Biochemical, structural and genetic studies on the cytoskeletal proteins Fibril and MreBs from *Spiroplasma*

A thesis

submitted in partial fulfillment of the requirements

of the degree of

Doctor of Philosophy

By

Shrikant R. Harne

20133258

Indian Institute of Science Education and Research, Pune

2019

DECLARATION

I declare that this written submission represents my ideas in my own words and where others' ideas have been included, I have adequately cited and referenced the original sources. I also declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in my submission. I understand that violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

Shrikant Ramdas Harne

Reg. No. 20133258

Date: 25 July 2019

Indian Institute of Science Education and Research Pune
Dr. Homi Bhabha Road, Pashan, Pune 411 008.

Dr. Gayathri Pananghat
Assistant Professor

CERTIFICATE

Certified that the work incorporated in the thesis entitled **"Biochemical, structural and genetic studies on the cytoskeletal proteins Fibril and MreBs from *Spiroplasma*"** submitted by **Mr. Shrikant Ramdas Harne** was carried out by the candidate, under my supervision. The work presented here or any part of it has not been included in any other thesis submitted previously for the award of any degree or diploma from any other university or institution.

Dr. Gayathri Pananghat
Supervisor

Date: 25 July 2019

Tel.: +91 020 2590 8128; Fax: +91 020 2586 5315 Web: www.iiserpune.ac.in/~gayathri

Acknowledgement

PhD has been a great experience for me. It was not possible to sail through this journey without support from many people and funding agencies. Thus, I would like to acknowledge people and agencies who have contributed to my overall development during this period.

To begin with, I would like to express my sincere gratitude to my PhD advisor Dr. Gayathri Pananghat for accepting me as a graduate student in her lab where I learnt science and (lab) management. Due to the exposure on various cytoskeletal proteins and techniques, I could learn about cytoskeleton from the scratch. Being her first PhD student, I was fortunate enough to learn techniques from her 'first hand' and manage lab. She has always been open to discuss ideas and explore new techniques. My PhD has been an amazing experience, thanks to the scientific freedom I enjoyed in her laboratory. I also thank her for preparing codon optimized full length fibril clones before I joined the lab. I am thankful to her for reviewing my project & fellowship proposals, for letting me drive projects and for critical evaluation of experimental data. She has been supportive mentor throughout my tenure in her lab.

I wish to thank my research advisory committee members Dr. Richa Rikhy (IISER, Pune) and Dr. Janesh Kumar (NCCS, Pune) for their valuable inputs and guidance on my work on regular basis. I am thankful to Dr. Saikrishnan Kayarat (IISER, Pune) for his critical evaluation and feedbacks on my work during the lab meetings throughout these 6 years. I thank Prof. L. S. Shashidhara for his support.

My sincere thanks also goes to Dr. Laure Béven (INRA, Bordeaux, France) and Dr. Martin Pilhofer (ETH, Zürich, Switzerland) for accepting me as short-term student in their laboratories and making resources available for the collaborative work. Special thanks to Ms. Sybille Duret (INRA, Bordeaux, France) and Dr. Laure Béven for their help with S. citri cultures, experiments and for continuing these experiments beyond my tenure at INRA, Bordeaux. I am also thankful to Dr. Martin Pilhofer and Dr. Piotr Szwedziak (ETH, Zürich, Switzerland) for their help with electron cryotomography experiments. In absence of required resources in India, the experiments performed in Chapter 4 and 5 of my thesis were impossible without support from these collaborators. I am thankful to Dr. Ashok Giri, (CSIR-NCL, Pune) for providing resources related to work on P. pastoris.

I am obliged to IISER Pune (PhD fellowship), EMBO (short-term fellowship), CEFIPRA (Raman-Charpak fellowship), Biophysical society and Infosys Foundation for financial support. I am thankful to electron microscopy facilities and associated staff at IISER Pune, NCBS & IISc Bangalore, ETH Zürich, IECB Bordeaux for facilitating processing of my samples.

I am grateful to Dr. Ramanujam Srinivasan (NISER, Bhubaneshwar), Dr. Thomas Pucadyil and Dr. Nishad Matange for useful scientific discussions related to my work. The discussions (not related to work) with Dr. Sudha Rajamani and Dr. Anand Krishnan were encouraging and refreshing.

I would like to thank my friends Sachin, Amey, Kunalika, Gregor and Manasi for valuable scientific discussions. I would also like to thank Rajnandani Kashyap for her help with preparation and characterization of short constructs of Fibril. I am thankful to Mrinmayee Bapat for the help with one of the clone preparation. Thanks to my summer/semester trainees for helping me learn to be a good mentor.

The long list of friends whom I want to thank begins with my 'Ajinkyan' friends who have always supported me. My special thanks to 'non-Ajinkyan' friends Manish, Manasi, Amey, Kunalika, Neha, Neelesh, Jyoti, Devika, Maithilee, Aboli, Ron, Mukul, Swati and my batchmates (2013 PhD batch) for making this journey enjoyable.

I am thankful to my fellow labmates from G3 and SK labs for their co-operation. I will always cherish the discussions I had over tea with Jyoti, Mahesh, Om, Vinayak, Pratima and Basila. I acknowledge Academic office staff, Biology admin staff and library staff for their prompt help.

I am indebted to my parents for educating me at the cost of their comforts. I am obliged to my brothers and sisters in-law for their constant support. I am also thankful to my in-laws for the encouragement and support. I am grateful to my boss at home, Kunalika, for all the support and for always being with me. She has always helped me out in times of trouble. Thanks to my sister and her family for the support. My nieces Sanvi, Siya and Niya deserve a special mention in the thesis for they have been my stress relievers on a regular basis.

Shrikant R. Harne

25th July 2019

Table of Contents

Chapter 2: Purification and characterization of full-length Fibril filaments

List of Figures

Chapter 4

Chapter 5

List of Tables

Synopsis

Biochemical, structural and genetic studies on the cytoskeletal proteins Fibril and MreBs from *Spiroplasma*

Name: Shrikant R. Harne

Registration number: 20133258

Supervisor: Dr. Gayathri Pananghat

Department: Biology, IISER, Pune

Date of registration: August 2013

Spiroplasma is a group of cell wall-less bacteria belonging to class *Mollicutes* (Rottem and Razin, 1973). The characteristic helical shape and kinking motility distinguishes *Spiroplasma* from other bacteria under class *Mollicutes*. The bacteria are also unique in possessing Fibril, a protein with no homolog in available gene or protein database (Cohen-Krausz et al., 2011). *Spiroplasma* species have been reported to infect a wide range of hosts including *Drosophila*, crabs, honey bees, aphids, shrimps, plants, etc. (Bové et al., 2003; Clark et al., 1985; Markham et al., 1974; Nunan et al., 2004; Saillard et al., 1987). Due to the maintenance of a distinct helical shape and motility in absence of external appendages, *Spiroplasma* has attracted attention from scientific community since 1970s. Studies have speculated the role of Fibril and MreB proteins in shape determination and motility of *Spiroplasma* (Kürner et al., 2005; Razin, 1978). These proteins filaments have been proposed to produce kinks essential for *Spiroplasma* motility (Kürner et al., 2005; Shaevitz et al., 2005).

Based on cryo electron microscopy studies on Fibril filaments (Cohen-Krausz et al., 2011), a model was proposed for role of Fibril in *Spiroplasma* motility. According to the model, Fibril filaments are homo-polymers of a monomeric 59 kDa protein. The repeating unit in the filament is the Fibril tetramer. The tetramers undergo conformational changes and result into net variation in filament lengths. The co-ordinated variation of lengths between two sets of Fibril filaments leads to formation of kinks in the cell body (Cohen-Krausz et al., 2011). A pair of kinks drive the

cell forward (Shaevitz et al., 2005). However, the exact nature of conformational changes occurring in Fibril and the forces that bring about conformational changes are unknown.

MreB, a bacterial actin, is responsible for positioning cell wall synthesis machinery in cell walled bacteria, thereby imparting a rod shape to the bacterium. However, the mechanism of how MreBs function in cell wall-deficient bacteria is unknown. Since *Spiroplasma* possesses at least 5 copies of *mreB* gene (Ku et al., 2014), it is likely that these proteins perform different roles in shape determination and motility of *Spiroplasma.*

Studies to answer these interesting questions have been limited due to the lack of suitable molecular genetic tools to modify *Spiroplasma* genome. The only strategy available for generation of a target gene knockout mutant of *Spiroplasma* is by use of a self-replicating *oriC*-based disruption vectors (Duret et al., 1999; Lartigue et al., 2002). However, the method involves extensive passaging and usually requires at least about a year to generate a single gene deletion mutant.

With this information available from the literature, I decided to perform biochemical and structural characterization of Fibril in polymerized and monomeric form, by heterologous expression approach. Further, to obtain insights into the proteins involved in shape determination and motility, comparative studies were performed on a naturally occurring non-helical, non-motile mutant of *Spiroplasma* with wildtype cells. Thus gene disruption mutants of *Spiroplasma* to delineate functions of each MreB and Fibril protein were prepared.

The aim of my thesis is to understand roles of Fibril and MreB proteins in shape determination and motility of *Spiroplasma*. To achieve this aim, the work is divided into following objectives- 1) Perform biochemical and structural characterization of Fibril filaments, 2) Identification of domain boundaries of Fibril to understand roles of each domain, 3) Comparative studies on *S. citri* GII-3 (wildtype) and ASP-1 (naturally occurring non-helical, non-motile) cells 4) Gene disruption experiments on *Spiroplasma* to decipher roles of each MreB and Fibril. Work done towards achieving each of these objectives is organized as chapters in my thesis. Following is the list of chapters and summary of relevant work carried out as part of my PhD research.

Chapter 1: Introduction

This chapter summarizes the information available in the literature on *Spiroplasma* beginning from their discovery to pathogenicity. The chapter also describes the current model proposed for *Spiroplasma* motility. Further discussed are the tools available for specific gene disruption. It also enlists the unanswered questions in the field and highlights why *Spiroplasma* can be employed as a model organism to study shape determination and motility in a cell wall-deficient bacterium.

Chapter 2: Purification and characterization of full length Fibril filaments

In this chapter I have discussed the need for heterologous expression of *Spiroplasma* proteins. The chapter also summarizes multiple protocols that I have standardized for purification of Fibril using different heterologous systems and challenges associated with it. An attempt to obtain partially purified Fibril from the native organism *Spiroplasma citri* is also described. The Fibril filaments purified from *E. coli* and *S. citri* cells have been visualized using electron microscopy.

Chapter 3: Identification of domain boundaries of Fibril

Since Fibril does not have a homolog in the available databases and no structural information is available on it, the domain boundaries of Fibril are unknown. Identification of domain boundaries is useful for design of non-polymerizing soluble constructs, and for delineating the minimal segment required for polymerization. In this chapter, I have described the logic for preparation of short constructs based on information obtained from purification of full length Fibril filaments and also on the basis of sequence alignment. My attempts to purify short constructs suggest that the shorter constructs consisting of the N-terminal domain can also polymerize on their own.

Chapter 4: Comparative characterization of *Spiroplasma citri* GII-3 and ASP-I cells

This chapter summarizes the comparative studies between a naturally occurring non-helical, non-motile mutant and helical, motile wildtype cells of *S. citri*. The genome comparison during this study suggested that a point mutation in one of the *mreB* genes in the mutant cells made them defective. Simultaneously, preliminary electron cryotomography studies of the whole organism revealed that the mutant cells have only one set of filaments present underneath the membrane as compared to three sets reported for *S. melliferum* wildtype cells. Complementation experiments involving vector-based expression of the functional MreB5 in *S. citri* ASP-I cells has instated

helicity and motility in these cells, thus highlighting the role of MreB in rod-to-helical shape transition.

Chapter 5: Genetic studies on *Spiroplasma citri* to understand the role of Fibril and MreB proteins in shape determination and motility

Here I summarize the strategy employed and work carried out towards preparation of *mreB* or *fibril* gene disruption mutants of *Spiroplasma*. To investigate the roles of Fibril and MreB, gene disruption mutants of *Spiroplasma* lacking either one of the *mreBs* or *fibril* were designed. Since *Spiroplasma mreBs* share high sequence identity, obtaining a single gene disruption mutant at a time is challenging. I have cloned the variable regions from *mreB* and *fibril* genes into suitable disruption vector and obtained *Spiroplasma* transformants. Upon confirmation of gene disruption, the mutants will be characterized for their shape and motility defects, if any.

Chapter 6: Summary and future prospects

In this chapter, I have summarized the work done on biochemical and structural characterization of Fibril, comparative characterization of wildtype versus mutant cells of *Spiroplasma* as well as genetic studies carried out to decipher functions of MreBs and Fibril. I have proposed further characterization of full length Fibril and alternate ways of purifying short constructs for structural studies. As an alternate approach to preparation and characterization of short constructs, an EGFP-tagging based screen has been proposed for identification of polymerization interface of a novel protein. Experiments to identify interacting partner(s) of MreB in the cell are suggested. Also proposed are experiments that can be performed on gene disruption mutants to gain insights into *Spiroplasma* physiology. The insights obtained using these studies will help us understand the underlying molecular mechanisms of shape determination and motility in *Spiroplasma*.

Bibliography

Bové, J.M., Renaudin, J., Saillard, C., Foissac, X., and Garnier, M. (2003). *Spiroplasma citri*, a plant pathogenic *Mollicute*: Relationships with its two hosts, the plant and the leafhopper vector. Annu. Rev. Phytopathol. *41*, 483–500.

Clark, T.B., Whitcomb, R.F., Tully, J.G., Mouches, C., Saillard, C., Bove, J.M., Wroblewshi, H., Carle, P., Rose, D.L., Henegar, R.B., et al. (1985). *Spiroplasma melliferum*, a new species from the honeybee (*Apis mellifera*). Int. J. Syst. Bacteriol. *35*, 296–308.

Cohen-Krausz, S., Cabahug, P.C., and Trachtenberg, S. (2011). The monomeric, tetrameric, and fibrillar organization of Fib: the dynamic building block of the bacterial linear motor of *Spiroplasma melliferum* BC3. J. Mol. Biol. *410*, 194–213.

Duret, S., Danet, J.-L., Garnier, M., and Renaudin, J. (1999). Gene disruption through homologous recombination in *Spiroplasma citri*: An scm1-disrupted motility mutant is pathogenic. J. Bacteriol. *181*, 7449–7456.

Ku, C., Lo, W.-S., and Kuo, C.-H. (2014). Molecular evolution of the actin-like MreB protein gene family in wall-less bacteria. Biochem. Biophys. Res. Commun. *446*, 927–932.

Kürner, J., Frangakis, A.S., and Baumeister, W. (2005). Cryo–electron tomography reveals the cytoskeletal structure of *Spiroplasma melliferum*. Science *307*, 436–439.

Lartigue, C., Duret, S., Garnier, M., and Renaudin, J. (2002). New plasmid vectors for specific gene targeting in *Spiroplasma citri*. Plasmid *48*, 149–159.

Markham, P.G., Townsend, R., Bar-Joseph, M., Daniels, M.J., Plaskitt, A., and Meddins, B.M. (1974). *Spiroplasmas* are the causal agents of citrus little-leaf disease. Ann. Appl. Biol. *78*, 49–57.

Nunan, L.M., Pantoja, C.R., Salazar, M., Aranguren, F., and Lightner, D. V (2004). Characterization and molecular methods for detection of a novel *Spiroplasma* pathogenic to Penaeus vannamei. Dis. Aquat. Organ. *62*, 255–264.

Razin, S. (1978). The mycoplasmas. Microbiol Rev *42*, 414—470.

Rottem, S., and Razin, S. (1973). Membrane lipids of *Mycoplasma hominis*. J. Bacteriol. *113*, 565–571.

Saillard, C., Vignault, J.C., Bove, J.M., Raie, A., Tully, J.G., Williamson, D.L., Fos, A., Garnier, M., Gadeau, A., Carle, P., et al. (1987). *Spiroplasma phoeniceum* sp. nov., a new plant-pathogenic species from Syria. Int. J. Syst. Bacteriol. *37*, 106–115.

Shaevitz, J.W., Lee, J.Y., and Fletcher, D. A. (2005). *Spiroplasma* swim by a processive change in body helicity. Cell *122*, 941–945.

Awards, funding and fellowships

- Awarded Raman Charpak Fellowship 2017 to perform collaborative research in the laboratory of Dr. Laure Béven, at INRA, Bordeaux, France (April 2018 - September 2018).
- Awarded EMBO Short-Term Fellowship to perform collaborative research in the laboratory of Dr. Martin Pilhofer at ETH, Zurich, Switzerland (November 2017-January 2018).
- Awarded Infosys foundation Travel Award for attending an International Conference on "Bacterial Morphogenesis, Survival and Virulence: Regulation in 4D" held between 29 Nov - 1 Dec 2016 at IISER, Thiruvananthapuram.
- Received funding from Biophysical Society, Rockville, Maryland towards attending "The 1st Cryo-Electron Microscopy and 3D Image Processing of Macromolecular assemblies and Cellular Tomography (CEM3DIP)" held at IISER Thiruvananthapuram between 2 July - 13 July 2016.

List of publications

- Baranwal, J., Lhospice, S., Kanade, M., Chakraborty, S, Gade, P. R., **Harne, S.**, Herrou, J., Mignot, T., and Gayathri, P. (2019). Allosteric regulation of a prokaryotic small Ras-like GTPase contributes to cell polarity oscillations in bacterial motility. PLoS Biol. 17(9): e3000459

- Adhav, A. S*., **Harne, S***., Bhidey, A. J., Giri, A. P., Gayathri, P., and Joshi, R. S. (2018). Mechanistic insights into enzymatic catalysis by trehalase from the insect gut endosymbiont *Enterobacter cloacae*. The FEBS Journal. *286*: 1700-1716 (* equal first authors)
- Palani, S., Chew, T. G., Ramanujam, S., Kamnev, A., **Harne, S.**, Chapa-y-Lazo, B., Hogg, R., Sevugan, M., Mishra, M., Gayathri, P., Balasubramanian, M. K. (2017). Motor activity dependent and independent functions of myosin II contribute to actomyosin ring assembly and contraction in *Schizosaccharomyces pombe*. Current Biology. *27 (5)*:751-757
- Gayathri, P., and **Harne, S.** (2017). Structure and dynamics of actin-like cytomotive filaments in plasmid segregation. Prokaryotic cytoskeletons, Subcellular Biochemistry Series. Springer International Publishing AG. Vol 84: 299-321 (Ed: Jan Löwe and Linda A Amos)

Potential future manuscripts

- **Harne, S.,** Beven, L[#]., and Gayathri, P[#]. *Spiroplasma* as a model organism to understand bacterial physiology: Opportunities and challenges ([#]Corresponding authors)

  In this review, we summarize the information on cellular machinery in *Spiroplasma* that facilitates genetic manipulation, based on available sequenced genomes. We then discuss the tools available for genetic modification and characterization of *Spiroplasma* to make it a model organism to study important physiological processes such as shape determination, motility, cell division etc. in cell wall-deficient bacteria.

- **Harne, S.** and Gayathri, P. Heterologous expression and purification of *Spiroplasma* Fibril filaments using *E. coli* as the model system.

  Multiple protocols standardized for expression and purification of Fibril filaments will be described in this manuscript. (Work described in Chapter 2 of the thesis)

- **Harne, S.,** Hatti, K., and Gayathri, P. Structure of alcohol dehydrogenase from *Pichia pastoris*.

Here we summarize structures of nickel and cobalt ion bound structures of alcohol dehydrogenase purified from *P. pastoris* in 3 different space groups and insights obtained from these structures.

- **Harne, S***., Bapat, M.*, Kashyap, R., and Gayathri, P. A fluorescence tag based screen for identification of polymerization interfaces of a novel cytoskeletal protein. (* equal contribution).

  This manuscript summarizes a method for identification of polymerization interface of proteins that polymerize in a nucleotide-independent manner. This study will prove useful in advancing structural studies on challenging proteins such as bactofilin, crescentin and intermediate filaments.

- **Harne, S.,** Duret, S., Szwedziak, P., Pilhofer, M[#]., Beven, L[#]., and Gayathri, P[#]. A comparative study of wildtype versus non-helical, non-motile mutant strains of *Spiroplasma citri.* ([#]Corresponding authors)

  In this manuscript, we summarize our results obtained through comparative genomic, proteomic and electron cryotomography studies of wildtype versus mutant cells of *Spiroplasma* that provide hints towards novel functions of MreB protein. (Work described in Chapter 4 of the thesis).

Curriculum vitae

Name : **Shrikant Ramdas Harne**
E-mail id : **hshrikant30@gmail.com/harne.shrikant@yahoo.com**
Contact No. : **+91 787 506 0094**

Education & Research experience

Qualification	University / Institute	Duration
Post-Doctoral fellow	Indian Institute of Science Education and Research, (IISER) Pune, Maharashtra, India	August 2019 – present
Ph.D. (Structural biology)	Indian Institute of Science Education and Research, (IISER) Pune, Maharashtra, India	August 2013 – July 2019
M.Sc. Microbiology	Department of Microbiology, University of Pune, Maharashtra, India	2010 – 2012
M.B.A. Marketing and Health care	Department of Management Science, University of Pune, (PUMBA) Maharashtra, India	2008 – 2010
B.Sc. Microbiology	Fergusson college, Pune, Maharashtra, India	2005 – 2008

August 2013- present: PhD thesis (Supervisor Dr. Gayathri Pananghat, IISER, Pune, India)

Topic: *Biochemical, structural and genetic studies on the cytoskeletal proteins Fibril and MreBs from Spiroplasma*

Abstract: *Spiroplasma* is a group of cell wall-deficient, helical bacteria with characteristic kinking motility. In absence of any external appendages, motility in *Spiroplasma* is achieved by dynamic changes in helicity. These changes are brought about by filaments formed of Fibril, a cytoskeletal protein of novel fold, and MreBs, the bacterial actins. The repeating unit in Fibril forms very stable, nucleotide-independent, constitutive filaments which makes it difficult to isolate. The presence of five MreB homologues in *Spiroplasma*, an organism lacking cell wall, makes it an interesting system to explore novel functions of MreBs. In the absence of convenient molecular genetic tools, limited information is available on shape determinaion and motility in *Spiroplasmas*.

We have purified Fibril filaments using heterologous expression systems for the first time. We have also found that the shorter constructs of Fibril can form polymers. In the absence of structural information on Fibril, we have prepared multiple short constructs and performed their characterization to identify its domain boundaries. To understand shape determination and motility in *Spiroplasma*, we performed comparative studies between the wildtype and a naturally occurring mutant (non-helical, non-motile) strain. Our studies reveal that the mutant cells are deficient in one of the MreBs. Our studies reveal the novel role of MreB and suggest that MreBs may have non-overlapping functions in *Spiroplasma*. Thus, to decipher the role of MreBs and Fibril, we have initiated the process of obtaining *Spiroplasma* mutants with each of the *mreBs* and *fibril* genes disrupted. Results from these experiments will help us delineate the functional significance of each MreB and Fibril.

Skills: Cloning and purification of recombinant proteins using heterologous systems, purification of macromolecular complexes, electron microscopy, Electron cryo tomography, protein characterization techniques such as circular dichroism and thermal shift assay, light microscopy etc.

Research articles

Baranwal, J., Lhospice, S., Kanade, M., Chakraborty, S, Gade, P. R., **Harne, S.**, Herrou, J., Mignot, T., and Gayathri, P. (2019). Allosteric regulation of a prokaryotic small Ras-like GTPase contributes to cell polarity oscillations in bacterial motility. PLoS Biol. 17(9): e3000459

Adhav, A. S*., **Harne, S***., Bhidey, A. J., Giri, A. P., Gayathri, P., and Joshi, R. S. (2018). Mechanistic insights into enzymatic catalysis by trehalase from the insect gut endosymbiont *Enterobacter cloacae*. The FEBS Journal. 286: 1700-1716. (* equal first authors)

P Gayathri and **Harne S** (2017) Structure and dynamics of Actin-like cytomotive filaments in plasmid segregation. Prokaryotic cytoskeletons, Subcellular Biochemistry Series. Springer International Publishing AG. Vol 84: 299-321 (Ed: Jan Löwe and Linda A Amos)

Palani S, Chew T G, Ramanujam S, Kamnev A, **Harne S**, Chapa-y-Lazo B, Hogg R, Sevugan M, Mishra M, Gayathri P, Balasubramanian M K (2017) Motor activity dependent and independent functions of myosin II contribute to actomyosin ring assembly and contraction in *Schizosaccharomyces pombe*. Current Biology. 27 (5):751-757

Sant D G, Gujarathi T R, **Harne S R**, Ghosh S, Kitture R, Kale S, Chopade B A, Pardesi K R (2013) *Adiantum philippense* L. frond assisted rapid green synthesis of gold and silver nanoparticles. Journal of Nanoparticles 2013:1-9

Harne S, Sharma A K, Dhaygude M, Joglekar S, Kodam K, Hudlikar M (2012) Novel route for rapid biosynthesis of copper nanoparticles using aqueous extract of *Calotropis procera* L. latex and their cytotoxicity on tumour cells. Colloids Surf B Biointerfaces. 95: 284-288

Patents

Indian Product patent on the invention titled "Novel antibacterial culture medium composition". Inventors: **Shrikant R. Harne** and S. L. Laware. (Patent grant no 260402)

Indian Product patent invention titled "Novel formulation to inhibit bacterial blight disease of pomegranate [*Punica granatum* (LINN.)]". Inventors: **Shrikant R. Harne** and S. L. Laware. (Awaiting grant, application no 2070/MUM/2010)

Academic Achievements

Awarded **Raman Charpak fellowship** 2017 to perform collaborative research in the lab of Dr. Laure Béven, at INRA, Bordeaux, France (March 2018-September 2018)

Awarded **EMBO Short-Term fellowship** to perform collaborative research in the lab of Dr. Martin Pilhofer at ETH, Zurich, Switzerland (November 2017-January 2018)

Awarded **Infosys foundation Travel Award** for attending an International Conference on "Bacterial Morphogenesis, Survival and Virulence: Regulation in 4D" held between 29 Nov-1 Dec 2016 at Thiruvananthapuram, India.

Received **funding from Biophysical Society, Rockville, Maryland** towards attending "The 1st Cryo-Electron Microscopy and 3D Image Processing of Macromolecular assemblies and Cellular Tomography (CEM3DIP)" held at Thiruvananthapuram, India between 2 and 13 July 2016.

Secured 10th rank in M.Sc. Microbiology exam (April 2012) at Pune University, India

Qualified CSIR-UGC National Eligibility Test (NET) June 2012 for Lecturership

Qualified Graduate Aptitude Test in Engineering (GATE) 2012

Conferences, courses and workshops attended

EMBO workshop Frontiers in cytoskeleton research (29 Oct - 1 Nov 2017; IISER Pune, India)

GIAN practical course: The 1st Cryo Electron Microscopy and 3D Image Processing of Macromolecular Assemblies and Cellular Tomography (CEM3DIP) (2 July – 13 July 2016; IISER Thiruvananthapuram, India)

EMBO conference series Bacterial morphogenesis, survival and virulence: regulation in 4D (27 Nov – 1 Dec 2016; Thiruvananthapuram, India)

Centre of Excellence workshop on Biomolecular interactions (25 - 28 Nov 2015; NCBS, Bangalore, India)

First electron microscopy workshop on Electron tomography (16 - 19 Oct 2016; NCBS, Bangalore, India)

International Symposium-cum-workshop Frontiers of structural biology New advances in X-ray diffraction and cryo-electron microscopy (15 – 17 Dec 2014; INSA, New Delhi, India)

Other research experience

January 2013- July 2013: Research project assistant in the lab of Dr. Ashok Giri at CSIR-NCL, Pune, India.
Topic: Discovery of therapeutic molecule.

August 2011 – April 2012: Research project as a part of M.Sc. thesis under the guidance of Prof. (Dr.) B. A. Chopade, Department of Microbiology, University of Pune, India
Topic: Synthesis and characterization of gold and silver nanoparticles using *Zingiber officinale* extracts.

July 2007 – April 2010: Research projects under the guidance of Dr. S. L. Laware, Fergusson college, Pune, India
Topics: Novel antibacterial culture medium composition and
Novel formulation to inhibit bacterial blight disease of pomegranate [*Punica granatum*]

Membership of professional bodies

- Life member of Electron Microscope Society of India (EMSI), New Delhi, India (LM-957)

References

1) Dr. Gayathri Pananghat
Assistant Professor,
Indian Institute of Science Education and Research (IISER),
Dr. Homi Bhabha road, Pashan, Pune-411 008, India

E-mail id: gayathri@iiserpune.ac.in

2) Dr. Saikrishnan Kayarat
Associate Professor,
Indian Institute of Science Education and Research (IISER),
Dr. Homi Bhabha road, Pashan, Pune-411 008, India
E-mail id: saikrishnan@iiserpune.ac.in

3) Dr. Janesh Kumar
Scientist D,
National Centre for Cell Science (NCCS),
Pune, India
E-mail id: janesh@nccs.res.in

4) Dr. Richa Rikhy
Associate Professor,
Indian Institute of Science Education and Research (IISER),
Dr. Homi Bhabha road, Pashan, Pune-411 008, India
E-mail id: richa@iiserpune.ac.in

Abstract

Spiroplasmas are cell wall-deficient, helical bacteria with characteristic kinking motility. In absence of any external appendages, motility in *Spiroplasma* is achieved by dynamic changes in helicity. These changes are brought about by filaments formed by Fibril, a cytoskeletal protein of novel fold, and MreBs, the bacterial actins. The repeating unit in Fibril forms very stable, nucleotide-independent, constitutive filaments. The presence of five MreB homologues in *Spiroplasma*, an organism lacking cell wall, makes it an interesting system to explore novel functions of MreBs. In the absence of convenient molecular genetic tools, limited information is available on shape determinaion and motility in spiroplasmas. My work was aimed at understanding the molecular mechanism of shape determination and motility by Fibril and MreBs, through a combination of biochemical, structural and genetic studies.

To perform structural characterization, Fibril filaments were heterologously expressed in *E. coli* and successfully purified, as observed using electron microscopy. Domain-wise dissection of Fibril based on sequence alignments and secondary structure prediction provided insights for designing non-polymerizing soluble constructs and a polymerizing construct of minimal length. To understand shape determination and motility in *Spiroplasma*, comparative studies between the wildtype and a naturally occurring mutant (non-helical, non-motile) strain were performed. These studies revealed that the mutant cells are deficient in one of the MreBs and the MreBs have non-overlapping functions in *Spiroplasma*. The comparative analysis and complementation studies provided the first evidence for a new role of MreB in rod-to-helical shape transition. To decipher the role of MreBs and Fibril, the process of obtaining *Spiroplasma* mutants with each of the *mreB* and *fibril* gene disrupted has been initiated. Results from these experiments will help us delineate the functional significance of each MreB and Fibril.

Key words: *Spiroplasma*, cytoskeleton, Fibril, MreB, gene disruption, cell shape determination.

List of abbreviations

Abbreviation	Full form
µF	Microfarad
µg/mL	Micrograms per milliliter
µL	Microlitre
µm	Micrometer
ºC	Degree Celsius
2X	Two times concentrated
3D	Three-dimensional
Å	Angstrom
AOX	Alcohol oxidase
ATCC	American Type Culture Collection
ATP	Adenosine triphosphate
BLAST	Basic local alignment search tool
Bp	Base pairs
BSE	backscattered electrons
C. crescentus	*Caulobacter crescentus*
CRISPR	Clustered regularly interspaced short palindromic repeats
Cryo-CLEM	Cryo-correlative light and electron microsocpy
CryoEM	Electron cryomicroscopy
C-terminal	Carboxy-terminal
CV	Column volume
D-Ala	D-alanine
DGC	Density gradient centrifugation
DMMA	2,3-dimethylmaleic anhydride
DNA	Deoxyribonucleic acid
DNase	Deoxyribonuclease
DSMZ	Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (German Collection of Microorganisms and Cell Cultures GmbH)
e^-	Electron

E. coli	*Escherichia coli*
e/A^2	Electrons per angstrom square
ECT	Electron cryotomography
EDTA	Ethylenediaminetetraacetic acid
EF-Tu	Elongation factor-Tu
EGFP	Enhanced green fluorescent protein
EM	Electron microscopy
EMBO	European molecular biology organization
ETH	Eidgenössische Technische Hochschule
FBS	Fetal bovine serum
FE-SEM	Field Emission-Scanning Electron Microscope
FP	Fluorescent protein
Fps	Frames per second
GFP	Green fluorescent protein
gm^r	Genetic determinant encoding gentamicin resistance
g/cm^3	Grams per cubic centimeter
GTP	Guanosine triphosphate
H. pylori	*Helicobacter pylori*
H. sapiens	*Homo sapiens*
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HIC	Hydrophobic interaction chromatography
His_6	Hexahistidine
HS buffer	HEPES and sucrose containing buffer
INRA	Institut national de la recherche agronomique
KDa	Kilodalton
keV	Kiloelectron volts
LB	Luria Bertani
LSLB broth	Low salt- Luria Bertani broth
LSS	N-lauroyl sarcosine sodium salt
M	Mole
MC	Methyl-cellulose

mL	Millilitre
Mm	Millimolar
MMC	Multimodal cation-exchanger
MTAN	5'-methyladenosine/S-adenosylhomocysteine nucleosidase
NCBI	National Centre for Biotechnology Information
NEB	New England Biolabs
Ni-NTA	Nickel-Nitrilotriacetic acid complex
nm	Nanometer
NMR	Nuclear magnetic resonance
N-terminal	Amino-terminal
OD	Optical density at 600 nm
ORF	Open reading frame
OriC	Origin of replication
P. pastoris	*Pichia pastoris*
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase chain reaction
PDB	Protein data bank
PG	Peptidoglycan
PROMALS3D	PROfile multiple alignment with predicted local structures and 3D constraints
PS	*spiralin* gene promoter sequence
RFC	Restriction-free cloning
RoI	Region of interest
RPM	Revolutions per minute
SCARP	*Spiroplasma citri* adhesion-related proteins
SDC	sodium deoxycholate
SDS	Sodium dodecyl sulphate
SE	secondary electrons
SEM	Scanning electron microscope
SPA	Single particle analysis
Spiroplasma^{RecA}	*Spiroplasma* cell with a functional RecA protein
TCS	Two-component system

TE	Tris and EDTA
TEK	Tris, EDTA and KCl containing buffer
TEM	Transmission Electron Microscope
tetM	Tetracycline resistance protein M
TF	*fibril* gene terminator sequence
TLC	Thin layer chromatography
tnpR	Transposon's gene coding for Resolvase protein
Tris	Tris(hydroxymethyl)aminomethane
Tuf	Elongation factor-Tu
UV	Ultraviolet
V	Volt
v/v	Volume/Volume
Van-Alexa568	Vancomycin-Alexa568 conjugate
w/v	Weight/Volume
YFP-MreB	Yellow fluorescent protein-MreB conjugate
YPD agar	Yeast extract, peptone, dextrose agar
Δ (Delta)	Deletion
Ω	Ohm

Chapter 1

Introduction

1.1) Cytoskeleton

Cytoskeleton is a network of protein filaments in the cell involved in diverse cellular functions ranging from cell shape determination and chromosome segregation to controlling cargo transport (reviewed in Wagstaff and Löwe, 2018). The three major cytoskeletal elements of the eukaryotic cell comprise actin, tubulin and intermediate filaments. While actin and tubulin form nucleotide-dependent [Adenosine triphosphate (ATP) and Guanosine triphosphate (GTP) respectively] polymers that disassemble upon nucleotide hydrolysis, intermediate filaments form constitutive structures without involving any nucleotide (Pollard, 2003). Molecular motors such as kinesin and dynein interact with actin and tubulin filaments to bring about intra-cellular cargo transport (Vale and Milligan, 2000).

While an intricate network of cytoskeleton and motor proteins were known to play a part in organization of eukaryotic cells, prokaryotes were thought to be devoid of cytoskeleton. However, since early 1990s proteins that constitute prokaryotic cytoskeleton have been

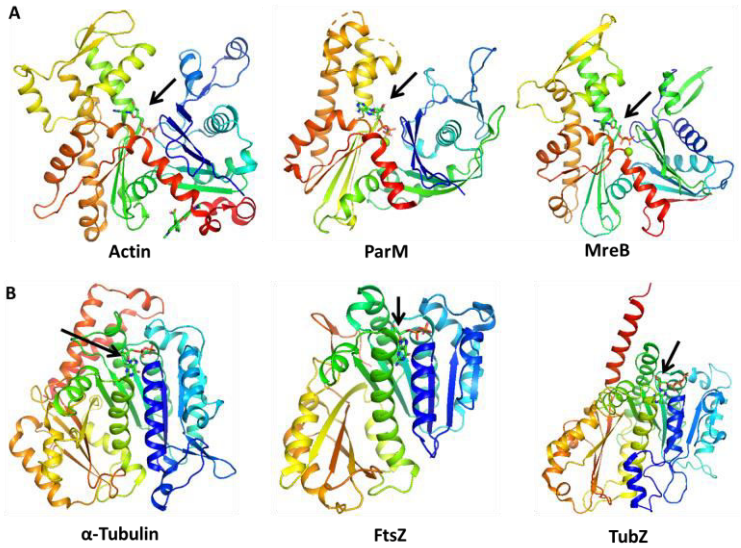

Figure 1.1 Crystal structures of eukaryotic cytoskeletal proteins actin, tubulin and their bacterial homologs. A. Structures of actin (PDB ID 1J6Z) and its bacterial homologs ParM (PDB ID 4A61) and MreB (PDB ID 4CZL). B. Structures of eukaryotic α-tubulin (PDB ID 5SYF) and its bacterial homologs FtsZ (PDB ID 3VO8) and TubZ (PDB ID 3M89). All the structures are shown as cartoon representation and colored blue to red from amino to carboxy terminus. The black arrows point to bound nucleotide. The figures were prepared using PyMOL software (https://pymol.org/2/).

discovered (Table 1.1). These proteins form ATP/GTP dependent polymers and exhibit structural homology with eukaryotic cytoskeletal elements (Bork et al., 1992; Van den Ent et al., 2001; Löwe and Amos, 1998) (Figures 1.1 & 1.2). Each cell possesses at least one homolog each of an actin-like and tubulin-like protein in addition to other cytoskeletal proteins that may not have a eukaryotic homolog.

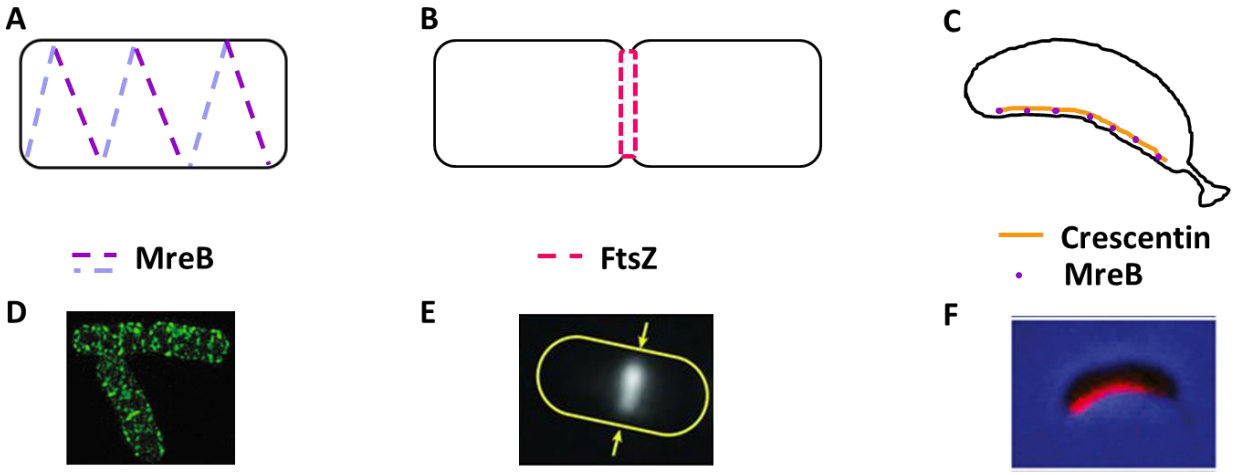

Figure 1.2 Intra-cellular localization of bacterial cytoskeletal proteins. A - C) Pictorial representation for localization of MreB (A), FtsZ (B) and Crescentin (C). The purple lines in image (A) represent MreB patches on the inner membrane of top surface of the cell whereas, the blue lines represent projections of MreB patches on the inner membrane at the bottom of the cell. D) Structured illumination microscopy image of *E. coli* cell showing MreB localization (image adapted from Colavin et al., 2018). E) Maximum intensity projection of an *E. coli* cell showing FtsZ localization at mid-cell region (image adapted from Yang et al., 2017). F) Fluorescence microscopy image of a *C. crescentus* cell showing localization of GFP labeled Crescentin polymers (image adapted from Charbon et al., 2009).

One of the key differences between eukaryotic and prokaryotic cytoskeletal systems is the absence of molecular motors; no proteins equivalent to molecular motors (like dynein, kinesin) have been reported from prokaryotes yet. Table 1.1 summarizes some of the prokaryotic cytoskeletal proteins, their properties and functions. With the advances in microscopy, genetics and biochemical studies, the role of bacterial cytoskeleton have been found to be essential for several cellular processes such as motility, intra-cellular transport, cell division, and DNA segregation similar to their eukaryotic counterparts. Shape determination, one of the major functions performed by bacterial cytoskeleton, is discussed in the following section.

Table 1.1 Prokaryotic cytoskeletal proteins and their function.

Protein(s)	Organism	Function	Polymerization/ oligomerization	Reference
Actin-like proteins				
MreB	Most of the rod-shaped bacteria	Cell wall synthesis machinery positioning	ATP-dependent polymerization/de-polymerization	(Bork et al., 1992; van den Ent et al., 2001)
ParM	Many bacteria	Plasmid segregation		(Møller-Jensen et al., 2003)
FtsA	Most bacterial species	Cell division		(van den Ent and Löwe, 2000)
MamK	Bacteria possessing magnetosomes	Organization of magnetosomes within the cell		(Komeili et al., 2006)
Tubulin-like proteins				
FtsZ	Most bacterial species	Cell division	GTP-dependent polymerization/de-polymerization	(Bi and Lutkenhaus, 1991; RayChaudhuri and Park, 1992)
TubZ	Many bacteria	Plasmid segregation		(Larsen et al., 2007)
BtubA/B	*Prosthecobacter* spp.	Unknown		(Pilhofer et al., 2011)
Proteins that do not have a homolog in eukaryotes				
ParA	Many bacteria	DNA segregation	Walker A-type cytoskeletal ATPase that form dimers	(Davey and Funnell, 1997)
MinD	Most bacterial species	Cell division machinery positioning		(Hayashi et al., 2001)
Fibril	*Spiroplasma* spp	Shape determination and motility	Nucleotide-independent, constitutive filaments	(Townsend et al., 1980a)
Bactofilin	*M. xanthus*, *C. crescentus*	Scaffold-like function		(Kühn et al., 2010)
Crescentin	*Caulobacter* spp	Shape determination		(Ausmees et al., 2003)
DivIVA	Few gram-positive bacteria	Polar growth		(Oliva et al., 2010)
CrvA	*Vibrio* spp	Helping in cell curvature		(Bartlett et al., 2017)
PopZ	Few gram-negative bacteria	Multiple functions at poles		(Ebersbach et al., 2008)

1.2) Shape determination in cell walled bacteria

Bacteria exist in many different shapes and arrangements such as cocci, rods, comma, spiral, and pleomorphic. Often the cell shape is a characteristic of the given species and is used as a taxonomic tool. The shape is determined and maintained by the cell wall, a rigid continuous structure made of peptidoglycan strands surrounding cell membrane. The cell wall also provides protection against turgor pressure (Shi et al., 2018). Many gene products are involved in synthesis of peptidoglycan, a polymer of sugar-crosslinked amino acids that forms cell wall. Cell wall components are synthesized in the cytosol, transferred outside the cell and inserted into existing cell wall. Peptidoglycan synthases play an important role in synthesis and positioning peptidoglycan strands. Bacteria have evolved different mechanisms of cell wall growth to ensure that the cell shape is maintained. All such mechanisms utilize cytoskeletal proteins to direct peptidoglycan insertion sites.

In spherical bacteria, the cytoskeletal protein FtsZ assembles as a ring at the mid-cell region and

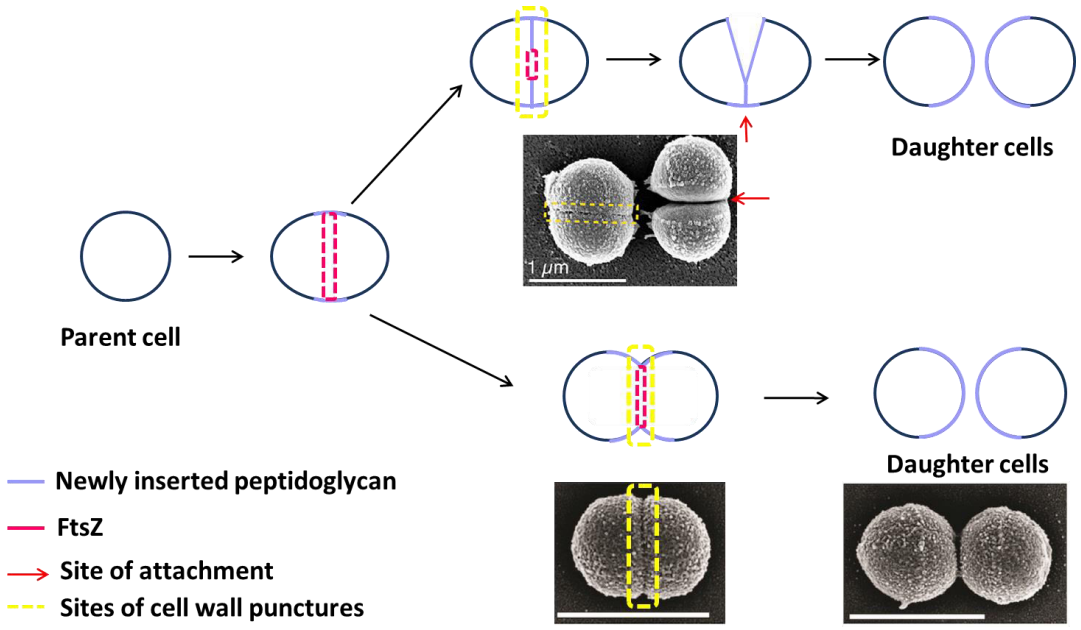

Figure 1.3 FtsZ-dependent incorporation of the cell wall and growth in spherical bacteria. Two models for growth and division in spherical bacteria have been shown as pictorial representation. In *S. aureus* during cell wall synthesis at the septum, the daughter cells start to separate abruptly from one end (top panel) as opposed to uniform cell constriction accompanying cell wall synthesis at septum in case of *Jeotgalicoccus* spp ATCC 8456 (bottom panel). Both the panels shown above are pictorial representations of cell wall synthesis and cell division while below are scanning electron microscopic images of cells (images adapted from Zhou et al., (2016). The red arrows in top panel point to corresponding positions where the daughter cells are attached to each other. Sites of cell wall puncture are shown with rectangles of discontinuous lines. Scale bars in the bottom panel represent 1 µm.

positions the peptidoglycan synthesis machinery to facilitate cell wall growth and cell division (Figure 1.3).

In rod-shaped bacteria, peptidoglycan synthases are positioned along the inner surface of cell membrane by the bacterial actin MreB when it is bound to the membrane with negative curvature (Carballido-López, 2006; Carballido-López et al., 2006; Shih and Rothfield, 2006; Ursell et al., 2014). *mreB* was one of the initial set of genes to be identified as shape determinants in *E. coli* and *B. subtilis* (Abhayawardhane and Stewart, 1995; Doi et al., 1988; Levin et al., 1992). The circumferentially moving MreB patches facilitate local synthesis and uniform insertion of newly synthesized peptidoglycan components into existing cell wall thereby maintaining rod shape (Figure 1.4) (Dominguez-Escobar et al., 2011; Garner et al., 2011; van Teeffelen et al., 2011; Ursell et al., 2014).

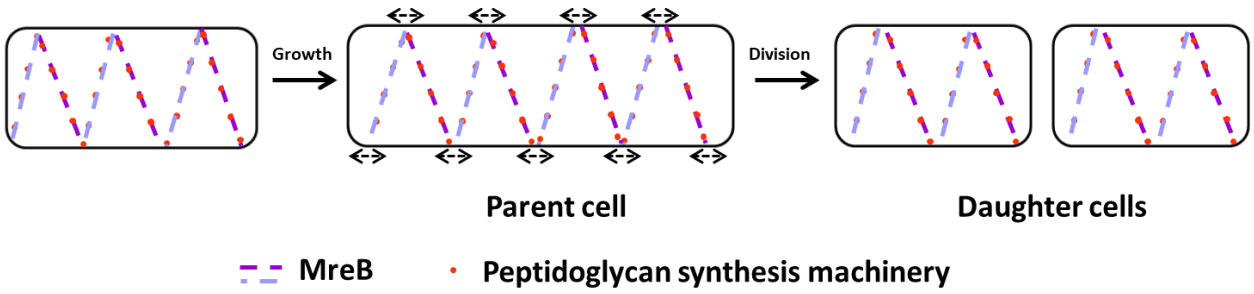

Figure 1.4 MreB-dependent growth mechanism in cell walled, rod-shaped bacteria. Rod-shaped bacteria such as *E. coli* exhibit uniform lateral growth. The peptidoglycan synthesis machinery (orange dots) is positioned along the cell membrane by short patches of MreB (purple and blue lines). The bi-directional arrows indicate uniform growth of the cell wall.

Thus, MreB acts as a scaffold for the cell elongation and shape determination machinery in most of the rod-shaped bacteria. Cells with *mreB* depletion exhibit spherical/lemon shaped morphology that eventually lyse (Bendezu and Boer, 2008; Figge et al., 2004).

Most of the cell walled, non-spherical bacteria have at least one copy of *mreB* gene or its homologs (Daniel and Errington, 2003; Jones et al., 2001). However, there are two groups that are exceptions to this; first are some of the Cyanobacteria and Chlamydia that are often spherical in spite of possessing *mreB* genes (Daniel and Errington, 2003). The other group of exceptions belongs to members of *Mycoplasma*, *Corynebacterium*, *Mycobacteria* and *Rhizobium* that are rod-shaped without the presence of an MreB (Daniel and Errington, 2003). Most of these bacteria exhibit polar growth by an MreB-independent pathway involving DivIVA-dependent

mechanism (Figure 1.5; Table 1.1) (members of Actinobacteria; Brown et al., 2011) or using polarity determinant proteins (Cameron et al., 2015).

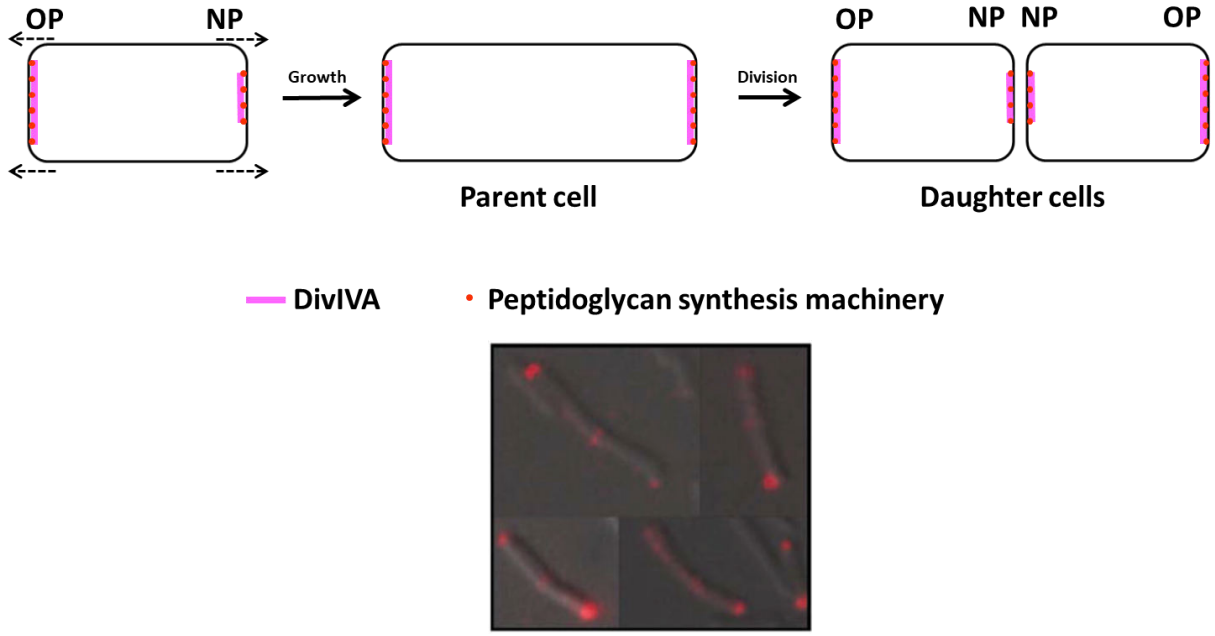

Figure 1.5 MreB-independent mechanism involving polar growth in cell walled bacteria. Bacteria such as *M. smegmatis* exhibit MreB-independent polar growth with the help of peptidoglycan synthesis machinery (orange dots) positioned by DivIVA (pink line), a protein localizing at the cell poles. OP and NP indicate old pole and new pole respectively. The discontinuous arrows indicate direction of cell growth. Bottom panel: Visualization of insertion of newly synthesized peptidoglycan in *M. smegmatis* cells by incubating the cells with vancomycin conjugated to Alexa568 dye (Van-Alexa568) followed by visualization using fluorescence microscopy. Since vancomycin binds to the peptidoglycan precursor at the terminal D-Ala-D-Ala, sites of new peptidoglycan insertions are identified by visualizing Van-Alexa568. Image adapted from Kang et al., (2008).

While sphere and rod are the predominant cell shapes in bacteria, a few bacterial species exhibit modification of the rod shape such as crescent, vibroid, spiral and helical among others. The curved shape of *Caulobacter crescentus* is attained with the help of a coiled-coil protein called Crescentin (Ausmees et al., 2003). The elastic assemblies of Crescentin are anchored to the cytoplasmic membrane along inner curvature of the cell (Figure 1.6) (Cabeen et al., 2009). The uniform growth of cells along the side walls, partly facilitated by turgor pressure, leads to extension of Crescentin filaments that are present in a tensed state. The tensed Crescentin filaments exert compression on the peptide bridges stretched due to turgor pressure, thereby preventing insertion of new peptidoglycan along the surface where Crescentin is present.

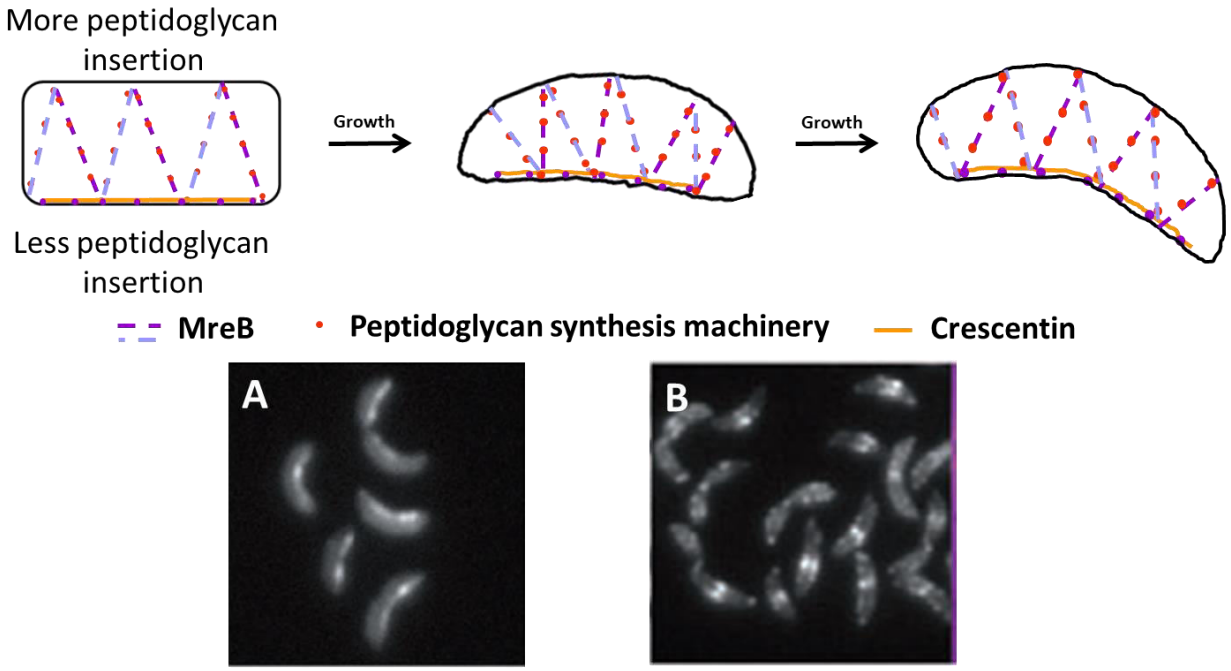

Figure 1.6 Modification of rod shape into crescent. *C. crescentus* attains crescent shape by asymmetric growth of the peptidoglycan. The bias in peptidoglycan is created by Crescentin polymers (orange line), that prevents peptidoglycan insertion in the surface where it is localized. The intra-cellular organization of Crescentin polymers is facilitated by MreB (purple dots). Bottom panel: A) Fluorescence microscopy image of *C. crescentus* cells expressing GFP-Crescentin to visualize localization of Crescentin. B) Fluorescence microscopy image of *C. crescentus* cells expressing YFP-MreB showing localization of MreB within cells. Images adapted from Dye et al., (2011).

As you move away from point of Crescentin binding the compressive effect decreases and insertion of new peptidoglycan components is facilitated. Thus, Crescentin creates a peptidoglycan insertion gradient along the short axis of cell and results into formation of crescent shaped cells (Woldemeskel and Goley, 2017). Crescentin can show its effect only when it is associated with the membrane in a polymerized form (Cabeen et al., 2009). The cellular organization of Crescentin is performed by MreB (Charbon et al., 2009). While Crescentin does not seem to be an essential protein for survival as its deletion leads to formation of rod-shaped cells (Ausmees et al., 2003), depletion of *mreB* leaves Crescentin detached from the cell wall thereby deviating from crescent shape of *C. crescentus* cells (Charbon et al., 2009; Figge et al., 2004). Thus both MreB and Crescentin seem to be necessary for crescent shape of *C. crescentus* cells by causing a gradient of new peptidoglycan insertion.

Another interesting case is that of *Helicobacter pylori* which attains helical shape by modulating cell wall growth using a multiprotein complex by local inflection of rates of PG synthesis (Blair

et al., 2018). The helical shape of *H. pylori* is an outcome of a much complicated mechanism which is not understood yet. According to literature, co-ordinated action of over 30 proteins results into the helical morphology of *H. pylori* (Krzyżek and Gościniak, 2018; Specht et al., 2011; Yang et al., 2019). The group of proteins includes endopeptidases, carboxypeptidase, peptidoglycan precursor synthases, bactofilin, CcmA (curved cell morphology), Ccrp (coiled-coil rich protein), ATP synthase and proteins of unknown function. Deletion of any of these proteins leads to morphological defects in *H. pylori* (Ghachi et al., 2011; Specht et al., 2011; Sycuro et al., 2010; Waidner et al., 2009; Yang et al., 2019). While some of these proteins are known to interact with peptidoglycan and influence its modification, the others perform regulatory function. Thus all these proteins are vital for helical shape in *H. pylori*. Interestingly MreB does not seem to participate in helical shape formation in *H. pylori* since *ΔmreB* cells are helical with the same diameter as wildtype cells (Waidner et al., 2009). However, absence of *mreB* does affect cell length, chromosome segregation and pathogenicity in *H. pylori* (Waidner et al., 2009).

Taken together, the studies on cell walled bacteria suggest that the cell shape is determined and maintained by ensuring regulated peptidoglycan insertion (Cabeen and Jacobs-wagner, 2005). Thus, it becomes interesting to investigate as to how cell wall deficient bacteria, classified under class *Mollicutes*, have a definite shape. The mechanism of shape determination in cell wall deficient bacteria is not understood yet. Among *Mollicutes* only *Spiroplasma* cells have a distinct helical shape and hence serve as an interesting system to understand shape determination in cell wall-less organisms.

My work therefore has been to understand shape determination mechanism in *Spiroplasma*. The next few sections introduce *Spiroplasma*, its pathogenicity, the proposed mechanisms for its shape determination and motility, and why it is a model organism to study these processes.

1.3) *Spiroplasma*: a bacterium lacking cell wall

Spiroplasma were discovered as Mycoplasma-like helical bacteria associated with citrus stubborn and corn stunt diseases (Davis et al., 1972; Lafleche and Bove, 1970). Biochemical characterization of these bacteria led to their classification as spiroplasmas (Saglio et al., 1973). *Spiroplasma* were proposed to be classified as cell wall-lacking bacteria under class *Mollicutes* due to their sensitivity to osmotic shock and presence of a single lipid bilayer (Razin et al., 1973). Initial investigations of *Spiroplasma* cells included electron microscopy studies to

visualize locomotory organelles and cell wall. These studies concluded the absence of cell wall and motility appendages (Cole et al., 1973). Genome analysis of spiroplasmas suggests that the genes coding for proteins involved in cell wall synthesis are absent. The lack of cell wall renders spiroplasmas sensitive to changes in environmental conditions such as pH, osmotic pressure or media composition during laboratory growth conditions.

Over the years, many species of *Spiroplasma* have been isolated and studied. A look at the 33 genomes available in the database suggests that the genome lengths of *Spiroplasma* vary between 0.891 (*S. monobiae* MQ-1) and 2.3 (*S. poulsonii* MSRO) million base pairs. Genome analysis studies proposed that spiroplasmas may have evolved from gram-positive bacteria (*Clostridia*) with low G+C content in the DNA, by genome reduction through evolution (Razin et al., 1998; Weisburg et al., 1989). The genome reduction is thought to be associated with their parasitic life cycle leaving behind only the genes essential for survival (Pereyre et al., 2009). Though members of genus *Spiroplasma* lack the canonical genes coding for chemotaxis and motility, they are motile and exhibit chemotactic behavior (Daniels et al., 1980; Liu et al., 2017; Markham et al., 1974; Terahara et al., 2017; Trachtenberg and Gilad, 2001). While motility and chemotaxis is needed for obtaining nutrients and in response to external stimuli, they do not seem to be essential for pathogenicity (Duret et al., 1999).

1.4) *Spiroplasma* as a pathogen

The discovery of *Spiroplasma* as a pathogen for diseases of citrus plants and crabs has led to it being recognized as a notable pathogen. In fact, most of the *Spiroplasma* species discovered until now have been reported to be pathogenic to plants and insect vectors. While the role of *Spiroplasma* species as a causative agent of transmissible spongiform encephalopathies (TSE; Bastian et al., 2007) is reported, recent reports suggest their ability to infect immunocompromised humans (Etienne et al., 2018; Mueller et al., 2015). Table 1.2 summarizes *Spiroplasma* species involved in various infections.

The ability of spiroplasmas to grow in two different hosts, plant and insect, has made them a model organism to study relationship between *Mollicutes* and their hosts (Duret et al., 2005). In addition, the genus *Spiroplasma* has attracted attention from cell biologists due to their motility despite the lack of external appendages and a distinct helical shape in absence of a cell wall (Cole et al., 1973; Saglio et al., 1973).

Table 1.2 *Spiroplasma* species associated with various diseases and infections.

***Spiroplasma* species**	**Diseases (host)**	**Reference**
Spiroplasma citri	Citrus stubborn (citrus plants), (*Catharanthus roseus*), Brittle root disease (horseradish)	(Duret et al., 2003; Fletcher et al., 1981)
Spiroplasma melliferum	Infects honey bees	(Clark et al., 1985)
Spiroplasma eriocheris	Tremor disease (Chinese mitten crab) and infects most of the artificially breeding crustaceans	(Liu et al., 2017; Wang et al., 2011)
Spiroplasma apis	May disease (Honey bee)	(Mouches et al., 1983)
Spiroplasma floricola	Pathogenic to larvae of *Galleria mellonella*	(Eskafi et al., 1987)
Spiroplasma turonicum	Systemic infection in humans	(Aquilino et al., 2015; Etienne et al., 2018)
Spiroplasma spp.	Corn stunt	(Davis and Worley, 1973)
Spiroplasma poulsonii	Selective male killing (*Drosophila melanogaster*)	(Williamson et al., 1999)
Spiroplasma spp.	Infection and mortality (*Penaeus vannamei*)	(Nunan et al., 2004)
Spiroplasma spp.	Cataract (Suckling mouse)	(Tully et al., 1976)
Spiroplasma spp.	Systemic infection (*Procambarus clarkii*)	(Wang et al., 2005)
Spiroplasma spp.	(*Macrobrachium rosenbergii*)	(Liang et al., 2011)
Spiroplasma kunkelii	Corn stunt disease (Maize plants), Insect-borne plant pathogen	(Ozbek et al., 2003; Whitcomb et al., 1986)
Spiroplasma phoeniceum	Plant pathogen (*Catharanthus roseus*)	(Saillard et al., 1987)

1.5) Shape determination in *Spiroplasma*

Plasma membrane is the only barrier between intra-cellular components and environment in *Mollicutes*. It plays an important role in pathogenesis and interaction with the host (Beven and Wroblewski, 1997). Since the plasma membrane is composed of lipids and proteins that are not as rigid as peptidoglycan, deciphering the mechanism of how do such cell wall-deficient bacteria maintain specific cell shapes becomes interesting. Due to the natural absence of cell wall synthesis machinery in *Mollicutes*, they likely possess the minimal components required to divide or sculpt a membrane-bound cell. It is thought that in absence of cell wall the cytoskeleton of these bacteria must be responsible for maintaining the cell shape (Cabeen and Jacobs-Wagner, 2005). However, such bacteria cannot take broad range of shapes other than spherical and

oblong. Only exception to this is the genus *Spiroplasma* that forms distinct helical cells (Figure 1.7 A).

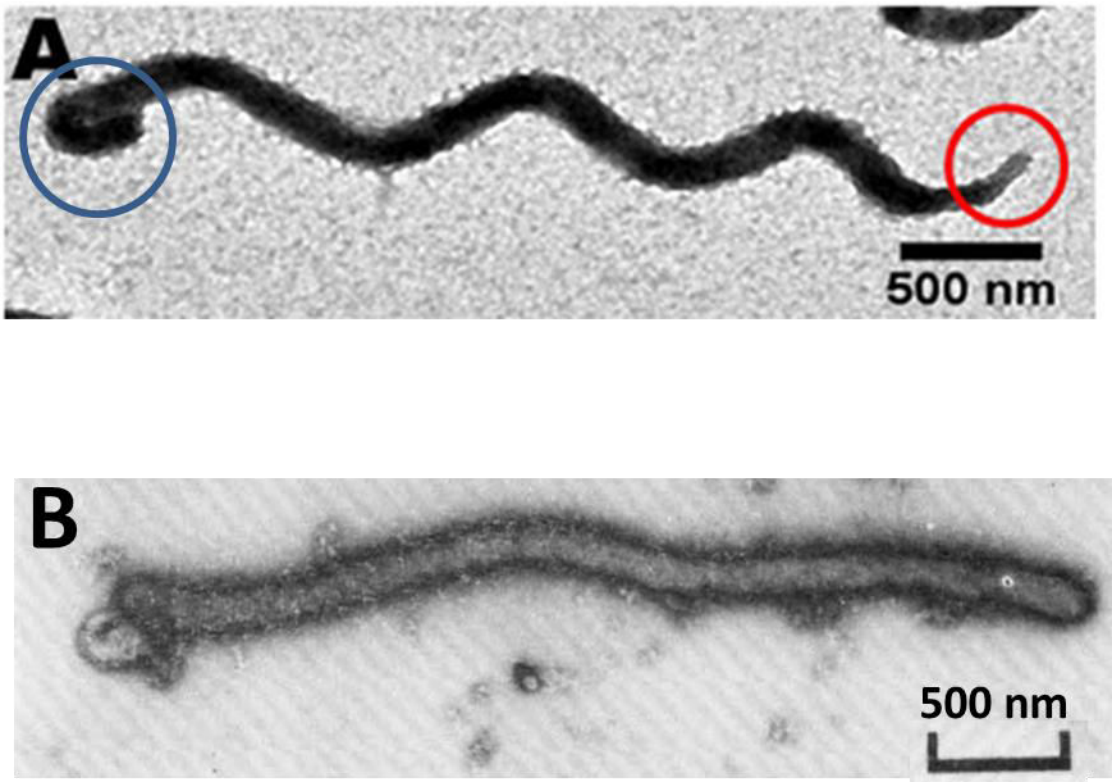

Figure 1.7 Geometry of *Spiroplasma* cells. A and B) Negatively stained electron microscopy images of *S. citri* GII-3 (wildtype; helical, motile) and *S. citri* ASP-I (naturally occurring mutant; non-helical, non-motile) cells respectively. Note the tapered/leading pole in red circle and the circular/lagging pole shown in blue circle in *S. citri* GII-3 cell. The images are adopted from Liu et al., 2017 (*S. citri* GII-3) and Townsend et al., 1977 (*S. citri* ASP-I).

Ever since the discovery of spiroplasmas the cell biologists were puzzled by their spiral shape in the absence of a cell wall. Based on the information available on spiroplasmas and other closely related bacteria, following hypotheses were proposed.

1) Studies by Freeman et al., (1976) on lipid composition and metabolism of *Spiroplasma citri* suggested that alteration in the growth medium composition led to changes in helical morphology. While the presence of palmitic acid as media supplement resulted in cell with helical morphology, stimasterol and ergosterol resulted in loss of motility, helicity and other morphological defects during whole or part of the life cycle. Thus, it was proposed that the helical cell shape must be a result of rigid and tightly packed 16-carbon lipids such as palmitic acid. However, this hypothesis was later weakened when *Mycoplasma hominis*, an organism

with lipid composition similar to *Spiroplasma* was not found to be helical (Razin, 1978; Rottem and Greenberg, 1975; Rottem and Razin, 1973).

2) Studies on shape determination in a helical mutant of *Bacillus licheniformis* provided the basis for an alternate model. Initial studies to understand helical shape determination were carried out on *B. subtilis* B1S, a strain with multiple mutations in its genome including suppressed cell division. Germination and growth of *B. subtilis* B1S spores on potato agar resulted into helical phenotype when the poles were attached to the surface. An explanation in support of such a phenotype was the synthesis and insertion of peptidoglycan as a helix (Mendelson, 1976). The possibility that helical shape might be a result of attaching cell poles to a surface that restricted rotation of poles to release stress, could not be ruled out. Thus, another model was proposed for helical shape formation in bacteria. Based on studies using a Triton-X (non-ionic detergent) resistant, helical mutant of *B. licheniformis* (TR61), Tilby (1977) proposed a model. According to the model, peptidoglycan in gram-positive bacteria is synthesized in the form of right- or left-handed fibres. The cell assumes a helical shape when fibres with one type of handedness dominate or if the fibres were compressed or in a state of tension. Thus, fibres constituting cell wall of bacteria were thought to be responsible for shape determination.

Williamson (1974) observed the release of protein fibres from *Spiroplasma* upon treatment with detergent. The discovery of *Spiroplasma* Fibril prompted Razin (1978) to extend the hypothesis for role of these fibres in shape determination to *Spiroplasma* helicity. According to the hypothesis, Fibril are present helically along the membrane at the cytoplasmic side. When the Fibril are in a tensed or compressed state, they can pull the poles towards each other and the cell becomes helical (Razin, 1978). This is still a hypothesis in the field that needs verification by genetic studies. Further research in the past three decades has led to proposition of following proteins to be involved in *Spiroplasma* helicity.

i) Fibril

One of the prominent features seen while visualizing *Spiroplasma* under electron microscope was the presence of a layer along the cytoplasmic side of cell membrane (Cole et al., 1973). Treatment of *Spiroplasma* cells with 0.5 % sodium deoxycholate (anionic detergent) resulted in destruction of cell membrane and release of Fibril (Williamson, 1974). The role of these intra-cellular Fibril was unknown. Razin (1978) proposed the role of Fibril in conferring helical shape

to *Spiroplasma* cells. Later studies found Fibril protein to be unique to members of *Spiroplasma* genus. Biochemical characterization of *Spiroplasma* Fibril involved its treatment with colchicine (an alkaloid that prevents microtubule polymerization) and dialysis against ATP to identify if Fibril exhibit de-polymerization dynamics like actin or tubulin (Townsend et al., 1980b). Since Fibril remained unaffected after these treatments, it was proposed that they polymerize in a manner different from actin or tubulin, without involvement of ATP or GTP (Townsend et al., 1980b). Thus, Fibril polymers were considered to be constitutive cytoskeletal filaments that do not exhibit polymerization- depolymerization dynamics.

Genetic studies to create *fibril* deletion mutants of *Spiroplasma* have not been reported in literature. This could be due to non-viability of *fibril* deficient mutant, in case it is an essential gene. However, the role of Fibril in *Spiroplasma* requires further investigation.

ii) MreBs

The role of actin in eukaryotic cell shape determination is known for long (Cooper and Pollard, 1982). Also, MreB, the bacterial actin-like protein, is known to be responsible for determination of cell shape in bacteria. The presence of actin-like protein in *Spiroplasma* cells has been proposed long ago (Bove et al., 1989). However the MreB filaments in *Spiroplasma* were first observed in 2005 by Kürner et al., (2005). Though the identification of these filaments were difficult based on tomography data, inter-filament spacing, sequence analysis and western blotting proved useful in ascertaining them to be MreB. The authors propose that MreB plays a role in conferring elongated shape to the *Spiroplasma* cell. However, there is no experimental evidence in support of this.

iii) A 39 kDa protein of unknown identity

Apart from the intra-cellular filaments thought to be responsible for shape determination in *Spiroplasma*, clues were obtained from *S. citri* ASP-I, a non-helical, non-motile strain of *S. citri* (Figure 1.7 B). *S. citri* ASP-I was isolated by Townsend et al., (1977). DNA homology, serology tests and other studies confirmed ASP-I to be a strain of *S. citri*. Studies to identify defects associated with loss of helicity and motility involved partial purification of membrane proteins from *S. citri* ASP-I and comparison with *S. citri* SP-A (helical, motile) cells. The study revealed that Fibril protein, at that time considered to be responsible for helicity, was present in both cell

types whereas a 39 kDa protein was missing from ASP-I cells. Thus, the 39 kDa protein was proposed to be one of the components of machinery involved in shape maintenance in *Spiroplasma*. Till date, the identity of this 39 kDa protein remains enigmatic.

Apart from the role of Fibril, MreBs and the 39 kDa protein in shape determination, they have also been proposed to be involved in characteristic kinking motility of *Spiroplasma* (Kürner et al., 2005; Townsend et al., 1977).

1.6) Components involved in *Spiroplasma* motility

Members of genus *Spiroplasma* are motile without external motility appendages (Cole et al., 1973). A recent study involving visualization of *Spiroplasma* motility using high resolution video microscopy revealed that the *Spiroplasma* cells produce kinks for movement (Shaevitz et al., 2005).

Often bacterial cell shape seems to facilitate motility. For instance, spiral cells swim efficiently through viscous fluids than rod-shaped cells (Ferrero and Lee, 1988). Many bacterial species have evolved specialized motility organelles such as flagella, pili, and focal adhesion complexes (Jarrell and McBride, 2008). Pili and flagella are protein filament extensions outside the cell powered by the multi-protein motor complexes embedded in the cell envelope. Contrary to this, focal adhesion complexes are multi-protein complexes on cell surfaces that help in gliding motility of *Myxococcus* species.

Members of genus *Spiroplasma* are some other examples of bacteria that are motile without external appendages. They are motile by using cytoskeleton; cytoskeletal protein filaments of Fibril and MreB have been proposed to constitute linear motor responsible for *Spiroplasma* motility (Gilad et al., 2003; Trachtenberg et al., 2003a). The source of energy driving the motor has been identified in most other bacteria whereas, the energy source of linear motor of *Spiroplasma* remains to be identified. Thus, spiroplasmas become an interesting system to explore novel motility mechanisms.

Fibril and MreB proteins have been speculated to be involved in *Spiroplasma* motility since long (Williamson et al., 1991). However, their organization within the cell is unknown. Two models have been proposed for organization of MreB and Fibril filaments within the cell. Kürner et al., (2005) observed a set of MreB filaments sandwiched between two sets of Fibril filaments just

beneath the cell membrane by electron cryotomography (ECT) of *Spiroplasma* cells (Figure 1.8 A, C and D). Whereas, Trachtenberg and co-workers (2008) observed only a single 70 nm wide ribbon connected to the cell membrane via a set of MreB filaments (Figure 1.8 B and E). Thus, Trachtenberg et al., (2008) proposed that MreB filaments are present between Fibril filaments and cell membrane.

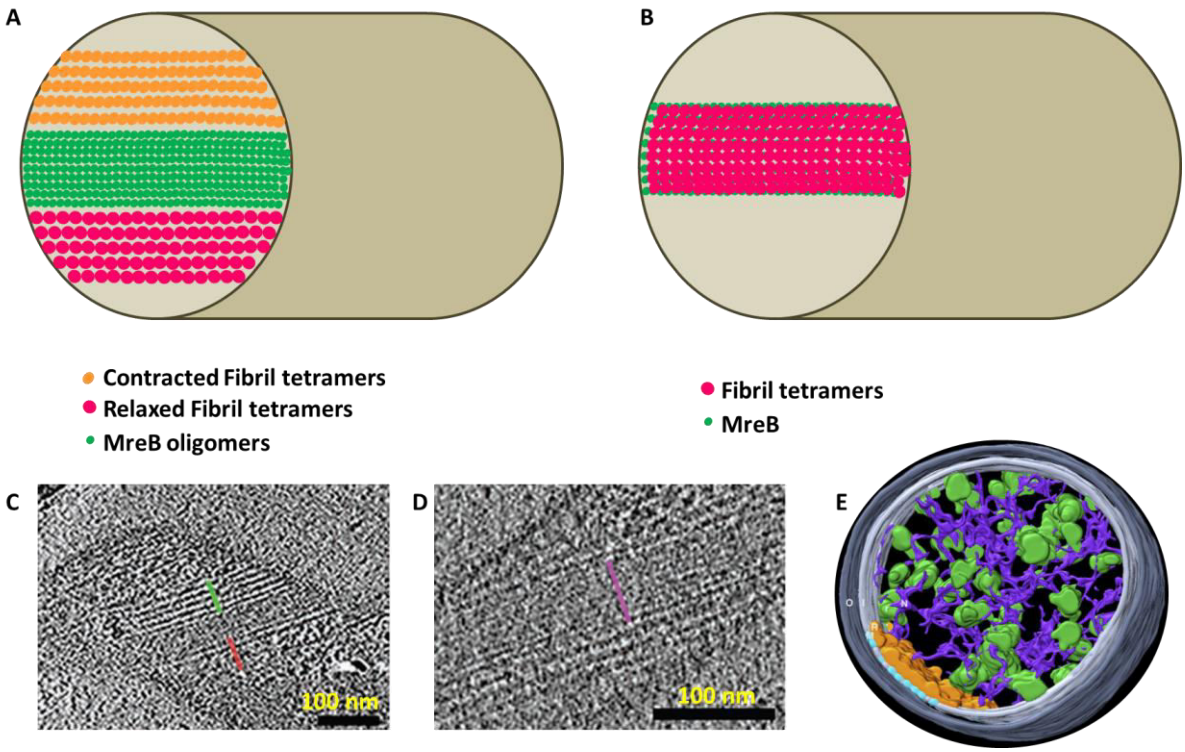

Figure 1.8 Molecular organization of *Spiroplasma* cytoskeleton. A) Pictorial representation of a longitudinal section of *Spiroplasma* cell showing organization of cytoskeletal protein filaments of Fibril and MreB as proposed by Kürner et al., 2005. Three sets of filament bundles are present just beneath the cell membrane. Nine filaments of MreB (lines of green circles) are flanked by two sets (lines of pink and orange circles) of Fibril bundles each consisting of 5 filaments. Note that one set of Fibril filaments (orange color) are contracted as compared to the other (pink filaments). The contraction is brought about by conformational changes in Fibril tetramers (orange and pink circles). C and D) Sections from tomograms of *S. melliferum* BC-3 cells. Two sets of Fibril filaments with 11 nm inter-filament spacing (C) sandwich the MreB filament set having inter-filament spacing of 4 nm (D). B) According to the model proposed by Trachtenberg et al., 2008, Fibril filaments may be linked to the cell membrane via MreBs. Fibril (shown in pink circles) may indicate relaxed or contracted state since the original authors do not mention about conformation in this respect. E) Top view of a surface rendered tomogram of cross-sectioned *S. melliferum* BC-3 cell showing intra-cellular organization. Dark and light gray colors represent outer and inner membrane leaflets. The cytoskeletal (Fibril) ribbon is shown in orange while membrane-associated densities (MreB) are shown in light blue (between orange and light grey). The green densities correspond to ribosomes and magenta is likely to be DNA. Images modified from Kürner et al., (2005) (C and D) and Trachtenberg et al., (2008) (E).

Since artefacts associated with chemical fixation of samples are well known (Kellenberger et al., 1992), the possibility of loss/re-arrangement of cytoskeletal filaments due to fixation cannot be ruled out. Thus, cytoskeletal organization of Fibril and MreB filaments remains unclear. Nevertheless, both the groups supported the hypothesis that these protein filaments (Fibril and MreB) constitute a linear motor and play a role in *Spiroplasma* motility.

The linear motor is composed of internal cytoskeletal protein filaments made of Fibril and MreB proteins (Kürner et al., 2005). Fibril filaments have been proposed to undergo co-ordinated conformational changes which are reflected as a kink in the cell body. The movement of a pair of kinks from one pole of the cell body to the other helps the cell move forward. The movement of *Spiroplasma* cells is not random, though proteins related to two component system are absent in their genome (Liu et al., 2017). However, the molecular mechanism bringing about conformational changes in Fibril and conveying them to the cell membrane remains unclear.

In addition to cytoskeleton, the helical shape may also be necessary for *Spiroplasma* motility since the non-helical mutant of *Spiroplasma* is also non-motile (Figure 1.7 B) (Townsend et al., 1977). The observation of *scm1* mutants that are helical but non-motile suggest that only cell helicity is not sufficient for *Spiroplasma* motility (Jacob et al., 1997). This also suggests that the machinery responsible for helicity and motility are likely to be independent of each other.

Apart from Fibril and MreB, some more proteins have been proposed to be involved in *Spiroplasma* motility. The basis for inclusion of these proteins in *Spiroplasma* motility machinery is discussed below.

i) *scm1* gene

In order to identify proteins involved in *Spiroplasma* motility, Jacob et al., 1997 performed transposon-based mutagenesis studies on *S. citri* to identify its motility mutants. The study involved transformation of *S. citri* GII-3 cells with Tn*4001* transpon (containing gentamycin resistance gene) to facilitate its random insertion in *S. citri* genomic DNA. The motility mutants were identified based on formation of non-diffuse, non-satellited colonies on agar (Figure 1.9).

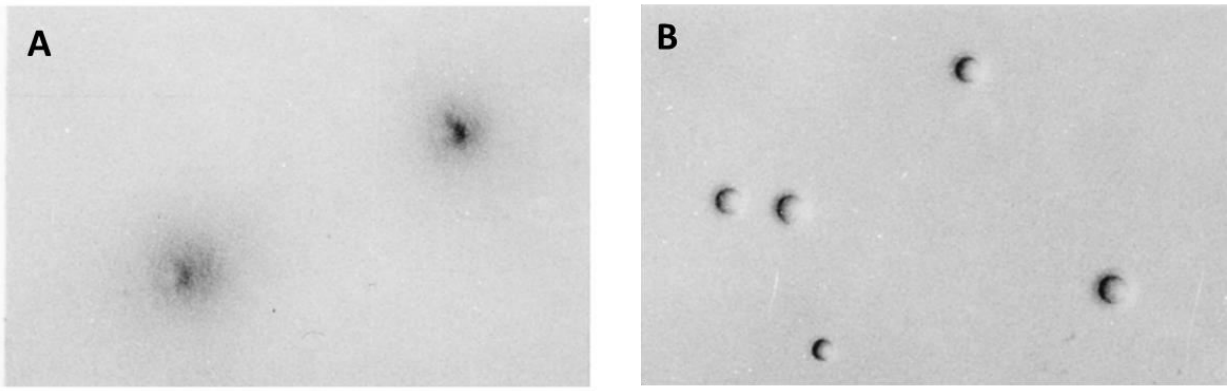

Figure 1.9 Translational motility and colony morphology of *Spiroplasma* cells. Colony morphology of A) *S. citri* GII-3 cells and B) *S. citri* G540 (artificially prepared helical, non-motile mutant) cells on SP4 medium containing 1 % agar. Note that the *S. citri* GII-3 cells form diffuse colonies (an indication of translational motility) as opposed to fried egg shaped colonies formed by *S. citri* G540 cells that are non-motile. Images of *Spiroplasma* colonies (15 X magnification) have been taken from Jacob et al., (1997).

Microscopic observation of motility mutants revealed that, unlike *S. citri* ASP-I, the cells were helical in shape. The mutant (G540) exhibited no rotation about helical axis, low frequency of flexing and twisting, as against wildtype cells. Analysis of the genomic DNA of the mutant cell revealed that the Tn*4001* was inserted into an ORF coding for a protein of 409 amino acids. Since the protein negatively affected ***S***. ***c****itri* **m**otility, it was named as SCM 1. The role of *scm1* in *Spiroplasma* motility was confirmed upon restoring the motility of mutant by transforming it with vector containing wildtype *scm1* gene. Transformation of *S. citri* ASP-I cells with vector containing *scm1* gene did not reinstate motility in them (Jacob et al., 1997). This suggests that *S. citri* ASP-I cells are defective in gene other than *scm1*. A database analysis showed that the *scm1* gene is present in all the spiroplasmas known at that time. However, the SCM1 protein does not share homology with proteins in other organisms. Though the protein SCM1 remains to be characterized further, the above study suggests its role in inducing rotation of *Spiroplasma* cells along the long axis.

ii) Dumbell shaped structure and 16 proteins

While the aforesaid proteins may be involved in *Spiroplasma* motility, what induces conformational changes in Fibril is unknown. The existence of distinct morphologies (tapered and round ends) at the cell poles was reported long ago (Ammar et al., 2004). However, its significance in *Spiroplasma* cells is not clear.

A recent publication by Liu et al., (2017) reports the observation of introduction of the kinks at the tapered end (leading pole). While the kinks move towards the other (round) cell pole, the cell propels in the direction of the tip (Liu et al., 2017; Shaevitz et al., 2005). The attachment of tip to cytoskeletal filament ribbon that extended to the other end, provoked the authors to propose the role of tip in motility (Liu et al., 2017). The absence of proteins related to two-component system (TCS) in *Spiroplasma* and swimming of cells with the tip in front, prompted the authors to propose that the tip may have a chemosensory role.

Spiroplasma Fibril are present in the 'Triton X-100 insoluble' fraction upon treatment of the cell lysate with Triton X-100 (Cohen-Krausz et al., 2011; Townsend et al., 1980b; Trachtenberg et al., 2003b). In order to check the proteins associated with Fibril, Miyata and co-workers extracted Triton X-100 insoluble proteins from *S. eriocheiris*. They obtained 11 other proteins apart from Fibril and four MreBs in the Triton X-100 insoluble fraction (Liu et al., 2017). The 11 proteins include ATP synthase subunit, lipoproteins, FtsH etc. The presence of ATP synthase subunit in Triton X-100 insoluble fraction prompted the authors to propose that it may be associated with Fibril as the source of energy for the linear motor of *Spiroplasma*.

Though exact composition of the motility machinery of *Spiroplasma* is unknown, it is proposed through multiple studies that Fibril is the main component of cytoskeletal ribbon. The positioning of Fibril filaments close to the plasma membrane directly or through other proteins may be necessary for motility (Kürner et al., 2005; Trachtenberg et al., 2008). There is ambiguity about localization of MreB filaments within the cell - whether it is an adapter between Fibril and the cell membrane.

1.7) Model for *Spiroplasma* motility

According to the current model, a cytoskeletal ribbon in *Spiroplasma* is responsible for its motility. The ribbon follows shortest path along the cytoplasmic face of the inner membrane between the two cell poles (Trachtenberg and Gilad, 2001). The ribbon is flat, monolayered structure which may be attached to the cell membrane directly or through MreB (Trachtenberg and Gilad, 2001; Trachtenberg et al., 2008). The ribbon is predominantly composed of Fibril filaments (Trachtenberg et al., 2003a).

The Fibril filaments are homo-polymers of a 55 - 59 kDa monomeric protein called Fibril (Fib; Townsend et al., 1980b; Trachtenberg and Gilad, 2001). A monomeric Fibril molecule is probably a bi-domained structure (Cohen-Krausz et al., 2011). *In vitro* studies suggest that the diameter of Fibril filaments is about 4.5 nm and an axial repeat of 8.2 nm (Figure 1.10 A). The individual filaments tend to pair and form filaments with the width of about 10 nm. The dimensions so formed are sufficient to accommodate a tetramer of 59 kDa protein, Fibril (Trachtenberg et al., 2003b). The repeating unit in Fibril filaments is the tetramer that is linked to the surrounding tetramers (Figure 1.10 A & B). Many such subunits assemble laterally and axially to form a ribbon of flat, parallel filaments organized in a helical fashion, called Fibril (Cohen-Krausz et al., 2011). The number for Fibril filaments constituting the ribbon *in vivo* may be 5 (Kürner et al., 2005) or 7 (Cohen-Krausz et al., 2011) (Figure 1.8 A).

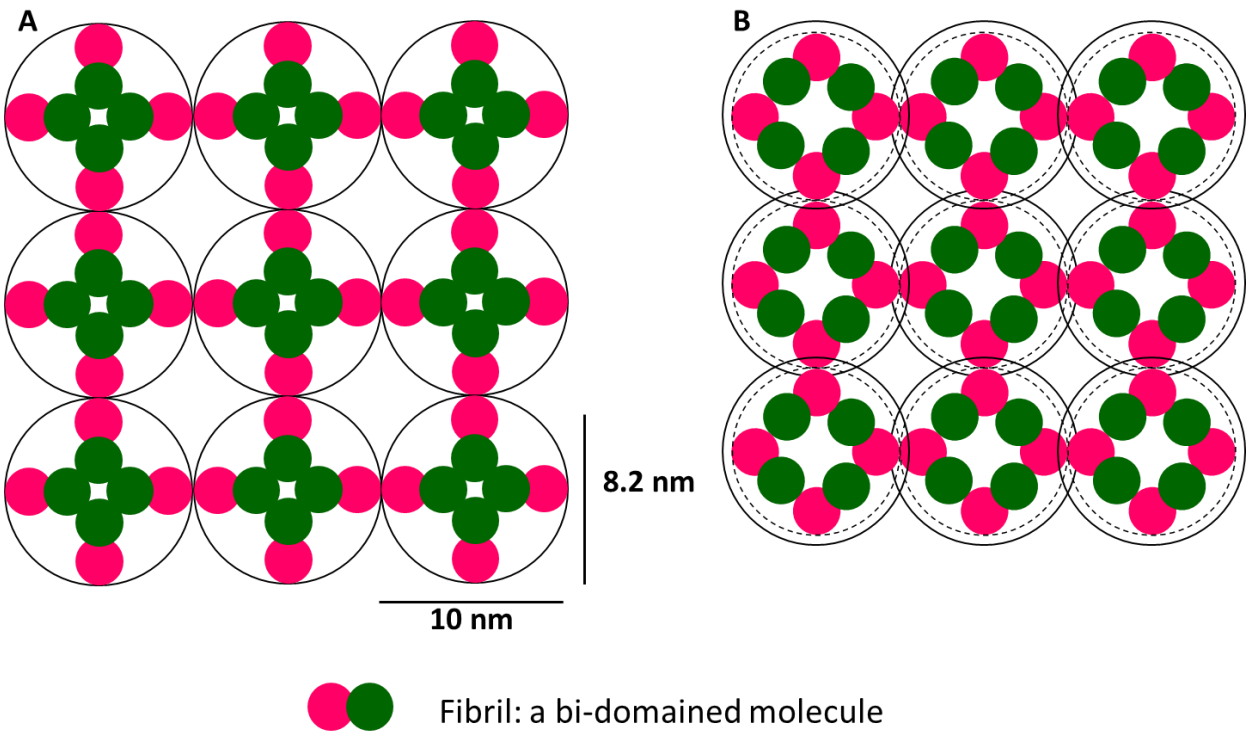

Figure 1.10 Organization of Fibril and its role in *Spiroplasma* motility. A and B) Fibril filaments are homo-polymers of monomeric protein Fibril. A Fibril monomer is probably a bi-domained structure (green and pink circles are two domains of a monomer). The repeating unit in Fibril filaments is the Fibril tetramer. The tetramers are arranged axially and laterally with a repeat of about 8.2 nm and 10 nm to form Fibril filaments. The Fibril tetramers can take up two stable conformations: A) cross-like and B) circular. A transition of all tetramers in a filament from cross-like (continuous circles) to circular conformation (discontinuous circles) results into reduction in net length and vice-versa (B). Co-ordinated conformational changes in Fibril filaments lead to change in handedness of the cell body and contribute to *Spiroplasma* motility.

The Fibril tetramers can exist in two stable conformations: cross-like and near-circular (Figure 1.10 A & B) (Cohen-Krausz et al., 2011; Trachtenberg and Gilad, 2001). The conformational changes can occur by rotational and diagonal shifts of the monomers through transition states and be facilitated by phosphorylation of Fibril (Cohen-Krausz et al., 2011). When many such tetramers are arranged axially and laterally, a change in conformation of all the tetramers from cross-like to near-circular state would result into net reduction in length of Fibril filaments and vice versa (Figure 1.10 A & B) (Cohen-Krausz et al., 2011; Trachtenberg and Gilad, 2001). Two set of Fibril ribbons were observed in a *Spiroplasma* cell (Figure 1.8 A) (Kürner et al., 2005). Co-ordinated length changes between the two ribbons such that one of them contracts keeping the other in a relaxed state would confer handedness to the cell. The alternate reversal of contraction and relaxation between the two ribbons would alternate handedness of the cell between left and right. The position in cell where handedness is switched would reflect as a kink in the cell body (Kürner et al., 2005). The kinks are introduced at the tapered end of the cell body and travels to the other end of cell (Shaevitz et al., 2005; Terahara et al., 2017). When the first kink reaches near the other end of cell body, another kink is introduced at the front (tapered end) (Shaevitz et al., 2005). The kink propagation through the cell body gives it the thrust that pushes the surrounding fluid backward and moves the cell forward (Figure 1.11) (Shaevitz et al., 2005). Thus, the motility of *Spiroplasma* becomes an interesting model system to explore functions of cytoskeleton in cell movement.

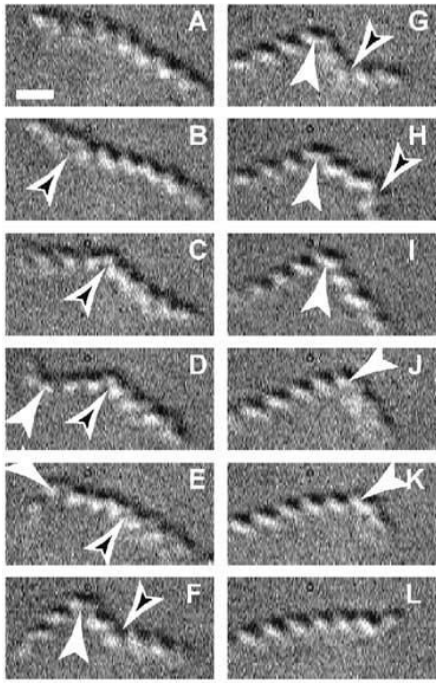

Figure 1.11 Kinking motility of *Spiroplasma*. High resolution video frames of swimming *S. melliferum* cells. Kinks are introduced at only one pole of the cell and travel towards the other pole. When the first kink reaches about half the cell length, second kink (white arrow) is introduced at the same cell pole. Through the frames B-K the kinks are seen propagating along the cell body. The handedness of the cell body on either side of the kink is opposite. Scale bar in frame A represents 1 µm. Image is adapted from Shaevitz et al., 2005.

1.8) *Spiroplasma* as a model organism

The characteristics of *Spiroplasma* such as distinct helical shape without cell wall and kinking motility using cytoskeleton make it an interesting system to understand shape determination and motility mechanism in bacteria. Research in the past two decades has uncovered some more proteins in addition to Fibril and multiple MreBs to be contributing to shape and motility of *Spiroplasma*. However, the roles of these proteins are yet to be ascertained with an experimental proof.

Fibril, a protein unique to members of genus *Spiroplasma* (Townsend and Archer, 1983), does not have a homolog. The proposed roles of Fibril in *Spiroplasma* shape determination and motility remains poorly understood. As discussed earlier, MreB is responsible for shape determination of cell walled bacteria by positioning cell wall synthesis machinery. It also plays role in chromosome segregation, gliding motility (Gitai et al., 2005; Lange et al., 2015; Madabhushi and Marians, 2009). Cell walled bacteria possess a maximum of 3 copies of *mreB* gene in the genome to carry out aforementioned functions. Interestingly, although *Spiroplasma* cells lack a cell wall, they have at least five copies of *mreB* gene (Carle et al., 2010; Ku et al., 2014). The presence of multiple *mreB* copies seem to have resulted from gene duplication events (Ku et al., 2014). The conservation of five copies of *mreB* despite the tendency to keep minimum

genes necessary for survival indicates that all of them must be functionally significant. In absence of cell wall, it becomes interesting to investigate if and how the MreBs are involved in shape determination of *Spiroplasma*.

Identification of role, localization and assigning function to proteins proposed to be involved in *Spiroplasma* motility and shape determination (enlisted above) would require verification by genetic studies. Although SCM1 protein is proposed to be involved in *Spiroplasma* motility with experimental evidence, its role and localization remains unknown. The current model for *Spiroplasma* motility has been proposed with Fibril and MreB proteins as the components of linear motor.

As mentioned earlier, *Spiroplasma* is the only *Mollicute* with a distinct helical shape and kinking motility. Members of genus *Spiroplasma* are also discrete from other bacteria in unique possession of *fibril* and more than 5 *mreB* genes (*Haloplasma contractile* is the other cell wall deficient bacteria to possess 5 *mreB* genes). The presence of Fibril and MreB filaments beneath the cell membrane is the circumstantial evidence for their role in morphology and motility of *Spiroplasma*. No experimental evidence to support the role of these proteins is available in literature. Experiments involving deletion/disruption of one of these genes at a time and characterization of mutants would verify the role of these proteins in *Spiroplasma* physiology. However, since spiroplasmas lack a functional RecA, they are not amenable to genetic modification using tools available for genetic studies in other organisms. The next section discusses in detail the development of tools for *Spiroplasma* genetics.

1.9) Tools available for *Spiroplasma* genetic studies

Gene deletion/truncation is often the method for ascertaining functional significance of a protein. In general, homologous recombination is often the method of choice for creating gene deletion mutants. The technique involves preparation of a DNA construct (clone) having an antibiotic resistance marker and a portion of the target gene to be deleted. The clone is transformed into the cells and grown for a few generations in the presence of antibiotic. During their growth, the cells repair mechanism incorporates the portion of the clone sequence into the genomic DNA at the homologous region, thus resulting in a truncated gene that produces non-functional protein (Christiansen and Griffith, 1986).

The recombination machinery, that facilitates homologous recombination required for gene disruption, consists of a protein complex consisting of a helicase and nucleotide-dependent nuclease as well as RecA protein that brings about recombination. Upon encountering the homologous region of DNA, RecA brings about strand exchange leading to recombination (Smith, 1988). Though there are multiple pathways for bringing about DNA recombination, all of them involve RecA as the central player (Amundsen and Smith, 2003). Thus, presence of RecA is crucial for a successful recombination event (Clark, 1973). Though RecA-independent recombination mechanisms are known, they are less efficient and occur at very low frequencies (Dutra et al., 2006).

Genetic studies on *Spiroplasma* are being carried out for about 3 decades now. The first induced gene mutant of *Spiroplasma* was obtained by irradiation of ultra-violet (UV) light on cells (Labarère and Barroso, 1989). Simultaneously, another group produced motility mutant of *S. melliferum* using nitrous acid (Cohen et al., 1989). Due to the difficulties associated with identification of site of mutation in the mutants obtained by above methods, transposon (Tn*4001*) was used to generate random mutations in *S. citri* genome. The use of transposon facilitated identification of site of mutation by southern blot hybridization using Tn*4001* specific probe P1 (Foissac et al., 1997b, 1997a). When generating mutants using transposons, in order for the mutation to be carried on over generations, it is necessary that the transposons are stably maintained. In mutants obtained by this method, the transposons inserted into genome can be lost in absence of selection pressure (Dybvig and Cassell, 1987). Although all the aforesaid methods produced mutants of *Spiroplasma*, these methods are not useful when targeting a specific gene.

As described above, specific gene targeting is possible through homologous recombination strategies. However, spiroplasmas were found to contain a truncated and probably a non-functional *recA* gene (Marais et al., 1996). The absence of a functional RecA protein in *Spiroplasma* genome reduces the chances of homologous recombination from occurring. Thus incompatibility of multiple *oriC*, intra-molecular recombination and low frequency homologous recombination were thought to be the possible causes of DNA recombination in *Spiroplasma* cells (Marais et al., 1996). Table 1.3 summarizes the advances in targeting specific genes in *Spiroplasma*.

Table 1.3 Advances in design of tools for genetic modifications in *Spiroplasma*.

Method/plasmid	Advantages/Drawbacks
UV/ nitrous acid treatment	Problems with identification of site of mutation.
Transposons eg. Tn*4001*	Difficulty in targeting specific genes, transposons are unstable.
Replicative viruses eg. SpV1-RF	Unstable in the cell.
Artificial plasmids with 2 kbp *oriC* region (*oriC*, *dnaA* gene and *dnaA* boxes) eg. pBOT1, pOT1, pOTPG, pOTPS, pSD4	Vector integration at chromosomal *oriC* rather than target site due to presence of full-length *dnaA* gene (next to *oriC*).
Artificial plasmids with 163 bp *oriC* region (*oriC* and *dnaA* boxes) Eg. pC2; pCL2	Presence of free plasmids in the cell (instead of integration into genomic DNA), requires more passaging for integration.
Aritificial plasmid containing promoter-less antibiotic resistance gene downstream of target gene fragment Eg. pC55	Growth of cells by expressing antibiotic resistance gene using gene sequence in the vector as promoter. Required passaging of cells in media supplemented with high antibiotic concentration (10-15 µg/mL) for integration to occur.
Artificial plasmids containing a gentamycin resistance gene under *spiralin* promoter and a promoter-less tetracycline resistance gene Eg. pGOT1 and pGOT18	The process is not 100% efficient since cells were found to grow in tetracycline containing medium without integration into genomic DNA. The tetracycline resistance gene was transcribed probably using regions from gene sequence present at its upstream. Subsequent gene deletion mutants cannot be created using the same antibiotic.
Artificial plasmids with promoter-less antibiotic resistance gene sandwiched between *resolvase* sites. Eg. pSD6 and pSD61. Vector-based expression of Resolvase protein removes antibiotic resistance gene and produces unmarked mutants	The method is not 100 % efficient since antibiotic resistant cells with free plasmids are obtained.

Gene manipulation in *Spiroplasma* was accidentally discovered through experiments involving expression of foreign proteins in *Spiroplasma* (Marais et al., 1993; Renaudin et al., 1995; Stamburski et al., 1991; Ye et al., 1994). Since spiroplasmas encompass viruses, a replicative form (RF) of SpV1 virus, called SpV1-RF (from *Spiroplasma citri* R8A2 strain) was used for

cloning foreign genes. The foreign genes were expressed upon transformation of recombinant viruses into *Spiroplasma* cells (Marais et al., 1993; Stamburski et al., 1991). However, a later study found that the recombinant viral DNA becomes unstable and is lost with successive passages (Marais et al., 1996). Since patches of viral DNA sequence are present in *Spiroplasma* genome, the recombinant viral DNA was thought to have undergone homologous recombination with host DNA. Thus, gene manipulation in *Spiroplasma* by recombination using SpVI-RF was thought to be a possibility.

Due to the instability of recombinant SpVI-RF DNA, artificial plasmids (pBOT1, pOT1, pOTPG, pOTPS) were created which contain *S. citri* chromosomal origin of replication (*oriC*) and tetracycline resistance marker gene, *tet*M (Renaudin et al., 1995). The *oriC* used in these plasmids is a 2 kbp region containing *dnaA* gene flanked by *dnaA* boxes. The 2 kbp region was used since it includes autonomously replicating sequences and may help in plasmid amplification and maintenance in cells. The plasmid pBOT1 and its derivatives could replicate in *E. coli* as well as *S. citri*. pBOT1 has intact *dnaA* sequence (an autonomously replicating sequence) and integrates into genomic DNA of *S. citri* in a few passages. Its derivatives with only the *dnaA* boxes (without the *dnaA* gene) remain extra-chromosomal, which implicated that the *dnaA* gene is not required while o*riC* and *dnaA* boxes are sufficient for replication of these plasmids. The plasmid pBOT1 was thought to integrate into genomic DNA by homologous recombination involving single crossing-over and is stably maintained irrespective of presence of selection pressure (Renaudin et al., 1995).

This discovery led to method development of disruption of target genes by cloning regions from target gene into pCJ32, a pBOT1 derived vector containing *oriC*, and transformation into *Spiroplasma* cells. During subsequent passages the vector got integrated into genomic DNA resulting in disruption of target gene. However, the use of this vector required extensive passaging and often led to integration of plasmid mostly at the *oriC* region of genomic DNA as compared to target gene (Duret et al., 1999). The recombination frequency at the *oriC* instead of target site was thought to be partly contributed by length of homologous sequence in the vector. Thus, new vectors (pC2; pCL2) were designed with reduced *oriC* sequence of 163 base pairs (consisting of two *dnaA* box regions and beginning of *dnaN* gene) and target gene fragment of 937 base pairs (bp). Reducing the length of *oriC* region in plasmid from 1986 base pairs (pSD4) to a minimum sequence that promoted plasmid replication (163 bp) increased recombination

frequency at the target site (Figure 1.12 A). However, extensive passaging may be necessary for achieving recombination at target site (Lartigue et al., 2002).

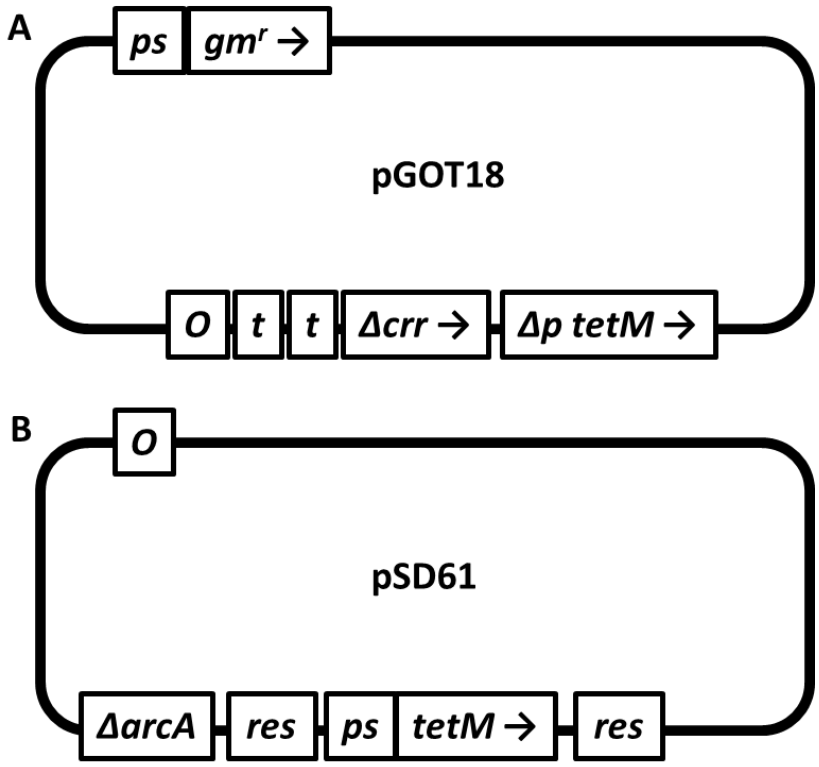

Figure 1.12 Tools available for genetic modification in *Spiroplasma*. A and B) Vector maps of pGOT18 and pSD61 vectors useful for target gene disruption in *Spiroplasma* cells. A) The pGOT18 vector has a promoter-less tetracycline resistance gene (Δp *tetM*) and gentamycin resistance marker (*gm*^{r}) under constitutively expressing *spiralin* promoter (*ps*). The portion of target gene to be disrupted (Δ*crr*) is preceded by *S. citri fibril* transcription terminator sequence and *oriC (O)*. B) The vector pSD61 has tetracycline resistance gene (*tetM*) under *spiralin* promoter (*ps*) flanked by *res* sites on either side. Partial sequence of the target gene (Δ*arcA*) is placed between one of the *res* sequence and *oriC* (*O*). Both the two vector maps have been modified from Duret et al., (2005).

The need for extensive passaging was overcome by designing another vector (pC55) with reduced *oriC* sequence and by placing the tetracycline resistance gene (*tetM*) devoid of promoter, downstream of target gene fragment. This strategy facilitated expression of *tetM* only from the target gene promoter upon recombination. When *Spiroplasma* transformants containing such a plasmid were grown in media supplemented with low concentrations of tetracycline (2 µg/mL), they were found to have free plasmids instead of integration at the target site. The growth of transformants at low tetracycline concentration was thought to be by expression of *tetM* gene using regions from target gene fragment as promoter. The transcription from the makeshift promoter facilitated low-level transcription, making it necessary for the cell to carry many copies

of the plasmid. Growing these cells with high copies of plasmid in media with higher tetracycline concentration (10-15 µg/mL) resulted into rapid integration of the plasmid at the target site (Lartigue et al., 2002). Another set of experiments were carried out to check the minimum homologous region required for disrupting the target gene. Use of a 220 bp fragment of a target gene in the vector did not yield target gene disruption mutants, thus indicating that longer (>220 bp) homologous sequence was necessary for successful recombination (André and Renaudin unpublished data; Duret et al., 2005).

Duret et al., (2005) designed new vectors (pGOT1 and pGOT18) with a two-step strategy to achieve target gene disruption (Figure 1.12 B). The vector has two different resistance marker genes: gentamycin and tetracyclin. Gentamycin resistance gene was placed under constitutively expressing *spiralin* promoter. Growing transformants in media supplemented with gentamycin allowed growth of only the cells carrying desired vector and facilitated increase in number of transformants as well as the copy number of plasmid. Since the efficiency of recombination is low, the high density of cells carrying many copies of plasmid allowed obtaining cells with recombination at target site.

Spiroplasma cells are known to be sensitive to only three antibiotics - tetracycline, gentamycin and chloramphenicol (Duret et al., 2005). However, genetic studies on *Spiroplasma* have shown that tetracycline resistance gene works as the best marker for selection of positive clones (Breton et al., 2011). This limits the options available for generating multiple gene disruption mutants since use of multiple antibiotics would further hamper the growth rate of *Spiroplasma* cells. Thus, removal of the resistance marker upon obtaining the first gene disruption mutant would allow usage of same resistance marker for disrupting the next gene. This is possible as is evident from experiments by Duret et al., (2005). The authors prepared vectors (pSD6 and pSD61) with the *tetM* gene and its promoter sandwiched between two resolvase (*res*) sites. Upon obtaining target gene mutants, the *tetM* gene from the genomic DNA of mutant was removed by vector-based expression of Resolvase protein (using a separate vector) and tetracycline sensitive cells with target gene disrupted were obtained. The enzyme Resolvase cuts the target DNA between the two *res* sites (Malaga et al., 2003). However, the method is not foolproof to obtain target gene disruption mutants since tetracycline-resistant cells carrying free plasmids were still observed (Duret et al., 2005).

Apart from gene disruption, another tool needed for molecular biology studies is a vector that can be used for expression of the desired protein. As mentioned earlier, initial attempts to express proteins in *Spiroplasma* involved use of replicative form (RF) of virus SpV1 (SpV1-RF), a natural infectant of most *Spiroplasma* species (Marais et al., 1993; Stamburski et al., 1991). In recombinant SpV1-RF, the gene of interest was transcribed using the promoter sequence present in the viral DNA (Stamburski et al., 1991). The proteins expressed using these strategies were obtained at low levels and the recombinant viral DNA was unstable (Marais et al., 1993). This led to a need for advanced vectors. Thus, *oriC* plasmids containing chromosomal origin of replication of *S. citri* were designed (Renaudin et al., 1995).

The *oriC* plasmids employed promoters of constitutively expressing genes such as *fibril* or *spiralin* genes and tetracycline resistance (*tetM*) as selection marker (Renaudin et al., 1995). Upon a few passages it was found that the vectors often integrated into genomic DNA. As a further modification, reduction/disruption of the *dnaA* gene in *oriC* vectors led to extra-chromosomal maintenance of the vectors (Renaudin et al., 1995). Hence, the vectors containing *S. citri oriC* and resistance marker in addition to cloning the gene under a constitutively expressing promoter were considered promising. However, the transformation efficiency of *S. citri* cells using *oriC* plasmids was low (Breton et al., 2008; Renaudin, 2002). This problem was overcome by use of vectors derived from pSci1-6 plasmids that are naturally present in *Spiroplasma citri* GII-3 cells. Derivatives of these plasmids were prepared by insertion of an antibiotic resistance marker *tetM* and have also been used for expression of proteins. Another advantage of these plasmids is that they can be tuned to be maintained extra-chromosomally (Breton et al., 2008). A recent study involving heterologous expression of protein in *Spiroplasma* showed that use of promoter sequence and ribosome binding site of a gene coding for protein with high expression level (*tuf*; elongation factor-Tu) resulted into over-expression of the desired protein (Renaudin et al., 2015). Thus, derivatives of pSci1-6 plasmids containing an antibiotic marker (*tetM*) under the promoter of a constitutively expressing gene (such as *spiralin*) and the desired gene under *tuf* promoter seems to be a good combination for expression of desired proteins in *Spiroplasma* cells. However, one must ensure that only the codon-optimized genes (not to use TGA codon as a STOP codon) are used for expression in *Spiroplasma* cells.

Thus, genetic tools available for *Spiroplasma* genome modification are not as efficient as the methods applicable for other bacterial species. This has hampered the progress of the field from

exploring *Spiroplasma* biology. *Spiroplasma* being a slow growing organism also slows down the progress of genetic experiments. Enlisted in Table 1.4 are the essential components in designing a plasmid for targeting specific genes.

Table 1.4 Essential components of gene disruption vectors in *Spiroplasma* (also refer Figure 1.12).

Abbreviation	Expansion	Significance
Ori	Origin of replication	Facilitates replication of the vector in *Spiroplasma*
Gmr/TetM*	Gentamycin/Tetracycline	Genes coding for proteins that provide resistance to respective antibiotics
PS	*spiralin* promoter	Facilitates expression of downstream genes
TF	*fibril* terminator	Prevents transcription of DNA downstream of the terminator sequence
Res	Resolvase site	Acts as restriction sites to facilitate removal of DNA between them by the action of protein resolvase
MCS	Multiple cloning site	Site for cloning DNA of interest
RoI	Region of interest/portion of target gene to be disrupted	Acts as a site of recombination with the genomic DNA

* Two antibiotic resistance cassettes may be used; one of them downstream of a promoter sequence while the other one downstream of the portion of target gene (RoI) to be disrupted.

The challenges associated with *Spiroplasma* genetics has restricted the number of groups working in the area to a few around the world.

1.10) Objectives

Many species of *Spiroplasma* have been identified as pathogens of plants, insects and arthropods and some species can infect animals and humans. Limited knowledge is available about the pathogenicity and cell biology of *Spiroplasma*. This may be attributed to lack of sophisticated tools for genetic studies and difficulties associated with *in vitro* and *in vivo* visualization of protein complexes. Nevertheless, cell biology studies over past five decades have revealed that *Spiroplasma* is the only cell wall deficient bacteria with distinct helical shape and move by kinking motility. *Spiroplasma* cells exhibit directed motility without the presence of conventional chemotaxis proteins.

In vitro and *in vivo* studies on *Spiroplasma* have led to proposition of a model for its motility. Though the model proposed by contribution from different groups through independent experiments may be true, some questions remain unanswered. Firstly, the model does not explain

how conformational changes are introduced in Fibril filaments. Like other bacterial motors that derive energy from the nucleotide hydrolysis (GTP/ATP), the energy source of the linear motor remains obscure. Conformational change in Fibril is the highlight of the model however the nature of changes and forces bringing about these changes remain unidentified. Also, unknown is the function of MreB in cell wall deficient bacteria and why multiple copies of *mreB* are necessary for *Spiroplasma* cells. These questions can be addressed by a combination of *in vitro* and *in vivo* characterization of the proposed proteins. Thus, the work was initiated with following objectives:

A) Structural studies on Fibril in filamentous form

Fibril has been proposed to undergo conformational changes. Thus, I decided to initiate work on purification and structural characterization of *Spiroplasma* Fibril filaments. Due to the requirement of high quantities of protein for standardization and structural studies, the protocols for heterologous expression and purification have been standardized. The purified protein filaments have been visualized using scanning as well as transmission electron microscopy.

B) Identification of domain boundaries in Fibril

Since Fibril forms constitutive filaments, it is likely to be a difficult sample for structural studies, especially using X-ray crystallography. Thus, attempts were made to identify non-polymerizing constructs of Fibril by dissecting domain boundaries. Several constructs were made and characterized to obtain information on domains constituting Fibril. The study also proved useful to understand if part of the protein is sufficient to form polymers.

C) Identification of protein(s) involved in rod-to-helical transition in Spiroplasma

To decipher the components involved in rod-to-helical transition of *Spiroplasma*, comparative studies were performed on wildtype and naturally occurring mutant *Spiroplasma* cells with the aim of identifying proteins in the mutant cell. Rescuing of the mutant cells by vector based-expression of missing protein confirmed the role of the protein.

D) Identification of the role of 5 MreBs and Fibril

The roles of Fibril and MreBs have only been speculated until now without any experimental evidence. To delineate the functions of each MreB and Fibril in *Spiroplasma*, gene deletion experiments have been carried out. Attempts have been made to generate gene disruption mutants of *Spiroplasma* that are deficient of either one of the five MreBs or Fibril. Upon

obtaining the mutants, their characterization can provide clues regarding roles of these proteins in shape determination and/or motility of *Spiroplasma*.

These studies would be useful to verify the role of identified proteins as well as to identify new proteins that might be important in *Spiroplasma* physiology and advance our understanding of *Spiroplasma* biology. Experiments towards achieving each of the above objectives are presented as separate chapters in the thesis.

1.11) References

Abhayawardhane, Y., and Stewart, G.C. (1995). *Bacillus subtilis* possesses a second determinant with extensive sequence similarity to the *Escherichia coli mreB* morphogene. J. Bacteriol. *177*, 765–773.

Ammar, E., Fulton, D., Bai, X., Meulia, T., and Hogenhout, S.A. (2004). An attachment tip and pili-like structures in insect- and plant-pathogenic spiroplasmas of the class *Mollicutes*. Arch. Microbiol. *181*, 97–105.

Amundsen, S.K., and Smith, G.R. (2003). Interchangeasble parts of the *Escherichia coli* recombination machinery. Cell *112*, 741–744.

Aquilino, A., Masiá, M., López, P., Galiana, A.J., Tovar, J., Andrés, M., and Gutierrez, F. (2015). First human systemic infection caused by *Spiroplasma*. J. Clin. Microbiol. *53*, 719–721.

Ausmees, N., Kuhn, J.R., and Jacobs-Wagner, C. (2003). The bacterial cytoskeleton: an intermediate filament-like function in cell shape. Cell *115*, 705–713.

Bartlett, T.M., Bratton, B.P., Duvshani, A., Miguel, A., Sheng, Y., Martin, Nicholas, R., Nguyen, J.P., Persat, A., Desmarais, S.M., VanNieuwenhze, M.S., et al. (2017). A periplasmic polymer curves *Vibrio cholerae* and promotes pathogenesis. Cell *168*, 172–185.

Bastian, F.O., Sanders, D.E., Forbes, W.A., Hagius, S.D., Walker, J. V, Henk, W.G., Enright, F.M., and Elzer, P.H. (2007). *Spiroplasma* spp. from transmissible spongiform encephalopathy brains or ticks induce spongiform encephalopathy in ruminants. J. Med. Microbiol. *56*, 1235–1242.

Bendezu, F.O., and Boer, P.A.J. De (2008). Conditional lethality, division defects, membrane

involution and endocytosis in *mre* and *mrd* shape mutants of *Escherichia coli*. J. Bacteriol. *190*, 1792–1811.

Beven, L., and Wroblewski, H. (1997). Effect of natural amphipathic peptides on viability, membrane potential, cell shape and motility of *Mollicutes*. Res. Microbiol. *148*, 163–175.

Bi, E., and Lutkenhaus, J. (1991). FtsZ ring structure associated with division in *Escherichia coli*. Nature *354*, 161–164.

Blair, K.M., Mears, K.S., Taylor, J.A., Fero, J., Jones, L.A., Gafken, P.R., Whitney, J.C., and Salama, N.R. (2018). The *Helicobacter pylori* cell shape promoting protein Csd5 interacts with the cell wall, MurF and the bacterial cytoskeleton. Mol. Microbiol. *110*, 114–127.

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.

Bove, J.M., Carle, P., Garnier, M., Laigret, F., Renaudin, J., and Saillard, C. (1989). Molecular and cellular biology of spiroplasmas. In The Mycoplasmas, R.F. Whitcomb, and J.G. Tully, eds. pp. 243–364.

Breton, M., Duret, S., Arricau-bouvery, N., Beven, L., and Renaudin, J. (2008). Characterizing the replication and stability regions of *Spiroplasma citri* plasmids identifies a novel replication protein and expands the genetic toolbox for plant-pathogenic spiroplasmas. Microbiology *154*, 3232–3244.

Breton, M., Duret, S., Béven, L., Dubrana, M., and Renaudin, J. (2011). I- Sce I-mediated plasmid deletion and intra-molecular recombination in *Spiroplasma citri*. J. Microbiol. Methods *84*, 216–222.

Brown, P.J.B., Kysela, D.T., and Brun, Y. V (2011). Polarity and the diversity of growth mechanisms in bacteria. Semin. Cell Dev. Biol. *22*, 790–798.

Cabeen, M.T., and Jacobs-wagner, C. (2005). Bacterial cell shape. Nat. Rev. Microbiol. *3*, 601–610.

Cabeen, M.T., Charbon, G., Vollmer, W., Born, P., Ausmees, N., Weibel, D.B., and Jacobs-

Wagner, C. (2009). Bacterial cell curvature through mechanical control of cell growth. EMBO J. *28*, 1208–1219.

Cameron, T.A., Zupan, J.R., and Zambryski, P.C. (2015). The essential features and modes of bacterial polar growth. Trends Microbiol. *23*, 347–353.

Carballido-lópez, R. (2006). Orchestrating bacterial cell morphogenesis. Mol. Microbiol. *60*, 815–819.

Carballido-López, R., Formstone, A., Li, Y., Ehrlich, S.D., Noirot, P., and Errington, J. (2006). Actin homolog MreBH governs cell morphogenesis by localization of the cell wall hydrolase LytE. Dev. Cell *11*, 399–409.

Carle, P., Saillard, C., Carre, N., Carrere, S., Duret, S., Eveillard, S., Gaurivaud, P., Gourgues, G., Gouzy, J., Salar, P., et al. (2010). Partial chromosome sequence of *Spiroplasma citri* reveals extensive viral invasion and important gene decay. Appl. Environ. Microbiol. *76*, 3420–3426.

Charbon, G., Cabeen, M.T., and Jacobs-Wagner, C. (2009). Bacterial intermediate filaments: in vivo assembly, organization, and dynamics of crescentin. Genes Dev. *23*, 1131–1144.

Christiansen, G., and Griffith, J. (1986). Visualization of the paranemic joining of homologous DNA molecules catalyzed by the RecA protein of *Escherichia coli*. Proc. Natl. Acad. Sci. U. S. A. *83*, 2066–2070.

Clark, A.J. (1973). Recombination deficient mutants of *E. coli* and other bacteria. Annu. Rev. Genet. *7*, 67–86.

Clark, T.B., Whitcomb, R.F., Tully, J.G., Mouches, C., Saillard, C., Bove, J.M., Wroblewshi, H., Carle, P., Rose, D.L., Henegar, R.B., et al. (1985). *Spiroplasrna melliferum*, a new species from the Honeybee (*Apis mellifera*). Int. J. Syst. Bacteriol. *35*, 296–308.

Cohen-Krausz, S., Cabahug, P.C., and Trachtenberg, S. (2011). The monomeric, tetrameric, and fibrillar organization of Fib: the dynamic building block of the bacterial linear motor of *Spiroplasma melliferum* BC3. J. Mol. Biol. *410*, 194–213.

Cohen, A.J., Williamson, D.L., and Brink, P.R. (1989). A motility mutant of *Spiroplasma melliferum* induced with nitrous acid. Curr. Microbiol. *18*, 219–222.

Colavin, A., Shi, H., and Huang, K.C. (2018). RodZ modulates geometric localization of the bacterial actin MreB to regulate cell shape. Nat. Commun. *9*, 1–11.

Cole, R.M., Tully, J.G., Popkin, T.J., and Bové, J.M. (1973). Morphology, ultrastructure and bacteriophage infection of the helical *Mycoplasma*-like organism (*Spiroplasma citri* gen. nov., sp. nov.) cultured from "Stubborn" disease of citrus. J. Bacteriol. *115*, 367–386.

Cooper, B.J.A., and Pollard, T.D. (1982). Methods to measure actin polymerization. In Methods in Enzymology, (Academic Press, Inc.), pp. 182–210.

Daniel, R. a., and Errington, J. (2003). Control of cell morphogenesis in bacteria: Two distinct ways to make a rod-shaped cell. Cell *113*, 767–776.

Daniels, M.J., Longland, J.M., and Gilbart, J. (1980). Aspects of motility and chemotaxis in spiroplasmas. J. Gen. Microbiol. *118*, 429–436.

Davey, M.J., and Funnell, B.E. (1997). Modulation of the P1 plasmid partition protein ParA by ATP, ADP, and P1 ParB. J. Biol. Chem. *272*, 15286–15292.

Davis, R.E., and Worley, J.F. (1973). *Spiroplasma*: motile, helical microorganism associated with corn stunt disease. Phytopathology *63*, 403–408.

Davis, R.E., Worley, J.F., Whitcomb, R.F., Ishijima, T. and, and L., S.R. (1972). Helical filaments produced by a *Mycoplasma*-like organisms. Science (80). *176*, 521–523.

Doi, M., Wachi, M., Ishino, F., Tomioka, S., Ito, M., Sakagami, Y., Suzuki, A., and Matsuhashi, M. (1988). Determinations of the DNA sequence of the *mreB* gene and of the gene products of the *mre* region that function in formation of the rod shape of *Escherichia coli* cells. J. Bacteriol. *170*, 4619–4624.

Dominguez-Escobar, J., Chastanet, A., Crevenna, A.H., Fromion, V., Wedlich-Soldner, R., and Carballido-Lopez, R. (2011). Processive movement of MreB-associated cell wall biosynthetic complexes in bacteria. Science (80). *333*, 225–228.

Duret, S., Danet, J.-L., Garnier, M., and Renaudin, J. (1999). Gene disruption through homologous recombination in *Spiroplasma citri*: An scm1-disrupted motility mutant is pathogenic. J. Bacteriol. *181*, 7449–7456.

Duret, S., Berho, N., Danet, J., Garnier, M., and Renaudin, J. (2003). Spiralin is not essential for helicity, motility or pathogenicity but is required for efficient transmission of *Spiroplasma citri* by its leafhopper vector *Circulifer haematoceps*. Appl. Environ. Microbiol. *69*, 6225–6234.

Duret, S., Andre, A., and Renaudin, J. (2005). Specific gene targeting in *Spiroplasma citri*: improved vectors and production of unmarked mutations using site-specific recombination. Microbiology *151*, 2793–2803.

Dutra, B.E., Sutera, V.A., and Lovett, S.T. (2006). RecA-independent recombination is efficient but limited by exonucleases. Proc. Natl. Acad. Sci. U. S. A. *104*, 216–221.

Dybvig, K., and Cassell, G.H. (1987). Transposition of Gram-positive transposon Tn916 in *Acholeplasma laidlawii* and *Mycoplasma pulmonis*. Science (80). *235*, 1392–1394.

Dye, N.A., Pincus, Z., Fisher, I.C., Shapiro, L., and Theriot, J.A. (2011). Mutations in the nucleotide binding pocket of MreB can alter cell curvature and polar morphology in Caulobacter. Mol. Microbiol. *81*, 368–394.

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.

van den Ent, F., and Löwe, J. (2000). Crystal structure of the cell division protein FtsA from *Thermotoga maritima*. EMBO J. *19*, 5300–5307.

van den Ent, F., Amos, L. a, and Löwe, J. (2001). Prokaryotic origin of the actin cytoskeleton. Nature *413*, 39–44.

Van den Ent, F., Amos, L.A., and Löwe, J. (2001). Prokaryotic origin of the actin cytoskeleton. Nature *413*, 39–44.

Eskafi, F.M., McCoy, R.E., and Norris, R.C. (1987). Pathology of *Spiroplasma floricola* in *Galleria mellonella* larvae. J. Invertebr. Pathol. *49*, 1–13.

Etienne, N., Bret, L., Brun, C. Le, Lecuyer, H., Moraly, J., Lanternier, F., Hermine, O., Ferroni, A., Lecuit, M., Pereyre, S., et al. (2018). Disseminated *Spiroplasma apis* infection in patient with agammaglobulinemia, France. Emerg. Infect. Dis. *24*, 2382–2384.

Ferrero, R.L., and Lee, A. (1988). Motility of *Campylobacter jejuni* in a viscous environment: comparison with conventional rod-shaped bacteria. J. Gen. Microbiol. *134*, 53–59.

Figge, R.M., Divakaruni, A. V, and Gober, J.W. (2004). MreB, the cell-shape determining bacterial actin homolog, coordinates cell wall morphogenesis in. Genomics *51*, 1–47.

Fletcher, J., Schultz, G.A., Davis, R.E., Eastman, C.E., and Goodman, R.M. (1981). Brittle root disease of horseradish: evidence for an etiological role of *Spiroplasma citri*. Phytopathology *71*, 1073–1080.

Foissac, X., Danet, J.L., Saillard, C., Gaurivaud, P., Laigret, F., Paré, C., and Bové, J.M. (1997b). Mutagenesis by insertion of Tn4001 into the genome of *Spiroplasma citri*: characterization of mutants affected in plant pathogenicity and transmission to the plant by the leafhopper vector *Circulifer haematoceps*. Mol. Plant-Microbe Interact. *10*, 454–461.

Foissac, X., Saillard, C., and Bove, J.M. (1997a). Random insertion of transposon *Tn*4001 in the genome of *Spiroplasma citri* strain GII3. Plasmid *37*, 80–86.

Freeman, B.A., Sissenstein, R., Mcmanus, T.T., Woodward, J.E., Lee, I.M., and B, M.J. (1976). Lipid composition and lipid metabolism of *Spiroplasma citri*. J. Bacteriol. *125*, 946–954.

Garner, E.C., Bernard, R., Wang, W., Zhuang, X., Rudner, D.Z., and Mitchison, T. (2011). Coupled, circumferential motions of the cell wall synthesis machinery and MreB filaments in *B. subtilis*. Science (80). *333*, 222–225.

Ghachi, M. El, Matteï, P., Ecobichon, C., Martins, A., Hoos, S., Schmitt, C., Colland, F., Ebel, C., Prévost, M., Gabel, F., et al. (2011). Characterization of the elongasome core PBP2 : MreC complex of *Helicobacter pylori*. Mol. Microbiol. *82*, 68–86.

Gilad, R., Porat, A., and Trachtenberg, S. (2003). Motility modes of *Spiroplasma melliferum* BC3: a helical, wall-less bacterium driven by a linear motor. Mol. Microbiol. *47*, 657–669.

Gitai, Z., Dye, N.A., Reisenauer, A., Wachi, M., and Shapiro, L. (2005). MreB Actin-mediated segregation of a specific region of a bacterial chromosome. Cell *120*, 329–341.

Hayashi, I., Oyama, T., and Morikawa, K. (2001). Structural and functional studies of MinD ATPase: implications for the molecular recognition of the bacterial cell division apparatus. Embo

J *20*, 1819–1828.

Jacob, C., Nouzières, F., Duret, S., Bové, J.M., and Renaudin, J. (1997). Isolation, characterization, and complementation of a motility mutant of *Spiroplasma citri*. J. Bacteriol. *179*, 4802–4810.

Jarrell, K.F., and McBride, M.J. (2008). The surprisingly diverse ways that prokaryotes move. Nat. Rev. Microbiol. *6*, 466–476.

Jones, L.J.F., Carballido-López, R., and Errington, J. (2001). Control of cell shape in bacteria: helical, actin-like filaments in *Bacillus subtilis*. Cell *104*, 913–922.

Kang, C., Nyayapathy, S., Lee, J., Suh, J., and Husson, R.N. (2008). Wag31 , a homologue of the cell division protein DivIVA, regulates growth, morphology and polar cell wall synthesis in mycobacteria. Microbiology *154*, 725–735.

Kellenberger, E., Johansen, R., Maeder, M., Bohrmann, B., Stauffer, E., and Villiger, W. (1992). Artefacts and morphological changes during chemical fixation. J. Microsc. *168*, 181–201.

Komeili, A., Li, Z., Newman, D.K., and Jensen, G.J. (2006). Magnetosomes are cell membrane invaginations organized by the actin-like protein MamK. Science. *229*, 424–427.

Krzyżek, P., and Gościniak, G. (2018). Morphology of *Helicobacter pylori* as a result of peptidoglycan and cytoskeleton rearrangements. Gastroenterol. Rev. *13*, 182–195.

Ku, C., Lo, W.-S., and Kuo, C.-H. (2014). Molecular evolution of the actin-like MreB protein gene family in wall-less bacteria. Biochem. Biophys. Res. Commun. *446*, 927–932.

Kühn, J., Briegel, A., Mörschel, E., Kahnt, J., Leser, K., Wick, S., Jensen, G.J., and Thanbichler, M. (2010). Bactofilins, a ubiquitous class of cytoskeletal proteins mediating polar localization of a cell wall synthase in *Caulobacter crescentus*. EMBO J. *29*, 327–339.

Kürner, J., Frangakis, A.S., and Baumeister, W. (2005). Cryo–electron tomography reveals the cytoskeletal structure of *Spiroplasma melliferum*. Science. *307*, 436–439.

Labarère, J., and Barroso, G. (1989). Lethal and mutation frequency responses of *Spiroplasma citri* cells to UV irradiation. Mutat. Res. - Fundam. Mol. Mech. Mutagen. *210*, 135–141.

Lafleche, D. and, and Bove, J.M. (1970). Mycoplasmes dans les argumes atteints de "Greening ", de “Stubborn” ou de maladies similaires. Fruits *25*, 455–465.

Lange, A.T., Macia, E., Guzzo, M., Hot, E., Faure, L.M., Jakobczak, B., Espinosa, L., Alcor, D., Ducret, A., Keilberg, D., et al. (2015). The small G-protein MglA connects to the MreB actin cytoskeleton at bacterial focal adhesions. J. Cell Biol. *210*, 243–256.

Larsen, R. a., Cusumano, C., Fujioka, A., Lim-Fong, G., Patterson, P., and Pogliano, J. (2007). Treadmilling of a prokaryotic tubulin-like protein, TubZ, required for plasmid stability in *Bacillus thuringiensis*. Genes Dev. *21*, 1340–1352.

Lartigue, C., Duret, S., Garnier, M., and Renaudin, J. (2002). New plasmid vectors for specific gene targeting in *Spiroplasma citri*. Plasmid *48*, 149–159.

Levin, P.A., Margolis, P.S., Setlow, P., Losick, R., and Sun, D. (1992). Identification of *Bacillus subtilis* genes for septum placement and shape determination. J. Bacteriol. *174*, 6717–6728.

Liang, T., Li, X., Du, J., Yao, W., Sun, G., Dong, X., Liu, Z., Ou, J., Meng, Q., Gu, W., et al. (2011). Identification and isolation of a *Spiroplasma* pathogen from diseased freshwater prawns, *Macrobrachium rosenbergii*, in China : A new freshwater crustacean host. Aquaculture *318*, 1–6.

Liu, P., Zheng, H., Meng, Q., Terahara, N., Gu, W., Wang, S., Zhao, G., Nakane, D., Wang, W., and Miyata, M. (2017). Chemotaxis without conventional two-component system, based on cell polarity and aerobic conditions in helicity-switching swimming of *Spiroplasma eriocheiris*. Front. Microbiol. *8*, 1–13.

Löwe, J., and Amos, L.A. (1998). Crystal structure of the bacterial cell-division protein FtsZ. Nature *391*, 203–206.

Madabhushi, R., and Marians, K.J. (2009). Actin homolog MreB affects chromosome segregation by regulating topoisomerase IV in *Escherichia coli*. Mol. Cell *33*, 171–180.

Malaga, W., Perez, E., and Guilhot, C. (2003). Production of unmarked mutations in mycobacteria using site-specific recombination. FEMS Microbiol. Lett. *219*, 261–268.

Marais, A., Bove, J.M., Dallo, S.F., Baseman, J.B., and Renaudin, J. (1993). Expression in *Spiroplasma citri* of an epitope carried on the G fragment of the cytadhesin P1 gene from

Mycoplasma pneumoniae. J. Bacteriol. *175*, 2783–2787.

Marais, A., Bove, J.M., and Ranaudin, J. (1996). *Spiroplasma citri* virus SpV1-derived cloning vector: deletion formation by illegitimate and homologous recombination in a Spiroplasmal host strain which probably lacks a functional *recA* gene. J. Bacteriol. *178*, 862–870.

Markham, P.G., Townsend, R., Bar-Joseph, M., Daniels, M.J., Plaskitt, A., and Meddins, B.M. (1974). Spiroplasmas are the causal agents of citrus little-leaf disease. Ann. Appl. Biol. *78*, 49–57.

Mendelson, N.H. (1976). Helical growth of *Bacillus subtilis*: A new model of cell growth (cell surface organization/cell elongation/tension-restricted structures/developmental biology of a procaryote). Proc. Natl. Acad. Sci. U. S. A. *73*, 1740–1744.

Møller-Jensen, J., Borch, J., Dam, M., Jensen, R.B., Roepstorff, P., and Gerdes, K. (2003). Bacterial mitosis: ParM of plasmid R1 moves plasmid DNA by an Actin-like insertional polymerization mechanism. Mol. Cell *12*, 1477–1487.

Mouches, C., Bove, J.M., Tully, J.G., Rose, D.L., McCoy, R.E., Carle-Junca, P., Garnier, M., and Saillard, C. (1983). *Spiroplasma apis*, a new species from the honey bee *Apis mellifera*. Ann. l'Institut Pasteur / Microbiol. *134*, 383–397.

Mueller, N.J., Tini, G.M., Weber, A., Gaspert, A., Husmann, L., Bloemberg, G., Boehler, A., and Benden, C. (2015). Hepatitis from *Spiroplasma* sp. in an immunocompromised patient. Am. J. Transplant. *15*, 2511–2516.

Nunan, L.M., Pantoja, C.R., Salazar, M., Aranguren, F., and Lightner, D. V (2004). Characterization and molecular methods for detection of a novel *Spiroplasma* pathogenic to *Penaeus vannamei*. Dis. Aquat. Organ. *62*, 255–264.

Oliva, M.A., Halbedel, S., Freund, S.M., Dutow, P., Leonard, T.A., Veprintsev, D.B., Hamoen, L.W., and Lowe, J. (2010). Features critical for membrane binding revealed by DivIVA crystal structure. EMBO J. *29*, 1988–2001.

Ozbek, E., Miller, S.A., Meulia, T., and Hogenhout, S.A. (2003). Infection and replication sites of *Spiroplasma kunkelii* (Class: Mollicutes) in midgut and malpighian tubules of the leafhopper

Dalbulus maidis. Int. J. Invertebr. Pathol. *82*, 167–175.

Pereyre, S., Sirand-Pugnet, P., Beven, L., Charrin, A., Renaudin, H., Barre, A., Avenaud, P., Jacob, D., Couloux, A., Barbe, V., et al. (2009). Life on arginine for *Mycoplasma hominis*: clues from its minimal genome and comparison with other human urogenital *Mycoplasmas*. PLoS Genet. *5*.

Pilhofer, M., Ladinsky, M.S., McDowall, A.W., Petroni, G., and Jensen, G.J. (2011). Microtubules in bacteria: ancient tubulins build a five-protofilament homolog of the eukaryotic cytoskeleton. PLoS Biol. *9*, e1001213.

Pollard, T.D. (2003). The cytoskeleton, cellular motility and the reductionist agenda. Nature *422*, 741–745.

RayChaudhuri, D., and Park, J.T. (1992). *Escherichia coli* cell-division gene *ftsZ* encodes a novel GTP-binding protein. Nature *359*, 251–254.

Razin, S. (1978). The *Mycoplasmas*. Microbiol Rev *42*, 414—470.

Razin, S., Hasin, M., Ne'eman, Z., and Rottem, S. (1973). Isolation, chemical composition, and ultrastructural features of the cell membrane of the *Mycoplasma*-like organism *Spiroplasma citri*. J. Bacteriol. *116*, 1421–1435.

Razin, S., Yogev, D., and Naot, Y. (1998). Molecular biology and pathogenicity of *Mycoplasmas*. Microbiol. Mol. Biol. Rev. *62*, 1094–1156.

Renaudin, J. (2002). Extrachromosomal elements and gene transfer. In Molecular Biology and Pathology of *Mycoplasmas*, S. Razin, and K. Herrmann, eds. (Plenum Publishers, New York), pp. 347–370.

Renaudin, J., Marais, A., Verdin, E., Duret, S., Foissac, X., Laigret, F., and Bove, J. (1995). Integrative and free *Spiroplasma citri oriC* plasmids: Expression of the *Spiroplasma phoeniceum spiralin* in *Spiroplasma citri*. J. Bacteriol. *177*, 2870–2877.

Renaudin, J., Béven, L., Batailler, B., Duret, S., Desqué, D., Arricau-bouvery, N., Malembic-maher, S., and Foissac, X. (2015). Heterologous expression and processing of the flavescence dorée phytoplasma variable membrane protein VmpA in *Spiroplasma citri*. BMC Microbiol. *15*,

1–12.

Rottem, S., and Greenberg, A.S. (1975). Changes in composition, biosynthesis, and physical state of membrane lipids occurring upon aging of *Mycoplasma hominis* cultures. J. Bacteriol. *121*, 631–639.

Rottem, S., and Razin, S. (1973). Membrane lipids of *Mycoplasma hominis*. J. Bacteriol. *113*, 565–571.

Saglio, P., Lhospital, M., Lafleche, D., Dupont, G., Bove, J.M., Tully, J.G., and Freundt, E.A. (1973). *Spiroplasma citri* gen. and sp. n.: A *Mycoplasma*-like organism associated with "Stubborn" disease of citrus. Int. J. Syst. Bacteriol. *23*, 191–204.

Saillard, C., Vignault, J.C., Bove, J.M., Raie, A., Tully, J.G., Williamson, D.L., Fos, A., Garnier, M., Gadeau, A., Carle, P., et al. (1987). *Spiroplasma phoeniceum* sp . nov., a new plant-pathogenic species from Syria. Int. J. Syst. Bacteriol. *37*, 106–115.

Shaevitz, J.W., Lee, J.Y., and Fletcher, D.A. (2005). *Spiroplasma* swim by a processive change in body helicity. Cell *122*, 941–945.

Shi, H., Bratton, B.P., Gitai, Z., and Huang, K.C. (2018). How to build a bacterial cell: MreB as the foreman of *E. coli* construction. Cell *172*, 1294–1305.

Shih, Y.-L., and Rothfield, L. (2006). The bacterial cytoskeleton. Microbiol. Mol. Biol. Rev. *70*, 729–754.

Smith, G.R. (1988). Homologous recombination in prokaryotes. Microbiol. Rev. *52*, 1–28.

Specht, M., Schatzle, S., Graumann, P.L., and Waidner, B. (2011). *Helicobacter pylori* possesses four coiled-coil-rich proteins that form extended filamentous structures and control cell shape and motility. J. Bacteriol. *193*, 4523–4530.

Stamburski, C., Renaudin, J., and Bove, J.M. (1991). First step toward a virus-derived vector for gene cloning and expression in spiroplasmas, organisms which read UGA as a tryptophan codon: Synthesis of chloramphenicol acetyltransferase in *Spiroplasma citri*. J. Bacteriol. *173*, 2225–2230.

Sycuro, L.K., Pincus, Z., Gutierrez, K.D., Biboy, J., Stern, C.A., Vollmer, W., and Salama, N.R. (2010). Peptidoglycan crosslinking relaxation promotes *Helicobacter pylori*'s helical shape and stomach colonization. Cell *141*, 822–833.

van Teeffelen, S., Wang, S., Furchtgott, L., Huang, K.C., Wingreen, N.S., Shaevitz, J.W., and Gitai, Z. (2011). The bacterial actin MreB rotates, and rotation depends on cell-wall assembly. Proc. Natl. Acad. Sci. *108*, 15822–15827.

Terahara, N., Tulum, I., and Miyata, M. (2017). Transformation of crustacean pathogenic bacterium *Spiroplasma eriocheiris* and expression of yellow fluorescent protein. Biochem. Biophys. Res. Commun. *487*, 488–493.

Tilby, M.J. (1977). Helical shape and wall synthesis in a bacterium. Nature *266*, 450–452.

Townsend, R., and Archer, D.B. (1983). A Fibril protein antigen specific to *Spiroplasma*. J. Gen. Microbiol. *129*, 199–206.

Townsend, R., Markham, P.G., Plaskitt, K.A., and Daniels, M.J. (1977). Isolation and characterization of a non-helical strain of *Spiroplasma citri*. J. Gen. Microbiol. *100*, 15–21.

Townsend, R., Burgess, J., and Plaskitt, K. a. (1980a). Morphology and ultrastructure of helical and nonhelical strains of *Spiroplasma citri*. J. Bacteriol. *142*, 973–981.

Townsend, R., Archer, D.B., and Plaskitt, K.A. (1980b). Purification and preliminary characterization of *Spiroplasma* fibrils. J. Bacteriol. *142*, 694–700.

Trachtenberg, S., and Gilad, R. (2001). A bacterial linear motor: Cellular and molecular organization of the contractile cytoskeleton of the helical bacterium *Spiroplasma melliferum* BC3. Mol. Microbiol. *41*, 827–848.

Trachtenberg, S., Gilad, R., and Geffen, N. (2003a). The bacterial linear motor of *Spiroplasma melliferum* BC3: from single molecules to swimming cells. Mol. Microbiol. *47*, 671–697.

Trachtenberg, S., Andrews, S.B., and Leapman, R.D. (2003b). Mass distribution and spatial organization of the linear bacterial motor of *Spiroplasma citri* R8A2. J. Bacteriol. *185*, 1987–1994.

Trachtenberg, S., Dorward, L.M., Speransky, V. V., Jaffe, H., Andrews, S.B., and Leapman, R.D. (2008). Structure of the cytoskeleton of *Spiroplasma melliferum* BC3 and its interactions with the cell membrane. J. Mol. Biol. *378*, 778–789.

Tully, J.G., Whitcomb, R.F., Williamson, D.L., and Clark, H.F. (1976). Suckling mouse cataract agent is a helical wall-free prokaryote (*Spiroplasma*) pathogenic for vertebrates. Nature *259*, 117–120.

Ursell, T.S., Nguyen, J., Monds, R.D., Colavin, A., Billings, G., Ouzounov, N., Gitai, Z., Shaevitz, J.W., and Huang, K.C. (2014). Rod-like bacterial shape is maintained by feedback between cell curvature and cytoskeletal localization. Proc. Natl. Acad. Sci. *111*, E1025–E1034.

Vale, R.D., and Milligan, R.A. (2000). The way things move: looking under the hood of molecular motor proteins. Science. *288*, 88–96.

Wagstaff, J., and Löwe, J. (2018). Prokaryotic cytoskeletons: protein filaments organizing small cells. Nat. Rev. Microbiol. *16*, 187–201.

Waidner, B., Specht, M., Dempwolff, F., Haeberer, K., Schaetzle, S., Speth, V., Kist, M., and Graumann, P.L. (2009). A novel system of cytoskeletal elements in the human pathogen *Helicobacter pylori*. PLoS Pathog. *5*, e1000669.

Wang, W., Gi, W., Ding, Z., Ren, Y., Chen, J., and Hou, Y. (2005). A novel *Spiroplasma* pathogen causing systemic infection in the crayfish *Procambarus clarkii* (Crustacea: Decapod), in China. FEMS Microbiol. Lett. *249*, 131–137.

Wang, W., Gu, W., Gasparich, G.E., Bi, K., Ou, J., Meng, Q., Liang, T., Feng, Q., Zhang, J., and Zhang, Y. (2011). *Spiroplasma eriocheiris* sp. nov., associated with mortality in the Chinese mitten crab, *Eriocheir sinensis*. Int. J. Syst. Evol. Microbiol. *61*, 703–708.

Weisburg, W.G., Tully, J.G., Rose, D.L., Petzel, J.P., Oyaizu, H., Yang, D., Mandelco, L., Sechrest, J., Lawrence, T.G., Van Etten, J., et al. (1989). A phylogenetic analysis of the *Mycoplasmas*: Basis for their classification. J. Bacteriol. *171*, 6455–6467.

Whitcomb, R.F., Chen, T.A., Williamson, D.L., Lia, C., Tully, J.G., Bove, J.M., Mouches, C., Rose, D.L., Coan, M.E., and Clark, T.B. (1986). *Spiroplasma kunkelii* sp. nov. : Characterization

of the etiological agent of corn stunt disease. Int. J. Syst. Bacteriol. *36*, 170–178.

Williamson, D.L. (1974). Unusual fibrils from the spirochete-like sex ratio organism. J. Bacteriol. *117*, 904–906.

Williamson, D.L., Renaudin, J., and Bove, J.M. (1991). Nucleotide sequence of the *Spiroplasma citri* Fibril protein gene. J. Bacteriol. *173*, 4353–4362.

Williamson, D.L., Sakaguchi, B., Hackett, K.J., Whitcomb, R.F., Tully, J.G., Carle, P., Bove, J.M., R, A.J., Konai, M., and Henegar, R.B. (1999). *Spiroplasma poulsonii* sp. nov., a new species associated with male-lethality in *Drosophila willistoni*, a neotropical species of fruit fly. Int. J. Syst. Bacteriol. *49*, 611–618.

Woldemeskel, S.A., and Goley, E.D. (2017). Shapeshifting to survive: Shape determination and regulation in *Caulobacter crescentus*. Trends Microbiol. 1–15.

Yang, D.C., Blair, K.M., Taylor, J.A., Petersen, T.W., Sessler, T., Tull, C.M., Leverich, C., Collar, A.L., Wyckoff, T.J., Biboy, J., et al. (2019). A genome-wide *Helicobacter pylori* morphology screen uncovers a membrane spanning helical cell shape complex. J. Bacteriol. 10.1128/JB.00724-18. [Epub ahead of print].

Yang, X., Lyu, Z., Miguel, A., Mcquillen, R., Huang, K.C., and Xiao, J. (2017). GTPase activity – coupled treadmilling of the bacterial tubulin FtsZ organizes septal cell wall synthesis. Science. *355*, 744–747.

Ye, F., Renaudin, J., Bové, J.M., and Laigret, F. (1994). Cloning and sequencing of the replication origin (*oriC*) of the *Spiroplasma citri* chromosome and construction of autonomously replicating artificial plasmids. Curr. Microbiol. *29*, 23–29.

Zhou, X., Halladin, D.K., and Theriot, J.A. (2016). Fast mechanically driven daughter cell separation is widespread in Actinobacteria. MBio *7*, 1–6.

Chapter 2

Purification and characterization of full-length Fibril filaments

2.1) Introduction

Current model for *Spiroplasma* motility proposes a role of conformational changes in Fibril to be responsible for the characteristic kinking motility (Kürner et al., 2005; Shaevitz et al., 2005; Trachtenberg and Gilad, 2001; Trachtenberg et al., 2008). This model has been proposed on the basis of 14-17 Å resolution data (Cohen-Krausz et al., 2011). Identification of conformational changes in proteins is impossible at this resolution. This necessitates the need for high resolution studies. This work was limited, in part, due to challenges in obtaining optimum samples of Fibril for electron microscopy reconstruction and the unavailability of advanced electron microscopes and image processing methods.

Advancements in electron cryo-microscopy (cryoEM) in the recent past has made it possible to obtain atomic resolution structures of proteins (Cheng, 2015). Single particle analysis (SPA) and helical reconstruction are widely used techniques to determine structures of protein complexes using cryoEM. In SPA, a pre-requisite for high resolution structure determination using cryoEM is obtaining a homogeneous sample with a wide range of orientations (Cheng et al., 2015). This can be achieved by careful biochemical studies and sample standardization using cryoEM. While SPA is predominantly used for homogenous, non-polymerizing samples, polymeric proteins with inherent helical symmetry are processed by helical reconstruction method (Paul et al., 2004). In order to identify conformational changes in Fibril filaments, we initiated its biochemical and structural characterization. Use of advanced technology will enable us to obtain atomic resolution information.

Fibril forms constitutive filaments with no known method to control polymerization (Cohen-Krausz et al., 2011). Thus, it is not possible to obtain ‘single particles’ for structural studies. In such a case, the strategy for structure determination can be decided by obtaining preliminary information about the property of Fibril filaments, the important question being whether they are helical or not. If the filaments are helical, helical reconstruction may be performed to obtain structure of the filament. Alternatively, if Fibril forms filament bundles, the bundles require to be

dissociated into individual filaments. These procedures require tremendous amounts of protein for biochemical standardization. Since Fibril content of *Spiroplasma* cells is low (about 1.4×10^4 molecules/cell; Trachtenberg et al., 2014), it may not be the best source to obtain Fibril for structural characterization. While protocol for purification of Fibril from *Spiroplasma* has already been standardized (Cohen-Krausz et al., 2011), a pure culture of any species of *Spiroplasma* is unavailable in India. *Spiroplasma* being a known pathogen, special permission is required from government of India for its import. However, since the import permission sanction is a lengthy procedure, alternate strategies for Fibril expression and purification are necessary. Further, genetic modifications, if any required, are difficult in *Spiroplasma* cells. Use of standard heterologous expression systems such as *Escherichia coli* and *Pichia pastoris* can help overcome these difficulties. Thus, work was initiated for overexpressing *fibril* gene of *Spiroplasma citri* in *E. coli* cells and standardization of expression and protein purification protocols.

This chapter describes the attempts for purification of Fibril filaments using *E. coli* and *P. pastoris* as heterologous expression systems as well as from *S. citri*, its natural source. Multiple protocols have been standardized for purification of Fibril using these organisms. Further, the standardization of purified protein for visualization using electron microscopy is described.

2.2) Material and methods

Since their discovery, Fibril polymers have been purified from *Spiroplasma* using Triton X-100 and sodium deoxycholate as detergents. The detergent (Triton X-100 and sodium deoxycholate)-insoluble nature of Fibril was used as a starting point for standardization of multiple protocols for purification of Fibril expressed in *E. coli*. The standardized protocols are explained in the next few sections.

2.2.1) Cloning

fibril gene (NCBI accession number CAK99396) was amplified using genomic DNA of *Spiroplasma citri* as the template (DSMZ accession no 21846). In spiroplasmas, the TGA codon codes for amino acid tryptophan (Citti et al., 1992) while in the organisms used as heterologous expression systems (*E. coli* and *P. pastoris*), it represents a stop codon. Fibril protein from *S. citri* has seven tryptophans of which six are coded by TGA. Thus, these six TGA codons

(corresponding to amino acid numbers 26, 200, 258, 316, 323 and 342) were modified to TGG in *fibril* gene amplified from *S. citri* genomic DNA, using appropriate primers (Table 2.1).

Table 2.1 List of primers used for amplification, mutation and cloning of *fibril* gene into pHIS17 and pPICZα vectors. Restriction sites in the primers are highlighted in red.

Sr. no	Name of the primer	Sequence (5'→ 3')	Purpose
1	ScF-f	GTTTAACTTTAAGAAGGAGATATACATATGATCGGTGTAATTTCAACAGCATATTTTACG	Cloning full length *fibril* with or without a hexahistidine (His_6) tag for expression in *E. coli*.
2	ScF-Nhisf	CTTTAAGAAGGAGATATACATATGCGTGGCCATCATCATCATCATCATCATATCGGTGTAATTTCAAC	
3	ScF-r	GCTTTTAATGATGATGATGATGATGGGATCCATCATTACTCACTTTTTAAACGAATTGTAACTTTG	
4	ScF-Chis-r	GCTTTTAATGATGATGATGATGATGGGATCCATCATTCTCACTTTTTAAACGAATTGTAACTTTG	
5	ScF-W26A-f	CAGTAAAAAAATATTGGTGGAAGAACTGTGTAATTC	Conversion of tryptophan codons from TGA to TGG
6	ScF-W200-f	AATGCATTATCACCATGGGATAATGATCCAATCCATA	
7	ScF-W258-f	GAAATTAAACATGACCAATGGATTAAATTATTTAAACC	
8	ScF-W316-322f	GTAATTTTAGGAGAGGATGAATGGAAAAATGCTCCTAAAAAATGGTTACGTAAATTATTA	
9	ScF-W316-322-r	TAATAATTTACGTAACCATTTTTTAGGAGCATTTTTCCATTCATCCTCTCCTAAAATTAC	
10	ScF-W342-r	CATATTTTGCTGATTTATTCCATAATAATTCATCATCAT	
11	pPICZalpha Kpn1-Fib-F	CGAGAAAAGAGAGGCTGAAGCTGGTACCATCGGCGTAATTTCAACAGC	Cloning full length *fibril* with a N-terminal α factor signal sequence and C-terminal C-myc epitope followed by His_6 tag for expression in *P. pastoris*.
12	PPICZalpha-Not1-Fib-R	GTTTTTGTTCTAGAAAGCTGGCGGCCGCTTCACTTTTTAAACGAATTG	

For cloning and expression of *fibril* in *E. coli*, codons for hexahistidine (His_6) tag were added at the 5' or 3' end of the *fibril* gene to obtain full length Fibril protein with His_6 tag at the N and C termini respectively. The gene was cloned into pHIS17 vector that has an ampicillin resistance marker (Addgene vector #78201), by restriction-free cloning (RFC) method (van den Ent and Löwe, 2006). Primers used for gene amplification, point mutations and restriction-free cloning

were purchased from Integrated DNA Technologies (IDT) and Sigma-Aldrich®, USA and have been listed in Table 2.1. Pfu Turbo® and Accuprime Pfx polymerase used for polymerase chain reactions (PCR) were purchased from Agilent technologies and Thermo Fischer Scientific respectively. All PCRs were performed using Eppendorf PCR machines.

fibril gene was cloned into pPICZα vector for expression in *P. pastoris* cells. The vector has a Zeocin™ resistance marker gene and is useful for expression and secretion of recombinant proteins using *P. pastoris*. The gene of interest can be expressed under alcohol oxidase 1 (AOX1) promoter by induction with methanol (Daly and Hearn, 2005; Ellis et al., 1985). Forward and reverse primers were designed with KpnI and NotI restriction sites respectively. Cloning was done by restriction digestion-ligation method as described in section 2.2.1.3).

2.2.1.1) *fibril* gene amplification and site-directed mutagenesis

fibril gene was amplified from *S. citri* genomic DNA using gene-specific primers. Appropriate primers were designed to incorporate point mutations to convert TGA to TGG. A series of extension PCRs were then performed to mutate the six TGA codons (corresponding to amino acid positions 26, 200, 258, 316, 323 and 342) in *fibril*, while amplifying the whole gene. The product so obtained with necessary mutations was cloned into pHIS17 vector by restriction-free cloning method. All the primers used for cloning experiments mentioned in this chapter have been tabulated in Table 2.1.

2.2.1.2) Restriction-Free cloning (RFC)

Primers for RFC were designed as described by van den Ent and Löwe, (2006). Briefly, the 5' end of primers contained sequence from target vector and the sequence at 3' was from the gene of interest. In the first round of PCR reaction, the gene of interest was amplified using these primers. In the next step, the amplified gene containing regions from vector at the 5' and 3' ends was used as megaprimer and the target vector, in this case pHIS17, was used as the template to obtain a plasmid containing the gene of insert. A reaction with template plasmid but no primers acted as the negative control. 0.5 µL of restriction enzyme DpnI was added to make a final volume of 20 µL each of the test and control reactions. The enzyme DpnI degrades methylated template DNA in contrast to non-methylated DNA obtained by PCR. DpnI treated PCR products were then transformed into *E. coli* NEB®Turbo cells by electroporation. The NEB® Turbo *E. coli*

[genotype F' *proA*$^{+}$*B*$^{+}$ *lacI*^{q} Δ*lacZM15 / fhuA2* Δ*(lac-proAB) glnV galK16 galE15 R(zgb-210::Tn10)Tet*^{S} *endA1 thi-1* Δ*(hsdS-mcrB)5*] cells were chosen for their high transformation efficiency (1-3 × 10^{9} cfu/µg pUC19 DNA; NEB Inc. catalogue number C29841). Electroporation conditions used for transformation were- 2500 V voltage, 25 µF capacitance and 200 Ω resistance. Equal volume of (2 X) Luria Bertani (LB) broth was added to the electroporation reaction followed by incubation at 37 ^{0} C for 1 hour under shaking condition. Subsequently the cells were spread onto LB agar plates supplemented with ampicillin at 100 µg/mL final concentration and incubated at 37 ^{0} C for 12-16 hours. Observation of colonies on test plate and no colonies on control plate indicated complete digestion of template DNA and successful transformation of PCR product. Colonies from the test plate were grown in LB broth supplemented with ampicillin and used for plasmid purification. The presence of gene of interest in plasmid was checked initially by restriction digestion of the plasmid. The region containing gene of interest was sequenced to confirm the clones. A clone with the correct sequence was selected for expression and purification of Fibril protein from *E. coli* cells.

2.2.1.3) Cloning of *fibril* gene in pPICZα by restriction digestion-ligation method and transformation into *P. pastoris* cells

For cloning by restriction digestion-ligation method, the forward and reverse primers were designed with KpnI and NotI restriction sites. *fibril* gene was amplified using *fibril* in pHIS17 clone as the template and appropriate primers listed in Table 2.1. The amplicon and empty pPICZα vector (Invitrogen; obtained from Dr. A P Giri's lab at CSIR-NCL, Pune, India) were incubated with KpnI and NotI restriction enzymes and Cutsmart™ buffer (New England Biolabs) to produce sticky ends. The restriction digestion reactions were subjected to PCR clean-up using a kit (Qiagen) to remove unwanted small DNA fragments, restriction enzymes and buffer. A ligation reaction was set up to stitch the insert with the vector using T4 DNA ligase enzyme (New England Biolabs). The reaction was incubated at 4 ^{0} C for 12-14 hours and then transformed into *E. coli* DH5α electro-competent cells (genotype: *fhuA2 (argF-lacZ)U169 phoA glnV44 80 (lacZ)M15 gyrA96 recA1 relA1 endA1 thi-1 hsdR17*; NEB catalogue number C2987H) by electroporation. *E. coli* DH5α electro-competent cells were chosen for their high transformation efficiency, high quality plasmid preparations as well as *recA* and *endA* deficiency. Equal volume of Low salt- Luria Bertani {LSLB; [1 % tryptone (w/v), 0.5 % NaCl (w/v), 0.5 %

yeast extract (w/v); adjust pH to 7.5; 1.5 % agar (w/v)]} broth was added to electroporation reaction and incubated at 37 ^{0} C for 1 hour with shaking. The recovered cells were plated onto LSLB agar plates supplemented with 25 µg/mL Zeocin™ and incubated at 37 ^{0} C for 12-14 hours. The colonies were checked for presence of insert by colony PCR. Colonies found positive by colony PCR were further grown in LSLB for plasmid extraction to check clones using restriction-digestion and confirmation by sequencing.

The transformation of *fibril* gene cloned in pPICZα vector was done as per protocol provided by Invitrogen (https://assets.thermofisher.com/TFS-Assets/LSG/manuals/ppiczalpha_man.pdf). Briefly, a positive clone was linearized by restriction digestion with PmeI (New England Biolabs) followed by PCR clean-up. The linearized clone was then transformed into freshly prepared electro-competent *P. pastoris* GS115 cells (genotype: *his4*) by electroporation. Electroporation conditions used were- 2000 V voltage, 25 µF capacitance, 200 Ω resistance with a 2 mm cuvette. The electroporated cells were recovered by addition of 1 mL sterile sorbitol (1 M) and incubating the mixture at 30 ^{0} C for 2 hours without shaking. 200 µL of the recovered cells were spread onto YPD agar [Yeast extract (1 % w/v), peptone (2 % w/v), D- glucose (2 % w/v) and agar (2 % w/v)] plates supplemented with Zeocin™ and incubated for 3 days at 30 ^{0} C. One colony was then inoculated into sterile YPD broth for preparation of glycerol stocks of the clone. Glycerol stocks were prepared by growing the 100 mL culture to mid-log phase. The cultures was then pelleted and re-suspended in 5 mL YPD media. 5 mL of sterile glycerol (100 %) was added to the re-suspended culture aseptically. The culture was mixed so as to obtain uniform solution which was aliquoted as 0.5 mL aliquots in sterile 1.5 mL tubes and flash frozen in liquid nitrogen. The frozen aliquots were stored at -80 ^{0} C until further use for expression and purification of Fibril from *P. pastoris*.

2.2.2) Expression

2.2.2.1) Expression of full-length *fibril* in *E. coli*

Upon confirmation of the positive clones by sequencing, the clones were used for expressing protein of our interest. All the clones prepared in pHIS17 were expressed in BL21(AI) [genotype: *F-ompT hsdSB (rB- mB-) gal dcm araB::T7RNAP-tetA*] strain of *E. coli*. While pHIS17 plasmid contains T7 promoter for expressing gene of interest, the gene coding for T7 polymerase is present on the chromosome of BL21(AI) cells. Expression of T7 polymerase in

this strain is tightly controlled by the *araBAD* operon that gets triggered only in presence of L-arabinose in the media. Since pHIS17 plasmid contains the ampicillin resistance marker gene, ampicillin was supplemented in the media used for growing transformants at all stages of expression so as to provide selection pressure for the cells to maintain the transformed plasmid.

The plasmid was transformed into BL21(AI) cells by heat-shock method (standard protocol). The cells were then spread on sterile LB agar plates supplemented with ampicillin (100 µg/ mL) and incubated for 10-12 hours at 37 ^{0} C. The parameters such as temperature, optical density (OD) for induction and post-induction incubation time were optimized to get maximum protein in the soluble fraction. For expression check, the 10 mL culture pellets were re-suspended in appropriate lysis buffer and lysed by sonication. The solubility was checked by spinning the lysate at 21,000 xg/4 ^{0} C/15 minutes. A fraction of total lysate and supernatant at 21,000 xg for uninduced and induced culture were checked on 12 % SDS-PAGE gel. The presence of the band of interest in the supernatant was indicative of soluble protein.

The standardized protocol for large scale expression is as follows. Five to eight colonies from the agar plates were inoculated in 250 mL sterile LB broth supplemented with ampicillin (100 µg/ mL) and incubated at 37 ^{0} C for 10-12 hours under shaking condition, which acted as the primary culture. 10 mL from the primary culture that attained OD (measured at 600 nm; OD_{600}) = 0.6 was transferred to a sterile 100 mL flask without addition of L-arabinose. This acted as uninduced control for confirmation of protein overexpression in cultures induced by addition of L-arabinose. From rest of the primary culture, 10 mL each was then inoculated into six 2 litre flasks containing 1 litre sterile LB broth (supplemented with 100 µg/ mL ampicillin). These cultures were then induced by addition of sterile L-arabinose solution (final concentration = 0.02 grams/ 100 mL). The induced as well as uninduced cultures were incubated under shaking condition at 25 ^{0} C for 6 hours. 10 mL each of the induced and uninduced cultures were pelleted separately and used for expression check. Rest of the induced culture pellet was washed with $T_{10}E_1$ buffer (10 mM Tris pH 7.6, 1 mM EDTA) and flash frozen in liquid nitrogen as 200 mL culture equivalent aliquots. The aliquots were stored at -80 ^{0} C until further use.

2.2.2.2) Expression of full-length *fibril* in *P. pastoris*

Heterologous expression of proteins in *P. pastoris* using pPICZα vector results into production of recombinant proteins with an N-terminal α-factor signal sequence followed by Kex2 protease site. C-myc epitope and His_6 tag are linked at the C-terminus of recombinant protein. The codons for α-factor signal sequence in the vector are incorporated from *Saccharomyces cerevisiae,* where it is necessary to secrete mating proteins. In cells, upon expression of the protein tagged with α-factor signal sequence, the two are separated by action of Kex2 protease in two steps. In the first step, the sequence EKREAEA is cleaved between R and E by Kex2 protease. In the second step the EA repeats are further cleaved by STE13 protease. The mature protein is then secreted into the growth medium (pPICZα manual by Invitrogen; https://assets.thermofisher.com/TFS-Assets/LSG/manuals/ppiczalpha_man.pdf). Thus, when a heterologous protein is expressed with α-factor signal sequence, the former is obtained in the growth media.

For expression of *fibril* in *P. pastoris*, a glycerol stock aliquot of *Pichia pastoris* GS115 cells with *fibril* (in pPICZα) clone integrated into genomic DNA was thawed on ice. The culture was inoculated initially into 50 mL of sterile BMMY media [yeast extract (1 % w/v), peptone (2 %), phosphate buffer (100 mM; pH 6.0), yeast nitrogen base (1.34 %), biotin (4×10^{-5} %), methanol (0.5 % v/v)] and incubated at 30 ^{0} C under shaking conditions for 24 hours. The culture was then aseptically transferred into a sterile 2 litre flask containing 450 mL BMMY media and incubated at 30 ^{0} C under shaking conditions for 72 hours. 1 mL methanol was added to the growing culture after every 24 hours to ensure supply of carbon source and keep cells induced for Fibril production.

2.2.2.3) Growing *Spiroplasma* cultures

Spiroplasma citri GII-3 culture from Dr. Laure Béven's lab at INRA, Bordeaux, France were used for purification of Fibril. This part of the work was carried out in Dr. Laure Béven's lab at INRA, Bordeaux, France. The protocol followed for growth of *Spiroplasma* cultures is as follows.

Modified SP4 media composition and preparation protocol

The modified SP4 media (Duret et al., 2005) used for growing *Spiroplasma* cultures was prepared in two parts: the first contains components that were sterilized by autoclaving (Table 2.2). The second part contains compounds that cannot withstand autoclaving and hence needed to be filter sterilized (Table 2.3).

Table 2.2 SP4 media base: to be autoclaved

Component	Quantity/ 300 mL SP4 media
Mycoplasma base broth	1 g
Peptone	1.6 g
Tryptone	3 g
Phenol red (9 mg phenol red/ mL 1 M NaOH)	1.2 mL
Adjust pH to 7.8	
MilliQ to make up the volume*	205 mL

* In case of preparation of modified SP4 agar, add 3 g of Noble agar/205 mL base after adjusting pH.

Table 2.3 Components to be filter sterilized and added to sterile SP4 media base**

Component	Stock concentration	Quantity/ 300 mL SP4 media
Fetal calf serum (FBS)***	-	50 mL
CMRL 1066	-	15 mL
Yeastolate	4 % (w/v)	25.5 mL
D-Glucose	50 % (w/v)	3 mL
Penicillin G	200,000 U/mL	1.5 mL

** To be filter sterilized aseptically initially using 0.45 µm filter initially and then through 0.22 µm filter.

*** To be thawed and then heat inactivated at 60 ^{0} C for 30-60 minutes in water-bath with shaking after every 5-10 minutes.

For preparation of 300 mL modified SP4 media, the base media components were added to 150 mL MilliQ and solubilized. The pH of mixture was adjusted to 7.8 and volume made to 205 mL by addition of MilliQ. This base mixture was sterilized by autoclaving at 121 ^{0} C at 15 lbs pressure for 15 minutes. The components to be filter sterilized were prepared as follows- FBS was allowed to thaw at room temperature. Upon thawing, FBS was heat inactivated in a water-bath by incubation at 60 ^{0} C for 30-60 minutes with shaking after every 5-10 minutes. The desired volume of heat inactivated FBS was mixed with other components to be filter sterilized,

in an aseptic environment. All the components were first passed through a 0.45 µm filter initially to get rid of bulk matter and any precipitated material. It was then passed through 0.22 µm filter into autoclaved, cooled base media and stored at room temperature until use. The modified SP4 media prepared by this protocol can be used for up to 60 days when stored at room temperature.

Frozen 0.5 mL culture of *S. citri* GII-3 cells was thawed on ice and aseptically added to 2 mL of fresh modified SP4 medium (composition mentioned above) in a sterile 5 mL screw capped glass tube. The culture was then incubated at 32 ^{0} C for about 18 - 20 hours without shaking. 0.5 ml of the growing culture was then re-inoculated into sterile 2 mL modified SP4 media and incubated at 32 ^{0} C for about 18 - 20 hours. Upon incubation, the 2.5 mL grown culture was added to 20 mL of sterile modified SP4 medium incubated at 32 ^{0} C for about 18 - 20 hours. The 22.5 mL culture was then inoculated into200 mL of sterile modified SP4 medium incubated at 32 ^{0} C for about 18 - 20 hours without shaking. Finally the 222.5 mL culture was added to 777.5 mL sterile modified SP4 medium and incubated at 32 ^{0} C for about 12 - 15 hours.

At each stage, the growth of cells was monitored by change in color of sterile modified SP4 media from red to yellow and confirmed by microscopic observation of cells and cell density from the growing culture. Well grown cells of *S. citri* GII-3 exhibited helical morphology, kinking motility and were small to medium in length (2 to 5 µm). Care was taken not to shake the cultures since aeration is not necessary for growing *Spiroplasma*.

Cultures were pelleted using JA14 rotor in a Beckman coulter high speed centrifuge. The first spin was performed to pellet down cells from the medium by spinning at 14,000 xg/18 ^{0} C/30 minutes. The supernatant was discarded and pellet washed thrice, or until the media was removed completely, by cycles of pellet re-suspension and centrifugation in cold condition. For washing the pellet, each pellet was re-suspended on ice initially in ice-cold 5 mL HS buffer (8 mM HEPES pH 7.4, 280 mM sucrose) and the volume made to 50 mL by addition of ice-cold HS buffer. The cell suspension was spun at 14,000 xg/18 ^{0} C/30 minutes, supernatant discarded and the washing cycle repeated. The completely washed pellets were transferred to sterile 2 mL microfuge tubes and stored at -80 ^{0} C until use.

2.2.3) Protocols for Fibril purification

Fibril purification has been attempted using two heterologous expression systems, *E. coli* and *P. pastoris*. According to published reports, Fibril has always been purified from *Spiroplasma*, a cell wall deficient organism. The protocol established for Fibril purification from *Spiroplasma* was modified for purification from *E. coli* and *P. pastoris*. Detergents were screened for their ability to solubilize Fibril to enable their purification. Based on the insights obtained from detergent screen, multiple approaches such as affinity chromatography, hydrophobic interaction chromatography and density gradient centrifugation were used for obtaining purified Fibril. The details of protocols and various modifications are described below.

2.2.3.1) Principle of density gradient centrifugation

Density gradient centrifugation (DGC) is a process of separation of molecules on the basis of their densities using a decreasing/increasing gradient of solution with higher density than the molecules to be separated. It is often used for purification of viruses, lipoproteins and filamentous proteins from other biomolecules (Chung et al., 1980; Redgrave et al., 1975; Reimer et al., 1967; Soellner et al., 1985). DGC is further classified as (i) Rate zonal centrifugation, which separates proteins based on the differences in their sizes and mass and is used for separation of proteins with similar density (ii) Isopycnic or buoyant density gradient centrifugation in which the separation is based on the differences in density of the proteins (Rødahl, 1998). The gradients can be prepared as continuous (linear) or discontinuous (step) gradient. The difference between the two is the presence of sharp differences in the density for a discontinuous gradient while in a continuous gradient, the density decreases in a linear fashion. A discontinuous gradient is not preferred for the rate zonal separations as sudden changes in the gradient viscosity and density leads to irregular separation of particles. A rate zonal centrifugation involves centrifugation of sample for shorter duration with shallow gradient, and is preferred when the protein is labile, osmotically sensitive. However, isopycnic runs allow for separation of large amounts of sample and leads to a separation of a wide variety of samples according to their density, as they can have steep gradients (Patsch and Patsch, 1986). The material to be used for preparation of the gradient depends on the density of the sample to be separated. Table 2.4 lists compounds commonly used for density gradient preparation and their densities.

Table 2.4 List of compounds used for density gradient preparation.

Compounds (concentration)	Density
Glycerol (100% v/v)	1.26 g/cm^3
Sucrose (100% w/v)	2.0 g/cm^3
Urografin (76% w/v)	1.52 g/cm^3
Caesium chloride (6M)	1.83 g/cm^3
Percoll	1.0 - 1.3 g/cm^3
Nycodenz (80% w/v)	1.426 g/cm^3

Since Fibril have been purified using isopycnic density gradient centrifugation (Williamson et al., 1991), I used linear density gradients of glycerol, urografin or Nycodenz for its purification.

2.2.3.2) Purification of Fibril using glycerol gradients

A 200 mL culture pellet of *E. coli* cells expressing *fibril* was washed with $T_{10}E_{10}$ [10 mM Tris (pH 7.6), 10 mM EDTA] buffer. The washed pellet was re-suspended in 100 mL of lysis buffer 1 (Appendix 1). Cells were lysed on ice for 3 minutes with 60 % amplitude and 5 seconds ON, 5 seconds OFF cycle using a VibraCell™ probe sonicator. The lysate was spun at 4600 xg/4 ^{0}C/15 minutes and cell debris in the form of pellet was discarded. Filamentous proteins can be pelleted from a solution by spinning at 100,000 xg (Mizuno, 1992; Woodring et al., 2001). Hence, supernatant of the 4600 xg spin was further centrifuged at 100,000 xg/4 ^{0}C/2 hours to pellet down Fibril filaments. Simultaneously, a 34 mL linear glycerol gradient was prepared using 17 mL each of 10 % and 50 % glycerol in $T_{10}E_{10}$ buffer to separate proteins from 100,000 xg pellet based on their density. At the end of the 100,000 xg spin, the pellet was re-suspended in 1.5 mL $T_{10}E_{10}$ buffer and loaded on top of the glycerol gradient. Sample loaded gradients were then spun at 79,500 xg/4 ^{0}C/2 hours to attain protein separation. The gradient was then fractionated using peristaltic pump assisted by a glass capillary. The glass capillary was dipped into bottom of the gradient and fractions collected (1.5 mL each). The fractions were sequentially labeled with the densest fractions (from the bottom) being 1 and the topmost was the last. Samples from all the protein fractions were electrophoresed and visualized on 12 % SDS-PAGE gel along with marker proteins to identify protein of interest. Samples containing protein of interest were pooled into one tube.

The samples are placed in a chamber under vacuum for visualization using an electron microscope (EM). Thus, the samples to be imaged need to be dried before processing in SEM. The protein purified using glycerol gradient has salts and glycerol associated with it. These compounds form a thick layer on silicon wafer when the samples are dried. Visualization of sample in presence of such layers becomes difficult. Thus, pooled protein was dialyzed against water before preparation of samples for visualization using Field Emission-Scanning Electron Microscope (FE-SEM).

2.2.3.3) Buffer standardization for solubility of Fibril protein

Six pellets of 200 mL culture were re-suspended in 45 mL of lysis buffer 2 (Appendix 1). Cells were lysed by sonication on ice for 18 minutes with 60 % amplitude and 1 second ON, 3 seconds OFF cycle using a probe sonicator. Any one of the following compounds were added to each of the lysate- N-Lauroyl sarcosine sodium salt (LSS; 0.5 % w/v), Sodium deoxycholate (1 % w/v), Tween-20 (1 % v/v), sodium dodecyl sulphate (SDS; 1 % w/v), EDTA (10 mM w/v), lysis buffer 2 (buffer base). Values in bracket indicate final concentration of respective compounds. Volume in each tube was made to 50 mL by addition of lysis buffer 2. The lysates were retained on ice for 30 minutes so as to allow the compounds to act. Post-incubation, all samples were initially spun at 4629 xg to remove cell debris as pellet. The supernatant was further spun at 159,000 xg/4 ^{0} C for 2 hours. The supernatants were transferred to fresh tubes and each pellet was re-suspended in 2 mL $T_{10}E_{10}$ buffer. 10 µL sample from each condition at all the stages was mixed with 10 µL SDS-PAGE loading dye (2X) and incubated at 99 ^{0} C for 10 minutes under shaking condition. For visualization of samples on SDS-PAGE gel, 7 µL each was loaded for total lysate and 159,000 xg. For 4629 xg supernatant and 159,000 xg supernatant, 15 µL each of the samples were loaded on the gel.

2.2.3.4) Purification of Fibril expressed in *E. coli* by Ni-NTA affinity chromatography

A 200 mL culture pellet of *E. coli* cells expressing *fibril* with an N-terminal His_6 tag was re-suspended in 50 mL lysis buffer 3 (Appendix 1). Cells were lysed on ice for 5 minutes by sonication with 60 % amplitude and 1 second ON, 3 seconds OFF cycle using a VibraCell™ probe sonicator. The lysate was spun at 4629 xg /4 ^{0}C/15 minutes to remove cell debris as pellet. Supernatant was further spun at 159,000 xg/4 ^{0}C/ 30 minutes to pellet down Fibril filaments. The

pellet was re-suspended using 1.8 mL binding buffer/ buffer A_{200} (50 mM Tris pH 7.6, 200 mM NaCl). 0.2 mL of 10 % SDS solution was added and mixed with the re-suspended pellet. The protein suspension was spun at 21,000 xg/25 ^{0}C/10 minutes to remove any large insoluble aggregates. The soluble fraction from this step was spun at 159,000 xg/25 ^{0}C/30 minutes and Fibril was obtained in the supernatant. All the steps after addition of SDS were performed at room temperature/ 25 ^{0} C since SDS precipitates at lower temperatures (4 ^{0} C).

The supernatant was mixed with 2 mL Ni-NTA resin (Ni-NTA agarose; Qiagen) pre-equilibrated with binding buffer/buffer A_{200} (50 mM Tris pH 7.6, 200 mM NaCl). Protein was allowed to bind to Ni-NTA resin by mixing on a rotomix (Tarsons) at room temperature for 30 minutes. Post incubation, the mixture was put into an empty column (GE Healthcare Life Sciences) with filter. The filter retained Ni-NTA resins and bound protein, allowing unbound proteins to flowthrough, which were collected as the unbound fraction. The Ni-NTA resin was washed by passing more binding buffer to remove unbound proteins and collected as wash. Elution buffer containing increasing concentrations (25 mM, 50 mM, 100 mM, 250 mM and 500 mM) of imidazole in binding buffer was passed through the resin in steps and each collected as a separate fraction. 10 µL from each of these fractions were added to 10 µL 2X SDS-PAGE loading buffer [Tris (100 mM; pH 6.8), SDS (4 % w/v), bromophenol blue (0.2 % w/v), dithiothreitol (200 mM) and glycerol (20 % v/v)] and incubated at 95 ^{0} C for 10 minutes with shaking at 300 revolution per minute (rpm). Samples were briefly spun and 10 µL loaded into each well of 12 % SDS-PAGE gel along with a protein marker. Resolving samples on SDS-PAGE followed by staining and de-staining allowed us to identify fractions containing protein of interest and purity of the protein. Fractions containing protein of interest were pooled and extensively dialyzed against distilled water to remove salts, detergents etc. Dialyzed protein was used for preparation of samples for visualization of filaments using Field Emission-Scanning Electron Microscopy (FE-SEM) or Transmission Electron Microscope (TEM) using protocols as described under 2.2.4.1) and 2.2.4.3) respectively.

2.2.3.5) Purification of Fibril expressed in *E. coli* using ion-exchange chromatography

Fibril was obtained in the soluble fraction by re-suspending 4 × 200 mL culture pellets of *E. coli* expressing *fibril* in 100 mL lysis buffer 4 (Appendix 1). The cells were lysed by sonication on ice. Lysate was spun at 100,000 xg/4 ^{0} C/30 minutes to remove cell debris and unwanted

proteins. The supernatant containing Fibril was used for ammonium sulphate precipitation. 22.6 g ammonium sulphate (22.6 g/100 mL; 40 % saturation) was added to the supernatant in steps over 1 hour at 4 ^{0} C and stirring condition. Upon completion of ammonium sulphate addition, the mixture was allowed to stir at 4 ^{0} C for 30 minutes. The turbid solution was then spun at 100,000 xg/4 ^{0} C/30 minutes. The supernatant was discarded and pellet containing Fibril was re-suspended in 50 mL sodium acetate buffer (100 mM; pH 4.63). The suspension was spun at 28,000 xg/4 ^{0} C/30 minutes to dilute out traces of ammonium sulphate. Fibril was obtained in pellet which was further re-suspended in 48 mL 50 mM Tris (pH 8.0). The suspension was spun at 28,000 xg to remove any insoluble aggregates as pellet. To the supernatant 4.62 grams of ammonium sulphate was added to attain final concentration of 700 mM. The mixture was again spun at 28,000 xg/4 ^{0} C/30 minutes to remove aggregates as pellet. The clear solution containing Fibril was transferred to a fresh tube and loaded onto HIC columns [GE Healthcare Life Science; in the order: Butyl HP, Phenyl (low substitution), Phenyl (high substitution) and Phenyl HP] connected in series and pre-equilibrated with binding buffer [50 mM Tris pH 8.0, 700 mM ammonium sulphate] in an Äkta cold cabinet (4 ^{0} C). The unbound protein was collected as flowthrough. Upon complete loading of the partially purified lysate, all the columns were disconnected from the system. The columns were individually connected to the system, washed with binding buffer and eluted with a linear gradient of decreasing concentrations of ammonium sulphate (binding buffer mixed with 50 mM Tris pH 8.0).

To attempt purification of Fibril using Capto MMC (multimodal cation exchanger) matrix the cultures were processed as follows. 4 × 200 mL culture pellets were re-suspended in lysis buffer 4. The cells were lysed by sonication on ice using a probe sonicator. Lysate was spun at 100,000 xg/4 ^{0} C/30 minutes to remove cell debris and unwanted proteins. The supernatant containing Fibril was used for ammonium sulphate precipitation. 22.6 g ammonium sulfate precipitation (40 % saturation) was added to the supernatant in steps over 1 hour at 4 ^{0} C and stirring condition. Upon complete addition of ammonium sulphate the mixture was allowed to stir at 4 ^{0} C for 80 minutes. The turbid solution was then spun at 100,000 xg/4 ^{0} C/30 minutes. Supernatant was discarded and pellet containing Fibril was re-suspended in 50 mL sodium acetate buffer (100 mM; pH 4.6). The suspension was spun at 28,000 xg/4 ^{0} C/30 minutes to dilute out traces of ammonium sulphate. Fibril was obtained in pellet which was further re-suspended in 24 mL of 50 mM Tris (pH 7.6). The suspension was spun at 27,200 xg/4 ^{0} C/30 minutes to remove any

insoluble aggregates as pellet. The supernatant was loaded onto Capto MMC column pre-equilibrated with 50 mM Tris (pH 7.6) using Äkta Prime in a cold cabinet (4 ^{0} C). Unbound protein was collected as flowthrough. The bound protein was washed with excess of binding buffer [50 mM Tris (pH 7.6)] and collected as wash. The bound protein was eluted using a 0 % to 100 % concentration of elution buffer [50 mM Tris (pH 9.83)] against binding buffer and collected as fractions.

30 µL sample from all the stages of purification and fractions was mixed with 30 µL SDS-PAGE loading dye (2X) and heated at 99 ^{0} C/10 minutes under shaking condition. All the samples were briefly spun and 10 µL loaded on 12 % SDS-PAGE gels.

2.2.3.6) Purification of Fibril expressed in *P. pastoris* by Ni-NTA affinity chromatography

200 mL culture pellet of *P. pastoris* cells expressing Fibril was lysed by cycles of freeze thawing and crushing using a mortar pestle. Liquid nitrogen was taken into mortar and droplets of culture pellet re-suspended in 1 mL lysis buffer 3 (Appendix 1) were dropped into nitrogen. The pellets were crushed using pestle until liquid nitrogen in the mortar evaporated. Fresh nitrogen was then added to mortar and crushing continued. Three cycles of liquid nitrogen addition and crushing were necessary for complete lysis of *P. pastoris* cells. Upon completion of lysis, the lysate was transferred into a clean 50 mL falcon tube (Tarsons) and volume made up to 50 mL by addition of lysis buffer 3 (Appendix 1). The mixture was vortexed to obtain uniform solution. The lysate was then transferred into a clean Ti-45 tube (Beckman Coulter) and spun at 100,000 xg/4 ^{0} C/30 minutes to get rid of cell debris, unlysed cells and unwanted proteins. The supernatant containing Fibril was incubated with pre-equilibrated Ni-NTA resins for 30 minutes at room temperature with brief shaking. The mixture was then put into an empty column and purification performed as mentioned in section 2.2.3.4) of this chapter.

2.2.3.7) Purification of Fibril from *Spiroplasma*

A 2 litre culture pellet of *Spiroplasma citri* GII-3 cells was rinsed with washing buffer (10 mM Tris pH 7.6, 400 mM NaCl). The pellet was then re-suspended in 5 mL lysis buffer 1 (Appendix 1). Mg^{2+}, DNase and protease inhibitor cocktail (Sigma) were added at final concentrations of 10 mM, 50 µg/mL and twice the company-recommended value respectively. The cell suspension was initially stirred at room temperature for 5 hours and then at 4 ^{0} C for 24 hours to facilitate

cell lysis. The lysate was then spun at 100,000 xg/4 ^{0} C/2 hours to separate the Triton X-100 insoluble fraction as pellet. Supernatant obtained after the spin was discarded and pellet rinsed using 200 µL TE buffer (10 mM Tris pH 7.6, 10 mM EDTA). The pellet was re-suspended in 500 µL TEK buffer (10 mM Tris pH 7.6, 10 mM EDTA, 1 M KCl) and stirred at 4 ^{0} C for a week to facilitate separation of individual filaments from bundles. Following a week long stirring, 250 µL of the enriched Fibril were loaded on top of cold 1 mL Nycodenz density step gradient prepared using 250 µL each of 0 %, 20 %, 30 % and 50 % (w/v) Nycodenz concentration in TE buffer (10 mM Tris pH 7.6, 10 mM EDTA) and stirred at 100,000 xg/20 ^{0} C/1 hour to separate proteins based on their densities. At the end of the spin, 5 different layers were seen in the tubes. Each individual layer was separated and transferred to fresh tubes. 10 µL sample from each layer was mixed in separate tubes with 10 µL SDS-PAGE loading dye (2X) and incubated at 95 ^{0} C for 3 minutes. All the samples were spun briefly and 10 µL each loaded onto a 12 % SDS-PAGE gel to check for fraction containing Fibril. Thus, layer 4 from gradient was re-suspended in 1 mL TE buffer (10 mM Tris pH 7.6, 10 mM EDTA) and spun at 100,000 xg/4 ^{0} C/2 hours to remove Nycodenz as supernatant. The pellet was re-suspended in 250 µL TE (10 mM Tris pH 7.6, 10 mM EDTA) buffer and stirred at 4 ^{0} C until visualization using electron microscope.

2.2.3.8) Preparation of urografin density gradient

Urografin solution (76% w/v; Cadila healthcare Ltd, Kundaim, India) was obtained. To prepare a 30.4 % solution, 10 mL of urografin was added to 15 mL of $T_{10}E_{10}$ buffer. 4.5 mL of 76 % urografin was put into a tube compatible with SW-41 rotor. 4.5 mL of 30.4 % urografin solution was layered on top of the 76 % solution in tube and the tube was left at 4 ^{0} C overnight (16 hours). The overnight incubation allows the formation of a linear gradient. The gradient was used further for separation of protein filaments.

Alternatively, a linear gradient was prepared using a density gradient mixer (Sigma-Aldrich). For preparation of a 9 mL gradient, 4.5 mL each of 76 % and 30.4 % urografin solutions were put into the two wells of the gradient mixture such that the well containing dense solution is closer to outlet. The gradient mixer was kept on top of a magnetic stirrer and a magnetic bead put into the well containing dense solution. Outlet tubing of the mixer was put into a fresh thin, clear tube (13 mL) compatible with SW-41 rotor. The valve connecting both the wells as well as the outlet

valve were opened while magnetic bead was stirring. The stirring facilitated gradual mixing of the two solutions with different densities. The solution exiting the outlet tubing had decreasing density with time thus forming a linear density gradient in the tube. The gradient formed by this method was immediately used for separation of protein mixture based on their densities.

2.2.3.9) Purification of Fibril from *E. coli* by separation on urografin density gradient

A 200 mL culture pellet of *E. coli* BL21(AI) cells expressing full length *fibril* with an N-terminal 6xHis tag or without a tag was re-suspended in 50 mL of any one of the lysis buffers from 1 or 5. Cells were lysed on ice by sonication for 20 minutes with 60 % amplitude and 5 seconds ON, 5 seconds OFF cycle using a VibraCell™ probe sonicator. The lysate was spun at 8,000 xg/4 ^{0}C/10 minutes and cell debris in the form of pellet was discarded. The supernatant so obtained was spun at 159,000 xg to separate Fibril filaments from the non-polymerized proteins. The pellet so obtained contained enriched Fibril filaments. Therefore the pellet was re-suspended in 1 mL $T_{10}E_{10}$ buffer and left for stirring on a magnetic stirrer at 4 ^{0} C for at least 10 days. 1 mL of stirred enriched Fibril was loaded on top of 9 mL urografin density gradient. The tubes containing gradient loaded with protein were spun using a swinging bucket rotor at 154,000 xg/4 ^{0}C/2 hours. The gradient was manually fractionated using a pipette into 10 fractions with 1 mL volume per fraction. The fractions were sequentially labeled with the topmost fractions being 1 and the densest was 10. Samples from all the protein fractions were electrophoresed and visualized on 12 % SDS-PAGE gel along with marker proteins to identify protein of interest.

To obtain urografin-free protein, fractions containing protein of interest were pooled and volume made to 60 mL by addition of 50 mM Tris (pH 7.6). The solution was spun at 159,000 xg/4 ^{0}C/30 minutes and Fibril obtained in the pellet. The pellet was re-suspended in 1 mL $T_{10}E_{10}$ buffer [50 mM Tris (pH 7.6), 10 mM EDTA] and used for sample preparation for TEM studies.

2.2.3.10) Purification of Fibril from *E. coli* without detergent, using urografin density gradient

A 200 mL culture pellet of *E. coli* BL21(AI) cells expressing untagged full length *fibril* was re-suspended in 50 mL of 10 mM Tris (pH 7.6). Cells were lysed on ice by sonication for 20 minutes with 60 % amplitude and 5 seconds ON, 5 seconds OFF cycle using a VibraCell™ probe sonicator. The lysate was spun at 8,000 xg/4 ^{0}C/10 minutes and cell debris in the form of

pellet was discarded. The supernatant so obtained was spun at 30,000 xg to get rid of other macromolecules as a pellet. The supernatant was further spun at 159,000 xg at 4 ^{0}C for 30 minutes to separate Fibril filaments from the non-polymeric cytosolic components that remain in the supernatant. The pellet so obtained contained Fibril filaments. Therefore the pellet was re-suspended in 1 mL $T_{10}E_{10}$ buffer and left for stirring on a magnetic stirrer at 4 ^{0} C for at least 10 days.

Following prolonged stirring the re-suspended pellet was loaded on top of the 9 mL urografin gradient (40 % to 100 %) and spun at 159,000 xg/ 4 ^{0}C / 120 minutes. The gradient facilitated separation of proteins based on their density. Upon spinning for 120 minutes, the gradient was fractionated using pipette into 10 equal fractions. The fractions were labeled 1 to 10 such that the topmost fraction was 1 and last was 10. Samples from all the fractions were visualized on 12 % SDS-PAGE gel. Fractions containing protein of interest were pooled and volume made to 60 mL by addition of 50 mM Tris (pH 7.6). The diluted protein was spun at 159,000 xg/4 ^{0}C/30 minutes to remove urografin in the supernatant and obtain enriched Fibril in the pellet. The pellet was re-suspended in 1 mL of 50 mM Tris (pH 7.6). The filaments were again washed by spinning at 21,000 xg/4 ^{0}C/20 minutes. The pellet so obtained was re-suspended into 1.5 mL of chloroform and methanol (2:1) mixture and vortexed for 2 minutes. It was then spun at 21,000 xg/4 °C/20 minutes and protein obtained as pellet. The supernatant was discarded and washed with chloroform + methanol mixture. The pellet so obtained after second round of washing with chloroform + methanol mixture was allowed to air dry in a dust-free environment for 30 minutes to allow evaporation of organic solvents. Upon drying, the pellet was re-suspended in 50 µL of 10 mM Tris (pH 7.6) and used for grid preparation for visualization of filaments using TEM.

2.2.4) Visualization of Fibril filaments

2.2.4.1) Principle of electron microscopy (EM)

An electron microscope (EM) is an instrument used for visualization of objects with dimension smaller than a few microns. Unlike light microscope, an EM uses electron (e^-) beam to image objects. An electron beam has wavelength about 100,000 times shorter than that of visible light and hence an EM can resolve the structures of much smaller objects with high intricate details as compared to a light microscope. Since electrons can travel only in vacuum, the samples are imaged under vacuum.

When an accelerated, focused electron beam hits the sample, the electrons can pass through or interact with the sample. The electrons interacting with the sample can either be scattered backwards (backscattered; original, high energy e^-) or interact with the sample and result into production of augur e^- (released from the atom due to its replacement by e^- from the beam), secondary (low energy) e^- , elastically (no loss of energy) or inelastically (energy is lost) scattered e^- or even production of X-rays. Depending on the resultant e^- detected for image formation, EMs are classified as- i) Transmission electron microscope (TEM) and ii) Scanning electron microscope (SEM). While TEM images are formed due to the electrons that have passed through the sample, SEM images are formed by the electrons that are scattered from the sample.

The energy of beam used in a SEM is between 0.2 and 40 keV (kilo electron volts) and the resolution it can attain is 20 nm or better. A scanning electron microscope is equipped with one or more detectors that can each scan secondary electrons (SE), X-rays and backscattered electrons (BSE). Most commonly used detector in a SEM detects secondary electrons which are generated by inelastic interactions between the beam and sample. Since the secondary electrons are generated from a few nanometers below the surface of the sample, they provide information regarding sample surface topology. This information is then used for image generation.

In contrast to SEM, a TEM has the detector placed below the sample and detects electrons that have passed through the sample with or without scattering. The electrons passing through the sample are focused by electromagnetic coils to form image. Depending on the electron density in the sample, the electrons are deviated from their original path to varying degrees. Thus, electron-rich areas of the sample appear dark compared to regions of sample with light atoms.

The voltage used in TEMs is in the range of 100-300 keV, much higher than used in SEM. While there are no limitations of the thickness of samples imaged in a SEM, the ideal thickness for imaging using a TEM is <100 nm.

All biological samples have C, O, N, H as the building blocks. Since all these elements are not electron dense, the contrast that can be obtained in images of biological samples using SEM and TEM is very less. Thus, electron dense material is usually layered on the sample to generate contrast. For imaging using SEM, metals such as gold or platinum are coated on the sample for improving contrast. Samples for TEM are ‘negatively stained’ with heavy metal solutions such

as uranyl acetate, phosphotungstic acid, ammonium molybdate and others. The negative staining leaves the heavy metal in the areas not occupied by sample. The heavy metal strongly scatters electrons which interact with those electrons scattered by sample and generate contrast in the image.

Electron cryomicroscopy (Cryo EM) is a modification of TEM, where the samples frozen in vitreous ice are imaged under cooling conditions. Imaging of samples using CryoEM is often used for structure determination of protein(s) of interest. To obtain high resolution structure of a protein using cryoEM, it is necessary to obtain images of the protein in all possible orientations. These images are then used for 3D reconstruction of structure (Chung and Jung, 2016; DeRosier and Klug, 1968). When using a monomeric/short-oligomeric form of protein, the molecule is free to take up all possible orientations on an electron microscopy grid. In a protein polymer the freedom of orientation may be limited by the contacts with neighboring subunits, making it unsuitable for structure determination by cryoEM. This problem is overcome in a filamentous protein with inherent helical symmetry; the helicity facilitates diverse orientations of the protein molecules within filament (Desfosses et al., 2014; Egelman, 2015). The flowchart of sample preparation, data collection and processing for structure determination using cryoEM is shown in Figure 2.1.

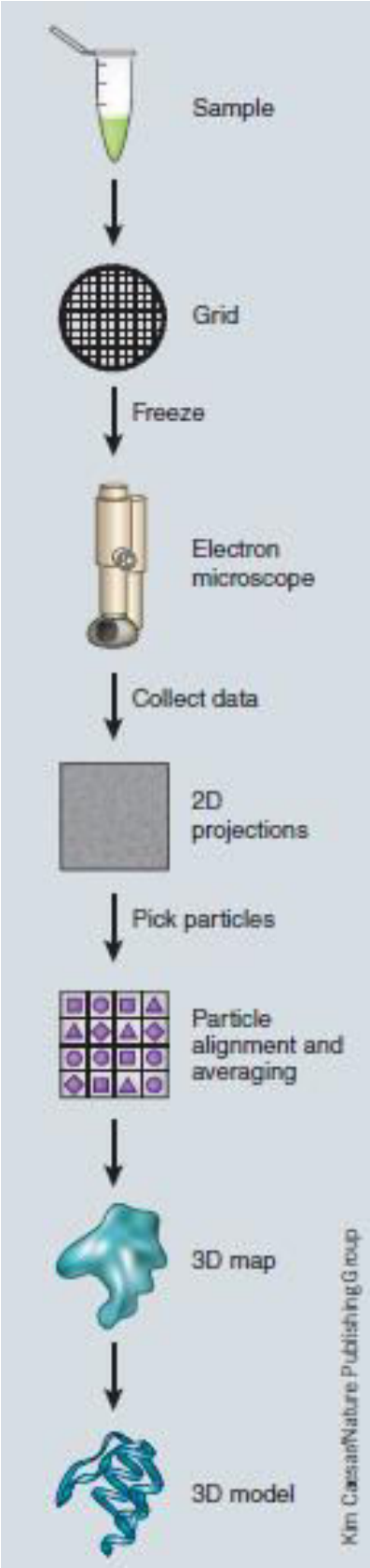

Figure 2.1 Work flow for structure determination by cryoEM. The purified samples are applied to and electron microscopy grid and plunge frozen to embed sample in vitreous ice. The grids containing samples are then visualized using a cryoEM. The cryoEM images are nothing but 2-dimensional (2D) projections of protein of interest in various orientations. The particles are selected from the images, aligned and used for averaging to obtain 2D averages. The 2D averages are then used for 3D reconstruction to obtain a 3D map of the protein. The high resolution 3D maps are useful for building 3D models of the protein. Image adapted from (Doerr, 2016).

2.2.4.2) Sample preparation for Field Emission-Scanning Electron Microscopy (FE-SEM)

A silicon wafer disc (Sigma-Aldrich) was cut into smaller pieces of 5 mm X 5 mm dimensions using a diamond cutter. These pieces were thoroughly washed with 100 % ethanol and stored in 70 % ethanol at room temperature until use.

For sample preparation, the silicon wafer pieces were removed from 70 % ethanol and allowed to air dry at room temperature to allow evaporation of residual ethanol and water mixture. 2-3 µL of the dialyzed protein was drop casted on dried silicon wafer pieces and allowed to dry in dust-free environment at 37 ^{0} C. 15 nm thick gold coating was then done on the silicon wafer pieces with completely dried protein samples, using sputter coater (Quorum Tech) before visualizing the samples under FE-SEM (ZEISS).

2.2.4.3) Sample preparation for Transmission Electron Microscopy (TEM)

Carbon-Formvar coated copper grids were purchased from TED-PELLA, INC and glow-discharged in plasma cleaner (Quorum technologies) just before use. Alternatively, light carbon-formvar coated grids (TED-PELLA., INC) were used without glow-discharging. 2 % (w/v) uranyl acetate (Sigma-Aldrich) was prepared by dissolving 200 mg powder in 10 mL sterile distilled water. For preparation of phosphor tungstic acid solution of pH 8.0, 200 mg phosphotungstate (Sigma) powder was dissolved initially in 5 mL sterile distilled water. pH of the solution was adjusted to 8.0 and volume made up to 10 mL. The solution was filtered using 0.22 µm filter and then distributed as 500 µL aliquots into 1.5 mL tubes. All the uranyl acetate containing tubes were covered with aluminum foil and stored at -80 ^{0} C until use while phosphotungstic acid containing tubes were stored at room temperature. For sample preparation, solution of desired concentration (0.5 % w/v) was prepared by diluting the stain solution with sterile water.

Samples were prepared by applying varying volumes (2-5 µL) of purified protein to grids without glow-discharging. The grids with sample were allowed to stand at room temperature for at least 30 seconds to facilitate settling down of the filaments. Excess buffer was back-blotted using Kimwipes® (Kimberly Clark Hygiene Products Pvt Ltd., Pune, India). Stain (0.5 % w/v uranyl acetate) solution was applied to the grid and immediately absorbed from bottom using blotting paper so as to prevent absorption of the stain by protein. The grids were allowed to dry

for 10-15 minutes at room temperature in a dust-free environment or were dried under table lamp for 2 hours. Dried grids were stored in a grid storage box in a dry, dust-free environment until observation in a TEM.

2.3) Results

Two heterologous expression systems were used for Fibril expression and purification. Multiple protocols have been standardized for Fibril purification from *E. coli.* Purified Fibril (molecular weight of a monomer = 59 kDa) have been visualized using electron microscopy. Following sections discuss the observations and advantages and disadvantages of different protocols.

2.3.1) Fibril purified from *E. coli* exists as filaments

Conventionally Fibril have been purified from *Spiroplasma,* for which a protocol has been standardized (Cohen-Krausz et al., 2011; Townsend et al., 1980; Trachtenberg and Gilad, 2001). This protocol had to be modified since *Spiroplasma* cells lack cell wall whereas I was working with a cell wall containing organism, *E. coli*. The modified protocol for Fibril purification from *E. coli* has been described in section 2.2.3.1) of this chapter.

During purification of Fibril from *E. coli* cell lysate using glycerol gradient, Fibril remained in supernatant when spun at 3200 xg. Subjecting the soluble fraction to 100,000 xg spin led to pelleting down of Fibril filaments along with contaminant proteins (Figure 2.2 A). Along with Fibril, the other prominent bands observed on SDS-PAGE gel correspond to molecular weight of about 37 kDa and 27 kDa (Figure 2.2 A). The pellet obtained by spinning clarified lysate at 100,000 xg was insoluble in the various buffers tested. Thus, the insoluble proteins from 100,000 xg pellet were subjected to separation on glycerol density gradient. In the gradient, a small fraction of Fibril remained in the top fractions (numbered 18-21) along with other protein impurities (Figure 2.2 B, C and D) while most of the protein was found at the bottom of the gradient as impure fraction. Fibril filaments were observed when the samples prepared using dialyzed fractions containing protein were visualized using FE-SEM (Figure 2.3 A and B). This confirmed that the purified Fibril indeed existed as filaments. Thus, enriched Fibril filaments could be obtained using the glycerol density gradient method.

However, certain drawbacks with enriching Fibril using the above discussed method are listed below. 1) The protein pellet obtained by spinning the lysate at 100,000 xg could not be re-solubilized into suitable buffer for further processing. 2) Most of the impurities did not get separated on glycerol gradients. Fractions from glycerol gradient containing Fibril had many contaminant proteins (lanes 18-21 in Figure 2.2 D), making its further purification necessary. 3) Visualization of protein from gradient fractions in FE-SEM showed that the Fibril filaments were

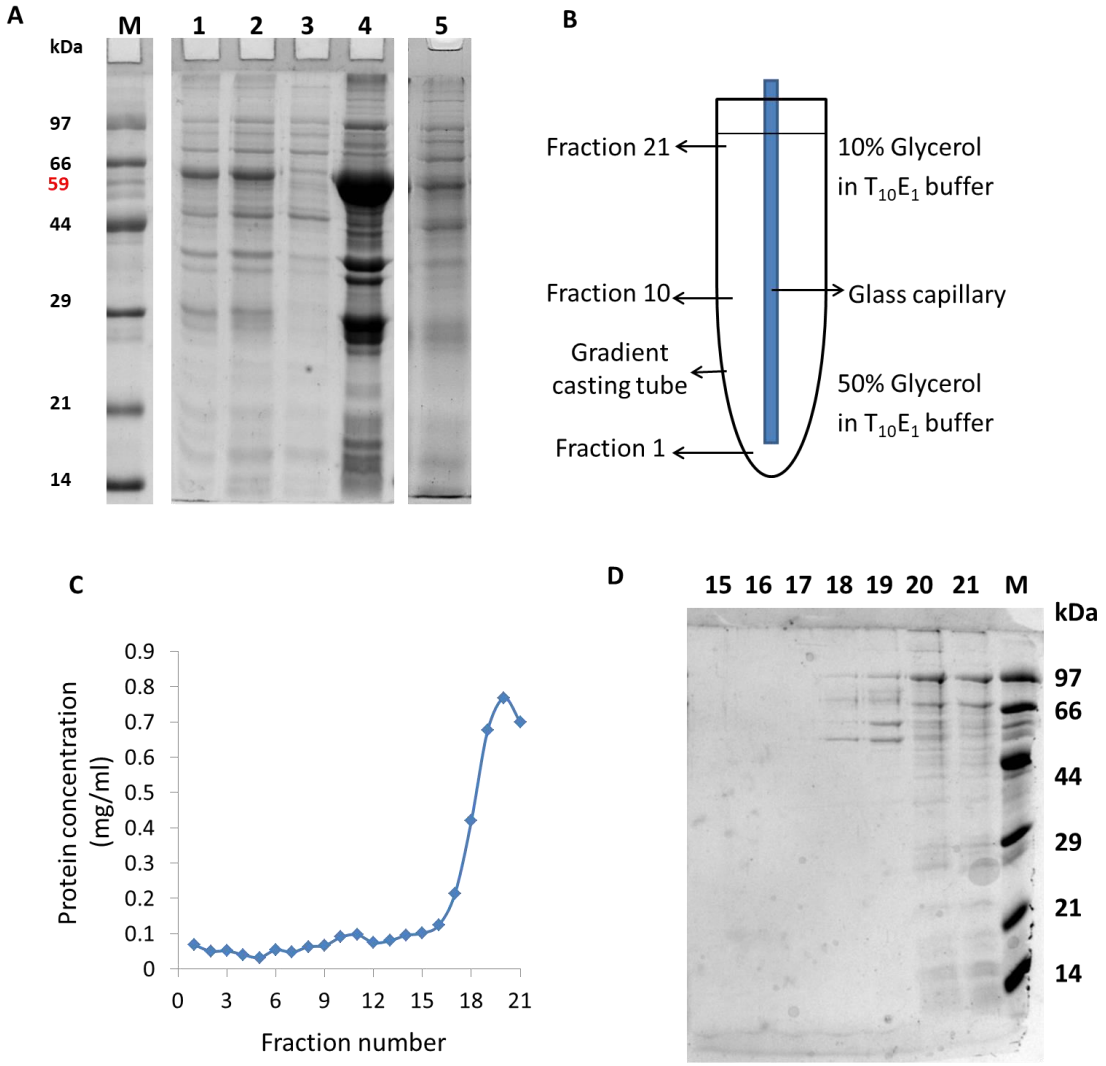

Figure 2.2 Purification of His_6-full length Fibril from *E. coli* using glycerol density gradient. A) 12 % SDS-PAGE gel showing purity of Fibril at different stages of pelleting. Lanes correspond to M- protein marker, 1- total lysate, 2- supernatant after 3,200 xg spin, 3 and 4- supernatant and pellet after 100,000 xg spin respectively and 5- 100,000 xg pellet partially re-suspended. B) Pictorial representation of glycerol density gradient fractionation. The initial fractions are most dense while the later fractions have less density. C) A plot showing protein concentration in different fractions after density gradient fractionation. D) Visualization of proteins in fractions obtained from density gradient. Proteins could be seen in fractions 18-21.

covered probably with compounds used for purification (detergents and/or glycerol) or protein impurities (Figure 2.3 A and B). In order to overcome these issues standardization of buffer conditions and 100,000 xg pellet re-solubilization was done.

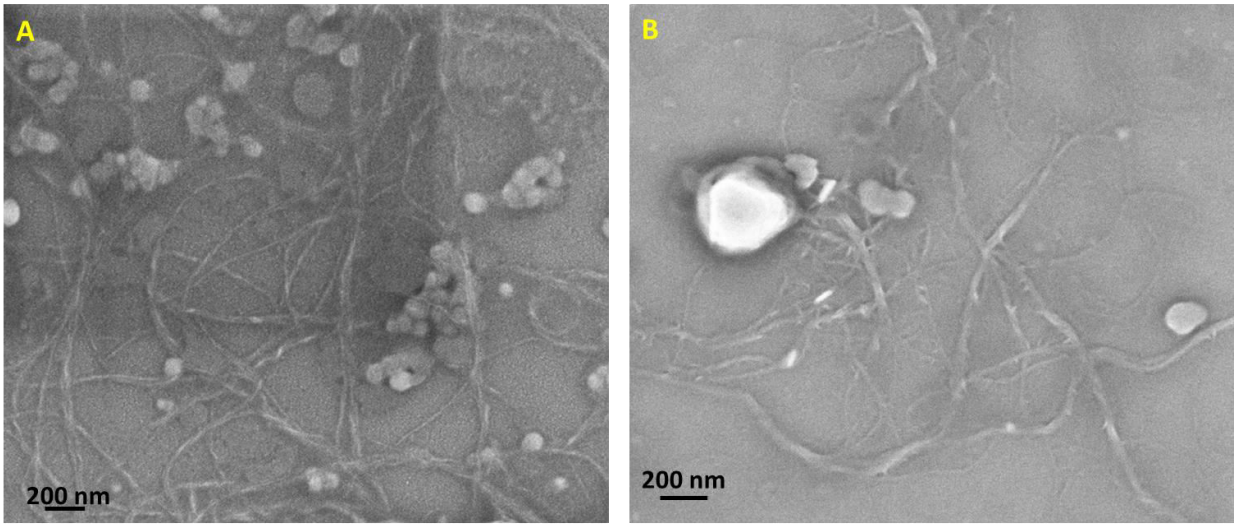

Figure 2.3 Visualization of enriched His_6-full length Fibril obtained from *E. coli* using glycerol density gradient. A & B) FE-SEM images of His_6-full length Fibril purified from *E. coli* by glycerol density gradient method shows filament bundles from which thinner filaments are emanating. Though filaments can be seen for Fibril purified by glycerol density gradient method, they are accompanied by many unwanted clumps that occasionally cover filaments.

2.3.2) SDS treatment of Fibril filaments facilitates Ni-NTA binding.

A few detergents/ compounds were tested for their ability to i) increase the recovery of protein in the supernatant fraction after lysis and ii) solubilise the protein obtained in the pellet after 100,000 xg spin.

Use of urea to reversibly unfold proteins is a well-known technique (Egan et al., 1993). Thus using urea to unfold Fibril and expose hexahistidine tag to facilitate Ni-NTA based affinity chromatography was thought as a possibility. However, based on literature it is known that Fibril are only slightly affected by 8 M urea treatment and are soluble in SDS (Townsend, 1983). In many cases, detergents have proven to be useful to solubilize or stabilize membrane proteins (Privé, 2007; Seddon et al., 2004). Thus, different compounds [0.5 % N-lauroylsarcosine sodium salt (LSS), 20 mM ethylenediaminetetraacetic acid (EDTA), 1 % sodium deoxycholate (SDC), 1

% sodium dodecyl sulphate (SDS), 1 % TWEEN 20] were used as additive to a buffer base to attempt solubilization of the 100,000 xg pellet. Complete pellet solubilization was observed only with the addition of SDS solution (at a final concentration of 1 % w/v). On the contrary, all other compounds did not have a significant effect on solubility of Fibril (Figure 2.4 A-C).

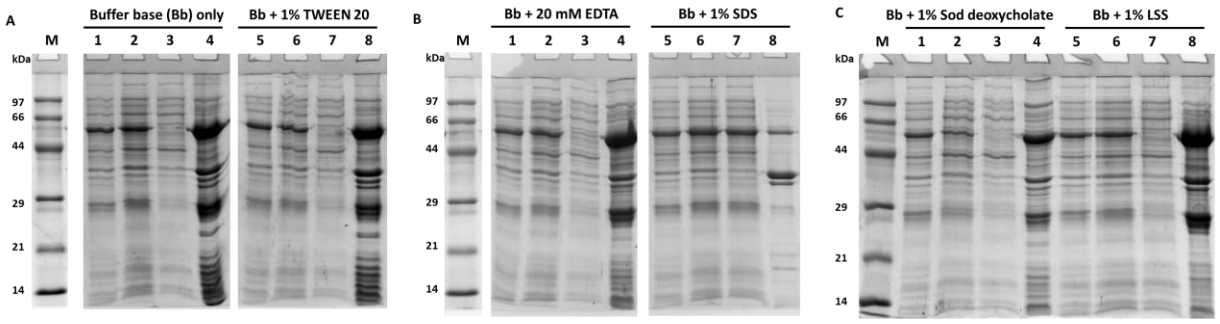

Figure 2.4 Detergent screen to solubilize His_6-full length Fibril obtained in 100,000 xg pellet. 12 % SDS-PAGE gels showing solubility of Fibril after spinning at 100,000 xg in lysis buffer 2 (buffer base) and upon treatment with 1 % TWEEN 20 detergent (A), in presence of 20 mM EDTA and upon treatment with 1 % SDS detergent (B), upon treatment with sodium deoxycholate (1 %) or Lauroyl sarcosine sodium salt (LSS; 0.5 %) (C). Fibril was solubilized only by treatment using SDS (lane 7, panel B). Lanes: M- Protein marker, 1 & 5- total lysate, 2 & 6- supernatant after 4,629 xg spin, 3 & 7- supernatant after 159,000 xg spin, 4 & 8- pellet after 159,000 xg spin.

Hence, SDS treatment step was incorporated to solubilize the pellet obtained after 100,000 xg spin. Attempts were made to separate Fibril filaments from non-polymeric proteins by spinning the clarified solution at 159,000 xg. Though Fibril filaments were expected to be in the pellet, they remained in supernatant (Figure 2.5 A, lane 6). This suggested that Fibril probably formed less dense filaments that do not pellet down at 159,000 xg or may have dissociated into short oligomers/monomers.

Since the SDS-solubilised Fibril possessed an N-terminal or a C-terminal hexahistidine (His_6) tag, I proceeded to purify them further using Ni-NTA affinity chromatography (Figure 2.5 A, lane 7). When the dialyzed Fibril were visualized using FE-SEM, bundles and twisted filaments with width ~ 50 nm were seen (Figure 2.5 B). The dialyzed protein was concentrated to about 16 mg/mL concentration and used for visualization by TEM. Individual filaments within the bundle could be seen under TEM (Figure 2.5 C and D). The purified Fibril were often associated with ~ 27 kDa band pairs which might be the breakdown product of full-length Fibril. Similar bands

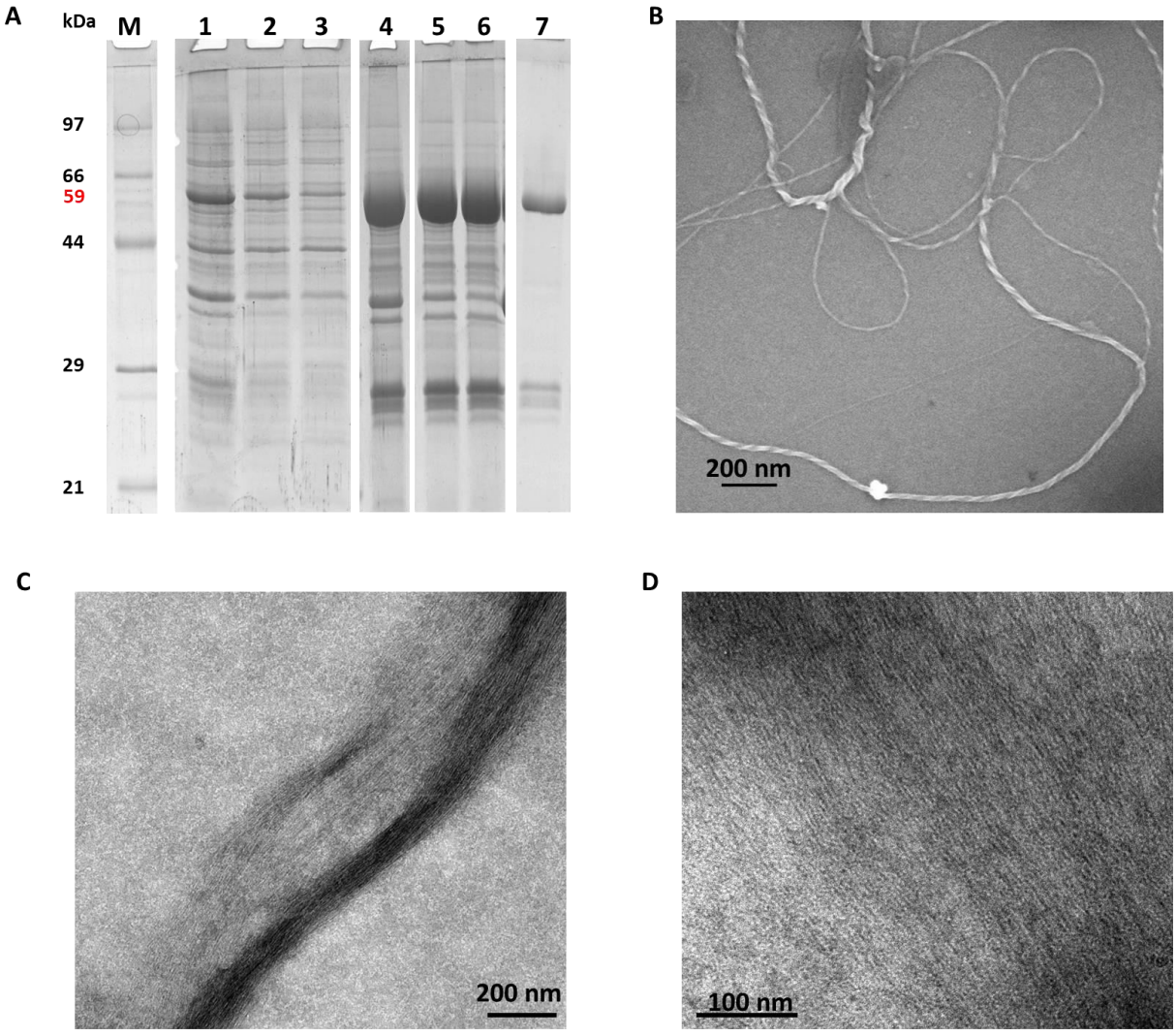

Figure 2.5 Purification and visualization of His_6-full length Fibril expressed in *E. coli* by using SDS-treatment and affinity chromatography. A) Solubility of Fibril at different stages of pelleting for purification by SDS treatment. Lanes correspond to M- Protein marker, 1- total lysate, 2- supernatant after 3,200 xg spin, 3- supernatant after 159,000 xg spin, 4- pellet after 159,000 xg spin, 5- supernatant after 21,000 xg spin upon SDS treatment, 6- supernatant after 159,000 xg spin upon SDS treatment, 7- protein purified using Ni-NTA affinity chromatography. B) FE-SEM micrograph of Fibril purified by SDS treatment followed by affinity chromatography assisted purification. C & D) TEM images of Fibril bundle purified by SDS treatment, affinity chromatography and stained with 1 % phosphotungstic acid. Individual filaments within a protein bundle are seen. D) Zoomed view of individual filaments in a bundle.

have been observed for purification of full length Fibril by glycerol density gradient method that did not employ use of SDS (Figure 2.2 A). The consistent occurrence of these bands with molecular weights of about 37 kDa and 27 kDa suggest that these bands have resulted from proteolysis of full length Fibril.

In spite of the reports that SDS can denature proteins (Giancola et al., 1997), Fibril bundles were observed upon SDS treatment. This suggests that Fibril might be able to withstand treatment with 1 % SDS. However, the possibility of partial denaturation by SDS treatment cannot be ruled out due to absence of a functional assay for Fibril. In such a case, use of SDS treated proteins for structural studies might result into determination of a physiologically irrelevant conformation of the protein. Structural comparison of protein purified using SDS versus protein purified without using SDS will be useful to know if Fibril remain unaffected by SDS treatment. Thus, it was decided to standardize alternate protocols to purify Fibril without using harsh conditions.

2.3.3) Fibril solubilized by LSS does not bind to ion-exchange and hydrophobic interaction chromatography matrices

Based on my earlier experiments it was known that Fibril pelleted by spinning at or above 100,000 xg were insoluble without SDS treatment. One of the ways to avoid use of SDS was by identification of a lysis buffer composition that did not let Fibril filaments pellet down at 100,000 xg or facilitated their solubilization from 100,000 xg pellet.To identify lysis buffer that fulfilled these requirements, solubility experiments were carried out with lysis buffers 6, 7 and 8 (Appendix 1). In lysis buffers 6 and 8 (containing Triton X-100) Fibril was obtained in insoluble fraction when the lysate was spun at 100,000 xg. However in lysis buffer 7 (not containing Triton X-100 but contains LSS), Fibril did not pellet down when spun at 100,000 xg/4 ^{0}C for 30 minutes and were observed in the supernatant (Figure 2.6 A).

This suggested that Fibril may not be in filamentous form or may be forming small oligomers that did not pellet down even at 159,000 xg. The other possibility is that, LSS keeps Fibril filaments in solution. Thus, lysis buffer 7 [with 10 mM Tris (pH 7.6), 0.5 g /100 mL LSS composition] was identified as the condition that kept Fibril in the soluble fraction. In order to purify soluble Fibril from rest of the lysate, affinity chromatography purification was attempted since Fibril was His_6 tagged. Unlike the SDS-treated Fibril filaments, the protein did not bind to beads in this case, probably due to an unexposed His_6 tag. Thus the possibility of purification of soluble Fibril using affinity chromatography was ruled out.

Ammonium sulphate precipitation is a routinely used technique to remove unwanted proteins (Wingfield, 2016). In some cases, it resulted into obtaining completely pure protein (Buchanan,

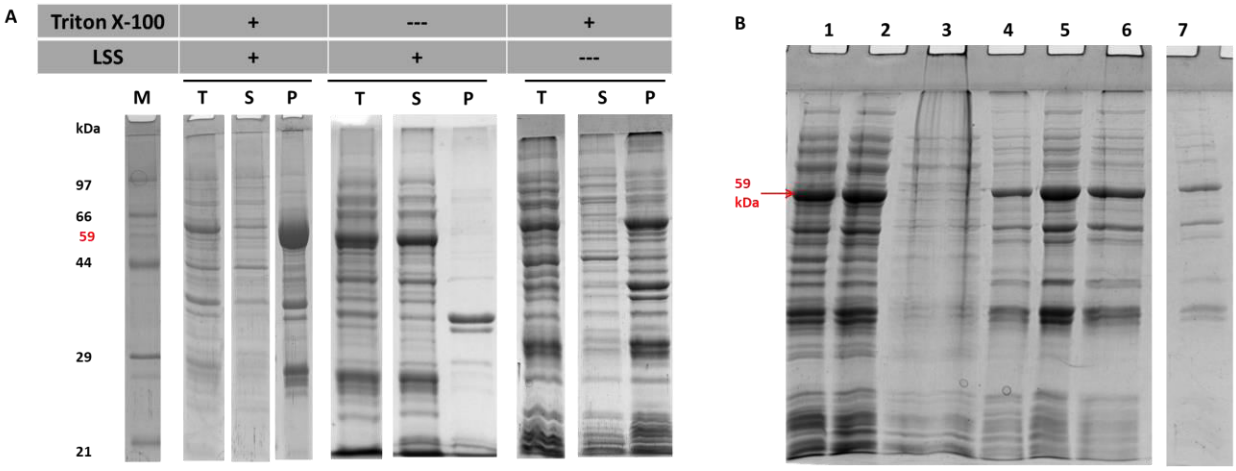

Figure 2.6 Purification of His_6-full length Fibril expressed in *E. coli* by ammonium sulfate precipitation followed by ion-exchange chromatography. A) A 12 % SDS-PAGE gel showing effect of Triton X-100 and Lauroyl sarcosine sodium salt on solubility of Fibril after spinning at 100,000 xg. Lanes correspond to M- protein marker, T- total cell lysate, S- supernatant after 100,000 xg spin, P- pellet after 100,000 xg spin. B) Partial purification of Fibril by precipitation using 50 % ammonium sulfate saturation. 1- total cell lysate, 2- supernatant after 100,000 xg spin, 3- supernatant obtained after precipitation of proteins at 40 % ammonium sulphate saturation followed by spinning at 100,000 xg, 4- supernatant obtained after precipitation of proteins at 40 % ammonium sulfate saturation followed by spinning at 100,000 xg was re-suspended in acetate buffer, 5- pellet obtained by spinning proteins in acetate buffer at 28,000 xg followed by re-suspension in Tris, 6- soluble protein obtained in buffer containing 50 mM Tris (pH 8.0) and 700 mM ammonium sulfate after two rounds of spinning at 28,000 xg. This is also referred to as input for HIC columns. 7- A fraction of protein eluted from phenyl (high substitution) HIC column using 100 % elution buffer. Note that the sample loaded in well 7 is double that of other wells in the gel.

1975; Mohanty and Elazhary, 1989). Hence, ammonium sulphate precipitation step was incorporated to obtain enriched Fibril from lysate. A variety of chromatography techniques including ion-exchange, multimodal cation-exchange and hydrophobic interactions were attempted to further purify Fibril. Very less quantity (~ 5% of input) of Fibril bound to some of the columns and eluted along with other impurities (Figure 2.6 B). Thus, it was concluded that Fibril solubilized by using LSS cannot be purified using these techniques.

Thus, of all the methods to purify Fibril from *E. coli*, only SDS treatment protocol resulted into purified Fibril filaments. While soluble Fibril obtained using LSS could not be purified, the experiments to purify Fibril using glycerol density gradient resulted into most of the protein going to the bottom of the gradient. This suggested that the density range used in glycerol density gradients was not sufficient to separate Fibril from protein clumps. Thus, it was decided to use the compounds with density higher than Fibril filaments for preparation of density gradients.

2.3.4) Fibril filaments can be purified from *E. coli* using urografin gradient

In the past, metrizamide and urografin have been used for Fibril purification from *Spiroplasma* cells (Cohen-Krausz et al., 2011; Townsend et al., 1980). Since urografin was readily available from medical stores, urografin was chosen.

It has been reported that prolonged stirring (~ 2 months) of Fibril in buffer containing 1 M KCl helps separate individual Fibril filaments from bundle (Cohen-Krausz et al., 2011). Hence, the enriched Fibril containing pellet obtained by spinning clarified lysate at 159,000 xg was stirred for 15 days and separated on urografin density gradient (Figure 2.7 A). The fractions 6 and 7 from urografin density gradient showed Fibril (59 kDa band) as the predominant protein. In these fractions, the other prominent bands observed, apart from Fibril, were with molecular weight around 27, 37 and 44 kDa. The presence of some of these bands was also seen during Fibril purification using glycerol gradient as well as SDS treatment protocol. Visualization of the enriched Fibril in TEM revealed the presence of Fibril bundles though the protein was sparsely distributed on the electron microscopy grid (Figure 2.7 B-D).

The distribution of protein on grid can be increased by concentrating the protein sample. However, Fibril filaments were often associated with many clumps that were thought to have resulted from the detergents (Triton X-100 and/or N-lauroylsarcosine sodium salt) and/or glycerol used in the lysis buffer and may have remained associated with protein bundles. The clumps could also be Fibril bundles that had absorbed excess stain. One of the ways to overcome these clumps was by avoiding glycerol in the lysis buffer and standardizing the buffer by screening detergents. As a result, detergents were removed from lysis buffer in step-by-step manner and the reduction in clumps accompanying Fibril during EM imaging was observed.

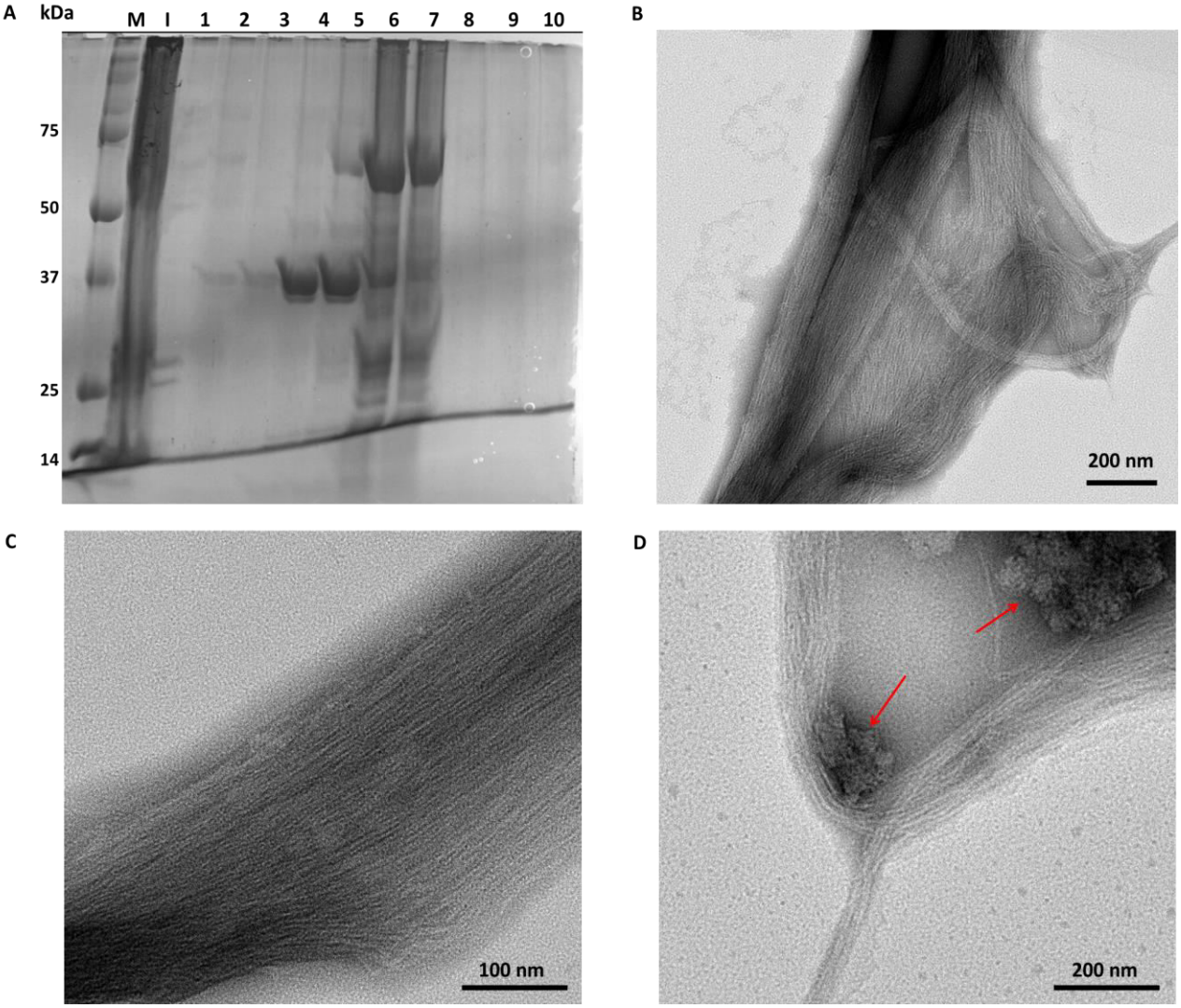

Figure 2.7 Purification and characterization of Fibril expressed in *E. coli*, using Triton X-100 and sodium deoxycholate detergent in lysis buffer and separation using urografin density gradient. A) 12 % SDS-PAGE gel showing distribution of Fibril bundles in different fractions of urografin gradient. The protein was pelleted using lysis buffer containing Triton X-100 and sodium deoxycholate detergents. The pellet was stirred and fractionated on urografin density gradient. Lanes correspond to M- protein marker; 1 to 10- fractions from urografin gradient from top to bottom in the increasing order of density. The problems associated with smeary migration of gradient fractions on gel are due to the presence of urografin, a dense material. B-D) Negatively stained TEM images of Fibril filaments from fractions 6 and 7. Individual filaments in the bundle can be seen. Red arrows point to the clumps associated with Fibril bundles.

In the first step, Fibril were purified using sodium deoxycholate as the only detergent and observation under TEM revealed Fibril bundles (Figure 2.8 A-D). The Fibril bundles were often associated with clumps (Figure 2.8 A-C).

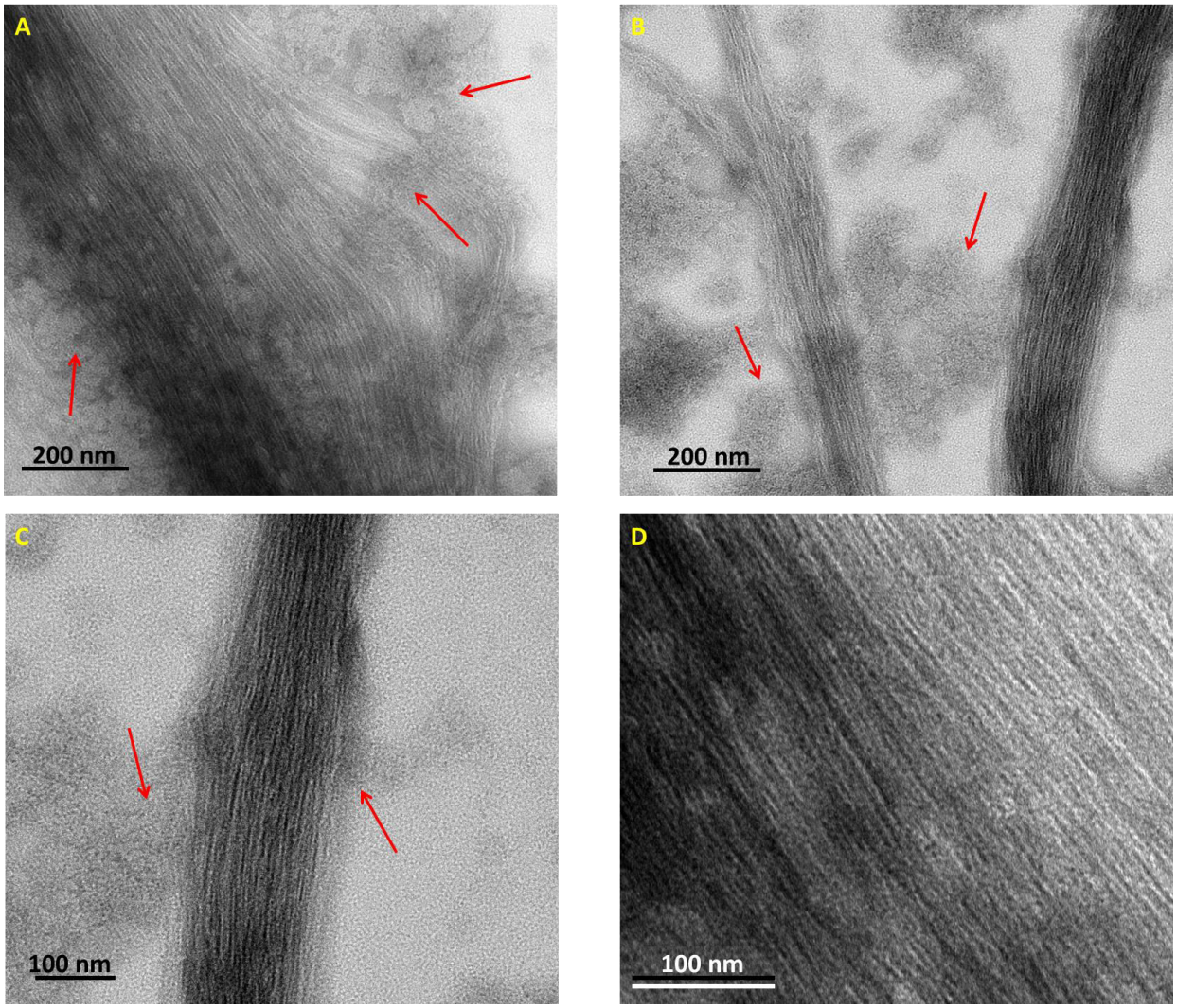

Figure 2.8 Purification and visualization of Fibril expressed in *E. coli*, using sodium deoxycholate detergent in lysis buffer and separation using urografin density gradient. A-D) Negatively stained TEM images of Fibril filaments purified using sodium deoxycholate as the only detergent in lysis buffer. Individual filaments in the bundle can be seen. Red arrows point to the clumps associated with Fibril bundles.

Next, Fibril purification was attempted without using any detergent or glycerol, by steps of differential centrifugation (Figure 2.9 A). Fibril obtained in pellet, after spinning clarified lysate at 159,000 xg, was insoluble in the TEK buffer (10 mM Tris pH 7.6, 10 mM EDTA, 1 M KCl).

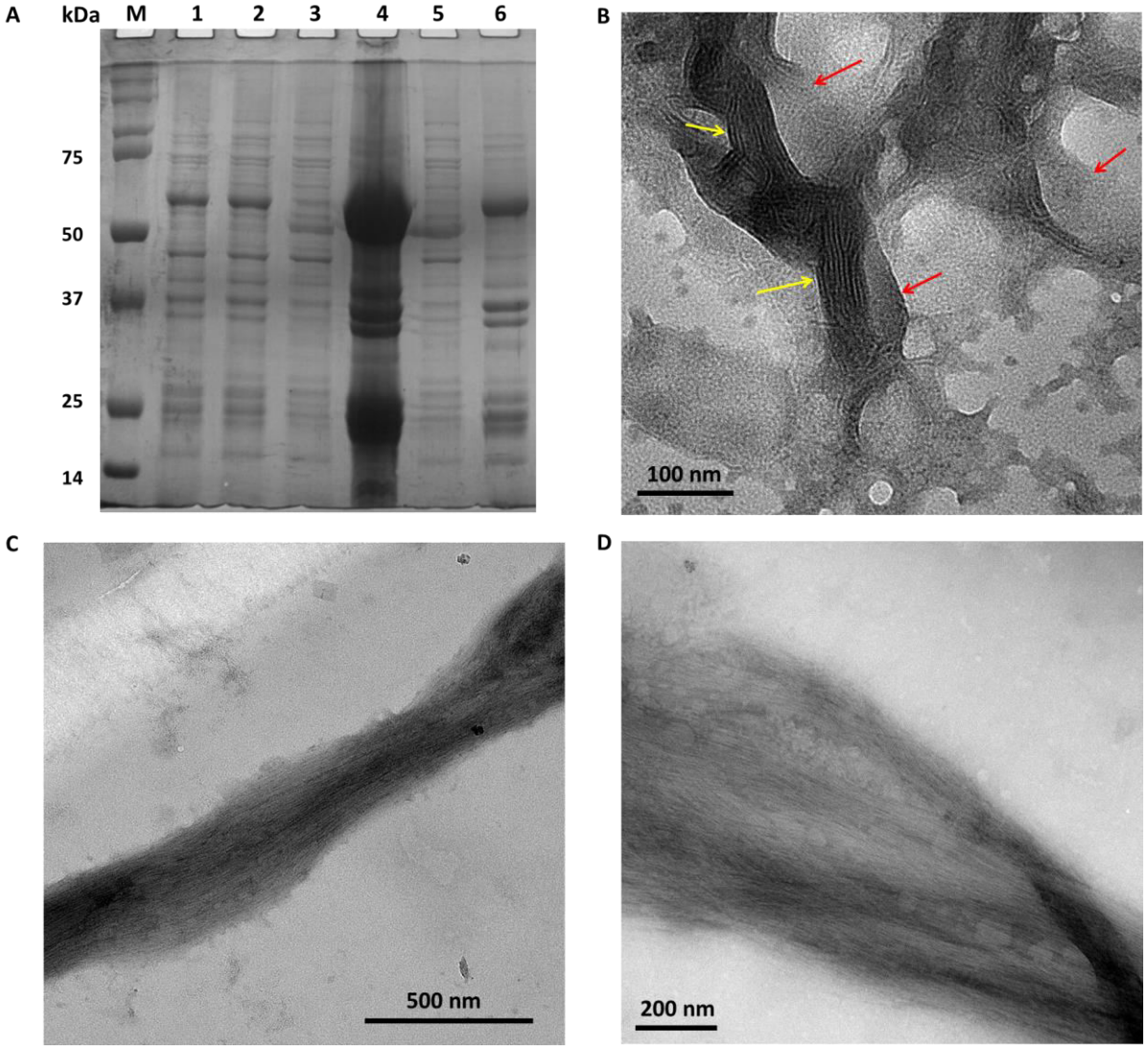

Figure 2.9 Purification and characterization of Fibril expressed in *E. coli*, without using any detergent in lysis buffer and separation using urografin density gradient. A) Distribution of Fibril filaments during purification without use of any detergent as visualized on a 12 % SDS-PAGE gel. Lanes correspond to M- protein marker, 1- total lysate; 2- supernatant obtained by spinning lysate at 8,000 xg; 3 & 4- supernatant and pellet respectively, obtained after spinning the 8,000 xg supernatant at 30,000 xg; 5 & 6- supernatant and pellet respectively, obtained after spinning the 30,000 xg supernatant at 159,000 xg. Negatively stained TEM images of Fibril purified without any detergent (B), followed by treatment with chloroform + methanol mixture (C and D). Yellow arrows point to protein filaments while associated clumps are shown with red arrows. In all the above images individual filaments within Fibril bundle are seen clearly.

Hence it was stirred in $T_{10}E_{10}$ (10 mM Tris pH 7.6, 10 mM EDTA) buffer to break down clumps before visualization of protein using TEM. The TEM images of Fibril purified by this protocol also revealed clumps associated with Fibril filaments (Figure 2.9 B). Hence, it is possible that the clumps were *E. coli* lipids associated with Fibril bundles.

The purified protein was further subjected to chloroform and methanol extraction to remove detergents/lipids. Protein purified by this method was re-suspended in low volume of buffer to keep the protein concentration high. Grids prepared using this protein revealed a good distribution of Fibril bundles in TEM (Figure 2.9 C and D).

To summarize the purification of Fibril from *E. coli,* multiple protocols have been standardized and enriched Fibril has been obtained. Visualization of Fibril using electron microscopy revealed that Fibril purified from *E. coli* exist as filaments. Except for the SDS treatment protocol (Figure 2.5 A), the Fibril purified by all other protocols revealed that Fibril was accompanied by unwanted proteins (Figures 2.2 A & D, 2.9 A).

The contaminant proteins may be avoided by use of an expression system that can secrete target proteins outside the cell. Thus, simultaneous experiments were performed for purification of Fibril using *Pichia pastoris* as another heterologous protein expression system.

2.3.5) Fibril is found in the cytoplasmic fraction in a secretion-based strategy for *Pichia pastoris* expression

In *P. pastoris* the target protein can be expressed with α-factor secretion signal that facilitates secretion of target protein outside the cell, into the growth medium. Since *P. pastoris* secretes only a few proteins, the purification of target protein from the media may be easier as compared to purification from the cytosol. As *fibril* was being expressed with an N-terminal α-factor signal sequence, it was expected to be secreted into the growth media (Figure 2.10 A). When the media was checked on a SDS-PAGE gel to monitor the secreted protein, an expected band corresponding to 61.5 kDa (59 kDa of Fibril with 2.5 kDa of C-myc epitope and His_6 tag) was not observed. Hence, the transformed *P. pastoris* cells were checked to investigate if Fibril was expressed. Presence of an overexpressed protein band corresponding to ~ 71 kDa was confirmed in the cell lysate as compared to uninduced cells (Figure 2.10 B). This corresponds to the expected size of the expressed protein along with the α-factor signal sequence. Thus it was

thought that in Fibril tagged with the α-factor signal sequence, the KeX2 site was not available for proteolysis probably due to Fibril polymerization and appeared as a 71 kDa band on gel. Since Fibril was not separated from α-factor signal, it remained inside the cells. Hence, attempts were made to purify Fibril from within *P. pastoris* cells.

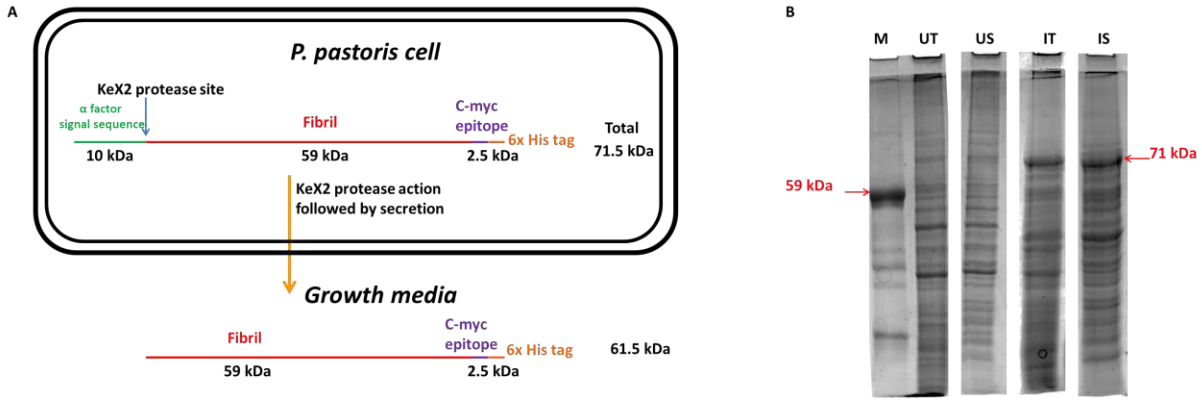

Figure 2.10 Expression of Fibril in *P. pastoris*. A) Schematic representation for expected expression and secretion of Fibril using *P. pastoris* cells and pPICZα vector. B) 12 % SDS-PAGE gel showing expression of Fibril only upon induction by methanol. Lanes correspond to M- protein marker with 59 kDa protein as the most prominent band, UT- total cell lysate of uninduced *P. pastoris* GS-115 cells, US- supernatant obtained by spinning lysate of uninduced *P. pastoris* GS-115 cells at 21,000 xg, IT- total cell lysate of P. pastoris GS-115 cells transformed with Fibril in pPICZα vector and induced with methanol, IS- the soluble fraction of proteins obtained by spinning IT at 21,000 xg. The 71 kDa band corresponds to un-cleaved α-factor signal sequence.

In contrast to Fibril expressed in *E. coli*, the α-factor signal sequence-tagged (71 kDa band) Fibril expressed in *P. pastoris* was soluble in lysis buffer 3. It is important to note that lysis buffer 3 does not have any detergent as its constituent. The 71 kDa protein did not pellet down upon spinning at 100,000 xg/4 ^{0} C/ 30 minutes (lanes 1 and 2 in Figure 2.11). Fibril purification was attempted using the cell lysate and Ni-NTA affinity purification.

Most of the Fibril did not bind affinity matrix and eluted as flowthrough or wash (lanes 2 and 3 respectively, Figure 2.11). Fractions eluted using increasing concentrations of imidazole showed very less amounts of 71 kDa protein in the bound fractions (lanes 5-10 in Figure 2.11). The inability of Fibril attached with α-factor signal sequence to bind Ni-NTA resin was consistent upon repetitions of the experiment. Thus, it was concluded that full-length Fibril expressed in *P. pastoris* with α-factor signal sequence at the N-terminus and C-myc epitope followed by 6xHis-tag at C-terminus does not bind to Ni-NTA resins probably due to unexposed 6xHis tag. This

finding is consistent with my experiments involving purification of N- or C- terminally hexahistidine tagged Fibril from *E. coli*. Thus, it is highly likely that the amino (N) as well as carboxy (C) termini of Fibril remain buried.

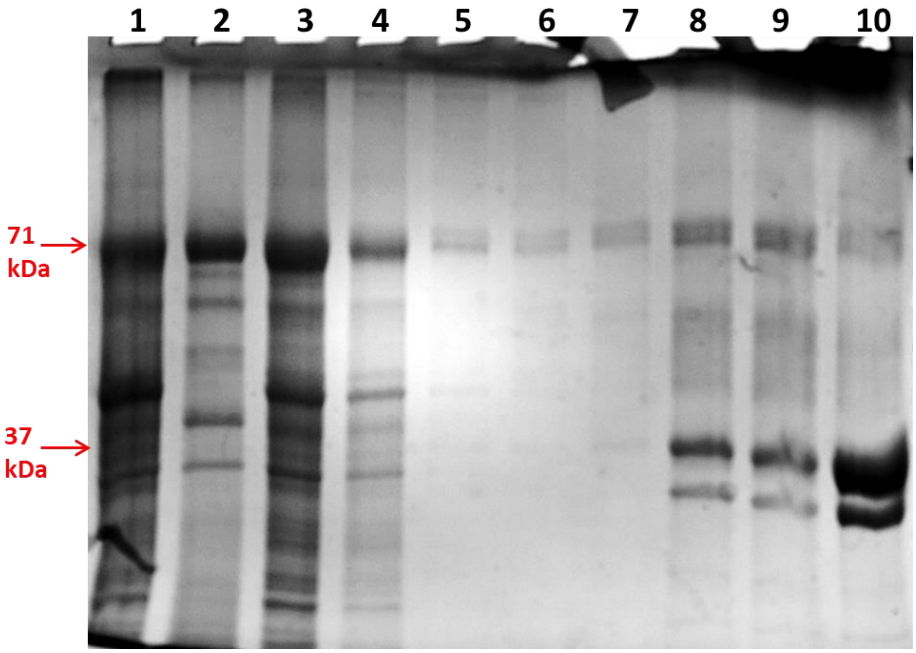

Figure 2.11 Purification of Fibril in *P. pastoris*. A 12 % SDS-PAGE gel showing Fibril at different stages during purification by affinity chromatography. Lanes: 1 - total cell lysate; 2 - soluble fraction of lysate obtained by spinning at 100,000 xg; 3 - flowthrough or unbound protein from the lysate; 4 - protein obtained by washing the column with binding buffer; 5 to 10 - fractions of protein bound to Ni-NTA resins, eluted with increasing concentrations of imidazole containing buffer. A small fraction of Fibril (71 kDa band) seems to bind the Ni-NTA matrix with very low affinity.

Interestingly, a protein of about 37 kDa mass from Fibril-expressing *P. pastoris* lysate bound tightly to Ni-NTA resins and eluted with higher concentrations of imidazole (lanes 8-10 in Figure 2.11). In Ni-NTA affinity based purification, normally the hexahistidine tagged protein binds the matrix with such high affinities. Thus, I proceeded with the purification of the protein corresponding to 37 kDa band binding with assumption that it might be a proteolytic product of Fibril with exposed hexahistidine tag.

Since no functional assays are available for Fibril, the truncated (37 kDa) protein was purified and used for biochemical and structural characterization. Structure determination of the protein revealed that it is alcohol dehydrogenase from *P. pastoris* (data not shown). The binding of alcohol dehydrogenase to Ni-NTA resins with high affinity can be explained by the fact that it is

a metalloprotein (Ni-binding protein). Details regarding characterization of this protein are not included in this thesis.

Hence, Fibril could not be purified using *P. pastoris* as the heterologous expression system and using the methods summarized above. Due to unavailability of *Spiroplasma* cultures at any research laboratories or microbial culture collection centers in India, Fibril purification from *Spiroplasma* species could not be carried out in India. However, an opportunity to spend 6 months in the lab of Dr. Laure Béven, INRA, Bordeaux (a lab that extensively works on *Spiroplasma*) provided chance to purify Fibril from *Spiroplasma citri* cells.

2.3.6) Enriched Fibril filaments are obtained from *Spiroplasma citri* GII-3 cells

Partially purified Fibril was obtained from *S. citri* cells in the Triton X-100 insoluble fraction

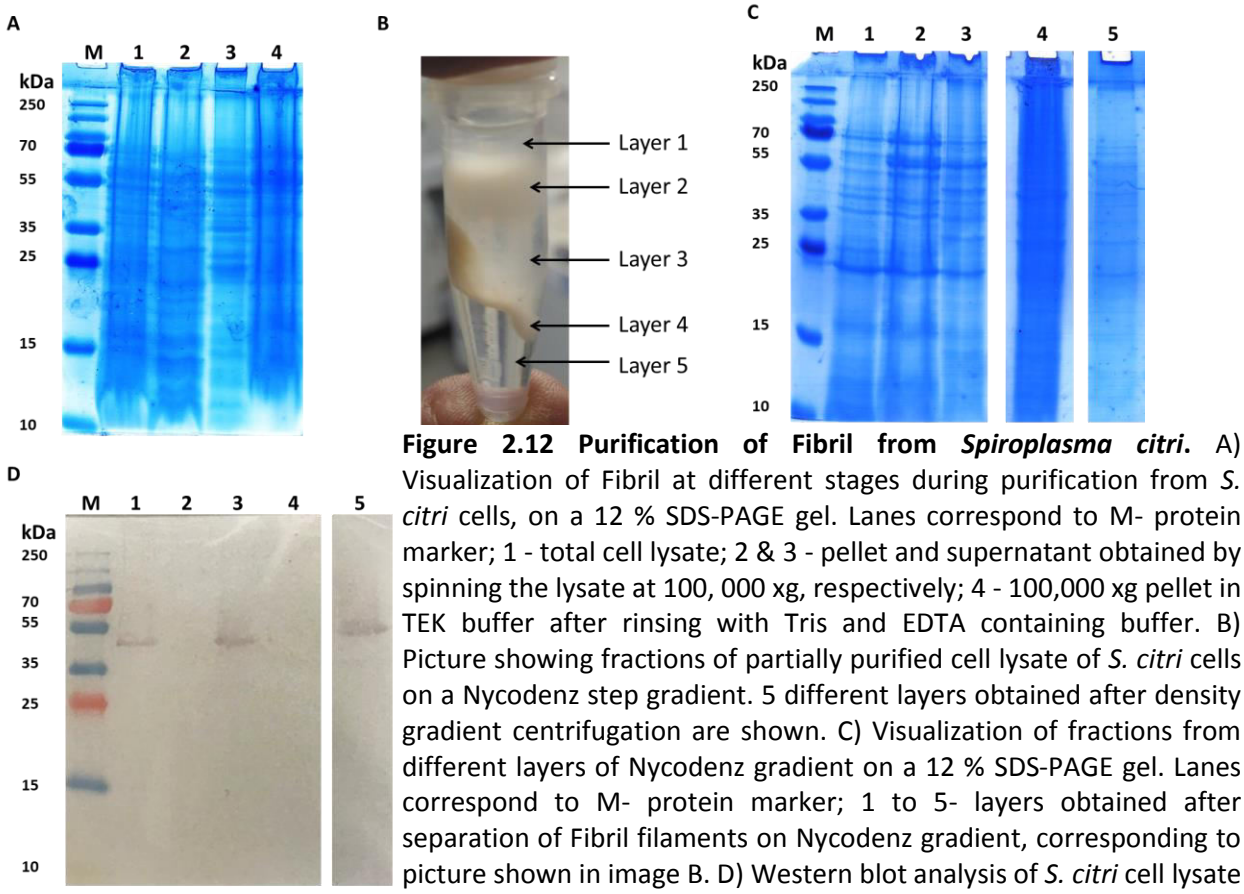

Figure 2.12 Purification of Fibril from *Spiroplasma citri*. A) Visualization of Fibril at different stages during purification from *S. citri* cells, on a 12 % SDS-PAGE gel. Lanes correspond to M- protein marker; 1 - total cell lysate; 2 & 3 - pellet and supernatant obtained by spinning the lysate at 100, 000 xg, respectively; 4 - 100,000 xg pellet in TEK buffer after rinsing with Tris and EDTA containing buffer. B) Picture showing fractions of partially purified cell lysate of *S. citri* cells on a Nycodenz step gradient. 5 different layers obtained after density gradient centrifugation are shown. C) Visualization of fractions from different layers of Nycodenz gradient on a 12 % SDS-PAGE gel. Lanes correspond to M- protein marker; 1 to 5- layers obtained after separation of Fibril filaments on Nycodenz gradient, corresponding to picture shown in image B. D) Western blot analysis of *S. citri* cell lysate at different purification stages using anti-Fibril antibody. M- protein marker; 1- total lysate; 2 & 3- supernatant and pellet respectively, obtained by spinning lysate at 100,000 xg; 4- layer 2 from nycodenz gradient; 5- layer 4 from Nycodenz gradient. Fibril is seen in total lysate, insoluble fraction upon spinning at 100,000 xg and in fraction 4 of Nycodenz gradient.

using standardized protocol (Cohen-Krausz et al., 2011). Presence of Fibril at various purification stages was confirmed by checking samples on SDS-PAGE gels (Figure 2.12 A).

Upon subjecting the Triton X-100 insoluble fraction to separation, Fibril were found in the fourth layer of Nycodenz gradient (Figure 2.12 B). Though enriched Fibril have been obtained, they are associated with many unwanted proteins (Figure 2.12 C). Presence of Fibril at various stages during purification was confirmed by Western blot analysis using anti-Fibril antibody (Figure 2.12 D).

The enriched protein was visualized by preparing negatively stained samples and imaging using an electron microscope. The TEM images revealed presence of Fibril bundles (Figure 2.13 A and B). Similar bundles have been seen for Fibril purified from *E. coli* using various protocols described in this chapter (Figures 2.5, 2.7, 2.8 and 2.9).

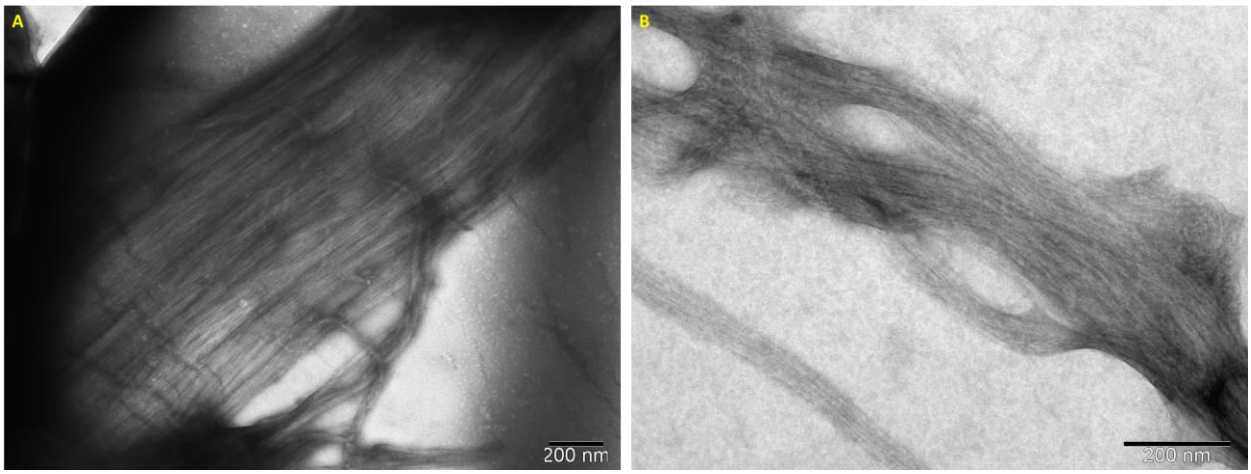

Figure 2.13 Visualization of Fibril purified from *Spiroplasma* cells. A and B) Negatively stained TEM images of Fibril purified from *S. citri* GII-3 cells. Fibril bundles can be seen emanating from big clumps of protein. The scale bars in both the images represent 200 nm.

At this stage the Fibril filaments/bundles obtained from *E. coli* and *Spiroplasma citri* look similar. Structure determination of Fibril by cryoEM will be the next step in the project. However, sample standardization to obtain well separated filaments is necessary before proceeding ahead.

2.4) Discussion

In absence of a pure culture of *Spiroplasma* in India, heterologous expression systems were explored for obtaining purified Fibril. The heterologous expression system also facilitates genetic

modifications of Fibril protein. The work was initiated by expressing codon-optimized *fibril* gene initially in *E. coli* and later in *P. pastoris*. Attempts made to attain extra-cellular secretion of Fibril filaments using *P. pastoris* were unsuccessful. Access to *Spiroplasma* cultures allowed us to attempt Fibril purification from *S. citri* cells.

In this study multiple protocols have been standardized for purification of Fibril filaments from *E. coli*. This work, for the first time, reports heterologous expression and protocols for purification of *Spiroplasma* Fibril filaments in *E. coli*. The standardization of Fibril purification from a heterologous expression system helps overcome the difficulties associated with gene manipulation in *Spiroplasma* to obtain higher quantity of protein required for structural studies.

Initially, protocol was standardized for overexpression of Fibril and for optimization of the amount of protein obtained in the soluble form upon cell lysis. Fibril filaments expressed in *E. coli* were first partially purified using glycerol density gradient. Most of the Fibril clumps could not be separated on glycerol gradient. The most probable reason for this is the inability of glycerol gradient [50% (w/v); density = 0.63 g/cm^3] to separate Fibril filaments due to the higher density [1.22 g/cm^3; (Townsend et al., 1980)] of latter. Further, the low density of glycerol could not separate Fibril from associated proteins. Nevertheless, the little quantity of Fibril that remained in top fractions of glycerol gradient probably formed thinner and less dense filaments which were visualized using electron microscopy. Since high quantity of purified protein is required for structural studies, a different approach was tried.

Affinity chromatography is one of the widely used techniques for purification of proteins of interest. It is known to help obtain pure protein in quantities sufficient for structural studies. Thus, attempts were made to purify N or C terminally hexahistidine tagged Fibril using Ni-NTA resins. However, the hexahistidine tagged Fibril did not bind to Ni-NTA resins in presence of Triton X-100 and/or LSS. The pelleted Fibril were solubilized only upon treatment with SDS and purified using affinity chromatography. It therefore appears that the SDS separated the contaminant proteins from Fibril (Figure 2.5 A), perhaps by denaturation of contaminants. It is important to note that out of the Fibril purified by different protocols from *E. coli*, only the SDS-treated bound to Ni-NTA resins through a His_6 tag. Unlike monomeric protein that needs exposed His_6 tag in every molecule to bind Ni-NTA, even a single exposed His_6 tag from a filament could possibly retain it on the affinity matrix.

A completely different possibility is that Fibril binds to membrane lipids through its N and C terminal residues. Addition of SDS may cause separation of Fibril from amphipathic molecules, thereby exposing its His_6 tag at the terminus. The exposed His_6 tag proved useful in Fibril purification using Ni-NTA matrix. However, the interaction between SDS and membrane lipids is not well characterized and hence, it should only be considered a speculation at this stage.

Both the aforementioned possibilities suggest that SDS may have affected Fibril filaments in a way that exposed its His_6 tag. The SDS-treated Fibril did not pellet down upon spinning at 159,000 xg, a property often used for separation of protein polymers from non-polymeric proteins. This led me to speculate if Fibril filaments were denatured to monomers or short oligomers. Visualization of the filaments using electron microscopes revealed that Fibril was present as bundle of filaments in the purified fractions (Figure 2.5 B-D). Thus, it is clear that Fibril was indeed present as filaments upon SDS treatment. Since Fibril filaments tend to clump up *in vitro*, the effect of denaturation by SDS may be limited to the molecules in the periphery of Fibril bundles. Exposure of His_6 tags from a few molecules denatured by SDS in a bundle can help the bundle to remain attached to affinity matrix. The treatment of Fibril filaments with SDS may have significantly reduced their density as compared to high density assemblies that can be formed on membranes. Thus, low density Fibril bundles formed due to SDS treatment did not pellet down upon spinning at 159,000 xg. However, further experiments must be carried out to test this hypothesis.

In order to avoid use of SDS-treated protein for structural characterization, SDS-free protocols were standardized. The major modification between the new and above mentioned approach is the separation of Fibril using urografin, a material with density (1.52 g/cm^3) higher than Fibril. The use of urografin enabled me to obtain relatively pure Fibril filaments from *E. coli*, without the need for SDS. The densities of the gradient at the end of spin were expected to be 1.15 g/ cm^3 and 1.52 g/cm^3 at the top and bottom respectively. This being a linear gradient, a decrease in density by 0.037 g/cm^3 per mL is expected. Since the reported density of Fibril bundles is 1.22 g/cm^3 (Townsend et al., 1980), they were expected to be present in Fraction number 3-4 from top. However, Fibril was obtained in my experiments in fraction number 7. This may mean that the gradient may not be linear or that the density of Fibril purified by this protocol is different than reported. Thus, it becomes necessary to verify the density of fraction containing Fibril by

alternate means such as estimation of refractive index. Nevertheless, visualization of Fibril purified by this protocol using electron microscopy confirms that the Fibril assemblies obtained from *Spiroplasma* cells are similar to the ones obtained from *E. coli*.

Electron microscopy observations suggested the presence of clumps of unknown composition with Fibril bundles in spite of avoiding detergents and glycerol for purification (Figures 2.7 D, 2.8 A to C and 2.9 B). Since the clumps did not have any features, like those seen for proteins, it is highly likely that they are clumps of membrane lipids. Extraction of Fibril samples using chloroform-methanol mixture led to reduction in number of clumps associated with Fibril bundles (Figure 2.9 C and D). This supports the notion that the clumps are made of lipids. Since no external lipids or detergents were used during purification, these lipids must be from *E. coli* cell membranes. This is also supported by my experiments involving purification of Fibril solubilized using LSS as the only detergent.

Soluble Fibril filaments obtained by lysis of *E. coli* cells in presence of LSS did not pellet down even when spun at 100,000 xg (Figure 2.6 A). However in absence of any detergents most of the Fibril from clarified lysate of *E. coli* cells pelleted at 30,000 xg (Figure 2.9 A). Thus, it is possible that LSS separated Fibril from *E. coli* membrane since LSS is known to disrupt cytoplasmic membrane of *E. coli* (Filip et al., 1973). The soluble Fibril bundles so formed in presence of LSS may not be dense enough to pellet down at 100,000 xg (Figure 2.6 A). Attempts to purify these bundles using various HIC and ion exchange chromatography failed since the protein did not bind to matrix of any of these columns. The possible explanation for this is, if Fibril is a membrane binding protein, its separation from the membrane would compel it to bind the detergents to stabilize itself. Since LSS was used at a concentration (0.5 g/ 100 mL; 17 mM) higher than its CMC (0.42 g/100 mL; 14.6 mM), it is also possible that LSS remained bound to Fibril bundles as micelles. In such a case, the charged surface of the protein might be inaccessible and will not bind to any of the charged chromatography matrices. An alternate possibility for Fibril not binding to column matrices is the exclusion of Fibril filaments (about >20 nm wide and several µm in length as observed in EM; Figures 2.1 D and 2.3 B) from entering chromatography matrix (beads of size <100 µm) due to their long lengths. In such a case, the Fibril filaments would come out of the column without interacting with the column matrix.

Triton X-100 and sodium deoxycholate detergents are used for purification of Fibril as detergent insoluble fraction since a long time (Townsend et al., 1980; Williamson et al., 1991). It is possible that these detergents extract Fibril by fragmenting *Spiroplasma* membrane with Fibril anchored to it. The reported density of Fibril (1.22 g/cm^3) is close to the density of cytoplasmic membranes of gram-positive [1.23 g/cm^3; (Dubochet et al., 1983)] and gram-negative [1.16 g/cm^3; (Osborn et al., 1972)] bacteria. The density of *Spiroplasma* lipids is unreported but not expected to be much different from 1.22 g/cm^3. Thus, it may not possible to separate membranes and Fibril using density gradient centrifugation method.

The data provided in this chapter is the first report of heterologous expression and purification of Fibril filaments using *E. coli* as the model organism. This work helps overcome the problems associated with importing a pathogen and tedious process of growing large cultures of *Spiroplasma*, a slow growing organism. High quantity of purified Fibril filaments required for structural studies can be obtained by above protocols. Further, my studies suggest that Fibril is likely to be a membrane–binding protein. However, further studies are necessary to provide conclusive evidence in support of this. Nevertheless, it supports the observation of Fibril filaments close to cell membrane by electron cryotomography studies (Kürner et al., 2005).

Irrespective of presence of detergent in lysis buffer, the purified Fibril bundles obtained from *E. coli* by all different methods described here possess random twists similar to the one observed in literature (Cohen-Krausz et al., 2011). This is a supporting evidence that the protein purified by protocols enlisted above is similar to the one obtained from *Spiroplasma* (Figure 2.13). In absence of a functional assay for Fibril, protein purified from *S. citri* will be useful as a control to check the structures of Fibril purified from *E. coli*. Experimental standardization is necessary to obtain uniformly distributed individual filaments from the bundle on electron microscopy grids for performing structural studies using cryoEM. Since there are no known ways to separate Fibril filaments from bundle, identification of condition suitable for obtaining separated filaments is likely to be a time consuming process and would require a constant access to TEM.

An alternative approach to this could be to obtain Fibril monomers for structure determination using X-ray crystallography. Since Fibril forms constitutive filaments, obtaining its monomers or short oligomers is impossible. Based on current model for role of Fibril in *Spiroplasma* motility, Fibril is proposed to be a bi-domained protein (Cohen-Krausz et al., 2011). Studying each

domain of Fibril may provide clues regarding their function. It may also prove useful to know if both the domains are involved in polymerization. If the residues involved in polymerization are identified, polymerization mutant of Fibril can be obtained for X-ray crystallography studies. However, the domain boundaries of Fibril have not been defined. Thus, the next chapter discusses approaches adopted for identification of domain boundaries of Fibril.

2.5) References

Buchanan, B.T.M. (1975). Antigenic heterogeneity of Gonococcal pili. J. Exp. Med. *141*.

Cheng, Y. (2015). Single-particle Cryo-EM at crystallographic resolution. Cell *161*, 450–457.

Cheng, Y., Grigorieff, N., Penczek, P.A., and Walz, T. (2015). A primer to single-particle Cryo-Electron Microscopy. Cell *161*, 438–449.

Chung, J.M., and Jung, H.S. (2016). Advanced Cryo-Electron microscopy technology: high resolution structure of macromolecules. Appl. Microsc. *46*, 1–5.

Chung, B.H., Wilkinson, T., Geer, J.C., and Segrest, J.P. (1980). Preparative and quantitative isolation of plasma lipoproteins : rapid, single discontinuous density gradient ultracentrifugation in a vertical rotor. J. Lipid Res. *21*, 284–291.

Citti, C., Marechal-drouard, L., Saillard, C., Weil, J.H., and Bove, J.M. (1992). *Spiroplasma citri* UGG and UGA tryptophan codons: Sequence of the two tryptophanyl-tRNAs and organization of the corresponding genes. J. Bacteriol. *174*, 6471–6478.

Cohen-Krausz, S., Cabahug, P.C., and Trachtenberg, S. (2011). The monomeric, tetrameric, and fibrillar organization of Fib: the dynamic building block of the bacterial linear motor of *Spiroplasma melliferum* BC3. J. Mol. Biol. *410*, 194–213.

Daly, R., and Hearn, M.T.W. (2005). Expression of heterologous proteins in *Pichia pastoris*: A useful experimental tool in protein engineenring and production. J. Mol. Recognit. *18*, 119–138.

DeRosier, D.J., and Klug, A. (1968). Reconstruction of three dimensional structures from electron micrographs. Nature *217*, 130–134.

Desfosses, A., Ciuffa, R., Gutsche, I., and Sachse, C. (2014). SPRING - An image processing package for single-particle based helical reconstruction from electron cryomicrographs. J. Struct. Biol. *185*, 15–26.

Doerr, A. (2016). Single-particle cryo-electron microscopy. Nat. Methods *13*, 23.

Dubochet, J., McDowall, A.W., Menge, B., Schmid, E.N., and Lickfeld, K.G. (1983). Electron microscopy of frozen-hydrated bacteria. J. Bacteriol. *155*, 381–390.

Duret, S., Andre, A., and Renaudin, J. (2005). Specific gene targeting in *Spiroplasma citri*: improved vectors and production of unmarked mutations using site-specific recombination. Microbiology *151*, 2793–2803.

Egan, D.A., Logan, T.M., Liang, H., Matayoshi, E., Fesik, S.W., and Holzman, T.F. (1993). Equilibrium denaturation of recombinant human FK binding protein in urea. Biochemistry *32*, 1920–1927.

Egelman, E.H. (2015). Three-dimensional reconstruction of helical polymers. Arch. Biochem. Biophys. *581*, 54–58.

Ellis, S.B., Brust, P.F., Koutz, P.J., Waters, A.F., Harpold, M.M., and Gingeras, T.R. (1985). Isolation of alcohol oxidase and two other methanol regulatable genes from the yeast *Pichia pastoris*. Mol. Cell. Biol. *5*, 1111–1121.

Van Den Ent, F., and Löwe, J. (2006). RF cloning: A restriction-free method for inserting target genes into plasmids. J. Biochem. Biophys. Methods *67*, 67–74.

Filip, C., Fletcher, G., Wulff, J.L., and Earhart, C.F. (1973). Solubilization of the cytoplasmic membrane of *Escherichia coli* by the ionic detergent Sodium-lauryl sarcosinate. J. Bacteriol. *115*, 717–722.

Giancola, C., De Sena, C., Fessas, D., Graziano, G., and Barone, G. (1997). DSC studies on bovine serum albumin denaturation effects of ionic strength and SDS concentration. Int. J. Biol. Macromol. *20*, 193–204.

Kürner, J., Frangakis, A.S., and Baumeister, W. (2005). Cryo–electron tomography reveals the

cytoskeletal structure of *Spiroplasma melliferum*. Science. *307*, 436–439.

Mizuno, K. (1992). Induction of cold stability of microtubules in cultured tobacco cells. Plant Physiol. *100*, 740–748.

Mohanty, J.G., and Elazhary, Y. (1989). Purification of IgG from serum with caprylic acid and ammonium sulphate precipitation is not superior to precipitation alone. Vine *12*, 153–160.

Osborn, M.J., Gander, J.E., Parisi, E., and J, C. (1972). Mechanism of assembly of the outer membrane of *Salmonella typhimurium*. J. Bacteriol. *247*, 3962–3972.

Patsch, J.R., and Patsch, W. (1986). Zonal ultracentrifugation. In Methods in Enzymology, pp. 3–26.

Paul, D., Patwardhan, A., Squire, J.M., and Morris, E.P. (2004). Single particle analysis of filamentous and highly elongated macromolecular assemblies. J. Struct. Biol. *148*, 236–250.

Privé, G.G. (2007). Detergents for the stabilization and crystallization of membrane proteins. Methods *41*, 388–397.

Redgrave, T.G., Roberts, D.C.K., and West, C.E. (1975). Separation of plasma lipoproteins by density-gradient untracentrifugation. Anal. Biochem. *65*, 42–49.

Reimer, C.B., Baker, R.S., VanFrank, R.M., Newlin, T.E., Cline, G.B., and Anderson, N.G. (1967). Purification of large quantities of influenza virus by density gradient centrifugation. J. Virol. *1*, 1207–1216.

Rødahl, E. (1998). Ultracentrifugation.

Seddon, A.M., Curnow, P., and Booth, P.J. (2004). Membrane proteins, lipids and detergents: Not just a soap opera. Biochim. Biophys. Acta - Biomembr. *1666*, 105–117.

Shaevitz, J.W., Lee, J.Y., and Fletcher, D.A. (2005). *Spiroplasma* swim by a processive change in body helicity. Cell *122*, 941–945.

Soellner, P., Quinlan, R.A., and Franke, W.W. (1985). Identification of a distinct soluble subunit of an intermediate filament protein : tetrameric vimentin from living cells. Proc. Natl. Acad. Sci.

U. S. A. *82*, 7929–7933.

Townsend, R. (1983). *Spiroplasma* fibrils. Yale J. Biol. Med. *56*, 447–452.

Townsend, R., Archer, D.B., and Plaskitt, K.A. (1980). Purification and preliminary characterization of *Spiroplasma* fibrils. J. Bacteriol. *142*, 694–700.

Trachtenberg, S., and Gilad, R. (2001). A bacterial linear motor: Cellular and molecular organization of the contractile cytoskeleton of the helical bacterium *Spiroplasma melliferum* BC3. Mol. Microbiol. *41*, 827–848.

Trachtenberg, S., Dorward, L.M., Speransky, V. V., Jaffe, H., Andrews, S.B., and Leapman, R.D. (2008). Structure of the cytoskeleton of *Spiroplasma melliferum* BC3 and its interactions with the cell membrane. J. Mol. Biol. *378*, 778–789.

Trachtenberg, S., Schuck, P., Phillips, T.M., Andrews, S.B., and Leapman, R.D. (2014). A structural framework for a near-minimal form of life: mass and compositional analysis of the helical *Mollicute Spiroplasma melliferum* BC3. PLoS One *9*, e87921.

Williamson, D.L., Renaudin, J., and Bove, J.M. (1991). Nucleotide sequence of the *Spiroplasma citri fibril* protein gene. J. Bacteriol. *173*, 4353–4362.

Wingfield, P.T. (2016). Protein precipitation using ammonium sulfate. Curr. Protoc. Protein Sci. *2016*, A.3F.1-A.3F.9.

Woodring, P.J., Hunter, T., and Wang, J.Y.J. (2001). Inhibition of c-Abl tyrosine kinase activity by filamentous actin. J. Biol. Chem. *276*, 27104–27110.

Chapter 3

Identification of domain boundaries of Fibril

3.1) Introduction

Conformational changes in a protein can be identified through structural studies. X-ray crystallography, nuclear magnetic resonance (NMR) and electron cryo microscopy (cryoEM) are the widely used techniques in structural biology. While X-ray crystallography is used for proteins that are amenable to crystallization, structures of macromolecular complexes and filamentous proteins that resist crystallization can be studied using cryoEM. A constitutively polymerizing protein, such as Fibril, cannot be crystallized and hence is not suitable for X-ray crystallographic studies unless a non-polymerizing construct is obtained. The X-ray crystal structure of a non-polymerizing construct will provide the three dimensional atomic arrangement of the monomeric conformation of the protein. This can be compared to the filamentous conformation to provide information on the conformational changes accompanying filament assembly.

Since there are no known means of controlling Fibril polymerization, the filament architecture can be deciphered by helical reconstruction of Fibril filaments, if helicity exists in the filaments. While efforts for structure determination of the filaments were in progress (Chapter 2), approaches for obtaining a non-polymerizing construct of Fibril were initiated. The non-polymerizing construct can be used to obtain structural information by X-ray crystallography studies. A structure of the Fibril monomer will be useful to understand filament assembly. Comparing the conformation of Fibril monomer within the filament and the crystal structure of monomeric Fibril will provide insights into the conformational changes accompanying polymerization.

Since Fibril forms constitutive, nucleotide-independent filaments that cannot be depolymerized without denaturation, obtaining Fibril monomers is challenging. In absence of the structure of a filament and since there are no identified homologs of Fibril that are known to polymerize, information on the polymerization interface required for design of mutants is not available. Identification of domain boundaries of Fibril can help in dissection of Fibril into domain-wise soluble constructs that do not polymerize. These constructs can be used for X-ray

crystallography studies. Hence, we initiated an approach of domain dissection of Fibril to identify non-polymerizing constructs and use them for X-ray crystallography studies.

Domain is a portion of protein that can exist and function independently of the rest of the protein chains (Richardson, 1981). Using electron cryo microscopy, modelling and sequence analysis of Fibril filaments, Cohen-Krausz et al., (2011) predicted it to be a bi-domained protein. Based on their analysis, MTAN (5'-methyladenosine/S-adenosylhomocysteine nucleosidase) from *Arabidopsis thaliana* (3BSF.PDB) was identified as a domain that shares 29.1% sequence identity with 1-234 amino acids from the N-terminus of Fibril. Thus, it is proposed that approximately 234 amino acids from N-terminus of Fibril form one domain. The C-terminal region of Fibril (amino acids 235-512) does not have a homolog and thus is predicted to form a domain of novel fold. However, the exact domain boundaries of Fibril have not been defined or investigated yet. Understanding the domain boundaries will help us to investigate if the individual domains can polymerize on their own. Also, non-polymerizing domains can be used for crystallization studies.

This chapter describes my attempts to purify non-polymerizing constructs of Fibril to obtain structural information using X-ray crystallography. Attempts were also made to obtain a minimum/smallest construct to identify domain boundaries of Fibril. These studies led to identification of a smaller construct that can potentially polymerize independent of the rest of the protein.

3.2) Material and methods

Genetic modifications are tough to perform in *Spiroplasma*. Thus, *E. coli* [BL21(AI) cells with genotype: *F-ompT hsdSB (rB- mB-) gal dcm araB::T7RNAP-tetA*] were used as the heterologous expression system to express and purify shorter constructs of Fibril. Following sections describe the methodology for investigation of domain organization in Fibril.

3.2.1) Basis for design of short constructs

In order to investigate the domain boundaries of Fibril, multiple shorter constructs were designed. The basis of designing these shorter constructs is explained below.

i) FibrilΔ(237-512)-His$_6$ and FibrilΔ(1-236)-His$_6$ constructs: An investigation of the proteins with sequences similar to *S. citri* Fibril protein using NCBI BLAST (https://blast.ncbi.nlm.nih.gov/Blast.cgi) resulted into identification of MTAN. Two short, complementary constructs from the N and C terminus of Fibril were prepared based on the alignment of Fibril and MTAN (from *E. coli*; 1JYS.PDB) sequence (Figure 3.1 A). *E. coli* MTAN shares 23 % sequence identity with the N-terminal 236 residues (1-236) of Fibril. Thus, FibrilΔ(237-512)-His$_6$ was prepared considering it as one domain and the rest

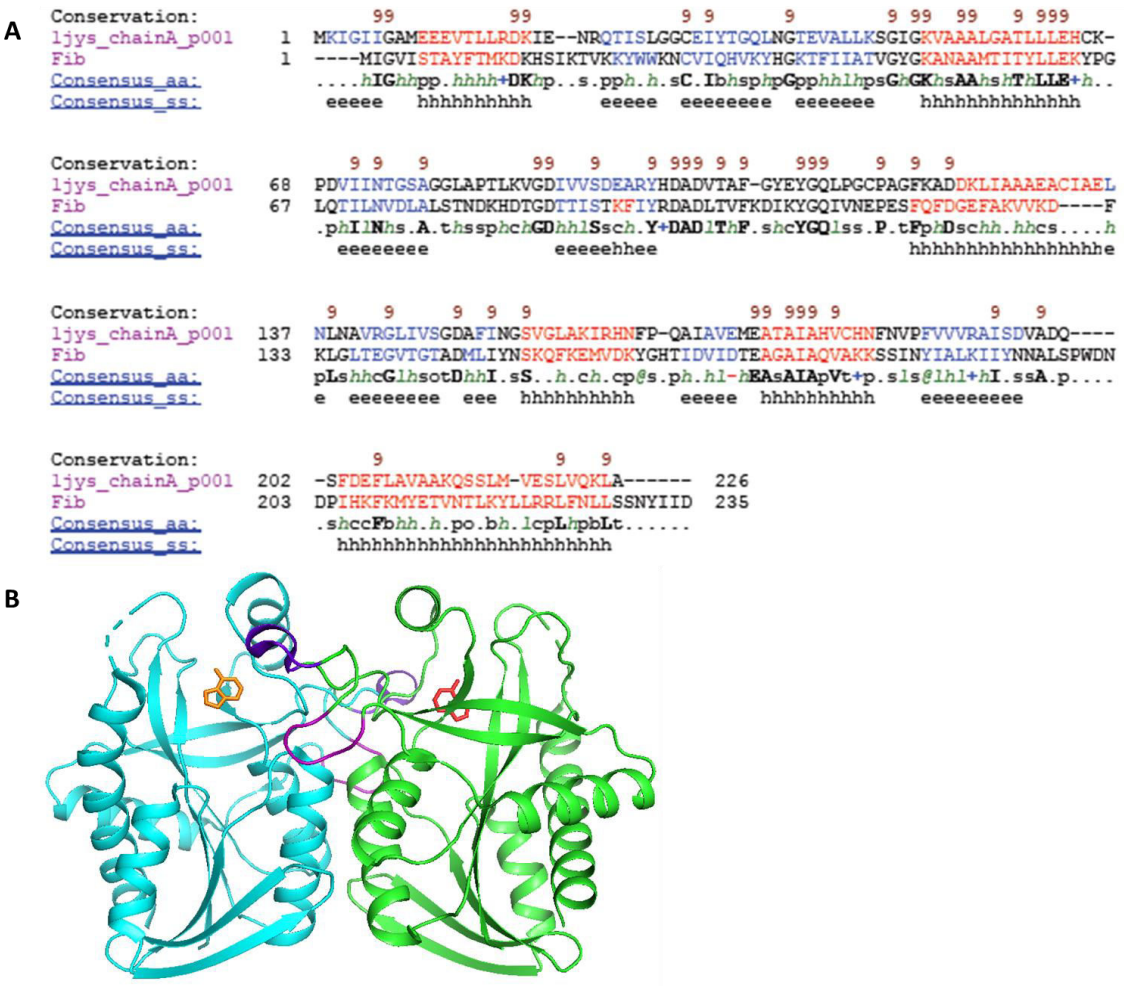

Figure 3.1 MTAN, a basis for design of short constructs of Fibril. A) Sequence alignment of Fibril and MTAN (PDB ID: 1JYS) showing conserved residues (conservation score of 9; consensus amino acids shown in bold, capital letters) and secondary structures. B) Crystal structure of dimeric MTAN (PDB ID: 1JYS) is shown as cartoon representation with loops participating in dimer formation highlighted in purple (amino acids 112-116) and purple-blue (amino acids 103-109). Bound adenine molecules are shown as stick representation and highlighted in orange and red. The sequence alignment was performed using PROMALS3D (http://prodata.swmed.edu/promals3d/promals3d.php).

[FibrilΔ(1-236)-His_6] as another domain.

ii) FibrilΔ(104-109) and FibrilΔ(112-116) constructs: Another two constructs were prepared based on alignment with MTAN. Since loops contribute to MTAN dimerization (Lee et al., 2001), Fibril constructs were prepared by deletion of the regions corresponding to the loops participating in MTAN dimerization (Figures 3.1 B). This could potentially disrupt polymerization if the dimer interface was necessary for polymerization.

iii) His_6-FibrilΔ(305-512), Strep-Fibril(290-479)-His_6 and Strep-FibrilΔ(1-301)-His_6 constructs: His_6-FibrilΔ(305-512) was prepared based on the observation of approximately 37 kDa bands on SDS-PAGE gels during purification of full-length Fibril (Figure 2.1 A and 2.3 A). Strep-Fibril(290-479)-His_6 and Strep-FibrilΔ(1-301)-His_6 were prepared based on observation of about 27 kDa band pairs during purification of full-length Fibril as well as FibrilΔ(1-236)-His_6. These bands were prominent even during purifications using SDS treatment protocol. Since an increase in intensity of these bands was observed with time during purification of FibrilΔ(1-236)-His_6 and full-length Fibril, these were hypothesized to be proteolysis products of full-length Fibril. Streptavidin (Strep) and His_6 tags were introduced at the two termini in two of these constructs to enable purification of the whole construct, and to remove impurities from the first step of purification.

iv) His_6-FibrilΔ(285-512), FibrilΔ(273-512)-His_6, His_6-FibrilΔ(253-512) and FibrilΔ(228-512)-His_6 constructs: Based on observation of filaments/protein polymers formed by one of the short constructs His_6-FibrilΔ(305-512), these shorter constructs were prepared to identify non-polymerizing constructs of Fibril. The characterization of these constructs was also thought to be useful for identification of minimum construct of Fibril that polymerizes independent of the rest of the protein.

Due to lack of structural information on Fibril, secondary structure prediction (Figure 3.2) was used as the basis for designing shorter constructs. While designing constructs care was taken to not dissect the protein in between a predicted secondary structure (α-helix and β-sheet).

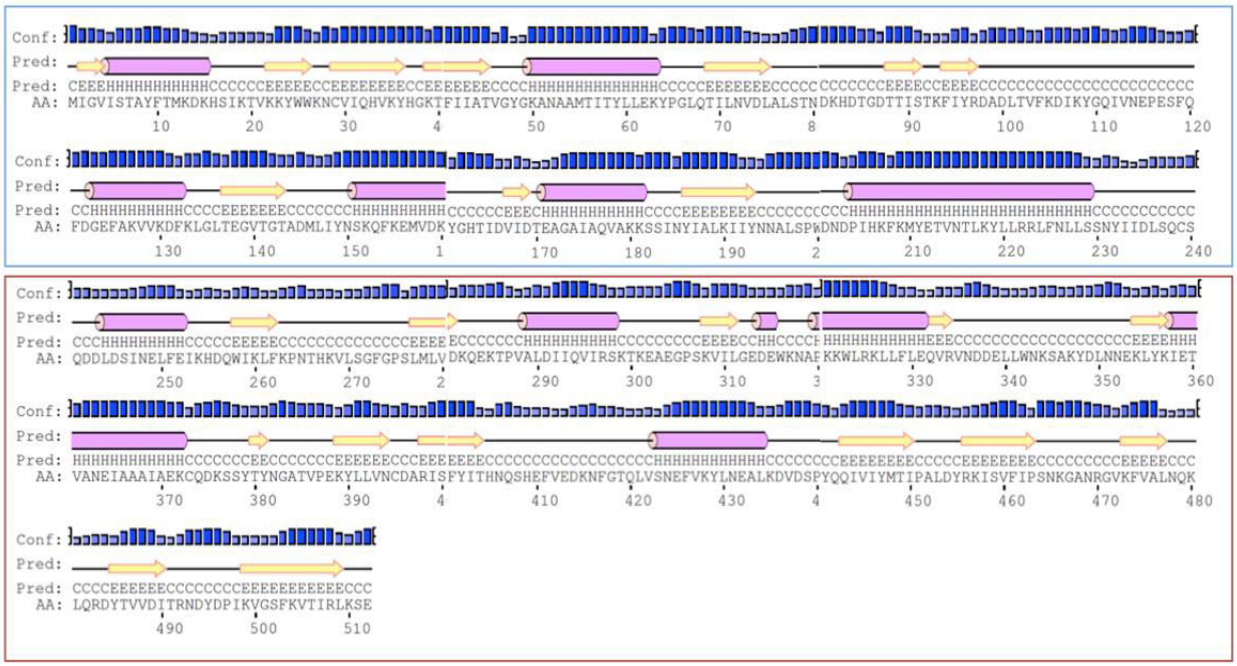

Figure 3.2 Secondary structure prediction of Fibril. Secondary structures in Fibril were predicted by using Psipred (http://bioinf.cs.ucl.ac.uk/psipred/). The regions predicted to form α helices, β sheets & coils have been shown as purple cylinders, yellow arrows and black lines respectively. The confidence of prediction for each residue is shown as blue bars above predicted secondary structures. The blue and red boxes represent the proposed N-terminal and C-terminal domains respectively.

3.2.2) Cloning

All the clones used in this chapter were prepared either by Restriction-Free cloning (RFC) or by restriction digestion-ligation method as described in sections 2.2.1.2 and 2.2.1.3 of Chapter 2 in this thesis. All the genes of insert (amplified from codon-optimized *fibril* gene described in Chapter 2 of this thesis) used in this chapter were cloned into pHIS17 vector and preliminary check of the clones performed by restriction digestion of the clones using NdeI and BamHI enzymes. All the clones were confirmed by sequencing and used for expressing proteins of interest. List of primers used for cloning all the short constructs is presented in Table 3.1.

Table 3.1 List of primers used for cloning short constructs of Fibril

Primer name	Sequence (5'→3')	Purpose
ScF-f	GTTTAACTTTAAGAAGGAGATATACATATGATCGGTGTAATTTCAACAGCATATTTTACG	Preparation of N terminal constructs without a tag at N-terminus
ScF-NHis-f	CTTTAAGAAGGAGATATACATATGCGTGGCCATCATCATCATCATCATATCGGTGTAATTTCAAC	Preparation of N terminal constructs with a N-terminal His_6 tag
ScF-r	GCTTTTAATGATGATGATGATGATGGGATCCATCATTACTCACTTTTTAAACGAATTGTAACTTTG	Preparation of C terminal constructs without a tag at C-terminus
ScF-CHis-r	GCTTTTAATGATGATGATGATGATGGGATCCATCATTCTCACTTTTTAAACGAATTGTAACTTTG	Preparation of C terminal constructs with a C-terminal His_6 tag
ScF-N304-r	GCTTTTAATGATGATGATGATGATGGGATCCTTCTGCTTCTTTTGTTTTTGAACG	Preparation of His_6-Fibril Δ(305-512) construct
ScF-N284_No tag	GCTTTTAATGATGATGATGATGATGGGATCCTTATTCTTGTTTATCAACTAAC	Preparation of His_6-FibrilΔ(285-512) construct
ScF-N272-His-r	GCTTTTAATGATGATGATGATGATGGGATCCTCCTGATAATACTTTATGTG	Preparation of FibrilΔ(273-512)-HIS_6 construct
ScF-N252-r	GCTTTTAATGATGATGATGATGATGGGATCCATTATTCAAATAATTCATTAATTG	Preparation of His_6-FibrilΔ(253-512) construct
ScF-Nterm236-r	GATGATGGGATCCATCATTATGATAAATCAATAATATAAT	Preparation of FibrilΔ(237-512)-His_6 construct
ScF-N228-His-r	GCTTTTAATGATGATGATGATGATGGGATCCTAAGTTAAATAATCTTCTTAATAAG	Preparation of FibrilΔ(228-512)-His_6 construct
ScF-Cterm235-f	TAAGAAGGAGATATACATATGGATTTATCACAATGTTCAC	Preparation of FibrilΔ(1-236)-His_6 construct
pHis-Nd-Strp-Fib290-f	GTTTAACTTTAAGAAGGAGATATACATATGTGGTCTCATCCTCAGTTCGAAAAATTAGATATTATTCAAG	Preparation of Strep-Fibril(290-479)-His_6 construct
pHisFib479-His-r	GCTTTTAATGATGATGATGATGATGGGATCCTTGATTTAAAGCCACAAATTTAAC	
Strp-Nd-Fib302-f	CTCATCCTCAGTTCGAAAAACATATGGAAGCAGAAGGACCAAG	Preparation of Strep-FibrilΔ(1-301)-His_6 construct
pHis-Nd-Strp-Fib302-f	GTTTAACTTTAAGAAGGAGATATACATATGTGGTCTCATCCTCAGTTCGAAAAAGAAGCAGAAGGACC	
Scf-del104-109-f	GATGCTGACTTAACAGTTGGACAAATTGTTAATGAACC	Preparation of FibrilΔ(104-109) construct
Scf-del112-116-f	GATATTAAATATGGACAAGAATCATTCCAATTCGATGG	Preparation of FibrilΔ(112-116) construct

3.2.3) Expression of short constructs of Fibril

The protocol described in section 2.2.2.1 of Chapter 2 in this thesis was followed for expression of all the constructs. Parameters such as optical density before induction, post-induction incubation temperature and time were standardized to obtain maximum over-expression and also to get maximum protein in soluble fraction. BL21(AI) cells were used for expression of all the constructs. Thus, all the cells were induced using sterile L-arabinose at a final concentration of 0.2 % (w/v).

3.2.4) Pelleting/polymerization assay

Pelleting assays for N-His full length Fibril, His_6-FibrilΔ(305-512), His_6-FibrilΔ(285-512), FibrilΔ(273-512)-His_6, His_6-FibrilΔ(253-512) constructs and untransformed BL21(AI) cells were performed by re-suspending a 400 mL culture pellet in ice cold 50 mL lysis buffer 8 (refer Appendix I for composition). Cells were lysed by sonication using a probe sonication with 60 % amplitude, 1 second ON, 3 seconds OFF for a total time of 18 minutes. EDTA was added in the lysate at a final concentration of 1 mM to depolymerize nucleotide-dependent polymers of proteins. The lysate was initially spun at 14,636 xg/4 ºC/15 minutes to remove cell debris as pellet. The soluble fraction was spun at 46,500 xg/4 ºC/30 minutes to pellet down Fibril filaments. Pellet was re-suspended in 1.8 mL $T_{10}E_1$ [10 mM Tris (pH 7.6), 1 mM EDTA] buffer and divided equally into 2 tubes. 100 µL of SDS (10 % w/v) solution was added to one of the tubes while 100 µL $T_{10}E_1$ was added to the other tube and contents mixed. Both the tubes were spun at 21,150 xg/25 ºC/15 minutes to pellet down insoluble aggregates. The soluble fractions were transferred to a fresh tube and spun at 159,000 xg for 15 minutes at 4 ºC (sample without SDS) or 25 ºC (sample with SDS) to pellet down filamentous proteins. The supernatant was transferred to a fresh tube while the pellet was re-suspended in 20 µL distilled water. 10 µL sample was taken at each step during purification and mixed with 10 µL SDS-PAGE loading dye (2X). Samples were heated at 99 ºC/10 minutes/400 rpm, spun briefly and 10 µL loaded on a 12 % SDS-PAGE gel. Samples in the gel were stained and then de-stained for visualization of protein bands.

3.2.5) Purification and characterization of short constructs

Purification of short constructs of Fibril was attempted using a variety of techniques. The standardized protocols are described below. The nomenclature of all the short constructs is presented in Table 3.2.

Table 3.2 Nomenclature of constructs of Fibril and their details.

Name of the Fibril construct	Molecular weight (kilo dalton)	pI
His_6-FibrilΔ(305-512)	35	8.24
His_6-FibrilΔ(285-512)	33	7.77
FibrilΔ(273-512)-His_6	32	8.27
His_6-FibrilΔ(253-512)	30	6.86
FibrilΔ(237-512)-His_6	28	8.83
FibrilΔ(228-512)-His_6	26.9	8.99
FibrilΔ(1-236)-His_6	33	7.28
His_6-Fibril	59.6	8.44
Fibril-His_6	59.6	8.44
Fibril	58.7	8.44
FibrilΔ(104-109)	57.8	8.27
FibrilΔ(112-116)	58.1	8.57
Strep-FibrilΔ(1-301)-His_6	26	7.27
Strep-Fibril(290-479)-His_6	24	7.13

3.2.5.1) Purification of FibrilΔ(1-236)-His_6

A 6 litre culture pellet of *E. coli* cells expressing FibrilΔ(1-236)-His_6 was thawed on ice and re-suspended in 100 mL ice cold lysis buffer 9 supplemented with 25 mM imidazole. A protease inhibitor tablet (cOmplete™, mini, EDTA-free Protease inhibitor cocktail, Sigma-Aldrich) was added to the re-suspended pellet and mixed. The culture was lysed on ice by sonication using a probe sonicator (VibraSonic) for 8 minutes with 5 seconds ON and 5 seconds OFF cycles and 60 % amplitude. The lysate was spun at 159,000 xg/4 ºC/40 minutes to get rid of cell debris and polymeric proteins in the form of pellet. Simultaneously, a 5 mL Ni-NTA column (GE Healthcare Life Science) was equilibrated with binding buffer [50 mM Tris (pH 8.0), 200 mM NaCl] using an Äkta Prime system in a cold cabinet (4 ºC). Supernatant, from the 159,000 xg

spin, containing protein of interest was loaded on to the pre-equilibrated Ni-NTA column using Äkta Prime system. Unbound protein was collected as flowthrough. The column was washed with 30 mL (5 column volumes) binding buffer to remove any poorly bound or unbound proteins and collected as wash. Buffer with increasing concentrations of imidazole (25 mM, 50 mM, 100 mM, 250 mM, 500 mM) was then passed through the column to elute bound proteins. 8 column volumes (CV) each of 25 mM, 50 mM were loaded on the column and elution collected as 40 mL fractions each. Subsequently, elution buffer with 5 CVs each of 100 mM, 250 mM, 500 mM imidazole concentration was passed through the column and collected as 5 mL fractions. EDTA and DTT were added at final concentrations of 1 mM each to the tubes into which fractions were collected. 10 µL each of the sample was mixed with 10 µL SDS-PAGE loading dye (2X) and 10 µL visualized on a 12 % SDS-PAGE gel.

Fractions containing protein of interest were then pooled and 2 rounds of dialysis were carried out against 2 litre of buffer A_{50} [50 mM Tris (pH 8.0), 50 mM NaCl] for 2 hours each at 4 ºC. Upon completion of dialysis the samples were spun at 27,200 xg/4 ºC/15 minutes to remove any precipitate. The clear supernatant was loaded on to a 5 mL anion-exchange chromatography column (Q-HP; GE Healthcare Life Science) pre-equilibrated with buffer A_{50}. Unbound proteins were collected as flowthrough. 5 CVs of binding buffer (buffer A_{50}) was passed through the column to remove any unbound and loosely bound proteins and collected as wash. A linear gradient of elution buffer [50 mM Tris (pH 8.0), 1000 mM NaCl] from 0 % to 50 % against buffer A_{50} with a total volume of 100 mL was passed through the column. Both the binding buffer as well as elution buffer was supplemented with EDTA and DTT at a final concentration of 1 mM to prevent protein precipitation. Eluted contents were collected as separate fractions each of 0.5 mL. Sample were taken at each step during purification and checked on a 12 % SDS-PAGE gel. Fractions with protein of interest were pooled and dialyzed. Two rounds of dialysis were performed against 500 mL of buffer A_{50} [50 mM Tris (pH 8.0), 50 mM NaCl] for 2 hours each at 4 ºC. Dialyzed protein was concentrated using concentrators at 4 ºC.

The concentrated protein was used for gel filtration run to check oligomeric status of the protein as well as to obtain homogenous protein for further characterization. A Superdex™ 200 column was equilibrated using buffer A_{50} [50 mM Tris (pH 8.0), 50 mM NaCl], protein injected into the system and run performed using buffer with same composition. The eluted proteins were

collected as 0.5 mL fractions. 10 µL sample for the fractions corresponding to peak at 280 nm were mixed with equal volume of SDS-PAGE gel loading dye (2X). The sample and dye mixture was heated at 99 ºC/10 minutes/350 rpm and spun briefly. 10 µL each from the sample and dye mixture was loaded on 12 % SDS-PAGE gels and resolved. The gels were stained, de-stained and checked for presence of protein of interest. Fractions containing protein of interest were pooled and concentrated to minimum possible volume (100 µL). The concentration of protein was estimated and protein aliquoted as 20 µL fractions in thin-walled PCR tubes. The aliquoted tubes were flash-frozen in liquid nitrogen and stored at -80 ºC until further use.

3.2.5.2) Purification of SDS-treated His_6-FibrilΔ(305-512) by glycerol density gradient centrifugation

A 400 mL culture pellet of *E. coli* cells expressing His_6-FibrilΔ(305-512) construct was thawed on ice and re-suspended in 100 mL lysis buffer 8 (Appendix 1). Cells were lysed by sonication for 16 minutes with 1 second ON, 3 seconds OFF and 60% amplitude. The lysate was initially spun at 7000 xg/4 ºC/15 minutes to separate cell debris as pellet. The supernatant containing Fibril was further transferred to a fresh tube and spun at 159,000 xg/4 ºC/30 minutes to obtain His_6-FibrilΔ(305-512) protein in pellet. The pellet was re-suspended in 1.9 mL buffer containing EDTA [$T_{10}E_1$; 10 mM Tris (pH 7.6) and 1 mM EDTA] to depolymerize proteins that form filaments in presence of nucleotides. Since the pellet could not be solubilized, 100 µL SDS solution (20 % w/v) was added at a final concentration of 1 % (w/v). Addition of SDS led to complete solubilization of protein. The solution was then spun at 21,000 xg/25 ºC/15 minutes to remove particulate matter as pellet. 1 mL of the supernatant was then loaded onto each pre-cast glycerol gradient. The details of glycerol gradient preparation are mentioned in Chapter 2.

Glycerol gradients with samples on top were then spun at 79,500 xg/4 ºC/5 hours. At the end of spin the gradients were fractionated (1.6 mL/fraction) using peristaltic pump as described in Chapter 1. 10 µL sample was taken at each step and also from each fraction and mixed with equal volume of SDS-PAGE loading dye (2X). The samples were heated at 99 ºC/ 10 minutes/ 350 rpm and spun briefly. 10 µL from each of the sample and dye mix was loaded onto a 12 % SDS-PAGE gel. Upon migration the gels were stained and de-stained to visualize proteins at different stages.

3.2.5.3) Purification of SDS-treated FibrilΔ(273-512)-His_6 using Ni-NTA affinity chromatography

A frozen 400 mL culture pellet of *E. coli* cells expressing truncated Fibril construct was thawed on ice. The pellet was re-suspended in 50 mL of lysis buffer 8 (Appendix 1). Cells were lysed by sonication using a probe sonicator for 18 minutes with cycles of 1 second ON, 3 seconds OFF with 60% amplitude. EDTA was added to the lysate at final concentration of 0.5 mM to prevent binding of proteins that bind Ni-NTA matrix non-specifically. The lysate was spun initially at 14,600 xg/4 ºC/15 minutes to remove cell debris as pellet. The soluble fraction containing protein of interest was further spun at 72,700 xg/4 ºC/30 minutes to separate Fibril filaments as pellet from other non-polymerized proteins. The pellet containing Fibril was re-suspended in 1.9 mL of $T_{10}E_1$ buffer [10 mM Tris (pH 7.6) and 1 mM EDTA]. 100 µL SDS (20 % w/v) solution was added to the re-suspension to solubilize the pellet. The solution was spun at 21,000 xg/25 ºC/15 minutes to remove insoluble particulate matter. The supernatant was further spun at 159,000 xg/25 ºC/15 minutes to separate smaller clumps as pellet. The supernatant containing Fibril was dialyzed three times against 2 litres of distilled water for 1 hour each time to remove SDS, EDTA and other salts.

The dialyzed protein was transferred to a fresh tube and Tris (pH 7.6) and NaCl were added at the final concentrations of 50 mM and 200 mM respectively. The protein was then mixed with Ni-NTA resins pre-equilibrated with binding buffer [50 mM Tris (pH 7.6) and 200 mM NaCl] and left for mixing on rotomix (Tarsons) at room temperature (25 ºC) for 3 hours. The mixing allows exposed His_6 tag of the proteins to bind to Ni-NTA resins. Following incubation, the mixture was loaded into a syringe column (10 mL) and eluted as explained in Chapter 2.

Fractions containing purified protein were pooled and dialyzed thrice against 2 litre distilled water for 1 hour each time to remove salts and imidazole. The dialyzed protein was concentrated using concentrators with a cut-off of 3 kDa. Small quantity of the dialyzed and concentrated protein was used for visualization using FE-SEM. Rest of the concentrated protein was fractionated as 20 µL aliquots into thin-walled PCR tubes, flash frozen in liquid nitrogen and stored at -80 ºC until further use.

3.2.5.4) Purification of FibrilΔ(273-512)-His_6 without use of SDS

A frozen 2 litre culture pellet of *E. coli* cells expressing FibrilΔ(273-512)-His_6 construct was thawed on ice. The pellet was re-suspended in 200 mL lysis buffer 10. Cells were lysed by sonication using a probe sonicator for 48 minutes with cycles of 1 second ON, 3 seconds OFF with 60% amplitude. The lysate was spun at 46,500 xg/4 ºC/15 minutes to remove cell debris as pellet. The supernatant containing FibrilΔ(273-512)-His_6 was loaded into a pre-packed Ni-NTA column (5 mL; GE) equilibrated with binding buffer [50 mM Tris (pH 8.0) and 200 mM NaCl] using Äkta system in a cold cabinet. The unbound protein was collected as flowthrough. 40 column volumes (200 mL) of binding buffer was passed through the column to remove any unbound proteins and collected as wash. Bound protein was eluted from column using increasing concentrations of elution buffer containing imidazole [binding buffer + 500 mM imidazole]. 8 column volumes each were collected with 10 mM and 20 mM imidazole containing buffer. 6 column volumes each were collected for buffer containing 50, 100, 250 and 500 mM imidazole containing buffer. The eluted protein was collected as 5 mL fractions. 10 µL sample was taken at each step and mixed with equal volume of SDS-PAGE loading dye (2X). The mixture was heated at 99 ºC/10 minutes/350 rpm and spun briefly before loading on 12 % SDS-PAGE gel. The samples were migrated on gel and visualized after staining followed by de-staining.

Fractions containing protein of interest were pooled and spun at 39,139 xg /4 ºC/30 minutes for removal of precipitate, if any. The supernatant protein containing protein of interest was transferred into a dialysis bag with 3 kDa cut-off. The dialysis was performed against 2 litres of buffer A_{50} [50 mM Tris (pH 7.6), 50 mM NaCl] for 1 hour to remove salts and imidazole. The dialyzed protein was again spun at 39,139 xg /4 ºC/30 minutes to get rid of precipitated proteins during buffer exchange. The soluble protein was loaded onto a 1.7 mL cation-exchanger column (MonoS; GE) pre-equilibrated with buffer A_{50} [50 mM Tris (pH 7.6), 50 mM NaCl]. Since the theoretical pI (iso-electric point) of FibrilΔ(273-512)-His_6 is 8.27, it will have a net positive charge at pH 7.6. Also, the cation-exchanger column will be negatively charged at pH 7.6. The opposite charges on protein [FibrilΔ(273-512)-His_6] and matrix will facilitate their binding and will allow other proteins (with net negative charge) to pass through. The bound proteins can then be eluted by passing buffer with increasing salt concentrations to dissociate interaction between bound proteins and matrix. These steps facilitate purification of protein of interest.

Upon completion of loading of protein, the unbound protein was collected as flowthrough. The bound protein was eluted by passing 20 column volumes (34 mL) of linear gradient between 0 % and 100 % of elution buffer [50 mM Tris (pH 7.6), 500 mM NaCl] against binding buffer [50 mM Tris (pH 7.6), 50 mM NaCl]. The eluted buffer was collected as 1 mL fraction each. Samples from each step were checked on 12 % SDS-PAGE gel as described above.

The flowthrough (unbound proteins) from MonoS column was spun at 39,139 xg /4 ºC/30 minutes to remove precipitated proteins. Since the protein did not bind to MonoS column, it was passed through MonoQ column to check if can bind to the latter. Accordingly, the clear supernatant was loaded onto an anion exchange column (MonoQ; GE Healthcare) pre-equilibrated with buffer A_{50} [50 mM Tris (pH 7.6), 50 mM NaCl]. The unbound proteins were collected as flowthough and the bound proteins were eluted using the same buffers and protocol as described for MonoS assisted purification. The fractions were visualized on 12 % SDS-PAGE gels. The fractions containing protein of interest were pooled and concentrated using concentrators. The concentrated protein was checked on a 12 % SDS-PAGE gel.

3.2.6) Crystallization of FibrilΔ(1-236)-His_6

Frozen, purified protein in buffer A_{50} [50 mM Tris (pH 8.0), 50 mM NaCl] was thawed on ice. Aliquots of protein were pooled into a fresh tube and spun at 21,000 xg/4 ºC/30 minutes to remove any precipitate. The supernatant was transferred to a fresh tube and used for crystallization.

Commercial crystallization screens from Molecular Dimension, Jena Bioscience and Hampton were used for crystallization. A total of (96×9) 864 conditions were used for setting up an initial crystallization by sitting drop vapour diffusion method. The list of conditions with which crystallization of FibrilΔ(1-236)-His_6 was attempted is mentioned in Table 3.3. A crystallization drop of 200 nl (100 nl protein + 100 nl reservoir solution) was set up using the Mosquito robotic liquid handler. The plates were sealed and left at 18 ºC. Plates were monitored regularly for crystal formation by visualization of crystallization drops under a light microscope.

Table 3.3 List of conditions used for FibrilΔ(1-236)-His_6 crystallization

Plate number	Crystallization screen
1	Structure Screen 1 & 2 HT-96 MD1-30 (Molecular Dimensions)
2	Crystal Strategy Screen I MD1-31 (Molecular Dimensions)
3	Crystal Strategy Screen II MD1-32 (Molecular Dimensions)
4	MIDAS™ MD1-60 (Molecular Dimensions)
5	JBScreen Classic HTS II (Jena Bioscience)
6	JBScreen Classic HTS I (Jena Bioscience)
7	MD1-46 (MD1-47 HT) Morpheus™ Conditions (Molecular Dimensions)
8	JCSG-*plus* HT-96 MD1-40 (Molecular Dimensions)
9	Crystal Screen HT™ HR2-130 (Hampton Research Corp.)

3.2.7) Visualization of Fibril filaments using Field Emission-Scanning Electron Microscopy (FE-SEM)

To investigate if any of the short constructs formed filaments, purified proteins were used for visualization using FE-SEM. FE-SEM was the choice due to non-availability of a constant access to TEM and because the samples were stable in water without the presence of any salt or buffer conditions. The details of sample preparation for FE-SEM studies have been mentioned in Chapter 2 of this thesis. The same protocols as used visualizing full length Fibril were followed.

3.3) Results

Multiple short constructs of Fibril have been prepared to investigate domain architecture of Fibril and to identify domain boundaries. The details on logic for preparation of these constructs, results obtained from their purification and characterization are as follows.

3.3.1) Standardized conditions for expression of short constructs of Fibril

The optimized conditions for over-expression of various constructs and to obtain maximum protein in the solution are as presented in Table 3.4.

Table 3.4 Standardized conditions for expression of smaller constructs of Fibril

Name of the Fibril construct	Incubation parameters of primary culture	Optical density of secondary culture at the time of induction	Post-induction incubation parameters of secondary culture
FibrilΔ(1-236)-His_6	37 ^oC/10-12 hours	0.6 – 0.7	25 ^oC/ 5 hours
His_6-FibrilΔ(305-512)	37 ^oC/10-12 hours	0.8	25 ^oC/5 hours
(His_6 FibrilΔ(285-512)	37 ^oC/10-12 hours	1.0	25 ^oC/5 hours
FibrilΔ(273-512)-His_6	37 ^oC/10-12 hours	0.8 – 1.0	18 ^oC/12-16 hours
His_6-FibrilΔ(253-512)	37 ^oC/10-12 hours	0.8 - 1.0	25 ^oC/5 hours
FibrilΔ(237-512)-His_6	Yet to be standardized		
FibrilΔ(228-512)-His_6			
FibrilΔ(104-109)			
FibrilΔ(112-116)			
Strep-FibrilΔ(1-301)-His_6			
Strep-Fibril(290-479)-His_6			

Upon expression of the proteins, further steps of purification or polymerization assays were attempted for the constructs which were in the soluble fraction after removal of cell debris. In absence of any information about functional or biochemical assays to characterize purified short constructs, the characterization criteria were limited to comparing their sedimentation properties upon ultracentrifugation. The ability of smaller constructs to polymerize was studied by performing polymerization/pelleting assays, a centrifugation process for separating protein polymers from the non-polymerized form. The full-length Fibril that forms constitutive filaments and pellets down when spun at force of 46,500 xg was used as the positive control for pelleting assays. The possibility of pelleting down *E. coli* proteins with a molecular weight close to proteins of interest was ruled out by use of uninduced BL21(AI) cells as negative control for the pelleting assays.

3.3.2) Standardized protocols for purification attempts for short constructs of Fibril

Purification of multiple short constructs of Fibril was performed. The details of purity of protein and their further characterization for polymerization properties are as follows.

3.3.2.1) FibrilΔ(1-236)-His_6 is purified using affinity and ion exchange chromatography

Based on purification of full-length Fibril and also from literature, the filamentous proteins pellet down upon spinning at force > 80,000 xg (Mizuno, 1992; Woodring et al., 2001). The short construct FibrilΔ(1-236)-His_6 was obtained in the soluble fraction even after spinning at 159,000 xg and thus was thought to be a non-polymerizing construct. Further, purification by affinity chromatography helped obtain the protein in enriched form (Figure 3.3 A). However, it was

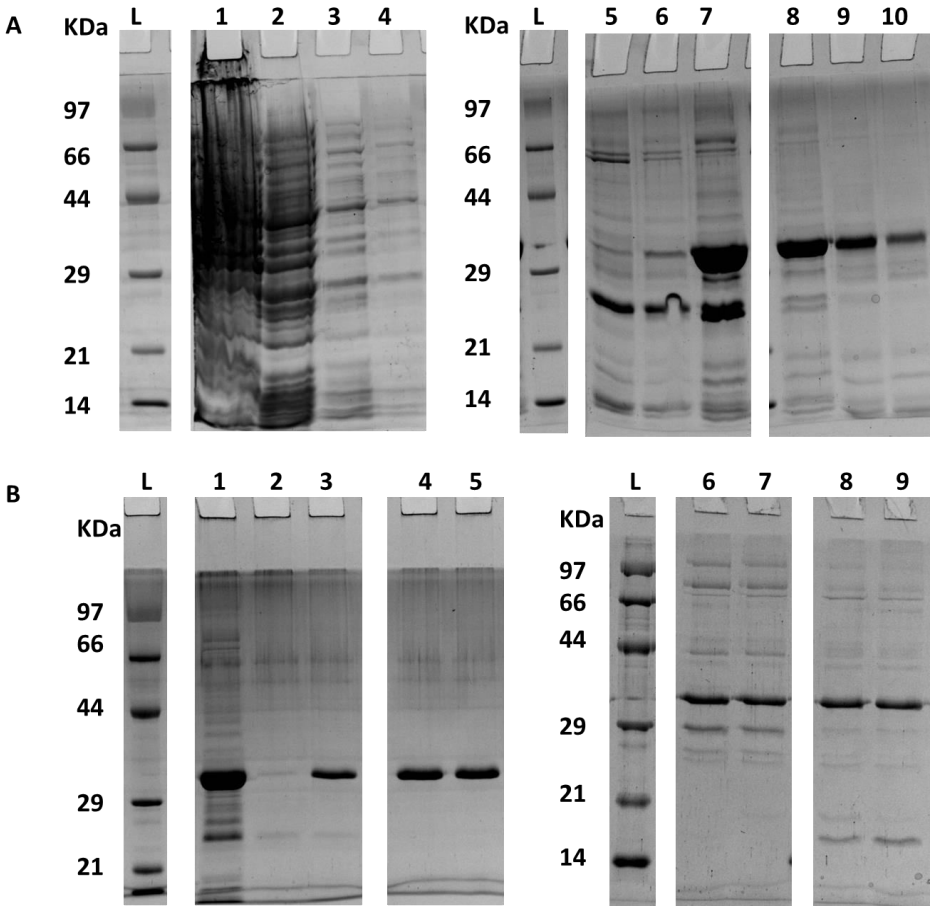

Figure 3.3 Purification of FibrilΔ(1-236)-His_6 (Mw 33 kDa). A) The protein was purified using Ni-NTA affinity chromatography. L- Protein marker; 1- total lysate; 2- supernatant after spinning at 159,000 xg loaded onto Ni-NTA column; 3- flowthrough; 4- wash; 5-10 fractions eluted using increasing concentrations of imidazole. B) The purity of protein was further improved by anion-exchange chromatography. L- Protein marker; 1- Input/protein loaded onto the column; 2- flowthrough; 3-9 fractions eluted from anion-exchanger column. Note the presence of impurities in fractions 6-9.

accompanied by unwanted proteins thus requiring further purification. Purification of the protein by anion exchange column led to further improvement in purity of the protein (Figure 3.3 B). Protein fractions occasionally contained unwanted bands with higher as well as lower molecular

weight. Most of the protein bands with molecular weight > 33 kDa, present in anion exchange column fractions with FibrilΔ(1-236)-His_6 (33 kDa) may be contaminants (Figure 3.3 B). These were separated from FibrilΔ(1-236)-His_6 using gel filtration chromatography. Gel filtration chromatography of purified protein revealed that the protein elutes at a volume of 15.69 mL (Figure 3.4 A & B). In a Superdex 200 (10/300 GL) column, this volume corresponds to a molecular weight (Mw) of about (33 kDa) (https://cdn.gelifesciences.com/dmm3bwsv3/AssetStream.aspx?mediaformatid=10061&destinationid=10016&assetid=16443). Thus, it was deduced that FibrilΔ(1-236)-His_6 protein (Mw 33 kDa) is present as monomer in solution.

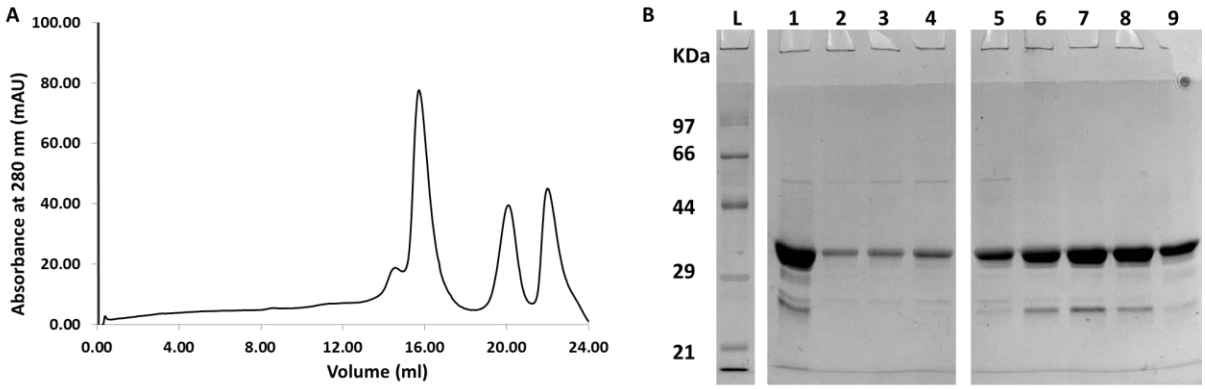

Figure 3.4 Gel filtration profile of FibrilΔ(1-236)-His_6 (Mw 33 kDa). A) The gel filtration profile of Fibril Δ(1-236)-His_6 construct on a Superdex 200 (10/300 GL) column. The protein eluted with peak centered at ~ 15.7 mL. B) Visualization of fractions corresponding to peak centered at 15.7 mL on a 12 % SDS-PAGE gel. L- Protein marker; 1- input/ protein loaded on the column; 2-9 fractions corresponding to the peak at 13.0 mL – 17.0 mL.

The FibrilΔ(1-236)-His_6 purified using anion exchange chromatography was also accompanied by low molecular weight proteins (about 29 and 25 kDa bands) (Figure 3.3 B). Since the intensity of these low molecular weight proteins seemed to increase with time (Figures 3.3 B and 3.4 B), they were thought to be proteolysis products of FibrilΔ(1-236)-His_6. Proteins with similar molecular weight were observed during purification of full-length Fibril filaments (Figures 2.2 A, 2.4 A-C and 2.5 A). This information was used for preparation of short constructs Strep-FibrilΔ(1-301)-His_6 (Mw 26 kDa) and Strep-Fibril(290-479)-His_6 (Mw 24 kDa).

Attempts were made to crystallize homogenous FibrilΔ(1-236)-His_6 protein obtained from a combination of chromatography techniques. However, no crystals could be obtained in the initial trials. Since the construct FibrilΔ(1-236)-His_6 could not be obtained in the soluble fraction in the subsequent trials, further work could not be carried out on it.

3.3.2.2) FibrilΔ(237-512)-His_6 is not be obtained in the soluble fraction

Since FibrilΔ(1-236)-His_6 was folded as evident from the gel filtration chromatography profile, it was thought to form the C-terminal domain of Fibril. Its complementary construct, FibrilΔ(237-512)-His_6, may form the N-terminal domain. Thus, simultaneously standardization of FibrilΔ(237-512)-His_6 was performed. However, the overexpressed protein could not be obtained in soluble fraction (Figure 3.5). Meanwhile, based on the information obtained from purification of full-length Fibril filaments, other N-terminal constructs were designed and purified to obtain insights into potential N-terminal domain of Fibril.

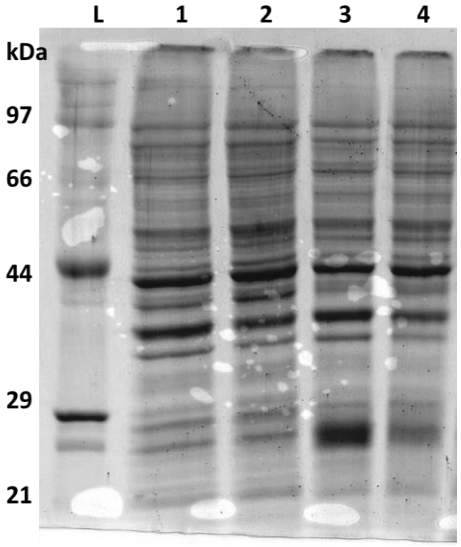

Figure 3.5 Expression profile of FibrilΔ(237-512)-His_6 (Mw 28 kDa). The protein of interest was over-expressed but could not be obtained in the soluble fraction. L- Protein marker; 1- total lysate of uninduced cells; 2- soluble fraction after spinning (1) at 21,000 xg; 3- total lysate of induced cells expressing protein of interest (28 kDa); 4- soluble fraction after spinning (3) at 21,000 xg. Note that an over-expression band can be seen in lane 3, just below 29 kDa. This band was absent in lanes 1 & 2 but small amount of it could be seen in lane 4.

3.3.2.3) His_6-FibrilΔ(305-512) is partially purified by SDS treatment and glycerol density gradient centrifugation

Hints about N-terminal domain were obtained from purifications of FibrilΔ(1-236)-His_6 and full-length Fibril. The observation of increasing intensity of about 27 kDa band pairs with time during FibrilΔ(1-236)-His_6 purification led me to postulate that the domain boundary could be around amino acid number 300 from N-terminus of Fibril. This was also consistent with purification of full-length Fibril where bands of about 37 kDa and 27 kDa were observed. Thus, His_6-FibrilΔ(305-512) and Strep-FibrilΔ(1-301)-His_6 were prepared as the complementary constructs. Also, Strep-Fibril(290-479)-His_6 construct was prepared considering the ~ 24 kDa band observed during purifications of FibrilΔ(1-236)-His_6 and full-length Fibril. The characterization of His_6-FibrilΔ(305-512) was initiated first.

As mentioned earlier, the aim was to obtain non-polymerizing construct of Fibril for X-ray crystallography studies. Thus, sedimentation properties of smaller constructs were compared with full-length Fibril; upon spinning at 46,500 xg, Fibril pellets down probably because it forms bundles. A non-polymerizing protein is not expected to pellet down upon spinning at > 80,000 xg unless it binds to membranes or forms a part of macromolecular complexes. As explained in Chapter 2 of this thesis, post SDS-treatment, the full-length Fibril remains maximally in the supernatant upon spinning at 159,000 xg. Thus both procedures with and without SDS treatment were followed for characterization of short constructs.

Like full-length Fibril, in absence of SDS His_6-FibrilΔ(305-512) also pelleted upon spinning at 159,000 xg (lanes 7 and 8 in Figures 3.6 A and B). Thus, His_6-FibrilΔ(305-512) was thought to form polymers or it may exist as aggregates. The differentiation between protein polymers and aggregates is possible only by visualization of samples using electron microscopes. However, the His_6-FibrilΔ(305-512) pellet obtained by spinning lysate at 159,000 xg was not soluble in any of the buffers attempted and was present as clumps. The pellet also had many other proteins as seen on SDS-PAGE gel (lane 8 Figure 3.6 B). Thus, the protein was purified further before visualization using an electron microscope.

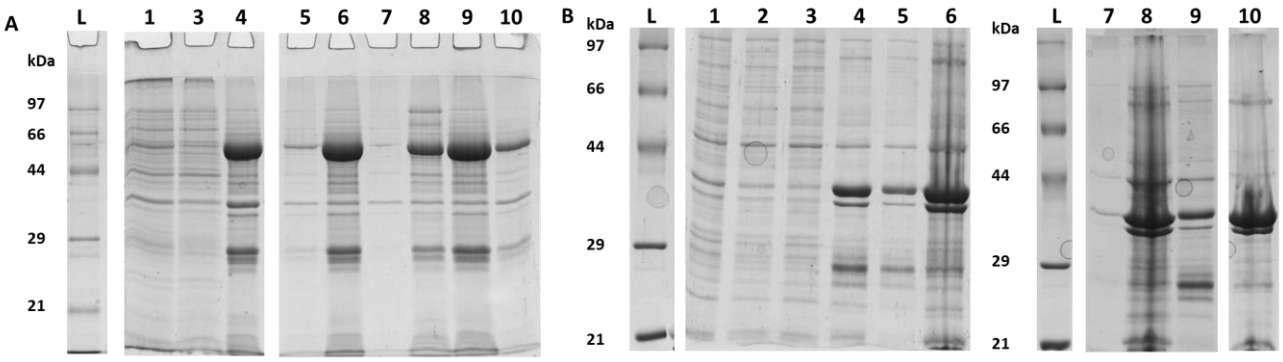

Figure 3.6 Pelleting /polymerization profiles of His_6-Fibril (Mw 59 kDa) and His_6-FibrilΔ(305-512) (Mw 35 kDa) with and without use of SDS. The differential centrifugation of lysate of *E. coli* cells expressing protein His_6-Fibril (A) or His_6-FibrilΔ(305-512) (B) revealed that these proteins were not soluble without SDS treatment (lane 5 in both the gels). However, addition of SDS drastically improved the solubility of both the proteins (lane 6 in both the gels). Centrifugation of the sample in lane 5 (without SDS treatment) at 159,000 xg resulted into very small quantity of the protein remaining in supernatant (lane 7) as compared to pellet (lane 8). In contrast to this, the protein solubilized using SDS (lane 6), when spun at 159,000 xg left most of the protein in supernatant (lane 9) as compared to pellet (lane 10). Note that for SDS treated samples (lanes 9 and 10) the amount of 159,000 xg pellet (lane 10) loaded on to the gel was much higher as compared to corresponding supernatant (lane 9). L-protein marker; 1- total lysate, 2- supernatant after 11,000 xg spin; 3 & 4- supernatant and pellet after 46,500 xg spin respectively; 5 & 6- soluble fraction of pellets from lane 4 re-suspended without and with SDS treatment followed by spin at 21,000 xg respectively; 7 & 8- supernatant and pellet respectively after spinning sample from lane 5 at 159,000 xg; 9 & 10 - supernatant and pellet respectively after spinning sample from lane 6 at 159,000 xg.

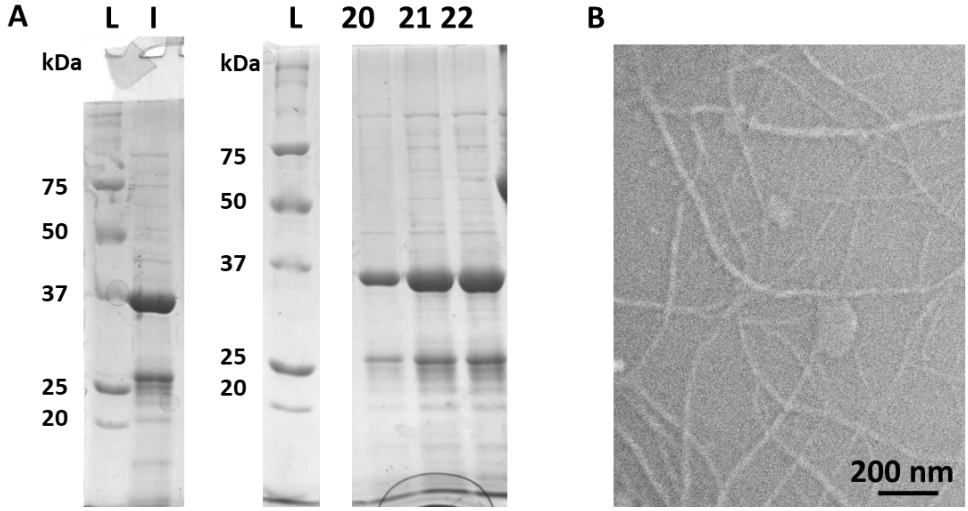

Figure 3.7 Purification and visualization of His_6-FibrilΔ(305-512) (Mw 35 kDa). A) The SDS-solubilized His_6-FibrilΔ(305-512) protein was separated from other proteins based on its density using glycerol density gradient. The fractions containing protein of interest were pooled, dialyzed and visualized using FE-SEM. L- Protein marker; I- input/ sample loaded on the glycerol gradient; 20-22 – (top most) fractions of glycerol density gradient containing protein of interest. B) FE-SEM image of His_6-FibrilΔ(305-512) purified by SDS-treatment followed by separation on glycerol density gradient.

Similar to full-length Fibril, the treatment of His_6-FibrilΔ(305-512) with SDS at final concentration of 1 % (w/v) resulted into complete solubilization of the pellet that did not pellet down upon spinning at 159,000 xg (lanes 9 and 10 in Figure 3.6 A and B). Separation of solubilized protein mixture was attempted using glycerol gradient. His_6-FibrilΔ(305-512) was found in the top 3 fractions of glycerol gradient, corresponding to the density about 0.13 g/cm^3. The fraction from glycerol gradient was slightly purer than the 159,000 xg pellet and it was also soluble. Visualization of the glycerol gradient fractions containing His_6-FibrilΔ(305-512) (Figure 3.7 A) using field emission scanning electron microscope (FE-SEM) revealed filamentous structures (Figure 3.7 B). However, unlike the characteristic twists seen in all EM images of full-length Fibril, filamentous structures seen for His_6-FibrilΔ(305-512) samples do not have any features. One of the possible reasons for this is the low resolution of the image.

Nevertheless, since filaments are seen for His_6-FibrilΔ(305-512) and the aim was to obtain a non-polymerizing construct of Fibril, a few more constructs were prepared. The approach for designing further smaller constructs was to reduce the length of His_6-FibrilΔ(305-512) by removing one secondary structure at a time from its carboxy-terminus. Thus, constructs His_6-FibrilΔ(285-512), FibrilΔ(273-512)-His_6, His_6-FibrilΔ(253-512) and FibrilΔ(228-512)-His_6 were prepared and characterized.

3.3.2.4) Constructs His_6-FibrilΔ(285-512), FibrilΔ(273-512)-His_6 and His_6-FibrilΔ(253-512) behave similar to full-length Fibril

The pelleting assay of His_6-FibrilΔ(285-512), FibrilΔ(273-512)-His_6 and His_6-FibrilΔ(253-512) from clarified (free of cell debris) lysate showed that all these protein constructs pellet down when spun at 46,500 xg/ 30 minutes/4 ºC when not treated with SDS (Figure 3.8 A-C). Protein bands corresponding to molecular weights of His_6-FibrilΔ(285-512) and FibrilΔ(273-512)-His_6 were not seen in uninduced BL21(AI) cell lysate (Figure 3.8 D). Thus, both these constructs of Fibril formed either protein filaments or aggregates.

Visualization of SDS-treated and dialyzed samples of His_6-FibrilΔ(285-512) and FibrilΔ(273-512)-His_6 constructs using FE-SEM did not reveal any filamentous structures. The possible reasons why filaments were not seen for these constructs are: i) both these constructs do not form filaments or ii) the filaments formed by these constructs were denatured during purification

procedure. Thus, attempts were made to purify these constructs without the use of SDS for further characterization. Since these constructs were prepared with a His_6 tag, the purification of these constructs was attempted using Ni-NTA matrix. The construct His_6-FibrilΔ(285-512) did

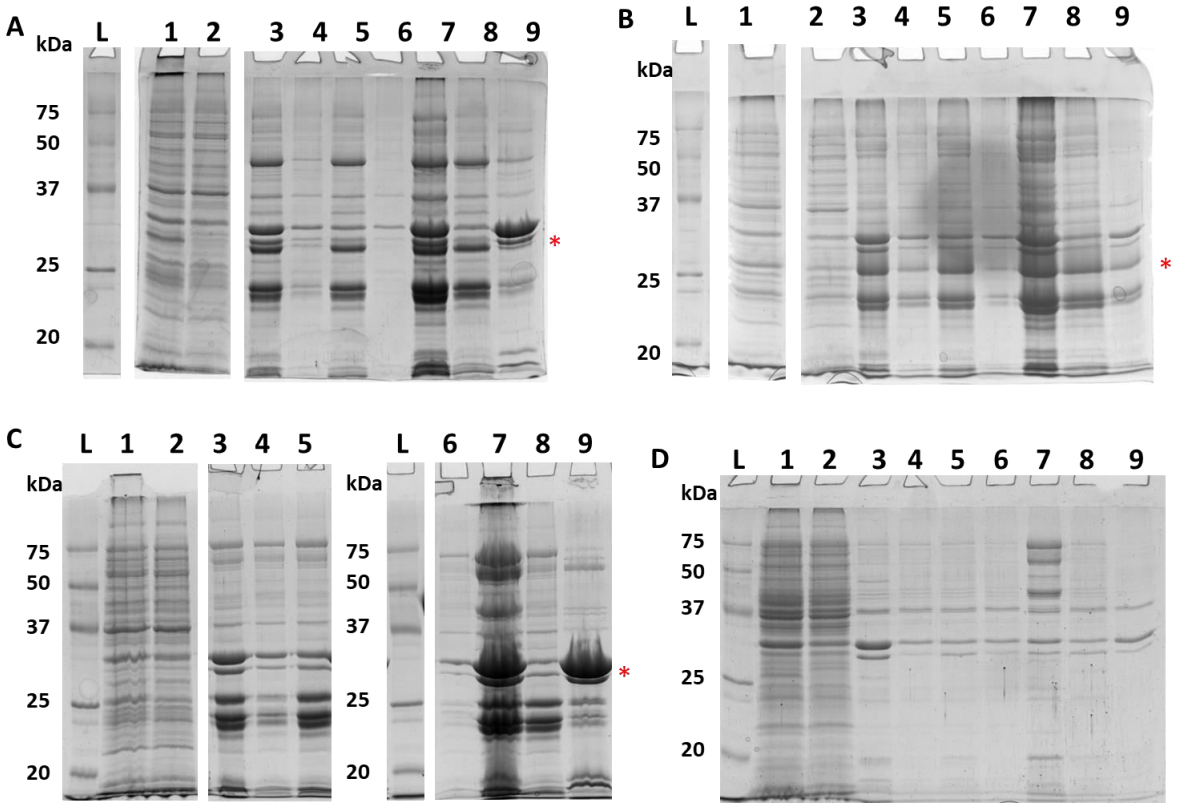

Figure 3.8 Polymerization/pelleting assay for short constructs of Fibril. A-C) Similar to His_6-Fibril, the short constructs His_6-FibrilΔ(285-512) (A; 33 kDa), FibrilΔ(273-512)-His_6 (B; 32 kDa) and His_6-FibrilΔ(253-512) (C; 30 kDa) were obtained in the pellet upon spinning clarified lysate at 46,500 xg (lane 3 in all gels). The pelleted short constructs were sparingly soluble in the buffer (lane 4) unless treated with SDS (lane 5). These short constructs pellet down upon spinning at 159,000 xg if not treated with SDS (lanes 6, 7). However, SDS treatment left these short constructs in supernatant upon spinning at 159,000 xg (lanes 8, 9). D) Pelleting profile of BL21(AI) cells as experimental control. Note that the amount of pellet loaded in lane 9 is much higher than the corresponding supernatant (lane 8). L- Protein marker; 1- total lysate; 2 - supernatant upon spinning lysate at 11,000 xg; 3- Pellet obtained upon spinning protein as in lane 2 at 46,500 xg; 4 & 5- soluble fraction obtained upon solubilization of the pellet (lane 3) without or with SDS respectively, followed by spinning at 21,000 xg; 6 & 7- supernatant and pellet respectively after spinning sample from lane 4 at 159,000 xg; 9 & 10 - supernatant and pellet respectively after spinning sample from lane 5 at 159,000 xg. The red asterisk indicates band of protein of interest.

not bind to the Ni-NTA affinity matrix (data not shown), hinting that the His_6 tag may not be exposed. Attempts were made for purification of FibrilΔ(273-512)-His_6 with and without SDS treatment.

3.3.2.5) FibrilΔ(273-512)-His_6 is purified with and without SDS-treatment

SDS-treated FibrilΔ(273-512)-His_6 could be purified using Ni-NTA affinity chromatography (Figure 3.9), to finally obtain a single band. However this protein was not stable during steps of dialysis or concentration using centricons for further experiments with size exclusion chromatography.

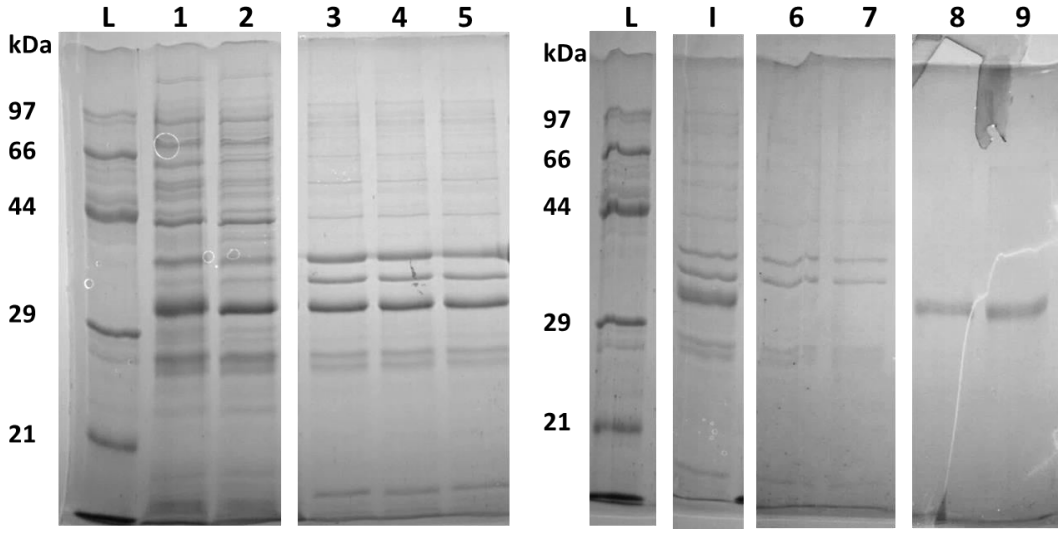

Figure 3.9 Purification of FibrilΔ(273-512)-His_6 by SDS treatment followed by affinity chromatography. FibrilΔ(273-512)-His_6 construct (Mw 32 kDa) was obtained in the soluble fraction upon SDS treatment followed by 159,000 xg spin. The protein was then purified by allowing its binding to Ni-NTA resins (affinity chromatography). The bound protein was eluted using buffers containing increasing concentrations of imidazole. The purified protein can be seen to elute with 250 mM and 500 mM imidazole containing buffer (lanes 8 & 9 respectively). L-Protein marker; 1- total lysate; 2- supernatant after spinning lysate at 11,000 xg; 3- Pellet obtained upon spinning protein as in lane 2 at 46,500 xg; 4- soluble fraction obtained upon solubilization of the pellet (lane 3) by treatment with SDS followed by spinning at 21,000 xg; 5 & I- supernatant obtained by spinning sample from lane 4 at 159,000 xg (I denotes input for affinity chromatography); 6 & 7 – flowthrough and wash (protein not bound to Ni-NTA matrix); 8 & 9- bound protein eluted using 250 mM and 500 mM imidazole containing buffer.

Further, attempts were made to purify this construct FibrilΔ(273-512)-His_6 by affinity chromatography without use of SDS. The protein bound to the affinity matrix and was obtained in partially purified form (Figure 3.10 A). Attempts were made to improve the purity of FibrilΔ(273-512)-His_6 using ion-exchange chromatography. However, the protein in spite of being positively charged at pH 7.6 (as deduced for its theoretical pI 8.27), did not bind the negatively charged matrix of MonoS column (Figure 3.10 A). Thus, it was checked if the protein binds positively charged matrix of MonoQ column.

The protein was found to bind anion exchange matrix and eluted as relatively pure fractions (Figure 3.10 B). However, it was accompanied by bands of proteins with higher molecular weight (> 33 kDa) than FibrilΔ(273-512)-His_6. The purified, concentrated FibrilΔ(273-512)-His_6 was visualized on SDS-PAGE gel (Figure 3.10 B). The FibrilΔ(273-512)-His_6 protein purified with and without SDS treatment will now be used for electron microscopy studies to check if they form polymers.

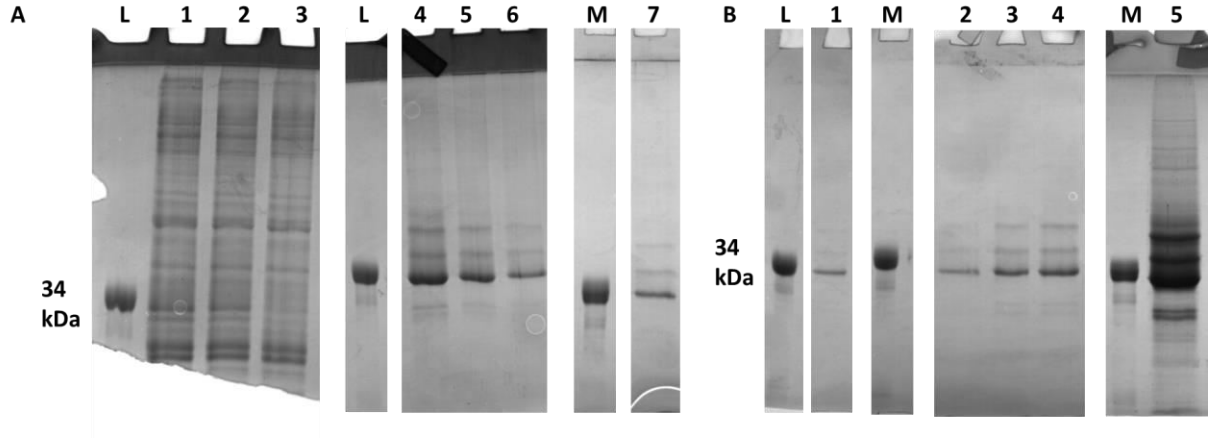

Figure 3.10 Purification profile of FibrilΔ(273-512)-His_6 without use of SDS. A) The protein FibrilΔ(273-512)-His_6 was purified (lanes 4-6) from clarified lysate using Ni-NTA affinity chromatography. The protein did not bind to cation exchange matrix (MonoS column) and eluted as flowthrough (lane 7). L- 34 kDa protein; 1- total lysate; 2- supernatant after spinning lysate at 46,500 xg; 3- flowthrough from Ni-NTA column; 4 to 6- protein bound to Ni-NTA column and eluted using high concentration of imidazole in the buffer; 7- flowthrough from cation exchange (MonoS) column. B) The flowthrough of MonoS column was passed through anion exchange (MonoQ) column, bound protein eluted and concentrated. L- 34 kDa protein, 1- flowthrough of MonoQ column; 2 to 4- fractions of protein bound to MonoQ column matrix and eluted with increasing salt concentration; 5- Pooled and concentrated fractions from MonoQ column containing protein of interest.

3.4) Discussion

This section summarizes attempts to identify domain architecture of Fibril. There are no reports on domain dissection of Fibril since genetic manipulation in *Spiroplasma* is tedious. Attempts of heterologous expression of Fibril have not been reported till date. Due to lack of structural information on Fibril, secondary structure prediction was used as the basis for preparation of short constructs. Multiple short constructs were prepared, purified, and characterized with respect to filament forming properties.

FibrilΔ(1-236)-His_6, a short construct of full-length Fibril, was purified and its crystallization attempted. Though crystals have not been obtained for FibrilΔ(1-236)-His_6, its purification has

provided hints as follows. The presence of FibrilΔ(1-236)-His_6 in folded, monomeric state suggests that it may form the C-terminal domain of Fibril. Further, since FibrilΔ(1-236)-His_6 did not polymerize it will be interesting to know if its complement, FibrilΔ(237-512)-His_6, can form polymers on its own. The characterization of FibrilΔ(237-512)-His_6 will help us know if it can polymerize independently of FibrilΔ(1-236)-His_6 or both the domains are required for polymerization. The information regarding role of N-terminal domain in polymerization of Fibril was obtained from characterization of the His_6-FibrilΔ(305-512) construct.

It is interesting to see that His_6-FibrilΔ(305-512), a truncated construct of Fibril, potentially formed polymers. However, the filaments did not have characteristics such as the twist or mesh-like appearance (due to repeating circular tetramers) seen for full-length Fibril. This raises two possibilities- i) The short construct, His_6-FibrilΔ(305-512) forms filaments with a different morphology or ii) these filaments are not really protein polymers and must be contamination. Further purification and characterization of His_6-FibrilΔ(305-512) is likely to provide hints to rule out one of these two possibilities. An important hint obtained from the purifications of FibrilΔ(1-236)-His_6 and His_6-FibrilΔ(305-512) is that the N-terminal domain may be sufficient for polymerization. Cryo EM studies of His_6-FibrilΔ(305-512) polymers and its comparison with full-length Fibril filaments will help decipher if the His_6-FibrilΔ(305-512) polymerization involved same interface as that of full-length Fibril. The comparison will help to verify the role of C-terminal domain in polymerization, if any.

Attempts to purify the N-terminal constructs His_6-FibrilΔ(305-512), His_6-FibrilΔ(285-512) and FibrilΔ(273-512)-His_6 revealed that only FibrilΔ(273-512)-His_6 could bind to Ni-NTA resins without SDS treatment. It should be noted that FibrilΔ(273-512)-His_6 has a C-terminal His_6 tag as opposed to His_6-FibrilΔ(305-512) and His_6-FibrilΔ(285-512) with His_6 tag at the N-terminal. Thus, the N-terminus of Fibril must be either buried within the N-terminal domain or at the polymerization interface. This is consistent with full-length Fibril; N-terminally His_6 tagged full-length Fibril does not bind to Ni-NTA matrix without SDS treatment.

Further, like full-length Fibril, all the three N-terminal constructs His_6-FibrilΔ(305-512), His_6-FibrilΔ(285-512) and FibrilΔ(273-512)-His_6 were obtained in the insoluble fraction in lysis buffer unless they were treated with SDS. In contrast, the C-terminal construct FibrilΔ(1-236)-His_6 was obtained in the soluble fraction as monomer without SDS treatment. Thus, it appears

that the N-terminal domain of Fibril may be responsible for interaction with membrane lipids and formation of polymers/filaments. Since the C-terminal region (beyond amino acid number 236 from amino terminus) of Fibril does not share homology with any of the known DNA or protein sequences, it is difficult to predict a function or interacting partners at this stage.

All the aforementioned observations are preliminary and need confirmation by further experiments in this direction. One of the major challenges in working with such proteins is the lack of an activity assay for monitoring the amount of functional protein. The information on activity of protein is useful to confirm that the right protein is being characterized. Another challenge of working with these short constructs is the inconsistency/non-reproducibility of expression and purification. Over-expression of short constructs and obtaining them in the soluble fraction has been inconsistent. The poor expression and insolubility of these short constructs may be due to the cleavage of protein between secondary structures (α-helix or β-sheet) making them the unstable.

Characterization of some of the most well behaved (in terms of expression and solubility) short constructs of Fibril is explained above. Additional constructs have been designed and cloned (Table 3.1). Further standardization of some of these constructs is necessary to obtain over-expression and maximum solubility to facilitate their purification and characterization. The present status of all the constructs of Fibril is summarized in Table 3.5).

Table 3.5 Current status of short constructs of Fibril.

Construct	Purification standardized	Filaments	
		Pelleting assay	Seen in FE-SEM
His_6-FibrilΔ(305-512)	Yes	Yes	Yes (once)
His_6-FibrilΔ(285-512)	No	Yes	No
FibrilΔ(273-512)-His_6	Yes	Yes	No
His_6-FibrilΔ(253-512)	No	Yes	-
FibrilΔ(237-512)-His_6	No	No	-
FibrilΔ(228-512)-His_6	No	No	-
FibrilΔ(1-236)-His_6*	Yes	No	-
FibrilΔ(104-109)	No	No	-
FibrilΔ(109-112)	No	No	-
Strep-FibrilΔ(1-301)-His_6	No	No	-
Strep-Fibril(290-479)-His_6	No	No	-

* Crystallization has been attempted.

In order to overcome challenges listed above, an alternate approach was designed to obtain non-polymerizing construct of full-length Fibril. The approach involved introduction of EGFP protein, a bulky molecule, at different sites in Fibril sequence. If the protein fails to fold properly, it will be degraded by the cell and total expression of the fusion protein will be low. If the EGFP-tagged Fibril folds properly, the EGFP may or may not occur at the polymerization of Fibril. If EGFP is at the polymerization interface, it will prevent Fibril polymerization and monomeric Fibril will be obtained. However, if the EGFP is at any site other than the polymerization interface, EGFP-tagged Fibril filaments will be obtained (Figure 3.11).

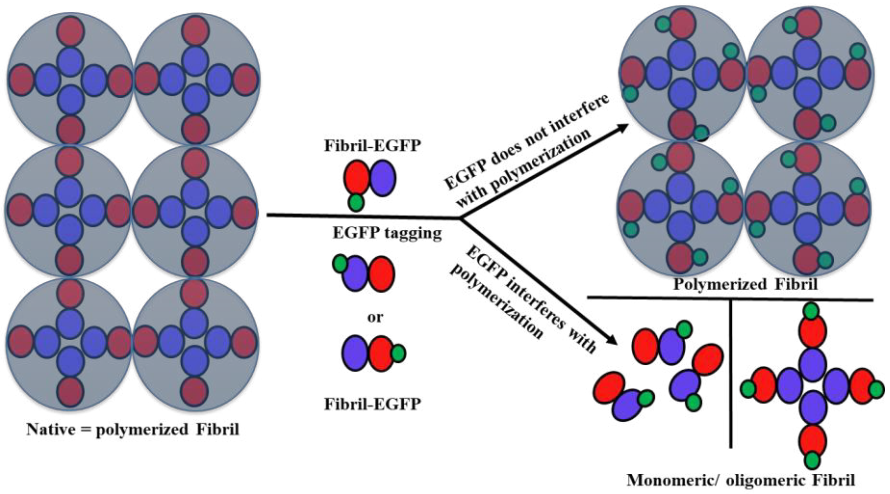

Figure 3.11 Pictorial representation of strategy to obtain non-polymerizing construct of Fibril. Tagging of EGFP into Fibril creates two possibilities- i) EGFP is away from the polymerization interface of Fibril, and provides a site for tagging Fibril without affecting its polymerization. ii) EGFP is placed at the polymerization interface of Fibril. This potentially results in obtaining monomeric/ non-polymeric Fibril.

In the former case, the information obtained on polymerization interface will be used to generate non-polymerizing constructs whereas, in the latter case, the information will be useful to visualize intra-cellular localization of Fibril within *Spiroplasma* cells by tagging it with a fluorescent protein.

25 sites have been identified within Fibril and EGFP-tagged constructs prepared. Initial characterization of these constructs revealed that some of the constructs are well expressed. Further studies on these constructs suggest that there are differences in fluorescence emitted by each construct. The constructs with low fluorescence constructs are likely to be resulting from interference by EGFP in Fibril folding and thus may not be at the polymerization interface. In

contrast, constructs with high fluorescence must be the folded protein. *In vitro* characterization of these constructs is in progress to identify if they are polymers or monomers.

Chapters 2 and 3 of my thesis described the in vitro investigation of Fibril, a proposed central player in shape determination and motility of *Spiroplasma*. The next chapter describes experiments performed to investigate components involved in *Spiroplasma* shape determination and motility by characterization of a naturally occurring non-helical, non-motile mutant of *Spiroplasma*.

3.5) References

Cohen-Krausz, S., Cabahug, P.C., and Trachtenberg, S. (2011). The monomeric, tetrameric and fibrillar organization of Fib: the dynamic building block of the bacterial linear motor of *Spiroplasma melliferum* BC3. J. Mol. Biol. *410*, 194–213.

Lee, J.E., Cornell, K.A., Riscoe, M.K., and Howell, L.P. (2001). Structure of *E . coli* 5'-methylthioadenosine/S-adenosylhomocysteine nucleosidase reveals similarity to the purine nucleoside phosphorylases. Structure *9*, 941–953.

Mizuno, K. (1992). Induction of cold stability of microtubules in cultured tobacco cells. Plant Physiol. *100*, 740–748.

Richardson, J.S. (1981). The anatomy and taxonomy of protein structure. Advances in protein chemistry, pp. 167–339.

Woodring, P.J., Hunter, T., and Wang, J.Y.J. (2001). Inhibition of c-Abl tyrosine kinase activity by filamentous actin. J. Biol. Chem. *276*, 27104–27110.

Chapter 4

Comparative characterization of Spiroplasma citri GII-3 and ASP-I cells

4.1) Introduction

The previous two chapters described structural characterization of Fibril in polymerized and non-polymeric forms using recombinant expression systems. In addition to Fibril, cytoskeletal protein MreB also has been proposed to be associated in *Spiroplasma* shape determination and motility. Due to challenges associated with gene manipulation in *Spiroplasma*, the role of these proteins has been hypothesized but remains to be verified through genetic studies. An alternative approach for deciphering gene function is to find naturally occurring mutants and identify the molecular determinants responsible for the defects.

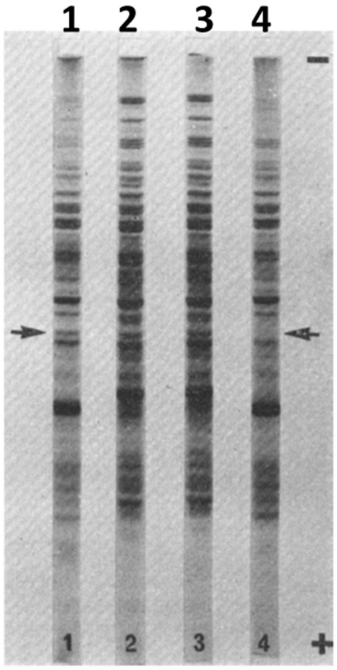

Figure 4.1 Protein profiles of *Spiroplasma citri* cells. Semi purified membrane proteins of *S. citri* strains as visualized on SDS-PAGE gel. 1: SP-A, 2: R8-A2, 3: C189 and 4: ASP-I. The arrows point to 39 kDa protein. Image adapted from Townsend et al., (1977).

A non-helical and non-motile strain of *Spiroplasma citri*, ASP-I, was reported by Townsend et al., (1977) (Figure 1.3 B). DNA homology, serology tests, production of toxin and pathogenicity confirmed that ASP-I was a strain of *S. citri*. Comparison of semi purified membrane fractions of *S. citri* SP-A (wildtype; helical and motile) and ASP-I cells was performed to decipher mechanism of helical shape determination in *Spiroplasmas*. Since Fibril was proposed to be an important determinant of helical shape in *Spiroplasma*, ASP-I cells were first investigated for presence of Fibril. The analysis showed that Fibril was in fact present in both cell types (Townsend et al., 1977). However, a protein with molecular weight of about 39,000 Da was missing from ASP-I as compared to SP-A cells

(Figure 4.1). Thus, it was hypothesized that the 39 kDa protein must be a structurally important component and an anchor for connecting Fibril to the cell membrane at regular intervals and its absence led to loss of helicity and motility (Townsend et al., 1977, 1980). The identity of the 39 kDa protein remains unknown till date and further studies have not been performed on *S. citri* ASP-I cells. Identification of the missing 39 kDa protein from *S. citri* ASP-I cells may provide mechanistic insights into components involved in *Spiroplasma* shape determination and motility. The study would also be useful for identification of protein(s) involved in motility and rod-to-helical transition of *Spiroplasma* morphology. Thus, characterization of *S. citri* ASP-I was carried out as a part of my Ph.D. research work.

This chapter describes the comparative analysis of the wildtype (helical and motile) isolate *S. citri* GII-3 and *S. citri* ASP-I, a naturally occurring non-helical, non-motile mutant. The characterization included genomic, proteomic studies and visualization of cytoskeletal protein filaments. Analysis of the genomes revealed defect in a gene of *S. citri* ASP-I cells that led to absence of the 39 kDa protein. Gene coding for the same 39 kDa protein in *S. citri* GII-3 cells was cloned into a vector and expressed in *S. citri* ASP-I cells. The transformed ASP-I cells exhibited helical morphology and kinking motility. The experimental approach followed and results obtained are described below in the following sections of the chapter.

4.2) Material and methods

S. citri ASP-I cultures are available in the lab of Dr. Laure Béven, INRA, Bordeaux, France. Since *S. citri* is a plant pathogen, the procedure for obtaining permits for importing cultures is a time consuming process and could not be completed within the time frame of my Ph.D. Thus the genomic and proteomic studies were carried out in lab of Dr. Laure Béven during the time spent in her lab as a Raman Charpak fellow for six months. In absence of electron cryotomography (ECT) facility in India, ECT studies were performed in the lab of Dr. Martin Pilhofer, ETH, Zürich, Switzerland, supported by an EMBO short-term fellowship for a period of three months.

Following sections describe the methodology followed for characterization of *S. citri* GII-3 and ASP-I cells. *Spiroplasma* cells can easily lose their shape and hence the cultures need to be handled with care. Also discussed are challenges associated with handling *Spiroplasma* cells for ECT studies.

4.2.1) Growing *Spiroplasma citri* GII-3 and ASP-I cells

Spiroplasma citri GII-3 as well as ASP-I cells were grown in modified SP4 medium as described by Duret et al., (1999) and explained in section 2.2.2.3 of this thesis. The grown cells were either frozen for ECT studies or used for genomic DNA isolation and proteomic studies.

4.2.2) Isolation and sequencing of genomic DNA

The genome sequences of *S. citri* GII-3 and ASP-I were obtained from Dr. Laure Béven.

4.2.3) Analysis of *S. citri* genome for identification of proteins with molecular weight of about 39 kDa

A list of genes coding for protein with molecular weight about 39 kDa was prepared with an intention to shortlist potential proteins that would be the missing in the non-helical ASP-I cells. Genome of *S. citri* is available in NCBI (https://www.ncbi.nlm.nih.gov/nuccore/CP013197.1). To identify all the possible genes in *S. citri* genome with a rough molecular weight of about 39 kDa, a range between 37 and 41 kDa was considered. The proteins in this range of molecular weights is expected to contain about 336 – 370 amino acids (considering average molecular weight of 110 Da/ amino acid). The gene length coding for proteins of 336-370 amino acids will be between 1050 and 1110 nucleotides. Thus, all genes in *S. citri* genome with a length between 1050 and 1110 nucleotides were identified and listed manually. *S. citri* R8-A2 genome was used for the analysis.

Interestingly, molecular weights of all the five MreBs present in *Spiroplasma* cells is close to 39 kDa. It is possible that the protein absent from *S. citri* ASP-I is one of the MreBs. Thus, pair-wise sequence alignment of corresponding *mreB* gene sequences from the two cell types (*S. citri* GII-3 and *S. citri* ASP-I) was performed. These sequences were determined by PCR amplification of the *mreB* genes using flanking primers in the genome.

The genome sequence analysis was also complemented with investigation of total protein in the cells using mass spectrometry.

4.2.4) Extraction of *Spiroplasma* proteins for mass spectrometry studies

Spiroplasma proteins were extracted, using the protocol from Renaudin et al., (2015), as follows. A 50 mL culture pellet of *Spiroplasma* cells (ASP-I and GII-3) was re-suspended into 1 mL Tris-buffered saline [TBS; 10 mM Tris-HCl (pH 7.4), 150 mM NaCl]. Cells were lysed on ice by sonication for 1 minute using a probe sonicator. 200 µL lysate was transferred to a fresh tube and stored at -80 ºC. 100 µL extraction buffer [10 mM Tris (pH 7.4), 154 mM NaCl] and 100 µL Triton X-114 (10 % w/v) were added to rest of the 800 µL lysate on ice. The mixture was left at 4 ºC for 1 hour under shaking condition without frothing to facilitate extraction of membrane proteins by Triton X-114. The mixture was then spun at 12000 xg/4 ºC/40 minutes to separate insoluble proteins as pellet from soluble ones. Since the supernatant contained aqueous soluble proteins in addition to the ones solubilized by Triton X-114, they were separated by incubating supernatant at 37 ºC for 15 minutes (or until the solution turned turbid). Since the cloud point for Triton X-114 is 23 ºC, at 37 ºC the solution appears turbid due to the presence of detergent micelles. Triton X-114-bound proteins and micelles were separated by spinning the turbid solution at 3,000 xg/37 ºC/3 minutes. The supernatant containing aqueous soluble proteins was transferred to a fresh tube. This stage onwards, the aqueous soluble protein containing fraction (upper phase) and detergent solubilized fraction (lower phase) were processed separately. Briefly, 3 washes were performed on both the fractions to remove unwanted proteins as follows.

The detergent-enriched lower phase was re-suspended in 900 µL cold extraction buffer and incubated on a rotary shaker at 4 ºC for 15 minutes. The solution was then transferred to 37°C for 15 minutes and then spun at 3,000 xg/37 ^{0}C/3 minutes. The upper phase was discarded and washing step repeated 2 more times. The 100 µL lower phase obtained after 3 washing steps was mixed with 900 µL ice cold methanol (100 %) and incubated overnight at -20 ^{0}C. The incubation with methanol under cold conditions led to removal of Triton X-114. Post-incubation, the solution was spun at 21,000 xg/4 ^{0}C/10 minutes to obtain precipitated proteins as a pellet, which was solubilized in Laemmli buffer lacking bromophenol blue (4% SDS, 20% glycerol, 10% β-mercaptoethanol, and 0.125 M Tris HCl, pH approx. 6.8). The solution was then transferred to a fresh tube and stored at -80 ^{0}C.

100 µL Triton X-114 (10 % w/v in TBS) was added to 900 µL supernatant containing aqueous soluble proteins. The mixture was allowed to incubate at 4 ºC for 15 minutes with shaking, and

then at 37°C for 15 minutes. It was then spun at 3,000 xg/37 ºC/3 minutes. The lower phase contained the remaining detergent soluble proteins while the upper phase contained aqueous soluble proteins. The upper phase was transferred to a fresh tube and lower phase discarded. The supernatant was washed twice in the same manner as above to remove detergent soluble proteins. The supernatant was transferred to a fresh tube and stored at -80 ºC.

Fractions containing total proteins, aqueous soluble proteins and Triton X-114 solubilized proteins were checked on a 12% SDS-PAGE gel. Proteins were identified by peptide mass fingerprint and tandem-mass spectrometry (MS/MS) sequencing from protease digested protein at the Proteome Platform, Functional Genomic Center of Bordeaux, University of Bordeaux. The samples were subjected to in-gel digestion by trypsin, nano-liquid chromatography (nano-LC) on a U3000 Thermo and tandem-mass spectrometry (MS/MS) using an Orbitrap Fusion™ Lumos™ Tribrid™ (ThermoFisher Scientific) to identify and compare proteins between *S. citri* GII-3 and ASP-I cells.

4.2.5) Analysis of proteins present in *Spiroplasma citri* GII-3 and ASP-I cells

The lists of proteins observed in aqueous and membrane-bound fractions of *S. citri* GII-3 as well as ASP-I cells was obtained from mass spectrometry analysis. Each list was arranged in descending order of relative abundance and used for manual analysis.

Proteins which were present in *S. citri* GII-3 but absent in ASP-I cells were identified and enlisted. About 130 proteins according to relative abundance were checked for their presence in *S. citri* GII-3 as well as ASP-I cells manually. To identify proteins with significantly different expression levels, the proteins were ranked according to their relative abundances such that the protein with maximum abundance was ranked first. Next, the ranks of a given protein were checked compared between *S. citri* GII-3 and ASP-I. A rank difference of 50 or more between the two for a candidate protein was considered significant. This exercise was performed for ~ 130 proteins analyzed in present study.

The aforementioned analyses enabled prediction of one of the proteins to be involved in shape determination in *Spiroplasma*. Thus, the gene was cloned under a constitutively expressing (elongation factor Tu: *tuf*) promoter for transformation into *S. citri* ASP-I cells.

4.2.6) Cloning *mreB5* of *S. citri* GII-3 expressible under *ef-tu* promoter into pUC18 vector

In the first PCR *tuf* (*ef-Tu*) promoter region from pSTP1 vector (Renaudin et al., 2015) was amplified using appropriate primers (pBSK(+)tuf-m5-F1 and tuf-m5-R1 or tuf-m5-R1mod;Table 4.1). The second PCR was performed to amplify *mreB5* gene from genomic DNA of *S. citri* GII-3 cells. The products of first and second PCR were purified using PCR clean-up kit (Qiagen) and eluted in 30 µL MilliQ water separately. An overlap-extension PCR (PCR 3) was set up to obtain *tuf promoter-mreB5* construct. The product was extracted from gel using a gel extraction kit (Qiagen) and re-amplified using end primers. The amplified product was cleaned up and digested with restriction enzymes. For restriction digestion, 1 µg each of PCR 3 product and pUC18 vector (Addgene plasmid # 49793) were incubated with 1 µL BamHI restriction enzyme each in a separate tube for 1 hour at 37ºC on a thermomixer. At the end of 1 hour of incubation, samples from the restriction reactions were checked on 1% agarose gel. The presence of single band on gel confirmed complete digestion of DNA. The restriction-digestion products [vector and insert (*tuf* promoter + *mreB5* insert] were then cleaned using PCR purification kit (Qiagen) and eluted with 30 µL MilliQ water. The insert was ligated into pUC18 vector as described in chapter 2 of this thesis. The ligation reaction was transformed into NEB® Turbo electrocompetent *E. coli* cells by electroporation (2500 V voltage, 25 µF capacitance and 200 Ω resistance). The NEB® Turbo *E. coli* [genotype F' $proA^{+}B^{+}$ $lacI^{q}$ Δ*lacZM15 / fhuA2* Δ*(lac-proAB) glnV galK16 galE15 R(zgb-210::Tn10)*Tet^{S} *endA1 thi-1* Δ*(hsdS-mcrB)5*] cells were chosen for their high transformation efficiency (1-3 × 10^{9} cfu/µg pUC19 DNA; NEB Inc. catalogue number C29841).The obtained transformants were grown in LB broth supplemented with ampicillin (final concentration of 100 µg/mL) and plasmid extracted as explained earlier. The plasmid was sequenced using T7 promoter primers to confirm insert sequence.

Ms. Mrinmayee Bapat assisted in ligating the PCR3 product into pUC18 vector.

Table 4.1 List of primers used for cloning *tuf* promoter and *mreB5* gene into pUC18 vector.

Primer name	Sequence (5' → 3')	Purpose
pBSK(+)tuf-m5-F1	CGATAAGCTTGATATC**GAATTC**GTTTAATCT ACCATTAC	Amplification of *tuf* (EF-Tu) promoter region
tuf-m5-R1	GGTCTAGTTTCTGGTCTCATTTTAATCTTTT CCTCC	
tuf-m5-R1mod	GTCAAGAGAAATAAATGGTCTAGTTTCTGGT CTCATTTTAATCTTTTCCTCC	
tuf-m5-F2	GGAGGAAAAGATTAAAATGAGACCAGAAACT AGACC	Amplification of *mreB5* gene of *S. citri* GII-3
tuf-m5-F2mod	GGAGGAAAAGATTAAAATGAGACCAGAAACT AGACCATTTATTTCTCTTGAC	
pBSK(+)tuf-m5-R2	GGATCCCCCGGGCTGCAG**GAATTC**TTATTTT CTTTTTTTACC	

Restriction site sequence in primer is highlighted in purple color.

4.2.7) Expression of wildtype *mreB5* gene from *S. citri* GII-3 in *S. citri* ASP-I cells

Since we do not have the permit to grow *Spiroplasma* cultures in India, the *S. citri* GII-3 *mreB5* gene cloned downstream of *tuf* promoter in pUC18 plasmid was sent to Dr. Béven's lab in INRA. There, the clone was restricted with BamHI enzyme and cloned into pSD4 (Lartigue et al., 2002), an expression vector for *Spiroplasma*.

4.2.8) Transformation of *Spiroplasma* cells

For transformation, *S. citri* ASP-I cells were grown as mentioned in section 2.2.2.3 of Chapter 2 in this thesis. The cells were transformed using the protocol described by (Stamburski et al., 1991). Briefly, 3 mL culture per transformation reaction was grown in SP4 media. The growth of cells was monitored by microscopy of the culture. Upon observation of the correct growth stage of cells, culture was pelleted at 17,400 xg/18 ºC/20 minutes. The pellet was washed to remove traces of media components by one cycle of re-suspension in HS buffer (8 mM HEPES pH 7.4, 280 mM Sucrose) and spinning at 17,400 xg/18 ºC/20 minutes. Care was taken to break cell clumps without damaging cells while re-suspending using a glass capillary with the help of a rubber bulb. The pellet was finally re-suspended in HS buffer in a final volume of 400 µL per reaction (number of transformation reactions X 400 µL).

Simultaneously, sterile electroporation cuvettes (4 mm; Cole-Parmer) were put on ice and 500 ng of plasmid (volume made to 15 µL using sterile MilliQ water) transferred to respective cuvettes

aseptically. A transformation reaction each was set using empty vector and sterile MilliQ water as positive and negative controls respectively for electroporation. All the cuvettes were allowed to chill on ice for 20 minutes. 400 µL of the washed *S. citri* ASP-I cell in HS buffer were added to each cuvette under aseptic conditions. All the cuvettes were further left on ice for another 2 minutes before electroporation using MicroPulser Electroporator (Bio-Rad) and following parameters- 2500 V Voltage, 3 µF Capacitance, 1000 Ω resistance.

After electroporation the transformation reactions were aseptically transferred to sterile 1.5 mL tubes. 800 µL of sterile SP4 liquid media was added to each tube containing transformation reaction and all the tubes incubated at 32 ºC, without shaking for 3 hours. Next, different dilutions (10^{-1}, 10^{-2}) of the recovered transformants were prepared using sterile SP4 liquid media. Two SP4 agar plates supplemented with tetracycline (final concentration of 2 µg/mL) were spread with 200 µL each of the undiluted, recovered cells. Two more plates of SP4 agar containing tetracycline were spread with 200 µL of 10^{-1} and 10^{-2} dilutions of recovered cells. Since the use of spreader can lead to erosion of agar surface and can cause confusion with *Spiroplasma* colonies, samples were spread by tilting plates. All the plates were sealed with parafilm M (Sigma-Aldrich) and incubated at 32 ºC until the development of colonies. The appearance of colonies was monitored by observation of plates, using binocular loupe, once every week until colonies have been grown enough. The colonies were stained with Dienes's stain using the method described by (Deeley et al., 1979) and imaged using binocular loupe.

The cloning of *tuf* promoter and *mreB5* fragment into pSD4 vector, its transformation into *S. citri* ASP-I cells, growth and analysis of transformants was carried out by Ms. Sybille Duret in Dr. Béven's lab in INRA, Bordeaux.

The transformant colonies were inoculated one each into 500 µL SP4 media supplemented with tetracycline (5 µg/mL) to allow further growth.

4.2.9) Confirmation of *S. citri* ASP-I transformants by PCR

The transformants obtained by transformation of *S. citri* ASP-I cells by pSD4 vector containing *tuf* promoter-*mreB5* gene were confirmed by three sets of PCRs. Genomic DNA was prepared from 2 mL of cells culture and resuspended in 20 µL H_2O as described by supplier (Wizard® Genomic DNA purification kit, Promega). The first PCR was performed using SR14 and SR16

primers (Duret et al., 2003) to check if any inhibitors of PCR are present in the DNA preparation. Second PCR was carried out using Tet1 and Tet2 primers (Duret et al., 1999) to confirm the presence of tetracycline-resistance gene carried on the pSD4 vector. The third PCR was performed using Scarp 235 primers (Berho et al., 2006) to check for the presence of genes coding for SCARPs (*Spiroplasma citri* adhesion-related proteins) carried on pSci vectors that are present in *S. citri* GII-3 cells but absent in ASP-I cells. The PCRs against *S. citri* GII-3 cells, *S. citri* ASP-I cells transformed with empty pSD4 vector were set up as negative or positive controls. The primers used for all the three PCRs are listed in Table 4.2.

Table 4.2 List of primers used for confirmation of *S. citri* ASP-I transformants.

Name of the primer	Sequence (5'→3')
SR14	TGTCGGATCCACATCAGTGGTTGCATGTAAT
SR16	CTTTTTGATCTGAATAGTTTCCCGCATCAC
Tet1	CTGCAAAAGATGGCGTAC
Tet2	CGTAAATGTAGTACTCCAC
Scarp 235F	TAAACATTGATATTGCCAACCCC
Scarp 235R	GGTTAAAGTTGCAGAATTATTATC

4.2.10) Imaging complemented *S. citri* cells

The morphology of *S. citri* GII-3, ASP-I as well as *S. citri* ASP-I cells transformed with empty pSD4 vector or pSD4 vector containing *tuf* promoter and *mreB5* gene was observed by dark field microscopy. The bacterial morphology was first observed during the exponential growth phase of *S. citri* GII-3 and ASP-I in modified SP4 under an Eclipse Ni (Nikon) microscope working in reflection and equipped with a dark field condenser. The Nikon oil immersion microscope objective was a 60X with a N. A. of 0.80. Pictures were taken with a camera Nikon Digital Sight DS-Qi1Mc (1280 × 1024 pixels). To determine whether the different strains were motile, the same microscope, objective and camera were used. However, motility was favored by increasing the viscosity of the medium following the protocol described in Boudet et al., (2018): during the exponential growth phase of the bacteria, one volume of the culture was diluted with one volume of Methyl-Cellulose (MC) dissolved in modified SP4. Here, 1 % (w/v) Methyl-Cellulose (from Sigma (M7027), Molecular weight 14000 g/mol) was used to adjust the viscosity of the solutions. Bacteria solutions were prepared between two sealed microscope slides, with a liquid thickness of 15 μm. Videos were recorded at a frame rate of 10 frames per second (fps).

4.2.11) Freezing of *Spiroplasma* cells for electron cryotomography (ECT) studies

Electron cryotomography was performed on *S. citri* GII-3 and ASP-I cells to check if there are differences with respect to organization of cytoskeletal proteins.

Electron cryotomography is an application of transmission electron microscopy (TEM). The ECT technique involves imaging the sample in its near-native, frozen hydrated state, at different angles/tilts to obtain a 'tiltseries'. The tiltseries is then used to prepare a high resolution three-

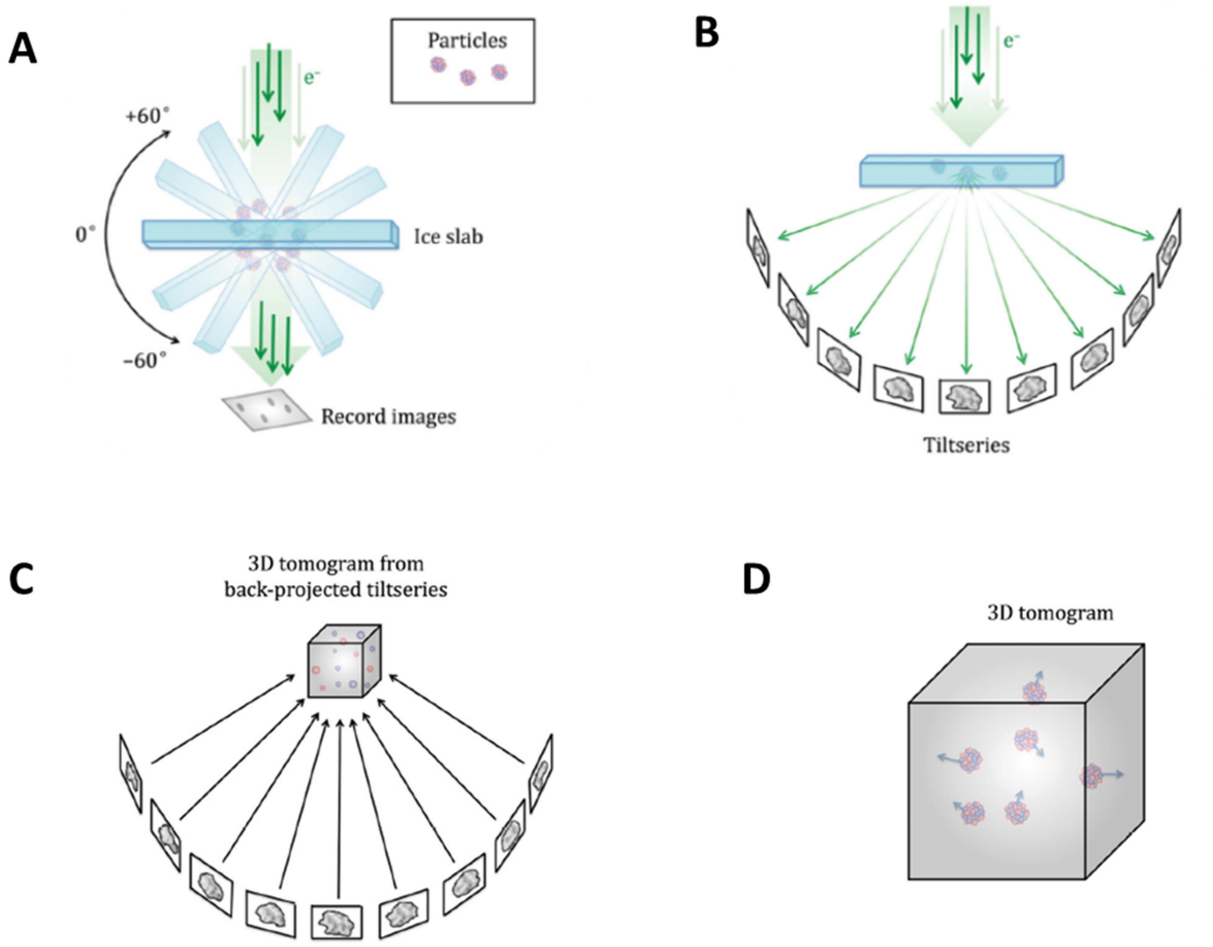

Figure 4.2 Schematic for data collection and tomogram reconstruction in electron cryotomography (ECT). A) The sample (particles/cells) embedded in vitreous ice is imaged at different tilt angles typically between -60° and + 60°. B) The series of images obtained at each tilt results into obtaining a tiltseries. The tiltseries is a collection of views of sample from different directions. C) The tiltseries images are then aligned and a 3D tomogram of the sample is generated. D) The tomogram provides high resolution 3D information about the sample. The schematic is modified from Galaz-Montoya and Ludtke, (2017).

dimensional (3D) tomogram to visualize intra-cellular organization of biomolecules (Basler et al., 2012; Briegel et al., 2006, 2012). The schematic followed for recording tiltseries and reconstruction of 3D tomogram is shown in Figure 4.2.

Recent advancements in the field of electron tomography have enabled visualization of macromolecular structures at a resolution better than 10 Å (Schur et al., 2016). One of the pre-requisites for ECT is the sample with minimum thickness to facilitate electrons to pass-through. Since *Spiroplasma* cells have only a cell membrane (and no cell wall) with a thickness of about 20 nm (Trachtenberg et al., 2014) and a total diameter of the cell of about 200 nm, it becomes appropriate for visualization of cytoskeleton using ECT. The ECT of *S. citri* GII-3 and ASP-I cells would be useful to know if there are differences with respect to organization of cytoskeleton in these cells.

Frozen 0.5 mL cell culture aliquots of *Spiroplasma citri* GII-3 and ASP-I cells were thawed at room temperature. The thawed culture was inoculated into fresh 2 mL SP4 media in a sterile tube under aseptic condition. Cultures were grown at 32 ºC for 20 hours and cell morphology observed by dark field microscopy of 20 µL sample. 0.5 mL of grown culture was inoculated in 2 ml fresh SP4 media for 20 hours. Cell morphology of the growing culture was observed by microscopy. 20 µL bovine serum albumin-coated colloidal gold nanoparticles (10 nm) were added to 80 µL of growing culture and mixed briefly. 5 µL of this mixture was applied to a glow-discharged EM grid (R2/2, 200 mesh, copper, Quantifoil) inside a Vitrobot™ (Marc IV, FEI) chamber at 32 ºC with 95 % humidity. The grids were blotted for 2 seconds and frozen in ethane + propane (37 % : 63 % vol/vol) mixture.

For freezing *Spiroplasma* cells by manual blotting procedure, the bacteria were grown as mentioned above. The culture was mixed with gold nanoparticles and applied to the grid for 30-60 seconds on table at room temperature (25 ºC). The grids held with tweezers were transferred to Vitrobot™. Excess media was absorbed by back-blotting manually using a blotting paper held with forceps (TED PELLA, INC.). The grid was plunge frozen in a mixture of ethane and propane (37 % : 63 % vol/vol) and then transferred to labeled grid storage box. Boxes containing frozen grids were stored in liquid nitrogen until further processing.

4.2.12) Screening, acquisition of tiltseries and tomogram reconstruction of *Spiroplasma* cells

The grid processing and data collection on *Spiroplasma* cells was done with help from Dr. Piotr Szwedziak, ETH, Zürich, Switzerland. The data was collected on Titan Krios (300 KeV; FEI) electron microscope with K2 Summit direct electron detector and Quantum LS imaging filter. The tiltseries were acquired automatically using SerialEM (Mastronarde, 2005). The data collection parameters are mentioned in Table 4.3.

For screening and data collection on *Spiroplasma* cells, the frozen grids were put into a pre-cooled loading station chamber. C-clip rings were clipped onto each grid to be loaded into Titan Krios using autoloader. The C-clipped grids were loaded into a cartridge and transferred to a pre-cooled multi-specimen holder. The holder was parked into the loading chamber of Titan Krios and cartridge loaded into the microscope. Grids were loaded one-by-one from this cartridge for screening/data collection.

Isolated, single cells were selected for data collection for ECT by screening multiple grids to identify regions with well dispersed cells and correct ice thickness. Regions on grids with isolated cells having right natural morphology (helical for *S. citri* GII-3 and non-helical rod shape for *S. citri* ASP-I cells) were selected for data collection.

The tomograms were reconstructed as described in section 3.6 of Weiss et al., (2017). Briefly, the images from tilt-series were aligned with each other to generate “tilt-series with motion correction”. The motion-corrected tilt-series was then used for 3D (3 dimensional) tomogram generation using Etomo as a part of IMOD software package (Mastronarde and Held, 2017).

The gold fiducials present in sample were selected manually and automatic image alignment by selecting “Track Beads” option. The gold beads were erased before generating final aligned stack. The final aligned stack was binned 4 times to generate the data with 1920 X 1854 (X × Y) pixels. Tomograms were generated by weighted back-projection in Fourier space (Radermacher, 2007). Contrast was generated in the tomogram using contrast slider option and entering appropriate values to better visualize cytoskeletal structures.

Table 4.3 Data collection parameters for electron cryotomography of *S. citri* GII-3 and ASP-I cells.

Parameter	*S. citri* GII-3 cells	*S. citri* ASP-I cells
Total dose	~120 e/A^2	
Increment	1° increment, from -20° to +60° and then from -21° to -60°	
Defocus	between -9 and -10 µm	between -7 and -8 µm
Instrument	Titan Krios equipped with energy filter	
Detector	K2 summit direct electron detector	

4.3) Results

Comparative genome and proteome analysis between the *S. citri* GII-3 and ASP-I cells proved useful to identify the protein with a molecular weight of about 39 kDa missing from the latter. Complementing the mutant cells with the functional protein that was lacking in it resulted into helical morphology and motility, thus proving that the effect was indeed due to the loss of the protein. The details are as follows.

4.3.1) Wildtype *S. citri* cells have 76 proteins with a molecular weight of about 39 kDa.

Genome analysis of another wildtype *S. citri* (R8-A2) was performed to identify proteins with molecular weight of about 39 kDa. The analysis showed that 76 genes in *S. citri* R8-A2 genome code for proteins with a molecular weight of about 39 kDa (between 37-41 kDa). In such a case, it is difficult to speculate which of these 76 proteins might be absent in non-helical non-motile ASP-I strain. In order to unambiguously determine the protein responsible for loss of helicity and motility of ASP-I cells, the genome comparison of GII-3 and the mutant ASP-I was necessary.

4.3.2) *S. citri* ASP-I cells have a point mutation in *mreB5* gene leading to a truncated, non-functional MreB5

Cytoskeletal proteins MreB and Fibril have been proposed to be involved in shape determination and motility of *Spiroplasma* (Kürner et al., 2005; Williamson et al., 1991). Since MreB proteins have a molecular weight of about 39 kDa, analysis of *mreB* genes from *S. citri* GII-3 and ASP-I genomes was performed first.

The pair-wise sequence alignment of *mreB* sequences from the genomes of *S. citri* GII-3 and ASP-I revealed that the genes coding for MreB1 to 4 were identical in ASP-I cells. However, a point mutation was observed in *mreB5* gene of ASP-I cells at 400th nucleotide (Figure 4.3).

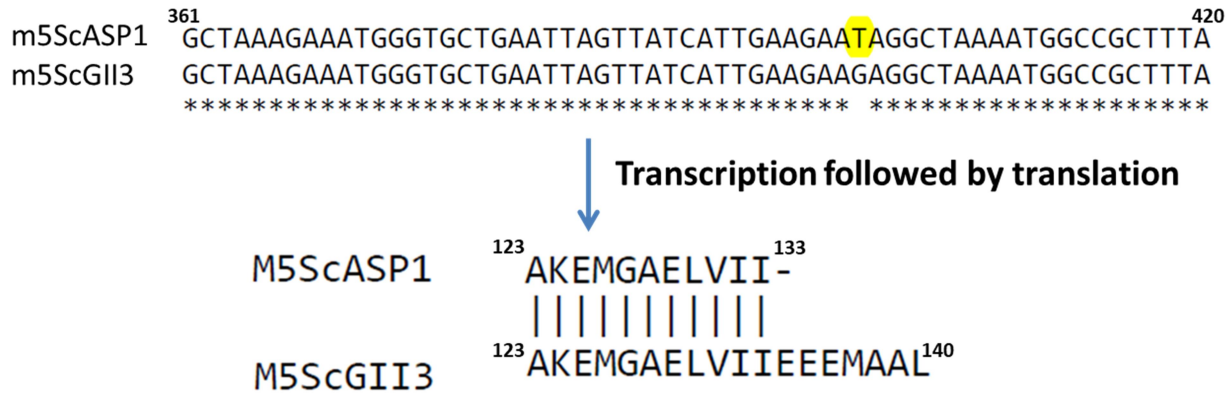

Figure 4.3 Detection of mutation in *mreB5* gene of *S. citri* ASP-I cells. *mreB5* gene in *S. citri* ASP-I cells have a point mutation at the 400th residue. The mutation leads to production of a truncated MreB5 protein with 133 amino acids, instead of 352 amino acids produced in wildtype cells by *mreB5* gene, upon transcription and translation. The mutated nucleotide and amino acid residue are highlighted in yellow. Amino acid and nucleotide numbers have been written in superscript. The sequence alignment was performed using EMBOSS Needle software (https://www.ebi.ac.uk/Tools/psa/emboss_needle/).

The mutation (G → T) results into conversion of a glutamic acid codon (GAG) into a stop codon (TAG). As a result, the *S. citri* ASP-I cells potentially possess a truncated MreB5 protein. Thus, the loss of a functional MreB5 protein could be responsible for the loss of helicity and motility. To identify role(s) of full-length MreB5 in *Spiroplasma* cells complementation study was performed on ASP-I cells.

Simultaneously a proteomic study was performed to check if full-length MreB5 protein is present in the two cell types. The proteomic study would also provide information if any other proteins are missing from *S. citri* ASP-I cells.

4.3.3) Comparative proteome analysis reveals that *S. citri* ASP-I cells have a truncated MreB5 in contrast to a full-length MreB5 present in *S. citri* GII-3 cells.

The visualization of protein samples prepared for mass spectrometry analysis, on SDS-PAGE revealed that a band of about 39 kDa was absent from Triton X-114 solubilized fraction (Figure 4.4). Thus mass spectrometry was performed on the samples to identify differences in protein composition of the two cell types.

The comparative proteomic analysis revealed that there were significant differences in levels of expression of certain proteins in the two cell types (Table 4.4). Interestingly, some proteins were absent in *S. citri* ASP-I cells but were present in GII-3 cells (Table 4.5). Since *S. citri* ASP-I cells lack pSci plasmids, which are present in *S. citri* GII-3 cells, the proteins coded by genes present on these plasmids were not considered during comparison.

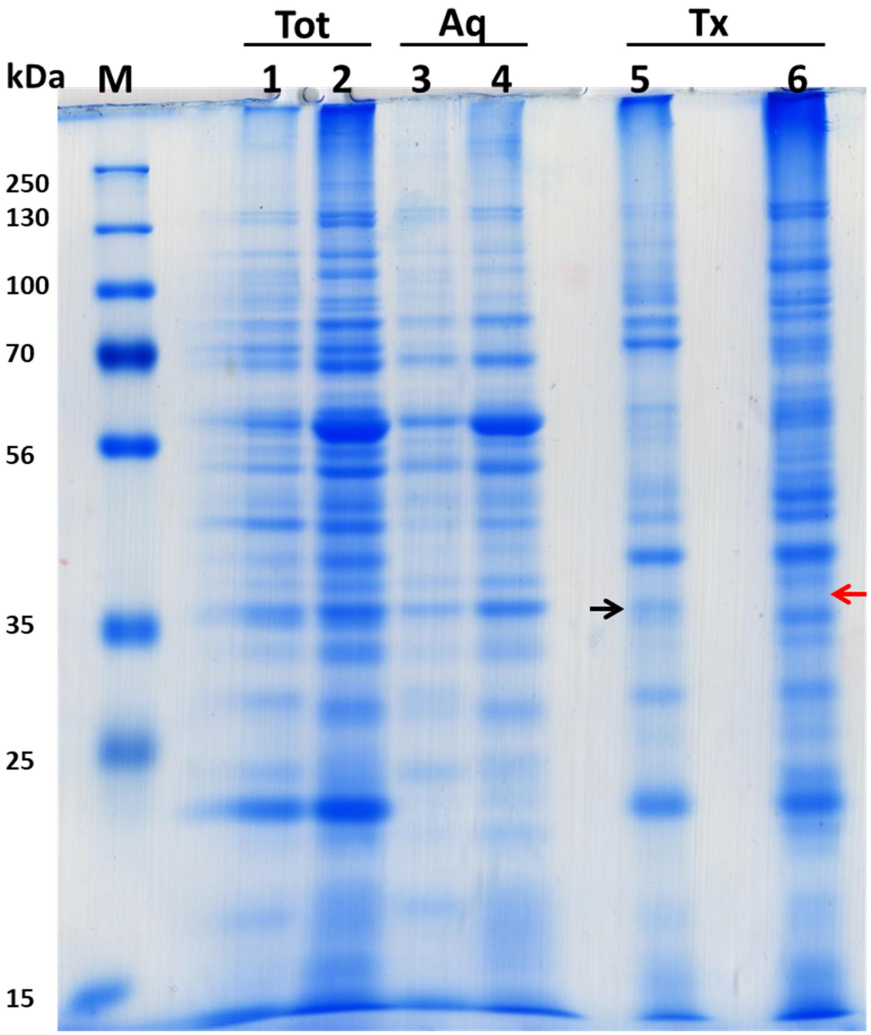

Figure 4.4 Comparative protein profiles of *S. citri* GII-3 and ASP-I strains. Electrophoretic mobility pattern of proteins extracted from *S. citri* GII-3 and ASP-I cells. Lanes 1, 3 and 5: *S. citri* GII-3 cells and lanes 2, 4 and 6: *S. citri* ASP-I cells. Tot: Total lysate, Aq: aqueous soluble fractions, Tx: Triton X-114 soluble fractions. The red arrows point to ~ 39 kDa protein missing from ASP-I cells while black arrow points to corresponding band present in GII-3 cells. It should be noted that the growth rates of GII-3 and ASP-I cells are different and hence, the number of cells used for protein extraction were different. This results into different quantity of proteins seen on the gel. Further, as seen on the gel, protein profiles of the cells under consideration shows several differences in protein pattern.

Table 4.4 List of proteins in *S. citri* ASP-I with significant reduction of expression levels compared to *S. citri* GII-3 cells.

Protein (annotation)	Protein (annotation)
Transketolase (SCITRIASP_01716)	Phosphate ABC transporter substrate-binding protein (SCITRIASP_00831)
Dihydrolipoyl dehydrogenase (SCITRIASP_00818)	ABC transporter permease (SCITRIASP_01316)
2-oxo acid dehydrogenase subunit E2 (SCITRIASP_00819)	Pre-protein translocase subunit SecA (SCITRIASP_00972)
Hypothetical protein (SCITRIASP_01777): lipoprotein	Hypothetical protein (SCITRIASP_01554)
Hypothetical protein (SCITRIASP_00520)	Hypothetical protein (SCITRIASP_01718)
Membrane protein (SCITRIASP_00512)	Hypothetical protein (SCITRIASP_00160)
ABC transporter permease (SCITRIASP_01238)	Hypothetical protein (SCITRIASP_01776)
Lysine--tRNA ligase (SCITRIASP_01173)	Hypothetical protein (SCITRIASP_00257)
ATP-dependent DNA helicase (SCITRIASP_00865)	Hypothetical protein (SCITRIASP_00050)
Chromosome segregation protein SMC (SCITRIASP_00832)	Hypothetical protein (SCITRIASP_01946)
Ribosome recycling factor (SCITRIASP_00748)	Rhomboid family intra-membrane serine protease (SCITRIASP_00467)
DNA topoisomerase IV subunit B protein (SCITRIASP_00340)	Hypothetical protein (SCITRIASP_00578): lipoprotein
Hypothetical protein (SCITRIASP_00323)	Probable dihydrolipoyl dehydrogenase component E3 of pyruvate dehydrogenase trans-membrane protein (SCITRIASP_00818)

Table 4.5 List of proteins not detected in *S. citri* ASP-I cells.

Protein (annotation)	Molecular weight (KDa)
Hypothetical protein (SCITRIASP_00280)	112.6
BMP family ABC transporter substrate-binding protein (SCITRIASP_01143)	51.5
Rod shape determination protein (MreB5)*	**38.6**
ABC transporter permease (SCITRIASP_00243)	31.7
ABC transporter ATP-binding protein (SCITRIASP_00244)	33.9

* Detected as peptides (n=2) that did not extend beyond 14.6 KDa fragment from N-terminus (SCITRIASP_01206)

The proteomic analysis revealed that the *S. citri* ASP-I cells expressed fragments of MreB5 that did not extend beyond the N-terminal 133 amino acids. However, the peptide fragments covering full-length MreB5 protein were found in *S. citri* GII-3 cells. This is consistent with the presence of a stop codon at 134th amino acid of *S. citri* ASP-I *mreB5* gene.

This prompted the analysis of important cytoskeletal proteins Fibril, MreBs 1-4, FtsA, FtsZ, and of the major membrane protein spiralin. All these proteins were found to be present in both the cell types.

This strengthened the hypothesis that absence of a full-length MreB5 protein led to loss of helicity and motility of *S. citri* ASP-I cells. To confirm that MreB5 is indeed responsible for the altered phenotype of *S. citri* ASP-I cells, complementation experiments were performed.

4.3.4) *S. citri* ASP-I cells became helical and motile upon vector-based expression of *S. citri* GII-3 *mreB5*.

The observation of defective *mreB5* gene in *S. citri* ASP-I led to the postulation that MreB5 protein contributes to helical shape determination and motility in *Spiroplasma*. Thus, wildtype *mreB5* gene (from *S. citri* GII-3) was cloned into pSD4 vector using the promoter of a constitutively expressing gene, *ef-tu* (elongation factor Tu). Sequence of the cloned *tuf* promoter is shown in Figure 4.5. Sequencing of the clone confirmed the presence of wildtype *mreB5* gene (from *S. citri* GII-3) downstream of the *tuf* promoter sequence. Hence, this was used for expressing full-length MreB5 in *S. citri* ASP-I cells.

```
5′GTTTAATCTACCATTACTAAAAAAAGATTGCAAAGTAAAAAGAATTAATTTAC
AATAGTTATGACATTTGTCTATTAATTAAAAATATAGGAGGAAAAGATTAAA 3′
```

Figure 4.5 Complementation of *S. citri* ASP-I cells. Sequence of *tuf* promoter (5′→3′) used for expression of *S. citri* GII-3 *mreB5* gene in ASP-I cells. The *mreB5* gene from *S. citri* GII-3 cells was cloned downstream of *tuf* promoter in pUC18 vector between BamHI sites.

The *S. citri* ASP-I cells transformed with empty pSD4 vector displayed colonies like a fried-egg without any satellite colonies (Figure 4.6 A and D). In contrast, *S. citri* ASP-I cells transformed with pSD4 vector containing *tuf* promoter-*mreB5* (pSD4- *mreB5*), displayed diffuse colony morphology (Figure 4.6 B and E) similar to motile *S. citri* GII-3 (Figure 4.6 C and F).

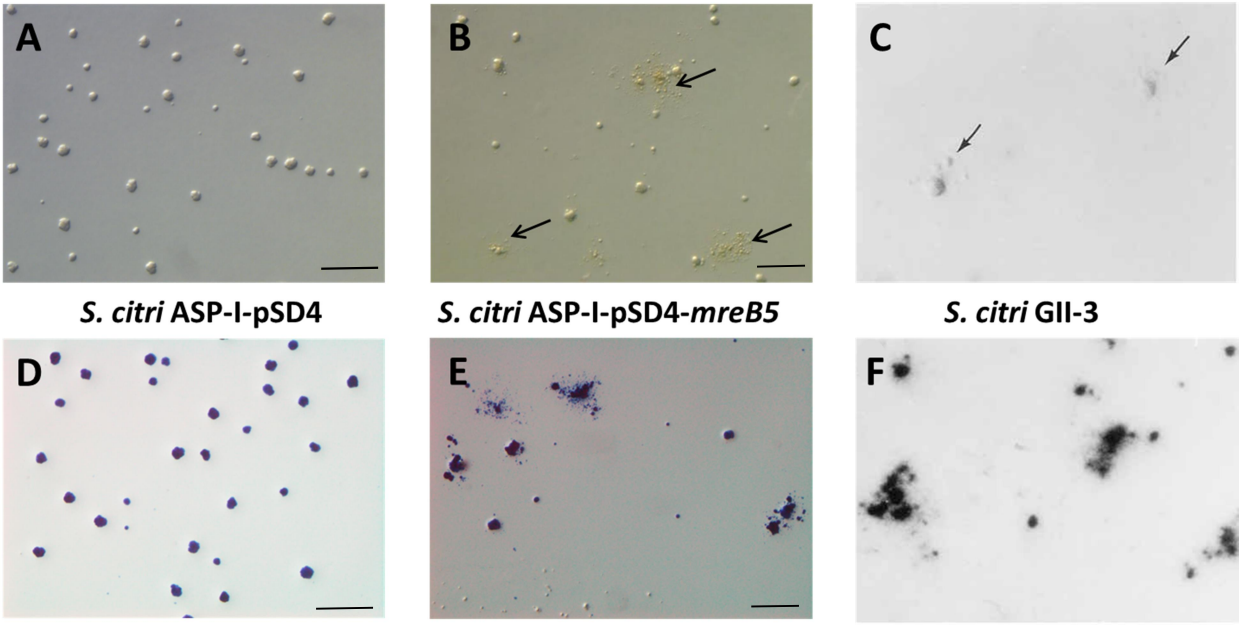

Figure 4.6 Colony morphology of *Spiroplasma citri* cells. Transformation of *S. citri* ASP-I cells with pSD4 empty vector resulted into formation of colonies with "fried-egg' morphology (A & D) similar to untransformed ASP-I cells. Vector-based expression of *mreB5* under *tuf* promoter resulted into formation of diffuse colonies by *S. citri* ASP-I cells (B & E). The diffuse colony morphology of rescued ASP-I cells was similar to wildtype *S. citri* GII-3 cells (C & F). The arrows point to satellite colonies that were formed as a result of translational motility in agar. The lower panel (D to F) shows *S. citri* colonies stained with Dienes stain. Images C and F have been adopted from (Jacob et al., 1997). Scale bars in A, B, D and E represents 100 µm.

These results indicate that the expression of a functional *mreB5* in *S. citri* ASP-I cells resulted into instating motility in them as evidenced from translational motility in agar leading to formation of diffuse colonies.

Microscopic observation of the cells *S. citri* ASP-I cells transformed with empty pSD4 vector show non-motile cells with non-helical rod-like morphology (Figure 4.7 A). In contrast, ASP-I cells transformed with pSD4-*mre*B5 revealed helical morphology (Figure 4.7 B) and kinking motility similar to *S. citri* GII-3 cells (Figure 4.7 C).

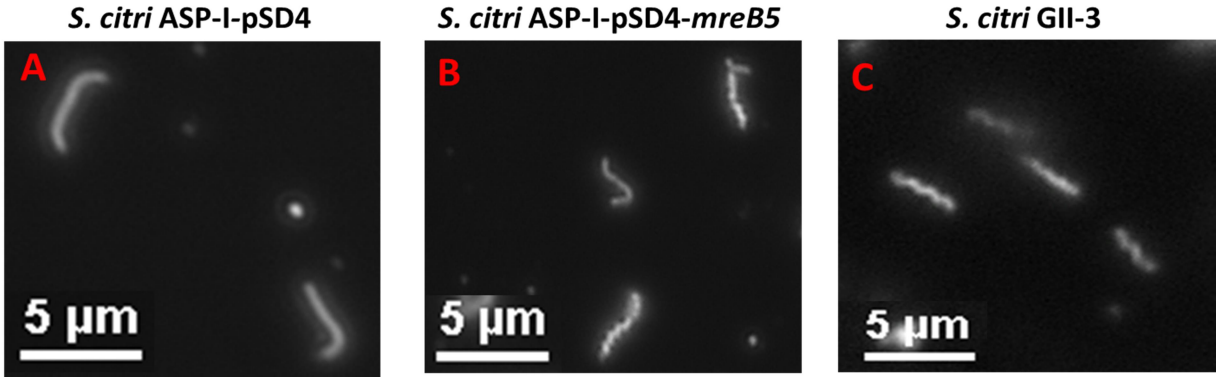

Figure 4.7 Visualization of morphology of *S. citri* cells by dark-field microscopy. *S. citri* ASP-I cells transformed with empty pSD4 vector displayed cells with rod-shaped, non-helical morphology (A). However, vector-based expression of functional *mreB5* (pSD4-*mreB5*) conferred helical morphology to *S. citri* ASP-I cells (B) which was similar to *S. citri* GII-3 cells (C). (Image courtesy: Laure Béven).

Since *S. citri* ASP-I transformants with striking similarity to *S. citri* GII-3 cells have been obtained, it became necessary to confirm that the helical and motile cells are complemented ASP-I cells and not a contamination *S. citri* GII-3 during experimental procedure.

4.3.5) PCR analysis confirmed that the complemented cells are indeed rescued *S. citri* ASP-I cells

The primers SR14 and SR16 are complementary to regions in *S. citri* genomic DNA separated by 574 bps. Amplification of the 574 bp region using these primers against *S. citri* GII-3, *S. citri* ASP-I-pSD4 and *S. citri* ASP-I-pSD4-*mreB5* cells was observed (Figure 4.8 A). Since the complementary regions of SR14 and SR16 primers are not present in TE or pSD4 vector, no amplification was observed using these as the template. Thus, this PCR confirmed that there are no PCR inhibitors in *S. citri* cells.

The possibility of *S. citri* ASP-I cell growth due to spontaneous resistance on media containing tetracycline was ruled out by performing PCR using primers (Tet1-Tet2) complementary to regions in *tetM* gene. As expected, no amplification was observed using *S. citri* GII-3 cells or TE as the template. The observation of 534 bp bands using pSD4 vector, *S. citri* ASP-I-pSD4 and *S. citri* ASP-I-pSD4-*mreB5* as template confirmed that these cells have not developed spontaneous resistance to tetracycline (Figure 4.8 B).

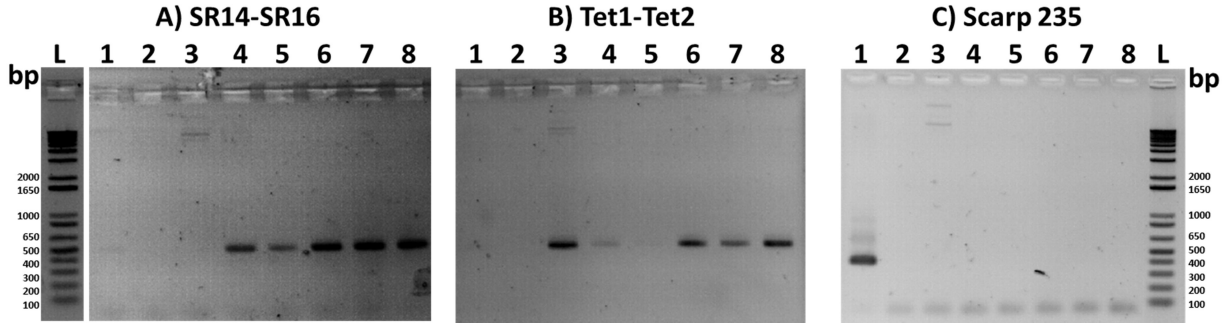

Figure 4.8 Confirmation of *S. citri* ASP-I transformants by PCR method. PCRs were performed using various templates- 1: *S. citri* GII-3 cells; 2: TE; 3: pSD4 empty vector; 4-5: *S. citri* ASP-I-pSD4 cells and 6-8: *S. citri* ASP-I-pSD4-*mreB5* cells. A) Amplification of 574 bp region from genomic DNA of *S. citri* GII-3 (faint band; lane 1), ASP-I-pSD4 cells (lanes 4 & 5) and *S. citri* ASP-I-pSD4-*mreB5* cells (lanes 6-8). No amplification was observed on using TE or pSD4 empty vector (lane 3). B) Confirmation of presence of *tetM* gene using gene specific primers. No amplification was observed for *S. citri* GII-3 cells (lane 1) or TE (lane 2) whereas, a band of 574 bp was observed for pSD4 vector (lane 3) and *S. citri* ASP-I cells transformed with empty pSD4 vector (lanes 4 & 5) or pSD4 vector containing *mreB5* gene (6-8). PCR using Scarp 235 primers and *S. citri* GII-3 cells (lane 1) resulted into amplification of 400 bp band. No amplification was observed in any other lane (lanes 2-8) confirming the absence of pSci plasmids in samples used as template (lanes 2-8), other than *S. citri* GII-3 cells. (Image courtesy: Sybille Duret)

S. citri ASP-I cells lack pSci plasmids (that carry genes coding for SCARPs) that are present in *S. citri* GII-3 cells. Thus, PCRs specific for these plasmids were performed to confirm that the transformants are not *S. citri* GII-3 cells. PCRs using Scarp 235 primers against *S. citri* GII-3 cells resulted into amplification of a major 379 bp band (specific to Scarp3 family). No amplification was found using TE, pSD4 vector, *S. citri* ASP-I-pSD4 and *S. citri* ASP-I-pSD4-*mreB5* as template (Figure 4.8 C). Thus, the possibility of contamination of *S. citri* ASP-I cells complemented with pSD4-*mreB5* was ruled out. This confirmed that the helicity and motility instated in *S. citri* ASP-I-pSD4-*mreB5* cells is due to expression of full-length MreB5.

4.3.6) Cytoskeletal organization of *S. citri* GII-3 and ASP-I cells appears different

MreB and Fibril filaments have been reported to be present just beneath *Spiroplasma* cell membrane (Kürner et al., 2005; Trachtenberg et al., 2008).Which of the MreBs constitute the cytoskeletal filaments is unknown. To decipher if there is any deviation in organization of cytoskeletal filaments in *S. citri* ASP-I cells as compared to GII-3 cells, electron cryotomography was performed on the two cell types. As explained earlier, ECT facilitates visualization of

intracellular macromolecular complexes, filaments. Thus attempts were made for visualization of cytoskeleton in *Spiroplasma* cells. The details are as follows.

4.3.6.1) *Spiroplasma* cells do not withstand blotting using Vitrobot™

Initial trials of freezing cells using Vitrobot™ resulted into disintegration of *Spiroplasma* cells into membrane vesicles (Figure 4.9). This was thought to have resulted from the force applied for blotting off excess of liquid from the grid, a standard procedure of freezing samples for ECT studies. Prior to blotting, samples were confirmed to contain helical (*S. citri* GII-3) or rod-shaped (*S. citri* ASP-I) cells. In absence of cell wall, the cells probably lost shape during sample blotting using Vitrobot™, where pressure is applied during blotting. Freezing samples using 0 (zero) blotting force also failed to preserve cell morphology. To overcome the problem associated with preservation of cell morphology during freezing, following modifications were carried out.

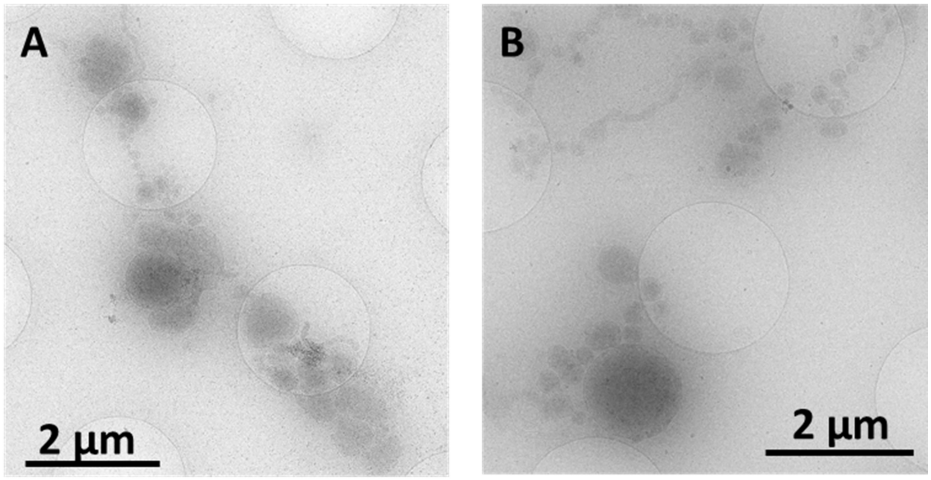

Figure 4.9 Automated blotting of *Spiroplasma* cells for electron cryotomography studies. *Spiroplasma* cells did not withstand blotting by Vitrobot™. *S. citri* GII-3 (A) as well as ASP-I (B) cells disintegrated into membrane vesicles when blotted using Vitrobot™

a) **Use of Teflon disc:** The blotting paper in the Vitrobot™ facing sample on grid was replaced with a teflon disc to avoid excessive blotting that might have resulted into loss of shape due to cell dehydration. However, the natural morphology of *Spiroplasma* cells could not be retained using Teflon disc.

b) **Use of Dextran:** Dextran has been reported to be a useful cryoprotectant (Matias et al., 2003). Thus, freezing *Spiroplasma* cells using 10% and 2% (w/v) dextran concentrations was attempted. However, the dextran concentration seemed to be too high to visualize samples

from the grid. Next, blotting the excess culture media from samples on grid was attempted manually.

4.3.6.2) *Spiroplasma* cells require manual blotting procedure for preserving cell morphology

Spiroplasma cells frozen on grids by manual blotting displayed their natural morphology in electron cryomicroscope; helical for *S. citri* GII-3 cells and straight rod-like for ASP-I cells (Figure 4.10). The images in Figure 4.10 have been shown as representative images for morphology preservation of *Spiroplasma* cells. However, cells such as the one shown in 4.10 (A) are dehydrated and hence unsuitable for data collection. Isolated single, hydrated cells frozen in vitreous ice on these grids were used for data collection.

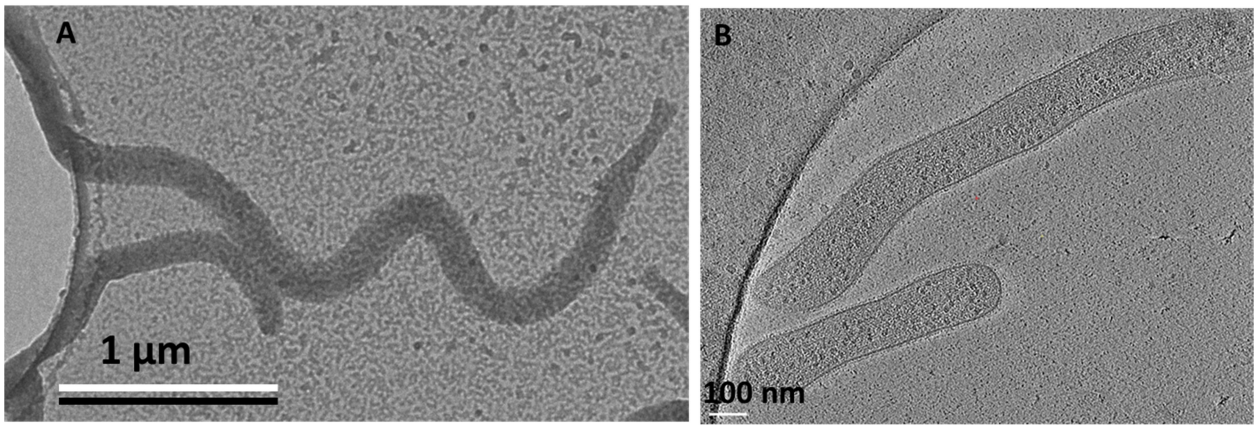

Figure 4.10 Manual blotting of *Spiroplasma* cells helps preserve natural morphology. *S. citri* GII-3 (A) and ASP-I (B) cells were obtained in their natural morphology by blotting manually and plunge-frozen using Vitrobot™. Isolated cells present on frozen grids are shown in the panels.

4.3.6.3) Cytoskeletal ribbon of *Spiroplasma citri* GII-3 cells passes through the shortest intra-cellular path

Initial data was collected on *S. citri* GII-3 cells at a defocus value of -10 µm. The tomograms reconstructed from this data showed the presence of filament sets along the cell length through shortest path (Figure 4.11 A & B).

However, the resolution of this data was insufficient to resolve individual filaments/ filament sets. Thus, data on *S. citri* ASP-I cells was collected by changing the defocus and at higher magnification.

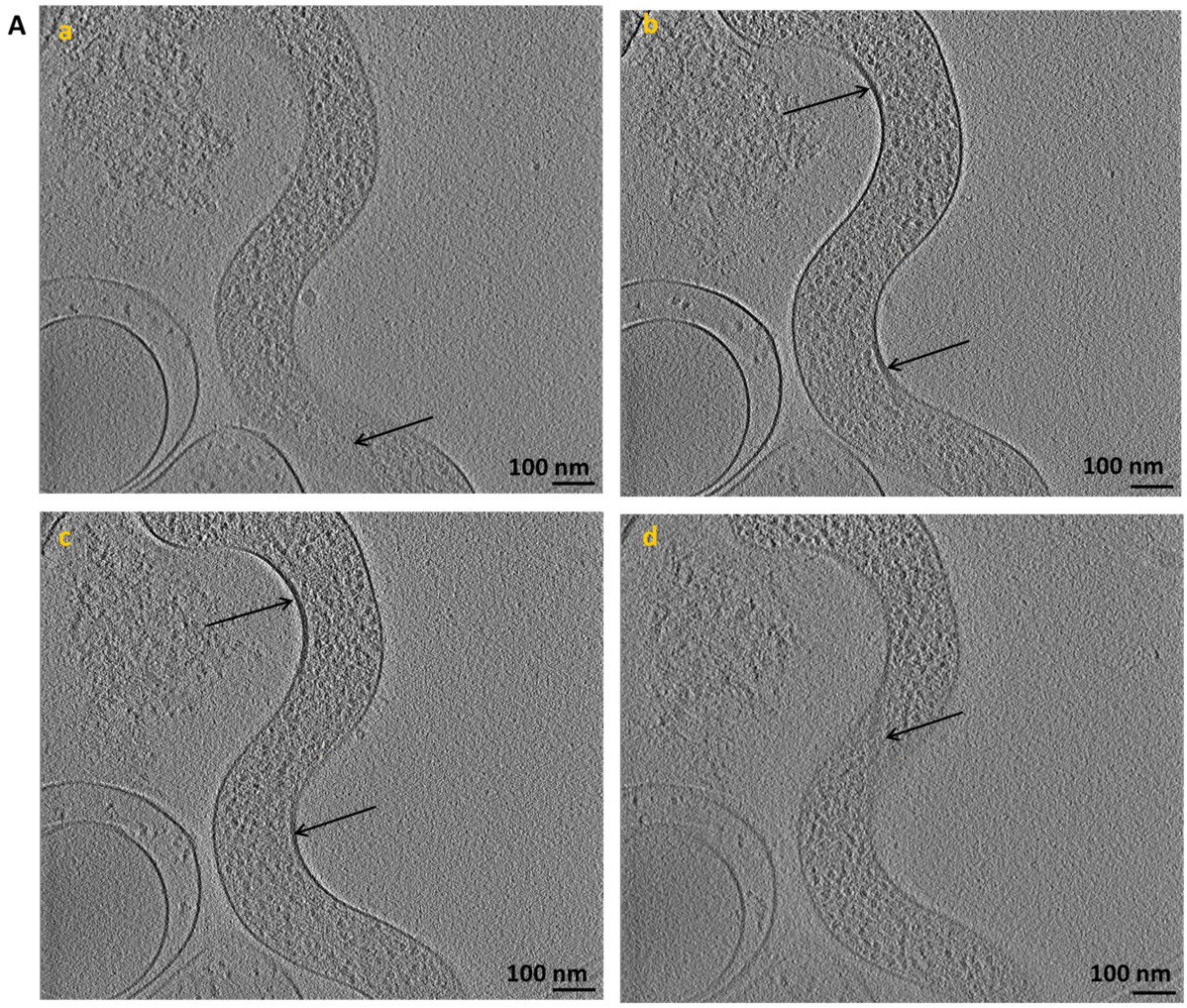

Figure 4.11 Cytoskeletal ribbon in *Spiroplasma citri* GII-3 follows the shortest path along the cell length. A) (above panel) and B) (refer next page) panels show sections (a to d) of the tomograms of *S. citri* GII-3 cells. The arrows point to the ribbon consisting of protein filaments. The ribbon is seen to pass through the shortest distance along cell length. Panel (B) of the figure is given in the next page.

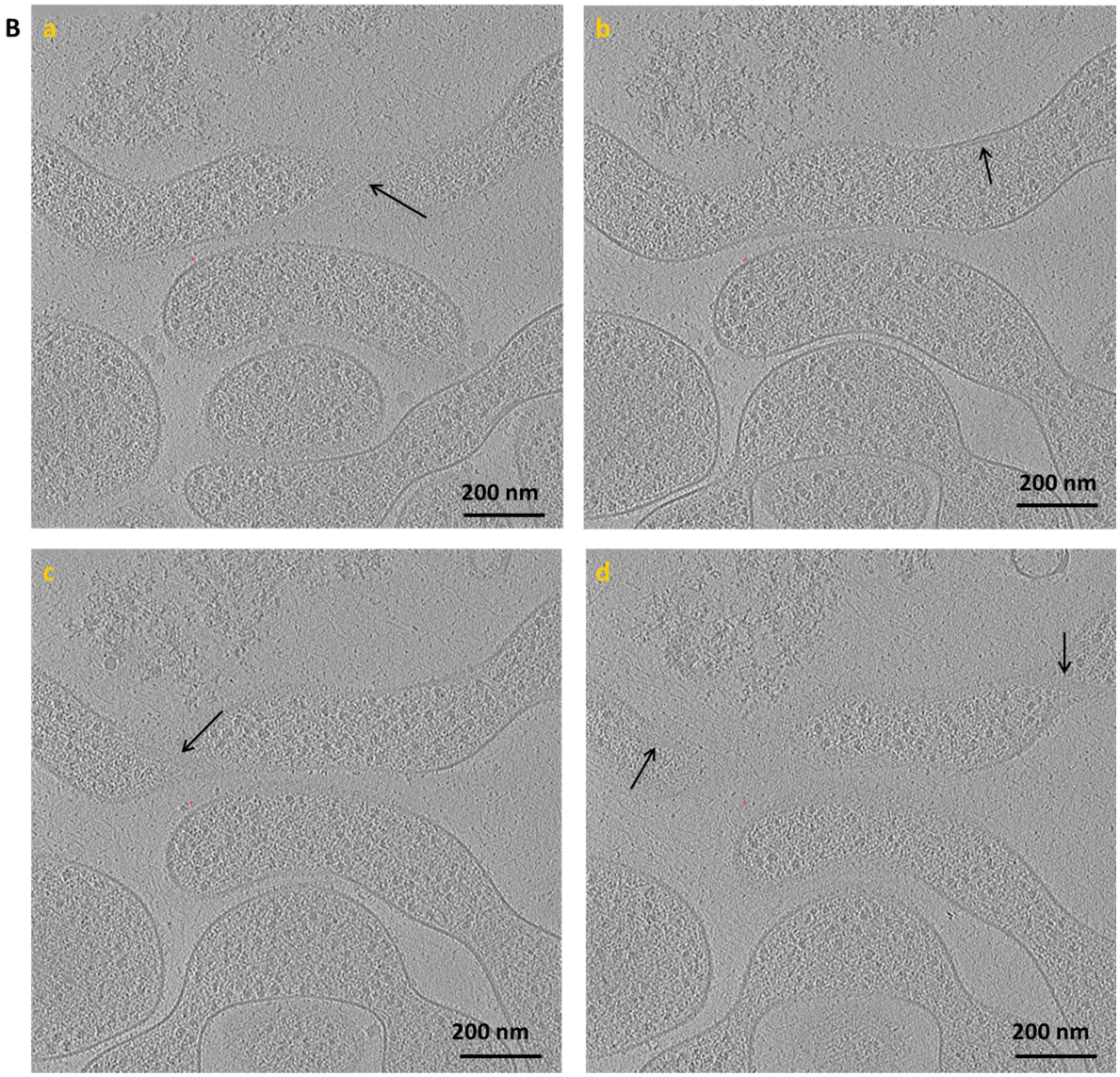

Figure 4.11 Continuation of the figure from previous page (panel B)

4.3.6.4) One set of filaments are seen inside *S. citri* ASP-I cells as compared to three sets reported for wildtype cells

In order to obtain detailed information regarding number of protein filaments with the cell, data was collected on *S. citri* ASP-I cells at a defocus value of -7 µm. The tomograms reconstructed using this data showed presence of only one set of filaments in the ASP-I cells (Figure 4.12 A).

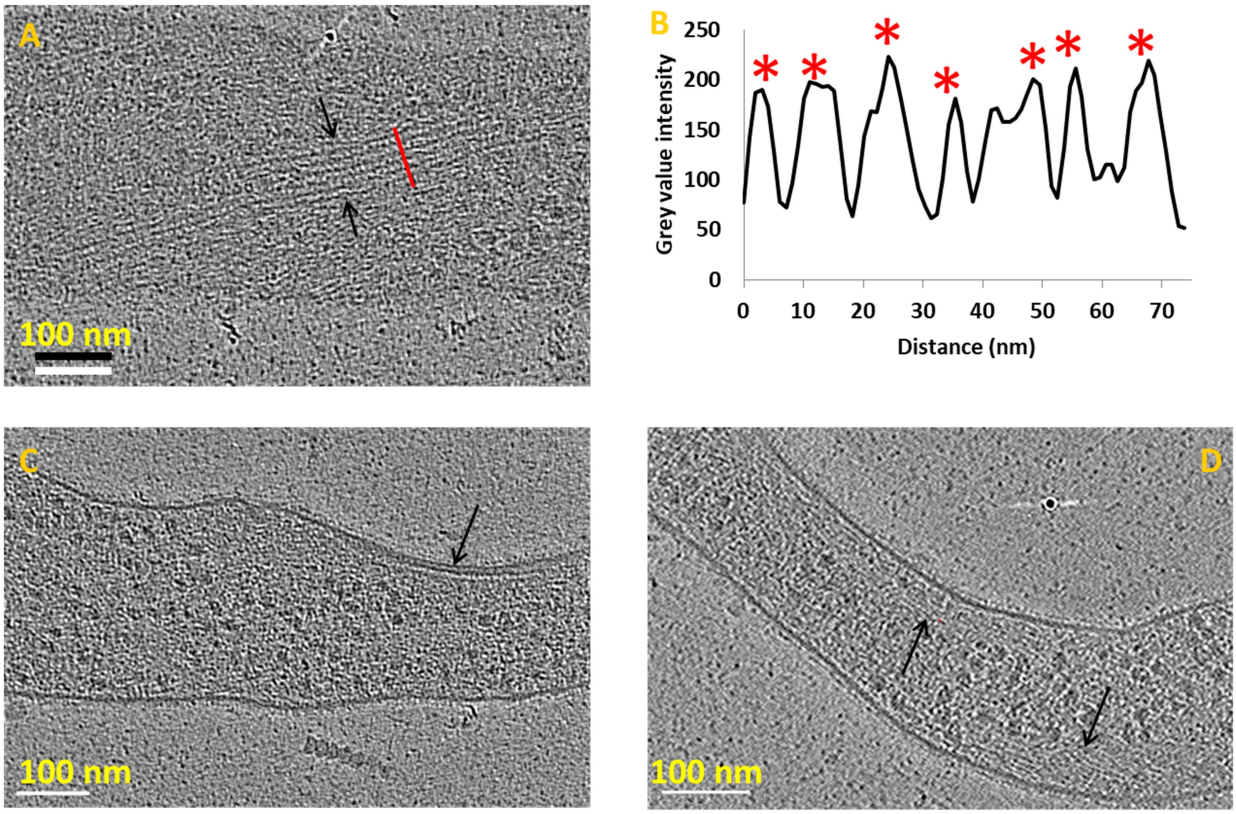

Figure 4.12 Visualization of *S. citri* ASP-I cytoskeleton. Sections from tomograms of multiple *S. citri* ASP-I cells showing localization of cytoskeletal ribbon (A, C and D). The filaments are clearly seen between the black arrows in (A). The intensity plot corresponding to red line in (A) is shown in (B). The filaments have an inter-filament spacing of ~ 10 nm as seen between the peaks (red asterisk) of intensity plot. The filament bundle close to the cell membrane, along the cell length is pointed by a black arrow in (C). The filaments in (D) are shown by arrows.

The inter-filament spacing between the filaments in these tomograms was observed to be about 10 nm (Figure 4.12 B). The inter-filament spacing of 11 nm has been earlier reported for Fibril filaments (Kürner et al., 2005). Thus, the filaments observed in *S. citri* ASP-I cells must be made of Fibril. No other filaments could be observed in tomograms of *S. citri* ASP-I cells.

Further studies for data collection on *S. citri* GII-3 cells at defocus value of -7 µm could not be carried out due to limited time availability on Titan Krios. Attempts to obtain high resolution tomograms to visualize and identify Fibril and MreB filament organization in *S. citri* GII-3 are in progress.

4.4) Discussion

Studies on cell walled bacteria have revealed that most of them acquire non-spherical shapes by regulating cell wall synthesis with the help of MreB. Bacteria such as *C. crescentus* and *H. pylori* modify their rod shape, into crescent and helical respectively, by creating a bias in peptidoglycan insertion. While *C. crescentus* utilizes crescentin and MreB as the components for creating bias, *H. pylori* follows an MreB-independent mechanism. However, how cell wall deficient bacteria maintain non-spherical shapes is unknown.

In the present study comparative genomic, proteomic and electron cryotomography studies on *S. citri* GII-3 (wildtype; helical, motile) and ASP-I cells (mutant; non-helical, non-motile) were performed. A truncated MreB5 protein, due to a point mutation in its gene, was identified and postulated as the cause of loss of helicity and motility in *S. citri* ASP-I cells. Expression of a functional MreB5 protein led to recovery of helicity and motility in *S. citri* ASP-I cells. This implicates that MreB5 protein plays a role in morphology as well as motility of *Spiroplasma*. It is currently not known if motility of *Spiroplasma* is dependent on helical cell shape. However, *prima facie* it appears that the characteristic kinks needed for *Spiroplasma* motility cannot occur without the helical shape of cells. This is also supported by the fact that all the reported naturally occurring motile *Spiroplasmas* are helical in shape. Thus, acquisition of helical shape must be a pre-requisite for *Spiroplasma* cells to exhibit motility. However, mere helical morphology is not sufficient to bring about *Spiroplasma* motility since, a helical but non-motile mutant was prepared by Jacob et al., (1997). Additional components, apart from those involved in helical shape determination, are likely involved in *Spiroplasma* motility machinery. One such protein which is likely involved in is the protein product of gene *scm1* (Jacob et al., 1997). Present study found the *scm1* gene to be present in *S. citri* ASP-I as well as *S. citri* GII-3 cells.

The electron cryotomography of *S. citri* ASP-I cells showed that these cells possess only one set of filaments. At this stage it is difficult to comment on differences in cytoskeletal organization of Fibril and MreB filaments between *S. citri* GII-3 and ASP-I cells since filaments in *S. citri* GII-3 cells could not be resolved. Obtaining high resolution tomograms of *S. citri* GII-3 would be useful to observe the organization of MreB5 and Fibril filaments as the cytoskeleton of *S. citri*. *S. citri* GII-3 tomograms would be valuable to verify models proposed by Kürner et al., (2005) (3

set of filaments) and Trachtenberg et al., (2008) (only one filament set) for organization of *Spiroplasma* cytoskeleton.

An important discovery of the present study is the point mutation in *mreB5* of *S. citri* ASP-I cells. To the best of my knowledge, this is the first report on function of an MreB protein in *Mollicutes*. It should be noted that among *Mollicutes*, only members of genus *Spiroplasma* and *Haloplasma contractile* have been reported to possess *mreB* genes (Ku et al., 2014). This finding provides following hints regarding the role of MreB in *Spiroplasma*, a cell wall-deficient bacteria:-

a) **MreB5 plays a novel function in *Spiroplasma*:** The expression of a full-length MreB5 protein in *S. citri* ASP-I cells results into gain of helical shape. This suggests that MreB5 does not play a role in rod-shape determination of Spiroplasma, but is essential for the rod-to-helical transition of *Spiroplasma*.
b) **MreBs have non-overlapping function in *Spiroplasma*:** Even though 4 of the 5 MreBs in *S. citri* ASP-I cells are similar to wildtype cells, none of these 4 MreBs could substitute for the function of MreB5. This indicates that MreBs may have non-overlapping functions in *Spiroplasma*. However, experiments need to be performed using single (*mreB*) gene deletion mutants to ascertain functions of each of the MreBs. Such a study would also be useful in identifying if any of the MreBs have partial or complete overlap of function(s).
c) **Role of MreB5 in *Spiroplasma* motility:** Microscopy observations of *S. citri* ASP-I cell suggested that they produce only erratic twitching movements in contrast to rotational movement and flexing as a part of motility Townsend et al., (1977).

   Observations in present study suggest that the *S. citri* ASP-I cells have severely compromised movements compared to the *S. citri* GII-3 cells. This is consistent with observations by Townsend et al., (1977). The recovery of wildtype *Spiroplasma*-like motility in ASP-I cells provided with full-length MreB5 also suggests that MreB5 may contribute to rotational and flexing movements required for *Spiroplasma* motility.

Thus, like cell walled bacteria, *Spiroplasma* also seems to utilize MreB for morphogenesis. As mentioned earlier, the inability of MreBs 1-4 to complement function of MreB5 in *S. citri* ASP-I cells prompted me to speculate that the MreBs have different functions. Hence, genetic studies

were performed on *S. citri* GII-3 cells to generate mutants that are deficient of one of the MreBs at a time. Also, to understand the role of Fibril, we decided to prepare *Δfibril* mutant of *Spiroplasma citri* GII-3 cells. The process of preparation of *Spiroplasma* mutants, their characterization and associated challenges have been summarized in Chapter 5 of this thesis.

4.5) References

Basler, M., Pilhofer, M., Henderson, G.P., Jensen, G.J., and Mekalanos, J.J. (2012). Type VI secretion requires a dynamic contractile phage tail-like structure. Nature *483*, 182–186.

Berho, N., Duret, S., Danet, J., and Renaudin, J. (2006). Plasmid pSci6 from *Spiroplasma citri* GII-3 confers insect transmissibility to the non-transmissible strain *S. citri* 44. Microbiology *152*, 2703–2716.

Boudet, J.F., Mathelie-Guinlet, M., Vilquin, A., Douliez, J.P., Béven, L., and Kellay, H. (2018). Large variability in the motility of spiroplasmas in media of different viscosities. Sci. Rep. 1–14.

Briegel, A., Dias, D.P., Li, Z., Jensen, R.B., Frangakis, A.S., and Jensen, G.J. (2006). Multiple large filament bundles observed in *Caulobacter crescentus* by electron cryotomography. Mol. Microbiol. *62*, 5–14.

Briegel, A., Li, X., Bilwes, A.M., Hughes, K.T., Jensen, G.J., and Crane, B.R. (2012). Bacterial chemoreceptor arrays are hexagonally packed trimers of receptor dimers networked by rings of kinase and coupling proteins. Proc. Natl. Acad. Sci. U. S. A. *109*, 3766–3771.

Deeley, J., Stevens, W.A., and Fox, R.T. V. (1979). Use of Dienes' stain to detect plant diseases induced by *Mycoplasma*-like organisms. Phytopathology *69*, 1169–1171.

Duret, S., Danet, J.-L., Garnier, M., and Renaudin, J. (1999). Gene disruption through homologous recombination in *Spiroplasma citri*: An scm1-disrupted motility mutant is pathogenic. J. Bacteriol. *181*, 7449–7456.

Duret, S., Berho, N., Danet, J., Garnier, M., and Renaudin, J. (2003). Spiralin Is Not Essential for helicity, motility or pathogenicity but is required for efficient transmission of *Spiroplasma citri* by its leafhopper vector *Circulifer haematoceps*. Appl. Environ. Microbiol. *69*, 6225–6234.

Galaz-Montoya, J.G., and Ludtke, S.J. (2017). The advent of structural biology in situ by single particle cryo- electron tomography. Biophys. Reports *3*, 17–35.

Jacob, C., Nouzières, F., Duret, S., Bové, J.M., and Renaudin, J. (1997). Isolation, characterization, and complementation of a motility mutant of *Spiroplasma citri*. J. Bacteriol. *179*, 4802–4810.

Ku, C., Lo, W.-S., and Kuo, C.-H. (2014). Molecular evolution of the actin-like MreB protein gene family in wall-less bacteria. Biochem. Biophys. Res. Commun. *446*, 927–932.

Kürner, J., Frangakis, A.S., and Baumeister, W. (2005). Cryo – electron tomography reveals the cytoskeletal structure of *Spiroplasma melliferum*. Science. *307*, 436–439.

Lartigue, C., Duret, S., Garnier, M., and Renaudin, J. (2002). New plasmid vectors for specific gene targeting in *Spiroplasma citri*. Plasmid *48*, 149–159.

Mastronarde, D.N. (2005). Automated electron microscope tomography using robust prediction of specimen movements. J. Struct. Biol. *152*, 36–51.

Mastronarde, D.N., and Held, S.R. (2017). Automated tilt series alignment and tomographic reconstruction in IMOD. J. Struct. Biol. *197*, 102–113.

Matias, R.F., Al-amoudi, A., Dubochet, J., and Beveridge, T.J. (2003). Cryo-transmission electron microscopy of frozen-hydrated sections of *Escherichia coli* and *Pseudomonas aeruginosa*. J. Bacteriol. *185*, 6112–6118.

Radermacher, M. (2007). Weighted back-projection methods. In Electron Tomography, pp. 245–273.

Renaudin, J., Béven, L., Batailler, B., Duret, S., Desqué, D., Arricau-bouvery, N., Malembic-maher, S., and Foissac, X. (2015). Heterologous expression and processing of the flavescence dorée phytoplasma variable membrane protein VmpA in *Spiroplasma citri*. BMC Microbiol. *15*, 1–12.

Schur, F.K.M., Obr, M., Hagen, W.J.H., Wan, W., Jakobi, A.J., Kirkpatrick, J.M., Sachse, C., Kräusslich, H., and Briggs, J.A.G. (2016). An atomic model of HIV-1 capsid-SP1 reveals

structures regulating assembly and maturation. Science. *353*, 506–508.

Stamburski, C., Renaudin, J., and Bove, J.M. (1991). First step toward a virus-derived vector for gene cloning and expression in spiroplasmas, organisms which read UGA as a tryptophan codon: Synthesis of chloramphenicol acetyltransferase in *Spiroplasma citri*. J. Bacteriol. *173*, 2225–2230.

Townsend, R., Markham, P.G., Plaskitt, K.A., and Daniels, M.J. (1977). Isolation and characterization of a non-helical strain of *Spiroplasma citri*. J. Gen. Microbiol. *100*, 15–21.

Townsend, R., Burgess, J., and Plaskitt, K.A. (1980). Morphology and ultrastructure of helical and nonhelical strains of *Spiroplasma citri*. J. Bacteriol. *142*, 973–981.

Trachtenberg, S., Dorward, L.M., Speransky, V. V., Jaffe, H., Andrews, S.B., and Leapman, R.D. (2008). Structure of the cytoskeleton of *Spiroplasma melliferum* BC3 and its interactions with the cell membrane. J. Mol. Biol. *378*, 778–789.

Trachtenberg, S., Schuck, P., Phillips, T.M., Andrews, S.B., and Leapman, R.D. (2014). A structural framework for a near-minimal form of life: mass and compositional analysis of the helical *Mollicute Spiroplasma melliferum* BC3. PLoS One *9*, e87921.

Weiss, G.L., Medeiros, J.M., and Pilhofer, M. (2017). In situ imaging of bacterial secretion systems by electron cryotomography. In Methods in Molecular Biology, pp. 353–375.

Williamson, D.L., Renaudin, J., and Bove, J.M. (1991). Nucleotide sequence of the *Spiroplasma citri* fibril protein gene. J. Bacteriol. *173*, 4353–4362.

Chapter 5

Genetic studies on Spiroplasma to understand the role of Fibril and MreB proteins in shape determination and motility

5.1) Introduction

In the previous chapter we have seen that the absence of a functional MreB5 protein led to loss of helicity and motility of *S. citri* ASP-I cells. Complementation of the cells by vector-based expression of full-length *mreB5* led to regaining helicity as well as motility in *S. citri* ASP-I cells. Thus, MreB5 seems to have a role in shape determination and motility of *Spiroplasma* cells. It is noteworthy that in absence of MreB5 none of the other MreBs (1- 4), predicted to be functional based on sequencing data, could compensate for its role. This suggests that each MreB may have a different role in *Spiroplasma* cells. Thus, it is interesting to delineate role of all the MreBs in *Spiroplasma*.

Cytoskeletal proteins Fibril and MreBs have long been proposed to be associated with *Spiroplasma* motility and shape determination (Kürner et al., 2005). While Fibril has been speculated to display conformational dynamics, MreB filaments are suggested to be responsible for providing elasticity needed by the cell for kink formation (Cohen-Krausz et al., 2011; Kürner et al., 2005). However, experimental evidence in support of roles of these proteins is lacking. In general, *Spiroplasma* cell biology remains poorly understood due to the unavailability of suitable molecular tools and techniques necessary for their genome modification.

Molecular biology tools are necessary to ascertain gene function by generating gene deletion/disruption mutants. Characterization of such mutants and rescuing them by supplying the deleted gene to retain original phenotype confirms the role of the gene. Mutants can be generated either artificially in laboratory or they can be isolated from nature. Naturally occurring mutants can be spotted easily if the phenotype is starkly different from wildtype cells. In cases of naturally occurring mutants, mutation(s) can be identified by genomic, proteomic and biochemical studies as mentioned in the previous chapter. However, it is impossible to always obtain a desired mutant from nature, thus creating a need for artificially producing desired mutants. When altering bacterial genomes, a variety of techniques are available such as exposure

to mutagens (eg ultraviolet light, ethidium bromide, nitrous acid), transposons, homologous recombination and CRISPR-Cas mediated approaches (Eggleston and West, 1996; Jiang et al., 2013; Setlow et al., 1963; Slonimski et al., 1968; Zimmermann, 1977). Since these techniques involve RecA mediated recombination, these options are not facile to perform genetic manipulations in *Spiroplasma* as it lacks RecA (Marais et al., 1996). Thus, genetic studies on *Spiroplasma,* the *Mollicute* of my interest, becomes challenging. As described in Chapter 1 of this thesis, the most efficient way employed till date for generating target protein mutants in *Spiroplasma* is by the use of a disruption vector containing a portion of the target gene (Duret et al., 2005).

This chapter describes my attempts to generate mutants of *Spiroplasma citri* GII-3 cells that are each deficient in one of the five *mreBs* or *fibril*. Characterization of the mutants will help to decipher functions of these proteins in *Spiroplasma* cells and their roles in cell shape determination and motility.

5.2) Material and methods

Generation of gene disruption mutants of *Spiroplasma* is a challenging and time consuming process due to incompatibility of advanced genetic tools and slow growth rate of spiroplasmas. Explained below are the strategies and approaches to generate *mreB* and *fibril* disruption mutants of *Spiroplasma* using the best tools available for their genome modification. The mutants will be useful to decipher functions of respective proteins in spiroplasmas. Due to unavailability of *Spiroplasma* cultures and genetic tools, the work incorporated in this chapter was carried out in the lab of Dr. Laure Béven, INRA, Bordeaux, France during the 6-months tenure of my Raman-Charpak fellowship and currently continued by Ms. Sybille Duret who is an expert in *Spiroplasma* genetics. *S. citri* GII-3 culture available in the lab of Dr. Laure Béven, INRA, Bordeaux, France was used for the study. This culture was grown on modified SP4 medium using growth conditions as described in Chapter 2 of this thesis.

5.2.1) Strategy for mutant generation

Disruption vectors pSD32 and pGOT-res1 were used for preparation of *Spiroplasma* mutants. The vectors were kindly provided by Ms. Sybille Duret, INRA, Bordeaux, France. The details of these vectors (unpublished) are described below.

The vector pSD32 has two resolvase target (*res*) sites (described in section 1.10 and Table 1.3 in Chapter 1) and a tetracycline resistance marker (*tetM*) gene expressible under *spiralin* promoter

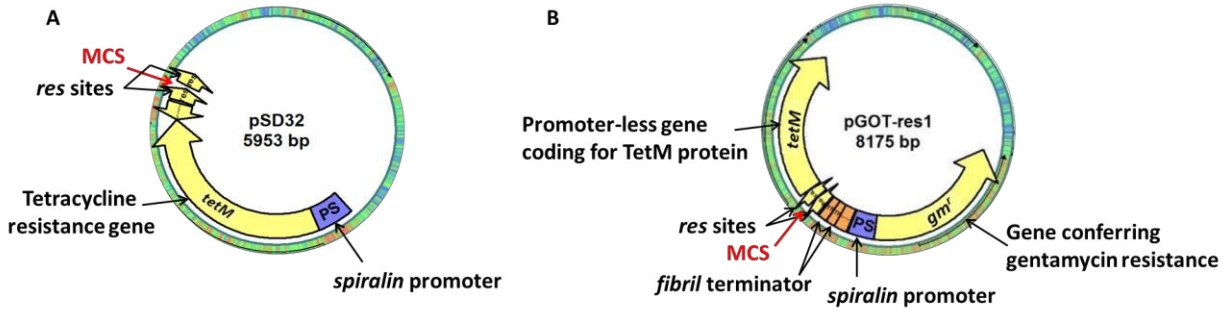

Figure 5.1 Vector maps of *Spiroplasma* gene disruption vectors. A) The pSD32 vector has a tetracycline resistance marker (*tetM*) expressible under *spiralin* promoter (PS). Two resolvase sites (*res*) are present downstream of the *tetM*. A multiple cloning site (MCS; highlighted in red) is sandwiched between the *res* sites. B) The pGOT-res1 vector has two antibiotic resistance markers: gentamycin (*gmr*) and tetracycline (*tetM*). While *tetM* is without a promoter, the *gmr* is expressible under *spiralin* promoter (PS). Upstream of the PS are two *fibril* terminator (TF) sequences to prevent transcription of *tetM* from PS. On the other side of TF are a couple of *res* sites between which is present a multiple cloning site (MCS). Downstream of the res sites is the promoter-deficient *tetM* gene. This arrangement facilitates tetracycline resistance only upon integration of the vector at target gene in genome.

(PS) (Figure 5.1 A). Since *spiralin* is a constitutively expressing gene, the use of *spiralin* promoter ensures transcription of *tetM* in the cell. As explained earlier, *res* sites are useful for removal of antibiotic resistance marker upon obtaining target disruption mutants. Different recombinant constructs of this vector are prepared each using a portion of one of the genes of interest. The recombinant vectors are then transformed into *Spiroplasma* cells and grown on modified SP4 media supplemented with tetracycline (2 µg/mL). The tetracycline resistant colonies from transformation plates are then grown in liquid modified SP4 media in the presence of tetracycline. Transformants are then extensively passaged with increasing concentrations of tetracycline. A passaging step refers to sub-culturing a growing culture from the current liquid medium into a fresh one. When using clones prepared in pSD32 vector, initial passages are carried out in media containing low concentration (2 µg/mL) of tetracycline to facilitate vector multiplication within the cells. After 5 passages, cells are transferred to media containing higher concentrations of tetracycline (15 µg/mL). The higher concentration of tetracycline increases selection pressure on the bacterium to maintain the vector and to facilitate its homologous recombination, with genomic DNA at the target site. Since *Spiroplasma* cells are deficient of a functional RecA, the recombination is a chance event. A fraction of cells (2 mL) from each transformant at every passage are retained to extract genomic DNA. PCRs using genomic DNA

as the template and primers that bind to neighboring regions of the target gene provide information about whether the recombination has occurred at target site or not. Obtaining mutants using this method requires close to 12 months since *Spiroplasma citri* is a slow growing organism. A quicker alternative to this method is to use an alternate vector, pGOT-res1.

pGOT-res1 vector has two antibiotic (gentamycin and tetracycline) resistance markers in addition to two *res* sites (Figure 5.1 B). While gentamycin resistance marker (*gmr*) is placed under constitutively expressing *spiralin* promoter, the tetracycline resistance gene (*tetM*) is deficient of a promoter. The two *res* sites are sandwiched between *tetM* (downstream) and two sets of *fibril* terminator sequences (upstream). This arrangement in the pGOT-res1 helps to prevent transcription of *tetM* in a free vector by *lacZ* promoter and facilitates *tetM* transcription using only the target gene promoter upon integration of the vector at target site in genomic DNA (Lartigue et al., 2002). Once the fragment of gene of interest is cloned between *res* sites, the vector is transformed into *Spiroplasma* cells and grown on media supplemented with gentamycin. The transformant colonies are then inoculated into liquid media containing gentamycin. 3-5 passages are carried out in gentamycin containing liquid media to facilitate amplification of the vector in transformants. The transformants are then transferred to media supplemented with tetracycline. Tetracycline resistance markers can be expressed only using promoter of target gene in the genome upon integration. This would help obtain mutants with recombinant vector integrated at the target site since *tetM* can only be transcribed using target gene promoter upon integration into genomic DNA. Obtaining gene disruption mutants by this method would require about 6 months, less than the time required using pSD32 vector. A major advantage of using pGOT-res1 vector over pSD32 is that even after 15 passages using latter, the cells without recombination can grow using free plasmids. However, the mutants obtained using pGOT-res1 vector will be able to grow in media containing high concentration (10-15 µg/mL) of tetracycline only if the vector integrates into genomic DNA. Mutants obtained in the present study by any of these methods can be used for understanding functions of target proteins.

Mutants obtained by integration of the disruption vector into genome (Figure 5.2 A) can be further processed for removal of resistance gene from genome by vector-based expression of Resolvase protein (tnpR) using a different vector (pSD26; Figure 5.2 B). Removal of the

resistance gene(s) helps obtain clean mutants that can be used for further mutation at another site, if required, using same resistance marker.

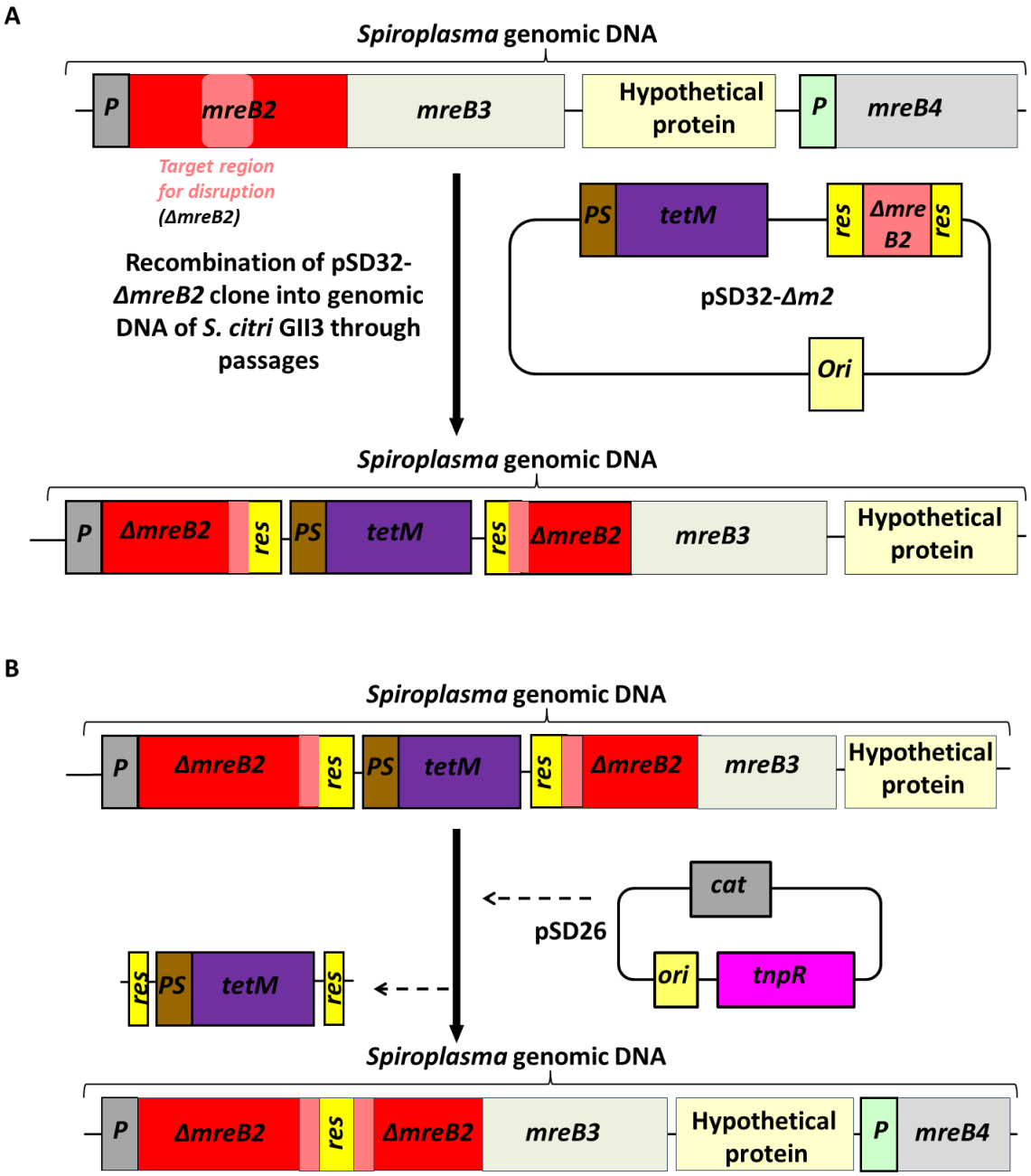

Figure 5.2 Flowchart for gene disruption in *Spiroplasma*. A) The region of interest (RoI) from *mreB2* (*ΔmreB2*; pink) is cloned in pSD32 vector between the two *res* sites to generate pSD32-Δm2 clone. Multiple passages of *S. citri* cells carrying pSD32-Δm2 clone facilitates integration at the *mreB2* gene in the genomic DNA. The integration results into disruption of *mreB2* gene. B) The antibiotic resistance marker integrated into genomic DNA can be removed by expression of Resolvase protein using pSD26 clone containing *tnpR* (gene coding for Resolvase) and chloramphenicol resistance marker (*cat*). The Resolvase protein cleaves at the *res* sites and removes DNA between two *res* sites. This leads to production of unmarked mutants.

To obtain mutants, portion from gene of interest must first be cloned into the disruption vectors. A total of 6 mutants are to be prepared: 5 mutants each with one of the *mreBs* (NCBI gene accession numbers CAK98230, AAL51057, AAL51058, CAK98230 and CAK98231 for mreBs 1 to 5 respectively) disrupted and the sixth one with *fibril* (NCBI accession number CAK99396) disruption. Since *Spiroplasma mreBs* share high sequence identity (at least 55 % identical; Table 5.1) with each other, the regions to be disrupted were carefully selected as described in the following section.

Table 5.1 Percent identity matrix of *S. citri mreB* sequences obtained using Clustal Omega tool (https://www.ebi.ac.uk/Tools/msa/clustalo/).

	Sc_*mreB*1	Sc_*mreB*2	Sc_*mreB*3	Sc_*mreB*4	Sc_*mreB*5
Sc_*mreB*1	100.00	63.98	55.68	80.69	65.12
Sc_*mreB*2		100.00	57.95	65.71	68.79
Sc_*mreB*3			100.00	56.10	61.85
Sc_*mreB*4				100.00	65.70
Sc_*mreB*5					100.00

5.2.2) Selection of region of interest for targeting a specific *mreB*

MreB is a bacterial homolog of eukaryotic actin responsible for maintenance of non-spherical cell shape in most bacterial species (van den Ent et al., 2001). Eukaryotic actin and some of its bacterial homologs have been well characterized (Wagstaff and Löwe, 2018). The information regarding conserved motifs and residues critical to functions of actin is available in literature (Bork et al., 1992). Therefore, sequences of *S. citri* GII-3 MreB were aligned with eukaryotic (*H. sapiens*) actin and its bacterial (*E. coli*) homologue MreB to identify conserved motifs (Figure 5.3). Targeting conserved motifs from MreB for disruption would be the best approach to obtain MreB-deficient mutants. Since *S. citri* GII-3 *mreB* genes share at least 55% sequence identity (Table 5.1), regions with sequence variations were identified for cloning into disruption vectors. Due consideration was also given to the minimum length (> 220 bp) of target gene required for gene disruption to work in *Spiroplasma* (Duret et al., 2005). Taking into account all the aforementioned criteria, a 263 bp region around nucleotides coding for phosphate 2 was selected as target region for disruption (Figure 5.4).

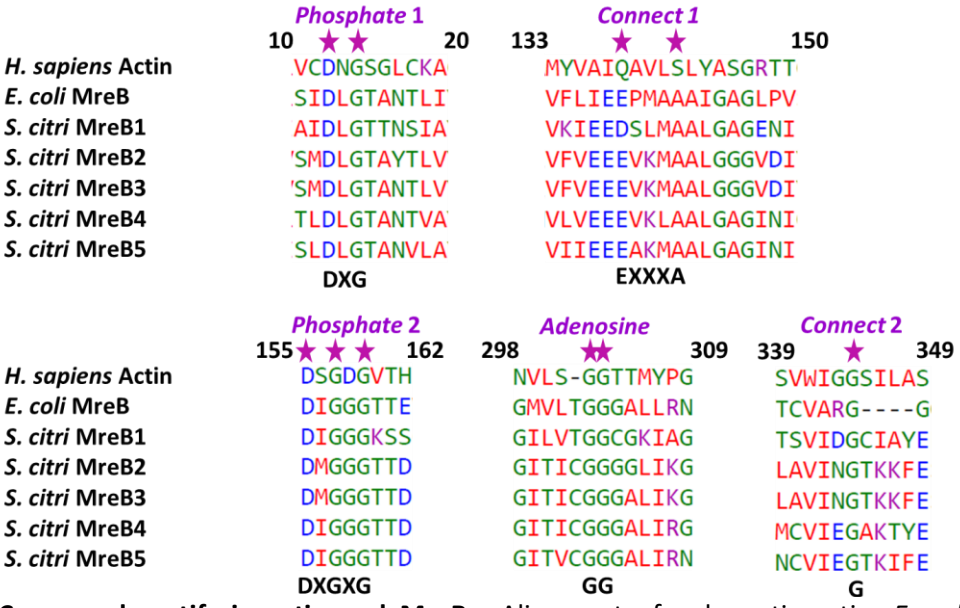

Figure 5.3 Conserved motifs in actin and MreBs. Alignment of eukaryotic actin, *E. coli* MreB and *Spiroplasma citri* MreB sequences with conserved motifs and residues highlighted using Kalign (https://www.ebi.ac.uk/Tools/msa/kalign/). Purple asterisks indicate residues involved in nucleotide binding. Presented in black, bold letters below the alignment are the conserved motifs and corresponding labels are written in purple on top. Residue numbering is according to actin sequence.

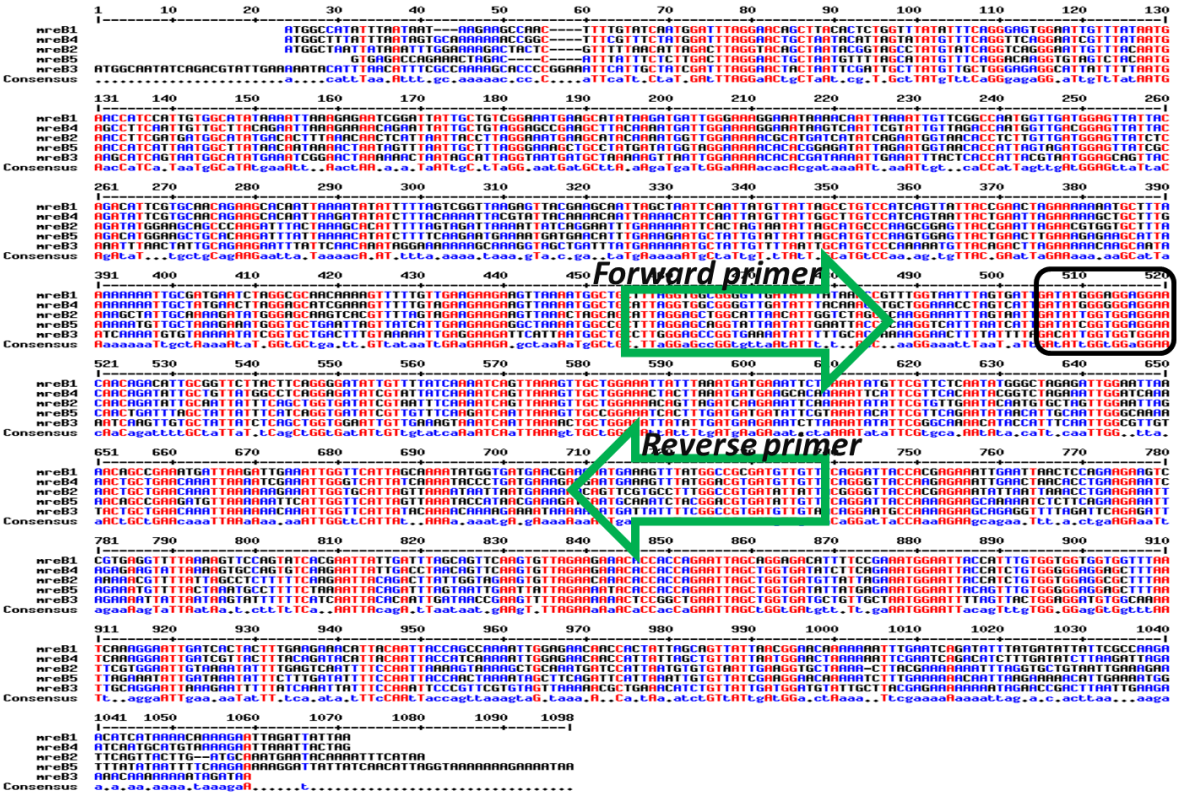

Figure 5.4 Selection of target sequences of *mreB*s for achieving homologous recombination at the desired gene. Alignment of *Spiroplasma citri* GII-3 *mreB* sequences for selection of region with minimum identity. The sites of primer design are indicated by green arrows. The nucleotides coding for phosphate 2 motif are shown within the black box.

5.2.3) Selection of region of interest for targeting *fibril*

Unlike MreB, Fibril does not have a homolog in the available databases (PDB and GenBank), except for sequences of spiroplasmas. Also, in absence of structural or biochemical information on Fibril, it is difficult to select its region based on functional importance for cloning into disruption vectors (pSD32, pGOT-res1) for generating *fibril* disruption mutant.

A query in NCBI (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins) using Fibril sequence resulted into identification of 5'-methylthioadenosine nucleosidase (MTAN; PDB ID 1JYS; Figure 3.1 B) as a closest homolog with approximately 30 % identity. Upon alignment of Fibril and MTAN sequences, the latter was found to align with the 236 residues at N-terminus of Fibril (Figure 3.1 A).

Thus, the 236 amino acids at the N-term of Fibril may constitute one domain whereas, rest of the protein was thought to form second domain (Cohen-Krausz et al., 2011). Taking clues from this information, we selected the target region of 373 bp, which were part of the N-terminal domain, in *fibril* for cloning into disruption vector (Figure 5.5).

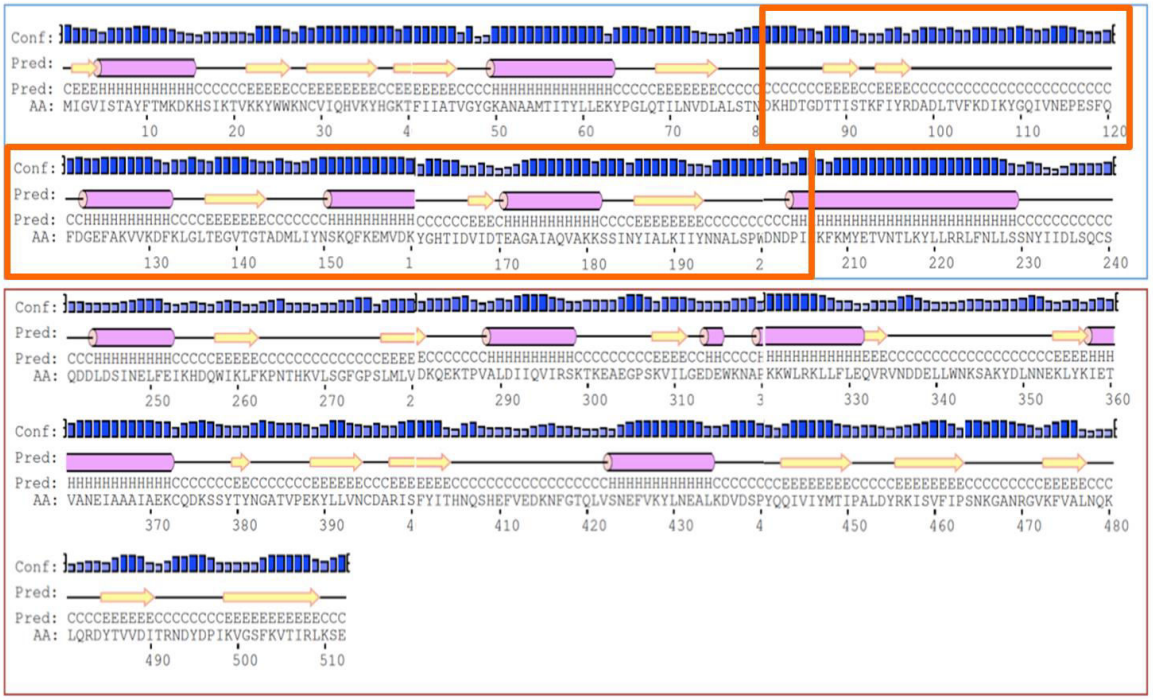

Figure 5.5 Selection of target sequence in Fibril for achieving homologous recombination. Secondary structure of Fibril predicted using Psipred (http://bioinf.cs.ucl.ac.uk/psipred/). The target sequence for disruption has been highlighted in orange boxes.

5.2.4) Primer design

Primers were designed to amplify target region/region of interest (RoI) of *mreB* and *fibril* genes for cloning into pSD32 as well as pGOT-res1 vectors within a BamHI site situated between the two *res* sites (Table 5.2). All the primers were designed by taking 19-22 bases from the vector, followed by a restriction site and 19-29 bases from the gene. This design of the primers enabled cloning attempts by restriction digestion-ligation method as well as by In-Fusion /restriction-free cloning methods.

Table 5.2 List of primers used for amplification of RoI and cloning into target vectors*.

Target gene	Primer name	Sequences (5'→3')
mreB1	**1IF-F**	CATTTTTATGAGATCCTCTAGAGGATCCGGGGTTGATATTTATAAACCGTTTGG
	1IF-R	CGGTTGCTCTAGAACTAGTGGATCTCGGCCATAAACTTTCATTTTTCG
mreB2	**2IF-F**	CATTTTTATGAGATCCTCTAGAGGATCCGGCATTAACATTGGTCTAGCTCAAGG
	2IF-R	CGGTTGCTCTAGAACTAGTGGATCTCGGCCAAAGGCACGAACTG
mreB3	**3IF-F**	CATTTTTATGAGATCCTCTAGAGGATCCGGTGAAAATATTTTTGCACCAAAAG
	3IF-R	CGGTTGCTCTAGAACTAGTGGATCTCGGCCGAAAATAATCATTTTTTTATTTTC
mreB4	**4IF-F**	CATTTTTATGAGATCCTCTAGAGGATCCGGGGTTGATATTTACAAACC
	4IF-R	CGGTTGCTCTAGAACTAGTGGATCTCGTCCATAAACTTTCATTCTTC
mreB5	**5IF-F**	CATTTTTATGAGATCCTCTAGAGGATCCGGTATTAATATTGAATTACCTCAAGGTC
	5IF-R	CGGTTGCTCTAGAACTAGTGGATCTCGTCCGTAGATTTGCATTGATC
Fibril	**Fib-F**	CATTTTTATGAGATCCTCTAGAGGATCCCGATAAGCATGATACAGG
	Fib-R	CGGTTGCTCTAGAACTAGTGGATCCTGGATCATTATCCCATGG

*For cloning by restriction digestion-ligation method, the sequences highlighted in pink were mutated to GGATCC (BamHI restriction site).

Regions highlighted in turquoise are from the target vector pSD32. The same sequence is also present in pGOT-res1 vector as well. Region highlighted in yellow is the BamHI restriction site. Sequences without background were selected from the regions of interest of target gene to be disrupted.

In order to identify if the disruption vector got integrated at the target site, primers were designed from the nucleotide sequences flanking genes of interest (Table 5.3). The end primers of *fibril* gene can be used to check recombination of disruption vector in the middle of *fibril* gene.

Table 5.3 List of primers designed for confirmation of integration of target disruption clone into genomic DNA.

Target gene	Primer name	Sequences (5'→3')
mreB1	***mreBF***	ATTCGTGCAACAGAAGCAC
	mreBR	TCACGGACTTCTTCTGGAG
mreB2	**114**	GGAAAAGACTACTCGTTTTTAAC
	R14	ACATTAATGGATCATTTGCAGC
mreB3	**F24**	ATTTAACATTTCGCCAAAAGCAC
	R24	TTCAGCGTTTTTAACTACACG
mreB4	**F34**	TACATGGCTTTATTTAATAGTGC
	R342	CATGCATTGATTCTAATCTTAAG
mreB5	**F44**	GAGGACAAAAACGTGAGACC
	344	GGTGTATTTTCTAATAATTCAATTAC
oriC	**EV13**	GGAAGATCGCACATTATTTTCCACG
	EV14	TTTTGATCCTTACTTTAGTATATTCTG

5.2.5) Cloning into pSD32 and pGOT-res1 vectors

RoI were amplified using freshly extracted genomic DNA of *S. citri* GII-3 as the template and forward and reverse primers specific for target gene. 4 µL each of the PCR products were checked on a 1.5% agarose gel by migrating samples along with a DNA marker. Replicate PCR reactions for a given gene were pooled and cleaned up using PCR clean-up kit (Promega) as per manufacturer's directions. All the amplicons were eluted in sterile MilliQ water and used as inserts for In-Fusion cloning reactions (section 5.2.7).

5.2.6) Preparation of a linearized vector by restriction digestion

500 ng each of pSD32 and pGOT-res1 were taken in separate, clean 1.5 mL microcentrifuge tubes. 5 µL buffer E (Promega) and 20 units of BamHI (Promega) were added to the tube. All the components were mixed gently by pipetting and spun briefly. The mixture was incubated for 100 minutes on a thermomixer with the temperature set to 37 ºC. Restriction enzyme BamHI was heat-inactivated by incubating the reactions at 65 ºC for 15 minutes on a thermomixer. 2 µL of the reaction was loaded onto a 1.5% agarose gel, DNA was migrated and visualized.

5.2.7) In-Fusion cloning

For each In-Fusion cloning reaction, 9 ng of linearized vector and 35 ng of insert were added to a fresh 1.5 mL microfuge tube on ice containing 1 µL of 5X In-Fusion HD Enzyme premix (Takara Bio USA, Inc.). Sterile milliQ was added to make up the volume to 5 µL. A negative control reaction was set up without addition of the insert to aforementioned components.

All the reactions were briefly spun and incubated for 15 minutes in a water bath pre-set to 50 ºC to facilitate cloning. The reactions were transferred to ice at the end of 50 ºC incubation to cool the products for 5 minutes. 40 µL of STELLAR chemical competent cells were added to each reaction aseptically and incubated on ice. The In-Fusion cloning product was transformed into STELLAR cells by heat shock method as per manufacturer's protocol (https://www.takarabio.com/assets/documents/User%20Manual/PT5055-2.pdf). At the end of recovery period, cells were plated onto Luria-Bertani (LB) agar plates supplemented with tetracycline at a final concentration of 10 µg/mL and incubated overnight (12-14 hours) at 37 ºC.

5.2.8) Clone check by restriction digestion method

Colonies selected for clone check were each inoculated aseptically into a sterile 20 mL tube containing 4 mL sterile LB broth supplemented with tetracycline (final concentration = 10 µg/mL). All the inoculated tubes were incubated at 37 ºC for 12-14 hours under shaking condition (200 rpm). Upon growth of cultures, cells were pelleted in separate tubes by spinning at 20,000 xg/25 ºC/10 minutes and supernatant discarded. Plasmid DNA was extracted from pellets using Wizard® Plus SV minipreps DNA purification systems kit and as per manufacturer's directions (https://www.promega.com/-/media/files/resources/protcards/wizard-plus-sv-minipreps-dna-purification-system-quick-protocol.pdf?la=en). Plasmid DNA was eluted using 50 µL sterile MilliQ water.

Restriction digestion reactions were set using 3 µL of extracted plasmid DNA, 1 µL buffer A (10X), 1 µL KpnI enzyme (Promega) and volume made up to 10 µL using sterile MilliQ water. A control reaction was set up using empty vector instead of recombinant plasmid. All the reactions were mixed gently by pipetting, spun briefly and incubated at 37 ºC for 1 hour on a thermomixer. Upon incubation, all the reactions were spun briefly to pull down reactants stuck to the walls of tube. 2 µL of gel loading dye was added to each tube, samples mixed and entire

volume loaded into 1.5 % agarose gel. The DNA samples were migrated on gel and visualized using a gel documentation system.

5.2.9) Clone confirmation by sequencing

1 µg each of the positive clones as evident from restriction digestion reactions was used for confirmation by sequencing using T7 promoter primer (GENEWIZ) before proceeding for further experiments.

5.2.10) Transformation of disruption vectors into *Spiroplasma* cells

Spiroplasma citri GII-3 cells were transformed with *mreB* and *fibril* RoI containing clones in pSD32 vector by the protocol as mentioned in section 4.2.8 of Chapter 4. A transformation reaction each was set up using pSD32 empty vector and sterile MilliQ water as positive and negative controls respectively for electroporation. The obtained transformants were observed and cultured as follows.

5.2.11) Growing *Spiroplasma citri* transformants and confirmation of mutants

Colonies were marked on the back of plate by visualization under microscope and picked using sterile glass capillary micropipettes in sterile environment. Each colony was put into sterile, screw-capped 1.5 mL tube under sterile conditions. 500 µL of sterile SP4 media supplemented with tetracycline (2 µg/mL) was added to each tube containing a colony. All the tubes were capped and incubated at 32 ºC in an incubator until growth was observed. The growth of transformants could be indicated by change in color of the media from red to yellow as compared to an un-inoculated control and confirmed by observation of culture using microscope.

5.3) Results

Next few sections describe the results obtained towards generating gene disruption mutants of *Spiroplasma citri* GII-3.

5.3.1) *fibril* and *mreB* RoI are cloned in pSD32 vector

The regions of interest (RoI) from *fibril* and all the 5 *mreB* genes of *S. citri* GII-3 were amplified (Figure 5.6). High sequence identity among five *mreBs* of *Spiroplasma* makes it difficult to amplify the desired *mreB* gene from genome. However, our strategy described in sections 5.2.1 - 5.2.3 helped us to amplify portions from desired genes and clone them into disruption vectors. The RoI were cloned into linearized pSD32 vector (Figure 5.6) by In-Fusion cloning method.

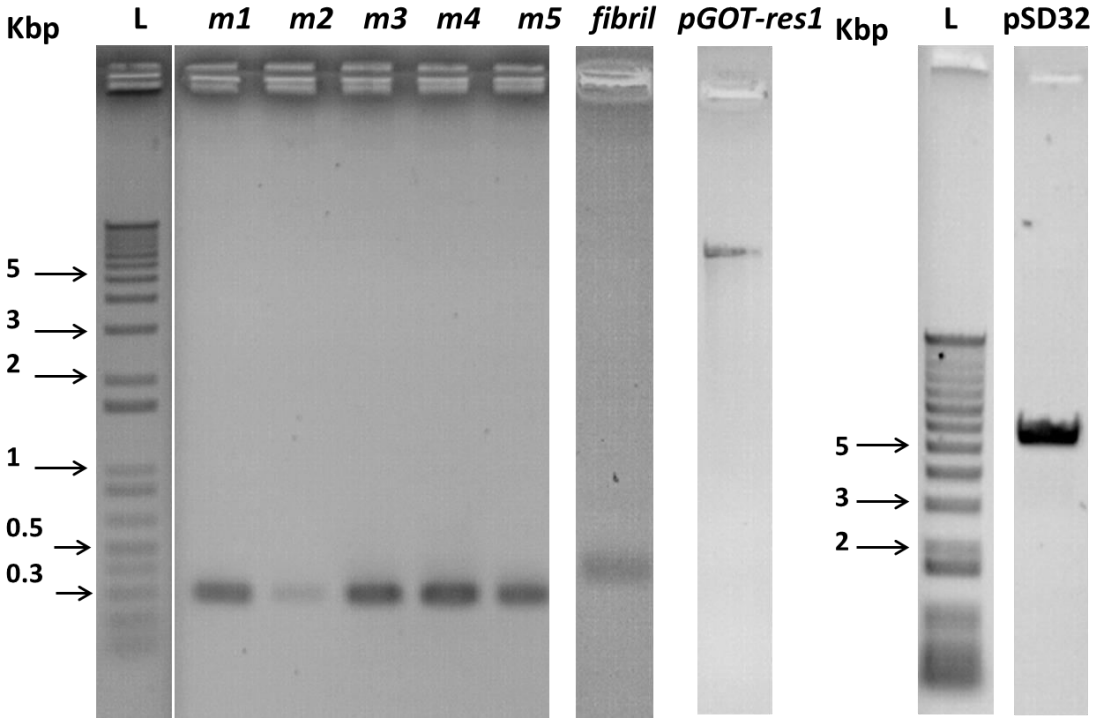

Figure 5.6 Preparation of *mreB* and *fibril* disruption clones. Amplification of regions of interest (RoI) and restriction digestion of vectors. Amplified regions of interest from *fibril* and *mreB* are visualized on 1.5% agarose gel. Expected sizes are: *m1-m5* (304 bp), *fibril* (414 bp), pGOT-res1 (8175) and pSD32 (5953 bp). BamHI-restricted vectors are shown in two lanes labeled accordingly. L: DNA marker, *m1-m5*: RoI from *mreB1* - *mreB5*.

The clones were initially checked by restriction digestion method (Figure 5.7) and bands corresponding to insert and vector were obtained as per the expected size.

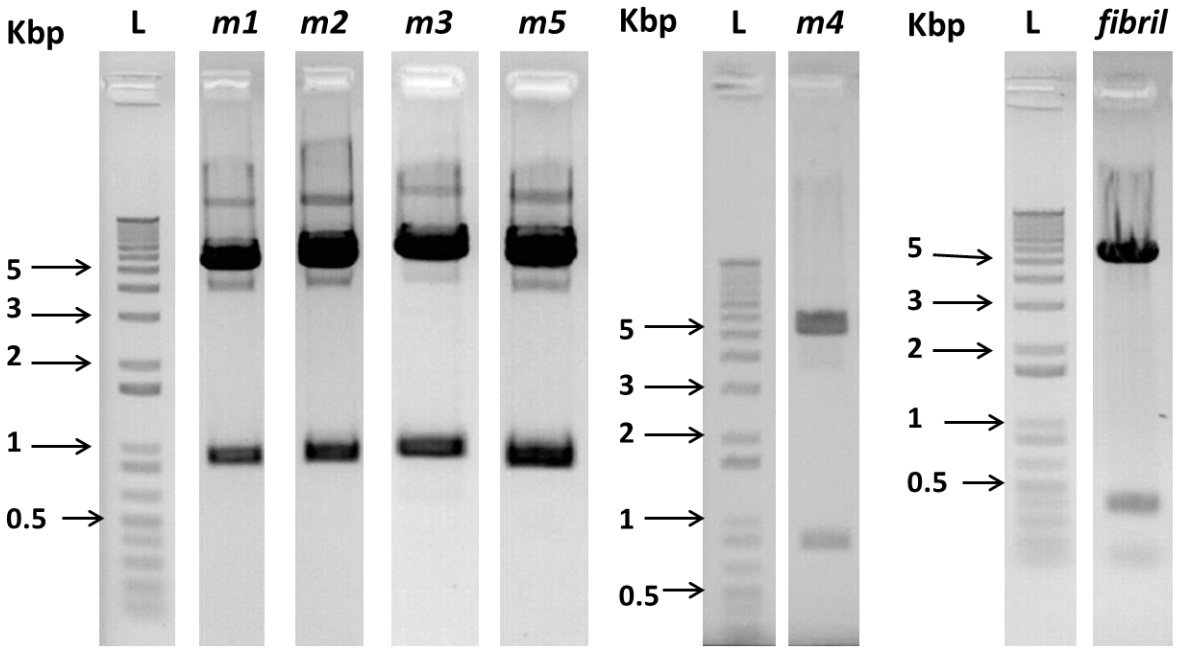

Figure 5.7 Clone check by restriction digestion. Clones obtained by In-Fusion cloning were checked by restriction digestion using KpnI restriction enzyme. For a *mreB* RoI positive clone, the expected sizes are 5.35 kbp and 869 bps. The sizes expected for a *fibril* RoI positive clone are 5.8 kbp, 580 bps. Due to the presence of a EcoRI site in *fibril* RoI, two bands are seen, each around 400 bps and 150 bps.

The high sequence identity between *Spiroplasma mreBs* makes it necessary to confirm the clones by sequencing to ensure that only the RoI from target *mreB* is present. Aligning the sequences of RoI in clones with respective gene sequences from *S. citri* GII-3 genomic DNA confirmed that indeed the clones have RoI from designated *mreB or fibril* (Figures 5.8 A to F). Sequence alignment showed point mutations in *mreB2* (682: A→C) and *fibril* (600: A→G; 606: T→A) genes. Since gene recombination employs long fragments of DNA, these point mutations are not expected to affect recombination experiments.

A

```
ScGII3 mreB1     436 GGGGTTGATATTTATAAACCGTTTGGTAATTTAGTGATTGATATGGGAGG 485
                     |||||||||||||||||||||||||||||||||||||||||||||||||||
mreB1 RoI          1 GGGGTTGATATTTATAAACCGTTTGGTAATTTAGTGATTGATATGGGAGG  50

ScGII3 mreB1     486 AGGAACAACAGACATTGCGGTTCTTACTTCAGGGGATATTGTTTTATCAA 535
                     |||||||||||||||||||||||||||||||||||||||||||||||||||
mreB1 RoI         51 AGGAACAACAGACATTGCGGTTCTTACTTCAGGGGATATTGTTTTATCAA 100

ScGII3 mreB1     536 AATCAGTTAAAGTTGCTGGAAATTATTTAAATGATGAAATTCTAAAATAT 585
                     |||||||||||||||||||||||||||||||||||||||||||||||||||
mreB1 RoI        101 AATCAGTTAAAGTTGCTGGAAATTATTTAAATGATGAAATTCTAAAATAT 150

ScGII3 mreB1     586 GTTCGTTCTCAATATGGGCTAGAGATTGGAATTAAAACAGCCGAAATGAT 635
                     |||||||||||||||||||||||||||||||||||||||||||||||||||
mreB1 RoI        151 GTTCGTTCTCAATATGGGCTAGAGATTGGAATTAAAACAGCCGAAATGAT 200

ScGII3 mreB1     636 TAAGATTGAAATTGGTTCATTAGCAAAATATGGTGATGAACGAAAAATGA 685
                     |||||||||||||||||||||||||||||||||||||||||||||||||||
mreB1 RoI        201 TAAGATTGAAATTGGTTCATTAGCAAAATATGGTGATGAACGAAAAATGA 250

ScGII3 mreB1     686 AAGTTTATGGCCG 698
                     |||||||||||||
mreB1 RoI        251 AAGTTTATGGCCG 263
```

B

```
ScGII3 mreB2     439 GGCATTAACATTGGTCTAGCTCAAGGAAATTTAGTAATTGATATTGGTGG 488
                     |||||||||||||||||||||||||||||||||||||||||||||||||||
mreB2 RoI          1 GGCATTAACATTGGTCTAGCTCAAGGAAATTTAGTAATTGATATTGGTGG  50

ScGII3 mreB2     489 AGGAACAACAGATATTGCAATTATTTCAGCTGGTGATATCGTAATTTCAA 538
                     |||||||||||||||||||||||||||||||||||||||||||||||||||
mreB2 RoI         51 AGGAACAACAGATATTGCAATTATTTCAGCTGGTGATATCGTAATTTCAA 100

ScGII3 mreB2     539 AATCAGTTAAAGTTGCTGGAAAACAGTTAGATCAAGAAATTCAAAAATAT 588
                     |||||||||||||||||||||||||||||||||||||||||||||||||||
mreB2 RoI        101 AATCAGTTAAAGTTGCTGGAAAACAGTTAGATCAAGAAATTCAAAAATAT 150

ScGII3 mreB2     589 ATTCGTGTTGAATACAATGTGCTAGTTGGAATTAGAACTGCTGAACAAAT 638
                     |||||||||||||||||||||||||||||||||||||||||||||||||||
mreB2 RoI        151 ATTCGTGTTGAATACAATGTGCTAGTTGGAATTAGAACTGCTGAACAAAT 200

ScGII3 mreB2     639 TAAAAAAGAAATTGGTGCATTAGTTAAAATAATTAATGAAAAAACAGTTC 688
                     ||||||||||||||||||||||||||||||||||||||||||.||||||
mreB2 RoI        201 TAAAAAAGAAATTGGTGCATTAGTTAAAATAATTAATGAAAAACCAGTTC 250

ScGII3 mreB2     689 GTGCCTTTGGCCG 707
                     |||||||||||||
mreB2 RoI        251 GTGCCTTTGGCCG 263
```

Figure 5.8 Confirmation of clones prepared in pSD32 vector by sequencing. A-F) Alignment of regions of interest (RoI) cloned into pSD32 vector with the sequences of corresponding genes from genomic DNA (*Sc*GII-3). Point mutations are seen in *mreB2* (682 A→C) and *fibril* (600 A→G; 606 T→A) genes**.**

Panels C-F of this figure are given on next two pages.

C

```
ScGII3 mreB3     466 GGTGAAAATATTTTTGCACCAAAAGGAACTTTTATTTTAGACATTGGTGG 515
                     |||||||||||||||||||||||||||||||||||||||||||||||||||
mreB3 RoI          1 GGTGAAAATATTTTTGCACCAAAAGGAACTTTTATTTTAGACATTGGTGG  50

ScGII3 mreB3     516 TGGAAAATCAAGTTGTGCTATTATCTCAGCTGGTGGAATTGTTGAAAGTA 565
                     |||||||||||||||||||||||||||||||||||||||||||||||||||
mreB3 RoI         51 TGGAAAATCAAGTTGTGCTATTATCTCAGCTGGTGGAATTGTTGAAAGTA 100

ScGII3 mreB3     566 AATCAATTAAAACTGCTGGAAATTATATTGATGAAGAAATCTTAAAATAT 615
                     |||||||||||||||||||||||||||||||||||||||||||||||||||
mreB3 RoI        101 AATCAATTAAAACTGCTGGAAATTATATTGATGAAGAAATCTTAAAATAT 150

ScGII3 mreB3     616 ATTCGGGCAAAACATACCATTTCAATTGGCGTTGTTACTGCTGAACAAAT 665
                     |||||||||||||||||||||||||||||||||||||||||||||||||||
mreB3 RoI        151 ATTCGGGCAAAACATACCATTTCAATTGGCGTTGTTACTGCTGAACAAAT 200

ScGII3 mreB3     666 TAAAAAACAAATTGGTTCATTATACAAAACAAAAGAAAATAAAAAAATGA 715
                     |||||||||||||||||||||||||||||||||||||||||||||||||||
mreB3 RoI        201 TAAAAAACAAATTGGTTCATTATACAAAACAAAAGAAAATAAAAAAATGA 250

ScGII3 mreB3     716 TTATTTTCGGCCG 728
                     |||||||||||||
mreB3 RoI        251 TTATTTTCGGCCG 263
```

D

```
ScGII3 mreB4     439 GGGGTTGATATTTACAAACCTGCTGGAAACCTAGTCATTGATATGGGGGG 488
                     ||||||||||||||||||||||||||||||||||||||||||||||||||
mreB4 RoI          1 GGGGTTGATATTTACAAACCTGCTGGAAACCTAGTCATTGATATGGGGGG  50

ScGII3 mreB4     489 AGGAACAACAGATATTGCTGTTATGGCCTCAGGAGATATCGTATTATCAA 538
                     ||||||||||||||||||||||||||||||||||||||||||||||||||
mreB4 RoI         51 AGGAACAACAGATATTGCTGTTATGGCCTCAGGAGATATCGTATTATCAA 100

ScGII3 mreB4     539 AATCAGTTAAAGTTGCTGGAAACTACTTAAATGATGAAGCACAAAAATTC 588
                     ||||||||||||||||||||||||||||||||||||||||||||||||||
mreB4 RoI        101 AATCAGTTAAAGTTGCTGGAAACTACTTAAATGATGAAGCACAAAAATTC 150

ScGII3 mreB4     589 ATTCGTTCACAATACGGTCTAGAAATTGGATCAAAAACTGCTGAACAAAT 638
                     ||||||||||||||||||||||||||||||||||||||||||||||||||
mreB4 RoI        151 ATTCGTTCACAATACGGTCTAGAAATTGGATCAAAAACTGCTGAACAAAT 200

ScGII3 mreB4     639 TAAAATCGAAATTGGGTCATTATCAAAATACCCTGATGAAAGAAGAATGA 688
                     ||||||||||||||||||||||||||||||||||||||||||||||||||
mreB4 RoI        201 TAAAATCGAAATTGGGTCATTATCAAAATACCCTGATGAAAGAAGAATGA 250

ScGII3 mreB4     689 AAGTTTATGGACG 701
                     |||||||||||||
mreB4 RoI        251 AAGTTTATGGACG 263
```

Figure 5.8 Panels C and D.

E

```
ScGII3 mreB5     427 GGTATTAATATTGAATTACCTCAAGGTCATTTAATCATTGATATCGGTGG 476
                     ||||||||||||||||||||||||||||||||||||||||||||||||||
mreB5 RoI          1 GGTATTAATATTGAATTACCTCAAGGTCATTTAATCATTGATATCGGTGG  50

ScGII3 mreB5     477 AGGAACAACTGATTTAGCTATTATTTCATCAGGTGATATCGTTGTTTCAA 526
                     ||||||||||||||||||||||||||||||||||||||||||||||||||
mreB5 RoI         51 AGGAACAACTGATTTAGCTATTATTTCATCAGGTGATATCGTTGTTTCAA 100

ScGII3 mreB5     527 GATCAATTAAAGTTGCCGGAAATCACTTTGATGATGATATTCGTAAATAC 576
                     ||||||||||||||||||||||||||||||||||||||||||||||||||
mreB5 RoI        101 GATCAATTAAAGTTGCCGGAAATCACTTTGATGATGATATTCGTAAATAC 150

ScGII3 mreB5     577 ATTCGTTCAGAATATAACATTGCAATTGGGCAAAAAACAGCCGAAGATGT 626
                     ||||||||||||||||||||||||||||||||||||||||||||||||||
mreB5 RoI        151 ATTCGTTCAGAATATAACATTGCAATTGGGCAAAAAACAGCCGAAGATGT 200

ScGII3 mreB5     627 TAAAAAATTCATTGGTTCATTAGTTAAATACCATAACGAAAGATCAATGC 676
                     ||||||||||||||||||||||||||||||||||||||||||||||||||
mreB5 RoI        201 TAAAAAATTCATTGGTTCATTAGTTAAATACCATAACGAAAGATCAATGC 250

ScGII3 mreB5     677 AAATCTACGGACG 689
                     |||||||||||||
mreB5 RoI        251 AAATCTACGGACG 263
```

F

```
ScGII3 fibril    240 CGATAAGCATGATACAGGAGATACAACTATCTCAACAAAATTTATTTATC 289
                     |||||||||||||||||||||||||||||||||||||||||||||||||||
fibril RoI         1 CGATAAGCATGATACAGGAGATACAACTATCTCAACAAAATTTATTTATC  50

ScGII3 fibril    290 GTGATGCTGACTTAACAGTTTTTAAAGATATTAAATATGGACAAATTGTT 339
                     |||||||||||||||||||||||||||||||||||||||||||||||||||
fibril RoI        51 GTGATGCTGACTTAACAGTTTTTAAAGATATTAAATATGGACAAATTGTT 100

ScGII3 fibril    340 AATGAACCTGAATCATTCCAATTCGATGGTGAATTCGCTAAAGTTGTTAA 389
                     |||||||||||||||||||||||||||||||||||||||||||||||||||
fibril RoI       101 AATGAACCTGAATCATTCCAATTCGATGGTGAATTCGCTAAAGTTGTTAA 150

ScGII3 fibril    390 AGACTTTAAATTGGGGTTAACAGAAGGAGTTACTGGAACAGCAGATATGT 439
                     |||||||||||||||||||||||||||||||||||||||||||||||||||
fibril RoI       151 AGACTTTAAATTGGGGTTAACAGAAGGAGTTACTGGAACAGCAGATATGT 200

ScGII3 fibril    440 TGATTTATAACTCAAAACAATTTAAGGAAATGGTTGATAAATATGGGCAT 489
                     |||||||||||||||||||||||||||||||||||||||||||||||||||
fibril RoI       201 TGATTTATAACTCAAAACAATTTAAGGAAATGGTTGATAAATATGGGCAT 250

ScGII3 fibril    490 ACTATTGATGTAATTGATACTGAAGCAGGGGCAATTGCGCAAGTTGCTAA 539
                     |||||||||||||||||||||||||||||||||||||||||||||||||||
fibril RoI       251 ACTATTGATGTAATTGATACTGAAGCAGGGGCAATTGCGCAAGTTGCTAA 300

ScGII3 fibril    540 AAAATCAAGTATTAATTATATTGCTTTAAAAATTATTTATAATAATGCAT 589
                     |||||||||||||||||||||||||||||||||||||||||||||||||||
fibril RoI       301 AAAATCAAGTATTAATTATATTGCTTTAAAAATTATTTATAATAATGCAT 350

ScGII3 fibril    590 TATCACCATGAGATAATGATCCA 612
                     |||||||||.|||||.||||||
fibril RoI       351 TATCACCATGGGATAAAGATCCA 373
```

Figure 5.8 Panels E and F.

Multiple attempts of cloning *mreB* and *fibril* RoI in pGOT-res1 vector by In-Fusion cloning as well as restriction digestion-ligation method failed to give positive clones (Figures 5.9).

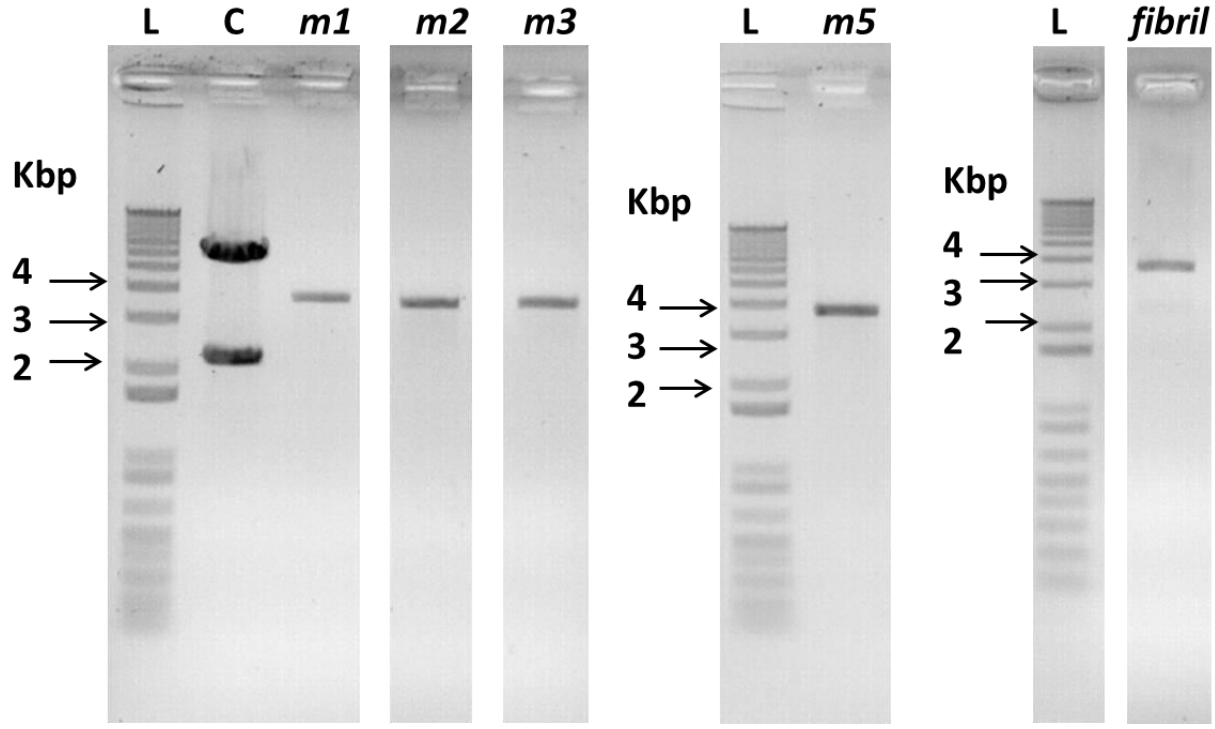

Figure 5.9 Checking of clones prepared in pGOT-res1 vector. Potential clones of *mreB* and *fibril* RoI in pGOT-res1 vector were checked by restriction digestion using enzymes BamHI and EcoRI. The expected sized for positive *mreB* clones are: 6 kbp and 2.4 kbp while for positive *fibril* clone it is 6 kbp and 2.5 kbp. An empty pGOT-res1 vector when restricted with BamHI and EcoRI gives bands corresponding to 6 kbp and 2.1 kbp. From the gel it is seen that the empty vector produces expected bands whereas, the clones do not produce expected bands. Thus, the clones are negative. L: DNA marker, C: empty pGOT-res1 vector.

Thus, *Spiroplasma* cells were transformed with RoI clones in pSD32 vector while simultaneously attempting cloning in pGOT-res1.

5.3.2) *mreB* RoI transformation mutants of *S. citri* GII-3 are obtained

After 50 days of incubating transformation reactions on SP4 agar containing tetracycline, *Spiroplasma* transformant colonies were obtained. Absence of colonies on negative control plates (without addition of any DNA) suggests that the colonies on test plate resulted from presence of clones prepared in pSD32 plasmid. About 49 colonies/100 ng clone per

transformation reaction were obtained for cells recovered upon electroporation. The colonies were very small with a diameter of about 20-30 µm (Figure 5.10).

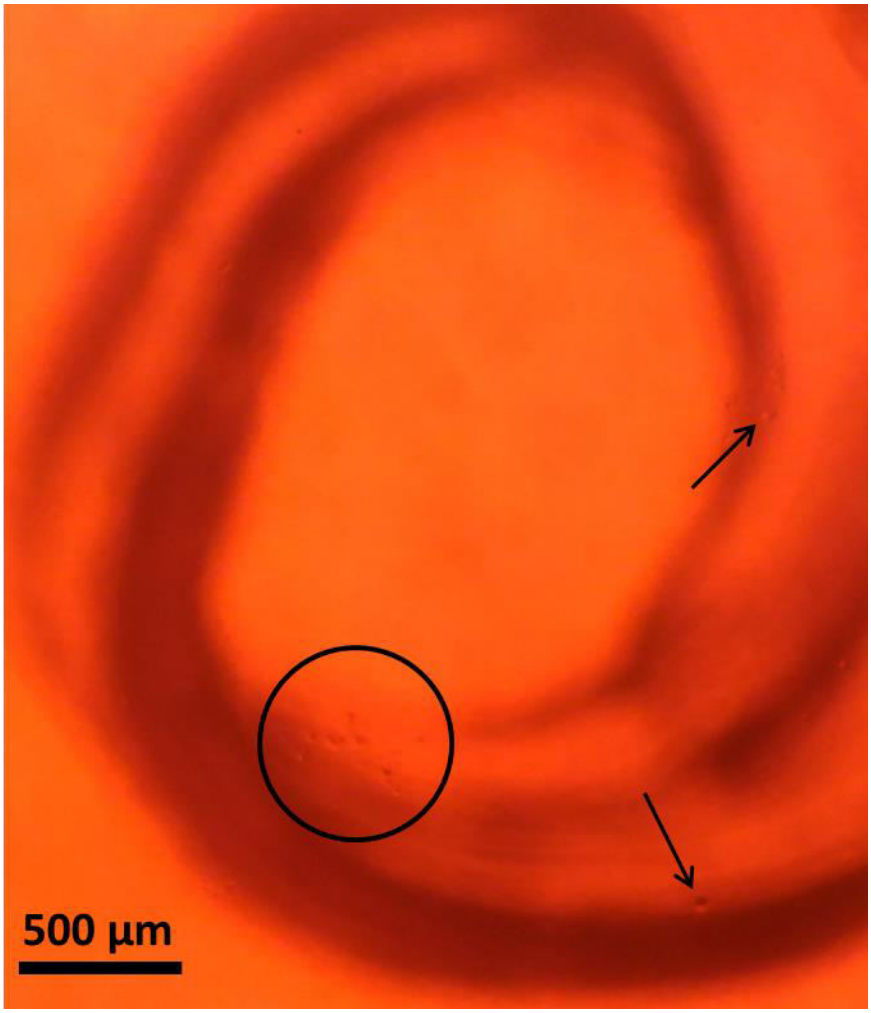

Figure 5.10 Visualization of *Spiroplasma* transformant colonies. *Spiroplasma* colonies on modified SP4 media supplemented with tetracycline (2 µg/mL) as seen under a dark-field microscope. A cluster of colonies are encircled while individual colonies are shown by arrows.

Inoculation of the transformant colonies into liquid modified SP4 media supplemented with tetracycline and incubation resulted into growth of one of the *mreB4* RoI in pSD32 clones. The cell growth was indicated by change in color of media in the inoculated tube as compared to an un-inoculated tube. The growth of this transformant was observed after about 3 months of incubation of the culture. This growth rate is slower than exhibited by wildtype *Spiroplasma* cells; doubling time of about 7 hours is observed for wildtype *Spiroplasma citri* growth in a different experiment. Thus, the growth of culture must be a result of possession of *mreB4* RoI in pSD32 clones and is hampered by the presence of tetracycline in growth media.

Since other transformants did not exhibit growth in two months, fresh transformation was performed again and transformant colonies have been inoculated into modified SP4 liquid media supplemented with tetracycline.

5.4) Discussion

Spiroplasma are cell wall-less bacteria with minimal genome essential for survival. The conservation of multiple *mreB* genes in such an organism indicates that these genes may have an essential function for survival. While some *Spiroplasma* species have seven *mreBs*, other species have only five. The reason for this variation in number of *mreB* genes is unknown. The role of MreB proteins in *Spiroplasma* and in general in cell-wall less bacteria remains unknown. As discussed in Chapter 4 of this thesis, studies on *S. citri* ASP-I, a naturally occurring non-helical and non-motile mutant of *Spiroplasma,* have revealed that MreB5 is involved in *Spiroplasma* morphogenesis and motility. This encouraged investigation of the role of other MreBs in *Spiroplasma*. Also, proposed roles of Fibril need verification by genetic studies. Thus, genetic experiments were performed on *S. citri* GII-3 to obtain gene disruption mutants of *fibril* or one of the *mreBs*.

The absence of functional RecA in *Spiroplasma* limits the recombination-based genetic modification for gene disruption to a chance event (Lartigue et al., 2002; Marais et al., 1996). Thus, it necessitates that the plasmid with portion of target gene to be disrupted, is carried by cells in high copy numbers over several generations to bring about recombination. The most efficient vectors available for specific gene disruption in *Spiroplasma* utilize a 163 bp *oriC* region, and a 220 bp target gene region in addition to antibiotic resistance marker(s) (Duret et al., 2005; Lartigue et al., 2002). However, these methods of recombination in *Spiroplasma* are not 100% efficient (Duret et al., 2005; Lartigue et al., 2002), making genetic modifications in *Spiroplasma* challenging.

In the present experiments two vectors (pSD32 and pGOTres1), which are modified from pSD6 and pGOT1 respectively (described in Chapter 1 of this thesis), were used for disruption of *mreB*s and *fibril*. Since *Spiroplasma mreB* genes have high sequence identity, the possibility of recombination at *mreB* genes other than the target *mreB* cannot be avoided. In order to increase chances of recombination at the target *mreB* genes, region of minimum sequence identity were

chosen (Figure 5.2 B). *S. citri* GII-3 transformants were obtained with rate of about 800 transformants per µg of plasmid. This transformation rate is comparable to that obtained by Duret et al., (2005). One of the colonies (out of 100 inoculated) grew in liquid modified SP4 media supplemented with tetracycline in about two month incubation. However, the other colonies did not grow in liquid modified SP4 media containing antibiotic. This raises the possibility that tetracycline hindered growth of *Spiroplasma*, a slow growing organism. Fresh set of transformants have been prepared and inoculated in liquid media for further growth and passaging. Since Fibril and MreB proteins have been proposed in essential cellular processes, it is possible that disruption of some of these genes is lethal to the cell. In such a case, the gene disruption must be accompanied by vector-based inducible expression of those proteins. This approach will be useful to investigate roles of those proteins at varying concentrations and to find if the gene is essential.

Once the gene disruption mutants of *Spiroplasma* are obtained, it will be exciting to compare their colony morphologies to the wildtype and ASP-I cells. This will give preliminary information about role of the gene in motility. Further, microscopic visualizations can be performed to observe if cell helicity is affected and quantitate swimming speeds. Simultaneously, electron cryotomography studies of the mutants will help understand if the absence of these proteins affects organization of cytoskeletal filaments.

As a part of experiments to rescue these mutants, the functional gene can be tagged with a fluorescent protein (FP) at an appropriate site. The site of FP insertion in MreB proteins can be decided based upon crystal structures of MreBs available in protein data bank (PDB). Choosing the site for tagging Fibril will be tricky since no structural information is available on it. Once the FP-tagged protein is expressed in *Spiroplasma* cells, the localization of that protein can be studied using Cryo correlative light electron microscopy (Cryo-CLEM) of the cells. This set of experiments may be useful to verify the hypotheses proposed by Kürner et al., (2005) and Trachtenberg et al., (2008) regarding organization of Fibril and MreB filaments. These investigations will provide first reports demonstrating roles of multiple MreBs in cell wall-less bacteria and also verify different functions proposed for MreB and Fibril since early 1970s.

5.5) References

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.

Cohen-Krausz, S., Cabahug, P.C., and Trachtenberg, S. (2011). The monomeric, tetrameric, and fibrillar organization of Fib: the dynamic building block of the bacterial linear motor of *Spiroplasma melliferum* BC3. J. Mol. Biol. *410*, 194–213.

Duret, S., Andre, A., and Renaudin, J. (2005). Specific gene targeting in *Spiroplasma citri* : improved vectors and production of unmarked mutations using site-specific recombination. Microbiology *151*, 2793–2803.

Eggleston, A.K., and West, S.C. (1996). Exchanging partners: recombination in *E. coli*. Trends Genet. *12*, 20–26.

Van den Ent, F., Amos, L.A., and Löwe, J. (2001). Prokaryotic origin of the actin cytoskeleton. Nature *413*, 39–44.

Jiang, W., Bikard, D., Cox, D., Zhang, F., and Marraffini, L.A. (2013). RNA-guided editing of bacterial genomes using CRISPR-Cas systems. Nat. Biotechnol. *31*, 233–239.

Kürner, J., Frangakis, A.S., and Baumeister, W. (2005). Cryo – electron tomography reveals the cytoskeletal structure of *Spiroplasma melliferum*. Science *307*, 436–439.

Lartigue, C., Duret, S., Garnier, M., and Renaudin, J. (2002). New plasmid vectors for specific gene targeting in *Spiroplasma citri*. Plasmid *48*, 149–159.

Marais, A., Bove, J.M., and Renaudin, J. (1996). Characterization of the *recA* gene regions of *Spiroplasma citri* and *Spiroplasma melliferum*. J. Bacteriol. *178*, 7003–7009.

Setlow, R.B., Swenson, P.A., and Carrier, W.L. (1963). Thymine dimers and inhibition of DNA synthesis by ultraviolet irradiation of cells. Science *142*, 1464–1466.

Slonimski, P.P., Perrodin, G., and Croft, J.H. (1968). Ethidium bromide induced mutation of Yeast mitochondria: complete transformation of cells into respiratory deficient non-chromosomal "Petites." Biochem. Biophys. Res. Commun. *30*, 232–239.

Trachtenberg, S., Dorward, L.M., Speransky, V. V., Jaffe, H., Andrews, S.B., and Leapman, R.D. (2008). Structure of the cytoskeleton of *Spiroplasma melliferum* BC3 and its interactions with the cell membrane. J. Mol. Biol. *378*, 778–789.

Wagstaff, J., and Löwe, J. (2018). Prokaryotic cytoskeletons: protein filaments organizing small cells. Nat. Rev. Microbiol. *16*, 187–201.

Zimmermann, F.K. (1977). Genetic effects of nitrous acid. Mutat. Res. *39*, 127–147.

Chapter 6

Summary and Future prospects

The work summarized in this thesis describes, for the first time, heterologous expression and purification of Fibril and its short constructs in *E. coli*. The work also includes identification of MreB5 to be responsible for rod-to-helical transition of *Spiroplasma* cells. Genetic studies to identify roles of each of the five MreBs and Fibril protein in *Spiroplasma* have been initiated. Further experiments on all these studies need to be carried out to better understand role of cytoskeletal proteins in *Spiroplasma* physiology. Discussed below are some of the experiments that can be performed in future.

Chapter 2 describes multiple protocols standardized for purification of Fibril from *E. coli*. The purified Fibril filaments have been obtained and observed using electron microscopy. Since these Fibril filaments are present as bundles, their standardization is necessary to obtain separated filaments for structure determination. Apart from structural characterization of Fibril, their biochemical characterization can also be performed. The first study could be verification of the proposal (this thesis) that Fibril is a membrane binding protein. This can be performed by thin layer chromatography (TLC) experiments of Fibril purified without any detergents compared with the protein purified further by methanol-chloroform extraction method. Standard lipids present in *E. coli* cell membranes can be used as control experiments to identify lipids bound to Fibril, if any. If Fibril is found to be associated with membrane lipids of *E. coli*, similar experiments of Fibril purified from *Spiroplasma* cells would confirm if Fibril is a membrane-binding protein. These experiments may prove useful in obtaining samples suitable for cryoEM studies by removal of lipids to obtain individual Fibril filaments.

Fibril has been proposed to bind DNA (Townsend and Archer, 1983) but this remains to be verified. Thus, DNA binding studies can be performed using purified Fibril to verify the hypothesis. Since *Spiroplasma* genomes have a high AT content, AT-rich DNA should be used for these experiments.

The next Chapter describes attempts to understand domain organization of Fibril and to obtain non-polymerizing constructs of Fibril. The study involved design of multiple short constructs

and their characterization. Some of these short constructs have been prepared and insights obtained from them. However, the work is stalled due to the difficulties associated with obtaining protein in soluble form at lysis stage. Thus, it is necessary to revisit these constructs and perform their characterization. The constructs Fibril Δ(1-236)-His_6 and His_6-Fibril Δ(305-512) may be prioritized since the crystallization has been attempted for the former and filaments have been seen for the latter. This work may be performed simultaneously with the characterization of EGFP-tagged constructs prepared for identification of polymerization interface of Fibril. Upon identification of polymerization interface of Fibril, a non-polymerizing construct can be obtained for X-ray crystallography studies. Mutation of residues involved only in lateral or axial interactions can provide insights into Fibril polymerization process and bundling.

Another approach to understand Fibril polymerization is to attempt complete dissociation of purified Fibril filaments by treating them with DMMA (2,3-dimethylmaleic anhydride) (Townsend and Plaskitt, 1985). The dissociated sub-units could be used for X-ray crystallography and Fibril association/polymerization studies. However, lack of functional biochemical assays of Fibril may mislead the investigation in this direction.

The fourth Chapter led to identification of lack of functional (full-length) MreB5 in *S. citri* ASP-I cells, leading to their loss of helical shape and motility. Thus, MreB5 is important in rod-to-helical transition of *Spiroplasma* cells. The proteomic analysis suggests that some lipoproteins are expressed at very low levels in *S. citri* ASP-I cells as compared to *S. citri* GII-3 cells. Based on this, it is appears that MreB5 may have a role in *Spiroplasma* shape determination through its interaction with lipoproteins. Thus, the characterization of interaction between MreB5 and lipoproteins may provide insights into shape determination mechanism of *Spiroplasma*.

The vector-based expression of full-length MreB5 in *S. citri* ASP-I led to instating motility in these cells, suggesting that MreB5 has a role in *Spiroplasma* motility as well. Additionally, Fibril was proposed to be a central player in *Spiroplasma* motility (Trachtenberg et al., 2003). Thus, it is possible that Fibril and MreB5 interact with each other. Verification of interactions of MreB5 with Fibril as well as lipoproteins can help us identify if MreB5 is an adapter-like protein. These experiments would also be useful to identify additional candidates involved in *Spiroplasma* motility.

In Chapter 5, experiments to generate *mreB* and *fibril* gene disruption mutants have been described. Further, once the *mreB* or *fibril* disruption mutant *Spiroplasma* cells are obtained, these can be characterized initially by their colony morphology and motility rates, in case they are motile. Electron cryotomography experiments can be performed on these cells to check if the cytoskeletal organization is disrupted.

The experiments to rescue these (target gene disrupted) cells can be performed by vector-based expression of a fluorescently-tagged target protein. The fluorescently tagged protein would be useful to identify localization of the target protein by fluorescence microscopy as well as electron cryotomography experiments.

Based on our experiments, literature and information obtained from lab involved in *Spiroplasma* genetics experiments, the delay in obtaining gene disruption mutants is because of a truncated RecA protein. Thus, alternate strategy for accelerating recombination in *Spiroplasma* by expression of RecA protein might be promising approaches for future experiments. One of the challenges of such as approach is the codon optimization of genes to be expressed in *Spiroplasma. Spiroplasma* being a *Mollicute*, TGA codon codes for tryptophan instead of a stop.

According to the literature, while *S. citri* has a truncated *recA* gene (Marais et al., 1996a), *S. melliferum* has a complete *recA* gene with a STOP codon in the N-terminal region that produces a truncated protein (Marais et al., 1996b). These results suggested that, a functional RecA must have been present in the ancestors of these species at some point of time. Thus, I propose two strategies to overcome low recombination frequencies in spiroplasmas as follows.

1) **Preparation of *Spiroplasma* strain (*Spiroplasma*^{RecA}) with a functional RecA protein:**

To begin with, the sequence analysis of *recA* gene from *Spiroplasma* species with closely related bacteria such as members of genus *Clostridium* that have a functional RecA protein, can be performed to identify residues critical for their function. Next, the *recA* gene can be amplified from genome of any one of the *S. melliferum* strains BC3, BW, B88, G1 and AS576. A point mutation need to be performed on the amplified gene to convert the STOP codon to a suitable amino acid (based on sequence alignment results) before it can be incorporated into disruption vector. Once a functional *recA* gene is obtained, it needs to be preceded by a promoter sequence. The promoter can be selected from any one of the constitutively expressing genes in *Spiroplasma*

such as *fibril*, *elongation factor-Tu*, *spiralin* or the native promoter of *Spiroplasma recA* gene can be used.

An artificial vector can be prepared by incorporation of a functional *recA* gene with a promoter into any one of the existing disruption vectors. The multiple cloning site of such a vector must be flanked by *res* sequences on either side. This arrangement ensures expression of *recA* gene upon transformation of the disruption vector into *Spiroplasma* cells. The presence of *res* sites is useful for removal of antibiotic marker at a later stage.

A functional RecA protein can then help bring about homologous recombination at the truncated *recA* region in the genome. This helps obtain recombination-efficient *Spiroplasma* (*Spiroplasma*^{RecA}) cells. In the next step, the antibiotic cassette needs to be removed from genomic DNA of *Spiroplasma*^{RecA} cells by expression of Resolvase protein using a different vector to obtain clean mutants. The clean *Spiroplasma*^{RecA} mutants should be compared with wildtype cells for traits such as growth rate, colony formation, motility, cell shape, pathogenicity etc. If no major differences are found, such a mutant can be used for genetic studies.

At this stage the *Spiroplasma*^{RecA} cells can be used for disruption of the target gene by using method as described in Chapter 5 of this thesis. Briefly, portion of target gene needs to be cloned into a disruption vector and transformed into *Spiroplasma* (in this case, *Spiroplasma*^{RecA}) cells. The functional RecA can help rapidly obtain gene disruption mutants.

Following is an alternate strategy to this.

2) **Preparation of a vector containing functional *recA* gene as well as portion of the target gene to be disrupted:** A *Spiroplasma* mutant can be prepared by cloning a functional *recA* gene with promoter as well as portion of the target gene to be disrupted, into a disruption vector. The expression of functional RecA can facilitate integration of the vector at the target gene in genomic DNA, yielding gene disruption mutant.

Spiroplasma mutant with functional RecA (*Spiroplasma*^{RecA}) obtained by these strategies would be useful for gaining insights into *Spiroplasma* cell biology.

Thus, the aforementioned experiments would provide insights into *Spiroplasma* cell biology, a field poorly understood due to lack of suitable molecular biology tools.

References

Marais, A., Bove, J.M., and Ranaudin, J. (1996a). *Spiroplasma citri* virus SpV1-derived cloning vector: deletion formation by illegitimate and homologous recombination in a spiroplasmal host strain which probably lacks a functional *recA* gene. J. Bacteriol. *178*, 862–870.

Marais, A., Bove, J.M., and Renaudin, J. (1996b). Characterization of the *recA* gene regions of *Spiroplasma citri* and *Spiroplasma melliferum*. J. Bacteriol. *178*, 7003–7009.

Townsend, R., and Archer, D.B. (1983). A Fibril protein antigen specific to *Spiroplasma*. J. Gen. Microbiol. *129*, 199–206.

Townsend, R., and Plaskitt, K.A. (1985). Immunogold localization of p55-Fibril protein and p25-Spiralin in *Spiroplasma* Cells. Microbiology *131*, 983–992.

Trachtenberg, S., Gilad, R., and Geffen, N. (2003). The bacterial linear motor of *Spiroplasma melliferum* BC3: from single molecules to swimming cells. Mol. Microbiol. *47*, 671–697.

Appendix-I

A) List of lysis buffers and their components

1) **Lysis buffer 1**
   10 mM Tris (pH 7.6), 1 % sodium deoxycholate, 1 % Triton X-100 and 2 M glycerol
2) **Lysis buffer 2**
   10 mM Tris (pH 7.6), 1 % Triton-X 100 and 18.4 % glycerol
3) **Lysis buffer 3**
   50 mM Tris pH 7.6, 200 mM NaCl and 10 % glycerol
4) **Lysis buffer 4**
   10 mM Tris pH 7.6 and 0.5 % LSS
5) **Lysis buffer 5**
   10 mM Tris (pH 7.6) and 1 % (w/v) Sodium deoxycholate
6) **Lysis buffer 6**
   10 mM Tris pH 7.6 and 1 % Triton X-100
7) **Lysis buffer 7**
   10 mM Tris pH 7.6 and 0.5 % LSS
8) **Lysis buffer 8**
   10 mM Tris pH 7.6, 1 % Triton X-100 and 0.5 % LSS
9) **Lysis buffer 9**
   50 mM Tris pH 8.0, 200 mM NaCl and 10 % glycerol
10) **Lysis buffer 10**
   50 mM Tris pH 8.0, 200 mM NaCl and 0.5 % LSS

Appendix II

Reprint of publications

The FEBS Journal

FEBS PRESS science publishing by scientists

Mechanistic insights into enzymatic catalysis by trehalase from the insect gut endosymbiont *Enterobacter cloacae*

Anmol Adhav[1], Shrikant Harne[2], Amey Bhide[3], Ashok Giri[3], Pananghat Gayathri[2] and Rakesh Joshi[1]

1 Institute of Bioinformatics and Biotechnology, Savitribai Phule Pune University, India
2 Indian Institute of Science Education and Research, Pune, India
3 Division of Biochemical Sciences, CSIR National Chemical Laboratory, Pune, India

Keywords
Enterobacter cloacae; *Plutella xylostella*; structural analysis; trehalase; validoxylamine A

Correspondence
R. Joshi, Institute of Bioinformatics and Biotechnology, Savitribai Phule Pune University, Ganeshkhind, Pune 411007, Maharashtra, India
Tel: +91 20 25691333
E-mail: rakesh.joshi@unipune.ac.in
P. Gayathri, Indian Institute of Science Education and Research, Dr. Homi Bhabha Road, Pune 411008, Maharashtra, India
Tel: +91 20 25908128
E-mail: gayathri@iiserpune.ac.in
and
A. Giri, Division of Biochemical Sciences, CSIR National Chemical Laboratory, Dr. Homi Bhabha Road, Pune 411008, Maharashtra, India
Tel: +91 20 25902710
E-mail: ap.giri@ncl.res.in

Anmol Adhav and Shrikant Harne contributed equally to this article.

(Received 5 July 2018, revised 21 October 2018, accepted 15 January 2019)

doi:10.1111/febs.14760

Energy metabolism in the diamondback moth *Plutella xylostella* is facilitated by trehalase, an enzyme which assists in trehalose hydrolysis, from the predominant gut bacterium *Enterobacter cloacae*. We report the biochemical and structural characterization of recombinant trehalase from *E. cloacae* (Px_EclTre). Px_EclTre showed K_M of 1.47 (±0.05) mM, k_{cat} of 6254.72 min^{-1} and V_{max} 0.2 (±0.002) mM·min^{-1} at 55 °C and acidic pH. Crystal structures of Px_EclTre were determined in the ligand-free form and bound to the inhibitor Validoxylamine A. The crystal structure of the ligand-free form, unavailable until now for any other bacterial trehalases, enabled us to delineate the conformational changes accompanying ligand binding in trehalases. Multiple salt bridges were identified that potentially facilitated closure of a hood over the substrate-binding site. A cluster of five tryptophans lined the −1 substrate-binding subsite, interacted with crucial active site residues and contributed to both trehalase activity and stability. The importance of these residues in enzyme activity was further validated by mutagenesis studies. Many of these identified residues form part of signature motifs and other conserved sequences in trehalases. The structure analysis thus led to the assignment of the functional role to these conserved residues. This information can be further explored for the design of effective inhibitors against trehalases.

Introduction

Plutella xylostella (diamondback moth) infestation is one of the prominent biotic stresses on cruciferous crops. It exhibits remarkable adaptability to its environment and has developed resistance against most of the existing pest control strategies [1]. Recently, it was found that a few bacterial species from gut microbiota of *P. xylostella* are vital for their digestive and detoxification processes. *P. xylostella* gut hosts 148 bacterial

Abbreviations
Ec_Tre, *E. coli* trehalase; Px_EclTre, *Enterobacter cloacae* trehalase; Sc_Tre, yeast trehalase; Suc, sucrose; Unlig, unliganded; ValA, validoxylamine A.

species belonging to 10 different phyla. Among these, *Enterobacter cloacae* was found to be the most abundant species of larval gut microbes. These bacteria produce various carbohydrate-modifying and glycolytic enzymes like cellulases, trehalases and other glucosidases, which participate in food digestion and nutrition acquisition process of the host [2]. Thus, structural characterization of these enzymes of the symbiont provides scope to explore them as novel targets for future biological or chemical insect control methods.

Among different carbohydrate metabolic pathways, trehalose metabolism has a variety of physiological roles in most organisms. Trehalose serves as a major energy source, and also functions as stress protectant, in phytophagous insects [3]. Trehalase catalyses the breakdown of trehalose into two glucose molecules (Fig. 1A) and the high amount of energy released in the process is utilized for various physiological activities [4]. Trehalase inhibition in insects leads to abnormalities like hampered growth, unsuccessful stress recovery, weight loss, hypoglycaemia, reduced chitin synthesis, lethal metamorphosis and decreased flight capacity [5,6]. Hence, trehalase inhibitors, Validoxylamine A being one of them (Fig. 1A), will be a good pest control strategy against insects. With this goal, we have initiated biochemical and structural characterization of trehalase from *E. cloacae* (Px_EclTre).

Trehalases are classified as either acidic or neutral depending on their pH optima and also, as soluble or membrane bound, based on their cellular localization [5]. Crystal structures of bacterial and yeast trehalases cocrystallized with various ligands and activation partners are summarized in Table 1. Crystal structure of trehalase from *Escherichia coli* (PDB ID: 2WYN) showed an $(\alpha/\alpha)_6$ barrel fold similar to other α-toroidal glycosidases. The complexes of *E. coli* trehalase with Validoxylamine A (Ec_Tre-ValA; 2JF4) and 1-thiatrehazolin (Ec_Tre-TTZ; 2JG0) revealed the interactions in the −1 and +1 subsites, where −1 subsite is 'catalytic site' and +1 subsite is 'recognition site' for the two glucose subunits in trehalose [7,8]. The corresponding unliganded structure of *E. coli* trehalase has not been determined. The requirement of an activation partner for catalysis differentiates a yeast trehalase from the bacterial counterpart. The yeast trehalase crystal structures include two unliganded forms and complexes with activation partners, however, an active conformation with the specific substrate and activation partner is unavailable. Hence, despite the presence of multiple crystal structures, features contributing to the molecular mechanism of trehalose catalysis are still unexplored.

In the present report, we have performed biochemical and structural characterization of trehalase (Px_EclTre) from the *P. xylostella* bacterial endosymbiont *E. cloacae*. We determined the crystal structures of Px_EclTre in the unliganded and Validoxylamine A-bound forms. A comprehensive structural analysis of the unliganded and liganded structures of Px_EclTre was performed by comparing them with the *E. coli* and *Saccharomyces cerevisiae* trehalase structures, to understand the salient features accompanying substrate recognition and catalysis. Characterization of mutants validated the role of conserved residues involved in catalysis and structural stability of Px_EclTre.

Results

Px_EclTre exhibited acidic α, α-trehalase activity with high optimum temperature

Px_Ecltre gene coding for periplasmic trehalase (Refseq Accession ID: WP_045355849.1) was recombinantly expressed in *E. coli* and the protein purified by affinity chromatography. The purified protein showed a single protein band at ~ 65 kDa on a 12% SDS/PAGE gel (Fig. 1B). The activity of purified Px_EclTre was measured by colorimetric estimation of product formed upon reaction of glucose with 3,5-dinitrosalicylic acid (DNSA) [9]. Px_EclTre exhibited a K_M of 1.47 (±0.05) mM, k_{cat} of 6254.72 min^{-1} and V_{max} 0.2 (±0.002) $mM{\cdot}min^{-1}$ (Fig. 1C). Px_EclTre displayed optimum activity at pH 5 (Fig. 1D), similar to the characteristic feature of acidic trehalases (pH 5–7) [6,10–13]. The optimum temperature for Px_EclTre activity was found to be 55 °C (Fig. 1E), which is higher as compared to previously reported trehalases from other organisms such as *Zunongwangia* sp. (50 °C), *Apis mellifera* L. (47 °C), *S. cerevisiae* (40 °C) and *Rhodotorula rubra* (30 °C) [13–16].

Validoxylamine A showed strong competitive inhibition of Px_EclTre activity

Validoxylamine A, a trehalose analogue, is an inhibitor of trehalases, with demonstrated inhibitory activity against bacterial and insect trehalases (Fig. 1A) [17]. In the present study, inhibition kinetics of Validoxylamine A against Px_EclTre at optimum conditions (pH 5, 55 °C) was measured to estimate its trehalase inhibitory activity. The concentration-dependent decrease in the activity of Px_EclTre was observed with increasing concentrations of Validoxylamine A (Fig. 1F). Enzyme kinetic analysis suggested competitive inhibition of Px_EclTre by Validoxylamine A

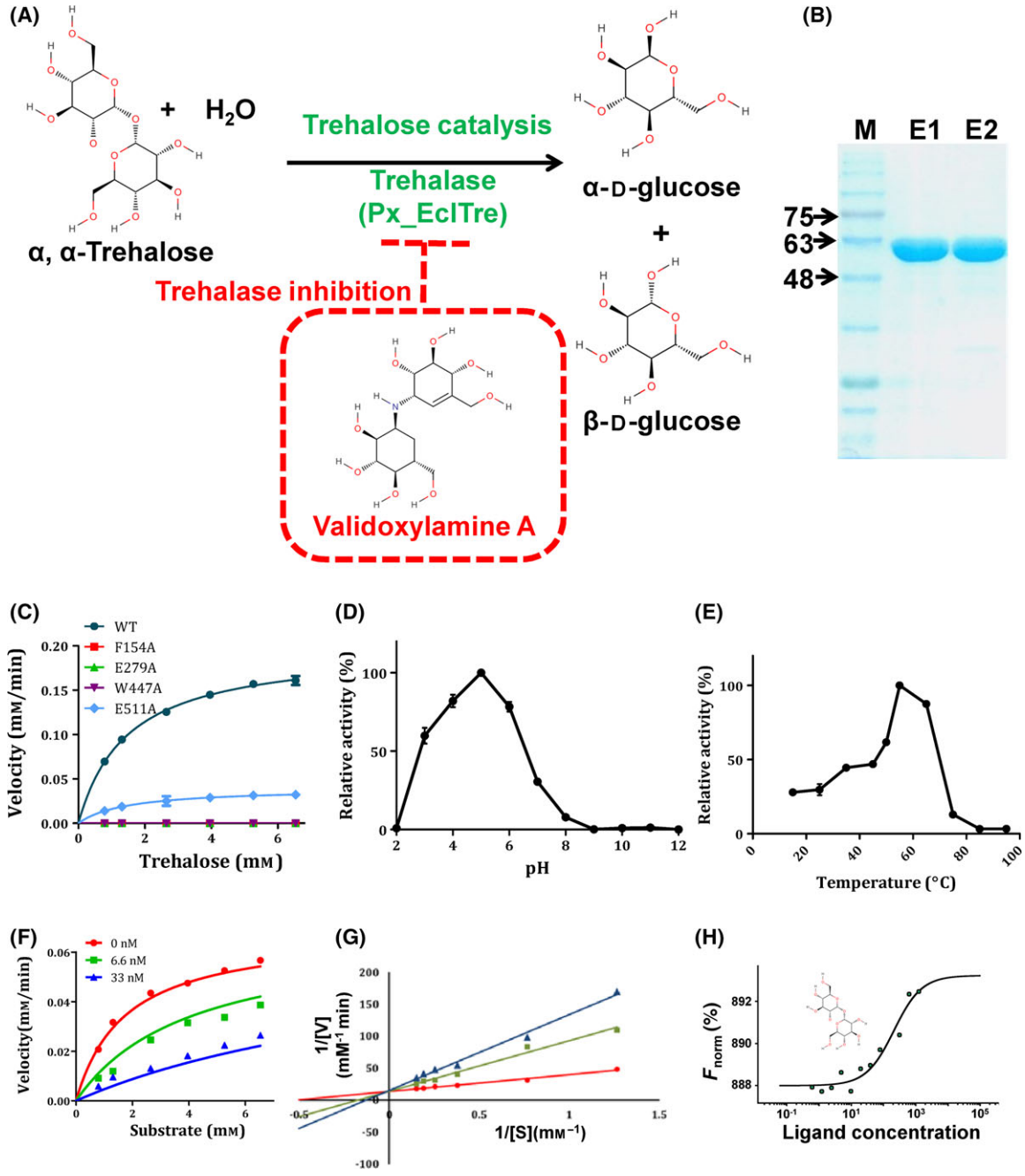

Fig. 1. Biochemical characterization of Px_EclTre. (A) Px_EclTre catalyses the hydrolysis of α, α-trehalose to break it down to α-D-glucose and β-D-glucose. This reaction is inhibited by the trehalose analogue, Validoxylamine A. (B) 12% SDS/PAGE gel showing purified Px_EclTre protein. (C) The rate of reaction as a function of substrate (trehalose) concentration for wild-type and mutants of Px_EclTre. (D) Effect of pH on trehalase activity by Px_EclTre. Px_EclTre shows optimum activity at pH 5. (E) Effect of temperature on trehalase activity by Px_EclTre. The optimum temperature for Px_EclTre activity was found to be 55 °C. (F) Enzyme inhibition of Px_EclTre activity by Validoxylamine A. Velocity vs trehalose concentration for different concentrations of Validoxylamine A has been plotted. (G) Lineweaver–Burk plot shows that Validoxylamine A functions as a competitive inhibitor. (H) MST analysis of Validoxylamine A binding to Px_EclTre. The chemical formula for Validoxylamine A is given in inset. In all the figures, error bars represent ±standard error for the observations of each treatment.

Table 1. Summary of bacterial and yeast trehalase structures.

PDB ID (abbreviation)	Bound ligand(s)/protein partner	Source of the enzyme	Occupancy of +1/−1 subsite	Position of		Angle between planes of W^{447} and W^{453} side chains	RMSD of $(\alpha/\alpha)_6$ domain[a]
				Side loop	Lid loop		
2JG0 (Ec_Tre-TTZ)	1-Thiatrehazolin (TTZ)	*Escherichia coli*	α-D-Glucopyranosilamine/ Tetrahydrocyclopentathiazole	Closed	Closed	5.12°	0.489 (378)
2WYN (Ec_Tre-LG9+Glu)	LG9 (Casuarine analogue) and α-D-Glucose (LG9+Glu)	*E. coli*	α-D-Glucose/LG9 (Casuarine analogue)	Closed	Closed	2.56°	0.389 (342)
2JF4 (Ec_Tre-ValA)	Validoxylamine A (ValA)	*E. coli*	C_7-Cyclitol/C_7-Cyclitol	Closed	Closed	6.27°	0.498 (422)
2JJB (Ec_Tre-Casu+Glu)	Casuarine and α-D-Glucose (Casu+Glu)	*E. coli*	α-D-Glucose/Casuarine	Closed	Closed	7.69°	0.530 (397)
5M4A (Sc_Tre-Trehalose)	Trehalose	*Saccharomyces cerevisiae*	α-D-Glucopyranose/α-D-Glucopyranose	Closed	Open	11.47°	1.317 (253)
5N6N (Sc_Tre-Suc+Bmh1)	Sucrose/Protein Bmh1 (Suc+Bmh1)	*S. cerevisiae*	β-D-Fructofuranose/α-D-Glucopyranose	Open	Intermediate	10.88°	1.259 (289)
5NIS (Sc_Tre-Unlig2)	None	*S. cerevisiae*	None/None	Open	Not modelled	8.88°	1.406 (265)
5JTA (Sc_Tre-Unlig1)	None	*S. cerevisiae*	None/None	Open	Not modelled	8.5°	1.372 (263)
5Z66 (Px_EclTre-ValA)	Validoxylamine A (ValA)	*Enterobacter cloacae*	C_7-Cyclitol/C_7-Cyclitol	Closed	Closed	7.25°	0.217 (374)
5Z6H (Px_EclTre-Unlig)	None	*E. cloacae*	None/None	Open	Open	3.62°	–

[a]Root-mean-square deviation of $(\alpha/\alpha)_6$ domain calculated by superposing the structures onto Px_EclTre-Unlig (5Z6H) structure. The number of superposed Cα atoms is indicated in bracket.

(Fig. 1G). Furthermore, IC_{50} and K_i of Validoxylamine A for Px_EclTre was found to be 7.5 (±0.2) and 1.9 (±0.075) nM respectively (Fig. 1F). A similar K_i value (1.9 nM) was previously reported for inhibition of trehalase from *Rhizoctonia solani* by Validoxylamine A [17,18]. Microscale thermophoresis (MST) analysis showed that Validoxylamine A bound strongly to Px_EclTre with a dissociation constant (K_D) of 206 (±10) nM (Fig. 1H). Furthermore, structural studies were carried out on Px_EclTre in unliganded form (Px_EclTre-Unlig) and cocrystallized with Validoxylamine A (Px_EclTre-ValA) to investigate its catalytic mechanism.

Crystal structures of Px_EclTre in unliganded form and Validoxylamine A-bound state

Crystal structures of Px_EclTre in unliganded (Px_EclTre-Unlig) and Validoxylamine A-bound (Px_EclTre-ValA) forms were determined to resolutions 2.3 and 1.8 Å (PDB IDs: 5Z6H and 5Z66) respectively. Data collection and refinement statistics are given in Table 2. The protein consists of $(\alpha/\alpha)_6$ toroidal fold, with a hood-like domain that closes over the ligand-binding pocket (Fig. 2A). Strong electron density for the ligand was observed, confirming that the inhibitor was indeed bound to trehalase (Fig. 2B–D). Interestingly, in the substrate-binding pocket of the unliganded form (Px_EclTre-Unlig; 5Z6H), a glycerol moiety from the crystallization condition was observed, whose atoms superposed very well with corresponding atoms of the bound ligand (Fig. 2E). In the absence of a ligand, a glycerol moiety and water molecules substituted for satisfying the hydrogen bonds formed by the active site and ligand-binding residues. Since glycerol is a non-specific ligand with very few contacts with the protein compared to the substrate (trehalose) and trehalase activity is not affected (specific activity of 1392.7 (±22.27) $U \cdot mg^{-1}$ with 10% glycerol compared to 1421.5 (±18.32) without glycerol), we consider glycerol as part of the solvent and the glycerol-bound form as the unliganded form of Px_EclTre (Px_EclTre-Unlig; 5Z6H). Structure determination of a ligand-free form of the enzyme (Px_EclTre-Unlig; 5Z6H) allowed us to compare the specific conformational changes accompanying ligand binding.

Structural comparison of trehalases in unliganded and ligand-bound states

The Px_EclTre shares about 79% sequence identity with *E. coli* trehalase over the entire sequence length and 29% with *S. cerevisiae* trehalase comparing the homologous region of the hood domain. While all four structures of trehalase from *E. coli* are in the ligand-bound forms, *S. cerevisiae* trehalase structures in the ligand-free and liganded forms are available. Since yeast trehalases are activated in the presence of an activator Bmh1, the complexed structure with Bmh1 was also included in the structural analysis [19]. The availability of these structures permitted us to carry out a comprehensive comparison of different conformations of trehalases, including ligand derivatives with different sugar linkages and activator bound structures. Table 1 contains the root-mean-square deviation (RMSD) of $(\alpha/\alpha)_6$ domain of all the unliganded and liganded conformations of trehalases with respect to $(\alpha/\alpha)_6$ domain of Px_EcTre-Unlig (5Z6H).

The extensive structural comparison of liganded/unliganded and active/inactive structures of trehalases led

Table 2. Data collection and refinement statistics.

	Trehalase – unliganded	Trehalase – validoxylamine A
PDB ID	5Z6H	5Z66
Data collection statistics		
Collected at	DLS I04	DLS I04
Wavelength (Å)	0.979	0.979
Space group	$P2_12_12_1$	$C121$
a, *b*, *c* (Å)	55.25, 127.91, 156.14	122.03, 51.45, 96.4
α, β, γ (°)	90, 90, 90	90, 106.95, 90
Resolution (Å)[a]	33.32–2.30 (2.38–2.3)	30.5–1.8 (1.84–1.8)
Number of unique reflections[a]	48 818 (4363)	53 185 (3172)
R_{merge} (%)[a]	12.1 (59.2)	9.9 (62.8)
R_{pim} (%)[a]	8.6 (42.1)	8.2 (52.1)
CC_{half}[a]	0.98 (0.685)	0.992 (0.634)
Mean $I/\sigma I$	9.5 (2.5)	9.4 (1.9)
Completeness (%)[a]	97.43 (96.0)	97.9 (100)
Redundancy[a]	4.2 (4.2)	3.0 (3.1)
Refinement statistics		
Resolution (Å)	33.32–2.3	30.49–1.8
Number of unique reflections (test set)	48 768 (2422)	53 176 (2698)
R_{work}/R_{free} (%)	16.84/22.96	16.28/19.36
Average B-factor (Å^2)	31	21
Wilson B-factor (Å^2)	27.1	15.6
RMS deviations		
Bond lengths (Å)	0.007	0.004
Bond angles (°)	0.926	0.729
Ramachandran map statistics		
Favoured (%)	96.0	96.47
Allowed (%)	3.6	3.14
Outliers (%)	0.4	0.39

[a]Values in parentheses denote last resolution shell.

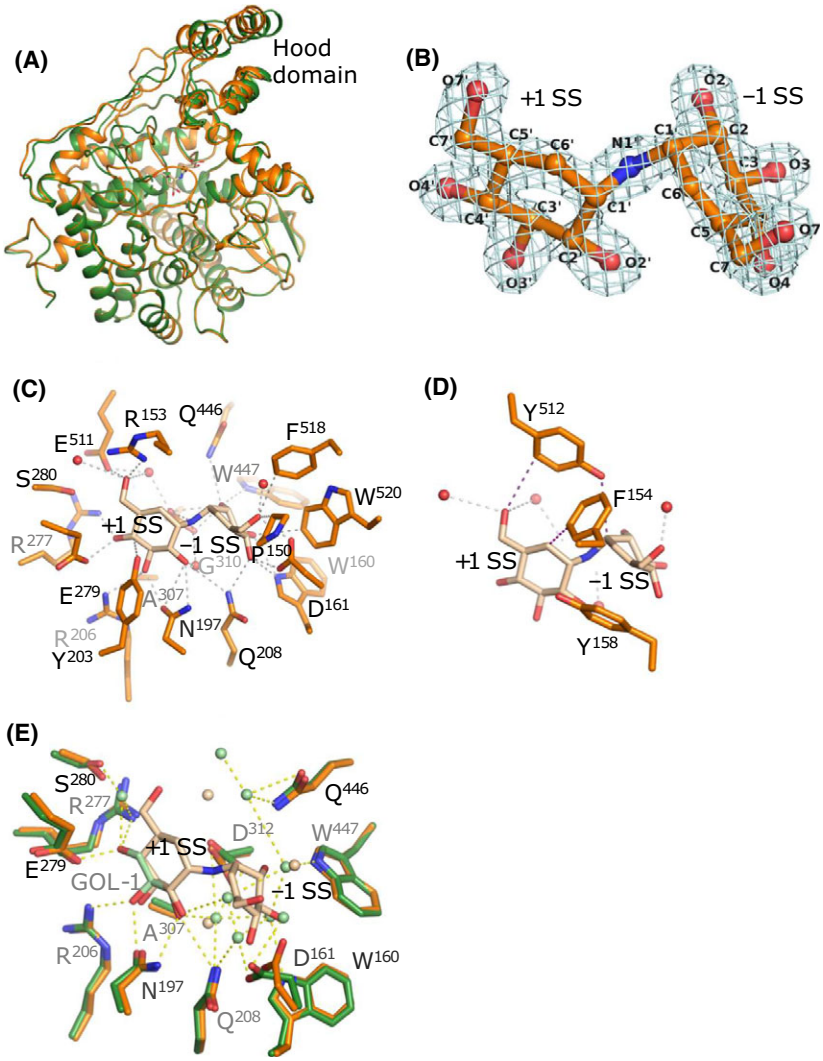

Fig. 2. Crystal structures of Px_Ecl trehalase in its unliganded form and bound to the inhibitor Validoxylamine A. (A) Superposition of unliganded (green) and ligand-bound (orange) forms of Px_EclTre. (B) The electron density (composite omit map of $F_0 - F_c$ at 2.5σ) of Validoxylamine A from the liganded form of Px_EclTre is shown with atoms labelled and −1 and +1 subsites (SS) marked. (C, D) Ligand binding site of Px_EclTre showing residues (orange) interacting with Validoxylamine A (wheat). For clarity, interactions of the residues shown above and below the ligand are shown in two separate figure panels (C, D), with the ligand maintained in the same orientation. (E) Ligand binding site of superposed Px_EclTre unliganded and ligand-bound structures. The glycerol moiety and water molecules present in unliganded structure contact the protein in a similar manner as Validoxylamine A in the liganded form. Colour scheme: Validoxylamine A (wheat), glycerol (light green), the unliganded form of Px_EclTre (dark green) and interactions between ligands and protein (dotted green/yellow lines). Subsite has been abbreviated as 'SS' in all the figure panels.

us to identify the following changes upon ligand binding – (a) subtle loop movements, (b) salt bridges specific to ligand binding, (c) interaction between conserved residues lining active site (+1 and −1 subsite), (d) predict roles in substrate recognition and hydrolysis for some of the residues conserved across trehalases. Most of these findings were not discussed in earlier reports on trehalases due to unavailability of a structure in catalytically active form with the specific substrate (for yeast) and unliganded form (for *E. coli*) in the respective organisms.

Conformational changes and ordering of the 'lid loop' upon ligand binding

Comparison of the ligand-free and inhibitor-bound structures illustrated that the hood domain Q^{36}-L^{100}, and loop comprising residues Y^{147}-Y^{159} and G^{506}-G^{519} (hereafter referred to as side loop and lid loop respectively) might undergo conformational changes upon ligand binding (Fig. 3A). We compared the lid loop, side loop and hood domain positions with other available structures of trehalases (Table 1). Liganded structure of Px_EclTre (Px_EclTre-ValA; 5Z66) superimposed on *E. coli* trehalase structures complexed with Validoxylamine A (Ec_Tre-ValA; 2JF4) and 1-thiatrehazolin (Ec_Tre-TTZ; 2JG0) showed RMSD of 0.334 (389 Cα atoms) and 0.311 Å (359 Cα atoms) respectively. The lid loop, side loop and the hood domain were oriented identically in these structures. Compared to the other inhibitor-bound trehalase structures of *E. coli*, one of the interesting observations was the difference in conformation of the lid loop in the structures liganded with derivatives of casuarine-6-*O*-α-D-glucoside (Ec_Tre-LG9+Glu;

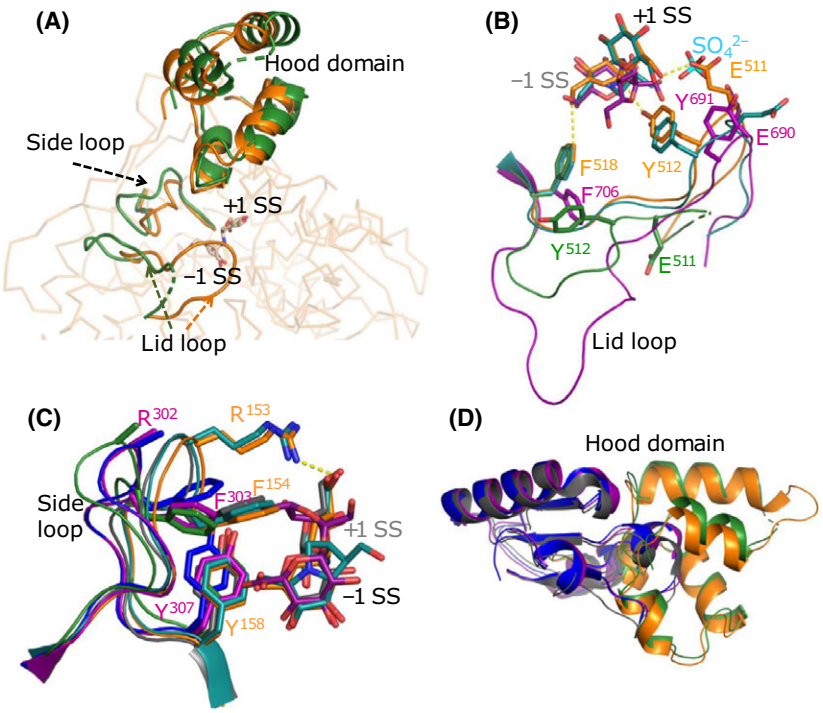

Fig. 3. Conformational changes accompanying ligand binding in trehalases (A) Superposition of the unliganded and ligand-bound forms of Px_EclTre highlight conformational changes upon ligand binding. Lid loops, side loops and the hood-like domains that undergo closure over the active site are shown as cartoon while rest of the structure is shown as the ribbon. (B) Lid loop is ordered and closes over the active site when bound to correct ligand. The residues making contacts with respective substrate have been shown as sticks. Sulphate bound in 2WYN is shown as ball and stick in yellow (sulphur) and red (oxygen). Sc_Tre-Suc+Bmh1 (5N6N; magenta), the only yeast structure with ordered lid loop, Ec_Tre-LG9+Glu (2WYN; cyan: bacterial trehalase structure in which the lid loop is in a different conformation due to bound sulphate ion), and the two Px_EclTre structures are shown as representative structures. Ec_Tre-LG9+Glu (2WYN) and Ec_Tre-Casu+Glu (2JJB) have the same conformation of the lid loop, while the lid loops of Ec_Tre-ValA (2JF4) and Ec_Tre-TTZ (2JG0) match exactly to Px_EclTre-ValA (5Z66; orange) and are not shown. (C) Occupancy of the +1 subsite is a determinant of the correct position of the side loop in the liganded state. The residues contacting ligand have been shown as sticks. The steric hindrance between F^{154} and the sugar occupancy in the +1 subsite due to the 1,4 linkage of sucrose prevents closure of the loop. 2WYN (Ec_Tre-LG9+Glu) represents the four trehalase structures from *Escherichia coli*. 5NIS (Sc_Tre-Unlig2), 5M4A (Sc_Tre-Trehalose) and 5N6N (Sc_Tre-Suc+Bmh1) represent conformations that are unliganded, bound to specific substrate (trehalose), and nonspecific substrate (sucrose), respectively. (D) Hood-like domains of bacterial and yeast trehalases belong to different folds. Unliganded (5NIS; Sc_Tre-Unlig2) and nonspecific ligand-bound (5N6N; Sc_Tre-Suc+Bmh1) structures from yeast represent the open conformation of the hood and the liganded state (5M4A; Sc_Tre-Trehalose) represents the closed conformation in yeast trehalase. 5Z6H (Px_EclTre-Unlig) and 5Z66 (Px_EclTre-ValA) represents the open and closed conformations in bacterial trehalase. [Colouring scheme for protein and respective ligands: Px_EclTre unliganded: green, Px_EclTre-Val: orange, Ec_Tre-LG9+Glu (2WYN): cyan, Sc_Tre-Suc+Bmh1 (5N6N): magenta, Sc_Tre-Trehalose (5M4A): grey, Sc_Tre-Unlig2 (5NIS): blue]. Subsite has been abbreviated as 'SS' in all the figure panels.

2WYN and Ec_Tre-Casu+Glu; 2JJB) [20,21]. The presence of a sulphate ion at the active site disrupted the interaction between the corresponding glutamate (E^{511}) and the substrate, thereby leading to incomplete lid loop closure in Ec_Tre-Casu+Glu (2JJB) and Ec_Tre-LG9+Glu (2WYN) (Fig. 3B). The lid loop was disordered in all the yeast trehalase structures except for Sc_Tre-Suc+Bmh1 (complexed with sucrose and Bmh1; 5N6N), where it was not closed completely (Fig. 3B). In Sc_Tre-Suc+Bmh1 (5N6N), the presence of nonspecific substrate (sucrose) probably kept the corresponding glutamate (E^{690}) distant from the ligand. The superposed structures showed that lid loop residues E^{511}, Y^{512} and F^{518}, which are conserved across trehalases, contacted the ligand, and binding of the specific ligand ensured complete loop closure. Thus, we hypothesized that residues from lid loop play a major role in recognizing the correct substrate (Fig. 3B). Mutation of the glutamate (E^{511}) to alanine in Px_EclTre resulted in 23% activity of the mutant enzyme compared to the wild-type (WT) Px_EclTre (specific activity of 397.4 $U \cdot mg^{-1}$ compared to 1734 $U \cdot mg^{-1}$ for wild-type enzyme; Fig. 1C), confirming the importance of E^{511} in substrate recognition and catalysis.

In addition to the lid loop, the side loop (Y^{147}-Y^{159}) also underwent a conformational change upon substrate binding (Fig. 3C). Out of the available yeast neutral trehalase structures (Table 1), the loop position was open in the unliganded structures (Sc_Tre-Unlig1; 5JTA and Sc_Tre-Unlig2; 5NIS). This loop superposed well with the corresponding loop position in the unliganded structure of Px_EclTre (5Z6H) (Fig. 3C). In the liganded form (5Z66), residues R^{153}, F^{154} and Y^{158} interacted with the substrate, all of which are conserved residues. Loop position in the liganded state appeared to be driven not only by the presence of a substrate in the pocket but also by the specificity of the correct ligand at the +1 subsite (Table 1). In the sucrose-bound yeast trehalase structure Sc_Tre-Suc+Bmh1 (5N6N), despite the substrate-binding site being occupied, the loop position matched with that of the unliganded conformations (Px_EclTre-Unlig; 5Z6H) (Fig. 3C). Since the glucoside linkage between the monosaccharide units is different in sucrose compared to trehalose, the +1 subsite occupancy of sucrose sterically prevented F^{303} (corresponding to F^{154} in Px_EclTre) from closing over the ligand. Hence, the loop closure is one of the features that contributes towards recognition of α, α (1→1) linkage sugar in the active site pocket of trehalases. In order to investigate the role of F^{154} in catalysis, we mutated it to alanine (F154A) in Px_EclTre. Although the mutant was well folded (Fig. 4A–D), it showed complete loss of trehalase activity (undetectable levels of activity for the mutant, Fig. 1C), emphasizing the significance of F^{154} in the enzyme function.

The hood domain closed over the ligand-binding pocket in the Validoxylamine A-bound Px_EclTre structure (5Z66). An RMSD of 1.125 Å (91 Cα atoms of the hood domain, upon superposition of the toroid domain) highlighted the domain movement compared to the unliganded structure (Fig. 3A,D). Although there were no direct interactions of hood domain residues with the ligand, interactions of the hood domain with residues of the side loop appeared to drive the conformational changes. Specifically, R^{153} from side loop formed a salt bridge with D^{59} from the hood domain. The position of the hood domain matched with those of all liganded structures of *E. coli* trehalase (Ec_Tre-TTZ; 2JG0, Ec_Tre-LG9+Glu; 2WYN, Ec_Tre-ValA; 2JF4 and Ec_Tre-Casu+Glu; 2JJB) (Table 1). Interactions that led to domain closure were dependent on occupancy of the specific ligand in the +1 subsite, and conformation of the side loop. These observations prompted us to look at the specific interactions of each residue that drive active site closure in trehalase structures.

A cluster of conserved salt bridges lines the substrate-binding pocket and mediates the active site closure

Salt bridges lining the substrate-binding pocket (Fig. 5A) constituted major determinants for closed conformation of the active site. A salt bridge between R^{277} and E^{511} is formed when the lid loop closes over the active site. Both R^{277} and E^{511} formed bridging hydrogen bonds with O7′ and O4′ of Validoxylamine A (+1 subsite), respectively, that led to substrate binding and loop closure (Fig. 5B,C). Furthermore, as discussed earlier, the decrease in activity of E511A mutant verified the crucial role of the residue in Px_EclTre glycosidic activity (Fig. 5C). The salt bridge also serves as a determinant of substrate specificity. In the *E. coli* trehalase structure bound to casuarine-6-*O*-α-D-glucoside (2WYN), the presence of a sulphate ion bound at the active site disrupted the salt bridge between E^{511} and R^{277} and resulted in an intermediate position of the lid loop.

The side loop closure over the substrate-binding site is accompanied by the formation of another salt bridge between R^{153} from the side loop and D^{59} from the hood domain. R^{153} formed a hydrogen bond with O7′ of Validoxylamine A, which coordinated the substrate binding with the closure of both side loop and hood domain. R^{206} interacted with Validoxylamine A (O3′ and O4′ atoms respectively) and E^{279}, thus forming a bridging set of interactions with the ligand (Fig. 5B, C). Mutation of E^{279} in Px_EclTre affected the activity of the enzyme (undetectable activity, Fig. 1C). Two other conserved residue pairs that link the toroid and hood domains are K^{61} and E^{156}, D^{65} and R^{200} respectively. These residues formed a cluster of salt bridges in both unliganded and ligand-bound states and did not interact with ligand (Fig. 5D). Thus, they may not contribute to the switch between conformational states accompanying ligand binding.

Structural superimposition of bacterial and yeast trehalases demonstrated that despite the difference in hood domain fold (Fig. 3D), the network of salt bridges was conserved at the corresponding positions in yeast trehalases also (Fig. 5C). This evolutionary conservation of salt bridge network between prokaryotic and eukaryotic trehalases, despite differences in the fold of the hood domain, indicates their significance in the catalytic activity of the enzyme. In yeast trehalase, R^{277} that is part of a salt bridge interaction with the conserved E^{511} is replaced with a proline. Interestingly, guanidium group of R^{473} in yeast occupied the same position as the side chain of R^{277} in Px_EclTre and was positioned optimally to conserve the salt bridge

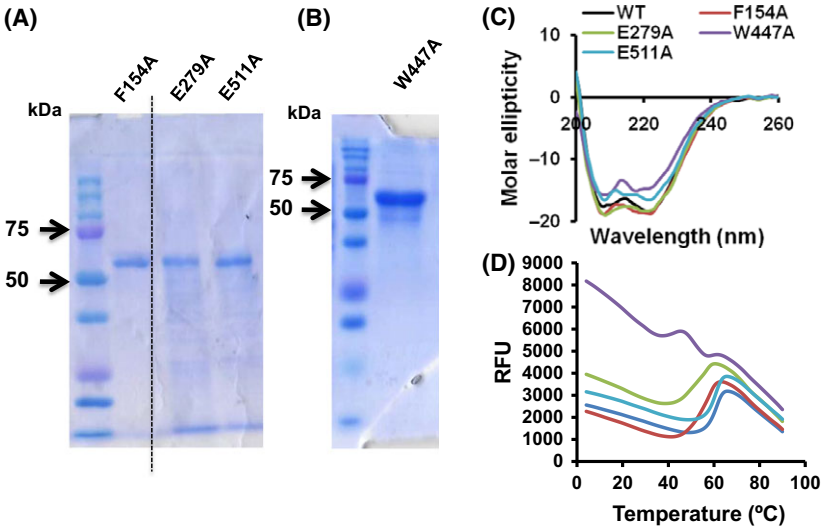

Fig. 4. Stability of Px_EclTre and its mutants (A) 12% SDS/PAGE gel showing purified mutants F154A, E279A and E511A and (B) W447A proteins. W447A mutant appears to behave differently compared to the wild-type while F154A, E279A and E511A appear to be folded. In (A) the dotted line indicates removal of an empty lane. (C) CD spectra and (D) Melting curves of the wild-type Px_EclTre (WT) and its different point mutants.

interaction as well as interaction with the ligand [e.g. yeast trehalase structure complexed with trehalose (Sc_Tre-Trehalose; 5M4A; Fig. 5C)]. The salt bridge was not formed in Sc_Tre-Suc+Bmh1 (5N6N: yeast trehalase structure complexed with sucrose and Bmh1) due to the presence of the nonspecific substrate, although the lid loop was ordered. R^{302} and E^{426} were poised to form the salt bridge pair in yeast trehalase for the side loop closure. The potential salt bridge interaction was not observed among yeast trehalase structures, probably due to them being either catalytically inactive or unliganded conformations. The salt bridge by the R^{206} and E^{279} pair and interaction with the ligand is conserved in the yeast trehalase structures (R^{355} and E^{424} in yeast). The interaction with the ligand (sucrose) was disrupted in Sc_Tre-Suc+Bmh1 (5N6N), and the salt bridge interaction also was not observed. D^{212}-$R^{349}$and K^{214}-E^{305} in yeast trehalases formed pairs of conserved salt bridges corresponding to D^{65}-R^{200}, and K^{61}-E^{156}, respectively, in bacterial trehalases. D^{212} and K^{214} from the hood domain of yeast were spatially conserved with K^{61} and D^{65} of bacterial trehalases, although the hood domains are of different folds.

All the salt bridges described above coordinated ligand binding and closure through recognition of the monosaccharide moiety at the +1 subsite. Further analysis was carried out to understand the effect of ligand binding at the −1 subsite.

Substrate-binding pocket at the −1 subsite is lined by tryptophan residues

Analysis of contacts to −1 subsite of the ligand highlighted the presence of conserved tryptophans in the ligand-binding pocket. The active site of trehalases is surrounded by five tryptophans, namely W^{160}, W^{311}, W^{447}, W^{453} and W^{520}, out of which all except W^{311} are conserved (Fig. 6A,B). Nε2 atom of W^{447} formed a hydrogen bond with O2 of Validoxylamine A, while W^{447} and W^{453} formed approximately parallel stacking interactions (Fig. 6B,C; the angle between the planes of side chains < 12°; Table 1) in close proximity to the substrate. W^{453} formed hydrophobic interactions with K^{497} (Fig. 6C). Only when the correct ligand or its analogue was bound, K^{497} oriented such that its terminal amino group formed crucial hydrogen bonds with Q^{446} and E^{496}. The conserved Y^{512} also interacted with Q^{446} and E^{496} side chains in catalytically active trehalase conformations (Fig. 6C). Since the density of K^{675} in yeast trehalase structures (corresponding to K^{497} in bacterial trehalase) was unavailable beyond Cβ, the presence of hydrogen bonds has been inferred only on basis of bacterial trehalase structures (PDB IDs: 2WYN, 2JG0, 2JJB, 2JF4, 5Z66, 5Z6H) and yeast trehalase structure 5N6N (Sc_Tre-Suc+Bmh1) where the side chain is structured for K^{675}. W^{453} was sandwiched between the lysine (K^{497}) and W^{447} that interacts with the ligand. These interactions might serve towards recognition of the cognate ligand for optimally orientating the active site cluster. Px_EclTre mutant W447A showed structural destabilization (Fig. 4C,D) and complete loss of activity (undetectable activity, Fig. 6C).

W^{311}, adjacent to the W^{447} side chain, formed a cation–π interaction with R^{316}, while R^{316}, in turn, formed salt bridging interactions with D^{336} and D^{389}. W^{311} is replaced by a histidine (H^{477}) in yeast trehalase, whose imidazolium ring coincides with the five-

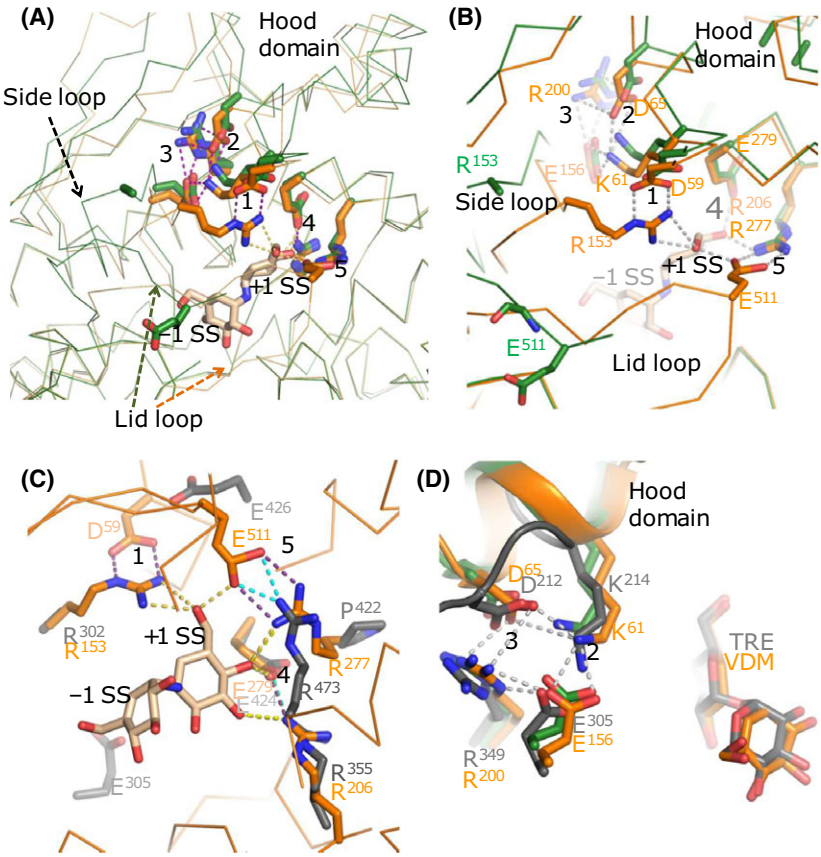

Fig. 5. Clusters of salt bridges stabilize ligand binding and accompanying domain closure. (A)The cluster of five salt bridges surrounding the ligand-binding pocket in Px_Ecl trehalase structures. (B) Zoomed view of salt bridges (labelled 1–5) in Px_EclTre structures. Salt bridges 1 (D^{59}-R^{153}) and 5 (R^{277}-E^{511}) are formed upon binding of correct substrate, leading to the closure of side loop, lid loop and hood domain. (C) Salt bridges 1, 4 and 5 interact with the ligand at the +1 subsite. In yeast trehalases (Sc_Tre-Trehalose; 5M4A is shown in grey), R^{302} and E^{426} probably form a salt bridge, equivalent to D^{59} and R^{153} in bacterial trehalases. Protein–ligand interactions are shown with yellow dotted lines and magenta dotted lines represent salt bridge interactions. The R355-E424 salt bridge is shown with a green dotted line while cyan dotted lines show that R473 is positioned to conserve salt bridge in Sc_Tre-Trehalose (5M4A). (D) Salt bridges 2 (K^{61}-E^{156}) and 3 (D^{65}-R^{200}) form a cluster in bacterial and yeast trehalase irrespective of the presence of a ligand, and do not interact with the ligand. Px_EclTre-ValA (5Z66) and Sc_Tre-Trehalose (5M4A) structures are shown as representatives of bacterial and yeast trehalase structures. The colour scheme is as mentioned in Fig. 3. Subsite has been abbreviated as 'SS' in the figure panels.

membered ring of tryptophan side chain (Fig. 6D). These interactions are conserved across all trehalases and might be involved in structure stabilization. Two other tryptophans in the vicinity of the active site and which interact with the ligand are W^{160} and W^{520}, whose side chains are perpendicularly oriented (Fig. 6B).

Although many of the residues discussed above did not interact with the ligand directly, they appear to form crucial interactions that modulate allosteric regulation of enzyme activity. A conservation analysis of the implicated residues among 185 sequences showed the level of conservation of each of these residues, highlighting the maintenance of the interactions across trehalases (Fig. 6E).

Discussion

Molecular mechanism of trehalose recognition and hydrolysis

Glycoside hydrolases catalyse the hydrolysis of sugar by two mechanisms – one involves retention of the anomeric configuration while the other occurs through an inversion step [22]. Trehalases are known to show the inversion of anomeric configuration of glucose, that is, it forms both α-glucose and β-glucose upon trehalose hydrolysis [23–25]. In inverting enzymes, an average distance of ~ 10 Å between the catalytic residues is observed [23]. In the case of liganded form and unliganded structures of Px_EclTre, the distances between

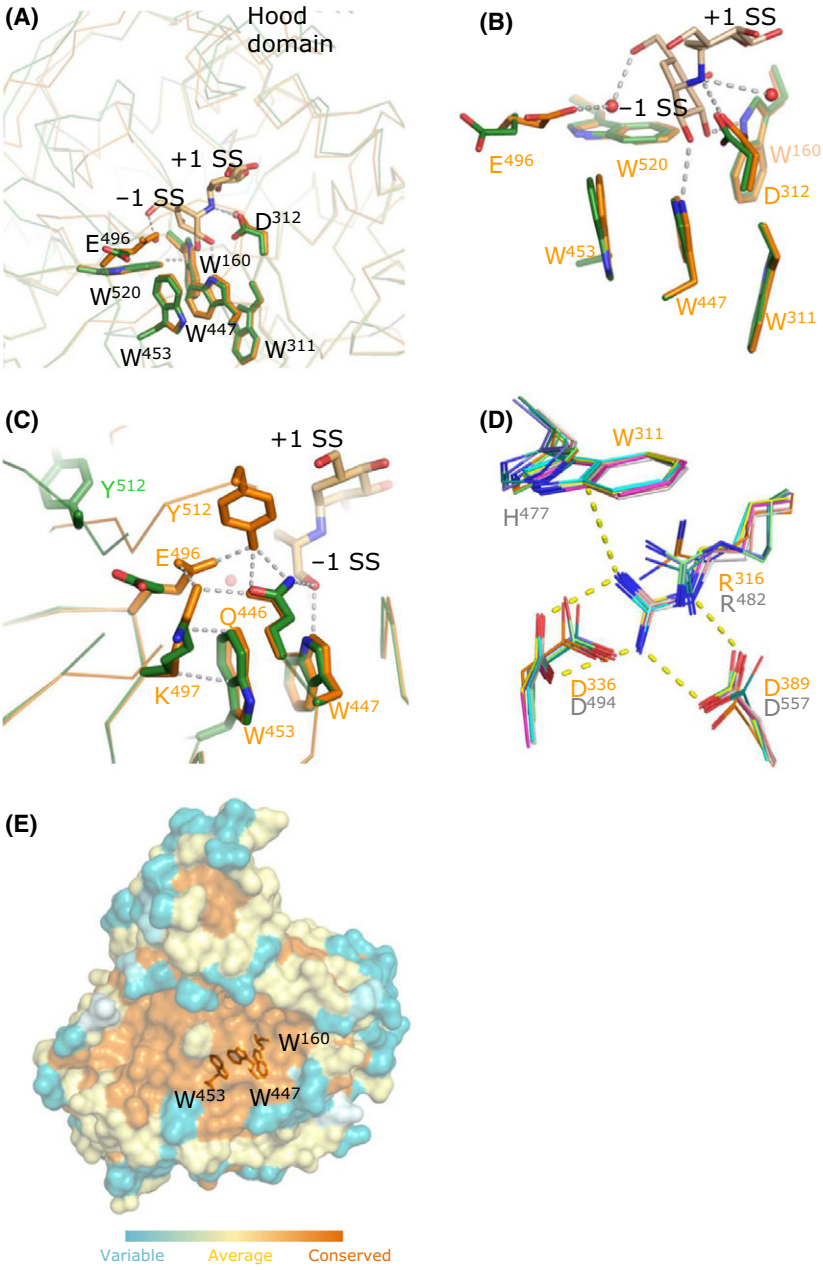

Fig. 6. Conserved tryptophans surround the −1 subsite of the ligand. (A) A cluster of five conserved tryptophans is present close to −1 subsite of the ligand-binding pocket. The catalytic residues D^{312} and E^{496} are also highlighted. (B) A zoomed view of the tryptophans close to −1 subsite of the ligand: W^{447} and W^{453} form an almost parallel stack while W^{160} and W^{520} are perpendicularly oriented. Also shown are the catalytic residues (D^{312} and E^{496}). (C) Residues K^{497} and Y^{512} form hydrogen bonds with Q^{446} and E^{496} in a catalytically active conformation. (D)R^{316} forms a cation–π interaction with W^{311} and salt bridges with D^{336} and D^{389}. In yeast trehalases, all the corresponding residues are conserved, except W^{311}that is replaced by H^{477}. Only Px_EclTre structures are shown in the subfigures A–C, while all the 10 structures are shown in 5D to highlight the clustering. (E) Surface representation of ligand-bound Px_EclTre structure, in which the colour gradient depicts the level of conservation at different parts of the protein, derived from Consurf analysis [50]. Subsite has been abbreviated as 'SS' in the figure panels.

the Cγ and Cδ atoms of the respective catalytic residues D^{312} and E^{496} were 9.2 and 11.6 Å respectively. This suggests that Px_EclTre possesses an inverting type of mechanism of catalysis. Interestingly, it was also observed that the catalytic D^{312} is in a cis-peptide conformation in both the liganded and unliganded structures.

The atoms from the substrate at the −1 subsite are surrounded by five tryptophans (W^{160}, W^{311}, W^{447}, W^{453} and W^{520}), four of which (except W^{311}) are conserved (Fig. 6A,B). Tryptophans have been observed in other carbohydrate-metabolizing enzymes such as endoglucanase CelA and β-mannanase. In these examples, the aromatic residue is near the active site and stacks with the sugar ring of the substrate ring by orienting the aromatic ring of the side chain in parallel [26,27]. However, in trehalases, substrate recognition does not involve stacking interactions of aromatic rings. The aromatic residue (W^{447}) interacts with the substrate by formation of a hydrogen bond with the –OH group at the C2 position of the ligand. Interestingly, earlier studies with derivatives of trehalose have shown that substitutions in the C2 position affect the

catalytic activity of trehalase, implying that the interaction of W^{447} plays a significant role in the catalytic activity [20]. Interaction of a tryptophan residue with the substrate that assists substrate binding in active site pocket has been demonstrated in other glycosyl hydrolases, including lysozyme and chitinases. In these hydrolases, tryptophan residues were involved in hydrogen bonding with the substrate through the Nε2 atom of the side chain (as observed in Px_EclTre) or formed a stacking interaction with sugar rings of the substrate [28–30].

In the present study, we have highlighted a set of five salt bridges. While three of these salt bridges were formed upon binding of the correct ligand at the active site, the remaining four residues formed a network of constitutive salt bridges (present independent of ligand binding). Apart from salt bridges, the residues from side loop and lid loop are involved in substrate recognition and interaction at the +1 subsite. The binding of a correct ligand at the active site led to reorienting the residues Q^{446}, E^{496}, K^{497} and Y^{512} such that residues K^{497} and Y^{512} formed H-bonds with Q^{446} and E^{496}. These hydrogen bonds were not formed when there is a nonspecific ligand at the active site or when there is no ligand. Hence, the interactions identified as specific to the active conformation of Px_EclTre could be signatures of the active form of trehalases, regulated by either the correct substrate or by allostery.

Role of conserved motifs and signature sequences in trehalose catalysis

Two signature motifs (PGGRFIEFYYWDSY and QWDFPNVWPP) and a glycine-rich region (GGG GEY) are conserved across all the insect trehalases [31]. Additionally, the dipterans also have been reported to contain five conserved sequences (DSKTFVDMK, IPNGGRV/IYY, RSQPPF/LL, GP RPESYREDI and ELKAG/AAESGMDFSSRWFV) [5,31]. Although the trehalase sequences from multiple sources (bacteria, yeast, fungi, insects, dipterans, etc.) belonging to different life forms are known, a comprehensive analysis correlating the conserved sites and sequence motifs to their function was not reported earlier.

Upon aligning trehalase sequences from 100 different sources, as picked up from UniProt, we find that the conserved motifs from the N terminus to C terminus are: $[D/R]^{59}X[K/Q/Y]XXXDXX^{67}$, $P^{150}GGRFX EXYXWDXY^{163}$, $N^{197}XXRXXYXX[R/H]SQPP^{210}$, $S^{309}G[W/H]DX[S/T]XR^{316}$, $P^{334}[I/V/A][E/D]LN^{338}$, $[D/Y]^{389}[Y/F/W][N/D]^{391}$, $Q^{446}WDXPXXWXP^{455}$ and $E^{511}YXXXXGFGWTN^{522}$ (numbering according to Px_EclTre). These also correspond to the conserved sequence stretches in dipteran trehalase sequences. The conserved motif $E^{511}YXXXXGFGWTN^{522}$ is preceded by GGGG, a 'Glycine-rich' region, in most of the trehalases. Apart from the aforesaid motifs, a few other residues are also conserved: H^{118}, W^{123}, Y^{158}, G^{168}, L^{169}, M^{215}, E^{237}, Y^{269}, $[R/P]^{277}$, E^{279}, L^{337}, Y^{488}, E^{496}, K^{497}. As mentioned earlier, in yeast structures where there is a P^{422}, corresponding to R^{277} in bacterial trehalases, the salt bridge is formed by R^{473}. Interestingly, we observe that these sequence motifs contain most of the residues implicated in the salt bridges and tryptophan residues identified as residues of potential functional significance. While E^{496} is a conserved catalytic base, the binding of the correct substrate led to the formation of hydrogen bonds by E^{496} and Y^{512} with Q^{446} and K^{497}. The residues R^{277}, E^{279} were involved in the formation of salt bridges with residues from conserved motifs. The rest of the conserved residues formed interacting pairs (H^{118} & Y^{488}, W^{123} & L^{169}, Y^{269} & L^{337}, Y^{158} & E^{237}) and are probably involved in structural stability. Mutational studies on these residues can provide further insights into their function. Conserved residues and their functions are summarized in Table 3.

Based on our structural, biochemical and mutational studies, we have assigned a functional role to the conserved residues in trehalases. The study highlights the role of salt bridges and a cluster of tryptophans in substrate specificity and catalysis. These strong interactions surrounding the active site might contribute to the optimal activity of the enzyme at a high temperature and acidic pH.

Design of validoxylamine A or validamycin A-based trehalase inhibitors

Many life forms including bacteria, yeast, plants, mice and humans have been reported to produce trehalase enzyme [19,32–35]. Interestingly, the sensitivity of trehalase from different sources to inhibitors such as validamycin A or Validoxylamine A seems to vary. For instance, feeding rats with 500 and 1% m/v validamycin A or Validoxylamine A is not toxic and no abnormalities were observed in F_0 and F_1 generations [36]. In normal humans, the expression of trehalase is at very low levels. Elevated expression of trehalase is observed only in cases of kidney disorders or untreated coeliac disease [37]. The use of trehalase inhibitors Validoxylamine A and Validamycin A as antifungal agents against plant pathogenic fungi has already been reported [18,38]. Validoxylamine A, a breakdown product of Validamycin A, can be naturally degraded to

Table 3. Residues conserved across trehalases and summary of their functional roles.

Residue[a]	Function
D^{59}& R^{153}, K^{61}& E^{156}, D^{65}& R^{200}, R^{206} & E^{279}, and R^{277}& E^{511}	Salt bridging pairs
R^{153}, R^{277}, E^{279}, E^{511}	Conserved residues interacting at +1 site
W^{160}, W^{311} [b], W^{447}, W^{453} and W^{520}	Conserved tryptophans near −1 site
D^{312} and E^{496}	Catalytic residues
R^{153}, F^{154}, Y^{158}, E^{511}, Y^{512}, F^{518} Q^{446}, E^{496} and K^{497}	Substrate recognition
H^{118}& Y^{488}, W^{123}& L^{169}, Y^{269}& L^{337}, Y^{158}& E^{237} R^{316}, D^{336} and D^{389}	Interacting pairs of residues, function not assigned

[a]Numbering according to bacterial trehalase residues in Px_EclTre. [b]Replaced by a histidine in some and methionine in one structure.

nonhazardous products Valienamine and Validamine [35].

Thus, in addition to its use as an antifungal agent, the use of validamycin A or validaxylamine A-based trehalase inhibitors appears to be a reasonable choice against insect pests, based on our study of inhibition of trehalase from an essential insect gut bacterium by Validoxylamine A [39]. The information about the interaction of Validoxylamine A with Px_EclTre-binding pocket can be further explored for the structure–activity relationship to optimize the inhibition potential and specificity of inhibitor molecule. Understanding of binding mode of trehalases with novel inhibitors that disrupt the salt bridges and the tryptophan interactions can be utilized for fine tuning enzyme–inhibitor interaction for the design of more effective and specific trehalase inhibitors.

Materials and methods

Production and purification of recombinant Px_EclTre

The sequence for the Px_Ecl_*tre* (Px000318) was obtained from the Diamondback moth Genome Database (http://iae.fafu.edu.cn/DBM) [40]. This sequence codes for an 890 aa protein. The trehalase coding sequence from Px000318 (1599 bp) was identified. Signal peptide from the amino acid sequence of Px_EclTre was predicted using SignalP 4.1 server and was excluded while designing primers. RNA was isolated from 100 mg of larval tissue using TriZol (Thermo Fisher Scientific, Waltham, MA, USA). RNA was treated with DNase enzyme to remove DNA contamination (Promega, Fitchburg, WI, USA). Furthermore, cDNA was prepared using High-Capacity cDNA Reverse Transcription kit (Applied Biosystems, Foster City, CA, USA) as per manufacturer's instructions. Random primers were used for cDNA preparation from total RNA. Gene was amplified using specific primers and cloned in Zero Blunt® vector (Thermo Fisher Scientific), followed by cloning in pET28a (Novagen, Madison, WI, USA) vector with EcoRI and HindIII restriction sites. Sequencing confirmed that cloned sequence corresponds to *E. cloacae* periplasmic trehalase (Refseq Accession ID: WP_045355849.1). cDNA prepared using oligo dT primers do not show amplification for this gene, which in turn indicates its bacterial origin.

BL21 Star (DE3) cells transformed with Px_EclTre construct were grown in Luria–Bertani (LB) media containing 50 $\mu g \cdot mL^{-1}$ Kanamycin to OD_{600} 0.6–0.8. The cells were then induced with isopropyl-β-D-1-thiogalactopyranoside at the final concentration of 0.5 mM and incubated at 16 °C for 10 h. Cells were harvested by centrifugation and lysed by sonication in lysis buffer (50 mM Tris pH 8.0, 500 mM NaCl and 10% Glycerol). The lysate was centrifuged at 18 500 ***g*** for 15 min at 4 °C. The clear supernatant thus obtained was allowed to bind with 700 μL Pro Bond Ni–nitrilotriacetic acid resin (Thermo Fisher Scientific) pre-equilibrated with binding buffer (50 mM Tris pH 8.0 and 500 mM NaCl) and then loaded on Bio-Rad gravity column (Bio-Rad, Hercules, CA, USA). Nonspecifically bound proteins were washed away with 25 mL wash buffer (50 mM NaH_2PO_4, 300 mM NaCl and 20 mM Imidazole pH 8.0). The protein bound to resin was eluted with 200 mM Imidazole in the wash buffer. The protein was concentrated by passing through 10 kDa cut-off membrane (Millipore, Billerica, MA, USA). The concentrated protein was further purified by size exclusion chromatography using Superdex-75 column (GE Healthcare, Buckinghamshire, UK) on ÄKTA system (GE Healthcare) and eluted into 50 mM Tris pH 8.0 and 50 mM NaCl. Purified fractions were pooled and concentrated for further activity studies and crystallization experiments. The mutants were purified using a Ni-nitrilotriacetic acid affinity chromatography as described earlier (Fig. 4A,B).

Biochemical characterization of Px_EclTre and mutants

Trehalase activity was estimated by monitoring release of glucose from α, α-Trehalose using DNSA reagent [9]. A premix containing 50 mM Tris-HCl pH 7.0 (130 μL) and equal units (32 nM) of purified Px_EclTre protein (20 μL) was prepared to which 150 μL of Trehalose (0.25%) was added. The reaction was incubated for 15 min at 37 °C and then stopped by adding 500 μL of DNSA reagent. The reaction tubes were incubated in a boiling water bath for 5 min and absorbance was measured at 540 nm. One unit of trehalase was defined as the amount of enzyme required to release 1 μM glucose·min^{-1} at 37 °C from trehalose under the given assay conditions.

Enzyme activity of Px_EclTre was measured at different pH values ranging from 2 to 12. Equal units of enzyme in citrate-phosphate, sodium phosphate, Tris and glycine-NaOH buffers (50 mM each) were used for the activity assay at specified pH values 2–6, 7, 8–9 and 10–12 respectively. For determining the optimum temperature for trehalose hydrolysis, the activity of purified Px_EclTre was measured at pH 5.0 using 50 mM citrate sodium phosphate buffer at various temperatures ranging from 15 °C to 95 °C.

Substrate saturation assay (Michaelis–Menten kinetics) was performed to calculate K_M and V_{max} of the recombinant enzyme. The assay was performed at 37 °C and pH 5 using various concentrations of trehalose (0.8–6.5 mM) as substrate. K_M and V_{max} for Px_EclTre were calculated by fitting the data to a Michaelis–Menten plot using GRAPHPAD PRISM v6.0 (GraphPad Software, San Diego, CA, USA). For determination of inhibition kinetics and K_i of Validoxylamine A (obtained by acid hydrolysis of Validamycin A), enzyme activity assays were performed at 6.6 and 33 nM inhibitor concentration. Trehalase inhibition values obtained were used to generate Michaelis–Menten and double reciprocal plot to determine K_i and mode of inhibition respectively. Results were analysed by plotting double reciprocal plots and data fitting into competitive inhibition model of GRAPHPAD PRISM v6.0. This model uses the equation of $K_M^{Obs} = K_M \times (1 + [I]/K_i)$, where K_M^{Obs} is apparent K_M and [I] is inhibitor concentration.

All assays were performed as two independent experiments in triplicate and obtained data were subjected to analysis of variance (ANOVA) using GRAPHPAD PRISM v6.0 (GraphPad Software). The level of significance between different groups was judged by *F*-test, while data were compared using the least significant differences at $P \leq 0.05$. In all the figures, error bars represent ±standard error for the observations of each treatment.

Crystallization of Px_EclTre

Purified Px_EclTre (11 mg·mL^{-1} in the buffer) was used to set up a crystallization screen of 480 commercially available conditions by sitting drop method with the help of Mosquito Nanodrop Crystallization Robot (TTP Labtech, Melbourn, UK). About 100 nL of protein was mixed with 100 nL of the crystallization condition and the plates were incubated at 18 °C. Among the various initial hits obtained, one of the conditions (0.17 M Ammonium Sulphate, 25% poly(ethylene glycol) 4000 and 15% glycerol) was optimized further. The optimized crystallization condition consisted of 0.5 M ammonium sulphate, 25% poly (ethylene glycol) 4000 and 15% glycerol. Equal volumes of optimized crystallization condition and protein (31 mg·mL^{-1}) were mixed and incubated at 18 °C to obtain crystals. Plate-like crystals were observed around 48–72 h. A mixture of 1 mM (final concentration) Validamycin A (Duchefa Biochemie B.V., Haarlem, Netherlands) and purified Px_EclTre (10 mg·mL^{-1}) was used for crystallization of the liganded complex. Diffraction quality crystals were obtained in the condition with 0.2 M ammonium sulphate and 28% poly(ethylene glycol) 4000. The cryo protectant for unliganded Px_EclTre consists of 0.17 M ammonium sulphate, 25.5% poly(ethylene glycol) 4000, 20% glycerol while 0.2 M ammonium sulphate, 28% poly(ethylene glycol) 4000 and 20% glycerol was used for Px_EclTre crystals in complex with Validamycin A.

Data collection, structure determination and refinement

X-ray diffraction data were collected at beamline I04 of Diamond Light Source (DLS) Synchrotron facility, United Kingdom. Data were processed using XDS [41] and SCALA [42]. The structure was determined by molecular replacement using PHASER [43], with the trehalase structure 2JG0 (Ec_Tre-TTZ) as a model. Model building and structure refinement were carried out using COOT and PHENIX respectively [44,45]. Data collection and refinement statistics are summarized in Table 2.

Interestingly, upon solving the cocrystal structure, we found that the protein had in fact Validoxylamine A bound, instead of Validamycin A used for setting up crystallization. Composite omit map calculated using PHENIX unambiguously showed that the bound ligand is indeed Validoxylamine A [46]. We believe this is possibly because of hydrolysis of Validamycin A to Validoxylamine A under acidic conditions present in the crystallization mixture for the prolonged time period.

Analysis of functional sites, mutagenesis and activity

Site-directed mutagenesis of Px_EclTre for F154A, E279A, W447A and E511A was carried out using Quick Change Lightning Site-Directed mutagenesis kit (Agilent Technology, Santa Clara, CA, USA) as per manufacturer's protocol. The primers were designed as described in the manual, and the mutation was confirmed by sequencing. Mutant plasmids were transformed in *E. coli* BL21 (DE3) cells and protein was expressed. The protocol as described for wild-type protein purification was followed for purification of mutants. Enzyme activity measurements were also carried out as described above and compared with wild-type Px_EclTre.

Thermal shift assay

Thermal shift assay was performed to gain information about the folded state of Px_EclTre mutants by comparing their thermal denaturation curves as a function of temperature, with respect to the wild-type protein. The assay was set up by adding following components to wells in a

Bio-Rad Multiplate® PCR plate™ 96-well: protein (final concentration 2 μM), Sypro Orange (Sigma-Aldrich, St. Louis, MO, USA) (diluted to 50× in DMSO: buffer 1 : 10) dye at a final concentration of 5×, buffer (20 mM sodium phosphate buffer pH 8.0 and 100 mM KCl) was added to make the final reaction volume to 25 μL. The wells were sealed with adhesive seals (Microseal® 'B' seals; Bio-Rad). Plates were briefly spun and put into a BIO-RAD CFX96 Real-Time System RT-PCR machine. The temperature was increased from 4 °C to 90 °C with an increment of 0.4 °C/20 s. Fluorescence was monitored using FRET channel 6 (excitation λ = 450–490 nm and detection λ = 560–580 nm). Melting temperature (T_m) of wild-type and the mutants was estimated from the plots of relative fluorescence vs temperature.

Circular dichroism spectroscopy

Circular dichroism (CD) assays were carried out using Jasco spectropolarimeter (Jasco, Tokyo, Japan) under constant nitrogen flow [47]. Far-UV spectra were recorded using 0.1-cm path length quartz cuvette. Five consecutive measurements were accumulated to obtain mean spectra. The observed ellipticities were converted into molar ellipticities (θ) based on molecular mass per residue of wild-type and mutants.

MicroScale thermophoresis of the Px_EclTre interaction with validoxylamine A

Microscale thermophoresis experiments were carried out using the Monolith NT 115 instrument from Nanotemper technologies [48]. His-tag labelled Px_EclTre was kept constant at a concentration of 5 nM, whereas the concentration of unlabelled partner Validoxylamine A varies from 0.61E-3 to 1.25 μM in 1× PBS. LED power of 100% and MST power of 60% were used for this experiment. Data analysis was carried out using NTanalysis (Nanotemper technologies, Munchen, Germany) [49].

Conservation analysis

In order to identify the residues conserved across trehalases, Consurf server was used [50]. Conservation pattern was visualized and depicted using The PYMOL MOLECULAR GRAPHICS SYSTEM, Version 2.0, Schrödinger, LLC (New York City, NY, USA).

Acknowledgements

The project work is supported by the research grant from the Department of Science and Technology – Science and Engineering Research Board (DST-SERB), Government of India under ECR/2015/000502 grant and Board of College and University Development, Savitribai Phule Pune University, Pune 411007, Maharashtra, India. Work in the lab of PG is supported by INSPIRE Faculty Fellowship (IFA12-LSBM-52); Indian National Science Academy Research grant; Women Excellence Award, Department of Science and Technology; and Innovative Young Biotechnologist Award, Department of Biotechnology. AA acknowledges UPE II for fellowship. SH acknowledges IISER Pune for fellowship. Authors also acknowledge Dr Sivaramaiah Nallapeta and Dr Saji Menon, NanoTemper Technologies for their assistance in Microscale thermophoresis analysis. Synchrotron facilities at ESRF, Grenoble, and Diamond Light Source, UK are acknowledged. We thank EMBL-ESRF and DBT India for facilitating data collection at ID29, ESRF; and beamline scientists at I03 and I04 (MX14198, MX15369, MX16452 and MX17251) at the Diamond Light Source, UK.

Conflict of interest

The authors declare no conflict of interest.

Author contributions

AA, SH, AB and RJ designed and carried out the experiments, SH and PG solved crystal structures, PG provided structural insights, AG, PG and RJ supervised overall work and wrote paper with comments from all authors.

References

1 Furlong MJ, Wright DJ & Dosdall LM (2013) Diamondback Moth ecology and management: problems, progress, and prospects. *Annu Rev Entomol* **58**, 517–541.
2 Xia X, Gurr GM, Vasseur L, Zheng D, Zhong H, Qin B, Lin J, Wang Y, Song F, Li Y *et al.* (2017) Metagenomic sequencing of diamondback moth gut microbiome unveils key holobiont adaptations for herbivory. *Front Microbiol* **8**, 663.
3 Becker A, Schlöder P, Steele JE & Wegener G (1996) The regulation of trehalose metabolism in insects. *Experientia* **52**, 433–439.
4 Wyatt GR (1957) The chemistry of insect hemolymph: II. Trehalose and other carbohydrates. *J Gen Physiol* **40**, 833–847.
5 Barraza A, Sánchez F, Barraza A & Sánchez F (2013) Trehalases: a neglected carbon metabolism regulator? *Plant Signal Behav* **8**, e24778.
6 Forcella M, Cardona F, Goti A, Parmeggiani C, Cipolla L, Gregori M, Schirone R, Fusi P & Parenti P

(2010) A membrane-bound trehalase from *Chironomus riparius* larvae: purification and sensitivity to inhibition. *Glycobiology* **20**, 1186–1195.
7 Silva MCP, Terra WR & Ferreira C (2004) The role of carboxyl, guanidine and imidazole groups in catalysis by a midgut trehalase purified from an insect larvae. *Insect Biochem Mol Biol* **34**, 1089–1099.
8 Gibson RP, Gloster TM, Roberts S, Warren RAJ, Storch De Gracia I, García Á, Chiara JL & Davies GJ (2007) Molecular basis for trehalase inhibition revealed by the structure of trehalase in complex with potent inhibitors. *Angew Chem Int Ed Engl* **46**, 4115–4119.
9 Miller GL (1959) Use of dinitrosalicylic acid reagent for determination of reducing sugar. *Anal Biochem* **31**, 426–428.
10 Forcella M, Mozzi A, Bigi A, Parenti P & Fusi P (2012) Molecular cloning of soluble trehalase from *Chironomus riparius* larvae, its heterologous expression in *Escherichia coli* and bioinformatic analysis. *Arch Insect Biochem Physiol* **81**, 77–89.
11 Shukla E, Thorat L, Bhavnani V, Bendre AD, Pal JK, Nath BB & Gaikwad SM (2016) Molecular cloning and in silico studies of physiologically significant trehalase from *Drosophila melanogaster*. *Int J Biol Macromol* **92**, 282–292.
12 Silva MCP, Terra WR & Ferreira C (2010) The catalytic and other residues essential for the activity of the midgut trehalase from *Spodoptera frugiperda*. *Insect Biochem Mol Biol* **40**, 733–741.
13 Cheng Q, Gao H & Hu N (2016) A trehalase from *Zunongwangia* sp.: characterization and improving catalytic efficiency by directed evolution. *BMC Biotechnol*, **16**, 9.
14 Lee J-H, Saito S, Mori H, Nishimoto M, Okuyama M & Kim D (2007) Molecular cloning of cDNA for trehalase from the European honeybee, *Apis mellifera* L., and its heterologous expression in *Pichia pastoris*. *Biosci Biotechnol Biochem* **71**, 2256–2265.
15 Biswas N & Ghosh AK (1996) Characterisation of an acid trehalase of *Saccharomyces cerevisiae* present in trehalase-sucrase aggregate. *Biochim Biophys Acta* **1290**, 95–100.
16 Mansure JJ, Silva JT & Panek AD (1992) Characterization of trehalase in *Rhodotorula rubra*. *Biochem Int* **28**, 693–700.
17 Asano N (2003) Glycosidase inhibitors: update and perspectives on practical use. *Glycobiology* **13**, 93–104.
18 Asano N, Takuji Y, Yukihiko K & Katsuhiko M (1987) Effect of validamycins on glycohydrolases of *Rhizoctonia solani*. *J Antibiot (Tokyo)* **40**, 526–532.
19 Alblova M, Smidova A, Docekal V, Vesely J, Herman P, Obsilova V & Obsil T (2017) Molecular basis of the 14-3-3 protein-dependent activation of yeast neutral trehalase Nth1. *Proc Natl Acad Sci USA* **114**, 9811–9820.
20 Cardona F, Goti A, Parmeggiani C, Parenti P, Forcella M & Fusi P (2010) Casuarine-6-O-α-d-glucoside and its analogues are tight binding inhibitors of insect and bacterial trehalases. *Chem Commun* **46**, 2629–2631.
21 Cardona F, Parmeggiani C, Faggi E, Bonaccini C, Gratteri P, Sim L, Gloster TM, Roberts S, Davies GJ, Rose DR *et al.* (2009) Total syntheses of casuarine and its 6-O-α-glucoside: complementary inhibition towards glycoside hydrolases of the GH31 and GH37 families. *Chemistry* **15**, 1627–1636.
22 Koshland DE (1953) Stereochemistry and the mechanism of enzymatic reactions. *Biol Rev* **28**, 416–436.
23 Davies G & Henrissat B (1995) Structures and mechanisms of glycosyl hydrolases. *Structure* **3**, 853–859.
24 Henrissat B, Callebaut I, Fabrega S, Lehn P, Mornon JP & Davies G (1995) Conserved catalytic machinery and the prediction of a common fold for several families of glycosyl hydrolases. *Proc Natl Acad Sci USA* **92**, 7090–7094.
25 Zechel DL & Withers SG (2000) Glycosidase mechanisms: anatomy of a finely tuned catalyst. *Acc Chem Res* **33**, 11–18.
26 Alzari PM, Souchon H & Dominguez R (1996) The crystal structure of endoglucanase CelA, a family 8 glycosyl hydrolase from *Clostridium thermocellum*. *Structure* **4**, 265–275.
27 Hilge M, Gloor SM, Rypniewski W, Sauer O, Heightman TD, Zimmermann W, Winterhalter K & Piontek K (1998) High-resolution native and complex structures of thermostable beta-mannanase from *Thermomonospora fusca* – substrate specificity in glycosyl hydrolase family 5. *Structure* **6**, 1433–1444.
28 Phillips DC, Cheetham JC & Ox O (1992) Refinement of an enzyme complex with inhibitor bound at partial occupancy. Hen egg-white lysozyme and tri-N-acetylchitotriose at 1.75 A resolution. *J Mol Biol* **224**, 613–628.
29 Maenaka K, Matsushima M, Kawai G, Kidera A, Watanabe K, Kuroki R & Kumagai I (1998) Structural and functional effect of Trp-62 Gly and Asp-101 Gly substitutions on substrate-binding modes of mutant hen egg-white lysozymes. *Biochem J* **333**, 71–76.
30 Terwisscha van Scheltinga AC, Armand S, Kalk KH, Isogai A, Henrissat B & Dijkstra BW (1995) Stereochemistry of chitin hydrolysis by a plant chitinaselysozyme and X-ray structure of a complex with allosamidin: evidence for substrate assisted catalysis. *Biochemistry* **34**, 15619–15623.
31 Shukla E, Thorat LJ, Nath BB & Gaikwad SM (2015) Insect trehalase: physiological signi fi cance and potential applications. *Glycobiology* **25**, 357–367.

32 Kendall EJ, Adams RP & Kartha KK (1990) Trehalase activity in plant tissue cultures. *Phytochemistry* **29**, 2525–2528.
33 Sacktor B (1968) Trehalase and the transport of glucose in the mammalian kidney and intestine. *Proc Natl Acad Sci USA* **60**, 1007.
34 Sasai-Takedatsu M, Taketani S, Nagata N, Furukawa T, Tokunaga R & Kojima TKY (1996) Human trehalase: characterization, localization, and its increase in urine by renal proximal tubular damage. *Nephron* **73**, 179–185.
35 Chen X, Lu Y, Fan Y & Shen Y (2017) An introduction to validamycins and their derivatives. In Validamycin and Its Derivatives, 1st edn (Chen X, Lu Y, Fan Y & Shen Y, eds), pp. 1–8. Elsevier, Amsterdam.
36 Moriya M, Ohta T, Watanabe K, Miyazawa T, Kato K & Shirasu Y (1983) Further mutagenicity studies on pesticides in bacterial reversion assay systems. *Mutat Res* **116**, 185–216.
37 Alpers DH (2003) Carbohydrates: digestion, absorption, and metabolism. In *Encyclopedia of Food Sciences and Nutrition,* 2nd edn (Caballero B, ed.), pp. 881–887. Academic Press, Oxford.
38 Wu Q, Zhang L, Xia H, Yu C, Dou K, Li Y & Chen J (2017) Omics for understanding synergistic action of validamycin A and *Trichoderma asperellum* GDFS1009 against maize sheath blight pathogen. *Sci Rep* **7**, 40140.
39 Asano N, Takeuchi M, Kameda Y, Matsui K & Kono Y (1990) Trehalase inhibitors, validoxylamine A and related compounds as insecticides. *J Antibiot (Tokyo)* **43**, 722–726.
40 You M, Yue Z, He W, Yang X, Yang G, Xie M, Zhan D, Baxter SW, Vasseur L, Gurr GM *et al.* (2013) A heterozygous moth genome provides insights into herbivory and detoxification. *Nat Genet* **45**, 220–225.
41 Kabsch W (2010) XDS. *Acta Crystallogr D* **66**, 125–132.
42 Collaborative Computational Project, Number 4 (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D* **50**, 760–763.
43 McCoy AJ (2007) Solving structures of protein complexes by molecular replacement with Phaser. *Acta Crystallogr D* **63**, 32–41.
44 Emsley P & Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D* **60**, 2126–2132.
45 Adams PD, Grosse-Kunstleve RW, Hung LW, Ioerger TR, McCoy AJ, Moriarty NW, Read RJ, Sacchettini JC, Sauter NK & Terwilliger TC (2002) PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr D* **58**, 1948–1954.
46 Bhat TN (1988) Calculation of an OMIT map. *J Appl Crystallogr* **21**, 279–281.
47 Mishra M, Joshi RS, Gaikwad S, Gupta VS & Giri AP (2013) Structural–functional insights of single and multi-domain *Capsicum annuum* protease inhibitors. *Biochem Biophys Res Commun* **430**, 1060–1065.
48 Jerabek-Willemsen M, André T, Wanner R, Roth HM, Duhr S, Baaske P & Breitsprecher D (2014) MicroScale thermophoresis: interaction analysis and beyond. *J Mol Struct* **1077**, 101–113.
49 Adhav AS, Kokane SR & Joshi RS (2018) Functional characterization of *Helicoverpa armigera* trehalase and investigation of physiological effects caused due its to inhibition by validamycin A formulation. *Int J Biol Macromol* **112**, 638–647.
50 Glaser F, Pupko T, Paz I, Bell RE, Bechor D & Martz E (2003) ConSurf: identification of functional regions in proteins by surface mapping of phylogenetic information. *Bioinformatics* **19**, 163–164.

Report

Current Biology

Motor Activity Dependent and Independent Functions of Myosin II Contribute to Actomyosin Ring Assembly and Contraction in *Schizosaccharomyces pombe*

Highlights

- In many eukaryotes, cytokinesis requires an actomyosin-based contractile ring
- The role of motor activity of myosin II in cytokinesis is a topic of active debate
- We isolate a new allele of *S. pombe* Myo2, an essential myosin heavy chain
- We show motor activity-dependent and -independent roles for Myo2

Authors

Saravanan Palani, Ting Gang Chew, Srinivasan Ramanujam, ..., Mithilesh Mishra, Pananghat Gayathri, Mohan K. Balasubramanian

Correspondence

s.palani@warwick.ac.uk (S.P.), m.k.balasubramanian@warwick.ac.uk (M.K.B.)

In Brief

Cytokinesis in many eukaryotes requires an actomyosin-based contractile ring. The role of the motor protein Myosin II in cytokinesis is actively debated. Palani et al. identify a new motor activity-defective allele of fission yeast myosin II and report that the motor activity is dispensable for ring assembly but is essential for ring contraction.

Palani et al., 2017, Current Biology *27*, 1–7
March 6, 2017
http://dx.doi.org/10.1016/j.cub.2017.01.028

CellPress

Please cite this article in press as: Palani et al., Motor Activity Dependent and Independent Functions of Myosin II Contribute to Actomyosin Ring Assembly and Contraction in *Schizosaccharomyces pombe*, Current Biology (2017), http://dx.doi.org/10.1016/j.cub.2017.01.028

Current Biology
Report

Motor Activity Dependent and Independent Functions of Myosin II Contribute to Actomyosin Ring Assembly and Contraction in *Schizosaccharomyces pombe*

Saravanan Palani,[1,*] Ting Gang Chew,[1] Srinivasan Ramanujam,[2] Anton Kamnev,[1] Shrikant Harne,[5] Bernardo Chapa-y-Lazo,[1] Rebecca Hogg,[1] Mayalagu Sevugan,[3] Mithilesh Mishra,[3,4] Pananghat Gayathri,[5] and Mohan K. Balasubramanian[1,6,*]

[1]Division of Biomedical Sciences, Warwick Medical School, University of Warwick, Coventry CV4 7AL, UK
[2]School of Biological Sciences, National Institute of Science Education and Research (NISER), Odisha 752050, India
[3]Temasek Life Sciences Laboratory, 1. Research Link, National University of Singapore, Singapore 117604, Singapore
[4]Department of Biological Sciences, Tata Institute of Fundamental Research (TIFR), Mumbai, Maharashtra 400005, India
[5]Biology Division, Indian Institute of Science Education and Research (IISER), Pune, Maharashtra 411008, India
[6]Lead Contact
*Correspondence: s.palani@warwick.ac.uk (S.P.), m.k.balasubramanian@warwick.ac.uk (M.K.B.)
http://dx.doi.org/10.1016/j.cub.2017.01.028

SUMMARY

Cytokinesis depends on a contractile actomyosin ring in many eukaryotes [1–3]. Myosin II is a key component of the actomyosin ring, although whether it functions as a motor or as an actin cross-linker to exert its essential role is disputed [1, 4, 5]. In *Schizosaccharomyces pombe*, the *myo2*-E1 mutation affects the upper 50 kDa sub-domain of the myosin II heavy chain, and cells carrying this lethal mutation are defective in actomyosin ring assembly at the non-permissive temperature [6, 7]. *myo2*-E1 also affects actomyosin ring contraction when rings isolated from permissive temperature-grown cells are incubated with ATP [8]. Here we report isolation of a compensatory suppressor mutation in the lower 50 kDa sub-domain (*myo2*-E1-Sup1) that reverses the inability of *myo2*-E1 to form colonies at the restrictive temperature. *myo2*-E1-Sup1 is capable of assembling normal actomyosin rings, although rings isolated from *myo2*-E1-Sup1 are defective in ATP-dependent contraction in vitro. Furthermore, the product of *myo2*-E1-Sup1 does not translocate actin filaments in motility assays in vitro. Superimposition of *myo2*-E1 and *myo2*-E1-Sup1 on available rigor and blebbistatin-bound myosin II structures suggests that *myo2*-E1-Sup1 may represent a novel actin translocation-defective allele. Actomyosin ring contraction and viability of *myo2*-E1-Sup1 cells depend on the late cytokinetic *S. pombe* myosin II isoform, Myp2p, a non-essential protein that is normally dispensable for actomyosin ring assembly and contraction. Our work reveals that Myo2p may function in two different and essential modes during cytokinesis: a motor activity-independent form that can promote actomyosin ring assembly and a motor activity-dependent form that supports ring contraction.

RESULTS AND DISCUSSION

The product of the *myo2*-E1 allele is predicted to harbor a substitution of glycine at position 345 with arginine (Figures S1A and S1B). Cells carrying this mutant allele are capable of colony formation at 24°C but are severely compromised for colony formation at 36°C (Figure 1A) due to defective actomyosin ring assembly [6, 7, 9, 10]. The *myo2*-E1 mutation resides between α-helix HL and β sheet S1D, which is part of the upper 50 kDa sub-domain in the head of Myo2p (Figure S1B). Previous work has shown that *Myo2*-E1p (product of myo2-E1) does not bind or move actin filaments and has a very low ATPase activity in vitro [10, 11]. The presence of a bulky arginine side chain between helices HL and HO in the upper 50 kDa sub-domain of this mutant might introduce constraints to the conformational changes in the Myo2p head domain during the actomyosin cycle, resulting in the observed phenotypes. To further understand the role of Myo2p in cytokinesis, we isolated genetic suppressors that restored the ability of *myo2*-E1 cells to form colonies at 36°C (Figure 1A). One suppressor, *myo2*-E1-Sup1, is described in this study. Genetic crosses between *myo2*-E1-Sup1 and wild-type cells only produced progeny that were able to form colonies at 36°C, suggesting that the suppressor mutation was intragenic or very tightly linked to *myo2*. Nucleotide sequence determination revealed that *myo2*-E1-Sup1 contained the original G345R mutation and also had additional mutations (Q640H and F641I) (Figures S1A and S1B). Furthermore, no sequence alterations were found in the neighboring *rgf3* gene (data not shown), which has also been implicated in cytokinesis [12, 13]. Therefore, we concluded that the sequence alteration Q640H F641I was responsible for the suppression of *myo2*-E1. Interestingly, Q640H and F641I are located in the HW region of the Myo2p head (within the lower 50 KDa sub-domain), which is at a significant distance (~36 Å) from HL and S1D, the region where the original mutation resides, suggesting

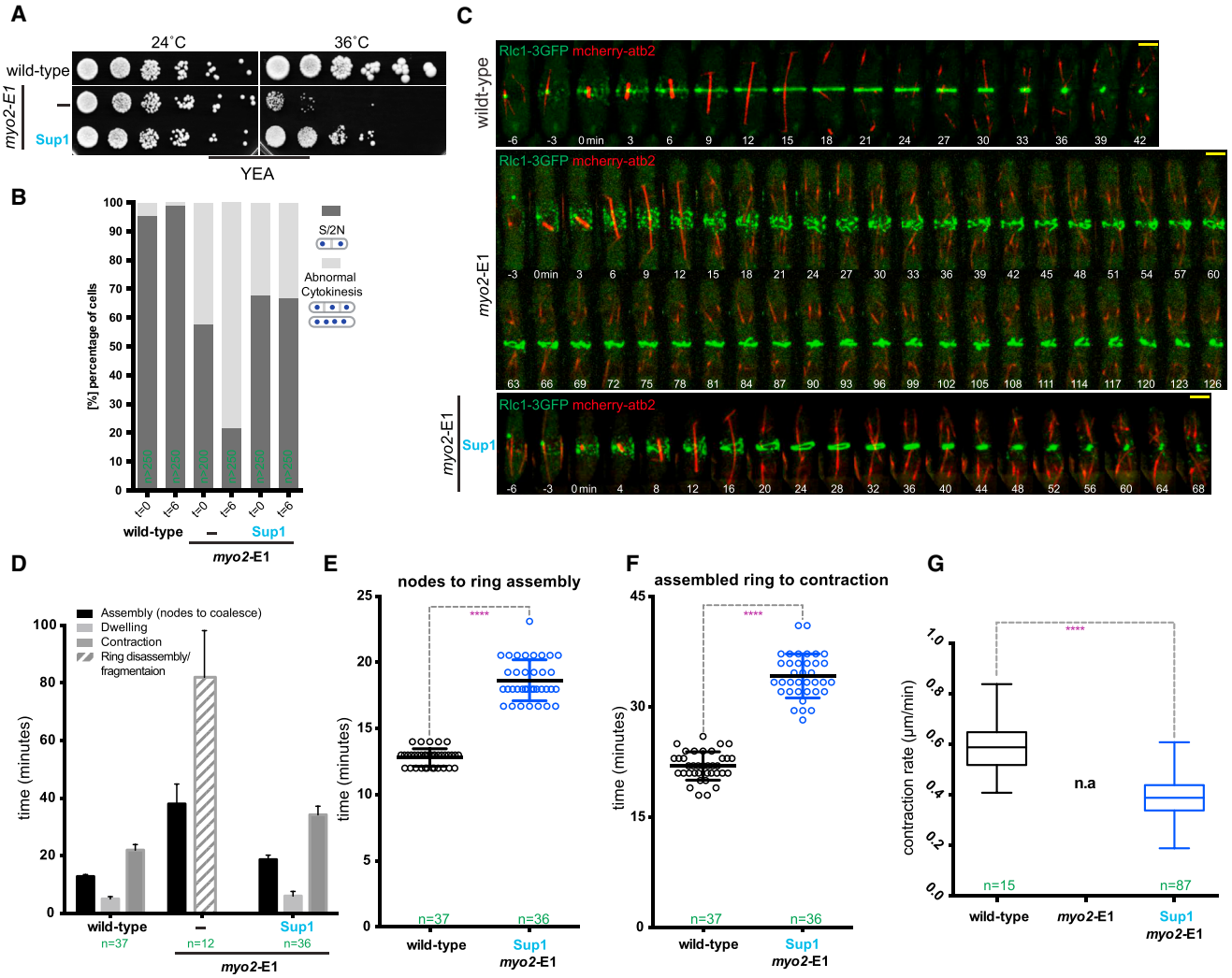

Figure 1. *myo2*-E1-Sup1 Restores Actomyosin Ring Assembly and Partial Ring Contraction

(A) Serial dilutions (10-fold) of wild-type, *myo2*-E1, and the intragenic suppressor *myo2*-E1-Sup1 were spotted onto yeast extract agar (YEA) plates and grown for 3 days at 24°C and 36°C.

(B) Quantification of DAPI and anillin blue staining used to visualize the nucleus and septum of wild-type, *myo2*-E1, and *myo2*-E1-Sup1 cells, respectively. Phenotypes of the mutants were categorized into two types: septa with two nuclei (S/2N) and cells with abnormal cytokinesis, revealed by the presence of multiple septa and nuclei (MS/>2N).

(C) Time-lapse series of wild-type, *myo2*-E1, and *myo2*-E1-Sup1 cells expressing 3GFP-tagged myosin regulatory light chain (Rlc1-3GFP) as a contractile ring marker and mCherry-tagged tubulin (mCherry-atb2) as a cell-cycle stage marker. Cells were grown at 24°C and shifted to 36°C for 3–4 hr before imaging at 36°C (t = 0 indicates the time before Rlc1-3GFP nodes localize to the cell middle). Images shown are maximum-intensity projections of z stacks. Scale bars represent 3 μm.

(D) Timing of contractile ring assembly, maturation/dwelling, and contraction. Quantification of (C) is shown. Error bars represent SD.

(E) Timing of actomyosin ring assembly from nodes. Quantification of (C) is shown (asterisks indicate the statistical significance of the difference between the two genotypes). Statistical significance was calculated by Student's t test (****$p < 0.0001$). Error bars represent SD.

(F) Timing of actomyosin ring contraction. Quantification of (C) is shown. Statistical significance was calculated by Student's t test (****$p < 0.0001$). Error bars represent SD.

(G) Constriction rate determined from a graph of ring circumference versus time. Statistical significance was calculated by Student's t test (****$p < 0.0001$). Error bars represent SD.

See also Figure S1.

potential allosteric mechanisms, rather than a simple reversal of original mutation, may operate in the suppression.

Following a 6 hr shift to 36°C, nearly 80% of *myo2*-E1 cells became multinucleate and had either improper septa with a wavy and patchy appearance or did not have a septum (Figure 1B). By contrast, only ~35% of *myo2*-E1-Sup1 cells contained such defects, while those defects were rarely seen in wild-type cells (Figures 1B and S1C). Since the ingressing actomyosin ring guides division septum assembly, we investigated the dynamics of the actomyosin ring component Rlc1p-3GFP

in wild-type, *myo2*-E1, and *myo2*-E1-Sup1 strains; mCherry-tubulin served as a cell-cycle marker in these experiments. In wild-type cells, actomyosin rings were assembled in metaphase/anaphase A in ~12.8 ± 0.6 min and contracted following spindle breakdown in ~22 ± 1.9 min, with an intervening dwell phase of 5 ± 0.8 min during which the actomyosin ring was stably maintained (Figures 1C–1E and S1D). As expected, all aspects of cytokinesis were slower in myo2-E1 mutants compared to wild-type cells: improper ring assembly took ~38 ± 6.9 min and improper contraction/disassembly lasted ~82 ± 16.2 min at 36°C (Figures 1C and S1D). Imaging *myo2*-E1-Sup1 cells revealed that they assembled actomyosin rings of normal appearance (Figure 1C, time point 24 min, ending on views in Figure S1D), with a significantly accelerated kinetics for both ring assembly (~18.6 ± 1.5 min) and contraction (~34.2 ± 3 min) compared to the original *myo2*-E1 mutant. Nevertheless, both steps were marginally slower in *myo2*-E1-Sup1 compared to wild-type cells (Figures 1C–1E and S1D). Whereas actomyosin rings in wild-type cells contracted at ~0.6 ± 0.1 μm/min, contraction rate in *myo2*-E1-Sup1 cells was ~0.4 ± 0.08 μm/min at 36°C. These experiments established that *myo2*-E1-Sup1 assembled contractile rings of normal appearance, although both ring assembly and ring contraction took ~1.5 times longer compared to wild-type cells.

Two type II myosin heavy chains participate in cytokinesis in *S. pombe* [14–17]. We therefore investigated the possibility that Myp2p, which is normally non-essential for ring assembly, assisted in actomyosin ring assembly and contraction in the *myo2*-E1-Sup1 strain through a potential ectopic upregulation. Toward this goal, we generated a double mutant of the genotype *myo2*-E1-Sup1 myp2Δ. Although this strain was viable at 24°C, surprisingly, it was inviable at 36°C (Figure 2A). Time-lapse microscopy was performed on wild-type, *myo2*-E1 *myp2*Δ, *myo2*-E1-Sup1 *myp2*Δ, and *myp2*Δ strains to investigate aspects of actomyosin ring function. The time taken for ring assembly and contraction and the ring contraction rate were comparable in wild-type and *myp2*Δ cells (Figures 2B–2F), clarifying that Myp2p is not important for either ring assembly or contraction at 36°C when Myo2p is fully functional. *myo2*-E1 *myp2*Δ assembled abnormal actomyosin rings that underwent abnormal disassembly (Figures 2B and 2C). *myo2*-E1-Sup1 *myp2*Δ assembled actomyosin rings of normal appearance, and the assembly of these rings took ~6 min more than wild-type and *myp2*Δ cells (Figures 2B–2D). Ring contraction was dramatically affected in *myo2*-E1-Sup1 *myp2*Δ (Figures 2B, 2C, 2E, and 2F). Contraction and disassembly took more than twice the amount of time compared to wild-type cells, while the ring contraction rate was less than half of that observed in wild-type cells (Figures 2E and 2F). Furthermore, contraction was frequently asymmetric and led to rings disassembling abnormally and often to the fragmentation of the ring into two or more clusters (Figures 2B, time points 48–72 min, and 2C). Since *myo2*-E1-Sup1 *myp2*Δ and *myo2*-E1-Sup1 were capable of actomyosin ring assembly but showed appreciable defects in ring contraction, we conclude Myo2p activity is essential for ring assembly and contraction, whereas Myp2p plays an ancillary role in promoting inefficient contraction when Myo2p motor activity is compromised at 36°C (compare ring contraction times and rates between *myo2*-E1-Sup1 and *myo2*-E1-Sup1 *myp2*Δ in Figures 2E and 2F).

Analysis of three-dimensional structures of rigor myosin (actin bound: 4A7F) and blebbistatin-bound myosin (actin unbound: 1YV3) suggested that the amino acid substitutions in *myo2*-E1-Sup1 may result in increased binding affinity toward F-actin (Figure S2; see the Supplemental Experimental Procedures for a detailed description of the structural analysis). This in turn may lead to defective actomyosin ring contraction due to *myo2*-E1-Sup1 being tightly bound to actin, leading to an actin filament translocation defect.

We have already developed methods to isolate ATP-dependent contraction-competent actomyosin rings [8, 18]. We therefore used this system to test if isolated actomyosin rings in cell ghosts from *myo2*-E1-Sup1 were capable of ATP-dependent contraction. Actomyosin rings were isolated from wild-type, *myo2*-E1, *myp2*Δ, *myo2*-E1 *myp2*Δ, *myo2*-E1-Sup1, and *myo2*-E1-Sup1 *myp2*Δ cells grown at the permissive temperature of 24°C. Actomyosin rings isolated from wild-type and myp2Δ cells underwent normal and rapid contraction upon ATP addition (Figures 3A and 3B). As previously reported [8], upon the addition of 0.5 mM ATP, actomyosin rings isolated from *myo2*-E1 and *myo2*-E1 myp2Δ either contracted slowly or underwent fragmentation (Figures 3A and 3B). Interestingly, despite the moderate delay in ring assembly timing, actomyosin rings of normal appearance assembled in *myo2*-E1-Sup1 and *myo2*-E1-Sup1 *myp2*Δ at the restrictive temperature. However, rings isolated from these strains did not contract normally, even at the permissive temperature for *myo2*-E1 (24°C). Instead, rings from these strains remained stable and broke into large fragments. These experiments established that, consistent with in vivo results, rings isolated from *myo2*-E1-Sup1 and *myo2*-E1-Sup1 *myp2*Δ cells are defective in ATP-dependent contraction in vitro. These results were consistent with the idea that the product of *myo2*-E1-Sup1 is defective in its motor activity and actin filament translocation, but not in actin filament binding, which in turn may explain the ability of *myo2*-E1-Sup1 to support actomyosin ring assembly, but not contraction. However, it was possible that the actin translocation defect in *myo2*-E1-Sup1 was due to allosteric effects on other unidentified components of the actomyosin ring that affect ring contraction, rather than a direct effect of *myo2*-E1-Sup1 on actin filament translocation.

To distinguish between these possibilities, we purified the products of *myo2*$^{+}$, *myo2*-E1, and *myo2*-E1-Sup1 using an expression system developed by Lord and Pollard [11]. *Myo2*-E1-Sup1p was more difficult to purify (potentially due to its tight binding to actin) and was eventually isolated from Latrunculin A-treated cells (Figure S3A). We then performed actin motility assays as described in Lord and Pollard [11]. In brief, Myo2p and the mutant versions were immobilized on nitrocellulose-coated coverslips, overlaid with rhodamine-phalloidin-stabilized rabbit actin filaments, and incubated with ATP (Figures 4A, 4B, S3B, and S3C; Movies S1, S2, S3, and S4). We found that wild-type Myo2p was able to bind and translocate actin filaments at ~0.72 ± 0.13 μm/s when incubated with ATP. Also, as previously reported [11], Myo2-E1p did not attach to actin filaments (Movie S2). Interestingly, unlike the product of *myo2*-E1, the product of *myo2*-E1-Sup1 bound actin tightly, since these filaments were either severely affected for motility or were non-motile (gliding velocity was

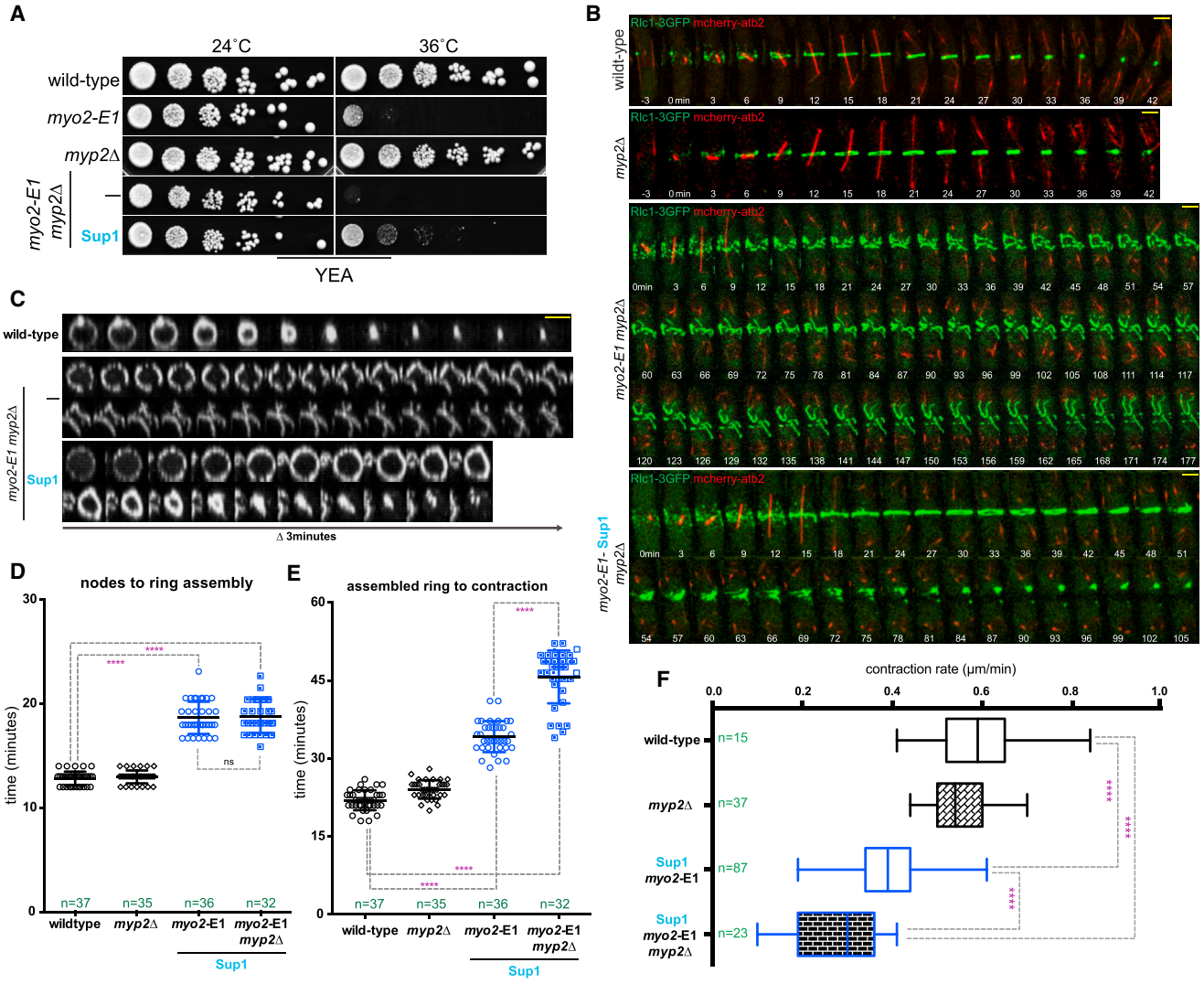

Figure 2. *myo2*-E1-Sup1 Fails in Actomyosin Ring Contraction in the Absence of the Non-essential Myosin Heavy Chain Myp2p

(A) Serial dilutions (10-fold) of wild-type, *myo2*-E1, *myp2*Δ, *myo2*-E1 *myp2*Δ, and *myo2*-E1-Sup1 *myp2*Δ were spotted onto YEA plates and grown for 3 days at 24°C and 36°C.

(B) Time-lapse series of wild-type, *myp2*Δ, *myo2*-E1 *myp2*Δ, and *myo2*-E1-Sup1 *myp2*Δ cells expressing 3GFP-tagged myosin regulatory light chain (Rlc1-3GFP) as a contractile ring marker and mCherry-tagged tubulin (atb2-mCherry) as a cell-cycle stage marker. Cells were grown at 24°C and shifted to 36°C for 3–4 hr before imaging at 36°C (t = 0 indicates the time before Rlc1-3GFP nodes localize to the cell middle). Images shown are maximum-intensity projections of z stacks. Scale bars represent 3 μm.

(C) Kymographs of a 3D-projected ring from wild-type, *myo2*-E1 *myp2*Δ, and *myo2*-E1-Sup1 *myp2*Δ cells. Scale bars represent 3 μm.

(D) Timing of actomyosin ring assembly from nodes. Quantification of (B) is shown. Asterisks indicate the statistical significance of the difference between the different genotypes compared to the wild-type. Statistical significance was calculated by Student's t test (****$p < 0.0001$). Error bars represent SD.

(E) Timing of actomyosin ring contraction. Quantification of Figure 1C and (B) is shown. Statistical significance was calculated by Student's t test (****$p < 0.0001$). Error bars represent SD.

(F) Constriction rate determined from a graph of ring circumference versus time. Contraction rates of Figure 1C and (B) are shown. Statistical significance was calculated by Student's t test (****$p < 0.0001$). Error bars represent SD.

See also Figure S2.

~0.06 ± 0.04 μm/s). Myo2-E1-Sup1p also had a dominant effect when mixed with wild-type Myo2p. The mixture bound to actin filaments but these filaments were non-motile. The fact that Myo2-E1-Sup1p did not support motility, despite binding actin filaments and its dominant-negative effect on motility over wild-type Myo2p, suggests that Myo2-E1-Sup1p is most likely a novel rigor mutant of Myo2p.

Our work reported in this study establishes that the type II myosin, Myo2p, plays two distinct and essential roles. Since cells harboring the novel rigor mutant allele *myo2*-E1-Sup1 assemble normal actomyosin rings, despite the defective contraction in vitro and in vivo, it is possible that actomyosin ring assembly depends on the ability of Myo2p to cross-link actin filaments. Actomyosin ring assembly in *myo2*-E1-Sup1 cells is

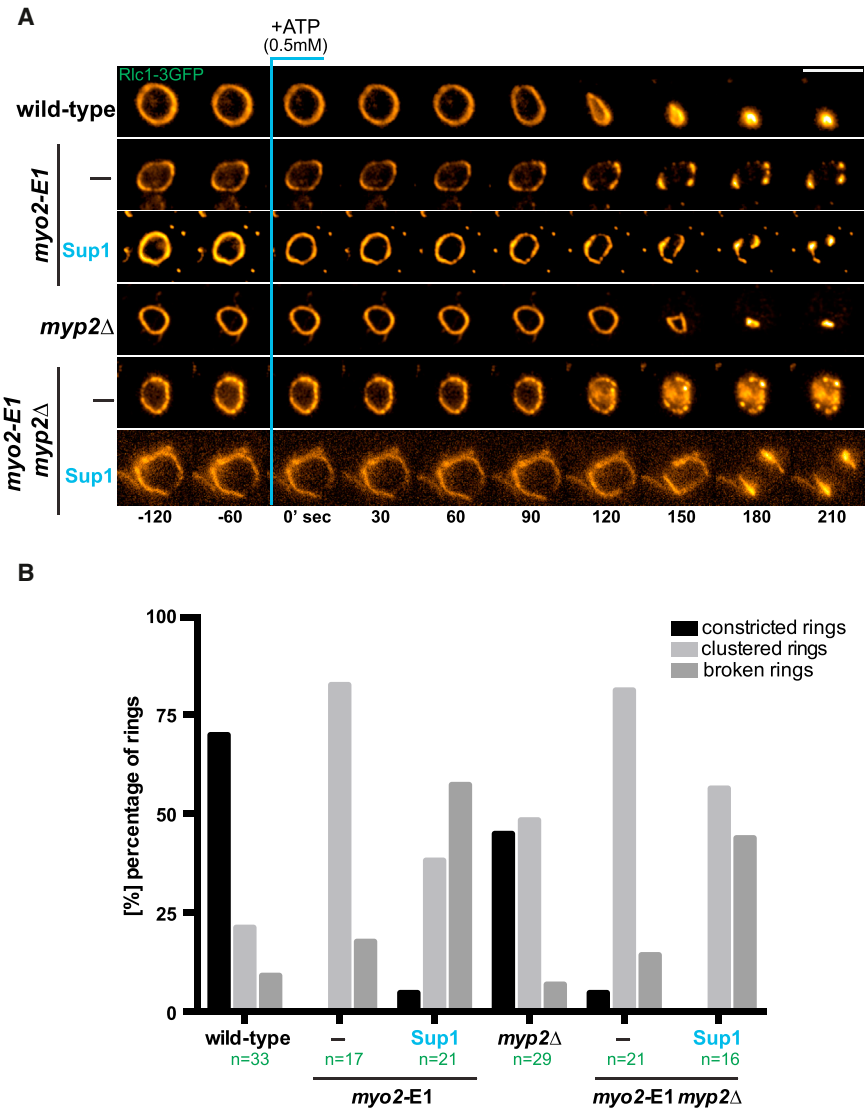

Figure 3. Isolated Actomyosin Rings of *myo2*-E1-Sup1 Do Not Undergo ATP-Dependent Contraction

(A) Cell ghosts were prepared from wild-type, *myo2*-E1, *myp2Δ*, *myo2*-E1 *myp2Δ*, *myo2*-E1-Sup1, and *myo2*-E1-Sup1 *myp2*Δ grown at 24°C. Ring contraction experiments were performed at 24°C and contraction was activated by the addition of 0.5 mM ATP. Images shown are maximum-intensity projections of z stacks. Scale bars represent 5 μm.

(B) Graph showing percentage of contracted, clustered, and broken rings. Quantification of (A) is shown.

See also Figure S2.

slower than in wild-type cells (possibly due to cross-linking and tighter binding of Myo2-E1-Sup1p with actin), suggesting that myosin II motor activity may also play a role in actomyosin ring assembly, as previously proposed [19, 20]. It is possible that clustering of cytokinetic precursor nodes can occur through tension generated by myosin II-dependent cross-linking of actin filaments. This view is consistent with aspects of the work of Ma and colleagues who have proposed that actin translocation activity of myosin II is not essential for cytokinesis [4]. Inconsistent with the work of Ma and colleagues, however, are our findings that actomyosin rings in *myo2*-E1-Sup1 cells do not contract normally, that actomyosin rings isolated from those cells fail to undergo ATP-dependent contraction, and that one-step-purified Myo2-E1-Sup1p does not support ATP-dependent actin filament motility in vitro. These observations suggest that myosin II motor activity is essential for actomyosin ring contraction.

Thus, through the analysis of novel myosin II mutant alleles, we have been able to discriminate between myosin II motor activity-dependent and -independent steps in cytokinesis. Published work in *S. cerevisiae* and mammalian cells [4, 5, 21] has questioned the role of myosin II motor activity in cytokinesis. It is likely that in some cell types, tension generated by actin filament cross-linking and filament disassembly alone may suffice for cytokinesis, whereas in others such as *S. pombe*, cytokinesis may depend on motor activity-dependent and -independent functions of myosin II.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, three figures, and four movies and can be found with this article online at http://dx.doi.org/10.1016/j.cub.2017.01.028.

AUTHOR CONTRIBUTIONS

S.P. conceived and designed experiments, acquired data, performed analysis and interpretation of data, and drafted/revised the article. S.R. and M.M. generated yeast strains and performed preliminary analysis. T.G.C., A.K., S.H., B.C.L., M.S., and R.H. performed analysis and interpretation of data

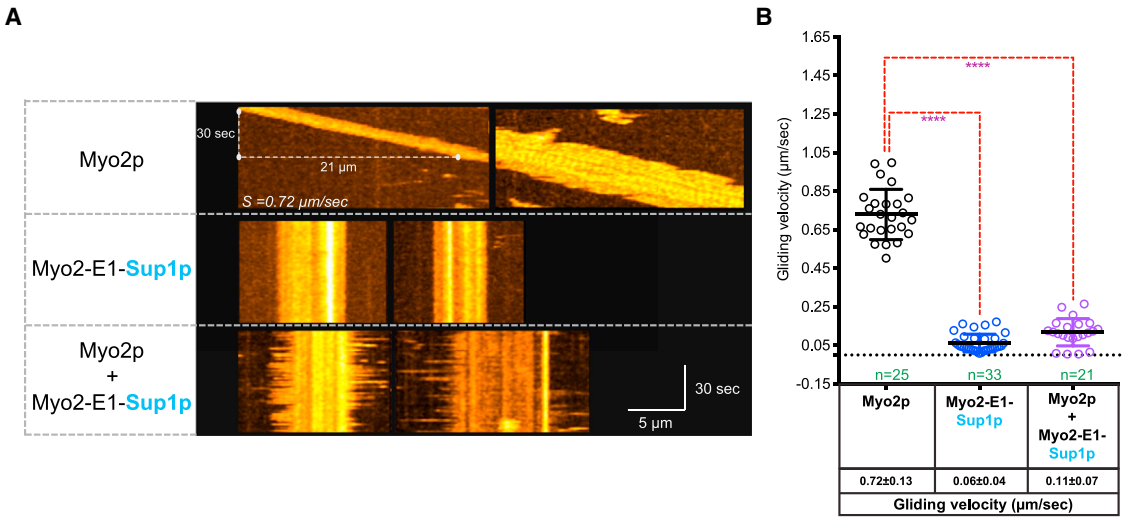

Figure 4. Myo2-E1-Sup1p Showed Tighter Actin Binding but No Motility

(A) Type II Myosin-based actin filament-gliding assay. Representative kymographs of time-lapse fluorescence micrographs of actin filaments labeled with rhodamine-phalloidin are shown. Scale bars represent 5 µm.

(B) Quantification of the actin filament-gliding assay of (A). Different myosins (Myo2p, Myo2-E1-Sup1p, and Myo2p + Myo2-E1-Sup1p) were tested for gliding velocity (µm/s) using rhodamine-phalloidin-labeled actin.

See also Figure S3 and Movies S1, S2, S3, and S4.

and generated yeast strains and reagents. P.G. performed structural analysis and interpretation of data and drafted/revised the article. M.K.B. conceived the project, conceived and designed experiments, and performed analysis and interpretation of data. S.P. and M.K.B. wrote the manuscript. All authors reviewed the manuscript.

ACKNOWLEDGMENTS

We thank Matt Lord, Kathy Trybus, and Luther Pollard for yeast strains and plasmids. Many thanks are due to members of the Balasubramanian laboratory for discussion and Rob Cross for critical comments. This work was funded by Warwick Medical School, Royal Society Wolfson Merit Award, and Wellcome Trust (WT101885MA). The early part of the work (described in Figure 1A) was performed in Temasek Life Sciences Laboratory, Singapore. P.G. acknowledges fellowships from INSPIRE, Department of Science and Technology, Government of India and an Innovative Young Biotechnologist Award (IYBA), Department of Biotechnology. S.H. acknowledges IISER Pune for a PhD fellowship. M.M. is an Intermediate Fellow of the Wellcome Trust-DBT India Alliance (IA/I/14/1/501317). M.M. acknowledges the India Alliance and the DAE/TIFR for funds.

Received: April 12, 2016
Revised: November 21, 2016
Accepted: January 16, 2017
Published: February 23, 2017

REFERENCES

1. Cheffings, T.H., Burroughs, N.J., and Balasubramanian, M.K. (2016). Actomyosin Ring Formation and Tension Generation in Eukaryotic Cytokinesis. Curr. Biol. *26*, R719–R737.
2. Pollard, T.D., and Wu, J.Q. (2010). Understanding cytokinesis: lessons from fission yeast. Nat. Rev. Mol. Cell Biol. *11*, 149–155.
3. Green, R.A., Paluch, E., and Oegema, K. (2012). Cytokinesis in animal cells. Annu. Rev. Cell Dev. Biol. *28*, 29–58.
4. Ma, X., Kovács, M., Conti, M.A., Wang, A., Zhang, Y., Sellers, J.R., and Adelstein, R.S. (2012). Nonmuscle myosin II exerts tension but does not translocate actin in vertebrate cytokinesis. Proc. Natl. Acad. Sci. USA *109*, 4509–4514.
5. Mendes Pinto, I., Rubinstein, B., Kucharavy, A., Unruh, J.R., and Li, R. (2012). Actin depolymerization drives actomyosin ring contraction during budding yeast cytokinesis. Dev. Cell *22*, 1247–1260.
6. Balasubramanian, M.K., McCollum, D., Chang, L., Wong, K.C., Naqvi, N.I., He, X., Sazer, S., and Gould, K.L. (1998). Isolation and characterization of new fission yeast cytokinesis mutants. Genetics *149*, 1265–1275.
7. Wong, K.C., Naqvi, N.I., Iino, Y., Yamamoto, M., and Balasubramanian, M.K. (2000). Fission yeast Rng3p: an UCS-domain protein that mediates myosin II assembly during cytokinesis. J. Cell Sci. *113*, 2421–2432.
8. Mishra, M., Kashiwazaki, J., Takagi, T., Srinivasan, R., Huang, Y., Balasubramanian, M.K., and Mabuchi, I. (2013). In vitro contraction of cytokinetic ring depends on myosin II but not on actin dynamics. Nat. Cell Biol. *15*, 853–859.
9. Kitayama, C., Sugimoto, A., and Yamamoto, M. (1997). Type II myosin heavy chain encoded by the myo2 gene composes the contractile ring during cytokinesis in Schizosaccharomyces pombe. J. Cell Biol. *137*, 1309–1319.
10. Stark, B.C., James, M.L., Pollard, L.W., Sirotkin, V., and Lord, M. (2013). UCS protein Rng3p is essential for myosin-II motor activity during cytokinesis in fission yeast. PLoS ONE *8*, e79593.
11. Lord, M., and Pollard, T.D. (2004). UCS protein Rng3p activates actin filament gliding by fission yeast myosin-II. J. Cell Biol. *167*, 315–325.
12. Morrell-Falvey, J.L., Ren, L., Feoktistova, A., Haese, G.D., and Gould, K.L. (2005). Cell wall remodeling at the fission yeast cell division site requires the Rho-GEF Rgf3p. J. Cell Sci. *118*, 5563–5573.
13. Davidson, R., Laporte, D., and Wu, J.Q. (2015). Regulation of Rho-GEF Rgf3 by the arrestin Art1 in fission yeast cytokinesis. Mol. Biol. Cell *26*, 453–466.
14. Laplante, C., Berro, J., Karatekin, E., Hernandez-Leyva, A., Lee, R., and Pollard, T.D. (2015). Three myosins contribute uniquely to the assembly and constriction of the fission yeast cytokinetic contractile ring. Curr. Biol. *25*, 1955–1965.

15. Bezanilla, M., Forsburg, S.L., and Pollard, T.D. (1997). Identification of a second myosin-II in Schizosaccharomyces pombe: Myp2p is conditionally required for cytokinesis. Mol. Biol. Cell *8*, 2693–2705.
16. Motegi, F., Nakano, K., Kitayama, C., Yamamoto, M., and Mabuchi, I. (1997). Identification of Myo3, a second type-II myosin heavy chain in the fission yeast Schizosaccharomyces pombe. FEBS Lett. *420*, 161–166.
17. Bezanilla, M., Wilson, J.M., and Pollard, T.D. (2000). Fission yeast myosin-II isoforms assemble into contractile rings at distinct times during mitosis. Curr. Biol. *10*, 397–400.
18. Huang, J., Mishra, M., Palani, S., Chew, T.G., and Balasubramanian, M.K. (2016). Isolation of cytokinetic actomyosin rings from Saccharomyces cerevisiae and Schizosaccharomyces pombe. Methods Mol. Biol. *1369*, 125–136.
19. Vavylonis, D., Wu, J.Q., Hao, S., O'Shaughnessy, B., and Pollard, T.D. (2008). Assembly mechanism of the contractile ring for cytokinesis by fission yeast. Science *319*, 97–100.
20. Wu, J.Q., Sirotkin, V., Kovar, D.R., Lord, M., Beltzner, C.C., Kuhn, J.R., and Pollard, T.D. (2006). Assembly of the cytokinetic contractile ring from a broad band of nodes in fission yeast. J. Cell Biol. *174*, 391–402.
21. Lord, M., Laves, E., and Pollard, T.D. (2005). Cytokinesis depends on the motor domains of myosin-II in fission yeast but not in budding yeast. Mol. Biol. Cell *16*, 5346–5355.

Chapter 10
Structure and Dynamics of Actin-Like Cytomotive Filaments in Plasmid Segregation

Pananghat Gayathri and Shrikant Harne

Abstract One of the well-known functions of the bacterial cytoskeleton is plasmid segregation. Type II plasmid segregation systems, among the best characterized with respect to the mechanism of action, possess an actin-like cytomotive filament as the motor component. This chapter describes the essential components of the plasmid segregation machinery and their mechanism of action, concentrating on the actin-like protein family of the bacterial cytoskeleton. The structures of the actin-like filaments depend on their nucleotide state and these in turn contribute to the dynamics of the filaments. The components that link the filaments to the plasmid DNA also regulate filament dynamics. The modulation of the dynamics facilitates the cytomotive filament to function as a mitotic spindle with a minimal number of components.

Keywords ParM • Actin-like • Alps • Plasmid segregation type II • Centromere • ParMRC system • R1 plasmid • Cytomotive filament • Treadmilling • ParR • PALM • ParRC helix • AlfA • AlfB • Alp12A • Alp7A • pSK41 • Bipolar spindle

Plasmid Segregation

Unlike high-copy number plasmids, low-copy number plasmids need to ensure that they are inherited to both daughter cells before or during bacterial cell division. To achieve this, they encode dedicated machinery that keeps the replicated plasmids apart within the bacterial cell. This greatly enhances the probability of equal partitioning to the daughter cells and facilitates the maintenance of the plasmids, extrachromosomal genetic material, across generations of bacteria.

P. Gayathri (✉) • S. Harne
Indian Institute of Science Education and Research (IISER),
Dr. Homi Bhabha Road, Pune 411008, India
e-mail: gayathri@iiserpune.ac.in

J. Löwe, L.A. Amos (eds.), *Prokaryotic Cytoskeletons*, Subcellular Biochemistry 84, DOI 10.1007/978-3-319-53047-5_10

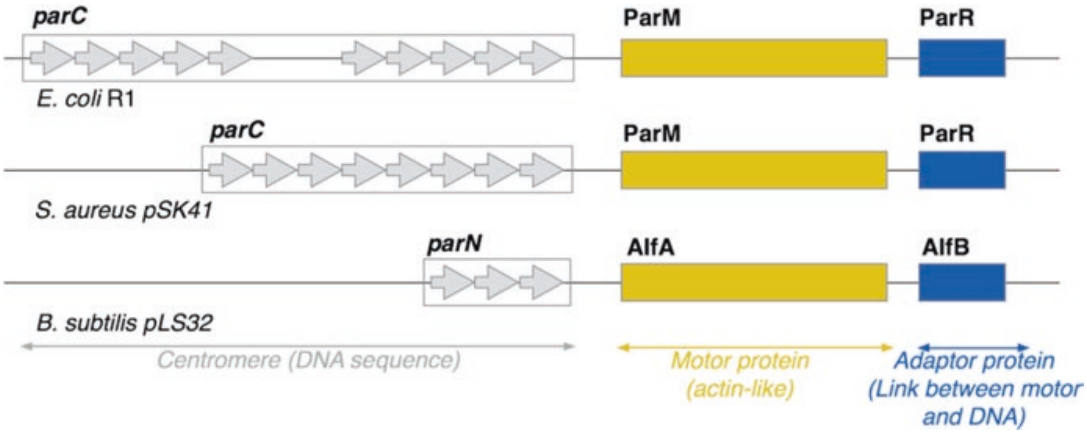

Fig. 10.1 Organization of *par* loci encoding actin-like protein homologs. In *E. coli*, *parC*, the centromere, is formed by two sets of five 11-base pair (bp) repeats separated by a 39 bp region. In *S. aureus* pSK41 and *B. subtilis* AlfA, the centromere is a continuous region formed by eight and three short repeats (iteron) respectively. The centromere region (each *arrow* represents an iteron) is followed by genes that encode a motor protein of the actin fold (*yellow*) and an adapter protein (*blue*). In general, the promoter region of the motor protein is located within the centromere sequence. Adapter proteins can thus repress the expression of motor proteins by binding to the centromere region. The lengths of the DNA segments shown for the centromere sequences and the protein-encoding regions are not drawn to scale

Components of a Plasmid Segregation System

Plasmid segregation systems were discovered in the early 1980s, well before the existence of filamentous proteins was established, and found to confer stability to plasmids containing them (Gerdes and Molin 1986). The first plasmid segregation locus to be identified was the ParMR*C* plasmid stability locus, first named *stb* (for stability). Studies carried out to understand the plasmid stability locus of *E. coli* R1 plasmid identified three components – (i) ParM – an actin-like protein that assembles into filaments in the presence of ATP as nucleotide (ii) *parC* – a centromeric DNA segment on the plasmid where ParM is attached, and (iii) ParR – an adaptor protein that couples ParM to *parC,* and also acts as a transcriptional repressor of ParM (Gerdes and Molin 1986; Dam and Gerdes 1994).

Later, characterization of a variety of plasmid and bacterial chromosome segregation loci led to the conclusion that the minimal number of necessary and sufficient components for a plasmid or DNA segregation system (Fig. 10.1) include – (i) a motor component that can actively push or pull the DNA, by utilizing energy derived from nucleotide binding/hydrolysis, (ii) a point of attachment on the plasmid DNA, a locus that functions as a centromere, and (iii) an adaptor protein that serves as a link between the DNA and the motor (Gerdes et al. 2010).

Motor

In some of the plasmid or DNA segregation systems, the motor component functions through assembly of the monomeric protein into a filamentous structure (Møller-Jensen et al. 2002). Typically, the adaptor protein mediates the binding of the centromeric DNA to the end of the filament (Jensen and Gerdes 1997). The elongation/shortening of the filament facilitates the movement of the DNA, thereby resulting in a pushing or a pulling mechanism (Møller-Jensen et al. 2002, 2003). Since it is the filament itself that acts as a linear motor, as opposed to the molecular motors of eukaryotic cytoskeletal systems such as myosins, kinesins or dyneins that walk on the cytoskeletal tracks, the bacterial cytoskeletal filaments involved in DNA segregation fall under the category of cytomotive filaments (Löwe and Amos 2009). The dynamic assembly and disassembly of the filaments is driven by changes in the nucleotide state. Interactions with the adaptor and/or the adaptor/centromere complex modulate these dynamics (Møller-Jensen et al. 2003; Garner et al. 2004).

The segregation systems have been classified into different types (I – III) based on the fold of the motor protein. The motor proteins of Type I, II and III comprise ATPases of the WACA family (deviant Walker A motif containing Cytoskeletal ATPases), ATPases of the actin family (Fig. 10.2), and GTPases of the tubulin family, respectively (Gerdes et al. 2010). Type I is the most common system that also functions in bacterial chromosome segregation but is the least understood. Type II systems will be described in detail in the current chapter on actin-like cytomotive filaments in plasmid segregation, while Type III will be discussed in the next chapter (Chap. 11).

Centromere

The locus on the plasmid that allows for the attachment of the filament motor via an adaptor protein can be regarded a centromere-like region of the plasmid DNA (Dam and Gerdes 1994). The centromeres of plasmid segregation systems comprise repetitive segments of DNA, termed iterons. In different systems there are 3–12 iterons in their centromeric sequences, each of 6–12 bp length (Figure 10.1). Each iteron binds a functional unit (either monomer or dimer) of the adaptor protein (Dam and Gerdes 1994; Møller-Jensen et al. 2007; Schumacher et al. 2007). The arrangement of the iterons allows for cooperative binding of the adaptor on the plasmid DNA (Møller-Jensen et al. 2003). The resulting protein-DNA complex is sometimes called the ‘segrosome’. The structure of the segrosome (Fig. 10.3) is typically organized with the DNA on the outer surface of an extended super-helical structure of the adaptor proteins (Møller-Jensen et al. 2007; Schumacher et al. 2007).

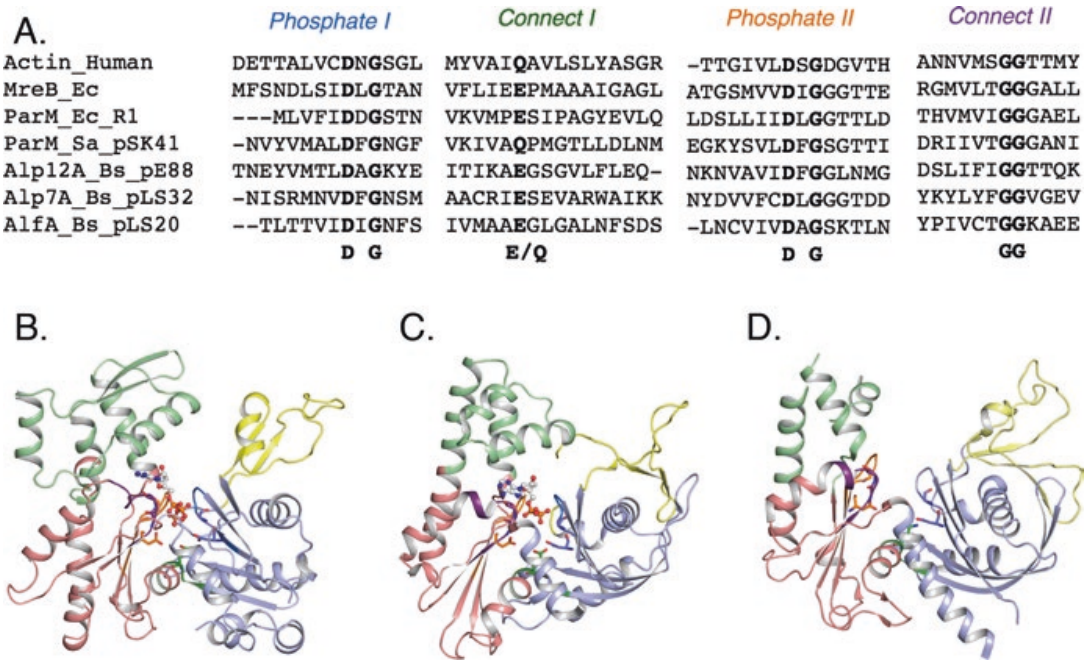

Fig. 10.2 Motor proteins in Type II plasmid segregation systems are actin-like. (**a**) Sections of the structure-based sequence alignment obtained using PROMALS3D (Pei et al. 2008) of human and bacterial actins (UNIPROT ID's- Human actin: P68133, *E. coli* MreB: P0A9X4, *E. coli* R1 ParM: P11904, *S. aureus* pSK41 ParM: O87364, *B. subtilis* pE88 Alp12A: Q89A01, *B. subtilis* pLS32 Alp7A: C7F6X5, *B. subtilis* pLS20 AlfA: S5DTP8) shows that the signature motifs of the ASKHA fold (Bork et al. 1992) are conserved in the bacterial actins involved in plasmid segregation. Monomeric structures of (**b**) actin (PDB: 1YAG), (**c**) R1 ParM (PDB: 1MWM) and (**d**) pSK41 ParM (PDB: 3JS6). The four domains are color coded as IA/1 (*blue*), IB/2 (*yellow*), IIA/3 (*red*), IIB/4 (*green*). The signature motifs are colored according to the labels in (a), and the conserved residues highlighted in stick representation. Bound nucleotides, if any, are shown in ball and stick representation

Adaptor Protein

Adaptor proteins, as the name implies, mediate the attachment of the filament motors to the centromere-like sequences on the DNA. They have a surface that interacts with the DNA, and another binding site for the motor protein (Salje and Löwe 2008). Cooperative binding of the adaptor protein to the repetitive sequences of the centromere results in the formation of the segrosome complex (Fig. 10.3). The adaptor protein, on its own and in the centromere-bound form, is capable of altering the dynamics of assembly of the cytomotive filament (see later sections) (Møller-Jensen et al. 2003; Garner et al. 2004; Gayathri et al. 2012; Garner 2007). This is a significant feature of plasmid segregation systems, which helps with dispensing of alternative nucleation factors for filament assembly, checkpoints for spindle assembly, and disassembly factors. In many if not all cases, the centromeric sequence overlaps with the promoter-binding region for the motor. Hence, the adaptor protein also acts as a transcriptional repressor for the motor and adaptor proteins (Dam and Gerdes 1994). The repression activity regulates the number of monomers

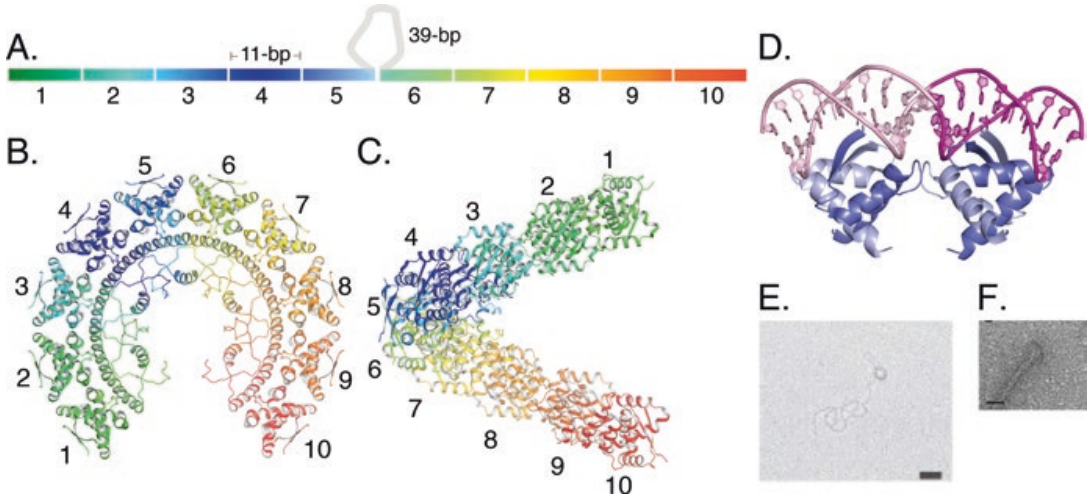

Fig. 10.3 Architecture of the segrosome complex in plasmid segregation. (**a**) Organization of the centromeric sequence into 10 iterons. The 39-bp insertion between the iterons is shown as a grey loop. (**B–C**) pB171 ParR (PDB ID: 2JD3) forms a super-helical structure resembling a lockwasher. A dimer of ParR binds to each iteron of *parC* (color-coded and numbered correspondingly) (**b**) View of ParR helix along the screw axis and (**c**) side view. (**d**) ParR binds DNA as a dimer to one iteron of the centromeric sequence, as shown by the crystal structure of the N-terminal region of dimeric ParR (PDB ID: 2Q2K) of pSK41 plasmid (*light* and *dark blue*) in complex with a 20-bp DNA sequence of two iterons of *parC* (magenta and pink). (**e**) Electron micrographs of R1 ParR assembled on *parC*-containing plasmid (Adapted from Møller-Jensen et al. (2007)). (**f**) Negatively stained EM image of R1 ParR bound to the tip of a ParM filament (Adapted from Salje and Löwe (2008))

of the motor protein available in the cell, and thus controls filament dynamics. This leads to an efficient mechanism of disassembly and/or control over excessive filament formation and bundling.

Actin-Like Cytomotive Filaments in Plasmid Segregation

A large number of actin-like proteins have been identified in bacterial genomes (Derman et al. 2012). Some of these proteins have been shown to be involved in plasmid partitioning while the functions of others remain elusive. The best-characterized plasmid segregation system involving an actin-like cytomotive filament is the ParMR*C* system (Salje et al. 2010). Other actin-like plasmid segregation systems are AlfA, Alp7A, Alp12, pB171 and pSK41. The following sections discuss in detail how the structure and dynamics of the cytomotive filaments lead to plasmid partitioning as, currently, the function of the ParMRC system is probably the best-understood in bacterial systems involving cytomotive filaments.

ParMRC System of E. coli *R1 Plasmid*

A number of studies utilizing cytological investigations (Gerdes and Molin 1986; Dam and Gerdes 1994), in vitro biochemical assays (Møller-Jensen et al. 2002; Jensen and Gerdes 1997; Møller-Jensen et al. 2003, 2007; Schumacher et al. 2007; Salje and Löwe 2008), reconstitution experiments in vitro (Garner et al. 2004; Garner 2007), live cell fluorescence microscopy (Campbell and Mullins 2007), X-ray crystallography (Gayathri et al. 2012; van den Ent et al. 2002), electron microscopy and tomography experiments (Salje et al. 2009; Orlova et al. 2007; Popp et al. 2008; Bharat et al. 2015) have led to deep mechanistic insights into the action of the ParMR*C* system of plasmid segregation.

Structural Features of ParMR*C*

The structural organization of the three components of the plasmid segregation machinery has been elucidated for ParMR*C*.

Architecture of the ParR*C* Complex

The *parC* locus of the ParMR*C* system consists of 10 iterons of 11-base pair sequences of DNA, five each on either side of a 39-base pair region (Fig. 10.3a) (Dam and Gerdes 1994). The structure of ParR demonstrates that it forms a lockwasher-like super-helical structure at high concentrations and in the presence of DNA containing the *parC* sequence (Fig. 10.3b–d) (Møller-Jensen et al. 2007). The positively charged surface on the outside of the lockwasher provides a complementary electrostatic surface for the negatively charged DNA backbone, as well as a repeating structural match to the iterons; each dimer of ParR binds to an iteron comprising 11-base pairs of *parC* (Møller-Jensen et al. 2007).

Architecture of ParM Filament

In the presence of ATP, ParM forms helical filaments with a left-handed twist (Gayathri et al. 2012; Orlova et al. 2007; Popp et al. 2008). These can also be described as a two-start helix with two protofilaments/strands (Fig. 10.4a). The arrangement of the monomers within the protofilament of ParM is similar to that in actin, though the helical parameters and the twist are different. The rotation between adjacent subunits in the protofilament is 27° for actin (right-handed), and −30° (left-handed) for ParM (Gayathri et al. 2012; Fujii et al. 2010). Table 10.1 contains a summary of the helical parameters of all the actin-like filaments in plasmid segregation for which a reliable reconstruction is available. The polar structure (filament

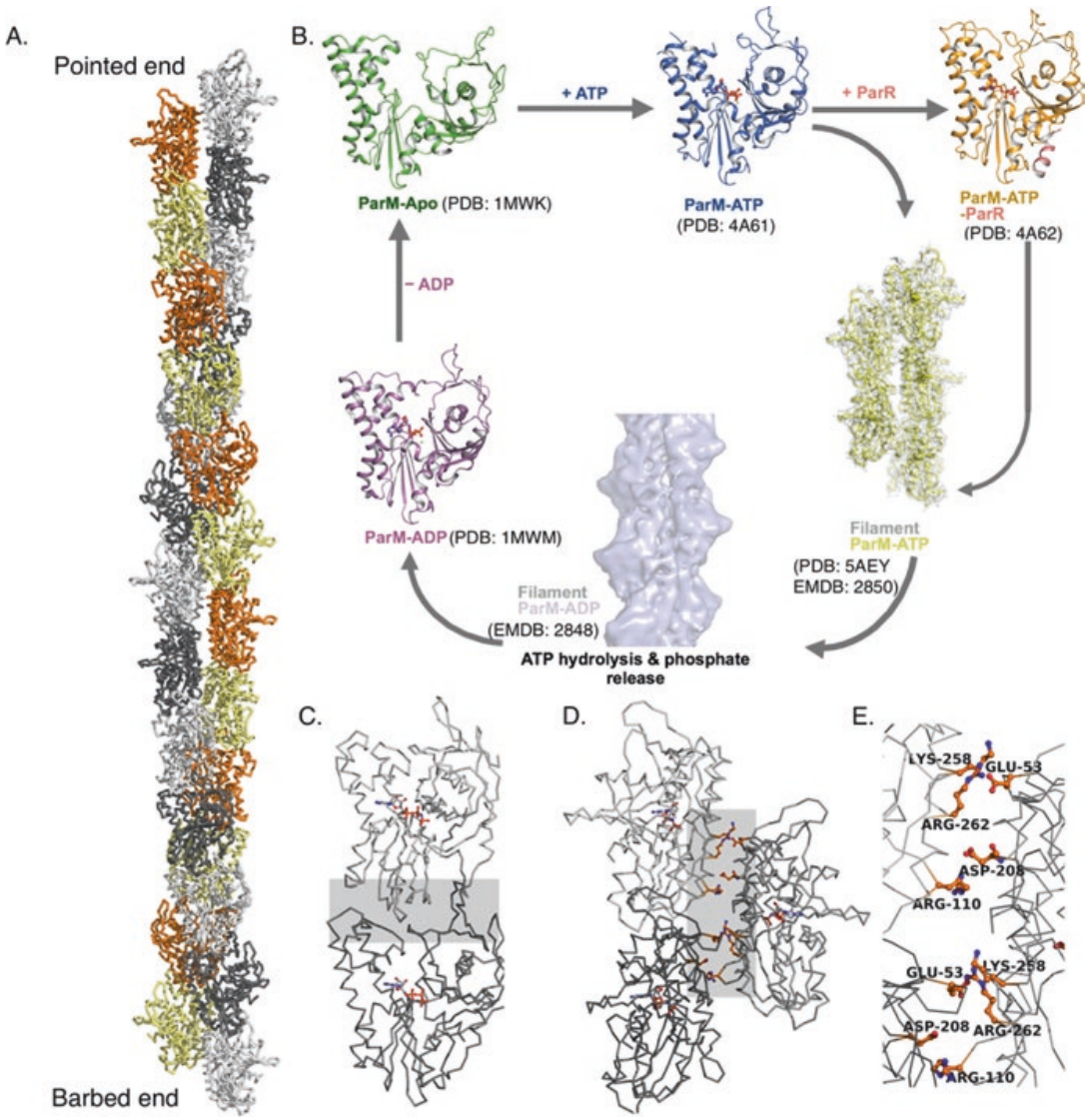

Fig. 10.4 Conformational cycle of ParM. (**a**) The ParM filament is composed of two protofilaments (PDB ID: 4A6J). Alternate monomers are colored as grey and black in one protofilament/strand and yellow and orange in the other. Monomers in one protofilament are offset by approximately half their length (~ 24 Å) with respect to the other protofilament – the filament is staggered. The filament is polar, with one end termed as the barbed end and the other as the pointed end. The standard view of the monomers, shown in Fig. 10.2, and in (**b**), is such that the bottom of the monomer faces the barbed end of the filament. (**b**) The nucleotide-dependent conformational cycle of ParM: ParM in its apo-form binds ATP resulting in a ~25° domain closure between domains 1 and 2. The conformational change in ParM-ATP can either allow its assembly to form filaments or it binds with ParR to forms a ParM-ATP-ParR complex. The ParR-bound conformation of ParM and the monomer conformation within the ParM-ATP filament are equivalent – a more compact conformation where ATP hydrolysis is stimulated. Upon release of phosphate, the filaments dissociate to form ParM-ADP monomers, which can release ADP to go back to the apo-form. The ATP-bound conformations have been captured using AMPPNP, a non-hydrolysable analogue of ATP. (**c**) Intra-protofilament contacts between R1 ParM monomers with the polymerization interface highlighted in *grey*. (**d**) Inter-protofilament contacts between monomers from adjacent protofilaments/strands. (**e**) Enlarged view of the contacts in (**d**) shows that the inter-protofilament contacts include a few salt bridges (PDB ID for C-E: 5AEY)

Table 10.1 Helical parameters and Dynamics of actin-like filaments

Actin-Like Protein	Structure (helical parameters)						Dynamics		
	PDB	EMDB	Resolution (Å)	Handedness	Phi (°)[a]	Azimuthal increment (Å)	Instability ratio (ccADP/ccATP)[b]	Behavior	Nucleus
Actin	3MFP	5168	6.6	Right	26.8	27.6	1.6 (Pollard 1986)	Treadmilling	3 (Pollard 1986)
R1 ParM	5AEY	2850 (AMPPNP)	4.3	Left	30	23.4	>160	Dynamic instability	3 (Garner et al. 2004)
	–	2848 (ADP)	11	Left	30	24.7			
pB171 ParM (Rivera et al. 2011)	–	–	19	Left	27.8	24.2	>140	Dynamic instability	2
AlfA (Polka et al. 2009)	–	–	15	Left	46	25	4.2	Treadmilling	3–4
Alp12A (Popp et al. 2012)	4APW	2068	19.7	Right	13.24	42.8	No polymerization with ADP	Dynamic instability	4

[a]Phi denotes the angle of rotation between adjacent subunits within a protofilament
[b]ccADP/ccATP denotes ratio of the critical concentrations of ADP to ATP

with dissimilar ends) of the ParM filament is described by the terms barbed end and pointed end, borrowing from the actin filament terminology (Fig. 10.4a).

Conformational Cycle of ParM

The different conformations of monomeric ParM in the nucleotide-free and in the bound states with the hydrolyzed (ADP) and non-hydrolyzed (AMPPNP, a non-hydrolysable analog of ATP) nucleotide states have been captured using X-ray crystallography (Gayathri et al. 2012; van den Ent et al. 2002). The conformations of the filament form of ParM in the different nucleotide states have been observed recently by high-resolution electron microscopic reconstructions (Bharat et al. 2015). These give a seemingly complete picture of the conformational cycle that ParM undergoes during nucleotide hydrolysis-dependent dynamics (Fig. 10.4b).

Upon nucleotide binding, there is ~25° domain closure between domains 1 and 2 (van den Ent et al. 2002). Comparison between ATP and ADP monomeric conformations does not reveal any prominent conformational changes. However, the assembly of the ATP-bound monomers into the filament form results in a further domain closure resulting in the formation of the most compact structure of ParM (Bharat et al. 2015). The conformation of ParM in the ParR peptide-bound state matches well the conformation of ParM monomers within the filament (Gayathri et al. 2012). This structure also shows that it binds to a pocket at the polymerization interface (Fig. 10.4b, c). Hence, ParR can be accommodated only at the barbed end of the ParM filament.

ParM Filament Dynamics

ParM filaments are preferentially formed in the presence of ATP. The critical concentration for filament formation is very high in the presence of ADP, and even higher without any nucleotide (Garner et al. 2004). Studies have shown that ParM can also utilize GTP for polymerization (Popp et al. 2008). Filament growth does not require external nucleation factors, and filaments are capable of nucleating on their own, above the critical concentration. Though the filaments are structurally asymmetric, they are kinetically symmetric, and grow, alone, equally at the barbed end and the pointed end (Garner et al. 2004). However, in the presence of ParR*C* complex, the critical concentration of ParM monomers for filament formation is lower, and the growth at the barbed end, which is the ParR*C*-bound end, is faster than at the pointed end (Fig. 10.5a) (Gayathri et al. 2012).

ParM filaments exhibit dynamic instability (Garner et al. 2004). They grow in the presence of ATP-bound monomers. Hydrolysis of ATP destabilizes the filaments, resulting in disassembly. Binding of ParR*C* prevents the disassembly of the filaments from the barbed end (Gayathri et al. 2012). ParM monomers are added by insertional polymerization at the ParR*C*-bound end (Møller-Jensen et al. 2003; Garner 2007). The pointed end is inhibited from disassembly by lateral stabilization

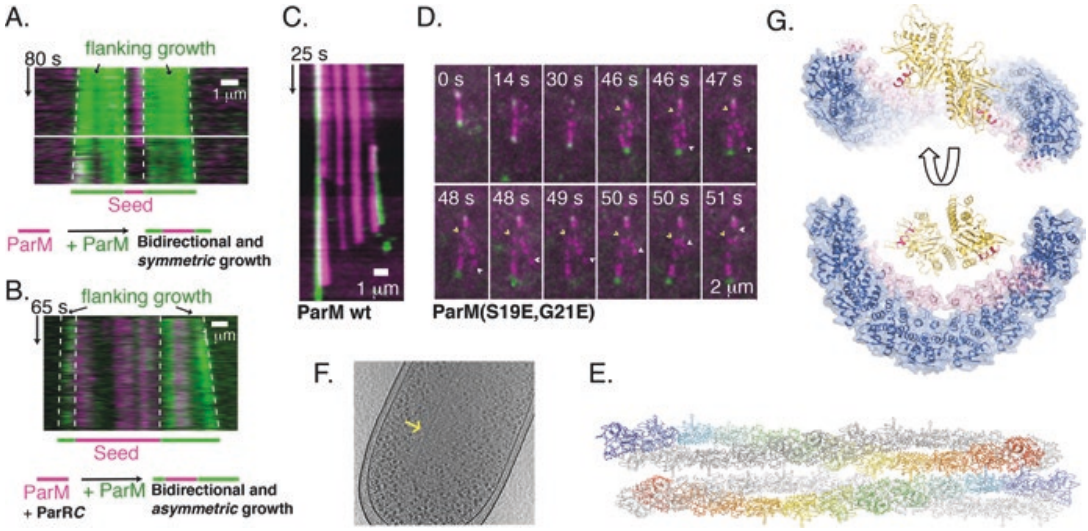

Fig. 10.5 Effect of ParR*C* on ParM filament assembly and bipolar spindle formation. Kymographs of ParM filaments by dual-label experiments (*purple* ParM seed and *green* growing filament) in the absence (**a**) and presence (**b**) of ParR*C* using TIRF microscopy. ParM filaments, in absence of ParR*C*, exhibit symmetric growth at both ends. The bidirectional growth becomes asymmetric upon addition of ParR*C*, which rapidly adds ParM monomers at one end, the barbed end. C & D) ParM filaments are stabilized by the binding of ParR*C* at the barbed end (**c**) and by pairing at the pointed end (**d**). Loss of ParR*C* (*green*) at the barbed end (**c**) leads to disassembly of the ParM (*purple*) filament. The white arrow highlights the initiation of disassembly, observed by the slope. The dissociation of a spindle (**d**), due to repulsive charges in a mutant of ParM (ParM-S19R, G21R), shows that a ParM spindle (*purple*) is composed of more than one filament, that they are oriented anti-parallel, and that un-pairing triggers disassembly from the pointed end. White and yellow arrowheads highlight the ends of each filament within the disassembling spindle. (**e**) Atomic model comprising one repeat of the antiparallel ParM filament pair. (**f**) A view of the bipolar spindle captured in a electron cryotomogram of an *E. coli* cell. The filament doublet shows a clear lack of superhelicity, consistent with the model in (**e**). (**g**) The overall dimensions of the ParR*C* complex and the helical pitch of ParM filaments are similar. Thus ParR*C* can facilitate nucleation and accelerate ParM growth at the barbed end. (**A–G** adapted from Gayathri et al. (2012) and Bharat et al. (2015))

upon pairing with another ParR*C*-bound filament in an antiparallel orientation (Fig. 10.5c–e) (Gayathri et al. 2012). Thus, an antiparallel arrangement of two filaments forms a stable bipolar spindle of ParMR*C* (Fig. 10.5d–e). An interesting feature of ParMR*C* spindle dynamics is that ParR*C* interacts only with the ATP-bound conformation of ParM (Møller-Jensen et al. 2003). This ascertains that ParR*C* remains bound only when there are sufficient free ParM monomers within the cell, where they will normally be ATP-bound. Decrease in the concentration of ATP-bound ParM monomers leads to the loss of ParR*C* from the tip, and subsequent disassembly of the bipolar spindle (Møller-Jensen et al. 2002; Gayathri et al. 2012).

Information on filament dynamics has been obtained by in vitro reconstitution experiments using fluorescently labeled proteins and DNA containing *parC* (Garner et al. 2004; Gayathri et al. 2012; Garner 2007). Light scattering experiments provide information on the kinetics of filament nucleation, growth and disassembly (Table 10.1) (Garner et al. 2004; Popp et al. 2008; Galkin et al. 2009). Plasmid

segregation dynamics have also been observed using live cell imaging within *E. coli* using fluorescently labeled components (Campbell and Mullins 2007). A detailed account of the various methods utilized to study actin-like filaments is provided in reference (Petek and Mullins 2014).

Correlation Between ParM Structure and Dynamics

Observation of the complete conformational cycle provides a clear understanding of how structural features of the actin-like filament contribute to dynamics of assembly (Fig. 10.4b). ParM monomers assemble into the filament form in the presence of ATP. The conformational transition from the monomeric to the filament form stimulates ATP hydrolysis within the filament (Møller-Jensen et al. 2003; Gayathri et al. 2013). The monomers are held in the filament conformation through longitudinal lattice contacts involving adjacent monomers even post ATP hydrolysis, as is observed from the ParM filament reconstruction in presence of ATP and vanadate (Bharat et al. 2015). However, ParM filaments formed in the presence of ADP have a different conformation of monomers within the filament. Comparison of the ParM-ADP and ATP filament conformations shows that the inter-protofilament and intra-protofilament contacts are not optimal for the stability of the ADP-bound filament. Figure 10.4c–e shows the contacts that hold the filament in the AMPPNP-bound state of ParM (Bharat et al. 2015). The sub-optimal contacts lead to the disassembly of the filaments post ATP hydrolysis and loss of phosphate, unless the lattice contacts with ATP-bound ParM monomers is maintained at the tips, where ATP hydrolysis may not have occurred yet.

Dynamic instability exhibited by the ParM filaments implies that they remain stable as long as the ends of the filament are held in the ATP-bound filament conformation. If the rate of addition of monomers slows down, presumably due to reduction in the concentration of ATP-bound ParM monomers, the conversion of ATP to ADP conformation occurs in the ParM monomers towards the ends of the filament, leading to rapid disassembly. The observation of disassembly in the ParM spindles show that the process is directional (Fig. 10.5c, d) and proceeds from the destabilized end of the filament (barbed end where the destabilization is by the loss of ParR*C* (Fig. 10.5c), pointed end if the destabilization is by decoupling of a paired filament) (Fig. 10.5d) (Gayathri et al. 2012).

The structural arrangement of the ParR*C* helix might explain its processivity when polymerizing ParM filaments and also the insertional polymerization mechanism of ParR*C*-driven ParM filament elongation (Fig. 10.5g). The binding site at the polymerization interface always ensures that the ParR*C* helix is at the tip of the filament, thus providing a rationale for insertional polymerization (Gayathri et al. 2012). ParR*C* facilitates the conversion to the filament conformation and prevents disassembly by speeding up the rate of addition of ATP-bound monomers. The multimerization involved due to the repetitive segments of the ParR binding sites at the iterons of *parC* provides multiple ParM-binding sites for the segrosome complex. This probably contributes towards the processivity of movement of ParR*C* at the

barbed end of ParM filaments. The amount of ATP-bound ParM monomers in the cell may also affect processivity. Finally, the architecture of the ParR*C* complex suggests that nucleation and acceleration of growth at the barbed end may occur due to similarity in the helical pitch of the ParM filament and that of the ParR*C* superhelix (Fig. 10.5g).

Antiparallel bipolar spindle assembly is facilitated by ParM's filament architecture. The groove of one filament fits into the groove of another in an antiparallel manner (Fig. 10.5e). Hence, surface residues on the filament contribute to the efficient functioning of the bipolar spindle, such that the interaction is not too strong to hinder the dynamics by antiparallel pairing, but not too weak to prevent them from forming a stable spindle, as has been demonstrated by mutational studies (Gayathri et al. 2012). These bipolar spindles have been observed in vitro (Fig. 10.5c, d) and also in vivo using electron tomography of *E. coli* cells (Fig. 10.5f) (Bharat et al. 2015). The electron tomography observations in vivo point out that the plasmid segregation is most likely 'asynchronous' – each bipolar spindle segregates a different pair of plasmid sisters (Bharat et al. 2015). The process of segregation is also not coordinated with the cell cycle of the bacterium. These facts point out that the mechanism of plasmid segregation in ParMR*C* strives to separate pairs of plasmid copies, and thus ensures that they are apart during cell division.

AlfA of Bacillus Subtilis pLS32

A partitioning system, active during vegetative growth and sporulation in *B. subtilis* was discovered on pBET31, the mini-plasmid of pLS32, and AlfA and AlfB were identified to be the actin-like protein and the adaptor protein components respectively (Becker et al. 2006).

Organization of the AlfA Segregation System

The *alf* operon consists of AlfA, forming an *a*ctin-*l*ike *f*ilament (*Alf*), AlfB, a DNA-binding protein that acts as the adapter coupling the filament to the plasmid, and *parN*, a DNA locus containing three binding sites for AlfB (Becker et al. 2006; Polka et al. 2009; Tanaka 2010). AlfA was found to be actin-like based on the identification of nucleotide binding and hydrolysis motifs (Fig. 10.2a) (Becker et al. 2006). The interaction between AlfB and *parN* has been characterized and the first two iterons are essential for the binding (Tanaka 2010). Since the *parN* sequence is located within the promoter region of *alfA*, binding of AlfB to *parN* represses its expression. It has also been shown using yeast two hybrid experiments that AlfB binds to AlfA. Thus, the three components AlfA, AlfB and *parN* constitute a plasmid segregation system for pLS32 of *Bacillus subtilis* (Tanaka 2010).

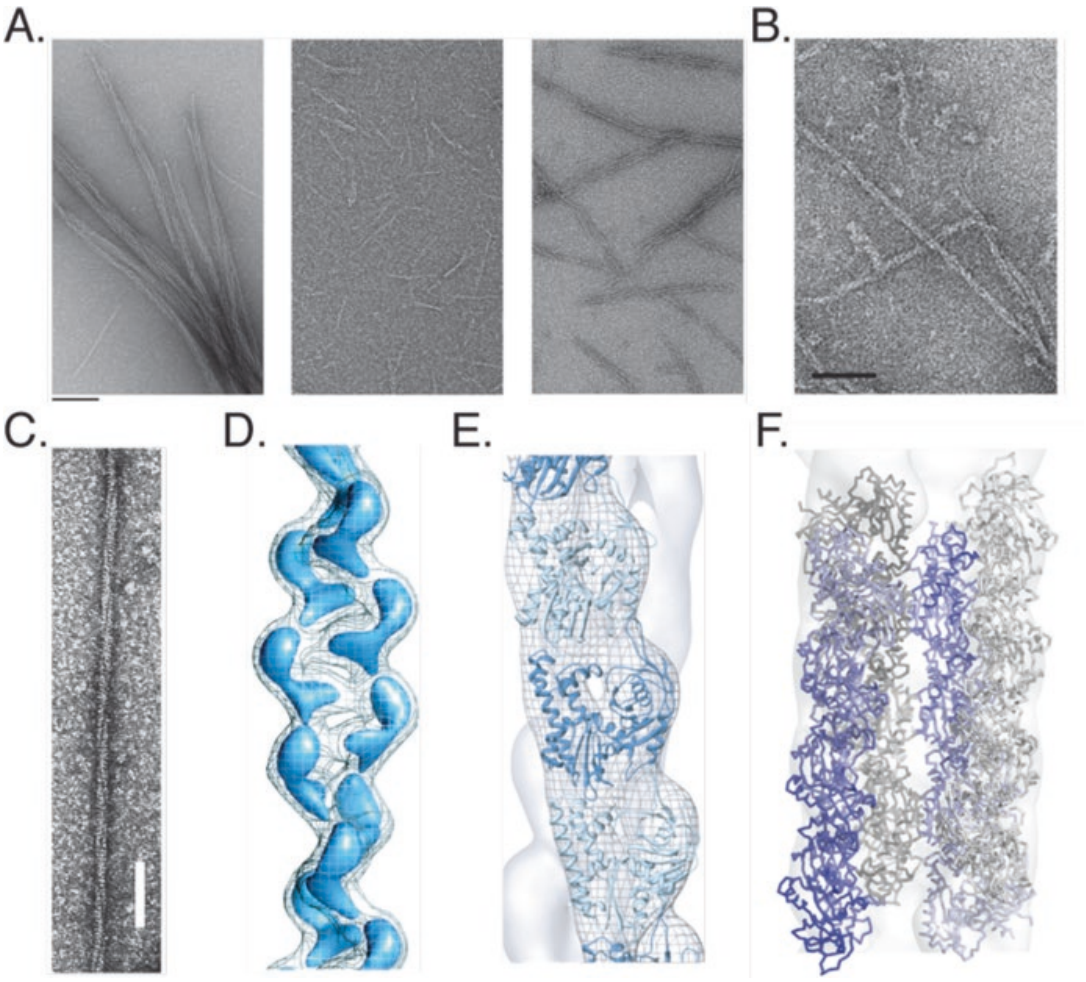

Fig. 10.6 Electron microscopy of actin-like proteins from Type II plasmid segregation systems. (**a**) AlfA alone forms bundles (*left*), the presence of AlfB debundles AlfA and decreases its length (*centre*), and bundles are restored upon addition of AlfB-parN complex (*right*) (Adapted from Becker et al. (2006)). (**b**) pB171 ParM forms filaments in presence of ATP (Adapted from Rivera et al. (2011)). (**c**) Filaments of Alp12A (Adapted from Popp et al. (2012)). (**d–f**) EM reconstruction of AlfA filaments (**d**, adapted from Polka et al. (2009)); pB171 ParM filament (e, adapted from Rivera et al. (2011)); Alp12A filaments (f, PDB: 4APW; EMDB: 2068). Alp12A exists as a four-stranded antiparallel filament. Each filament is shown as two protofilaments colored in *blue* and *grey* respectively. The shade of blue of the monomers varies from dark towards light from the barbed end to the pointed end, highlighting the antiparallel arrangement

Structure of AlfA

Though the monomeric structure of AlfA has not been determined yet, it is expected to be of actin fold. Contrary to the ParM filaments that exhibit dynamic instability, AlfA was observed to form stable filaments within the cell and also in vitro (Polka et al. 2009). The polymerization occurred in the presence of ADP or ATP, and also GDP or GTP, although filaments formed with ATP were more stable compared to those with GTP. Electron microscopy images showed the presence of bundles of AlfA filaments (Fig. 10.6a). The bundling of AlfA was dependent on the salt

concentration. The bundles were disrupted in the presence of higher concentration of KCl (Polka et al. 2009).

Three-dimensional reconstruction of electron microscopy images of negatively-stained AlfA showed that the filaments are left-handed, similar to ParM (Fig. 10.6). The helical parameters vary between AlfA and ParM. AlfA exhibits a higher twist of 46° between adjacent monomers in the filament compared to 30° in ParM (Table 10.1). A monomer of ParM in the apo-state (with an open nucleotide cleft) was observed to be the best fit into the EM reconstruction map (Polka et al. 2009). However, a high-resolution EM reconstruction from electron cryomicroscopy images will be required in order to provide more detailed information.

AlfA Filament Dynamics

AlfA forms polymers in the presence of ATP, ADP, GTP or GDP (Polka et al. 2009). The kinetics of polymerization show that the critical concentration of ADP-bound AlfA monomers is only about 4 fold higher than that of ATP-bound monomers. This implies that AlfA most likely does not function by a dynamic instability driven mechanism as observed for ParM. The ability to form filaments with ADP or ATP results in the formation of stable bundles of AlfA filaments (Polka et al. 2009). Initial observations of stable filaments of AlfA that do not show dynamics were intriguing since a mechanism for plasmid segregation involving a stable filament is difficult to reconcile. For plasmid segregation to occur by a filament bound to a DNA there should be available a pool of monomers, maintained above the critical concentration for elongation. The pool is typically maintained by recycling monomers released by the breakdown of filaments that are not bound to DNA, as in the case of dynamic instability of unbound ParM filaments.

Later studies showed that AlfB reduces the bundle formation of AlfA polymers and destabilizes them (Fig. 10.6a). This effect was attributed to the sequestration of AlfA monomers by AlfB (Polka et al. 2014). Additionally, AlfB bound to *parN* was found to nucleate AlfA filaments. The filaments formed in the presence of the AlfB-*parN* complex were longer, and had a tendency to bundle (Fig. 10.6a). The AlfB-*parN* complex bound only to one end of the AlfA filaments (Polka et al. 2014). In vitro experiments showed that DNA moves processively along the growing tips of AlfA filaments, while disassembly occurs at the other end, giving the appearance of a comet-like structure (Polka et al. 2014). Thus, the combined effect of filament stabilization by the AlfB-*parN* complex at one end, and the destabilizing effect of free AlfB at the other end results in treadmilling of the AlfA filament. This treadmilling is unidirectional, and pushes the DNA in the direction of growth of the filament. Bundling of two filaments in antiparallel orientations results in movement of DNA in opposite directions. Reference-free averaging of images of AlfA filament pairs suggests that the filaments do show a preference for bundling in an antiparallel orientation, similar to that of ParM filaments (Polka et al. 2014).

Further studies of the structural details of how AlfB interacts with AlfA will be needed to understand the mechanism of action of AlfB as a destabilizing agent on its own, and as a nucleator and a promoter of filament elongation in the presence of *parN* DNA.

Other Actin-Like Filaments Involved in Plasmid Segregation

Many actin-like proteins (Alps) have been identified in bacteria (Derman et al. 2012). Some of these are found in extra-chromosomal elements such as plasmids, phage and conjugative DNA elements. The functions of most of these proteins are unknown. However, representative members of some of the actin-like protein families that have been characterized demonstrate that they form filaments in vitro and in vivo and have been implicated in plasmid segregation (Derman et al. 2012; Rivera et al. 2011; Popp et al. 2012). Other than ParM and AlfA, for which mechanistic details of plasmid segregation are emerging, the rest of the actin-like cytomotive filaments have only been partially characterized. A summary of the information on these systems is given in the following sections (Fig. 10.6 and Table 10.1).

Alp12A

Alp12A forms part of the plasmid segregation system of *Clostridium tetani* pE88 plasmid (Popp et al. 2012). In addition to Alp12A, some of the other genes harbored in pE88 are the tetanus toxin and the transcriptional repressor TetR. The DNA sequence equivalent to *parC* has not been identified yet.

Structure of Alp12A

Alp12A forms filaments in the presence of ATP or GTP. Interestingly, electron microscopy of Alp12A filaments formed in presence of ATP shows that they comprise four actin-like protofilaments or strands. Out of the four protofilaments, two of the protofilaments are arranged antiparallel in orientation to the other two. The arrangements of pairs of protofilaments that are oriented parallel are not similar to the inter-protofilament arrangement observed in other known actin-like filaments (Fig. 10.6c). A prominent cleft is found in between the two antiparallel pairs of filaments, which was observed as a dark line in the negatively stained images. The handedness has been estimated to be right-handed. It is interesting to note that there is an antiparallel arrangement of protofilaments, which will probably contribute towards moving DNA in opposite directions during plasmid segregation (Popp et al. 2012).

Alp12A Filament Dynamics

The kinetics of polymerization for Alp12A have been studied using light scattering experiments and kinetic analysis of the data. Accordingly, Alp12A exhibits dynamic instability. It utilizes either ATP or GTP for polymerization. However, similar to other actin-like proteins discussed so far, ATP-dependent polymerization is more efficient. Fitting of experimental light scattering data suggests that the nucleus for polymerization involves an initial dimer formation, followed by tetramerization (Table 10.1). This model for nucleation is consistent with the four-stranded antiparallel filament model (Popp et al. 2012).

Alp7A

Alp7A is an actin-like protein from the *Bacillus subtilis* plasmid pLS20. The dynamics of Alp7A and the associated Alp7R and *alp7C* have been observed in vivo (Derman et al. 2012). GFP labeling of Alp7A expressed in cells showed that these filaments grow and shrink, thus exhibiting dynamic instability. Mutation of the aspartate most likely involved in ATP hydrolysis resulted in the presence of long and stable filaments within the cell, and disrupted the dynamics, emphasizing the role of ATP hydrolysis in the process. Similar to ParR, Alp7R acts as a repressor for Alp7A expression and binds to a repetitive sequence of DNA, *alp7C*, positioned at the promoter region of Alp7A.

pSK41

pSK41 is a plasmid that confers multiple drug resistance to *Staphylococcus aureus*. ParM and ParR proteins and *parC* iterons have been identified for the pSK41 plasmid segregation system. Crystal structures of the pSK41 ParM (Popp et al. 2010), and a DNA-bound segrosome complex of the ParR (Schumacher et al. 2007) have been determined.

Structure and Dynamics of pSK41 ParM

The structure of pSK41 ParM shows that it is indeed an actin-like protein (Fig. 10.2d). The protein in its apo-state superposes best with the nucleotide-bound conformation of R1 ParM. The filaments of pSK41 have been observed by negative stain EM, although currently no near-atomic structure or 3D-reconstruction of the filaments is available. A mutational analysis has identified the probable residues involved at the interface, and this study highlights that the orientation of monomers in the pSK41 ParM will be similar to that of the monomers in other actin-like ParM protofilaments (Popp et al. 2010).

Light scattering studies to understand the dynamics have shown that pSK41 ParM polymerizes in the presence of ATP or GTP as nucleotide, and filaments were not observed with ADP or GDP. Hence, they exhibit dynamic instability similar to R1 ParM (Popp et al. 2010).

Structure of pSK41 ParR

Crystal structure of pSK41 ParR shows that the protein adopts a ribbon-helix-helix fold, and a dimer of ParR interacts with a 20-mer tandem repeat of 10 bp of parC DNA. It forms a super helical structure with the DNA wound around the outer surface of the helix (Fig. 10.3b–d). The electrostatic surface potential on the inside surface of the super helical structure is complementary to that of the pSK41 ParM monomer (Schumacher et al. 2007).

pB171 ParM

ParM from *E. coli* virulence plasmid pB171 shares a comparatively high degree of conservation with ParM from R1 plasmid, Though 41% identity between them is much less than the degree of conservation between eukaryotic actins, the structure and dynamics of pB171 ParM seem to be quite similar to that of R1 ParM. Fig. 10.6 and Table 10.1 summarize relevant features of pB171 ParM in comparison to the other members of the actin-like filaments in plasmid segregation (Rivera et al. 2011).

General Features of Actin-Like Plasmid Machineries

The sets of actin-like plasmid segregation machinery characterized till date permit us to compare with the eukaryotic actin cytoskeleton and associated proteins and highlight some of the features of the bacterial cytoskeleton required for the function of DNA segregation. The major significance emerging from studying these systems lies in their simplicity, in how the cytoskeletal systems are empowered to carry out DNA segregation faithfully with minimal numbers of components. Of course, this is not accidental in the case of extra-chromosomal genetic elements such as plasmids, mobile genetic elements, and phages because of their extremely limited genome size.

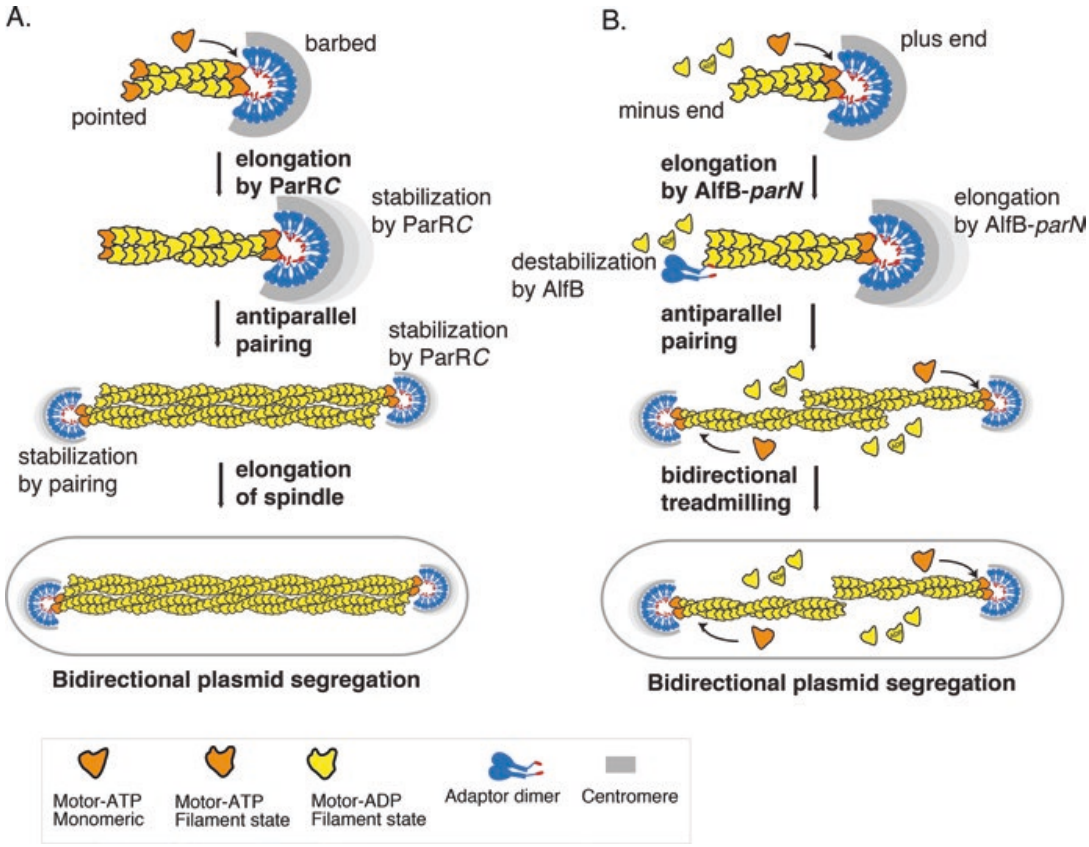

Fig. 10.7 Schematic representation of the mechanism of plasmid segregation in (**a**) R1 ParM, and (**b**), AlfA (Adapted from Gayathri et al. (2012) and Polka et al. (2014))

Possible Ways to Build a Bipolar Spindle

The actin-like cytomotive filaments functioning as DNA segregation machinery highlight that there are multiple ways in which a bipolar spindle can be constructed, using a basic building block or monomer of the same protein fold and the same protofilament architecture. Dynamics driven by nucleotide-hydrolysis play a major role in the formation of the bipolar spindle. The two major patterns of filament dynamics, as also observed in the eukaryotic cytoskeletal systems, are dynamic instability and treadmilling. How the plasmid segregation system manipulates either of these dynamics to achieve segregation is very interesting (Fig. 10.7).

A filament that exhibits dynamic instability requires both ends to be stabilized, else it will not survive. In the ParMR*C* system, the two ends of the filaments are rescued from dynamic instability by different means – (i) when ParR-bound DNA binds to one of the ends (ie, barbed end) (ii) when paired with another filament coming together in an antiparallel direction (at the pointed end). This arrangement ensures two requirements for DNA segregation – (i) capture of the DNA is achieved (ii) antiparallel pairing of two DNA-attached filaments facilitates coordinated movement of a pair towards the extremes of the cell, thus acting as a 'checkpoint'

for DNA replication. The antiparallel orientation and elongation in the DNA-bound end ensures separation of the attached pair.

In treadmilling, elongation occurs at one end, while the other end shortens. If DNA is attached to the growing end, as in AlfA-driven plasmid segregation, the system functions via a pushing mechanism and pairs separate when two such filaments come together in an antiparallel orientation. However, the antiparallel arrangement does not contribute to increase the stability of the paired end. In the case of AlfA, the dynamics at the growing end (plus end) and the shortening end (minus end) are both modulated by AlfB. *parN*-bound AlfB at the growing end elongates the filament, while free AlfB destabilizes the minus end and ensures recycling of monomers for the growing end to elongate.

Both of the cases of actin-based plasmid segregation mechanisms mentioned above work by a pushing mechanism at filament plus ends. An interesting case of plasmid attaching to the minus end of a treadmilling filament, implicating a pulling mechanism, has been observed in TubZR*C*, a tubulin-like cytomotive plasmid segregation system (Fink and Löwe 2015) (discussed in the next chapter). Many of the actin-like filaments appear to prefer an antiparallel arrangement of protofilaments, which is a desired or an intuitive configuration, for a pushing mechanism for the separation of two plasmids. Characterization of the interaction with the adaptor protein and the segrosome complex for more actin-like plasmid segregation systems will throw light on the different combinations of possible modulations of the filament dynamics, required for building a plasmid segregation system.

Design of a Minimal Bipolar Spindle

One of the endearing features of plasmid segregation systems is the minimal number of components required to carry out their functions. This is possible by exploiting the principles of filament dynamics for dispensing with many of the components such as nucleators, and disassembly factors needed in eukaryotic cytoskeletal systems. The insertional polymerization mechanism ensures that DNA is always bound to one end only and not to the sides of the filament, thus accomplishing segregation. Additionally, the amounts of proteins involved in plasmid segregation, and thus filament dynamics, is maintained by strategies such as transcriptional repression by the adaptor protein.

In the ParMR*C* system, the number of ParM molecules available is regulated by the transcriptional repression by ParR. Hence, when ParR is bound to DNA and initiates the formation of a spindle, production of more ParM monomers is shut down. The existing ParM monomers get incorporated into the stabilized spindle and this further reduces the amount of free monomers present in the cell. Processivity depends on the amount of ATP-bound ParM. When the level of ATP-bound ParM goes low, the spindle disassembles automatically due to the deficiency of ATP-bound ParM. This is facilitated by the design of an adaptor protein that binds only

to the ATP-bound ParM, and thus dispenses with the requirement of external disassembly factors.

Thus, the control of filament dynamics serves multiple functions: -(i) It facilitates the capture of DNA by nucleating filament formation, a role played by the segrosome complex. (ii) It dispenses with the requirement of additional factors for disassembly once the plasmids are moved apart within the cell. (iii) The requirement of two antiparallel filaments provides a measure for the presence of an even number of plasmids, as is the case after DNA replication.

Comparison with the Actin Cytoskeleton

The actin-like proteins in plasmid segregation and indeed all actin-like proteins of the bacterial cytoskeleton show a high degree of sequence divergence amongst themselves and with respect to the actin sequence. The sequence identities between these and actins are typically less than 20%. The sequence variation allows for the varying architectures of protofilaments/strands and the resulting filaments. Importantly, the structure of the intra-protofilament/strand interface is conserved throughout the family, including MreB (van den Ent et al. 2001), FtsA (Szwedziak et al. 2012), MamK (Ozyamak et al. 2013), ParM (Bharat et al. 2015), AlfA(Polka et al. 2009), Alp12A (Popp et al. 2012). Though a similar arrangement of the monomers is maintained within the protofilaments, the residues at the interface show variation, preventing the possibility of copolymerization among these filaments and also reflecting large evolutionary distances. The organization of the protofilaments gives rise to the different types of filament architectures. Each plasmid might have evolved and optimized its own plasmid partitioning system, since variability ensures that the machinery segregates only its own kind. The actin cytoskeleton in eukaryotes is highly conserved probably due to the actin pool required for a large number of functions within the cell, and the cell has to maintain a required pool of actin monomers for all these functions (Gunning et al. 2015). The number of protein interactions with actin is also very high, which results in a high degree of conservation. In contrast, plasmid segregation systems are autonomous, not interacting functionally with any cellular components and this means that the evolutionary pressure on the components involved is very low, as long as function is conserved.

An interesting feature that highlights a conserved feature of the actin fold in actin-like plasmid segregation systems is the region of interaction between ParM and ParR. ParR interacts with ParM through an amphipathic helix on the C-terminal end of ParR with a pocket in between the sub-domains 1A and 2A in ParM (Gayathri et al. 2012). The corresponding region in actin, between sub-domains 1 and 3, also called hydrophobic cleft, binds most of the modulators of actin polymerization such as nucleators (eg. formin, spire) and severing factors, for example cofilin and twinfilin, interact (Dominguez and Holmes 2011). For ParM, the same region in the fold has been exploited to regulate the filament dynamics in a bacterial actin. Determining the binding site on the actin fold will be required to confirm a similar mode of action

for the other adaptor proteins such as AlfB, which acts both as a severing factor and a nucleator. The existence of nucleating factors and formin-like proteins in the bacterial cytoskeletal systems was for the first time established for plasmid segregation systems (Gayathri et al. 2012).

To conclude, the actin-like plasmid segregation machinery has evolved to perform the function of DNA segregation with a minimal number of components. Further mechanistic insights being gained from the large variety of actin-like plasmid segregation machineries will uncover the means by which a spindle can be built, and also provide invaluable information on the evolution of cytoskeletal systems from bacteria, archaea and eukaryotes. Recent work reporting an archaeal segregation system is a timely example for progress in these directions (Schumacher et al. 2015).

Acknowledgements The work in the lab is supported by INSPIRE Faculty Research Grant, Department of Science and Technology (DST), India, and Innovative Young Biotechnologist Award, Department of Biotechnology, India. PG and SH acknowledge INSPIRE (DST) and IISER-Pune fellowships.

References

Becker E, Herrera NC, Gunderson FQ et al (2006) DNA segregation by the bacterial actin AlfA during Bacillus subtilis growth and development. EMBO J 25:5919–5931. doi:10.1038/sj.emboj.7601443

Bharat TAM, Murshudov GN, Sachse C, Löwe J (2015) Structures of actin-like ParM filaments show architecture of plasmid-segregating spindles. Nature 523:106–110. doi:10.1038/nature14356

Bork P, Sander C, 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

Campbell CS, Mullins RD (2007) In vivo visualization of type II plasmid segregation: bacterial actin filaments pushing plasmids. J Cell Biol 179:1059–1066. doi:10.1083/jcb.200708206

Dam M, Gerdes K (1994) Partitioning of plasmid R1 Ten direct repeats flanking the parA promoter constitute a centromere-like partition site parC, that expresses incompatibility. J Mol Biol 236:1289–1298. doi:10.1016/0022-2836(94)90058-2

Derman AI, Nonejuie P, Michel BC et al (2012) Alp7R regulates expression of the actin-like protein Alp7A in Bacillus subtilis. J Bacteriol 194:2715–2724. doi:10.1128/JB.06550-11

Dominguez R, Holmes KC (2011) Actin structure and function. Annu Rev Biophys 40:169–186. doi:10.1146/annurev-biophys-042910-155359

Fink G, Löwe J (2015) Reconstitution of a prokaryotic minus end-tracking system using TubRC centromeric complexes and tubulin-like protein TubZ filaments. Proc Natl Acad Sci U S A 2–7:E1845–E1870. doi:10.1073/pnas.1423746112

Fujii T, Iwane AH, Yanagida T, Namba K (2010) Direct visualization of secondary structures of F-actin by electron cryomicroscopy. Nature 467:724–728. doi:10.1038/nature09372

Galkin VE, Orlova A, Rivera C et al (2009) Structural polymorphism of the ParM filament and dynamic instability. Structure 17:1253–1264. doi:10.1016/j.str.2009.07.008

Garner EC (2007) Reconstitution of DNA Segregation. Science 315:1270–1274. doi:10.1126/science.1138527

Garner EC, Campbell CS, Mullins RD (2004) Dynamic instability in a DNA-segregating prokaryotic actin homolog. Science 306:1021–1025. doi:10.1126/science.1101313
Gayathri P, Fujii T, Møller-Jensen J et al (2012) A bipolar spindle of antiparallel ParM filaments drives bacterial plasmid segregation. Science 338:1334–1337. doi:10.1126/science.1229091
Gayathri P, Fujii T, Namba K, Löwe J (2013) Structure of the ParM filament at 8.5Å resolution. J Struct Biol 184:33–42. doi:10.1016/j.jsb.2013.02.010
Gerdes K, Molin S (1986) Partitioning of plasmid R1. J Mol Biol 190:269–279. doi:10.1016/0022-2836(86)90001-X
Gerdes K, Howard M, Szardenings F (2010) Pushing and pulling in prokaryotic DNA segregation. Cell 141:927–942. doi:10.1016/j.cell.2010.05.033
Gunning PW, Ghoshdastider U, Whitaker S et al (2015) The evolution of compositionally and functionally distinct actin filaments. J Cell Sci 128:2009–2019. doi:10.1242/jcs.165563
Jensen RB, Gerdes K (1997) Partitioning of plasmid R1. The ParM protein exhibits ATPase activity and interacts with the centromere-like ParR-parC complex. J Mol Biol 269:505–513. doi:10.1006/jmbi.1997.1061
Löwe J, Amos LA (2009) Evolution of cytomotive filaments: te cytoskeleton from prokaryotes to eukaryotes. Int J Biochem Cell Biol 41:323–329. doi:10.1016/j.biocel.2008.08.010
Møller-Jensen J, Jensen RB, Löwe J, Gerdes K (2002) Prokaryotic DNA segregation by an actin-like filament. EMBO J 21:3119–3127. doi:10.1093/emboj/cdf320
Møller-Jensen J, Borch J, Dam M et al (2003) Bacterial mitosis: ParM of plasmid R1 moves plasmid DNA by an actin-like insertional polymerization mechanism. Mol Cell 12:1477–1487. doi:10.1016/S1097-2765(03)00451-9
Møller-Jensen J, Ringgaard S, Mercogliano CP et al (2007) Structural analysis of the ParR/parC plasmid partition complex. EMBO J 26:4413–4422. doi:10.1038/sj.emboj.7601864
Orlova A, Garner EC, Galkin VE et al (2007) The structure of bacterial ParM filaments. Nat Struct Mol Biol 14:921–926
Ozyamak E, Kollman J, Agard DA, Komeili A (2013) The bacterial actin MamK: in vitro assembly behavior and filament architecture. J Biol Chem 288:4265–4277. doi:10.1074/jbc.M112.417030
Pei J, Tang M, Grishin NV (2008) PROMALS3D web server for accurate multiple protein sequence and structure alignments. Nucleic Acids Res 36:W30–W34. doi:10.1093/nar/gkn322
Petek NA, Mullins RD (2014) Bacterial actin-like proteins: purification and characterization of self-assembly properties. Methods Enzymol 540:19–34. doi:10.1016/B978-0-12-397924-7.00002-9
Polka JK, Kollman JM, Agard DA, Mullins RD (2009) The structure and assembly dynamics of plasmid actin AlfA imply a novel mechanism of DNA segregation. J Bacteriol 191:6219–6230. doi:10.1128/JB.00676-09
Polka JK, Kollman JM, Mullins RD (2014) Accessory factors promote AlfA-dependent plasmid segregation by regulating filament nucleation, disassembly, and bundling. Proc Natl Acad Sci 111:2176–2181. doi:10.1073/pnas.1304127111
Pollard TD (1986) Rate constants for the reactions of ATP-and ADP-actin with the ends of actin filaments. J Cell Biol 103:2747. doi:10.1083/jcb.103.6.2747
Popp D, Narita A, Oda T et al (2008) Molecular structure of the ParM polymer and the mechanism leading to its nucleotide-driven dynamic instability. EMBO J 27:570–579. doi:10.1038/sj.emboj.7601978
Popp D, Xu W, Narita A et al (2010) Structure and filament dynamics of the pSK41 actin-like ParM protein: implications for plasmid dna segregation. J Biol Chem 285:10130–10140. doi:10.1074/jbc.M109.071613
Popp D, Narita A, Lee LJ et al (2012) Novel actin-like filament structure from Clostridium tetani. J Biol Chem 287:21121–21129. doi:10.1074/jbc.M112.341016
Rivera CR, Kollman JM, Polka JK et al (2011) Architecture and assembly of a divergent member of the ParM family of bacterial actin-like proteins. J Biol Chem 286:14282–14290. doi:10.1074/jbc.M110.203828

Salje J, Löwe J (2008) Bacterial actin: architecture of the ParMRC plasmid DNA partitioning complex. EMBO J 27:2230–2238. doi:10.1038/emboj.2008.152

Salje J, Zuber B, Löwe J (2009) Electron cryomicroscopy of *E. coli* reveals filament bundles involved in plasmid DNA segregation. Science 323:509–512. doi:10.1126/science.1164346

Salje J, Gayathri P, Löwe J (2010) The ParMRC system: molecular mechanisms of plasmid segregation by actin-like filaments. Nat Rev Microbiol 8:683–692. doi:10.1038/nrmicro2425

Schumacher MA, Glover TC, Brzoska AJ et al (2007) Segrosome structure revealed by a complex of ParR with centromere DNA. Nature 450:1268–1271. doi:10.1038/nature06392

Schumacher MA, Tonthat NK, Lee J et al (2015) Structures of archaeal DNA segregation machinery reveal bacterial and eukaryotic linkages. Science 349:1120–1124. doi:10.1126/science.aaa9046

Szwedziak P, Wang Q, Freund SM, Löwe J (2012) FtsA forms actin-like protofilaments. EMBO J 31:2249–2260. doi:10.1038/emboj.2012.76

Tanaka T (2010) Functional analysis of the stability determinant AlfB of pBET131, a miniplasmid derivative of Bacillus subtilis (natto) plasmid pLS32. J Bacteriol 192:1221–1230. doi:10.1128/JB.01312-09

van den Ent F, Amos LA, Löwe J (2001) Prokaryotic origin of the actin cytoskeleton. Nature 413:39–44. doi:10.1038/35092500

van den Ent F, Møller-Jensen J, Amos LA et al (2002) F-actin-like filaments formed by plasmid segregation protein ParM. EMBO J 21:6935–6943. doi:10.1093/emboj/cdf672