741-7015-7-14
- Tariq, M., Wegrzyn, R., Anwar, S., Bukau, B., and Paro, R. (2013). *Drosophila* GAGA factor polyglutamine domains exhibit prion-like behavior. *BMC Genomics* 14:374. doi: 10.1186/1471-2164-14-374
- Tavares, A. C., Fernandes, P. B., Carballido-Lopez, R., and Pinho, M. G. (2015). MreC and MreD proteins are not required for growth of *Staphylococcus aureus*. *PLoS ONE* 10:e0140523. doi: 10.1371/journal.pone.0140523
- Tenaillon, O., Rodríguez-Verdugo, A., Gaut, R. L., McDonald, P., Bennett, A. F., Long, A. D., et al. (2012). The molecular diversity of adaptive convergence. *Science* 335, 457–461. doi: 10.1126/science.1212986
- Toombs, J. A., McCarty, B. R., and Ross, E. D. (2010). Compositional determinants of prion formation in yeast. *Mol. Cell. Biol.* 30, 319–332. doi: 10.1128/MCB.01140-09
- Toombs, J. A., Petri, M., Paul, K. R., Kan, G. Y., Ben-Hur, A., and Aand Ross, E. D. (2012). De novo design of synthetic prion domains. *Proc. Natl. Acad. Sci. U.S.A.* 109, 6519–6524. doi: 10.1073/pnas.1119366109
- Uversky, V. N. (2013). The most important thing is the tail: multitudinous functionalities of intrinsically disordered protein termini. *FEBS Lett.* 587, 1891–1901. doi: 10.1016/j.febslet.2013.04.042
- Uversky, V. N. (2015). Intrinsically disordered proteins and their (disordered) proteomes in neurodegenerative disorders. *Front. Aging Neurosci.* 7:18. doi: 10.3389/fnagi.2015.00018
- van Der Wel, P. C., Hu, K. N., Lewandowski, J., and Griffin, R. G. (2006). Dynamic nuclear polarization of amyloidogenic peptide nanocrystals. *J. Am. Chem. Soc.* 128, 10840–10846. doi: 10.1021/ja0626685
- Ventura, S., Zurdo, J., Narayanan, S., Parreño, M., Mangues, R., Reif, B., et al. (2004). Short amino acid stretches can mediate amyloid formation in globular

- proteins: the Src homology 3 (SH3) case. *Proc. Natl. Acad. Sci. U.S.A.* 101, 7258–7263. doi: 10.1073/pnas.0308249101
- Washburn, R. S., and Gottesman, M. E. (2011). Transcription termination maintains chromosome integrity. *Proc. Natl. Acad. Sci. U.S.A.* 108, 792–797. doi: 10.1073/pnas.1009564108
- Wickner, R. B. (1994). [URE3] as an altered URE2 protein: evidence for a prion analog in *Saccharomyces cerevisiae*. *Science* 264, 566–569. doi: 10.1126/science.7909170
- Wickner, R. B., Taylor, K. L., Edskes, H. K., Maddelein, M. L., Moriyama, H., and Roberts, B. T. (2001). Yeast prions act as genes composed of self-propagating protein amyloids. *Adv. Protein Chem.* 57, 313–334. doi: 10.1016/S0065-3233(01)57026-6
- Wren, B. W. (1991). A family of clostridial and streptococcal ligand-binding proteins with conserved C-terminal repeat sequences. *Mol. Microbiol.* 5, 797–803. doi: 10.1111/j.1365-2958.1991.tb00752.x
- Yuan, A. H., Garrity, S. J., Nako, E., and Hochschild, A. (2014). Prion propagation can occur in a prokaryote and requires the ClpB chaperone. *eLife* 3, 1–19. doi: 10.7554/eLife.02949
- Zambrano, R., O Conchillo-Sole, Iglesias, V., Illa, R., Rousseau, F., and Schymkowitz, J. (2015). PrionW: a server to identify proteins containing glutamine/asparagine rich prion-like domains and their amyloid cores. *Nucleic Acids Res.* 43, W331–W337. doi: 10.1093/nar/gkv490
- Zheng, J., Ma, B., and Nussinov, R. (2006). Consensus features in amyloid fibrils: sheet–sheet recognition via a (polar or nonpolar) zipper structure. *Phys. Biol.* 3, 1–1. doi: 10.1088/1478-3975/3/3/P01
- Zhou, Y., Smith, D., Leong, B. J., Brännström, K., Almqvist, F., and Chapman, M. R. (2012). Promiscuous cross-seeding between bacterial amyloids promotes interspecies biofilms. *J. Biol. Chem.* 287, 35092–35103. doi: 10.1074/jbc.M112.383737
- Zilhao, R., Naclerio, G., Henriques, A. O., Baccigalupi, L., Moran, C. P., and Ricca, E. (1999). Assembly requirements and role of CotH during spore coat formation in *Bacillus subtilis*. *J. Bacteriol.* 181, 2631–2633.

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Pallarès, Iglesias and Ventura. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

# Engineered bacterial hydrophobic oligopeptide repeats in a synthetic yeast prion, [REP-PSI<sup>+</sup>]

Fátima Gasset-Rosa<sup>†</sup> and Rafael Giraldo<sup>\*</sup>

Department of Cellular and Molecular Biology, Centro de Investigaciones Biológicas – Consejo Superior de Investigaciones Científicas, Madrid, Spain

## OPEN ACCESS

### Edited by:

Salvador Ventura,  
Universitat Autònoma de Barcelona,  
Spain

### Reviewed by:

Jesus R. Requena,  
University of Santiago de  
Compostela, Spain  
Galina Zhouravleva,  
Saint Petersburg State University,  
Russia

### \*Correspondence:

Rafael Giraldo,  
Department of Cellular and Molecular  
Biology, Centro de Investigaciones  
Biológicas – Consejo Superior de  
Investigaciones Científicas, C/ Ramiro  
de Maeztu 9, E-28040 Madrid, Spain  
rgiraldo@cib.csic.es

### <sup>†</sup>Present address:

Fátima Gasset-Rosa,  
Department of Neurosciences,  
Ludwig Institute for Cancer Research,  
University of California at San Diego,  
La Jolla, CA, USA

### Specialty section:

This article was submitted to Microbial  
Physiology and Metabolism, a section  
of the journal Frontiers in Microbiology

**Received:** 23 February 2015

**Accepted:** 29 March 2015

**Published:** 21 April 2015

### Citation:

Gasset-Rosa F and Giraldo R (2015)  
Engineered bacterial hydrophobic  
oligopeptide repeats in a synthetic  
yeast prion, [REP-PSI<sup>+</sup>].  
Front. Microbiol. 6:311.  
doi: 10.3389/fmicb.2015.00311

The yeast translation termination factor Sup35p, by aggregating as the [PSI<sup>+</sup>] prion, enables ribosomes to read-through stop codons, thus expanding the diversity of the *Saccharomyces cerevisiae* proteome. Yeast prions are functional amyloids that replicate by templating their conformation on native protein molecules, then assembling as large aggregates and fibers. Prions propagate epigenetically from mother to daughter cells by fragmentation of such assemblies. In the N-terminal prion-forming domain, Sup35p has glutamine/asparagine-rich oligopeptide repeats (OPRs), which enable propagation through chaperone-elicited shearing. We have engineered chimeras by replacing the polar OPRs in Sup35p by up to five repeats of a hydrophobic amyloidogenic sequence from the synthetic bacterial prionoid RepA-WH1. The resulting hybrid, [REP-PSI<sup>+</sup>], (i) was functional in a stop codon read-through assay in *S. cerevisiae*; (ii) generates weak phenotypic variants upon both its expression or transformation into [psi<sup>-</sup>] cells; (iii) these variants correlated with high molecular weight aggregates resistant to SDS during electrophoresis; and (iv) according to fluorescence microscopy, the fusion of the prion domains from the engineered chimeras to the reporter protein mCherry generated perivacuolar aggregate foci in yeast cells. All these are signatures of *bona fide* yeast prions. As assessed through biophysical approaches, the chimeras assembled as oligomers rather than as the fibers characteristic of [PSI<sup>+</sup>]. These results suggest that it is the balance between polar and hydrophobic residues in OPRs what determines prion conformational dynamics. In addition, our findings illustrate the feasibility of enabling new propagation traits in yeast prions by engineering OPRs with heterologous amyloidogenic sequence repeats.

**Keywords:** amyloid cross-seeding, prion variants/strains, RepA-WH1 prionoid, [REP-PSI<sup>+</sup>] prion, *Saccharomyces cerevisiae*, synthetic biology

## Introduction

Modularity is a basic principle of organization in proteins and their assemblies. In the case of the aggregation-prone amyloidogenic proteins, modularity comes from the existence of sequence stretches that become the  $\beta$ -strand building blocks in cross- $\beta$  sheets (reviewed in Eisenberg and Jucker, 2012). Such stretches can either have hydrophobic or polar average residue compositions, but usually not both at the same time, thus posing constraints to homotypic interactions, which require chemically compatible side chains to cross-aggregate. Polar glutamine/asparagine

(Q/N)-rich amyloidogenic sequences are a hallmark of prion domains in yeast (Alberti et al., 2009; reviewed in Liebman and Chernoff, 2012), in which cross-seeding has been well characterized, e.g., for the Rnq1p/[PIN<sup>+</sup>] prion nucleating the aggregation of Sup35p/[PSI<sup>+</sup>] (Vitrenko et al., 2007; Sharma and Liebman, 2013). In the translation releasing factor Sup35p, its prion-forming domain (N) includes up to five and a half oligopeptide repeats (OPRs; **Figure 1A**; reviewed in Tessier and Lindquist, 2009). OPRs, while still assembled in the amyloid fibers, have been proposed to be the targets for the disaggregase chaperone Hsp104p (Chernoff et al., 1995) that, with the aid of Hsp70 chaperones Ssa1-4p (Shorter and Lindquist, 2008; Winkler et al., 2012), generate [PSI<sup>+</sup>] propagons, i.e., oligomeric primordia that are readily diffusible to the progeny (Cox et al., 2003; Derdowski et al., 2010). The modularity of Sup35p OPRs has allowed their partial or total replacement by heterologous sequences, such as the octapeptide repeats in the mammalian prion protein PrP (Parham et al., 2001; Dong et al., 2007; Tank et al., 2007), while keeping the resulting chimeras their original function as epigenetic determinants of reading-through stop codons in yeast. With the accumulated knowledge on Sup35p/[PSI<sup>+</sup>], this is probably the most suitable model system to address the molecular determinants of the prion condition in proteins (Sabate et al., 2015).

RepA, the DNA replication protein of the *Pseudomonas* plasmid pPS10, is one of the members best characterized of a protein family spread across a large group of plasmids from Gram-negative bacteria (reviewed in Giraldo and Fernández-Tresguerres, 2004). Soluble RepA dimers, which function as transcriptional self-repressors of the *repA* gene, undergo a large conformational change upon binding to specific DNA sequences from the plasmid replication origin, resulting in their dissociation into replication-competent RepA monomers (Díaz-López et al., 2003, 2006). Such structural change consists in an increase in  $\beta$ -sheet at the expense of the  $\alpha$ -helical secondary structure component and affects WH1, the most N-terminal of the two 'winged-helix' domains in RepA, which is transformed from a dimerization domain into a secondary DNA binding module, ancillary to the main DNA recognition determinant, C-terminal WH2 (Giraldo et al., 1998, 2003). RepA monomers are aggregation-prone and, once a replication round is completed, the resulting plasmid copies are held together through interactions between RepA monomers bound to the replication origins of the two plasmid molecules, thus inhibiting further replication rounds (Gasset-Rosa et al., 2008a). Such complex interplay between transcriptional repression and DNA replication initiation/inhibition, focused further research on the molecular basis for the ligand (DNA)-modulated balance between RepA solubility and functional aggregation. In RepA-WH1, a single hydrophobic sequence stretch partially folded as a  $\alpha$ -helix changes its conformation, upon transient binding to specific DNA sequences, to assemble as a  $\beta$ -strand in amyloid fibers (Giraldo, 2007; Gasset-Rosa et al., 2008b). The RepA-WH1 domain, when expressed in *Escherichia coli* as a metastable fusion to the fluorescent protein mCherry, behaves as a synthetic prionoid, i.e., an amyloidogenic protein lacking infectivity (Aguzzi, 2009).

Unlike prions in yeast, RepA-WH1 triggers a sort of intracellular amyloid 'proteinopathy' in its host (Fernández-Tresguerres et al., 2010), where the 'vertical' propagation (i.e., from mother cell to daughter cells) and the conformational selection of 'strains' of the prionoid are dependent on the Hsp70 chaperone DnaK, rather than on the Hsp104 ortholog ClpB (Gasset-Rosa et al., 2014). A recent electron microscopy reconstruction of the RepA-WH1 fibers, as templated *in vitro* on soluble molecules of the protein by amyloid seeds preformed *in vivo*, showed that the fibers are composed of intertwined amyloid tubules built by distorted protein monomers (Torreira et al., 2015).

Since the amyloidogenic sequence stretch in RepA-WH1 has clearly distinct features to those naturally found in yeast prions, – i.e., a single hydrophobic amyloidogenic stretch in a properly folded domain, rather than multiple repeats with polar residue composition in the frame of an intrinsically disordered tail – the bacterial prionoid is a suitable source of radically heterologous sequences to explore their ability to supplant the OPRs in Sup35p/[PSI<sup>+</sup>]. We report here that chimeras between Sup35p and at least three repeats of the hydrophobic amyloidogenic stretch in RepA-WH1 constitute novel functional prions in yeast that we have named [REP-PSI<sup>+</sup>]. Furthermore, [REP-PSI<sup>+</sup>] can cross-seed the propagation of [PSI<sup>+</sup>], but generating [PSI<sup>+</sup>]<sup>WH1</sup>, a new strain with very low mitotic stability (i.e., epigenetically unstable), as expected for a weak prion with severe limitations on the generation of propagons. These findings validate the possibility of tailoring yeast epigenetics by engineering aggregation-prone hydrophobic amyloidogenic repeats within prion-forming domains.

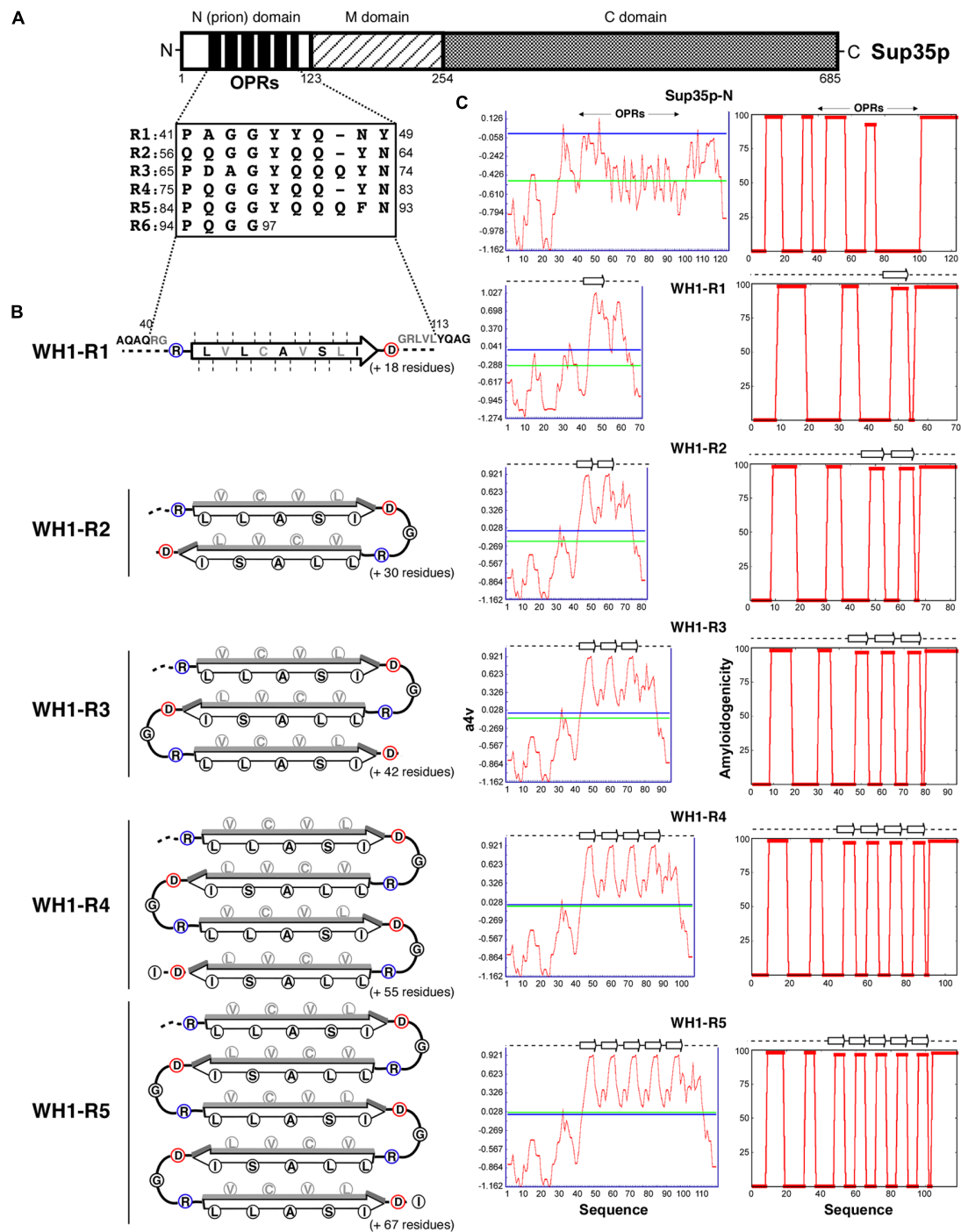
## Materials and Methods

### Plasmid Constructs

#### Substitution of Sup35p-OPRs by Tandem Repeats of the Amyloidogenic Stretch in RepA-WH1 (WH1-R1-5)

The reporter plasmid pUKC1620 (Parham et al., 2001; Von der Haar et al., 2007), harboring a WT copy of the *SUP35* gene with its natural promoter, was used as template for PCR-amplification (*Pfu* DNAPol) of *NΔR6-sup35* using primers with BamHI (3' to the promoter) and EcoRV (annealing after the 5th repeat) ends. After digestion with those enzymes, the product replaced the BamHI-EcoRV *N-SUP35* fragment in the original plasmid, leading to pUKC\_R5. On this vector, WH1-R1 was inserted through PCR, including the reverse primer (5'-CGAGgatacCGATCAAAGACACAGCGCATAGCACTAGATTGAATTGCTGCTGAT), after the EcoRV site (lowercase), the complement of the sequence coding for the nine core amino-acid residues (LVLCVAVSLI) in WH1(A31V) (bold; Giraldo, 2007). WH1-R2 was built in a similar way, but the reverse primer had that sequence repeated twice with a TCC/Gly spacer. For pUKC\_R0, a construction lacking the five natural OPRs in Sup35p, the PstI site previous to the first OPR in the parental plasmid was mutated to EcoRV (QuikChange kit, Stratagene), followed by digestion with this enzyme and in-frame religation of the vector. WH1-R1-3 were built





**FIGURE 1 | Chimeras between the yeast prion protein Sup35p and the amyloidogenic stretch in the bacterial prionoid RepA-WH1.**

**(A)** Schematic representation of *Saccharomyces cerevisiae* eRF3 (Sup35p), with its domain composition and the N-terminal 5½ OPRs highlighted.

**(B)** The R0 + WH1-R1-5 chimeras. Sup35p OPRs were totally replaced by 1–5 tandem repeats of the hydrophobic amyloidogenic stretch in the RepA-WH1 variant A31V (arrow; Giraldo, 2007). The heterologous sequences were engineered to fold as zigzag  $\beta$ -arcades (reviewed in Kajava et al., 2010), in which turns were made by inserting a Gly residue between the natural Asp (blue) C- and Arg (red) ‘gatekeeper’ N-ends from contiguous WH1 repeats.  $\beta$ -strand packing (faces of strands in black,

backs in gray) would be guided by interdigitation of chemically compatible side chains.  $\beta$ -arcades from distinct protein monomers would stack through main-chain hydrogen bonds into a parallel superpleated  $\beta$ -structure, as proposed for Sup35p/[PSI<sup>+</sup>] (reviewed in Kajava et al., 2010). In the resulting fiber, the engineered repeats should have to accommodate as a hydrophobic spine within an otherwise polar Q/N-rich axis. **(C)** The amyloidogenicity profiles (red plots) for the distinct chimeric N-domains, as estimated by the AGGRESCAN (left column; Conchillo-Solé et al., 2007) and WALTZ (right; Maurer-Stroh et al., 2010) algorithms, suggest that the synthetic peptide arrays assemble as amyloids. Horizontal lines: blue, aggregation hot spot threshold; green, sequence average.

by annealing the pre-phosphorylated (T4-PNK) oligos 5'-GGCCGCCTAGTGCTATGCGCTGTGTCTTTGATCGAT/5'-GCATAGCACTAGGCGGCCATCGATCAAAGACACAGC (complementary sequences in bold) followed by filling of the ends (Klenow DNAPol) and self-ligation, a step that generated an extra terminal half repeat. Each repeat in the concatemers includes the natural C-terminal Asp and N-terminal Arg found in the RepA-WH1 amyloidogenic stretch (Giraldo, 2007) plus a linker Gly. Upon ligation of the concatemers mix at the EcoRV site in pUKC\_R0 and transformation into *E. coli*, DNA sequencing identified clones with up to three RepA-WH1 repeats. Since WH1-R4-5 were recalcitrant to be obtained through such approach, a fragment including three RepA-WH1 repeats was made through chemical synthesis (ATG:biosynthetics, Germany) and inserted into pUKC\_R1-2, in which EcoRV had been regenerated by site directed mutagenesis, to give pUKC\_R4-5.

### Plasmids for Protein Expression in Yeast under the pGAL1 Promoter

The full-length WH1(Rn)-SUP35 chimeras were inserted into the pYef2 vector (Cullin and Minvielle-Sebastia, 1994) after amplification by PCR, using as templates the pUKC series (see above) and primers having BamHI and NotI ends. These chimeras carried a C-terminal hemagglutinin (HA)-tag from the vector. For fluorescence microscopy observation, fusions were made between the NM-domains from the chimeras and the monomeric red fluorescent protein mCherry at their C-termini. With this purpose, each NM-domain was amplified by PCR on the pUKC plasmids, with oligonucleotides carrying BamHI and BspEI ends, and then ligated with a BspEI (5')-NotI (3') mCherry fragment, amplified using as template pWH1(A31V)-mRFP (Fernández-Tresguerres et al., 2010). The BamHI-NotI fragments were then cloned into pYef2 as above.

### Plasmids for Protein Expression in *E. coli*

The fusions of NM-WT, NM-R0, NM-R0 + WH1-R2-4 to mCherry were cloned into the pRG vectors (*Ptac*, His<sub>10</sub> N-tag; Fernández-Tresguerres et al., 2010) by PCR amplification, performed with primers with SacII (5'-NM) and XbaI (3'-mCherry) ends. All constructs were verified by DNA sequencing.

### *Saccharomyces cerevisiae* Strains

For screening prion-dependent translation through stop codons, it was used a derivative of the strain 74D-694: MATa *ade1-14*<sup>UGA</sup> *trp1-289 his3-Δ200 ura3-52 leu2-3,112 sup35::loxP* [pYK810] [*PSI*<sup>+</sup>] [*PIN*<sup>+</sup>]. For fluorescence microscopy, the YJT28 strain (*ΔARS305::kanMX*, *ade2-1::ADE2*, W303-1a) was selected to suppress autogenous red fluorescence.

### Epigenetic Assay for Yeast Colony Color

The pUKCs derivatives (encoding the different chimeric alleles) were electroporated in a [*PSI*<sup>+</sup>] [*PIN*<sup>+</sup>] strain (see above) that initially carried pYK810, a plasmid bearing a copy of SUP35 (Von der Haar et al., 2007). For displacing the resident pYK810, thus assuring that the incoming pUKCs were the only source of Sup35p (or its chimeras with the WH1 repeats), colonies growing in SD-His were replicated on the same medium containing

0.1% 5-FOA, thus counter-selecting for cells that carried pYK810 (*URA3*). Colonies were then plated on 1/4 YPD and SD-adenine. For full development of color in colonies, agar plates were incubated for ≥72 h at 30°C and then transferred to 4°C for 24 h before photographic documentation.

### Protein Aggregate Transformation into [*psi*<sup>-</sup>] Cells

Overnight cultures inoculated from pUKCs-carrying red [*psi*<sup>-</sup>] colonies, that have spontaneously lost the [*PSI*<sup>+</sup>] phenotype, were diluted 1/8 into 60 ml of YPD and grown to OD<sub>600</sub> = 0.5. Cells were washed with water and 1 M sorbitol, resuspended in SCE buffer (1 M sorbitol, 100 mM Na-citrate pH 5.8, 10 mM EDTA, 10 mM DTT, 2 mg/ml lyticase), and then incubated at 30°C for 1 h. The resulting spheroplasts were centrifuged and resuspended in SCT buffer (1 M sorbitol, 10 mM Tris·HCl pH 7.5, 10 mM CaCl<sub>2</sub>) and 100 μl were co-transformed with 2 μM of pYef2 (Tanaka, 2010) as a marker (*URA3*), 10 μg of carrier ssDNA and 5 μl of whole cell extracts (i.e., the low-speed supernatants after cells lysis, see below) from pUKCs/[*REP-PSI*<sup>+</sup>] cells. Suspensions were incubated at room temperature (RT) with rotation mixing for 30 min. Then 44% PEG4000, 10 mM Tris·HCl pH 7.5, 10 mM CaCl<sub>2</sub> buffer was added and further incubated for 45 min. Spheroplasts were then sedimented, resuspended in SOS medium (1 M sorbitol, 25% YPD, 7 mM CaCl<sub>2</sub>), added to 10 ml of top agar (SD-URA, 2% dextrose, 0.8% agar, 1 M sorbitol, 2% YPD), and plated on SD-URA agar. Incubation proceeded for ≥72 h at 30°C. Large size colonies were selected and spotted on 1/4 YPD agar. To address the stability of [*PSI*<sup>+</sup>], white colonies obtained after transformation of Sup35p-WT [*psi*<sup>-</sup>] cells with R0 + WH1-R3-5 protein extracts were grown in YPD at 30°C overnight. They were then diluted to OD<sub>600</sub> = 0.001, grown for 24 h, and 30 μl plated on 1/4 YPD agar and incubated as above. Transformation and stability assays were performed independently twice.

### Aggregate Extraction and Sedimentation Assay

Two hundred ml cultures of yeast carrying the full-length protein chimeras cloned into pYef2s (see above) were grown overnight in selective medium (SD-Ura) with glucose. Then cultures were diluted to OD<sub>600</sub> = 0.07 in SD-Ura, but with 2% raffinose and 0.1% glucose, and grown to OD<sub>600</sub> = 0.2, when protein expression was induced by adding 2% galactose and further incubated until OD<sub>600</sub> = 2. Cells were then harvested and resuspended in 500 μl of 25 mM Tris·HCl pH 6.8, 250 mM NaCl, 5 mM EDTA, 10% glycerol (plus protease inhibitors; Roche). Lysis was then carried-out with glass beads (Lysing matrix C) in a MP FastPrep-24 homogenizer (five cycles, level 5, for 30 s at 4°C). Cell debris was removed by a low-speed sedimentation step (600 × g, 3 min). Two hundred microliter of the resulting whole cell extracts were ultracentrifuged at 50,000 rpm (100,000 × g), for 15 min at 4°C (Beckman Optima Max-XP, TLA100 rotor). Supernatants were collected and pellets were resuspended in 200 μl of the lysis solution. Proteins in equivalent volumes of supernatant and pellet fractions were analyzed by SDS-PAGE (10% polyacrylamide; 30 μg/lane) plus Western-blotting, using

an anti-HA antibody (Roche, 1:1,000) and chemiluminescence detection (ECL2; Pierce).

### Semi-Denaturing Detergent Agarose Gel Electrophoresis (SDD-AGE)

Total cell lysates (45  $\mu$ l, at 30 mg/ml) from yeast having the chimeras expressed from pYef2 (see above) were mixed with 15  $\mu$ l of loading buffer (TAE 2X, 20% glycerol, 8% sarkosyl, 0.5 g/l bromophenol blue, plus protease inhibitors). Samples were incubated at RT for 10 min, and electrophoresis performed in 1.5% agarose gels (TAE 1X, 0.1% SDS) at 100 V for 7.3 h, 10°C (Molina-García and Gasset-Rosa, 2014). Proteins were then transferred to a PVDF membrane in a Trans-Blot device (Bio-Rad) in TAE 1X, 0.1% SDS, at 16 V for 15 h, 10°C. Detection was performed with anti-HA (1:1,000).

### Visualization of Aggregates by Fluorescence Microscopy

#### Overexpression of the NM-mCherry Chimeras

pYef2s encoding the chimeras were transformed into the YJT28 strain and protein expression was carried out as described above. Culture aliquots were taken along 22 h for live cell observation.

#### Fluorescence Microscopy

It was performed with a Nikon Eclipse 90i microscope, equipped with CFI PLAN APO VC (NA 1.40) oil immersion objective and a Hamamatsu ORCA-R<sup>2</sup> CCD camera. A red filter with excitation 543/22 and emission 593/40 was used. Differential interference contrast (DIC) images were also captured.

### Purification of His<sub>6</sub>-NM-mCherry Chimeras

Protein expression and purification were performed as described for His<sub>6</sub>-RepA-WH1 (Giraldo, 2007), but extending the Ni<sup>2+</sup>-IMAC gradient to 0.5 M imidazole. Protein stocks (30  $\mu$ M) were kept at -70°C in 0.1 M Na<sub>2</sub>SO<sub>4</sub>, 20 mM Na<sub>2</sub>HPO<sub>4</sub> pH 6, 5 mM 2-mercaptoethanol, 10% glycerol.

### Amyloid Assembly of NM-mCherry Chimeras *In Vitro*

Protein chimeras (15  $\mu$ M) were assembled *in vitro* by still incubation at 5°C for a month, as described for RepA-WH1 (Giraldo, 2007), in 0.1 M Na<sub>2</sub>SO<sub>4</sub>, 60 mM Hepes pH 8, 8 mM MgSO<sub>4</sub>, 14% PEG4000, 6% MPD. Samples were examined in a JEOL JEM-1230 electron microscope.

### Circular Dichroism (CD) Spectroscopy

Spectra of the purified NM-mCherry chimeras (15  $\mu$ M) were acquired in 0.1 M Na<sub>2</sub>SO<sub>4</sub>, 15 mM Na<sub>2</sub>HPO<sub>4</sub> pH 6, at 5°C, as described (Giraldo, 2007).

### Analytical Ultracentrifugation

NM-mCherry chimeras were dialyzed in 0.1 M Na<sub>2</sub>SO<sub>4</sub>, 20 mM Na<sub>2</sub>HPO<sub>4</sub> pH 6, 5 mM 2-mercaptoethanol. Four hundred microliter of each sample were diluted to 0.8, 0.2, and 0.08 mg/ml and then centrifuged for 5 min at 13,000 rpm, 4°C. The clarified supernatants were studied by sedimentation velocity. Centrifugation was carried out in a Beckman-Coulter XLI

analytical ultracentrifuge, at 48,000 rpm and 20°C, measuring absorbance at 280 nm. Sedimentation coefficient distributions, *c*(*s*), were determined, with a confidence level of 0.68, using the SEDFIT 14.1 software (Schuck, 2000).

## Results

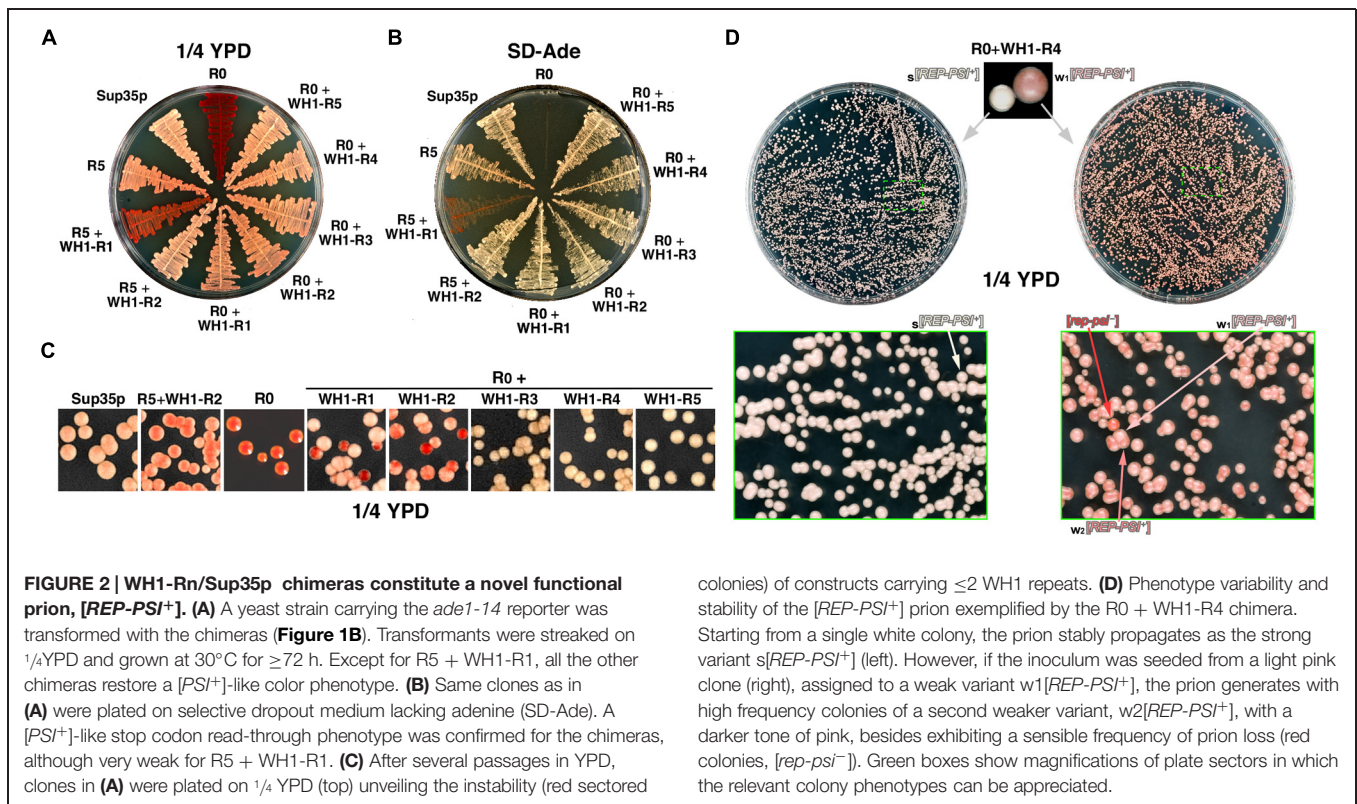
### Engineering Chimeras between the Amyloidogenic Stretch in RepA-WH1 and [PSI<sup>+</sup>]

Structural modeling of the amyloid fibers assembled by the N-domains of yeast prion proteins suggests that their basic building block might be a  $\beta$ -arch (reviewed in Kajava et al., 2010), in which adjacent Q/N-rich stretches would assemble as  $\beta$ -strands interdigitated through compatible side chains, while the intervening sequence would form a turn. If multiple stretches were present, as in Sup35p OPRs,  $\beta$ -arches would be further folded as  $\beta$ -arcades. The stacking of  $\beta$ -arcades from distinct protein monomers, stabilized through parallel main-chain hydrogen bonding, would result in a parallel superpleated  $\beta$ -structure (reviewed in Kajava et al., 2010). With a  $\beta$ -arcade based model in mind, we used a plasmid reporter system (Parham et al., 2001) in which all the OPRs in Sup35p were replaced by up to five tandem repeats of the hydrophobic amyloidogenic stretch found in the bacterial prionoid RepA-WH1 (Giraldo, 2007; WH1-R1-5; **Figure 1**).

### RepA-WH1 + Sup35p Chimeras are Functional in Yeast

The engineered chimeras were expressed in yeast from the *SUP35* promoter in a centromeric plasmid, and tested in an epigenetic red-white colony color assay (**Figure 2**): upon Sup35p aggregation as [PSI<sup>+</sup>], the reporter *ade1-14* allele, including a premature amber stop codon, is read-through by the ribosomes allowing for the synthesis of adenine, thus giving white color colonies on rich, unselective medium. Otherwise, soluble Sup35p efficiently terminates translation resulting in red [*psi*<sup>-</sup>] colonies by accumulation of the adenine precursor metabolite 5'-P-ribosyl-5-aminoimidazole (reviewed in Tessier and Lindquist, 2009). In the absence of endogenous Sup35p-WT, achieved in [PSI<sup>+</sup>] cells upon displacement of a resident *SUP35* plasmid by vectors encoding the chimeras, five of the native OPRs fused to two chimeric RepA-WH1 repeats (R5 + WH1-R2) were sufficient to yield a [PSI<sup>+</sup>]-like prion phenotype. However, with just one bacterial repeat (R5 + WH1-R1) the phenotype of this chimera was even weaker (more intense red color) than the R5 parental, suggesting that the insertion of a single hydrophobic stretch in the Q/N-repeats destabilized their assembly as amyloid. Interestingly, in the absence of any OPRs all the constructs including RepA-WH1 repeats (R0 + WH1-R1-5) gave a nearly WT phenotype, indicating that, in these chimeras, amyloids were successfully built with little interference between the RepA-WH1 and the Sup35p moieties (**Figure 2A**). In adenine-deficient media, all chimeras supported yeast growth, i.e., they read-through *ade1-14* thus restoring a functional pathway for adenine synthesis, albeit R5 + WH1-R1 led again to a poor phenotype (**Figure 2B**).





When, after several passages through liquid rich medium, yeast carrying the chimeras were plated on 1/4 YPD agar (Figure 2C) red sectored colonies appeared if the number of WH1 peptide repeats was ≤2, indicating that these were unstable chimeric prions. The chimeric prions including ≥3 WH1 repeats, which we termed [REP-PSI<sup>+</sup>], exhibited two clearly different phenotypes, namely white and pink colonies (Figure 2D, top). To explore the mitotic stability of both phenotypes, single colonies of each kind were subcultured and plated again on 1/4 YPD. White colonies showed a remarkable stability (frequency of conversion to red ≈10<sup>-6</sup>, matching that described for [PSI<sup>+</sup>]; as reviewed in Liebman and Chernoff, 2012; Figure 2D, left), whereas light pink colonies were very unstable, giving rise with high frequencies to dark pink (≈2 × 10<sup>-1</sup>) and red (≈10<sup>-2</sup>), or reverting to white (≈3 × 10<sup>-3</sup>), colonies (Figure 2D, right). Tentatively, the white phenotype was associated to a strong prion variant, s[REP-PSI<sup>+</sup>], the two pink tones to weak prion variants, w1 and w2[REP-PSI<sup>+</sup>] (light or dark pink, respectively), and the red colonies to cured (i.e., having lost the prion phenotype) [rep-psi<sup>-</sup>].

### [REP-PSI<sup>+</sup>] Prion is Infectious and Templates on Sup35p a New Weak, Unstable Variant, [PSI<sup>+</sup>]<sup>WH1</sup>

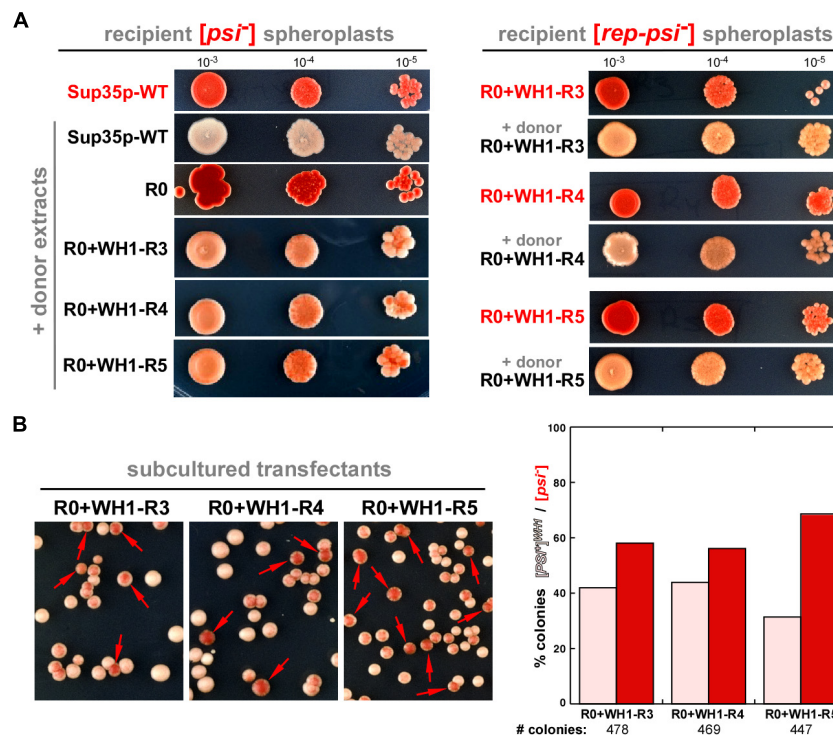
Yeast spheroplasts can be transformed with prion particles, either assembled *in vitro* or extracted from cultured cells (Tanaka, 2010). Spheroplasts prepared from red colonies, thus expressing either Sup35p-WT or the WH1-R3-5 chimeras in their non-prion form ([psi<sup>-</sup>] or [rep-psi<sup>-</sup>], respectively), were transformed with protein extracts from cells also carrying the chimeras

but grown from white colonies, thus with a [REP-PSI<sup>+</sup>] phenotype. When the epigenetic phenotype of the transformants was assayed (Figure 3A), all colonies showed the white or the light pink stop codon read-through phenotype characteristic of the incoming prion aggregates (Sup35p-WT or the chimeras, respectively). The demonstration of the ‘infectivity’ of the R0 + WH1-R3-5 chimeras through the transformation of their aggregated forms qualifies [REP-PSI<sup>+</sup>] as a new synthetic yeast prion. However, after serial sub-culturing without selective pressure, in those clones expressing Sup35p-WT and cross-transformed with the WH1-R3-5 [REP-PSI<sup>+</sup>] chimeras, besides the pink phenotype of the incoming aggregates, red [psi<sup>-</sup>] colonies appeared frequently (Figure 3B). These results indicated that the WH1-R3-5 chimeras can template their conformation on Sup35p-WT to generate a new very weak prion strain (or ensemble of weak strains) that we have named [PSI<sup>+</sup>]<sup>WH1</sup>.

### [REP-PSI<sup>+</sup>] Prion Assembles Amyloids Larger than [PSI<sup>+</sup>]

Biochemical analysis of the solubility of the engineered chimeric proteins, performed upon their overexpression from the *GAL1* promoter, showed that constructs R0 + WH1-R3-5 aggregated massively, whereas R0 + WH1-R1-2 were proteolytically more unstable (Figure 4A). Regarding the extra, higher mobility band observed in the SDS-PAGE of the whole cellular lysates for R0 + WH1-R1, including a single WH1 hydrophobic repeat in the context of the Sup35p Q/N-rich sequences might generate instability in the natural beta-arcades built by the prion,





**FIGURE 3 | [REP-PSI<sup>+</sup>] aggregates template a prion phenotype on recipient [psi<sup>-</sup>] and [rep-psi<sup>-</sup>] cells. (A)** Yeast spheroplasts prepared from red [psi<sup>-</sup>] (left) or [rep-psi<sup>-</sup>] (right) colonies that had spontaneously lost the prion phenotype (red labels) were transformed with extracts from either [PSI<sup>+</sup>] or [REP-PSI<sup>+</sup>] cells (black labels). Ten-fold serial dilutions of cultures from the transformed colonies, starting with OD<sub>600</sub> = 0.001, were plated on 1/4YPD agar. [REP-PSI<sup>+</sup>] extracts were able to confer a light pink, weak [PSI<sup>+</sup>]-like phenotype to the recipient cells expressing Sup35p-WT (left), or a [REP-PSI<sup>+</sup>] phenotype to those expressing the indicated R0 + WH1-R3-5 chimeras (right). **(B)** Cells

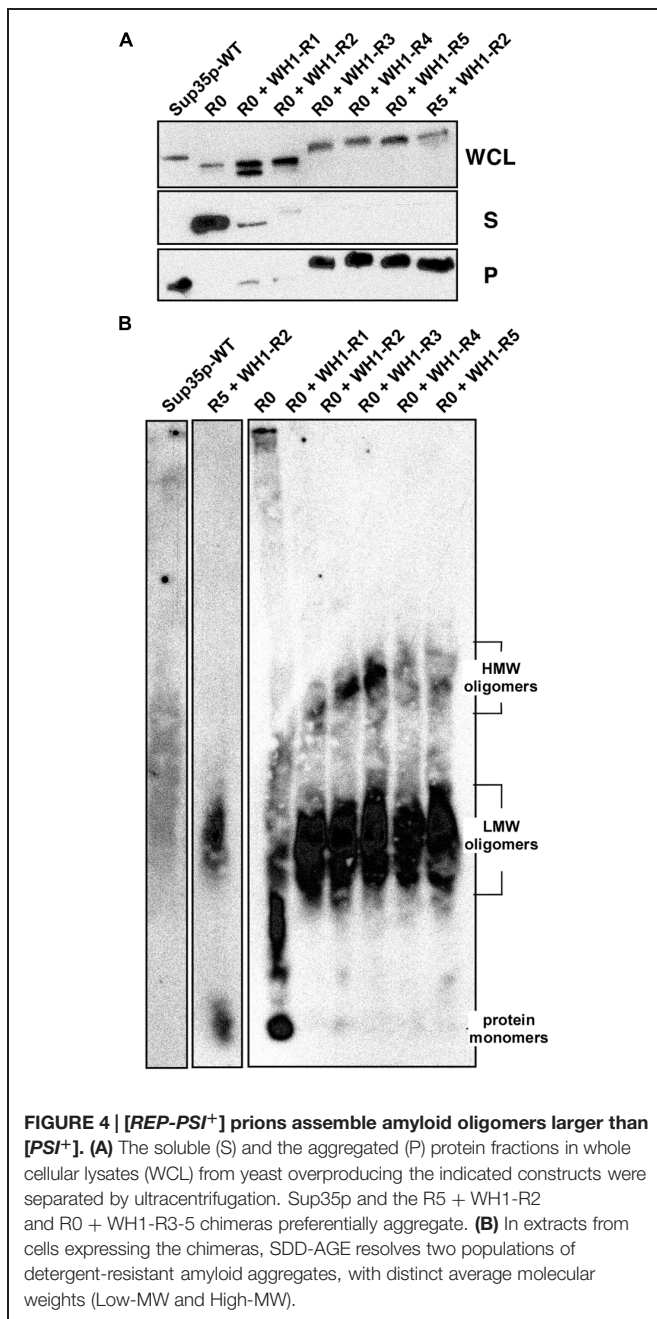
expressing Sup35p-WT and transformed with R0 + WH1-R3-5 [REP-PSI<sup>+</sup>] aggregates **(A, left)** were diluted 10<sup>-4</sup> in non-selective YPD medium and then plated on 1/4YPD agar. The high frequency of appearance of red colonies (arrows) besides the pink clones revealed the instability of a new prion variant, [PSI<sup>+</sup>]<sup>WH1</sup>, templated by the chimeras on Sup35p-WT. Right: a quantitative estimation of stability of the new prion variant, in the form of a histogram display of the mean values for the ratio between the number of pink [PSI<sup>+</sup>]<sup>WH1</sup> and red [psi<sup>-</sup>] colonies from two independent transformation rounds of the R0 + WH1-R3-5 aggregates with subsequent growth in YPD.

which will thus become proteolysis-prone. Following this speculation, that higher mobility band would correspond to molecules of the chimera that have lost the bit N-terminal to the single engineered WH1 repeat. The R0 + WH1-R1-2 chimeras were further degraded during the manipulation of the cell lysates for ultracentrifugation analysis, in spite of performing the experiment at low temperature and supplying the samples with protease inhibitors. Semi-denaturing detergent agarose gel electrophoresis (SDD-AGE), a technique for the detection of amyloid aggregates in a broad range of sizes (Bagriantsev et al., 2006), was carried out by immediately running in the gel the WCL fractions under conditions that should denature proteases, thus reducing degradation of the R0 + WH1-R1-2 chimeras to a minimum. SDD-AGE revealed (**Figure 4B**) that the chimeras form two populations of aggregates according to their electrophoretic mobilities. The species with higher molecular weights were evident in R0 + WH1-R1-5, showing sizes in direct correlation with the increasing number of RepA-WH1 repeats, but were barely detectable for Sup35p-WT, R0 or the construct carrying R5 + WH1-R2. This result points to the construction of a distinct type of assembly by the chimeras in which the hydrophobic repeats have completely replaced the Q/N-OPRs. The presence

of very high molecular weight species in SDD-AGE has been described in [PSI<sup>+</sup>] variants resistant to Hsp104 chaperone-promoted shearing, thus resulting in weak prion phenotypes with low mitotic stabilities (Derdowski et al., 2010; Alexandrov et al., 2012).

### RepA-WH1 Repeats in [REP-PSI<sup>+</sup>] Alter the Structure of the Amyloids Assembled by the NM-Domains

The results above pointed to a singular structural arrangement in [REP-PSI<sup>+</sup>] for the hydrophobic RepA-WH1 OPRs within the flanking polar Q/N-rich sequences, which remained unaltered from the Sup35p N-domain. The N and M domains (**Figure 1A**) in the [REP-PSI<sup>+</sup>] chimeras were then fused to the monomeric red fluorescent protein mCherry and expressed in *E. coli*. Those expressed at sufficiently high levels were purified and their assembly morphologies (**Figure 5A**), average secondary structure compositions (**Figure 5B**), and association states (**Figure 5C**) were physically analyzed. NM-mCherry and NM-R0-mCherry controls were able to assemble into fibers under standard conditions (Giraldo, 2007; Fernández-Tresguerres et al., 2010; **Figure 5A**). However, NM-mCherry fusions carrying WH1-R2 or WH1-R4



did not form fibers, but irregular oligomers whose average sizes ( $\leq 10$  and  $\approx 25$  nm, respectively), directly correlated with the number of hydrophobic repeats. Such oligomers were compatible with the smaller aggregates detected by means of SDD-AGE for the [REP-PSI<sup>+</sup>] chimeras (Figure 4B).

Circular dichroism (CD) analysis of the purified proteins (Figure 5B) indicated a net increase in  $\beta$ -sheet structure (broad band at  $\approx 220$  nm) for NM-R2/4-mCherry compared with NM-mCherry, whereas NM-R0-mCherry spectrum resembled the spectra of Q/N-rich peptides when forming coiled-coils (red-shifted band at  $> 225$  nm; Fiumara et al., 2010). Since sedimentation velocity experiments (Figure 5C) showed that purified

NM-R0-mCherry was a monomer ( $s = 2.3$  S), such coiled-coil should be intramolecular. On the contrary, poly-dispersed aggregation was evident as multiple peaks with increasing sedimentation coefficients for NM-R2/4-mCherry and the WT control NM-mCherry. These *in vitro* experiments indicated that the chimeric [REP-PSI<sup>+</sup>] prions, carrying hydrophobic OPRs of bacterial origin, assemble as amyloid oligomers rather than fibers, as Sup35p/[PSI<sup>+</sup>] does.

### [REP-PSI<sup>+</sup>] Prion Aggregates as Perivacuolar Foci *In Vivo*

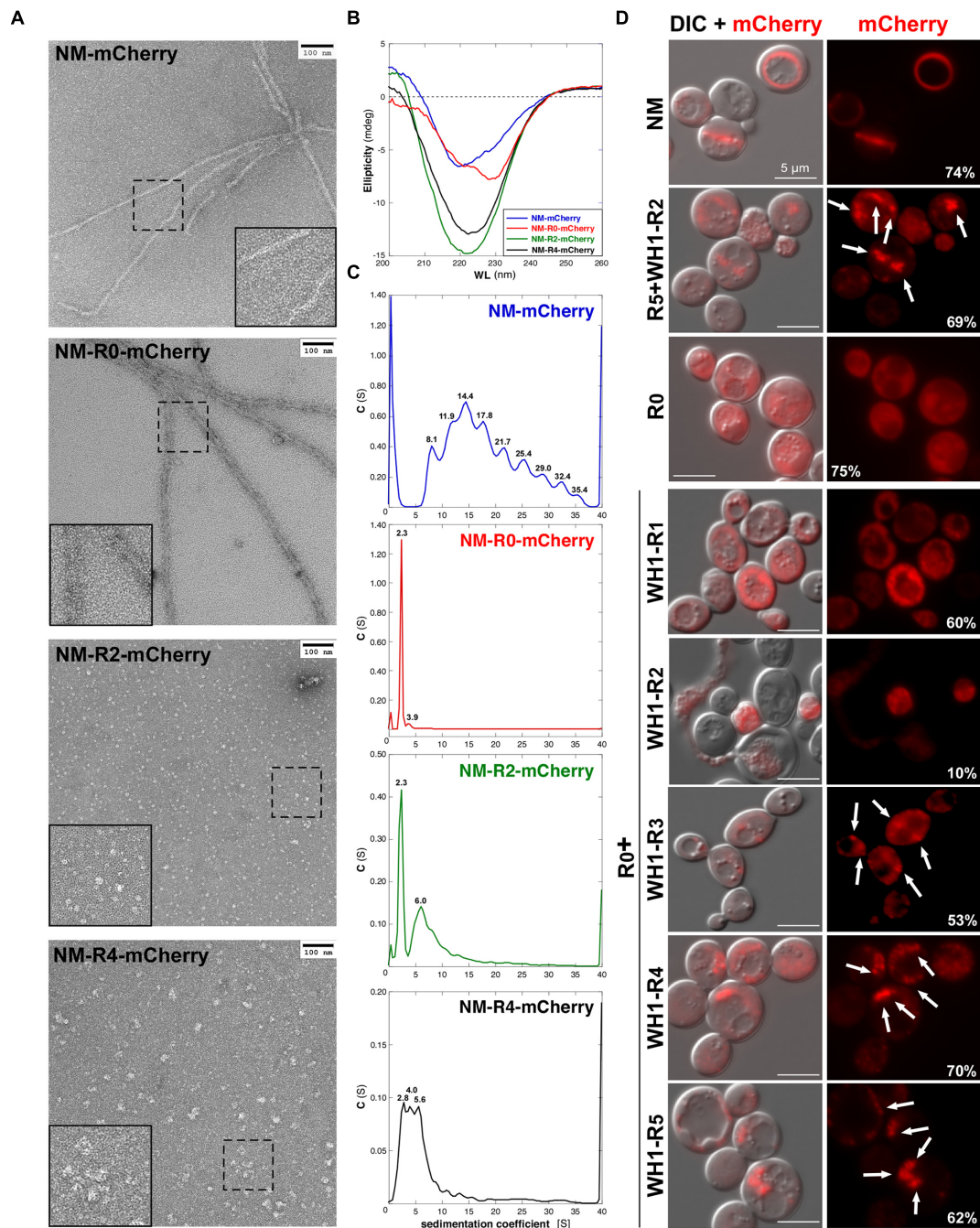
The NM-mCherry fusion proteins were then expressed in yeast from a Gal-inducible plasmid (Figure 5D). When the N domain carried the Sup35p wild-type sequence, the characteristic ring-like aggregates appeared in the cytoplasm (Tyedmers et al., 2010), whereas if it lacked all the OPRs (NM-R0-mCherry) fluorescence labeling was found diffused. If the constructs included the N domain of the distinct [REP-PSI<sup>+</sup>] chimeras, aggregation appeared as multiple dots/foci whose sizes increased with the number of RepA-WH1 repeats. They were disposed around the vacuole (IPOD compartment), as previously described for mature NM-GFP amyloids (Tyedmers et al., 2010). Interestingly, the fraction of yeast cells expressing the NM-R2-mCherry chimera was significantly reduced, and the lysis of many cells became evident. It has been proposed that the cellular toxicity of hydrophobic amyloidogenic peptides is linked to their ability to assemble as oligomeric pores upon insertion into lipid bilayers (reviewed in Butterfield and Lashuel, 2010). It might be the case that NM-R2-mCherry, lacking the ability to further assemble into a stable parallel superpleated  $\beta$ -structure (Figure 1B), would have a preference for membrane targeting.

## Discussion

In this work, we have generated chimeras (Figure 1) by replacing the polar OPRs in Sup35p with tandem repeats of the hydrophobic amyloidogenic stretch found in the bacterial prionoid RepA-WH1 (Giraldo, 2007; Gasset-Rosa et al., 2008a, 2014). The resulting synthetic [REP-PSI<sup>+</sup>] prions are functional in yeast, generating both a strong [PSI<sup>+</sup>]-like phenotype and various weak phenotypes (Figures 2 and 3), which are compatible with a cloud of prions strains (Bateman and Wickner, 2013). These chimeric prions assemble themselves *in vitro* as discrete size particles (Figure 5), which probably would suffice to act as competent propagons *in vivo*.

Our results suggest that cross-seeding of [REP-PSI<sup>+</sup>] through conformational templating, since the wild-type OPRs were absent from the chimeras, must be exerted by the flanking Q/N-rich sequences in the N-terminal domain that come from Sup35p-WT. This is compatible with findings showing that mutants including hydrophobic and/or aromatic residues at the N-terminus (residues 1–40) of Sup35p enhance [PSI<sup>+</sup>] nucleation, whereas aromatic side-chains at the OPRs promote chaperone-mediated propagation (Toombs et al., 2010, 2011; Alexandrov et al., 2012; Gonzalez-Nelson et al., 2014; MacLea et al., 2015). It is noteworthy that such studies were performed by





**FIGURE 5 | Chimeric NM domains from [REP-PSI<sup>+</sup>] fused to the mCherry reporter form oligomeric assemblies. (A)** Purified His<sub>6</sub>-tagged NM-mCherry fusions (15  $\mu$ M) were assembled *in vitro* and visualized by transmission electron microscopy. Insets are twofold magnifications of areas within dashed boxes. RepA-WH1 repeats drive the assembly of the NM-mCherry proteins into particles whose sizes increase with the number or repeats. **(B)** Circular dichroism (CD) spectra of the purified, preassembled NM-mCherry proteins (15  $\mu$ M) showed an increase in  $\beta$ -sheet structure (broad minima at  $\approx 220$  nm) for the R0 + WH1-R2/4 chimeras. **(C)** Sedimentation velocity experiments were performed in an analytical ultracentrifuge with the same purified NM-mCherry proteins (10  $\mu$ M) studied in **(A,B)**. The analyses indicate poly-dispersed aggregation for the chimeras (i.e., multi-peaked profiles), with a direct correlation between the number of

RepA-WH1 repeats and the sedimentation coefficient values/number of peaks. However, the profiles for the R0 + WH1-R2/4 chimeras were simpler than that for the NM-Sup35 fusion, reflecting the ability of the latter to assemble fibers **(A)**. **(D)** Epifluorescence microscopy imaging of yeast cells expressing the chimeric NM-mCherry fusions. Exposure times: 200 ms (Sup35-WT, R0, R0 + WH1-R1), 600 ms (R5 + WH1-R2, R0 + WH1-R3-5), and 2 s (R0 + WH1-R2). ND filter: 1/4. Left: superposition of the DIC and fluorescence images. Chimeras carrying R0 + WH1-R3-5 form foci (arrows) whose sizes increase with the number of WH1 repeats, whereas WH1-R1 is dispersed through the cytoplasm and R0 + WH1-R2 is cytotoxic (cell lysis). The fraction of cells expressing the chimera (i.e., those fluorescence-labeled) is indicated (%); 200–400 cells of each type were counted.

mutating a single OPR out of five in Sup35p, yielding constructs somehow analogous to our unstable R5 + WH1-R1-2 chimeras, whereas in this work we have built a complete assembly made of up to five hydrophobic repeats (R0 + WH1-R1-5). According to the results discussed here, there seems to be in  $[REP-PSI^+]$  a minimum threshold of three hydrophobic WH1 repeats in order to build a hydrophobic spine in the  $\beta$ -arcades sufficiently stable to surpass the control exerted by the surrounding natural polar sequences on prion nucleation and propagation. Thus, the engineered superpleated  $\beta$ -structure in  $[REP-PSI^+]$  seems to behave as an orthogonal synthetic module functional in prion propagation, in the sense proposed by Toombs et al. (2012). It has also been described that point mutations replacing polar Q/Y residues by charged Lys in OPRs positions confronted in the  $\beta$ -arcades, thus leading to electrostatic repulsion, result in new prion variants,  $[PSI^+]^{M1-M5}$ , some of which are unstable, and template on Sup35p-WT a non-epigenetically heritable conformation (Bondarev et al., 2013, 2014). The 'prion no more' mutant G58D, which affects the second OPR (R2) in Sup35p, can be incorporated in WT aggregates, leading to increased frequency of fragmentation, thus compensating for deficiencies in propagation of weak  $[PSI^+]$  variants but curing, as dominant-negative, strong prion variants (DiSalvo et al., 2011; Verges et al., 2011). We have shown here that even more extensive engineering of the OPRs can give way to a new prion,  $[REP-PSI^+]$ , that templates on Sup35p a new weak variant,  $[PSI^+]^{WH1}$ . Conversely, during the initial expression of the WH1-R1-5 chimeras from plasmids, the  $[PSI^+]$  prion resident in the yeast cells, before the displacement of its encoding (*SUP35*) vector, could influence through templating the compatibility and selection of the distinct  $[REP-PSI^+]$  variants described in this work. Cross-seeded aggregation between distinct RepA-WH1 variants has also been reported in bacteria (Molina-García and Giraldo, 2014). The possible contribution of the  $[PSI^+]$ -ancillary prion Rnq1p/ $[PIN^+]$  (Sharma and Liebman, 2013) to  $[REP-PSI^+]$  nucleation will require further studies.

Interestingly, the R0 + WH1-R3-5 chimeras build oligomers rather than fibers, as Sup35p does: most probably, having a large hydrophobic beta-arcade inserted in the polar Q/N-rich prion domain will impose a hindrance to the assembly of long, structurally regular fibers. Similarly, the natural, non-Q/N-rich prion  $[GAR^+]$  (Brown and Lindquist, 2009), which has recently been described to overcome glucose catabolite repression in *Saccharomyces cerevisiae* and other yeast (Jarosz et al., 2014), does not assemble as fibers but as oligomeric aggregates, as described here for the  $[REP-PSI^+]$  chimeras. In yeast, under normal conditions, binding of Ssa1p (an Hsp70 chaperone) to Sup35p targets Hsp104p to  $[PSI^+]$  aggregates for the generation of prion seeds (Winkler et al., 2012). Sup35p alleles defective in OPRs generate prions, such as  $[PSI^+]^{\Delta 22/69}$  (Borchsenius et al., 2001), that also build large aggregates behaving as unstable, weak prion variants (Tanaka et al., 2006; Derdowski et al., 2010). When the Q/N-rich

OPRs in *S. cerevisiae* Sup35p were replaced by heterologous non-Q/N sequences from other yeast species, the propagation of the resulting chimeric prion became independent on Hsp104p (Crist et al., 2003). The region between the OPRs and the initial residues in the medium (M) domain is the target recognized by Hsp104p in the Sup35p amyloids (Frederick et al., 2014). Besides this, the intermolecular contacts characteristic of weak  $[PSI^+]$  variants exhibit an increased dependence on Hsp70s for Hsp104p-driven shearing, compared with those found in strong variants (DeSantis and Shorter, 2012). The propagation of  $[GAR^+]$  is independent of the activity of the Hsp104 disaggregase, probably because the oligomeric nature of this prion allows for diffusion-driven propagation, but this becomes strictly dependent on Hsp70 (Brown and Lindquist, 2009; Jarosz et al., 2014). Interestingly, the bacterial Hsp104p orthologue, ClpB, does not contribute either to the propagation of the RepA-WH1 prionoid in *E. coli*, a function relying mainly on the Hsp70 chaperone DnaK (Gasset-Rosa et al., 2014). DnaK conformationally selects for an amyloid variant of RepA-WH1 with reduced toxicity and generates relatively small oligomeric particles, readily diffusible to the progeny (Gasset-Rosa et al., 2014). In the case of  $[REP-PSI^+]$ , the contribution of chaperones to its propagation in yeast remains to be explored.

Heterologous model systems have made fundamental contributions to the understanding of prion propagation. The expression of the yeast prion  $[PSI^+]$  in bacteria has revealed that Sup35p aggregates as inclusion bodies which retain the ability to nucleate distinct strains (Garritty et al., 2010; Espargaró et al., 2012), and that, as in its original host, it still needs nucleation by  $[PIN^+]$  (Garritty et al., 2010) and depends on Hsp104 for propagation (Yuan et al., 2014). In addition,  $[PSI^+]$  has also been propagated in mammalian cells in culture, exhibiting the hallmarks of cytoplasmic inheritance (Krammer et al., 2009; Hofmann et al., 2013). The synthetic  $[REP-PSI^+]$  prion, by expanding the repertoire of  $[PSI^+]$  variants and rewiring amyloidogenic parts of Sup35p with alien, non-Q/N-rich sequences of bacterial origin, is a proof of concept for the feasibility of generating new phenotypic traits in prions. Engineering the consortium between  $[REP-PSI^+]$  and its possible chaperone modulators will surely enable new trends in yeast epigenetics.

## Acknowledgments

We are grateful to M. Tuite for the gift of the pUKC1620 plasmid and the derivative of the strain 74D-694 used in the read-through assays, and J. A. Tercero for the strain YJT28. Thanks are also due to Y. Chernoff for advice on protein transformation into yeast and to J. R. Luque for performing the sedimentation velocity experiments at the CIB – CSIC analytical ultracentrifugation facility. This work has been supported by grants from Spanish MINECO (BIO2012-30852 and CSD2009-00088).

## References

Aguzzi, A. (2009). Beyond the prion principle. *Nature* 459, 924–925. doi: 10.1038/459924a

Alberti, S., Halfmann, R., King, O., Kapila, A., and Lindquist, S. (2009). A systematic survey identifies prions and illuminates sequence features of prionogenic proteins. *Cell* 137, 146–158. doi: 10.1016/j.cell.2009.02.044



- Alexandrov, A. I., Polyanskaya, A. B., Serpionov, G. V., Ter-Avanesyan, M. D., and Kushnirov, V. V. (2012). The effects of amino acid composition of glutamine-rich domains on amyloid formation and fragmentation. *PLoS ONE* 7:e46458. doi: 10.1371/journal.pone.0046458
- Bagriantsev, S. N., Kushnirov, V. V., and Liebman, S. W. (2006). Analysis of amyloid aggregates using agarose gel electrophoresis. *Methods Enzymol.* 412, 33–48. doi: 10.1016/S0076-6879(06)12003-0
- Bateman, D. A., and Wickner, R. B. (2013). The [PSI<sup>+</sup>] prion exists as a dynamic cloud of variants. *PLoS Genet.* 9:e1003257. doi: 10.1371/journal.pgen.1003257
- Bondarev, S. A., Shchepachev, V. V., Kajava, A., and Zhouravleva, G. A. (2013). Effect of charged residues in the N-domain of Sup35 protein on prion [PSI<sup>+</sup>] stability and propagation. *J. Biol. Chem.* 288, 28503–28513. doi: 10.1074/jbc.M113.471805
- Bondarev, S. A., Shirokolobova, E. D., Trubitsina, N. P., and Zhouravleva, G. A. (2014). Modification of [PSI<sup>+</sup>] prion properties by combining amino acid changes in N-terminal domain of Sup35 protein. *Mol. Biol.* 48, 270–277. doi: 10.1134/S0026893314020034
- Borchsenius, A. S., Wegryn, R. D., Newnam, G. P., Inge-Vechtomov, S. G., and Chernoff, Y. O. (2001). Yeast prion protein derivative defective in aggregate shearing and production of new 'seeds.' *EMBO J.* 20, 6683–6691. doi: 10.1093/emboj/20.23.6683
- Brown, J. C., and Lindquist, S. (2009). A heritable switch in carbon source utilization driven by an unusual yeast prion. *Genes Dev.* 23, 2320–2332. doi: 10.1101/gad.1839109
- Butterfield, S. M., and Lashuel, H. S. (2010). Amyloidogenic protein-membrane interactions: mechanistic insight from model systems. *Angew. Chem. Int. Ed. Engl.* 49, 5628–5654. doi: 10.1002/anie.200906670
- Chernoff, Y. O., Lindquist, S. L., Ono, B., Inge-Vechtomov, S. G., and Liebman, S. W. (1995). Role of the chaperone protein Hsp104 in propagation of the yeast prion-like factor [PSI<sup>+</sup>]. *Science* 268, 880–884. doi: 10.1126/science.7754373
- Conchillo-Solé, O., de Groot, N. S., Avilés, F. X., Vendrell, J., Daura, X., and Ventura, S. (2007). AGGRESCAN: a server for the prediction and evaluation of "hot spots" of aggregation in polypeptides. *BMC Bioinformatics* 8:65. doi: 10.1186/1471-2105-8-65
- Cox, B., Ness, F., and Tuite, M. (2003). Analysis of the generation and segregation of propagons: entities that propagate the [PSI<sup>+</sup>] prion in yeast. *Genetics* 165, 23–33.
- Crist, C. G., Kakayashiki, T., Kurahashi, H., and Nakamura, Y. (2003). [PHI<sup>+</sup>], a novel Sup35-prion variant propagated with non-Gln/Asn oligopeptide repeats in the absence of the chaperone protein Hsp104. *Genes Cells* 8, 603–618. doi: 10.1046/j.1365-2443.2003.00661.x
- Cullin, C., and Minvielle-Sebastia, L. (1994). Multipurpose vectors designed for the fast generation of N- or C-terminal epitope-tagged proteins. *Yeast* 10, 105–112. doi: 10.1002/yea.320100110
- Derdowski, A., Sindi, S., Klaipe, C. L., DiSalvo, S., and Serio, T. R. (2010). A size threshold limits prion transmission and establishes phenotypic diversity. *Science* 330, 680–683. doi: 10.1126/science.1197785
- DeSantis, M. E., and Shorter, J. (2012). Hsp104 drives 'protein-only' positive selection of Sup35 prion strains encoding strong [PSI<sup>+</sup>]. *Chem. Biol.* 19, 1400–1410. doi: 10.1016/j.chembiol.2012.09.013
- Díaz-López, T., Dávila-Fajardo, C., Blaessing, F., Lillo, M. P., and Giraldo, R. (2006). Early events in the binding of the pPS10 replication protein RepA to single iteron and operator DNA sequences. *J. Mol. Biol.* 364, 909–920. doi: 10.1016/j.jmb.2006.09.013
- Díaz-López, T., Lages-Gonzalo, M., Serrano-López, A., Alfonso, C., Rivas, G., Díaz-Orejas, R., et al. (2003). Structural changes in RepA, a plasmid replication initiator, upon binding to origin DNA. *J. Biol. Chem.* 278, 18606–18616. doi: 10.1074/jbc.M212024200
- DiSalvo, S., Derdowski, A., Pezza, J. A., and Serio, T. R. (2011). Dominant prion mutants induce curing through pathways that promote chaperone-mediated disaggregation. *Nat. Struct. Mol. Biol.* 18, 486–492. doi: 10.1038/nsmb.2031
- Dong, J., Bloom, J. D., Goncharov, V., Chattopadhyay, M., Millhauser, G. L., Lynn, D. G., et al. (2007). Probing the role of PrP repeats in conformational conversion and amyloid assembly of chimeric yeast prions. *J. Biol. Chem.* 282, 34204–34212. doi: 10.1074/jbc.M704952200
- Eisenberg, D., and Jucker, M. (2012). The amyloid state of proteins in human diseases. *Cell* 148, 1188–1203. doi: 10.1016/j.cell.2012.02.022
- Espargaró, A., Villar-Piqué, A., Sabaté, R., and Ventura, S. (2012). Yeast prions form infectious amyloid inclusion bodies in bacteria. *Microb. Cell Fact.* 11:89. doi: 10.1186/1475-2859-11-89
- Fernández-Tresguerres, M. E., de la Espina, S. M., Gasset-Rosa, F., and Giraldo, R. (2010). A DNA-promoted amyloid proteinopathy in *Escherichia coli*. *Mol. Microbiol.* 77, 1456–1469. doi: 10.1111/j.1365-2958.2010.07299.x
- Fiumara, F., Fioriti, L., Kandel, E. R., and Hendrickson, W. A. (2010). Essential role of coiled coils for aggregation and activity of Q/N-rich prions and poly-Q proteins. *Cell* 143, 1121–1135. doi: 10.1016/j.cell.2010.11.042
- Frederick, K. K., Debelouchina, G. T., Kayatekin, C., Dorminy, T., Jacavone, A. C., Griffin, R. G., et al. (2014). Distinct prion strains are defined by amyloid core structure and chaperone binding site dynamics. *Chem. Biol.* 21, 295–305. doi: 10.1016/j.chembiol.2013.12.013
- Garrity, S. J., Sivanathan, V., Dong, J., Lindquist, S., and Hochschild, A. (2010). Conversion of a yeast prion protein to an infectious form in bacteria. *Proc. Natl. Acad. Sci. U.S.A.* 107, 10596–10601. doi: 10.1073/pnas.0913280107
- Gasset-Rosa, F., Coquel, A. S., Moreno-del Álamo, M., Chen, P., Song, X., Serrano, A. M., et al. (2014). Direct assessment in bacteria of prionoid propagation and phenotype selection by Hsp70 chaperone. *Mol. Microbiol.* 91, 1070–1087. doi: 10.1111/mmi.12518
- Gasset-Rosa, F., Díaz-López, T., Lurz, R., Prieto, A., Fernández-Tresguerres, M. E., and Giraldo, R. (2008a). Negative regulation of pPS10 plasmid replication: origin pairing by zipping-up DNA-bound RepA monomers. *Mol. Microbiol.* 68, 560–572. doi: 10.1111/j.1365-2958.2008.06166.x
- Gasset-Rosa, F., Maté, M. J., Dávila-Fajardo, C., Bravo, J., and Giraldo, R. (2008b). Binding of sulphonated indigo derivatives to RepA-WH1 inhibits DNA-induced protein amyloidogenesis. *Nucleic Acids Res.* 36, 2249–2256. doi: 10.1093/nar/gkn067
- Giraldo, R. (2007). Defined DNA sequences promote the assembly of a bacterial protein into distinct amyloid nanostructures. *Proc. Natl. Acad. Sci. U.S.A.* 104, 17388–17393. doi: 10.1073/pnas.0702006104
- Giraldo, R., Andreu, J. M., and Díaz-Orejas, R. (1998). Protein domains and conformational changes in the activation of RepA, a DNA replication initiator. *EMBO J.* 17, 4511–4526. doi: 10.1093/emboj/17.15.4511
- Giraldo, R., Fernández-Tornero, C., Evans, P. R., Díaz-Orejas, R., and Romero, A. (2003). A conformational switch between transcriptional repression and replication initiation in the RepA dimerization domain. *Nat. Struct. Biol.* 10, 565–571. doi: 10.1038/nsb937
- Giraldo, R., and Fernández-Tresguerres, M. E. (2004). 20 years of the pPS10 replicon: insights on the molecular mechanism for the activation of DNA replication in iteron-containing bacterial plasmids. *Plasmid* 52, 69–83. doi: 10.1016/j.plasmid.2004.06.002
- Gonzalez-Nelson, A. C., Paul, K. R., Petri, M., Flores, N., Rogge, R. A., Cascarina, S. M., et al. (2014). Increasing prion propensity by hydrophobic insertion. *PLoS ONE* 9:e89286. doi: 10.1371/journal.pone.0089286
- Hofmann, J. P., Denner, P., Nussbaum-Krammer, C., Kuhn, P. H., Suhre, M. H., Scheibel, T., et al. (2013). Cell-to-cell propagation of infectious cytosolic protein aggregates. *Proc. Natl. Acad. Sci. U.S.A.* 110, 5951–5956. doi: 10.1073/pnas.1217321110
- Jarosz, D. F., Lancaster, A. K., Brown, J. C. S., and Lindquist, S. (2014). An evolutionary conserved prion-like element converts wild fungi from metabolic specialists to generalists. *Cell* 158, 1072–1082. doi: 10.1016/j.cell.2014.07.024
- Kajava, A. V., Baxa, U., and Steven, A. C. (2010).  $\beta$ -arcades: recurring motifs in naturally occurring and disease-related amyloid fibrils. *FASEB J.* 24, 1311–1319. doi: 10.1096/fj.09-145979
- Krammer, C., Kryndushkin, D., Suhre, M. H., Kremmer, E., Hofmann, A., Pfeifer, A., et al. (2009). The yeast Sup35NM domain propagates as a prion in mammalian cells. *Proc. Natl. Acad. Sci. U.S.A.* 106, 462–467. doi: 10.1073/pnas.0811571106
- Liebman, S. W., and Chernoff, Y. O. (2012). Prions in yeast. *Genetics* 191, 1041–1072. doi: 10.1534/genetics.111.137760
- MacLea, K. S., Paul, K. R., Ben-Musa, Z., Waechter, A., Shattuck, J. E., Gruca, M., et al. (2015). Distinct amino acid compositional requirements for formation and maintenance of the [PSI<sup>+</sup>] prion in yeast. *Mol. Cell. Biol.* 35, 899–911. doi: 10.1128/MCB.01020-14
- Maurer-Stroh, S., Debulpaep, M., Kuemmerer, N., Lopez de la Paz, M., Martins, I. C., Reumers, J., et al. (2010). Exploring the sequence determinants of amyloid

- structure using position-specific scoring matrices. *Nat. Methods* 7, 237–242. doi: 10.1038/nmeth.1432
- Molina-García, L., and Gasset-Rosa, F. (2014). Semi-denaturing detergent agarose gel electrophoresis (SDD-AGE). *Bio-protocol* 4:e1297.
- Molina-García, L., and Giraldo, R. (2014). Aggregation interplay between variants of the RepA-WH1 prionoid in *Escherichia coli*. *J. Bacteriol.* 196, 2536–2542. doi: 10.1128/JB.01527-14
- Parham, S. N., Resende, C. G., and Tuite, M. F. (2001). Oligopeptide repeats in the yeast protein Sup35p stabilize intermolecular prion interactions. *EMBO J.* 20, 2111–2119. doi: 10.1093/emboj/20.9.2111
- Sabate, R., Rousseau, F., Schymkowitz, J., and Ventura, S. (2015). What makes a protein sequence a prion? *PLoS Comput. Biol.* 11:e1004013. doi: 10.1371/journal.pcbi.1004013
- Schuck, P. (2000). Size distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and Lamm equation modeling. *Biophys. J.* 78, 1606–1619. doi: 10.1016/S0006-3495(00)76713-0
- Sharma, J., and Liebman, S. W. (2013). Exploring the basis of [PIN<sup>+</sup>] variant differences in [PSI<sup>+</sup>] induction. *J. Mol. Biol.* 425, 3046–3059. doi: 10.1016/j.jmb.2013.06.006
- Shorter, J., and Lindquist, S. (2008). Hsp104, Hsp70 and Hsp40 interplay regulates formation, growth and elimination of Sup35 prions. *EMBO J.* 27, 2712–2724. doi: 10.1038/emboj.2008.194
- Tanaka, M. (2010). A protein transformation protocol for introducing yeast prion particles into yeast. *Methods Enzymol.* 470, 681–693. doi: 10.1016/S0076-6879(10)70028-8
- Tanaka, M., Collins, S. R., Toyama, B. H., and Weissman, J. S. (2006). The physical basis of how prion conformations determine strain phenotypes. *Nature* 442, 585–589. doi: 10.1038/nature04922
- Tank, E. M., Harris, D. A., Desai, A. A., and True, H. L. (2007). Prion protein repeat expansion results in increased aggregation and reveals phenotypic variability. *Mol. Cell. Biol.* 27, 5445–5455. doi: 10.1128/MCB.02127-06
- Tessier, P. M., and Lindquist, S. (2009). Unraveling infectious structures, strain variants and species barriers for the yeast prion [PSI<sup>+</sup>]. *Nat. Struct. Mol. Biol.* 16, 598–605. doi: 10.1038/nsmb.1617
- Toombs, J. A., Liss, N. M., Cobble, K. R., Ben-Musa, Z., and Ross, E. D. (2011). [PSI<sup>+</sup>] maintenance is dependent on the composition, not primary sequence, of the oligopeptide repeat domain. *PLoS ONE* 6:e21953. doi: 10.1371/journal.pone.0021953
- Toombs, J. A., McCarty, B. R., and Ross, E. D. (2010). Compositional determinants of prion formation in yeast. *Mol. Cell. Biol.* 30, 319–332. doi: 10.1128/MCB.01140-09
- Toombs, J. A., Petri, M., Paul, K. R., Kan, G. Y., Ben-Hur, A., and Ross, E. D. (2012). De novo design of synthetic prion domains. *Proc. Natl. Acad. Sci. U.S.A.* 109, 6519–6524. doi: 10.1073/pnas.1119366109
- Torreira, E., Moreno-del Álamo, M., Fuentes-Perez, M. E., Fernández, C., Martín-Benito, J., Moreno-Herrero, F., et al. (2015). Amyloidogenesis of the bacterial prionoid RepA-WH1 recapitulates dimer to monomer transitions of RepA in DNA replication initiation. *Structure* 23, 183–189. doi: 10.1016/j.str.2014.11.007
- Tyedmers, J., Treusch, S., Dong, J., McCaffery, J. M., Bevis, B., and Lindquist, S. (2010). Prion induction involves an ancient system for the sequestration of aggregated proteins and heritable changes in prion fragmentation. *Proc. Natl. Acad. Sci. U.S.A.* 107, 8633–8638. doi: 10.1073/pnas.1003895107
- Verges, K. J., Smith, M. H., Toyama, B. H., and Weissman, J. S. (2011). Strain conformation, primary structure and the propagation of the yeast prion [PSI<sup>+</sup>]. *Nat. Struct. Mol. Biol.* 18, 493–499. doi: 10.1038/nsmb.2030
- Vitrenko, Y. A., Gracheva, E. O., Richmond, J. E., and Liebman, S. W. (2007). Visualization of aggregation of the Rnq1 prion domain and cross-seeding interactions with Sup35NM. *J. Biol. Chem.* 282, 1779–1787. doi: 10.1074/jbc.M609269200
- Von der Haar, Y., Jossé, L., Wright, P., Zenthon, J., and Tuite, M. F. (2007). Development of a yeast cell-based system for studying the aggregation of Alzheimer's disease-associated A beta peptides in vivo. *Neurodegener. Dis.* 4, 136–147. doi: 10.1159/000101838
- Winkler, J., Tyedmers, J., Bukau, B., and Mogk, A. (2012). Hsp70 targets Hsp104 chaperones to substrates for protein disaggregation and prion fragmentation. *J. Cell Biol.* 198, 387–404. doi: 10.1083/jcb.201201074
- Yuan, A. H., Garrity, S. J., Nako, E., and Hochschild, A. (2014). Prion propagation can occur in a prokaryote and requires the ClpB chaperone. *Elife* 3:e02949. doi: 10.7554/eLife.02949

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Gasset-Rosa and Giraldo. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

