Prosthecobacter BtubAB form bacterial mini microtubules

Xian Deng

Corpus Christi College
University of Cambridge

Thesis submitted to the University of Cambridge for the degree of Doctor of Philosophy

October 2017
Abstract

The tubulin/FtsZ superfamily contains a large set of proteins that spans through all kingdoms of life, with αβ-tubulins being the eukaryotic representatives and FtsZ being the best studied prokaryotic homologue. It is believed that all tubulin/FtsZ-related proteins have evolved from a common ancestor, however, members from this superfamily have diverged in many aspects. αβ-tubulins polymerise into giant and hollow microtubules in the presence of GTP. Despite the size of around 25 nm wide, microtubules display sophisticated dynamics. In particular, dynamic instability, the stochastic change between fast growth and rapid shrinkage, is a hallmark of microtubules. In contrast to αβ-tubulins, FtsZ lacks the C-terminal domain of tubulins and it probably functions as single homopolymeric protofilaments, possibly through treadmilling dynamics. There is strong divergence of the biological functions in the tubulin/FtsZ superfamily. Microtubules are involved in fundamental processes such as motility, transport and chromosomal segregation, whereas FtsZ is involved in bacterial cytokinesis (bacterial cell division), and the equivalent role of FtsZ is carried out by actin-based and ESCRTIII-based systems in eukaryotes. It seems that there is a big evolutionary gap between αβ-tubulins and FtsZ, and the only properties that are conserved within the tubulin/FtsZ superfamily are fold, protofilament formation and GTPase activity.

In 2002, a pair of tubulin-like genes, btuba and btubb were identified in Prosthecobacter bacteria, with higher sequence homology to eukaryotic tubulins than FtsZ or any other bacterial homologues. The crystal structures solved later revealed, again, a closer similarity to αβ-tubulins than to their prokaryotic equivalents. It has been known for a while that BtubAB form filaments in the presence of GTP, however, little knowledge has been available regarding the filament architecture. In this project, I determined the near atomic resolution structure of the in vitro BtubAB filament using cryoEM and cryoET, revealing a hollow tube that consists of four protofilaments. A closer look showed that BtubAB filaments have many conserved microtubule features including: an overall polarity, similar longitudinal contacts, M-loops in lateral interfaces, and the presence of the seam, a structural hallmark of microtubules. My study also shows that BtubC, a protein with a TPR fold, binds to the BtubAB filaments in a stoichiometric manner, similar to some MAPs on microtubules. Based on this work, I concluded that BtubAB from Prosthecobacter form bacterial ‘mini microtubules’, and my work provided interesting insight into the evolution of tubulin/FtsZ-related proteins.
**Declaration**

This thesis is the result of my own work and includes nothing which is the outcome of work done in collaboration, except Chapter 4.1 in discussion, which is based on the collaborative work done with my colleague Gero Fink, as indicated in the text.

This thesis is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution. I further state that no substantial part of my dissertation has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution.

This thesis does not exceed the prescribed word limit for the relevant Degree Committee.
Acknowledgements

First of all, I would like express my sincere gratitude to my supervisor Jan Löwe, for giving me this exciting project, providing scientific guidance and continuous support during my time in the LMB. Through Jan, I learnt the qualities needed to become a fully formed scientist. I am grateful to Tanmay Bharat, who introduced me to the world of electron microscopy, and taught me all the fundamentals about cryoEM, cryoET and helical reconstruction. My thanks also go to Gero Fink, for working on a parallel project without which the story would not be complete. I would like to thank Danguole Ciziene, for her support in my everyday lab work and help with protein crystallography. I consider myself lucky to be surrounded by the Löwe lab and other friends in the LMB, for all the helpful discussions and also joyful moments.

This project would not have been possible without the support from Shaoxia Chen, Christos Savva and Giuseppe Cannone from the LMB EM facility, Jake Grimmett and Toby Darling from Scientific Computing, Shaoda He and Sjors Scheres as developers of RELION, and Minmin Yu and Fabrice Gorrec from the X-ray Crystallography Facility. I have greatly benefited from the extraordinary cryoEM expertise and the friendly environment throughout LMB. I should also thank everyone in LMB Stores, for protecting my reputation generated from my excessive amount of Amazon parcels, and help with some inconvenience from living in Cambridge.

I am grateful to Boehringer Ingelheim Fonds, who generously supported my PhD studentship, and introduced me to other BIF members and alumni for stimulating and exciting discussions.

Finally, I would like to thank my parents, without whom I would not be here at all, and Tim Nierhaus, for his enduring patience and love for me.
Real Acknowledgements

I was informed that I have great autonomy in my thesis, therefore I decide to exercise this advantage and write this so-called ‘Real Acknowledgements’ to reflect. This is not to say that the previous acknowledgement is ‘fake’, like some of the news these days. I simply would like to express my gratitude again in a more personal way, as the PhD journey to me as a life experience has been of more value than scientific training.

My journey started with Jan Löwe’s decision of taking me as a PhD student, hence saved me from pursuing a doctorial study in Oxford. I am very grateful to Jan, not only for his scientific guidance, but also for his general support, for his understanding and encouragement even when I chose to pursue a future career that is different from science. I would also like to thank Danguole Ciziene, her caring was particularly touching since I am thousands of miles away from home. I enjoy very much my time working in the LMB, because of its excellent science, unique atmosphere, and the interesting people. LMB is such a diverse place, and I have benefited greatly from the cultural diversity of the Löwe lab and the LMB people. Meeting people from different backgrounds, having stimulating intellectual discussions and learning different opinions have made my PhD study a wonderful experience.

It is because of my parents’ foresight and bravery that I came to this country 11 years ago. Over time, I became more and more grateful to them, for their support, wisdom, and enabling me to achieve the life I have today. I also consider myself very fortunate to have many friends during my time in Cambridge. In particular, I would like to thank Tim Nierhaus, who has supported me through one of the most difficult times in my life. Also, I am very grateful to my extraordinary friendship with Jenny Boyang Liu over the past 10 years, for her being so understanding, honest and supportive.

Life always comes with a certain degree of chance and surprise. Without Jan taken me in the first place, I would not have spent my past three years with so many interesting people from the Löwe lab and the LMB. Although I have decided to follow a career outside science, my time in the LMB has been a wonderful experience, and I am truly grateful to everyone, and everything.
# List of abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>+TIPs</td>
<td>Plus-end tracking proteins</td>
</tr>
<tr>
<td>bklc</td>
<td>Bacterial kinesin light chain</td>
</tr>
<tr>
<td>btubA</td>
<td>Bacterial tubulin A</td>
</tr>
<tr>
<td>btubB</td>
<td>Bacterial tubulin B</td>
</tr>
<tr>
<td>CAMSAP</td>
<td>Calmodulin-regulated spectrin-associated protein</td>
</tr>
<tr>
<td>CLIP-170</td>
<td>Cytoplasmic linker protein of 170 kDa</td>
</tr>
<tr>
<td>cryoEM</td>
<td>Electron cryomicroscopy</td>
</tr>
<tr>
<td>cryoET</td>
<td>Electron cryotomography</td>
</tr>
<tr>
<td>CTF</td>
<td>Contrast transfer function</td>
</tr>
<tr>
<td>DCX</td>
<td>Doublecortin</td>
</tr>
<tr>
<td>DIC</td>
<td>Differential interference contrast</td>
</tr>
<tr>
<td>E. coli</td>
<td>Escherichia coli</td>
</tr>
<tr>
<td>EB1</td>
<td>End binding 1</td>
</tr>
<tr>
<td>EDTA</td>
<td>Ethylenediaminetetraacetic acid</td>
</tr>
<tr>
<td>EGTA</td>
<td>Triethylene glycol diamine tetraacetic acid</td>
</tr>
<tr>
<td>EM</td>
<td>Electron microscopy</td>
</tr>
<tr>
<td>EMTS</td>
<td>Thiomersal</td>
</tr>
<tr>
<td>ESCRT</td>
<td>Endosomal sorting complexes required for transport</td>
</tr>
<tr>
<td>FSC</td>
<td>Fourier shell correlation</td>
</tr>
<tr>
<td>GDP</td>
<td>Guanidine diphosphate</td>
</tr>
<tr>
<td>GTP</td>
<td>Guanidine triphosphate</td>
</tr>
<tr>
<td>GTP-g-S</td>
<td>Guanosine 5'-O-(gamma-thio)triphosphate</td>
</tr>
<tr>
<td>IPTG</td>
<td>Isopropyl β-D-1-thiogalactopyranoside</td>
</tr>
<tr>
<td>MAPs</td>
<td>Microtubule-associated proteins</td>
</tr>
<tr>
<td>MTOC</td>
<td>Microtubule-organising centre</td>
</tr>
<tr>
<td>Term</td>
<td>Definition</td>
</tr>
<tr>
<td>-----------------</td>
<td>-------------------------------------------------</td>
</tr>
<tr>
<td>NMR</td>
<td>Nuclear magnetic resonance</td>
</tr>
<tr>
<td>P. vanneervenii</td>
<td>Prosthecobacter vanneervenii</td>
</tr>
<tr>
<td>P. dejonjeii</td>
<td>Prosthecobacter dejonjeii</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase chain reaction</td>
</tr>
<tr>
<td>pdBtubAB</td>
<td>Prosthecobacter dejonjeii BtubAB</td>
</tr>
<tr>
<td>pdBtubC</td>
<td>Prosthecobacter dejonjeii BtubC</td>
</tr>
<tr>
<td>pvBtubC</td>
<td>Prosthecobacter vanneervenii BtubC</td>
</tr>
<tr>
<td>RMSD</td>
<td>Root-mean-square deviation</td>
</tr>
<tr>
<td>S. pombe</td>
<td>Schizosaccharomyces pombe</td>
</tr>
<tr>
<td>SAD</td>
<td>Single anomalous dispersion</td>
</tr>
<tr>
<td>SNR</td>
<td>Signal-to-noise ratio</td>
</tr>
<tr>
<td>TEM</td>
<td>Transmission electron microscope</td>
</tr>
<tr>
<td>TIRF</td>
<td>Total internal reflection fluorescence</td>
</tr>
<tr>
<td>TOG1</td>
<td>Tumor overexpressed gene 1</td>
</tr>
<tr>
<td>TPR</td>
<td>Tetratricopeptide repeat</td>
</tr>
</tbody>
</table>
Table of Contents

Abstract ........................................................................................................................... i

Declaration ...................................................................................................................... ii

Acknowledgements ...................................................................................................... iii

Real Acknowledgements .............................................................................................. iv

List of abbreviations ..................................................................................................... v

1 Introduction ................................................................................................................ 1

1.1 Microtubules .......................................................................................................... 1

1.1.1 Overview ........................................................................................................... 1

1.1.2 Tubulin dimer structure and nucleotide binding .............................................. 1

1.1.3 Microtubule structure ....................................................................................... 6

1.1.4 Protofilament curvature .................................................................................... 9

1.1.5 Dynamic instability .......................................................................................... 9

1.1.6 Dynamic polarity and treadmilling .................................................................. 12

1.1.7 Microtubule-associated proteins ..................................................................... 13

1.2 Prokaryotic tubulin homologues ......................................................................... 14

1.2.1 Overview ........................................................................................................... 14

1.2.2 FtsZ .................................................................................................................. 14

1.2.3 TubZ ................................................................................................................ 18

1.2.4 CetZ ................................................................................................................ 18

1.3 BtubAB from Prosthecobacter ............................................................................. 19

1.4 Aims of this work .................................................................................................. 21

2 Materials and Methods ............................................................................................. 22

2.1 General methods ................................................................................................... 22

2.2 Strains and plasmids ........................................................................................... 25

2.3 Plasmid construction ............................................................................................ 25

2.4 Protein expression and purification ..................................................................... 25

2.5 Pelleting assays .................................................................................................... 27

2.6 Negative stain electron microscopy sample preparation and imaging .............. 27

2.7 CryoET sample preparation and imaging ............................................................ 28

2.8 CryoET data analysis and subtomogram averaging ........................................... 29
2.9 CryoEM sample preparation and imaging .......................................................... 29
2.10 Helical reconstruction of BtubAB and BtubABC filaments, model building and refinement ........................................................................................................... 30
2.11 BtubC crystallisation, data collection and structure determination ............... 31
2.12 Coordinate and map depositions .................................................................. 32

3 Results ................................................................................................................. 33
3.1 BtubAB form 4 protofilament hollow tubes by cryoET ..................................... 33
3.2 CryoEM study of BtubAB filament and its limitation ........................................ 37
3.3 BtubC contains a TPR fold and binds to BtubAB filaments in a stoichiometric fashion ............................................................................................................. 42
3.4 BtubAB forms bacterial mini microtubules ...................................................... 48

4 Discussion ......................................................................................................... 57
4.1 Investigation of BtubAB filament dynamics by Gero Fink ................................. 57
4.2 BtubAB and a comparison to microtubules ...................................................... 64
4.3 Possible origins of the btub operon and the evolution of FtsZ/tubulin-related proteins ................................................................................................................. 71
4.4 BtubAB and microtubule dynamics .................................................................. 76

Appendix A Crystallography and cryoEM data statistics ................................. 90
Appendix B Plasmids and Strains ........................................................................ 91
Appendix C Growth media and general buffers .................................................. 92
Appendix D A brief introduction to electron microscopy .................................... 94
Appendix E Other experiments .......................................................................... 107
1 Introduction

1.1 Microtubules

1.1.1 Overview

As early as 1939, Wilhelm Joseph Schmidt observed birefringent mitotic spindles in marine eggs using polarising light microscopy (Schmidt, 1939). Sometime later, two groups discovered the $9 + 2$ arrangement of axonemes in cilia and flagella using transmission electron microscopy (Manton and Clarke, 1952; Fawcett and Porter, 1954; Brinkley, 1997). Today, it is well understood that mitotic spindles and the axoneme are constituted of microtubules, cylindrical filaments found in all eukaryotic cells. Microtubules are formed by the polymerisation of $\alpha$- and $\beta$-tubulin heterodimers, and are involved in many fundamental biological processes such as motility, transport and cell division. Despite being giant hollow tubes 25 nm wide, microtubules are highly dynamic. During mitosis, microtubules use a unique property termed dynamic instability, the stochastic change between slow growth and rapid shrinkage, to probe the surrounding environment until designated targets are found. Having central cellular functions, coupled with unique dynamic behaviours, the microtubule remains a fascinating object for biologists, biophysicists and structural biologists to study.

1.1.2 Tubulin dimer structure and nucleotide binding

The building blocks of microtubules are $\alpha\beta$-tubulin heterodimers, which are synthesised via a complex folding and assembly pathway involving a large number of chaperones and cofactors. Cytosolic chaperonin CCT is required to fold nascent tubulin polypeptides into folding intermediates, and the intermediates are subsequently captured by additional cofactors that facilitate the formation of native $\alpha\beta$-tubulin heterodimers (Lewis et al., 1997; Szymanski, 2002). Tubulin dimers are incredibly stable, with a dissociation half-time of ~9.6 hours and a dissociation constant of $10^{-11}$ M, explaining the scarcity of isolated tubulin monomers in cell extract (Caplow and Fee, 2002).

$\alpha$- and $\beta$-tubulin (~50 kDa each in size) are highly homologous with 30 – 40% sequence identity. Structurally, $\alpha$-tubulin has a root-mean-square deviation (RMSD) of 1.09 Å compared to $\beta$-tubulin, making them difficult to distinguish below 6 Å resolution (Nogales et al., 1995).
The high similarity between the \( \alpha \)- and \( \beta \)-tubulin implies their possible origins in an ancient gene duplication event, and a recent bioinformatics study suggests the presence of these two tubulins in the last common ancestor of eukaryotes (Findeisen et al., 2014). However, within the recently discovered ‘Asgard’ archaea, whose genomes are enriched for proteins formerly considered specific to eukaryotes including actin homologues, endosomal sorting complex required for transport (ESCRT) complex proteins and a variety of small GTPases, only Odinarchaeote seems to contain one copy of tubulin gene in its genome. (Zaremba-Niedzwiedzka et al., 2017)

Each tubulin monomer binds a guanine nucleotide, which is non-exchangeable when bound to the \( \alpha \) subunit at the N site, and exchangeable when bound to the \( \beta \) subunit at the E site (Weisenberg et al., 1968; MacNeal and Purich, 1978). A major contribution to our knowledge of tubulins and microtubules has come from studies using transmission electron microscopy (EM). The first high resolution structure of the tubulin dimer was obtained by electron crystallography studies of zinc-induced tubulin sheets stabilised with taxol, containing antiparallel tubulin protofilaments (Nogales et al., 1998b; Löwe et al., 2001) (Figure 1.1a). The 3.7 Å resolution map revealed that each monomer has three functional domains, namely the N-terminal nucleotide binding region, an intermediate (or activation) domain with the taxol binding pocket (only in \( \beta \)-tubulin), and the C-terminal domain, which motor proteins bind to. The N-terminal nucleotide binding domain consists of six parallel \( \beta \)-strands, S1 to S6, alternating with helices H1 to H6. The loops joining each \( \beta \)-strand with the next \( \alpha \)-helix are termed loop T1 to T6. The N-terminal region is connected to the next domain, the intermediate domain, via the core helix H7. The smaller intermediate domain contains three helices (H8 – H10) and four \( \beta \)-strands (S7 – S10), which is followed by two antiparallel helices (H11 and H12) constituting the C-terminal domain.

Along protofilaments, the longitudinal contact between subunits form the nucleotide binding and hydrolysis site. Both inter-dimer and intra-dimer interfaces are very similar, each consists of two extensive surfaces with highly complementary shapes and a nucleotide locked in between (Nogales et al., 1998a; Nogales et al., 1999) (Figure 1.1b). Specifically, the nucleotide is bound by loops T1 to T6 from the N-terminal domain and the beginning of core helix H7 of the tubulin monomer, and the nucleotide binding pocket is completed by loop T7 from the tubulin at the opposite side. The N-site (non-exchangeable) guanosine triphosphate (GTP) in \( \alpha \)-tubulin is buried deeply at the intra-dimer interface, while the E-site (exchangeable) nucleotide is exposed on the surface of the dimer, therefore explaining the exchangeability of
the E-site nucleotide. At the N-site, the $\gamma$-phosphate from the GTP is coordinated by Lys254 in $\beta$-tubulin, whereas the E-site nucleotide binding pocket is completed with Glu254 in $\alpha$-tubulin, which appears to be essential for polymerisation-dependent hydrolysis from yeast mutation studies (Anders and Botstein, 2001; Löwe et al., 2001). As the first high resolution tubulin structure was obtained from zinc-induced 2D sheets containing antiparallel tubulin filaments, detailed structural information about lateral interactions in microtubules was not revealed till later studies.
**Figure 1.1 Structure of αβ-tubulin by electron crystallography and cryoEM.**

(a) The first refined high resolution structure of αβ-tubulin by electron crystallography (PDB 1JFF) showing the three-domain architecture. Colour codes: slate – N-terminal GTP binding domain, yellow – central Helix H7, orange – intermediate domain, green – C-terminal domain containing two antiparallel helices that are involved in motor protein binding, black – GTP (in α-tubulin) and GDP (in β-tubulin), grey – taxol.

(b) Microtubule structure at 4.9 Å by cryoEM (PDB 3J6F); only a pair of heterodimers along one protofilament is shown here, with the zoom in views of the intra-dimer and inter-dimer interface showing the non-hydrolysable GTP at the nucleotide binding pocket of α-tubulin, and the GDP at the nucleotide binding pocket of β-tubulin.
1.1.3 Microtubule structure

A microtubule is a tube constructed from parallel protofilaments in which stable αβ-tubulin heterodimers polymerise in a head to tail fashion in the presence of GTP. This arrangement results in a polymer with a defined polarity, with ‘plus’ and ‘minus’ ends, that are structurally and dynamically distinct. Biochemistry studies showed that the plus end is capped by β-tubulin subunits exposing their nucleotide binding pockets to the solution, whereas the minus end is capped by α-subunits exposing their catalytic ends (Mitchison, 1993; Fan et al., 1996; Nogales et al., 1999). The polarity of microtubules has significant implications in terms of their dynamics, as discussed below. In cells, the minus end is typically anchored at microtubule organising centres (MTOCs), which microtubules grow from, whereas the plus end is exposed to the cytoplasm or attached to a specific target such as the kinetochore (Wu and Akhmanova, 2017).

A medium-resolution model of the microtubule was first obtained by docking the tubulin crystal structure described above into a 20 Å cryoEM reconstruction of microtubules (Nogales et al., 1999). The model revealed the orientation of tubulin monomers, showing that the C-terminal helices of the subunits (H11 and H12) form a crest on the outside surface of the protofilaments, enabling the binding of motor proteins. The lateral interaction between the protofilaments involves the M-loop from one subunit and loop H1-S2 and helix H3 from the other subunit. Later, a 4.7 Å map obtained by electron cryomicroscopy (cryoEM) in 2014 revealed unambiguously that the key lateral contact (between parallel protofilaments) involves the insertion of one aromatic residue (Y283 in β-β interaction, H283 in α-α interaction) from one tubulin into the hydrophobic pocket of the adjacent tubulin, and this interaction seems to be similar across different nucleotide hydrolysis states (Alushin et al., 2014). Given the small lateral contact (as compared to the longitudinal one), it is not surprising that a certain degree of microtubule architectural flexibility is tolerated. Indeed, when assembled in vitro using purified tubulins, microtubules with protofilament numbers ranging from 9 to 16 have been observed (Chrétien and Wade, 1991; Sosa and Milligan, 1996; Sui and Downing, 2010) (Figure 1.2). However, this heterogeneity is limited in vivo. Microtubules in cells grow out from nucleating complexes known as MTOCs, containing a templating ring structure formed by another tubulin homologue, γ-tubulin (Oakley and Oakley, 1989; Kollman et al., 2011; Oakley et al., 2015). The MTOCs establish microtubule polarity as well as serving as a size template. Under this controlled polymerisation regime, most of the microtubules in the cell have 13
protofilaments (Tilney et al., 1973; Nogales, 2000). An exception to this rule is found in the model organism *Caenorhabditis elegans*, where 11-protofilament microtubules are found in somatic cells and 15-protofilament microtubules are present in touch receptor neurons (Chalfie and Thomson, 1982). When composed of 13 protofilaments, microtubules run almost perfectly straight and non-helical (each protofilament axis is almost parallel to the microtubule axis), allowing microtubule-associated motor proteins such as kinesin and dynein to walk straight along a microtubule without rotating around it. In microtubules with fewer or more than 13 protofilaments, the lattices wind slowly around the axis (Chrétien and Wade, 1991; Amos and Schlieper, 2005; Sui and Downing, 2010).

A structural hallmark of 13- and 14-protofilament microtubules is the presence of a seam (Figure 1.2). Given the alternative arrangement of α- and β-subunits in a microtubule, most of the lateral interfaces are made between tubulins of the same type (α-α and β-β), known as the B-lattice. However, in microtubules with 13 or 14 protofilaments, at every helical turn an α-tubulin makes lateral contact with a β-tubulin, known as the A-lattice, or the seam (Song and Mandelkow, 1993; Kikkawa et al., 1994; Amos and Schlieper, 2005). The biological significance of the seam is still under investigation, nevertheless, one observation is that the presence of a seam provides a distinct contact surface for microtubule-associated proteins (MAPs). For example, doublecortin (DCX), a well-studied stabilising MAP, binds to 13-protofilament microtubules only at the B-lattice but not at the seam, whereas the *Schizosaccharomyces pombe* (*S. pombe*) end-binding protein 1 (EB1) homolog Mal3p stabilises microtubules by its preferential binding to the A-lattice seam (Sandblad et al., 2006; Fourniol et al., 2010).
Figure 1.2 Comparison of microtubules with 12-16 protofilaments, adapted from (Amos and Schlieper, 2005). For a microtubule containing 13 protofilaments, the protofilament lattice runs almost straight, whereas those with more or fewer protofilaments must wind slowly around the axis to allow the tubulin subunits to line up correctly. Only 13- and 14-protofilament microtubules have seams, where lateral contacts are formed between two different subunits ($\alpha$-$\beta$), as opposed to all other lateral interfaces that are formed by either $\alpha$-$\alpha$ or $\beta$-$\beta$. 
1.1.4 Protofilament curvature

Major efforts have been invested in solving structures of microtubules in different nucleotide states and in complex with various MAPs. It appears that tubulin dimers exist in a range of curvatures, from the ones in Zn$^{2+}$ induced 2D-sheet being completely straight, to that found in tumour overexpressed gene 1 (TOG1):αβ-tubulin complex with a 13° curvature (Nogales et al., 1998b; Ayaz et al., 2012). The curvature in this context refers to the overall bendiness of the microtubule, in regardless of its helical symmetry. For example, all 13-protofilament microtubules have the same helical parameters (rise and twist), however, the 13-protofilament microtubules can adopt different protofilament, or dimer curvatures depending on how much the microtubules are bent as whole.

Various microtubule stabilisers and destabilisers such as stathmin, kinesin-13 and Ndc80 have been shown to recognise or stabilise tubulin dimers in certain curvatures, however, a definitive link between protofilament curvature and microtubule dynamics is still under active investigation (Gigant et al., 2000; Alushin et al., 2010; Asenjo et al., 2013). Most of the microtubule structures available so far have been obtained in complex with MAPs and a free tubulin dimer structure representing a ‘monomeric state’ has yet to be reported. The different αβ-tubulin conformations in associations with MAPs have been extensively reviewed by Brouhard and Rice (Brouhard and Rice, 2014).

1.1.5 Dynamic instability

Microtubules are highly dynamic, they are capable of switching stochastically between growth and rapid shrinkage both in vivo and in vitro (Desai and Mitchison, 1997). This unique non-equilibrium behaviour, termed dynamic instability, was first proposed by Mitchison and Kirschner in 1984, and was later confirmed using dark field and differential interference contrast (DIC) microscopy (Mitchison and Kirschner, 1984a; Mitchison and Kirschner, 1984b; Horio and Hotani, 1986; Walker et al., 1988). The transition from polymerisation to depolymerisation is called catastrophe, and the reverse transition from depolymerisation to polymerisation is called rescue (Figure 1.3). It is generally accepted that the dynamic instability can be explained by the GTP-cap model. The model suggests a cap of GTP bound tubulin, the ‘GTP cap’, is present at the newly formed microtubule tip, which stabilises the microtubule and allows further growth, whereas microtubules with GDP bound tubulin at the plus end are intrinsically unstable. According to the GTP-cap model, as long as the GTP cap is
maintained, the microtubule continues to grow. The loss of the GTP cap by nucleotide hydrolysis triggers catastrophe and rapid shrinkage of the microtubule. Rescues are thought to be induced by the wave of catastrophic depolymerisation reaching ‘GTP islands’ present in the microtubule lattice or ‘rescue factors’ (Mitchison and Kirschner, 1984a; Desai and Mitchison, 1997; Akhmanova and Steinmetz, 2015). At a structural level, cryoEM studies suggest that changes in longitudinal interactions provide strong influence on microtubule stability. GTP hydrolysis and phosphate release lead to a compaction of the E-site and conformational changes in α-tubulin, in turn leading to global lattice rearrangements and the build-up of strain, which explains the metastable nature of the GDP-microtubule lattice (Hyman et al., 1995; Alushin et al., 2014; Zhang et al., 2015).

Although the GTP-cap model is the widely accepted model for microtubule dynamic instability, the exact mechanisms underlying the transitions between catastrophes and rescues are poorly understood. Catastrophes exhibit a complex ‘ageing’ effect in vitro, so that the microtubules that have been growing for longer (‘older’ microtubules) have a higher possibility to undergo catastrophes. This behaviour suggests that catastrophe requires more than one molecular event to occur, however, the nature of these events is largely unknown (Odde et al., 1995; Gardner et al., 2011; Coombes et al., 2013). Microtubule rescues are also not well understood. In vitro studies show that rescue events are not sensitive to tubulin concentration and therefore unlikely to be the result of the stochastic addition of GTP-tubulin at the shrinking end (Walker et al., 1988). The presence of ‘GTP islands’ of GTP-tubulins has been discovered and proposed to halt microtubule disassembly, and explain the observed rescue dynamics (Dimitrov et al., 2008). Structurally, cryoEM studies of microtubules bound with various MAPs, and crystallographic studies of tubulins sequestered by MAPs suggest possible effects of protofilament curvature in microtubule dynamics. However, the definitive links remain under active investigation (Brouhard and Rice, 2014).
Figure 1.3 Schematic explaining the assembly-disassembly cycle of microtubules, adapted from (Akhmanova and Steinmetz, 2015). Microtubules are known to exhibit the dynamic behaviour called dynamic instability: the property to switch stochastically between rapid growth and fast shrinkage. Currently, the GTP-cap model offers the most accepted explanation for this behaviour. It states that the growing tip of a microtubule contains tubulins with GTP that is yet to be hydrolysed, forming a ‘GTP cap’. As long as the cap is maintained, microtubule continues to grow, whereas GDP-tubulin in microtubule lattice is intrinsically unstable. Once the GTP cap is lost, microtubule undergoes catastrophe, the rapid depolymerisation process. This rapid shrinkage sometimes can be halted by rescue factors or GTP islands in the middle of microtubule lattice, subsequently, regrowth, or rescue, takes place and microtubules resume polymerisation.
As discussed in section 1.1.3, microtubules have a defined polarity with distinct structures and dynamics at each end. The plus end is capped by β-tubulin subunits exposing their nucleotide binding pocket to the solution, whereas the minus end is capped by α-subunits (Mitchison, 1993; Nogales, 1999). This structural polarity has profound impacts on microtubule dynamics. In vitro, both plus and minus ends can grow in solution and each can bear a GTP cap. However, the minus end grows more slowly and undergoes catastrophes less frequently compared to the plus end. When the stabilising cap is severed in vitro, the plus end rapidly depolymerises, whereas the minus end resumes growth at the normal minus end rate (Walker et al., 1989; Tran et al., 1997).

The structural and dynamic polarity of microtubules explains their ability to treadmill, that is filaments containing stationary subunits can appear to move across a surface by simultaneously growing at the plus end and shrinking at the minus end, at similar rates. Microtubule treadmilling was observed in vitro in the 1980s using electron microscopy and dark field microscopy methods (Margolis and Wilson, 1978; Rothwell et al., 1985; Hotani and Horio, 1988). Dynamics consistent with treadmilling have been reported in cells, however, such behaviours are often termed ‘poleward flux’ rather than ‘treadmilling’ due to the possibility that these dynamics are caused by the activities of microtubule motors (Mitchison, 1989; Mitchison and Salmon, 1992; Margolis and Wilson, 1998). Since most of the microtubules are believed to be anchored to MTOCs at the minus ends, the extent to which treadmilling contributes to microtubule dynamics in vivo is under debate. Despite that, it has been observed that microtubules sometimes detach from their nucleation sites and treadmill towards the periphery of the cell (Rodionov and Borisy, 1997; Rodionov et al., 1999; Vorobjev et al., 1999), although these events are relatively rare and short-lived.

Studies of Arabidopsis, a model organism for plant biology research, have revealed interesting microtubule arrangements in plant cells. In contrast to typical animal cells, in which the interphase microtubules radiate from the centrosome (MTOC) near the nucleus towards the cell periphery, microtubules in plant interphase cells form cortical transverse helical arrays, in order to guide the movement of the membrane-embedded cellulose synthase during cell wall synthesis (Lloyd et al., 2000; Baskin, 2001; Paredez et al., 2006). After nucleating on a pre-existing microtubule, the newly formed microtubule detaches from the old one it nucleated from, and the detached microtubules often exhibit dynamics at both ends. The plus end shows
polymerisation-biased dynamic instability, whereas the minus end shows a slow and intermittent depolymerisation. This hybrid treadmilling mechanism allows microtubules to migrate along the cell cortex, thus contributing to the microtubule repositioning and array organisation, required for plant cell morphogenesis (Shaw et al., 2003; Hashimoto, 2015).

1.1.7 Microtubule-associated proteins

MAPs have been the subject of intense research for decades. Numerous MAPs are known to regulate microtubule dynamics both spatially and temporally to facilitate the diverse functions of microtubules in different cells at different phases during the cell cycle. MAPs have a wide range of binding mechanisms. Some of them bind directly to the ends of microtubules. For example, cytoplasmic linker protein of 170 kDa (CLIP-170) is a plus-end tracking protein (+TIPs) (the first discovered of the more than 20 different +TIP families known today), whereas calmodulin-regulated spectrin-associated protein (CAMSAP) stabilises non-centrosomal microtubules by binding at the minus ends of growing microtubules (Perez et al., 1999; Akhmanova and Steinmetz, 2008; Akhmanova and Steinmetz, 2010; Jiang et al., 2014). In axons, MAP2/Tau is best known for its microtubule-stabilising activity by binding along the microtubule (Dehmelt and Halpain, 2004; Amos and Schlieper, 2005; Halpain and Dehmelt, 2006). Stathmin, a microtubule destabiliser, promotes depolymerisation by binding along a pair of tubulin heterodimers and caps the top α-tubulin subunit, thus preventing any further interactions (Gigant et al., 2000; Ravelli et al., 2004). Other MAPs include nucleation promoting proteins such as XMAP215, which binds free tubulin and target the microtubule ends by a diffusion-facilitated mechanism, resulting in an accelerated growth rate (Kinoshita et al., 2002; Brouhard et al., 2008).
1.2 Prokaryotic tubulin homologues

1.2.1 Overview

Tubulins, and the giant, hollow, yet dynamic microtubules were considered to be unique in eukaryotic cells for many decades, until the initial discovery of the bacterial tubulin homologue FtsZ in the 1990s (Bi and Lutkenhaus, 1991; De Boer et al., 1992; RayChaudhuri and Park, 1992; Löwe and Amos, 1998). Despite low sequence identity (10% to 18%), the crystal structure of FtsZ shows a fold that is similar to eukaryotic tubulin (Löwe and Amos, 1998). Since then, a surprising variety of bacterial and archaeal tubulin homologues have been subsequently identified (Jenkins et al., 2002; Tinsley and Khan, 2006; Aylett and Duggin, 2017). Research on these and other proteins over the past two decades has led to an appreciation that bacterial cells are highly organised in space and time. Just like their eukaryotic counterparts, prokaryotic tubulin proteins are involved in essential cellular processes such as DNA segregation, cell shape modulation and cell division (Amos et al., 2004; Møller-Jensen and Löwe, 2005; Erickson, 2007; Löwe and Amos, 2009).

1.2.2 FtsZ

The prokaryotic cell cycle used by most of the bacteria can be roughly divided into two steps, replication and segregation of DNA followed by cytokinesis and cell separation. The last step of bacterial cell division is septation, the constriction of the cell wall and cell membranes that leads to the formation of two daughter cells (Rothfield and Justice, 1997). FtsZ was first identified in genetic screens for bacterial cell division mutants, it is found in almost all bacteria and most archaea (Erickson, 1995; Erickson, 1997). Early studies have shown that FtsZ is a GTPase with weak sequence homology to eukaryotic tubulins, and as mentioned, the crystal structure revealed a close similarity between FtsZ and tubulins (De Boer et al., 1992; Löwe and Amos, 1998). Like tubulins, FtsZ has two domains, a GTPase domain and an activation domain connected by a long central helix (Figure 1.4a). Except for the two C-terminal helices that are missing in FtsZ, the general folds are well conserved between FtsZ and tubulin. (Figure 1.4b). Based on the structure and biochemical data, FtsZ became the first known bacterial tubulin homologue and indeed the first member of what is now called the Bacterial Cytoskeleton.

Several structure-based sequence alignments show that the amino acids highly conserved between FtsZ and eukaryotic tubulins are those involved in GTP binding and...
hydrolysis, whereas the residues in lateral interactions in tubulins and the residues after the activation domain are very divergent (Mukherjee and Lutkenhaus, 1994; Nogales et al., 1998a; Erickson, 2007). This explains why the only feature conserved between FtsZ and tubulins is the GTPase activity, and hence also the longitudinal interaction that makes protofilaments. Whereas tubulins form giant, hollow microtubules with 13 protofilaments exhibiting sophisticated dynamics, FtsZ probably exists as single homopolymeric protofilaments in cells (Erickson et al., 1996; Erickson et al., 2010; Szwedziak et al., 2014). In contrast to microtubules functioning as tracks for cellular transport and mitotic spindles during chromosomal segregation, FtsZ forms the so-called Z-ring, a ring-like structure at the bacterial division septum, which contracts during cytokinesis after recruiting other proteins to assemble into a macromolecular complex known as the divisome (Haeusser and Margolin, 2016). The exact mechanisms which FtsZ and the divisome employ to carry out cytokinesis and cell wall remodelling is still under investigation. Recent studies showed that short FtsZ filaments treadmill circumferentially around the division plane, and drive cell-wall remodelling divisomes with them. By moving, or treadmilling circumferentially, peptidoglycans are synthesised in increasingly smaller concentric rings, which ultimately leads to scission of the daughter cells (Bisson-Filho et al., 2017; Yang et al., 2017). The fact that FtsZ is able to adopt two distinct conformations (open and closed) provides a structural explanation for the treadmill behaviour of FtsZ filaments (Wagstaff et al., 2017).

FtsZ is also widely conserved in archaea with a role in cell division. However, the mechanism of archaeal FtsZ is poorly understood (Wang and Lutkenhaus, 1996; Aylett and Duggin, 2017). In addition, FtsZ systems are used by most plastids and many mitochondria for division (Kiefel et al., 2004; Osteryoung and Pyke, 2014).
a

GTP binding domain

Helix H7

Intermediate domain

b
Figure 1.4 The crystal structure of FtsZ and its comparison to β-tubulin.

(a) The crystal structure of FtsZ from *Methanocaldococcus jannaschii* at 2.8 Å resolution (PDB 1FSZ), showing a two-domain architecture with N-terminal GTPase domain and an intermediate domain connected by a long helix H7. Colour code: blue – GTPase domain, yellow – helix H7, orange – activation domain.

(b) Structural overlay of FtsZ (PDB 1FSZ) and β-tubulin (PDB 1JFF) in stereo. Except for the two C-terminal helices that are missing in FtsZ, the overall folds are very similar between FtsZ and eukaryotic tubulin. Colour code: Magenta ribbon – β-tubulin, green – FtsZ.
1.2.3 TubZ

TubZs form a diverse but likely monophyletic group of tubulin superfamily proteins mainly found on bacterial plasmids and in phage genomes (Fink and Aylett, 2017). They were first discovered on large virulence plasmids of Bacilli, and are so far mostly identified as components of type III plasmid segregation system (Okinaka et al., 1999; Berry et al., 2002). The phage-borne TubZ, known as PhuZ, is able to polymerise into filamentous arrays that are required for positioning phage DNA at the centre of the bacterial cell for optimal viral replication (Kraemer et al., 2012).

Several crystal structures have been solved for TubZ and PhuZ (Aylett et al., 2010; Hoshino and Hayashi, 2012; Kraemer et al., 2012; Aylett et al., 2013). In addition to the conserved GTPase domain and the activation domain that are signatory of tubulin/FtsZ-like proteins, TubZ and PhuZ have an additional C-terminal helix that extends to the next subunit along the protofilament, and is critical for robust filament formation and dynamics. While it is unknown how FtsZ might associate to form higher order structures in cells, TubZs from various species have been shown to form two-, three- and four-stranded helical filaments exhibiting interesting dynamics (Montabana and Agard, 2014; Zehr et al., 2014). In particular, TubZ from Bacillus thuringiensis pBtoxis plasmid is able to treadmill and pull its cargo DNA together in vitro (Fink and Löhwe, 2015).

1.2.4 CetZ

Unlike bacteria, many archaea possess additional FtsZ/tubulin-related proteins in addition to bona fide archaeal FtsZs that appear to function in cell division (Baumann and Jackson, 1996; Margolin et al., 1996; Vaughan et al., 2004). One example of these FtsZ co-existing proteins is CetZ, a distinct group of archaeal tubulin superfamily proteins that are found from the phylum Euryarchaeota (Vaughan et al., 2004). CetZ has features conserved in both FtsZs and eukaryotic tubulins. A recent study of CetZ1 from Haloferax volcanii shows CetZ1 to form dynamic filaments in vivo and that it is involved in cell shape maintenance (Duggin et al., 2015).
1.3 BtubAB from *Prosthecobacter*

For a long time, the giant, hollow, yet dynamic microtubules have been considered unique in eukaryotic system, and the sophisticated dynamic instability, although not fully understood at the mechanistic level, has rarely been observed in other cytoskeletal systems (Garner et al., 2004). Therefore, it is a particularly interesting discovery that the two tubulin-like genes, bacterial tubulin a (*btuba*) and bacterial tubulin b (*btubb*), in *Prosthecobacter* bacteria have much closer sequence homology with eukaryotic tubulins than FtsZ (Jenkins et al., 2002) (Figure 1.5a). Early microscopy and biochemistry studies showed that BtubAB proteins assemble into filaments in the presence of GTP at a 1:1 ratio, and together exhibit GTPase activity when mixed in equal molar ratio (Schlieper et al., 2005; Sontag et al., 2005). The crystal structure of the BtubAB heterodimer was solved not long after the initial discovery of the genes (Schlieper et al., 2005). BtubA and BtubB are closely related to each other, with 1.3 Å rmsd and 36% sequence homology. Remarkably, the BtubAB are strikingly similar to αβ-tubulins, the rmsd of Ca backbone of BtubA is 1.5 Å to α-tubulin and 2.7 Å to bacterial FtsZ. In particular, BtubAB have a three-domain architecture containing the two motor-protein-binding C-terminal helices that are conserved in tubulins, but missing from all other known prokaryotic FtsZ/tubulin-related proteins (Schlieper et al., 2005) (Figure 1.5b). Despite their bacterial origins, BtubAB are much more similar to eukaryotic tubulins than their prokaryotic homologues in both sequence and structure.

Given the striking similarities, it is fair to speculate that BtubAB might form microtubule-like structures. Several groups have seen filamentous structure *in vitro* via microscopy and light scattering (Schlieper et al., 2005; Sontag et al., 2005; Sontag et al., 2009). Using electron cryomicroscopy (cryoET) of *Prosthecobacter* bacteria and cryosectioning of *E. coli* cells overexpressing BtubAB, Pilhofer et al. observed 5-protofilament tubes formed by BtubAB, which were termed ‘bacterial microtubules’ (Pilhofer et al., 2011). However, due to the limitations of cryoEM and cryoET at the time, a high-resolution structure of BtubAB filament was not obtained.
Figure 1.5 BtubAB from Prosthecobacter are strikingly similar to eukaryotic αβ-tubulin.

(a) Schematic showing the btub operon from *Prosthecobacter dejonjeii*. This operon codes for three proteins, tubulin homologues BtubA and BtubB, and bacterial kinesin light chain (BKLC), which has been renamed as BtubC in this project.

(b) The crystal structure of BtubAB (PDB 2BTQ) showing a three-domain architecture very similar to eukaryotic tubulins. In this structure, the nucleotide binding pocket of BtubA contains a GDP, and the one in BtubB is empty. Colour code: blue – GTPase domain, yellow – helix H7, orange – intermediate domain, green – C-terminal domain.
1.4 Aims of this work

It is generally accepted that FtsZ and tubulins evolved from an ancient common ancestor (Erickson, 2007; Findeisen et al., 2014). Although the overall protein folds are very similar, there are very large evolutionary distances separating tubulin superfamily members. For instance, tubulins form high order polymers termed microtubules, whereas FtsZ functions probably as single homopolymeric protofilaments. In addition, tubulin superfamily members exhibit distinct dynamics and have extremely divergent cellular functions. The only features that have been conserved in the FtsZ/tubulin superfamily are the GTPase activity and protofilament formation, the longitudinal interaction which forms the active sites.

Study of *Prosthecobacter* BtubAB, with conserved eukaryotic features despite their bacterial origins, had already provided interesting insight into the evolutionary gap between prokaryotic and eukaryotic tubulins. However, not much information has become available about the higher-order assembly of BtubAB filaments. Previous attempts were limited by the technical difficulties associated with studying macromolecular complexes until the recent advances in cryoEM, which led to the so-called 'resolution revolution' where near-atomic resolution maps have been obtained for numerous biomolecules (Kühlbrandt, 2014). The recent invention of direct electron detectors and improvement in image analysis algorithms provided tools needed to investigate macromolecular structures such as BtubAB filaments at high resolution (McMullan et al., 2009; Scheres, 2012a), high enough for atomic interpretation.

Thus, the aim of this project was to obtain a high-resolution structure of the BtubAB filament in order to gain further insight into the evolution of FtsZ/tubulin superfamily proteins.
2 Materials and Methods

I would like to declare that this section is largely based on my first author paper (Deng et al., 2017), in which I have provided an extensive and detailed materials and methods section.

2.1 General methods

Preparation of electrocompetent cells

A freshly thawed stock of an appropriate Escherichia coli strain was streaked on a TYE agar plate and incubated at 37 °C overnight. A single colony from the plate was used to inoculate 50 ml 2xTY media and grown overnight at 37 °C. This culture was then used to inoculate 1 L 2xTY media at 1:100 v/v ratio and was left to grow at 37 °C until OD$_{600nm}$ reaches 0.4. The cells were cooled on ice for 2 hours before being centrifuged at 4000 g for 20 minutes at 4 °C. The supernatant was discarded and the pellets were resuspended gently with autoclaved double distilled water before centrifugation at 4000 g for 20 minutes at 4 °C. The washing-centrifugation cycle was repeated one more time. Next, the pellets were resuspended gently in ice cooled, sterile 10% v/v glycerol before centrifugation at 4000 g for 20 minutes at 4 °C. The final pellets were resuspended in ice cooled, sterile 10% v/v glycerol and aliquoted into 60 µl aliquots before flash frozen in liquid nitrogen and stored at -80 °C.

Polymerase chain reaction (PCR)

Genes of interests were amplified by PCR using Q5 Hot Start High-Fidelity 2x Master Mix (New England Biolabs) in this project. A typical PCR reaction consists of:

- 1x Q5 Hot Start High-Fidelity 2x Master Mix (New England Biolabs)
- 10 µM forward primer
- 10 µM reverse primer
- 1 µl DNA template

Top up to 25 µl final volume using MiliQ water. A standard PCR reaction cycle is as follows:

1. 30 seconds at 98 °C for initial denaturation.
2. 10 seconds at 98 °C for denaturation.
3. 30 seconds at 50 – 72 °C for annealing.
4. 30 seconds/kb at 72 °C for elongation.
5. Repeat step 2-4 30 times.
6. Store at 4 °C.

**Q5 mutagenesis**

Site-directed mutagenesis in this project were performed using Q5 site-directed mutagenesis kit (New England Biolabs). Vectors with gene of interests were amplified using PCR, using primers containing the desired mutations. The PCR reaction was then treated with KLD treatment typically consisting of:

- 1 µl PCR product
- 1x KLD reaction buffer (New England Biolabs)
- 1x KLD enzyme mix (New England Biolabs)
- MiliQ water up to 10 µl final reaction volume

The reaction was incubated at room temperature for 5 minutes and cleaned using QIAquick PCR Purification Kit (Qiagen) according to manufacturer's protocol before transformed into DH5α electrocompetent cells. The clean-up step with Qiagen kit was to remove the salt in the PCR and KLD reactions that would interfere with electroporation.

**Gibson assembly**

Gibson assembly was used to fuse amplified gene of interests with desired vector without using restriction enzyme. All the constructs used in this project were cloned using Gibson assembly (New England Biolabs). Typically, genes of interests and desired vectors were amplified using PCR with overlapping sequences according to the Gibson Assembly® protocol. Agarose gel electrophoresis was used to analyse the PCR reactions before assembly reaction that typically contains:

- 1x Gibson Assembly Master Mix (New England Biolabs)
- 1 µl amplified vector
- 1 µl amplified insert(s)
- MiliQ water to a final volume of 10 µl

Then the reaction was incubated in a thermocycler at 50 °C for 15 minutes before being diluted 3x using MiliQ water and transformed into DH5α electrocompetent cells. Again, the dilution with water was to lower salt concentrations to be compatible with electroporation.
**Agarose gel electrophoresis**

PCR reaction products were analysed by agarose gel electrophoresis before proceeding to the next step. Agarose gel was made by dissolving 1% w/v agarose (Biogene.com) in 1 x TBE buffer with 1:1000 v/v SYBR Safe (Invitrogen). Samples were mixed with DNA gel loading dye before loaded to the gel. The agarose gels were run at 100 V for 45 minutes before imaged with Gel Doc XR+ imaging system (Bio-rad).

**Transformation**

Purified plasmids or ligation mixture were transformed into electrocompetent *Escherichia coli* cells, C41(DE3) for expression and DH5α strain for plasmid amplification. Typically, 1 µl plasmids or cloning mixture was used to transform 50 µl electrocompetent cells added to a pre-chilled 2 mm electroporation cuvette on ice. A typical electroporation procedure had settings as 2.5 kV, 25 µF capacitance and 100 Ω resistance. After electroporation, the cells were immediately recovered by adding 200 µl SOB media and shaking for 30 minutes at 37 °C, before being spread onto TYE plates with appropriate antibiotics and incubated overnight at 37 °C.

**Plasmid purification**

If a cloning reaction mixture (Gibson assembly or mutagenesis) was used to transform *Escherichia coli* DH5α cells, after overnight incubation, a single colony was picked to inoculate 5 ml 2xTY media supplemented with appropriate antibiotic and left to grow overnight at 37 °C. The cells were harvested by centrifugation at 20,000 g for 2 minutes and the plasmids were extracted using a QIAprep Miniprep Kit (Qiagen) according to manufacturer’s protocol.

**SDS-PAGE**

Protein samples were often analysed using SDS-PAGE to separate proteins according to their molecular weights. Samples were mixed with protein loading dye and loaded onto an SDS-PAGE gel (10-20% gradient, Bio-Rad). The gel was typically run at 300 V for 30 minutes in 1x SDS-PAGE buffer before stained using Quick Coomassie stain (Generon) and imaged using a Gel Doc XR+ imaging system (Bio-rad).
2.2 Strains and plasmids

All strains and plasmids used in this work are listed in Appendix B.

2.3 Plasmid construction

To obtain a plasmid that codes for untagged proteins BtubA and BtubB from *Prosthecobacter dejongeii* DSM 12251 (PdBtubAB), a stop codon (TAA) was introduced before the C-terminal hexahistidine tag in the BtubAB bacterial T7 expression plasmid used previously (Schlieper et al., 2005) by Q5 mutagenesis (New England Biolabs). The plasmid maintains the intergenic region between BtubA and BtubB as present in the source genome. The gene *btubC* (previously known as the bacterial kinesin light chain BKLC) from *P. dejongeii* DSM 12251 was amplified from genomic DNA (based on the unpublished genome sequence, Igenbio, Chicago IL, USA) and was cloned into vector pET15b using Gibson Assembly (New England Biolabs). The resulting construct encoded an N-terminal hexahistidine tag followed by PdBtubC (MGSSHHHHHHSSGLPRGSH-1-257). The synthesised gene (Integrated DNA Technologies, IDT) *btubC* (previously known as the bacterial kinesin light chain BKLC) from *Prosthecobacter vanneervenii* (NCBI nucleotide database ID AM041148.3) was cloned into plasmid pHis17 using Gibson Assembly (New England Biolabs), yielding the construct encoding PvBtubC followed by a C-terminal hexahistidine tag (1-256-KLHHHHHHH).

2.4 Protein expression and purification

Untagged proteins BtubA and BtubB from *P. dejongeii* DSM 12251 were co-expressed and co-purified using a protocol adapted from previously described methods (Schlieper et al., 2005). The plasmid encoding untagged BtubAB was used to transform C41(DE3) *E. coli* cells (Lucigen) by electroporation. 60 ml 2xTY media supplemented with 100 µg/ml ampicillin were inoculated with a single colony from the plate, and were grown at 200 rpm, 37 ºC overnight. The culture was then used to inoculate 6 litres 2xTY media with 100 µg/ml ampicillin. After reaching an OD₆₀₀ of 0.6-1.0 at 200 rpm, 37 ºC, the expression was induced with 1 mM isopropyl β-D-thiogalactoside (IPTG) for 3 h at the same temperature, and the cells were harvested by centrifugation. The whole pellet was resuspended in 300 ml buffer A (20 mM Tris/HCl, 1 mM sodium azide, 5 mM magnesium acetate, pH 8.5) containing DNase I, RNase A (Sigma) and EDTA-free protease inhibitor tablets (Roche). Cells were disrupted at 25 kPSI in a cell disruptor (Constant Systems), and the lysate was cleared by centrifugation at 35,000
rpm in a 45 Ti rotor (Beckman) for 25 minutes at 4 °C. The cleared lysate was loaded onto a Q Sepharose HiLoad HP column (GE Healthcare), which was then washed with stepwise increases of NaCl in buffer A: 0, 100, 150, 250, 300 and 500 mM. The eluates were analysed by SDS-PAGE and the fractions containing BtubAB (mainly at 100 and 150 mM NaCl) were pooled and concentrated in Centriprep concentrators (30 kDa MWCO, Millipore). The concentrated proteins were loaded onto a Sephacryl S300 26/60 column (GE Healthcare) equilibrated with buffer B (20 mM Tris/HCl, 1 mM sodium azide, 1 mM EDTA, pH 7.5). Fractions containing pure BtubAB as determined by SDS-PAGE were concentrated in Centriprep concentrators (30 kDa MWCO, Millipore) to 20 mg/ml and flash frozen in liquid nitrogen.

*P. dejongei* BtubC (PdBtubC, MGSSHHHHHHSSGLLPRGSH-1-257) was expressed in C41(DE3) cells the same way as described above for BtubAB. For purification, the entire pellet was resuspended in 250 ml buffer C (50 mM Tris/HCl, 200 mM NaCl, 20 mM imidazole, 1 mM TCEP, 1 mM sodium azide, pH 8.0) supplemented with DNase I, RNase A (Sigma) and EDTA-free protease inhibitor tablets (Roche). Cells were disrupted at 25 kPSI in a cell disruptor (Constant Systems), and the lysate was cleared by centrifugation at 35,000 rpm in a 45 Ti rotor (Beckman) for 25 minutes at 4 °C. The cleared lysate was loaded onto a 5 ml HisTrap HP column (GE Healthcare), and was washed with stepwise increases of imidazole in buffer C: 0, 50, 100, 200, 500 and 1000 mM. The elution was analysed by SDS-PAGE and the fractions containing PdBtubC (mostly at 100 mM imidazole) were pooled and concentrated in Centriprep concentrators (30 kDa MWCO, Millipore). The concentrated proteins were loaded onto a Sephacryl S200 26/60 column (GE Healthcare) equilibrated with buffer D (50 mM Tris/HCl, 200 mM NaCl, 1 mM TCEP, 1 mM sodium azide, 1 mM EDTA, pH 8.0). Fractions containing pure PdBtubC as determined by SDS-PAGE were concentrated in Centriprep concentrators (30 kDa MWCO, Millipore) to 10 mg/ml and flash frozen in liquid nitrogen.

*P. vanneervenii* BtubC (PvBtubC) (1-256-KLHHHHHH) was also expressed in C41(DE3) cells as described above for BtubAB. The entire pellet was resuspended in 200 ml buffer E (50 mM CHES/NaOH, 500 mM NaCl, 20 mM Imidazole, 1 mM TCEP, 1 mM sodium azide, 5% (v/v) glycerol, pH 9.0) supplemented with DNase I, RNase A (Sigma) and EDTA-free protease inhibitor tablets (Roche). Cells were disrupted at 25 kPSI in a cell disruptor (Constant Systems), and the lysate was cleared by centrifugation at 35,000 rpm in a 45 Ti rotor (Beckman) for 25 minutes at 4 °C. The cleared lysate was loaded onto a 5 ml HisTrap HP column (GE Healthcare), and was washed with stepwise increases of imidazole in buffer E: 0,
30, 100, 200, 500 and 1000 mM. Eluates were analysed by SDS-PAGE and the fractions containing BtubC (mostly at 100 and 200 mM imidazole) were pooled and concentrated in Centriprep concentrators (30 kDa MWCO, Millipore). The concentrated proteins were loaded onto a Superdex 75 16/60 (GE Healthcare) equilibrated with buffer F (50 mM CHES/NaOH, 500 mM NaCl, 1 mM TCEP, 1 mM EDTA, 5% (v/v) glycerol, pH 9.0). Fractions containing pure BtubC as determined by SDS-PAGE were concentrated in Centriprep concentrators (30 kDa MWCO, Millipore) to 20 mg/ml and flash frozen in liquid nitrogen.

2.5 Pelleting assays

Each reaction contained 15 µM PdBtubAB (15 µM heterodimer: 15 µM BtubA and 15 µM BtubB co-expressed and co-purified), 0, 3, 15 or 60 µM PdBtubC in buffer containing 50 mM HEPES/NaOH (pH 7.5), 10 mM magnesium acetate, 100 mM potassium acetate and 1 mM Triethylene glycol diamine tetraacetic acid (EGTA) in a total volume of 120 µl. At the beginning, a 10 µl sample was taken from each reaction. For pre-spinning, all the reactions were centrifuged at 50,000 rpm for 10 min at 10 °C in a TLA100 rotor (Beckman), and a 10 µl sample was taken from the supernatant of each reaction. The supernatants after the pre-spin were used in subsequent pelleting assays, induced by the addition of 1 mM GTP and after 10-minute incubation at room temperature. The reactions were then centrifuged at 50,000 rpm for 20 min at 10 °C in a TLA100 rotor (Beckman). The supernatant from each reaction was removed and a sample of 10 µl was taken. Each pellet was resuspended in 100 µl 0.1% (w/v) SDS solution, and a sample of 10 µl was taken. All the 10 µl samples were analysed by SDS-PAGE after adding the same amount of loading buffer.

2.6 Negative stain electron microscopy sample preparation and imaging

PdBtubAB was polymerised in various buffers and each reaction was assessed by negative stain electron microscopy to select the condition that yield most number of long and straight filament, which would be subsequently used in cryoEM and cryoET. In total, seven buffer systems have been tested:

1. 50 mM HEPES/NaOH (pH 7.7), 5 mM magnesium acetate, 350 mM potassium acetate, 1 mM EGTA.
2. 50 mM HEPES/NaOH (pH 7.7), 5 mM magnesium acetate, 200 mM potassium acetate, 1 mM EGTA.
3. 50 mM HEPES/NaOH (pH 7.7), 5 mM magnesium acetate, 100 mM potassium acetate, 1 mM EGTA.
4. 80 mM PIPES/NaOH (pH 6.9), 5 mM magnesium chloride, 150 mM potassium chloride, 1 mM EDTA/EGTA 1:1 v/v mixture.
5. 100 mM PIPES/NaOH (pH 7.0), 5 mM magnesium chloride, 100 mM potassium chloride, 1 mM EDTA/EGTA 1:1 v/v mixture.
6. 100 mM PIPES/NaOH (pH 7.0), 5 mM magnesium chloride, 200 mM potassium chloride, 1 mM EDTA/EGTA 1:1 v/v mixture.
7. 100 mM PIPES/NaOH (pH 7.0), 5 mM magnesium chloride, 300 mM potassium chloride, 1 mM EDTA/EGTA 1:1 v/v mixture.

For filament assembly, 10 µM BtubAB (heterodimer: 40 µM BtubA and 40 µM BtubB co-expressed and co-purified) was added to each buffer system, and the polymerisation was induced by the addition of 1 mM GTP. Each reaction was incubated at room temperature for 30 minutes before a sample was taken for negative stain electron microscopy analysis.

For negative stain, 2 µl of the polymerisation reaction was added to a freshly glow-discharged CF400-Cu 400 mesh carbon film grid (Electron Microscopy Sciences), excess solution was blotted after 30 seconds of incubation time. The grid was quickly washed in a 2 µl droplet of 2% uranyl acetate solution before transferred to another 2 µl droplet of 2% uranyl acetate solution on a piece of Parafilm (Bemis). Excess uranyl acetate solution was blotted away after 30 seconds of staining time. The grids are imaged with an FEI Tecnai T12 electron microscope operating at 120 kV.

2.7 CryoET sample preparation and imaging

PdBtubAB was polymerised before being deposited on EM grids. The polymerisation reaction contained 40 µM BtubAB (heterodimer: 40 µM BtubA and 40 µM BtubB co-expressed and co-purified) in buffer containing 50 mM HEPES/NaOH (pH 7.7), 5 mM magnesium acetate, 100 mM potassium acetate and 1 mM EGTA in a total volume of 50 µl. The reaction was induced by the addition of 1 mM GTP and was left for 10 minutes to incubate at room temperature. Protein A conjugated with 10 nm colloidal gold (Cell Microscopy Centre, Utrecht University, Netherlands) was added to the polymerisation reaction at a ratio of 3:7 (gold:reaction, v/v) before 2.5 µl of the reaction was pipetted onto freshly glow-discharged Cu/Rh Quantifoil R2/2 200 mesh holey carbon grids (Quantifoil). The grids were blotted for 1.5 s with a blotting force of -15, a drain time of 0.5 s, and were flash frozen in liquid-nitrogen-
cooled liquid ethane using an FEI Vitrobot Mark IV (FEI). The Vitrobot chamber was set to 10 °C and 100 % relative humidity. Electron cryotomographic data were collected on an FEI Krios microscope operated at 300 kV, and equipped with a Quantum energy filter (Gatan), which was set 20 eV slit width around the zero loss peak. Data were acquired using SerialEM software on a K2 direct electron detector in counting mode at a calibrated pixel size of 2.17 Å. Tilt series were typically collected ± 60° with 3° tilt increment at 2.5–3 µm underfocus with a combined dose of about 80 e/Å² over the entire tilt series.

2.8 CryoET data analysis and subtomogram averaging

Tilt series were aligned and reconstructed in three dimensions using the IMOD package (Kremer et al., 1996). Subtomograms were picked along BtubAB filaments also using IMOD package. Each subtomogram was extracted as a 3D volume using RELION (Scheres, 2012a), with a box size of 160 x 160 x 160 voxels. CTF estimation for each sub-tomogram was conducted with CTFFIND (Mindell and Grigorieff, 2003). Finally, 3338 particles (volumes) were used for RELION subtomogram averaging refinement (Bharat et al., 2015). The resulting model was used to determine the helical parameters of BtubAB filaments by visualisation in UCSF CHIMERA (Pettersen et al., 2004), and the refined model was also used as the initial reference for the high resolution BtubAB filament helical reconstruction.

2.9 CryoEM sample preparation and imaging

BtubAB was polymerised and grids for cryoEM with helical reconstruction were prepared in the same way as for cryoET (above), except that fiducial gold was omitted. For BtubABC cryo-EM sample preparation, 40 µM PdBtubAB (40 µM heterodimer, 40 µM BtubA and 40 µM BtubB co-expressed and co-purified, untagged) and 45 µM PdBtubC (tagged) were polymerised in 50 µl buffer containing 50 mM HEPES/NaOH (pH 8.0), 5 mM magnesium acetate, 100 mM potassium acetate, 200 mM NaCl, 1 mM EGTA and 1 mM GTP. After 10 minutes incubation at room temperature, the polymerisation reaction was gently mixed and 2 µl of the reaction was pipetted onto freshly glow-discharged Quantifoil Au R2/2 holey carbon 200 mesh grids (Quantifoil). The grids were blotted for 3.5 s with a blotting force of -15, a drain time of 0.5 s, and were flash frozen in liquid-nitrogen-cooled liquid ethane using an FEI Vitrobot Mark IV (FEI). The Vitrobot chamber was set to 10 °C and 100 % humidity. CryoEM data collection for both BtubAB and BtubABC filaments followed the same protocol. Micrographs of BtubAB or BtubABC filaments were collected on a FEI Tecnai G2 Polara
microscope operating at 300 kV. Data were acquired on a Falcon III direct electron detector prototype in integrating mode at 30 frames per second and at a calibrated pixel size of 1.34 Å and a total dose of around 40 e/Å² using the automated acquisition software EPU (FEI). Images were collected at 1.5–3.0 μm underfocus over 1.5 seconds.

2.10 Helical reconstruction of BtubAB and BtubABC filaments, model building and refinement

For BtubAB filament reconstruction, 1394 movies were used from two separate 24 hour microscope sessions. Movie frames were combined into images after initial whole-image motion correction with MOTIONCORR (Li et al., 2013) and CTF parameters were estimated using GCTF (Zhang, 2016). All further processing was performed with RELION (Scheres, 2012a; Scheres, 2012b; He and Scheres, 2017). A few hundred helical segments were picked manually in boxes of 280 pixels and 2D classified. A few good classes were selected for automatic template-based autopicking in RELION with the helical option enabled so that the filaments were picked reliably. Overlapping segments 40 Å apart were picked. The autopicking algorithm also tracked which filaments individual particles were derived from, in order to be able to create two half sets for Fourier Shell Correlation (FSC) determination that separate only complete filaments. This yielded ~300,000 particles, which were assessed by one round of 2D classification, yielding 263,972 particles. Using rough pseudo-helical parameters of the BtubAB filaments as derived manually from the cryoET reconstruction (averaging BtubA and BtubB and ignoring the seam) of twist = -90 ° and rise = 10 Å a 3D refinement with helical averaging was performed. The cryoET reconstruction was used as the reference structure, filtered to 30 Å. The helical parameters were refined in RELION to twist = -90.7° and rise = 10.4 Å. No mask was used during 3D refinement. In postprocessing, the resulting two half maps were used to assess final resolution by gold standard FSC, combined, masked, and a sharpening B-factors was applied, yielding a final map at 4.2 Å resolution (FSC 0.143 criterion).

For BtubABC filament reconstruction, 6105 movies were used from four separate 24 hour microscope sessions. As above, whole image motion correction and CTF estimation were performed and initial classes for helical autopicking in RELION were produced. The same box size of 280 pixels was used. Autopicking yielded 345,777 particles, this time picked 80 Å apart. 2D classification was used to select the best particles from these, yielding 257,656 particles that were used in 3D refinement. The initial helical parameters were twist = -5.6 ° (-0.7° times
4 times 2) and rise = 80.0 Å (~ 40 Å times 2) that describe symmetry along one protofilament only, not averaging BtubA and BtubB and preserving the seam, but not using all symmetry available in the structure since the four protofilaments were not averaged. Helical parameters refined to twist = -5.54° and rise = 79.31 Å. Subsequent movie refinement in RELION as well as particle polishing yielded particles with higher signal to noise ratios since the final reconstruction and postprocessing (masking central 30 % of the map for determination of overall parameters) produced the highest resolution map at 3.6 Å, as assessed by the gold standard FSC procedure implemented in RELION (0.143 FSC criterion). For atomic fitting and refinement, a central portion of the map, covering BtubAB (not BtubC because of significantly weaker density) of the cryoEM filament density was cut out using REFMAC and fitted with previous crystal structures of BtubA and BtubB (PDB ID 2BTQ). Assignment of BtubAB to their corresponding parts of the density map was done through manual inspection of side chain densities and was very clear. The atomic model was manually adjusted with MAIN (Turk, 2013) and subsequent refinement against the cryoEM density was performed with REFMAC (Murshudov et al., 1997) in reciprocal space after back-transforming the cut out density into structure factors (REFMAC SFCALC mode) (Brown et al., 2015). Model quality was assessed using the standard R-factor and further in MOLPROBITY (Davis et al., 2007) for stereochemical plausibility. Statistics are summarised in Appendix A. To produce an atomic model of the entire BtubABC filament, four BtubAB heterodimers and the PvBtubC crystal structure (below) were fitted into the cryoEM densities of the four protofilaments by using CHIMERA (Pettersen et al., 2004) and then expanding helical symmetry (twist = -5.54° and rise = 79.31 Å).

2.11 BtubC crystallisation, data collection and structure determination

Initial crystallisation conditions for BtubC from Prosthecobacter vanneervenii (PvBtubC) were found using our in-house nanolitre crystallisation facility (Stock et al., 2005) using 20 mg/ml BtubC in 50 mM CHES/NaOH, 500 mM NaCl, 1 mM TCEP, 5% glycerol, pH 9.0 and 100 + 100 nl sitting drops in MRC crystallisation plates at room temperature. After optimisation, the best crystals were obtained in drops containing 5–8.1 % w/v PEG 8000, 5–8.1 % w/v PEG 1000, 200 mM lithium sulphate, 100 mM Tris/acetic acid at pH 8.5. Crystals were cryoprotected using 30 % v/v glycerol and flash frozen in liquid nitrogen for data collection. A native dataset to 2.5 Å was collected on beamline ID29 at the ESRF synchrotron (Grenoble, France). For phasing, a mercury derivative was prepared by soaking BtubC crystals.
in 5 % w/v PEG 8000, 5 % w/v PEG 1000, 420 mM lithium sulphate, 100 mM Tris/acetic acid, 600 mM NaCl, pH 8.5 supplemented with 0.5 mM thiomersal (EMTS) for 2 hours with subsequent cryoprotection (30% v/v glycerol) and freezing in liquid nitrogen for data collection. A derivative dataset was collected to 2.8 Å resolution on beamline ID30B at the ESRF synchrotron (Grenoble, France). The structure was solved by SAD (single anomalous dispersion) with CRANK2 (Pannu et al., 2011), leading to a rough model describing two BtubC chains in the asymmetric unit of the crystals. The initial model was refined against the native data at 2.5 Å resolution in REFMAC (Murshudov et al., 1997) and the resulting phases were used for further rounds of automated model building with BUCCANEER (Cowtan, 2006) and REFMAC refinement until a very complete model was obtained. The model was adjusted using MAIN (Turk, 2013) and refined with REFMAC and PHENIX (Adams et al., 2002) in several rounds.

2.12 Coordinate and map depositions

Coordinates and structure factors of the *Prosthecobacter vanneervenii* BtubC crystal structure were deposited in the Protein Data Bank (PDB) with accession code 5O01. The BtubABC filament structure was deposited in the PDB with accession code 5O09 and the corresponding 3.6 Å cryoEM 3D map was deposited in the EM Data Bank with accession code EMD-3726.
3 Results

3.1 BtubAB form 4-protofilament hollow tubes by cryoET

CryoET and subsequent subtomogram averaging were used to obtain an unbiased low-resolution model for BtubAB, so that information such as protofilament number and helical parameters could be derived de novo. Consistent with previously published work, purified untagged BtubAB formed filaments in the presence of GTP that can be observed using negative stain electron microscopy (Figure 3.1). Several polymerisation conditions were tested, and the reaction containing 40 µM BtubAB (heterodimer: 40 µM BtubA and 40 µM BtubB co-expressed and co-purified) in buffer containing 50 mM HEPES/NaOH (pH 7.7), 5 mM magnesium acetate, 100 mM potassium acetate and 1 mM EGTA was selected for subsequent cryoEM and cryoET studies, as this condition yielded the largest number of long, straight and single filaments (Figure 3.1 Error! Reference source not found.b).

The sample containing BtubAB filaments assembled in vitro using the optimised condition was used in cryoET (Figure 3.2a). Tomograms were reconstructed from the tilt series, and subtomograms containing segments of BtubAB filaments in three-dimensions were aligned and averaged without using any references (Kremer et al., 1996; Bharat et al., 2015). The resulting map, despite its low resolution, unambiguously revealed a left-handed four-protofilament architecture, which is different from the five-protofilament model reported previously (Pilhofer et al., 2011) (Figure 3.2b-c). Helical parameters, the relationship that explains how individual subunits are arranged in a helix, were simply defined by rise and twist in this project following the definition in RELION for helical reconstruction (He and Scheres, 2017). By manual inspection of the subtomogram-averaged map for BtubAB filament, approximate helical parameters (-90° twist, 10 Å rise) were deduced, not distinguishing BtubA and BtubB, following a 1-start helix.
Figure 3.1 Purified untagged BtubAB forms filament \textit{in vitro}.

(a) SDS-PAGE analysis of untagged PdBtubAB purification.

(b) Negative stain electron micrograph of BtubAB filament assembly \textit{in vitro} in the optimised condition. The reaction contained 40 µM BtubA and 40 µM BtubB co-expressed and co-purified, 50 mM HEPES/NaOH (pH 7.7), 5 mM magnesium acetate, 100 mM potassium acetate, 1 mM EGTA with 30-minute incubation time at room temperature after 1 mM GTP was added.
Figure 3.2 CryoET and subtomogram averaging of BtubAB filaments assembled *in vitro* clearly showing a four-stranded architecture.

(a) A sample tomographic slice of BtubAB filaments assembled *in vitro*.

(b) Top view of the subtomogram averaged BtubAB map.

(c) Side views of the subtomogram averaged BtubAB map. Approximate helical parameters (-90° twist, 10 Å rise) were deduced by manual inspection of the subtomogram-averaged map for BtubAB filament.
3.2 CryoEM study of BtubAB filament and its limitation

Having obtained an unbiased initial low resolution model of the BtubAB filament and derived approximate helical parameters, I proceeded with the cryoEM study of BtubAB filaments and collected 1394 cryoEM movies (Figure 3.3a). Initial data processing steps including whole-image motion correction and CTF parameters estimation were performed using MOTIONCORR and GCTF, respectively (Li et al., 2013; Zhang, 2016), and all further processing was done in RELION (Scheres, 2012a). Initially, I manually picked a few hundred helical segments along the BtubAB filaments in motion-corrected micrographs and performed reference-free 2D classification. A few good classes were selected as template for the automated template-based autopicking algorithm in RELION. After extraction of picked particles, several rounds of 2D classification were used for analysis and particle clean-up before the particles from good classes were submitted to 3D refinement. 3D classification, which is a routine step in single-particle cryoEM data processing, was not used in this project since the filamentous assembly generally tended to be rather homogenous, and the cryoET study did not yield models with protofilament numbers other than 4. In agreement with the results from cryoET and subtomogram averaging, the reference-free 2D class averages are compatible with the 4-protofilament architecture (Figure 3.3b). The subtomogram averaged model was used as the initial reference during 3D refinement, and a final map of 4.2 Å resolution with refined helical parameters (-90.7° twist and 10.4 Å rise) was obtained after postprocessing (Figure 3.3c). As this symmetry operator averaged BtubA and BtubB map regions, not many details beyond the shape of alpha helices can be observed.

Given that αβ-tubulins are arranged in an alternating manner in microtubules, and BtubAB forms heterodimers before or during filament assembly, it is highly likely that BtubAB are also arranged in an alternating manner (Figure 3.4a). If that is the case, with the helical parameters obtained, BtubAB filament must also have a seam, meaning that after applying the twist and rise four times around the helix, the identity of the next protein subunits must change from BtubA to B or B to A. As discussed in the introduction, the presence of a seam is a structural hallmark of eukaryotic 13- and 14- protofilament microtubules. The surface where the lateral contacts are formed by the same type of subunits (A-A and B-B lateral interactions) is called the B-lattice, and the one formed by two different subunits (A-B lateral interactions) is called the A-lattice.
Because of the likely presence of the seam, to obtain a correct reconstruction and structure, only helical symmetry along one protofilament, or treating the filament as a 4-start helix, must be applied (twist ~ -5.6°, rise ~80 Å) and the reconstruction algorithm must be able to distinguish BtubA and BtubB. Given the extreme similarity between BtubA and B with a RMSD of 1.3 Å, an additional tag that only binds to one of the subunit, or every 8 nm heterodimer repeat along each protofilament, was needed (Figure 3.4b).
Figure 3.3 CryoEM studies of BtubAB filaments.

(a) A typical electron micrograph of BtubAB filaments assembled in vitro. 
(b) Some selected 2D classes of segments of BtubAB filaments, which are compatible with a four-stranded model but incompatible with a five-stranded model. 
(c) CryoEM map of BtubAB filaments at 4.2 Å resolution. Helical symmetry (-90.7° twist and 10.4 Å rise) was applied during 3D refinement. As this symmetry operators averaged BtubA and BtubB map regions, not many details beyond the shape of alpha helices were observed.
Figure 3.4 Schematic of BtubAB filaments.

(a) Scheme summarising the symmetry of BtubAB mini microtubules as deduced by cryoEM and cryoET: four protofilaments with alternating filaments arrange into a hollow tube. The approximate helical parameters as deduced from the cryoET map dictate that the resulting structure must have a seam where lateral interactions change from B-lattice (A-A and B-B) to A-lattice (A-B).

(b) To enable the reconstruction algorithm to be able to distinguish BtubA and BtubB, an additional tag that only binds to one of the subunit, or every 8 nm heterodimer repeat along each protofilament, was needed.
3.3 BtubC contains a TPR fold and binds to BtubAB filaments in a stoichiometric fashion

During the search for a feature that binds along the BtubAB filament every 8 nm, my attention turned to the third gene in the btub operon, bacterial kinesin light chain (bklc), in Prosthecobacter genomes (Jenkins et al., 2002). Bklc was renamed during this study as btubC, as explained below.

I first cloned, expressed and purified Prosthecobacter dejongeii BtubC (PdBtubC), and added different amounts in pelleting assays to investigate whether BtubC binds to BtubAB filaments. In each reaction, 0, 0.25, 1 or 4 times molar excess BtubC was added to fixed amounts of BtubAB proteins. A pre-spin was performed to remove possible protein aggregation. Afterwards, GTP was added to allow filament polymerisation before the final centrifugation. If BtubC binds to BtubAB filaments with a sufficient affinity, it will end up in the pellet with the much heavier polymers, whereas un-polymerised BtubAB and excessive Btub should stay in the supernatant. The results were analysed by SDS-PAGE, which showed that BtubC binds to BtubAB filaments in a 1:1:1 ratio, as adding more BtubC beyond the number of BtubAB heterodimers in the filaments did not significantly increase the amount of BtubC that spun down (Figure 3.5). The pelleting assays suggested that BtubC may indeed be a feature that binds every 8 nm along the BtubAB filament, in other words, only binds once to each heterodimer.

Sequence based searches using BtubC reveal homology to the tetratricopeptide repeat (TPR repeat) domain of eukaryotic kinesin light chain but no structure is available in the Protein Data Bank. In order to obtain a structure, I initially used PdBtubC to set up crystallisation trials. Since no crystals were obtained after numerous attempts, I switched to work on BtubC from Prosthecobacter vanneervenii (PvBtubC), which has 76% sequence identity to PdBtubC. Microcrystals were found in several conditions in the Clear Strategy™ Screen (Molecular Dimensions) containing 10% PEG 8000, 10% w/v PEG 1000, 200 mM lithium sulphate, 100 mM Tris/acetic acid at pH 8.5/7.5. After optimisation, the best PdBtubC crystals were obtained in drops with 5–8.1% w/v PEG 8000, 5–8.1% w/v PEG 1000, 200 mM lithium sulphate, 100 mM Tris/acetic acid at pH 8.5 (Figure 3.6a). Crystals were cryoprotected using 30% v/v glycerol and flash frozen in liquid nitrogen for data collection (Figure 3.6b). Phasing information was obtained by single-wavelength anomalous dispersion (SAD) after soaking PvBtubC crystals with thiomersal (EMTS) for 2 hours. Eventually, the crystal structure of PvBtubC was solved to 2.5 Å resolution, showing the expected curved TPR repeat fold,
containing 11 helices and a C-terminal tail beyond residue 238 (Figure 3.7). Structural comparison revealed that BtubC is indeed similar to many TPR containing protein such as MamA as well as eukaryotic kinesin light chains (Zeytuni et al., 2012; Pernigo et al., 2013). Given that the *btubC* gene is not uniquely related to kinesin light chain, the word kinesin implies motor function, it is located in the same operon as *btubab*, and its ability to bind BtubAB filaments in a stoichiometric manner, I decided to rename this protein from its previous BKLC name to BtubC.
Figure 3.5 Pelleting assays showing stoichiometric binding of PdBtubC to PdBtubAB filaments. Four-fold excess of BtubC to BtubAB (molar, C to AB heterodimer) leads to roughly the same amount of BtubC in the pellet as compared to equimolar quantities. Excess BtubC protein remains in the supernatant and does not spin down, indicating a specific and stoichiometric interaction of BtubC with BtubAB heterodimers in the filaments.
Figure 3.6 X-ray crystallography study of PvBtubC.

(a) PvBtubC crystals from initial screen (left) and a sample PvBtubC crystal after optimisation (right).
(b) Diffraction pattern of PvBtubC crystal, showing spots up to 2.4 Å resolution.
Figure 3.7 Crystal structure of PvBtubC. PvBtubC contains a typically curved TPR repeat fold that is very similar to the TPR repeat domain in many other proteins such as eukaryotic kinesin light chain and MamA.
Encouraged by the pelleting assays suggesting BtubC binds to BtubAB filaments in a stoichiometric manner, I went on to collect cryoEM movies from samples containing BtubABC filaments (Figure 3.8a). In order to enable the reconstruction algorithm to distinguish BtubA from B, and thus obtain a high-resolution map, only helical symmetry along one protofilament (twist ~ -5.6°, rise ~80 Å) may be applied. Since the new symmetry is four-times less than the one used in BtubAB filament reconstruction, a much larger dataset (6105 movies over four separate 24-hour microscope session) was collected. This time, 345,777 particles were autopicked along BtubABC filaments with an inter-box distance of 80 Å, the approximate size of a heterodimer, after initial pre-processing with MOTIONCORR and GCTF (Li et al., 2013; Zhang, 2016). As before, several rounds of 2D classification were used to assess the data and select the best particles for subsequent 3D refinement. Already from the results of 2D classification, it became obvious that the added BtubC binds to the BtubAB filaments every 8 nm, or to each BtubAB heterodimer, confirming that BtubC is the feature that I was looking for (Figure 3.8b). With the help of BtubC, after iterative real space helical reconstruction in RELION with refined helical symmetry (twist -5.5°, rise 79.3 Å), a final map of the BtubABC filament was obtained at 3.6 Å resolution according to gold standard FSC (Henderson et al., 2012; He and Scheres, 2017) (Figure 3.8c, Figure 3.9d).

The high resolution and good quality of the map enabled me to assign BtubA and BtubB to their corresponding map regions by manual inspection of parts of the density where BtubA and B have very different side chains (Figure 3.9a-c). The crystal structure of BtubAB heterodimer (PDB 2BTQ) was fitted and refined against the best part of the filament density. Although the map revealed lots of details at near-atomic level for BtubAB, the density was less good for BtubC, probably due to a combination of partial occupancy and flexibility due to being at the surface of the filament. In addition, although my crystal structure of BtubC is from a different organism (Pv not Pd, where BtubAB were from), PvBtubC was fitted with high confidence as a rigid body against the PdBtubABC map (Figure 3.9c).

A complete atomic model of the BtubABC filament was obtained by expanding the refined and fitted BtubABC model according to the refined helical parameters returned by RELION. BtubC mainly binds to BtubB at lateral interfaces, and the whole structure shows a four-stranded hollow tube with a gentle twist (Figure 3.10a-b). As expected, BtubA and B subunits are arranged in an alternating manner along single protofilaments. Since all four
protofilaments run in parallel to each other, the resulting structure is polar. One feature that was immediately noticed was the presence of a seam, a structural hallmark of microtubules. Among three lateral interfaces that connect the four protofilaments within the BtubABC filament, by analogy to microtubules, both A-A and B-B contacts form the B-lattice, and one interface, or the seam, form the A-lattice with A-B being adjacent to each other (Figure 3.10b). A closer look at what holds the protofilaments together revealed a structure equivalent to the M-loop, the loop that is involved in lateral interactions in microtubules, and it is exclusively responsible for the lateral interaction in BtubAB filament through an interaction similar to that seen in microtubules. A key interaction is the insertion of F287 from BtubA to the hydrophobic pocket of the adjacent subunit (Figure 3.10c-d). Density for the M-loop in BtubA is observed clearly in both A-A interface and A-B interface, however, the corresponding density in BtubB is not ordered at interfaces in our cryoEM map and may not form a contact. Hence, the protofilaments in BtubAB filaments may solely be bound together through contacts made by BtubA.

The longitudinal contacts of BtubAB filaments consist of nucleotide binding pockets at both inter- and intra-dimer surfaces. In contrast to tubulin structures where the intra-dimer interface has a GTP at the α-tubulin nucleotide pocket, and GDP at the inter-dimer one, both BtubA and BtubB have a GDP in their nucleotide binding pockets, suggesting the added GTP being hydrolysed during or after filament formation (Figure 3.11).

BtubC’s binding site is formed predominantly by BtubB in-between protofilaments at both B-lattices and A-lattices, unlike MAP such as Mal3p that preferentially binds to the microtubule at the seam (Sandblad et al., 2006). The N-terminal helices of BtubC make contact with BtubB with some additional contacts to the other subunits surrounding the interface (Figure 3.12). Since the BtubC structure was solved from P. vanneervenii rather than P. dejongeii, PvBtubC was fitted only as a rigid body and detailed binding to BtubAB filaments is not discussed.
Figure 3.8 CryoEM study of BtubABC filament.

(a) A typical electron micrograph of BtubABC filaments assembled in vitro. PdBtubC is not seen easily in the micrograph due to its small size (~ 30 kDa).

(b) Selected 2D classes of BtubABC filaments, showing clearly that BtubC binds to the BtubAB filament every 8 nm, or every heterodimer. Arrows indicate where BtubC is located.

(c) Final cryoEM map of BtubABC filament obtained through iterative real space helical reconstruction in RELION to 3.6 Å resolution according to gold standard FSC.
Figure 3.9 The map of the BtubABC filament is of very good quality, allowing the assignment of BtubA and B to their corresponding map regions by manual inspection of side chain densities.

(a) Assignment of BtubA and BtubB to their corresponding map regions, which was done based on the excellent quality of the 3.6 Å resolution map that enabled the unambiguous assignment through inspection of many map regions. Panels on the left and right show two such map regions, superimposed with both the BtubA and BtubB model from the crystal structure, fitted into the map with CHIMERA. BtubA map and model are in red, blue for BtubB.

(b) CryoEM map of BtubABC filament fitted with model of one asymmetric unit containing one BtubA (red), one B (blue) and one PvBtubC (green).

(c) CryoEM density map with fitted atomic model. Map quality can be gauged by the separation of β-sheet (right) and the appearance of many side chain densities (left).

(d) BtubABC cryoEM Fourier shell correlation (FSC) of the half maps and atomic model. The blue curve describes the FSC as calculated by RELION during postprocessing, comparing the two gold standard half maps that have been kept separate the entire procedure before they are combined. The red curve describes the FSC of the map used to refine BtubAB heterodimer (omitting BtubC, both half maps averaged), cut around the atomic model, against the final atomic model. Using the suggested resolution estimates at FSC 0.5 for the model vs map and FSC 0.143 for the two gold standard half maps, identical values of 3.6 Å were obtained.
a -5.5° twist

79.3 Å

B A C

seam

180°

B-lattice

b top view

BtubC

1 2 3 4

C BtubAB (seam) lateral interface

M-loop

d BtubAA lateral interface

M-loop
Figure 3.10 Architecture of BtubABC filament and lateral interaction.

(a) Left: BtubABC filament. PdBtubA is in red, PdBtubB in blue, PvBtubC in green. The view shown highlights the seam (A-lattice, A-B lateral contacts). Right: rotated by 180°, showing the B-lattice (A-A and B-B lateral contacts). Helical parameters: twist = -5.5°, rise = 79.3 Å.

(b) View along the filament axis showing the four protofilaments end-on and also the protruding BtubC subunits in green.

(c) Close-up of the only lateral contact holding the protofilaments together. The contact is very similar to the equivalent contact in eukaryotic microtubules, with the M-loop of BtubA (residues ~ 280 - 290) reaching over to contact residues on the neighbouring protofilament (around 90 and also 55-65). Shown here is the seam contact (BtubA M-loop reaching over to BtubB). Note that the M-loop in BtubB is not ordered.

(d) Same as in (c), but B-lattice contact, BtubA contacting BtubA.
Figure 3.11 Longitudinal interactions in BtubAB filament. The longitudinal interactions in BtubAB filaments form and contain the nucleotide binding pockets. In contrast to microtubules, where α-tubulin binds a non-hydrolysable GTP and β-tubulin binds GDP, both BtubA and BtubB contain a GDP in the filament according to our cryoEM map, suggesting added GTP being hydrolysed during filament assembly in vitro.
Figure 3.12 BtubC mainly binds to BtubB through its N-terminal helices.
4 Discussion

4.1 Investigation of BtubAB filament dynamics by Gero Fink

Despite the difference in diameter, the structure of Prosthecobacter BtubAB filaments exhibits striking similarities to eukaryotic microtubules with conserved longitudinal and lateral interactions, and the presence of a seam. Therefore, Gero Fink, a postdoc in Jan Löwe’s lab investigated the dynamic properties of BtubAB filaments by total internal reflection fluorescence microscopy (TIRF microscopy). TIRF experiments were performed by Gero Fink.

The experimental setup is shown in Figure 4.1a and c. Upon the addition of GTP, BtubAB grew into long, rigid filaments with directed movement of smaller filaments. The movement could be completely abolished by the addition of GTP-\(\gamma\)-S, a slowly-hydrolysable analogue of GTP. To investigate the dynamics in more details, Gero performed dual colour experiments, where magenta subunits were added first, followed by green subunits so that the green BtubAB were seeded on the older, magenta filaments (Figure 4.1a). In this experiment, both magenta and green subunits were supplied with GTP so they are both dynamic. The results highlighted that the BtubAB filaments were capable of treadmilling, that is, they appear to move across the surface by growing at one end (the plus end), and shrinking at the other end (minus end) (Figure 4.1b). It was observed that the filaments disassembled from the magenta (old) end only, marking the minus ends, whereas growth appeared overwhelmingly at the green ends, marking the plus ends.

When performing the dual-colour experiments, Gero noticed that occasionally some filaments would rapidly disappear from the plus ends, suggesting possible dynamic instability, a hallmark of eukaryotic microtubule dynamics (see Section 1.1.5) (Figure 4.1b). Dynamic instability, the stochastic switch between fast growth and rapid shrinkage, is a unique dynamic behaviour of microtubules, which enables microtubules to probe their surroundings until designated targets such as kinetochores have been attached. Although the exact mechanism underlying this dynamic property is unknown, it is generally believed that the rapid shrinkage, or catastrophe, is caused by the loss of GTP cap at the growing end, whereas the resumed growth, or rescue, is initiated by the catastrophic disassembly reaching the GTP island present within the microtubule or by other MAPs. To investigate whether BtubAB filaments exhibit dynamic instability, Gero performed another dual-colour TIRF experiment. This time, the magenta subunits added first were stabilised by GMPCPP, another slowly-hydrolysable GTP
an analogue, before green subunits were added with GTP (Figure 4.1c). Since the magenta seeds are stabilised, the dynamics of the green filaments extended from the now stable plus ends of the magenta seeds can be studied more extensively. The results clearly show rounds of catastrophe and rescue on the green extensions, with a growth rate (0.38 μm min$^{-1}$ μM$^{-1}$) similar to the growth rates reported for mammalian microtubules (0.2 - 0.33 μm min$^{-1}$ μM$^{-1}$) (Walker et al., 1988; Hyman et al., 1992) (Figure 4.1d). Gero further analysed the data by plotting the time between catastrophe events (catastrophe time) against the extension length at which catastrophe occurred in cumulative plots. As expected, catastrophe time of BtubAB filaments was not dependent on protein concentration, whereas catastrophe length varied with BtubAB concentration due to the change in growth velocities (Figure 4.1e-f). In addition, the shape of the distribution of BtubAB catastrophe times were best fitted with a gamma function, suggesting catastrophes were caused by more than one events, rather than a simple exponential decay, which would be the case if there was only one single event causing catastrophe (Figure 4.1g-h). This complex dynamic behaviour of BtubAB filaments closely resembles the dynamics of microtubules, in which catastrophes are also best described as multistep processes (Gardner et al., 2011).

Finally, since my structural data clearly shows the binding of BtubC on BtubAB filaments, Gero investigated if BtubC has any influence on the dynamic properties of BtubAB filaments. Even at sub-stoichiometric ratios of BtubC, BtubAB filaments exhibited reduced catastrophe and increased rescue frequencies, and the filaments were fully stabilised and grew continuously when BtubC was added to near stoichiometric amounts (Figure 4.2a). Growth velocities did not change much in the presence of BtubC, however, catastrophe time was increased (Figure 4.2b-c). Further analysis revealed that the multistep nature of BtubAB catastrophe was unchanged in the presence of BtubC, and the increase in catastrophe times (the time since the filament has been growing until catastrophe occurs) in the presence of BtubC was caused by the lower catastrophe rate parameters (Figure 4.2e-f). Overall, BtubC seems to stabilise BtubAB filaments through the combination of a reduced catastrophe frequency and an increased rescue frequency (Figure 4.2d).
catastrophe times at different BtubAB concentrations fitted to a gamma distribution function

parameters obtained from fits of a gamma distribution for three different BtubAB concentrations
Figure 4.1 BtubAB filaments show dynamic instability and treadmilling (Deng et al., 2017).

(a) Sketch of TIRF microscopy setup used for (b). Elongation of Atto655-labelled BtubAB filaments (GTP seeds, magenta) upon addition of Atto488-labelled BtubAB monomers, also in GTP.

(b) Image series showing treadmilling filaments (asterisks). Seeds grow at the plus ends (green extension) while the opposing minus ends shrink (shortening of magenta seeds). Filaments display catastrophic shortening and regrowth from the plus ends, only (arrowheads), indicating dynamic instability. The kymograph (right) highlights treadmilling and dynamic instability.

(c) Sketch of TIRF microscopy setup used for (d). Stable GMPCPP seeds (magenta) were elongated by green BtubAB in GTP.

(d) Kymographs of three separate BtubAB filaments showing dynamic instability from their plus ends. Extensions show a sawtooth pattern since filaments undergo several cycles of growth and very rapid shrinkage, called catastrophe.

(e) Graphs showing cumulative distributions of catastrophe times and lengths (indicated by bars in the middle kymographs in D) for three different concentrations (n=3 experiments). Catastrophe times do not change with concentration and catastrophe length increases with BtubAB concentration.

(f) Scatter plot of growth velocities versus BtubAB concentration (n=3, 2.45 µM: n = 2). As expected, growth rate increases with concentration. Filaments grow at 0.96 ± 0.18 µm min⁻¹, mean ± std.; 0.81 ± 0.16 µm min⁻¹ and 0.5 ± 0.1 µm min⁻¹ for 2.45 µM, 2 µM and 1.5 µM BtubAB respectively. The growth velocities correspond to the observed changes in catastrophe length (E), meaning catastrophe is independent of growth rate.

(g) Cumulative distribution of catastrophe times fitted to a gamma function (Methods). A multistep process where n>1 is compared to a single step process (n=1) obtained by fitting a gamma distribution describes catastrophe of BtubAB filaments best over a range of concentrations. The goodness of the fits for different n (magenta) is given as the mean R² value for each condition (colour coded) in the corresponding graph.

(h) BtubAB concentrations have only small effects of on mean lifetimes, number of steps and rate parameters that have to take place before catastrophe. Values are given above each data point and error bars represent 95% confidence intervals.
e  catastrophe times at different BtubC concentrations fitted to a gamma distribution function

f  parameters obtained from fits of a gamma distribution in the presence of BtubC and constant BtubAB concentrations
Figure 4.2 The effect of BtubC on dynamics of BtubAB filaments (Deng et al., 2017).

(a) Kymographs of BtubAB filaments emanating from stabilised seeds (as in Error! Reference source not found.) in the presence of varying amounts of BtubC. BtubC interferes with dynamic instability and inhibits it at stoichiometric amounts. Filaments grown in the presence of BtubC do not completely disassemble instead they start regrowth from the non-stabilised (green) GDP/GTP ends (arrows).

(b) Scatter plots of BtubC-dependent growth velocities. BtubC does not significantly alter filament growth (compare to Error! Reference source not found.). Filaments grow at $0.8 \pm 0.14$ µm min$^{-1}$, mean ± std. and $0.81 \pm 0.16$ µm min$^{-1}$ for 2 µM BtubAB and $0.44 \pm 0.08$ µm min$^{-1}$ and $0.5 \pm 0.1$ µm min$^{-1}$ at a ratio of BtubC to BtubAB of 1:9 and 1:12, respectively.

(c) Cumulative catastrophe time distributions of filaments at different BtubC concentrations. Filament lifetimes are increased in the presence of BtubC, which is caused by a BtubC concentration dependent reduction in the rate parameter at which catastrophe occurs (e-f).

(d) Effect of BtubC on rescue/catastrophe ratios. In the presence of BtubC at ratios 1:9 and 1:11, 92 % ± 8.1 % of catastrophes are followed by rescue whereas without BtubC, this occurs in about 11 % ± 4.7 % of observed events. In summary, BtubAB mini microtubules are dynamically unstable and BtubC stabilises BtubAB filaments by increasing rescue events and reducing catastrophe frequency.

(e) BtubC has an effect on BtubAB lifetimes. Increasing molecule ratios (BtubC to BtubAB) shifts the distribution to older filaments, by reducing the rate parameters at which the steps have to occur to induce a catastrophe. Values are again given above each data point and error bars represent 95 % confidence intervals.

(f) Experimentally obtained distribution of catastrophe times for different ratios of BtubC to BtubAB, for which the concentration of BtubAB was kept constant at 2 µM. Catastrophe is still a multistep process with n=2 steps as seen for BtubAB alone. Again, fits for different n (magenta) are evaluated by $R^2$ values for each condition (colour coded).
4.2 BtubAB and a comparison to microtubules

Early bioinformatics studies and the crystal structure of BtubAB already revealed striking similarities in both structure and sequence between BtubAB and eukaryotic tubulins, despite their bacterial origin (Jenkins et al., 2002; Schlieper et al., 2005). BtubA has 38% sequence identity and 1.5 Å RMSD of its Cα backbone to α-tubulin, in comparison to 22% sequence identity and 2.7 Å RMSD to FtsZ. Remarkably, BtubAB contain regions equivalent to the two C-terminal motor protein-binding α helices in eukaryotic tubulins, which are missing in all other prokaryotic FtsZ/tubulin-related proteins. Mutagenic studies at both inter-dimer and intra-dimer interfaces showed that BtubAB form heterodimers first before polymerising to higher order filaments in the presence of GTP (Sontag et al., 2005; Sontag et al., 2009). Nevertheless, differences exist between BtubAB and αβ-tubulins. Eukaryotic tubulins follow a complex pathway for their folding, whereas BtubAB have been shown to display chaperone independent folding, similar to FtsZ (Lewis et al., 1997; Andreu et al., 2002; Schlieper et al., 2005). In addition, there is little differentiation between BtubA and BtubB, while eukaryotic tubulins evolved to be more specialised proteins with different functions. For instance, β-tubulin has a taxol binding pocket, which α-tubulin lacks, and αβ-tubulin polymerise into microtubules whereas γ-tubulin forms a ring complex that microtubule extends from (Oakley et al., 2015).

Although BtubAB were known to form filaments almost since their discovery, a high resolution structure of the filament was not available till now mainly due to the limitation of technology, and the absence of a marker for the heterodimer along the filaments. The recent advances in cryoEM have contributed to a paradigm shift in structural biology, enabling the study of much larger biomolecules at previously unthinkable resolutions. Inspired by the cryoET and cryo-sectioning study by Grant Jensen’s lab showing BtubA to form five-protofilament tubes in vivo, we decided to investigate the structure of BtubAB filaments to near-atomic resolution (Pilhofer et al., 2011).

My first structure of BtubAB filament was obtained by cryoET and subtomogram averaging. Despite the low resolution, this reference-free map revealed crucial information, including the protofilament number and rough helical parameters, as well as serving as an initial reference for subsequent cryoEM studies. The information obtained from the cryoET study led to the second structure in this project, the cryoEM structure of BtubAB filaments at 4.2 Å resolution. This map corresponded well with the subtomogram averaged model, showing
four protofilaments with traceable α helices. However, no further details could be obtained due

to the presence of the seam, which led to the averaging of BtubA and BtubB map regions, a

well-known issue in the microtubule field. Essentially, the seam breaks the overall helical

symmetry and an additional protein or tag is required for the reconstruction algorithm to
distinguish between BtubA and BtubB. In fact, all microtubule structures available so far are
in complex with either a MAP or a kinesin. The search for this marker or tag led to the discovery

of BtubC binding to BtubAB filaments at stoichiometric ratios. Previously, BtubC was termed

BKLC, due to its good sequence homology to eukaryotic kinesin light chain. However, my

crystal structure of PvBtubC revealed structural similarity not only to kinesin light chain, but
also to many other TPR repeat containing proteins such as MamA. Together with its ability to
bind BtubAB from the same operon, I decided to rename BtubC from its original name, BKLC.

After the discovery of this binding partner, the structure of BtubABC filament was obtained at
3.6 Å resolution. This map was of excellent quality, which enabled the fitting and all-atom
refinement of previous BtubAB crystal structures. PvBtubC was fitted as a rigid body since the
density of BtubC was weak and my crystal structure was from a different organism.

A complete atomic model of the BtubABC filament was generated by expanding helical
symmetry. My data unambiguously shows a four-stranded architecture instead of five, as
previously reported (Pilhofer et al., 2011). First of all, Pilhofer et al.’s structure was of much
lower resolution, as it was obtained from cellular tomograms where the noise level was very
high, and the study was done prior to the resolution revolution in cryoEM field. This difference
could also be explained by the possibility of BtubAB forming a range of filaments (like
microtubules), and my in vitro condition happen to yield four-stranded tubes whereas in cells,
BtubAB might exist as five-stranded filaments. Vastly improved cellular tomography,
followed by subtomogram averaging of these filaments in Prosthecobacter cells would offer
the ultimate solution, however, this would require an enormous amount of time on a state-of-
the-art electron microscope, given the low occurrence of the filaments and longer acquisition
time required for tomography data collection, rendering it almost infeasible.

Nevertheless, my in vitro structure of the BtubAB filament, again, showed remarkable
similarity to microtubules (Figure 4.3). BtubAB filaments consist of four parallel
protofilaments, resulting in an overall polar structure. The longitudinal interfaces contain the
nucleotide binding pockets, except both BtubA and BtubB contain GDP, whereas α-tubulin
has a non-hydrolysable GTP. Just like microtubule, the lateral interactions in BtubAB filaments
are maintained by the so-called M-loop. The lateral interface is relatively small, with the key
interaction being the insertion of one single amino acid into the hydrophobic binding pocket of the adjacent subunit. Presumably, overall stable filaments are formed nonetheless due to the avidity effect when many weak binding sites are combined. The most exciting similarity between Btub filament and microtubule to me, perhaps, is the presence of a seam. Although the biological significance of the seam is not fully understood in the microtubule field, it is certainly a peculiar hallmark of microtubules. Some people argue that the presence of the seam provides an alternative binding surface on the microtubule. Indeed, DCX, a microtubule stabilising MAP, only binds to 13-protofilament microtubules at the B-lattice but not at the seam, whereas the *S. pombe* EB1 homologue Mal3p stabilises microtubules by its preferential binding to the A-lattice (at the seam) (Sandblad et al., 2006; Fourniol et al., 2010). The binding of BtubC to the BtubAB filament resembles the regulatory functions of MAPs on microtubule dynamics, and our data shows that BtubC binds to the filament at both A- and B-lattice.

The biggest and most obvious structural difference between BtubAB filament and microtubules is the drastically smaller diameter, which might be more suitable in the much smaller bacterial cells. The BtubAB filament contains just four protofilaments, in contrast to microtubules with normally 13 protofilaments. The difference is mainly accommodated by changes in the M-loop. Although the M-loop is present in both BtubAB and tubulins, the amino acid composition and the actual position of the loop are very different, resulting in a twist of ~90° between adjacent subunits in BtubAB filaments, as oppose to ~60° in microtubules (Figure 4.4).

Despite the much smaller size, Gero Fink’s TIRF experiments showed that BtubAB filaments retain the complex dynamic behaviours of microtubules, exhibiting both treadmilling and dynamic instability, another hallmark of microtubules. Remarkably, Gero’s detailed analysis revealed that BtubAB filaments not only have dynamic instability, but also with similar growth rate and a complex ageing cap very similar to microtubules. The catastrophe times were best fitted by a gamma distribution, suggesting that catastrophe in BtubAB filaments is a multistep process, just as microtubule catastrophes. Lastly, the regulation of BtubAB filaments by BtubC resembles the relationship between various MAPs and microtubules.

Taken all together, I think it is justified to conclude that BtubAB form bacterial mini microtubules.
Figure 4.3 Schematic comparison of BtubAB mini microtubules and eukaryotic 13-protofilament microtubules. Apart from previously known similarity of the subunits themselves and the heterodimer, the two structures conserve overall polarity (chemically different ends because all subunits point in the same direction), the seam where the nature of lateral contacts between protofilaments change and alternating AB/αβ protofilaments. The M loop connecting protofilaments is also conserved. In contrast to microtubules, both BtubA and BtubB contain GDP in the polymerized and matured filament and, of course, the protofilament number is vastly different, leading to much thinner mini microtubules (and a small helical twist that is not depicted here).
BtubAB vs microtubule interprotofilament angle viewed along protofilament (filament) axes

~ 64.3°
(90° - 25.7°)

α-tubulin

outer microtubule wall

outer BtubAB wall

M-loops

BtubA / α-tubulin superimposed
(microtubule PDB ID 3JAT)

M-loops

protofilament

G42

S59

Rotated 90° around X-axis from A

protofilament

F287

R284

M-loops

protofilament

protofilament
Figure 4.4 Comparison of interprotofilament interfaces in microtubules and BtubAB mini microtubules.

(a) Superposition of BtubA (red) and α-tubulin (gray) when in 4-protofilament BtubAB mini microtubules and 14-protofilament microtubules (PDB ID code 3JAT). The protofilaments are rotated by ~26° almost perfectly around the M loop, the structure in both polymers that facilitates protofilament contacts.

(b) Detailed view of the same BtubAB/α-tubulin superposition showing the M loops. Although in roughly the same place, secondary structures are different and there is no obvious similarity of the contacts at the residue level. Divergent evolution probably changed the entire contact so to enable different inter-protofilament angles, leading to 4 instead of more than 10 protofilaments.

(c) This view is to (b), with a rotation of 90° around the x axis.
4.3 Possible origins of the btub operon and the evolution of FtsZ/tubulin-related proteins

Tubulins are among the most abundant proteins in eukaryotic cells, playing critical roles in biological processes such as chromosome segregation, cell motility and cellular transport. The prokaryotic tubulin homologue, FtsZ, is almost ubiquitous in bacteria and found in many archaea and is involved in cell division. Among all the FtsZ/tubulin-related proteins, the strongly conserved sequence and protein length in most eukaryotic α- and β-tubulins in comparison to γ- and other tubulin homologues suggest that α- and β-tubulin resulted from a relatively recent gene duplication event, presumably after the split of their ancestor from others, for example, FtsZ, γ- and ε-tubulin (Findeisen et al., 2014). In contrast to other FtsZ/tubulin-related proteins, which mainly function as homopolymers, α- and β-tubulin function cooperatively, they form a stable heterodimer first before assembling into the giant and hollow microtubule with sophisticated dynamic behaviours. Based on their supposed origins from a late gene duplication event and their cooperative functions, I will treat α- and β-tubulin as a single group of tubulins in this discussion.

The FtsZ/tubulin superfamily is a big group of proteins that spans through all three kingdoms of life, with αβ-tubulins being the best studied tubulins in the eukaryotic branch and FtsZs being the best studied protein the bacterial branch (archaeal FtsZ and CetZ are much less studied). It is generally accepted that both αβ-tubulins and FtsZ have evolved from a common ancestor, however, the evolutionary gap between them is vast and the only properties that are conserved are the GTPase activity and protofilament formation (Erickson, 2007; Löwe and Amos, 2009) (Figure 4.5). FtsZ has a two-domain architecture, missing the third tubulin domain, two C-terminal helices, which are involved in motor protein interactions. FtsZ probably functions as homopolymeric single protofilaments whereas αβ-tubulins assemble into higher-order microtubules. Cellular functions are also diverged. FtsZ forms the so-called ‘Z-ring’ at the division septum, and drives cell division by dragging the cell wall synthesis machinery along with treadmilling FtsZ filaments (Bisson-Filho et al., 2017; Wagstaff et al., 2017; Yang et al., 2017). In contrast, α- and β-tubulins work cooperatively to form microtubules which are involved in cellular transport, motility and chromosome segregation (Nogales, 2000). Interestingly, cytokinesis in eukaryotic cells today is carried out mostly by actin-based and ESCRT-based systems (Eggert et al., 2006; Schöneberg et al., 2017). It has been suggested that FtsZ may be an extremely ancient protein, due to the low occurrence of complex amino acids such as arginine, lysine, phenylalanine and histidine in its sequence.
Together with the wide spread of FtsZ/tubulin-related proteins across all domains of life, it is tempting to speculate that ancient FtsZ/tubulin was present in very early life forms probably involved in cell division, possibly before the split of prokaryotes and archaea. Surely, an FtsZ- or ESCRT-based membrane scission mechanism can facilitate cell division in an organised manner, but they may not be an absolute requirement for cell division. For example, L-form bacteria, a peptidoglycan cell wall deficient type of prokaryote, divide through a range of relatively disorganised membrane blebbing or vesiculation events, which provide a remarkably simple alternative mean for proliferation (Briers et al., 2012; Errington et al., 2016). Given that FtsZ is almost ubiquitous in prokaryotic kingdom and relatively wide spread among archaea, it is possible that an FtsZ-based cell division system was present in early life forms, while the ESCRT-based cytokinesis mechanism evolved roughly when eukaryotes branched out of the archaeal tree (as the ESCRT systems are present in Asgard superphylum) (Zaremba-Niedzwiedzka et al., 2017). As the eukaryotes evolved, the ESCRT-based cytokinesis was gradually complemented by the actin-based system to deal with the large size of the cell, and FtsZ/tubulin ancestor evolved, and split, to become the microtubule-forming αβ-tubulins and MTOC-forming γ-tubulin as known today. It is difficult to speculate whether ESCRT system replaced the FtsZ-based system in some early archaea or both systems evolved in parallel among early life forms, especially little is known about FtsZ/tubulin homologues in archaea, or archaea themselves in general.

Due to the large evolutionary gap between FtsZ and eukaryotic tubulins, the existence of BtubAB proteins, with so many conserved microtubule features despite their bacterial origins, is particularly interesting. The narrow distribution of the btub operon and the absence of other signature eukaryotic proteins in Prosthecobacter suggest the operon has likely originated from a horizontal gene transfer event into a Prosthecobacter ancestor from a eukaryotic organism (Schlieper et al., 2005; Staley et al., 2005). If this is the case, the next question would be, did the horizontal gene transfer occur before or after the gene duplication event in early eukaryotes that gave rise to α- and β-tubulin? I personally prefer the hypothesis that Prosthecobacter bacteria adopted two copies of tubulin genes from an eukaryotic ancestor, shortly after the gene duplication event but before they became specialised as the α- and β-tubulin known today. Findeisen et al. performed whole-genome-based analysis on a broad range of eukaryotic and prokaryotic tubulins, and they observed no trees in which all BtubAB form a clade grouping sister to the combined eukaryotic α- and β-tubulin, suggesting the horizontal gene transfer of BtubAB occurred after the split of eukaryotic α- and β-tubulin
In addition, the striking similarities between BtubAB filaments and microtubules at both structural and dynamic levels demonstrated by data from Gero and myself further support this hypothesis. Even within eukaryotic the world, presumed tubulin duplication events have resulted in diverged subfamilies. For example, γ-tubulin is the building block of MTOCs, where microtubules grow from, and δ-tubulin from *Chlamydomonas* is required for assembly of basal bodies of its flagella (Dutcher and Trabuco, 1998; Wu and Akhmanova, 2017). I would suggest that the probability of *Prosthecobacter* adopting one copy of tubulin, which duplicated and evolved into mini microtubule forming heterodimers, with conserved structural features and dynamic behaviours to microtubules, is extremely low. It seems more likely that the horizontal gene transfer occurred after the split of α- and β-tubulin, and they were already capable of forming microtubules, or microtubule-like structures by that time. However, since the similarity between BtubA and BtubB is much higher than to α- and β-tubulin, *Prosthecobacter* probably acquired the *btub* operon not too long after the αβ-tubulin duplication event in eukaryotic cells, in other words, before the bulk of the differentiation and specialisation of αβ-tubulins as seen today. In turn, this suggests that the formation of microtubule occurred relatively early in eukaryotic evolution, and possibly already existed in the last common ancestor of all eukaryotes.

Another interesting aspect in regards to the evolutionary gap between FtsZ and tubulins would be the function of BtubAB. As discussed above, FtsZ-based cell division mechanism may have existed in early life forms. It is possible that ESCRT-based cell division mechanism was already present at the time eukaryotes branched out of archaia, at the same time, FtsZ/tubulin ancestor started to evolve into microtubule-forming αβ-tubulins and MTOC-forming γ-tubulin as known today. Therefore, the function of BtubAB is also interesting in terms of giving insight of the divergent functions of FtsZ/tubulin-related proteins. Unfortunately, the biological function of BtubAB is not understood, mostly because *Prosthecobacter* is not genetically tractable right now. However, the presence of FtsZ in *Prosthecobacter* suggests that Btub proteins are unlikely to be involved in cell division (Schlieper et al., 2005). Tomography studies of these bacteria showed that BtubAB filaments were almost always localised in the stalk or in the transition zone between the stalk and cell body of the *Prosthecobacter* bacteria, suggesting Btub proteins could be involved in the stalk formation, or some other morphology related process (Pilhofer et al., 2011). To add to this idea, another recent tomographic study claims that BtubC has membrane binding properties, further supporting the possibility that Btub proteins are involved in membrane manipulation needed.
for stalk formation in *Prosthecobacter* (Akendengue et al., 2017). On the other hand, the biological function of BtubAB is not necessarily connected to its eukaryotic ancestors, as the bacteria might have co-opted these proteins for different purposes; although the general mechanism, that is the alteration of morphology through membrane attachment could be conserved. The origin of BtubC is more difficult to trace, since it only contains a widely-conserved fold that dominates the sequence.
Figure 4.5 Phylogenetic tree of tubulin protein family, adapted from (Findeisen et al., 2014). The phylogenetic tree suggests all tubulin/FtsZ-related proteins evolve from a common ancestor, and BtubA (bacterial $\alpha$-tubulin) and BtubB (bacterial $\beta$-tubulin) branched after the split of eukaryotic $\alpha$- and $\beta$-tubulin. The scale bar corresponds to estimated amino acid substitutions per site.
4.4 BtubAB and microtubule dynamics

Microtubules are fascinating objects not only because of their amazing structure, but also of their unique dynamics, in particular, dynamic instability. To be able to probe the three-dimensional surroundings, microtubules stochastically switch between rapid growth and fast shrinkage, until a designated target such as a kinetochore has been found. This behaviour, known as dynamic instability, has been observed extensively both in vivo and in vitro and is a key functional property of microtubules (Mitchison and Kirschner, 1984a). The current prevailing mechanistic theory supports the so-called ‘GTP cap model’, in which the growing microtubule tip contains tubulin heterodimers in which GTP hydrolysis is yet to occur, forming a ‘GTP cap’. The GTP cap has stabilising effects, and the loss of this tip structure (via hydrolysis of GTP to GDP happening faster than addition of new GTP bound heterodimers) leads to catastrophe, the rapid depolymerisation of the microtubule. This shrinkage does not occur indefinitely, often features such as ‘GTP islands’ present in microtubule can halt the catastrophe and trigger rescue, leading to the regrowth of the microtubule (Akhmanova and Steinmetz, 2015).

Despite general agreement upon the GTP cap model, the precise molecular mechanism and causes of dynamic instability are poorly understood. In vitro work shows that microtubules exhibit an ageing behaviour, where microtubules that have been growing for longer (hence are ‘older’) have a higher tendency to undergo catastrophe. A closer look at this ageing behaviour of microtubules revealed that catastrophe requires more than one event to occur (more than just a one-step decay), as the accumulated plot of microtubule catastrophe times follow a gamma distribution (Gardner et al., 2011). However, the nature of these molecular events is not known. Some evidence suggests the involvement of the size of GTP cap and the shape of microtubule tip, which seems to become more tapered as the microtubule grows (Coombes et al., 2013). From a structural point of view, there has been a longstanding interest in the link between protofilament curvature and microtubule dynamic instability. Earliest evidence for a link came from an electron microscopy study showing protofilaments peeling off from the microtubule tips, and the ever-growing number of tubulin structures bound with various MAPs showing all kinds of curvatures have added weight to the idea of a correlation between αβ-tubulin curvature and microtubule dynamics (Mandelkow et al., 1991; Brouhard and Rice, 2014). Recent high resolution structures of microtubules in different nucleotide states revealed a compaction at the longitudinal interface induced by GTP hydrolysis, which leads to conformational changes in
α-tubulin and eventually possible curvature. However, the lateral contacts in all nucleotide states remain relatively unchanged (Alushin et al., 2014).

What can we learn from BtubAB? It is truly remarkable that the BtubAB mini microtubule resembles the complex dynamic behaviour of its eukaryotic counterparts despite its bacterial origin and much small diameter. The growth rate of BtubAB filaments is similar to that of microtubules, and the ageing effect was also observed in Gero Fink’s TIRF data. The fact that the drastic reduction in protofilament number, from 13 to 4, does not seem to affect the dynamic behaviour of the filaments suggests there is little contribution of lateral interactions in catastrophe. On the other hand, the M-loops, although present in BtubAB filaments, are quite different from those in microtubules, for instance in the lateral contact angles and amino acid composition (Figure 4.4). A definitive study of lateral interactions in microtubule dynamics would be to measure the catastrophe parameters of microtubules with various protofilament numbers. Nevertheless, BtubAB filaments provide another system for studying microtubule-related dynamics, their bacterial origins make them a much more accessible system for mutations, as αβ-tubulins require a complex folding pathway. The immediate next step experiments could be structural studies of BtubAB filaments in different nucleotide states, together with a focus on the tips (now possible with cryoEM), to see if similar compaction, or curvature effects are present. An understanding of the native curvature of BtubAB heterodimers could be obtained by solving the dimer structure by cryoEM, treating them as single particles. In addition, pull-down assays can be performed to search for possible BtubAB regulating proteins, as MAPs have been proven to be valuable tools in studying microtubule dynamics. A to-be-made GTP-Btub specific antibody may be an alternative, in particular for recognising GTP lattices. A similar strategy was used by Dimitrov et al. to investigate the GTP islands in microtubules (Dimitrov et al., 2008). Another future direction is to investigate the stabilising mechanism of BtubC on BtubAB filaments, which may provide insights into the poorly understood mechanisms for many MAPs.
Reference


Caplow, Michael, and Lanette Fee. 2002. Dissociation of the tubulin dimer is extremely slow, thermodynamically very unfavorable, and reversible in the absence of an energy source. Molecular biology of the cell 13, no. 6: 2120-2131.


Gardner, MK, M Zanic, C Gell, V Bormuth, and J Howard. 2011. The depolymerizing kinesins Kip3 (kinesin-8) and MCAK (kinesin-13) are catastrophe factors that destabilize microtubules by different mechanisms. Cell 147, no. 5: 1092-1103.


Glaeser, RM. 1971. Low temperature electron microscopy-radiation damage in crystalline biological materials. JOURNAL OF MICROSCOPY


Huang, X, B Luan, J Wu, and Y Shi. 2016. An atomic structure of the human 26S proteasome. Nature Structural and Molecular Biology


Kraemer, JA, ML Erb, CA Waddling, EA Montabana, EA Zehr, H Wang, K Nguyen, DS Pham, DA Agard, and J Pogliano. 2012. A phage tubulin assembles dynamic filaments by an atypical mechanism to center viral DNA within the host cell. *Cell* 149, no. 7: 1488-1499.


Orlova, EV, and HR Saibil. 2011. Structural analysis of macromolecular assemblies by electron microscopy. *Chemical reviews*


Penczek, PA, J Zhu, R Schröder, and J Frank. 1997. Three dimensional reconstruction with contrast transfer compensation from defocus series. Scanning Microscopy


Rosier, DJ De, and A Klug. 1968. Reconstruction of three dimensional structures from electron micrographs. Nature


Voorhees, RM, IS Fernández, SHW Scheres, and RS Hegde. 2014. Structure of the mammalian ribosome-Sec61 complex to 3.4 A resolution. *Cell*


Williams, DB, and CB Carter. 1996. The transmission electron microscope. *Transmission electron microscopy*

Wu, Jingchao, and Anna Akhmanova. 2017. Microtubule-Organizing Centers. *Annual Review of Cell and Developmental Biology* no. 0:


Zehr, Elena A, James A Kraemer, Marcella L Erb, Joanna KC Coker, Elizabeth A Montabana, Joe Pogliano, and David A Agard. 2014. The structure and assembly mechanism of a novel three-stranded tubulin filament that centers phage DNA. *Structure* 22, no. 4: 539-548.


## Appendix A Crystallography and cryoEM data statistics

### Table 1 Crystallography and cryoEM data

<table>
<thead>
<tr>
<th>Statistics</th>
<th>Prosthecobacter vanneervenii PvBtubC EMTS derivative</th>
<th>Prosthecobacter vanneervenii PvBtubC native</th>
<th>Prosthecobacter dejongei PdBtubABC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample</td>
<td>CAP16140.1</td>
<td>CAP16140.1</td>
<td>AA012155.1, AA012159.2, this work</td>
</tr>
<tr>
<td>Constructs</td>
<td>1-256-KLHHHHHHH</td>
<td>1-256-KLHHHHHHH</td>
<td></td>
</tr>
<tr>
<td>Method</td>
<td>crystallography SAD</td>
<td>crystallography</td>
<td>crystallography, helical reconstruction</td>
</tr>
<tr>
<td>Data collection</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Beamline/microscope</td>
<td>ESRF ID30B</td>
<td>ESRF ID29</td>
<td>FEI Polara G2, Falcon III</td>
</tr>
<tr>
<td>Wavelength / energy</td>
<td>1.006 Å</td>
<td>0.977 Å</td>
<td>300 kV</td>
</tr>
<tr>
<td>Crystal / helical</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Space / point group</td>
<td>H32</td>
<td>H32</td>
<td></td>
</tr>
<tr>
<td>Cell (Å)</td>
<td>120.0, 120.0, 239.1</td>
<td>119.7, 119.7, 239.4</td>
<td>-5.54°, 79.31 Å</td>
</tr>
<tr>
<td>Data</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>2.8</td>
<td>2.5</td>
<td>3.6</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>100 (100)</td>
<td>99.4 (100)</td>
<td></td>
</tr>
<tr>
<td>Multiplicity</td>
<td>11.8 (11.3)</td>
<td>4.9 (5.2)</td>
<td></td>
</tr>
<tr>
<td>(I) / σ(I)</td>
<td>17.8 (2.4)</td>
<td>10.8 (2.7)</td>
<td></td>
</tr>
<tr>
<td>Rmerge</td>
<td>0.102 (1.036)</td>
<td>0.108 (0.674)</td>
<td></td>
</tr>
<tr>
<td>Rfree</td>
<td>0.044 (0.467)</td>
<td>0.054 (0.330)</td>
<td></td>
</tr>
<tr>
<td>CCC1/2</td>
<td>0.999 (0.726)</td>
<td>0.996 (0.690)</td>
<td></td>
</tr>
<tr>
<td>Anomalous correlation</td>
<td>0.24 (0.0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mercury sites</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Images, pixel size</td>
<td></td>
<td></td>
<td>6105, 1.34 Å</td>
</tr>
<tr>
<td>Defocus range, dose</td>
<td></td>
<td></td>
<td>-1.5 - -3 µm, 40 e/Å²</td>
</tr>
<tr>
<td>Helical segments</td>
<td></td>
<td></td>
<td>257,661</td>
</tr>
<tr>
<td>Refinement</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R / Rfree</td>
<td>0.186 / 0.237</td>
<td>0.32</td>
<td></td>
</tr>
<tr>
<td>FSC (REFMAC)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Models</td>
<td>2 chains/ASU: residues 2-255, 92 waters</td>
<td>1 BtubAB heterodimer: A residues 3-435, B 2-37; 46-273;281-426</td>
<td></td>
</tr>
<tr>
<td>Bond length rmsd (Å)</td>
<td>0.015</td>
<td>0.014</td>
<td></td>
</tr>
<tr>
<td>Bond angle rmsd (°)</td>
<td>1.716</td>
<td>1.556</td>
<td></td>
</tr>
<tr>
<td>Favoured (%)</td>
<td>100</td>
<td>99.5</td>
<td></td>
</tr>
<tr>
<td>Disallowed (%)</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>MOLPROBITY score</td>
<td>95th percentile</td>
<td>100th percentile</td>
<td></td>
</tr>
<tr>
<td>PDB/EMDB IDs</td>
<td>5001</td>
<td>5009, EMD-3726</td>
<td></td>
</tr>
</tbody>
</table>

1 Values in parentheses refer to the highest recorded resolution shell.
2 5% of reflections were randomly selected before refinement.
3 Percentage of residues in the Ramachandran plot (PROCHECK 'most favoured' and 'additionally allowed' added together).
Appendix B Plasmids and Strains

### Plasmids

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Description</th>
<th>Tag</th>
<th>Resistance</th>
<th>Promoter</th>
<th>Vector</th>
</tr>
</thead>
<tbody>
<tr>
<td>PdBtubAB</td>
<td>Full length BtubAB genes from <em>Prosthecobacter dejonjeii</em> with the intergenic region intact and unmodified for bicistronic coexpression.</td>
<td>No tag</td>
<td>Amp</td>
<td>T7</td>
<td>pHis17</td>
</tr>
<tr>
<td>PdBtubC</td>
<td>Full length BtubC gene from <em>Prosthecobacter dejonjeii</em>, MGSSHHHHHHHSSGLPRGSHPdBtubC(1-257).</td>
<td>N-term 6xHis</td>
<td>Amp</td>
<td>T7</td>
<td>pET15b</td>
</tr>
<tr>
<td>PvBtubC</td>
<td>Full length BtubC gene from <em>Prosthecobacter vanneervenii</em>, PvBtubC(1-256)-KLHHHHHHH.</td>
<td>C-term 6xHis</td>
<td>Amp</td>
<td>T7</td>
<td>pHis17</td>
</tr>
</tbody>
</table>

### Strains

<table>
<thead>
<tr>
<th>Strain</th>
<th>Genotype</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Escherichia coli</em> C41 (DE3)</td>
<td>BL21(DE3) derivative</td>
<td>(Miroux and Walker, 1996)</td>
</tr>
<tr>
<td><em>Escherichia coli</em> MAX Efficiency™</td>
<td>F- φ80lacZΔM15 Δ(lacZYA-argF) U169 recA1 endA1 hsdR17 (rK-, mk+) phoA supE44 λ- thi-1 gyrA96 relA1</td>
<td>ThermoFisher Scientific</td>
</tr>
</tbody>
</table>
Appendix C Growth media and general buffers

All chemicals are from Sigma Aldrich unless otherwise indicated.

TYE Plate

- 10 g tryptone
- 5 g yeast extract
- 8 g NaCl
- 15 g Agar

Final pH 7.0 at 25 °C, in 1 L ultra pure water then autoclave and pour, and antibiotics can be added after autoclaving at around 55 °C if needed.

SOB liquid media

- 28 g SOB ready mix (VMR)

Make final volume to 1 L, adjust pH to 7.0 before autoclaving.

2x TY liquid media

- 16 g tryptone
- 10 g yeast extract
- 5 g NaCl

Make final volume to 1 L using ultra pure water, and adjust pH to 7.4 before autoclaving.

1x SDS running buffer

- 30 g Trizma Base
- 144 g glycine
- 10 g SDS

Make up to 10 litres with distilled water.

Staining solution (for SDS-PAGE)

- Quick Coomassie Stain (Generon)

10x TBE

- 108 g Trizma Base
- 903 g Na₂EDTA x 2H₂O
- 55 g boric acid
Make up to 1 litre with MilliQ water

**DNA loading buffer**

- GelPilot DNA Loading Dye, 5x (Qiagen)

**Protein loading buffer**

- 0.6 g Trizma Base
- 0.77 g DTT
- 2 g SDS
- 10 g sucrose
- 0.095 g EDTA
- 0.25% w/v bromophenol blue

Make up to 50 ml with MilliQ water and adjust pH to 6.8.
Precise knowledge of the structure of macromolecules is essential for understanding their functions and the mechanisms which they utilise. Among other structural techniques including X-ray crystallography and nuclear magnetic resonance (NMR), biological electron microscopy (EM, mainly transmission electron microscopy) is becoming more and more popular for studying the architecture of cells, viruses and protein assemblies. Recent advances in microscope design, direct electron detectors, image processing algorithms and automation capabilities are yielding structure of biomolecules at near-atomic resolution, which leads the biological EM field through a resolution revolution (Kühlbrandt, 2014).

Generally speaking, biological EM can be divided into three techniques: negative stain EM, cryoEM and cryoET. The methods are different in suitable samples, specimen preparation, data collection and data processing to a certain extent, and a combination of EM techniques is often used during the whole project. A transmission electron microscope uses high-energy electrons (100 keV or higher) to form an image of very thin objects, and its principle is very similar to a light microscope, except the source of radiation is replaced from visible photons to electrons. In a typical TEM, an electron beam is generated by a source (a cathode or a field emission gun) under high vacuum, and then accelerated down the microscope column at accelerating voltages from 80 kV – 300 kV. The electrons are scattered as they pass through the thin specimen, which are focused by electromagnetic lenses within the microscope to form an image on a fluorescent screen, a charge-coupled device or a direct electron detector (Figure D.1). The most significant difference between a light microscope and a TEM is the resolving power, which is directly influenced by the source of radiation, and the resolution is limited to roughly $\frac{1}{2}$ of the wavelength of illumination beam. Compare to the 200 nm resolution of most light microscopes, a TEM can achieve a resolution of 0.04 Å. However, in reality, structures solved using TEM is far from reaching this theoretical limit due to the imperfections of the electromagnetic lenses. Objective lens aberrations, especially third-order spherical aberration, chromatic aberration and axial astigmatism, ultimately limit the resolution of the microscope, and the vulnerability of biological samples to radiation damage puts further restrain to the resolving power of TEM. Therefore, the advances in biological cryoEM have been mainly focused on reducing radiation damage and how to deal with the subsequent noisy image generated from low-dose imaging mode.
Figure D. 1 A schematic representation of a typical TEM, adapted from Williams and Carter, 1996.
**Negative stain EM**

A main focus for biological sample preparation for EM is to reduce the radiation damage by the use of heavy atom salt stain or liquid nitrogen temperature. The negative stain method was first introduced by Brenner and Horne in 1959 (Brenner and Horne, 1959). Typically, the biomolecule of interest is applied to a carbon-coated specimen grid and stained with an electron-dense heavy atom salt solution (typically uranyl acetate). The excess solution is blotted away and the grid is allowed to dry (Figure D.2a). Although the heavy atom salt solution produces high contrast, the negative stain electron micrograph contains only the cast of the solvent-accessible surface of the biomolecule being imaged. Consequently, only the shape of the molecule is reconstructed and the majority of internal structural information cannot be obtained using negative stain EM. In addition, artefacts including partial projections, as parts of the particle farthest away from the support stick out of the layer of stain, and flattening can occur depending on the specific interaction between the heavy atom salt stain and the molecule of interest. Despite the limits of negative stain EM, sample preparation using this method is rather fast and generally requires less concentrated sample compared to cryoEM. Therefore, negative stain EM these days is often used as a first step experiment in an EM project to assess the sample quality and investigate possible heterogeneity involved in the sample. Negative stain electron micrographs are typically imaged at room temperature and low voltages (80 kV – 120 kV).
Figure D. 2 Schematic overview of sample preparation process of (a) negative stain EM and (b) cryoEM method, adapted from Orlova and Saibil, 2011.
A quantum leap in the biological EM field was the development of cryoEM method, which consists of macromolecules being embedded in a thin layer of vitreous ice (Dubochet et al., 1988). To prepare a sample using the cryoEM method, a thin layer of aqueous solution containing the molecule of interest is applied to the specimen grid before the grid is plunged into liquid ethane at a temperature between −160° and −180° (Figure D.2b). Through this rapid cooling, instead of crystalline ice, which can be damaging to the sample, the thin layer of water film vitrifies into vitreous ice. As a result, the macromolecules stay fully hydrated using the cryoEM method, and the low temperature prevents evaporation of the sample inside the microscope as well as reduces radiation damage induced by electron beam (Glaeser, 1971). An electron micrograph obtained using the cryoEM method contains a collection of density projections of the molecule ideally adopting random orientations, surrounded by the lower density buffer. It is not difficult to imagine that cryo-electron micrographs have much lower contrast and lower signal-to-noise ratio (SNR), the typical noise powers in cryoEM images are 10 times larger than the signal (Nogales and Scheres, 2015). Nevertheless, a large enough dataset containing tens to hundreds of thousands of particles can be sufficient to solve structures to near atomic resolution as demonstrated by the 3.5 Å structure of human 26S proteasome, 3.4 Å structure of human γ-secretase and the 3.7 Å structure of the yeast spliceosome (Bai et al., 2015; Huang et al., 2016; Nguyen et al., 2016). While large macromolecular complexes can be routinely studied to obtain high-resolution structures by cryoEM, protein that is below 100 kDa present a particularly challenge as the particles becomes harder to align as the size decreases during image processing. A major breakthrough in tackling this problem is the development of the Volta phase plate, which greatly enhances the contrast of the molecules. An exciting paper from the laboratory of Baumeister and Danev demonstrated how they achieved the 3.2 Å structure of human haemoglobin (64 kDa) using the Volta phase plate (Khoshouei et al., 2017). Another common challenge in cryoEM, especially in single-particle studies, is the preferred-orientation problem. A cryoEM specimen grid needs to contain particles in various random orientations in order to allow 3D reconstruction of the macromolecule. However, in practice this is often not the case, presumably due to the surface properties of the molecules that cause specific regions to preferentially adhere to the air-water interface. This preferred orientation phenomenon often leads to the loss of information from certain angles and thus, limit the resolution parallel to the preferred orientation axis, or a ‘smearing effect’ on the map (Barth et al., 1989). Solutions to this issue include applying a
support layer on the specimen grid, such as graphene and graphene oxide, in order to disrupt the specific interaction between the air-water interface and the surface of the macromolecule (Pantelic et al., 2010; Voorhees et al., 2014; Bokori-Brown et al., 2016). In addition, tilt-based approaches involving tilting the specimen have also been employed in single-particle studies to tackle the preferred orientation issue (Tan et al., 2017).

Typically, cryoEM specimen grids are imaged at liquid nitrogen temperature (−180°) and at voltages from 120 kV to 300 kV. High resolution datasets are often collected at the state-of-the-art FEI Titan Krios electron microscope operating at 300 kV, coupled with either FEI Falcon III or Gatan K2 Summit direct electron detectors. As the incoming wave interacts with the sample, electrons are scattered at a wide range of angles and the exit waves are focused by electromagnetic lens at the image plane. Cryo biological samples are generally considered as ‘weak-phase object’, so that the images of these objects have both low amplitude and low phase contrast due to the weak scattering power from the light elements (carbon, nitrogen, oxygen and so on) from which biological macromolecules are built. However, defocus is often applied during data collection to enhance the contrast of the weak-phase biological samples and to retrieve more low frequency information. The effects of defocus, spherical aberration and wavelength (voltage) of radiation source on image contrast of biological samples are described by the contrast transfer function (CTF). The higher contrast obtained through defocus comes with a price, the applied defocus causes CTF signals to oscillate at various spatial frequencies, and information is lost at its zeros. Therefore, CTF estimation and correction of EM images must be done prior to further image processing, and micrographs collected at a range of defoci are combined to compensate the lack of information at its zeros (Figure D.3). In addition to the rapid oscillation of CTF at high frequencies, the intensity of CTF also decreases at high frequencies due to partial coherence, lens aberrations and stage or beam-induced motions. This dampening effect can be described by the envelope effect, which further limits the resolution EM images.

The CTF describes mathematically the exit waves that reach the final electron micrograph. Once the CTF-corrected micrographs are obtained, individual particles can be extracted. According to the central slice theorem, each particle on the micrograph is a 2D projection of the 3D object that corresponds to a slice through the centre of the 3D object in Fourier space. Ideally, thousands of 2D projections that encompass a wide range of orientations relative to the electron beam can be extracted from a cryoEM dataset. Once the orientation of each projection is assigned computationally, assuming sufficient orientations are sampled, the
3D object can be recovered by combining all projections with known orientations into the 3D transform and calculating an inverse Fourier transform of the combined 3D transform (Figure D.4). Typical orientation assignment algorithms include the ‘common line’ method and ‘projection-matching’ method (Nogales and Scheres, 2015). The ‘common line’ method exploits the observation that each pair of 2D projections shares a so-called ‘common line’ in the 3D Fourier transform to determine relative orientations between pairs of projections (van Heel et al., 1996). The ‘projection-matching’ approach is to compare each experimental particle image with computationally generated projections of an initial 3D reference structure in different directions and determine the orientation of each 2D projection (Penczek et al., 1994; Scheres, 2012c).
Figure D. 3 Theoretical contrast transfer functions, adapted from Penczek et al., 1997.

(a) Theoretical contrast transfer function for three defocus values: 1.5 μm (___), 2.0 μm (----), and 2.5 μm (-----).

(b) Theoretical contrast transfer function with envelope functions effects for three defocus values: 1.5 μm (___), 2.0 μm (----), and 2.5 μm (-----).
Figure D. 4 An overview of cryoEM structural determination process using the ‘projection-matching’ method, adapted from Nogales and Scheres, 2015.

(a) The central slice theorem states that the 2D projection of a 3D object in real space (left column) is equivalent to taking a central 2D slice out of the 3D Fourier transform of that object (right column). The real-space projection direction (left, dashed red arrows) is perpendicular to the slice (right, red frame).

(b) The software generates 2D projections from different angles of an initial low-resolution 3D reference.

(c) The software compares each experimental projection with all generated projections to find the best match of a given similarity measure.

(d) The software orients all experimental projections relative to the 3D structure.

(e) The projection-slice theorem then implies that the 3D reconstruction can be calculated by positioning many 2D slices (the 2D Fourier transforms of all experimental projections) into the 3D transform and calculating an inverse transform. Iterating steps (b)–(e) will gradually improve the orientations, and hence the resolution, of the reconstruction.
CryoET

Specimen grids for cryoET are prepared in a very similar way as cryoEM, except in the majority of samples, 10 – 20 nm gold fiducials are added to facilitate alignment during tomogram reconstruction. Compare to cryoEM, a much wider range of samples can be studied using cryoET including cryo-sectioned tissue, bacterial cells, viruses and filaments (Al-Amoudi et al., 2004; Oikonomou and Jensen, 2017). A typical cryoET data collection involves the use of the FEI Titan Krios electron microscope operating at 300 kV coupled with Gatan K2 Summit direct electron detector. Instead of taking a single image of an area of interest, as in the case of cryoEM, a series of images is collected. During data collection, the stage is tilted typically from $-60^\circ$ to $+60^\circ$ at $1^\circ$ or $3^\circ$ increment, and an electron micrograph is taken each time the stage is tilted. The whole set of tilted images is called the tilt series. With the help of fiducial markers added during sample preparation and back projection algorithm, a tomogram can be reconstructed from the tilt series representing the 3-dimensional volume of the area of interest (Figure D.5). Once the tomogram is obtained, particles in 3D volumes can be extracted and averaged again to produce high resolution structures through sub-tomogram averaging. Although the structures solved by sub-tomogram averaging rarely reach near-atomic resolution, cryoET provides a unique opportunity to study protein assemblies in their natural environment by averaging molecular machineries such as the type IV secretion apparatus extracted from tomogram of bacteria (Jeong et al., 2017). In addition, previously undiscovered structures may be revealed as more and more bacteria are being visualised using cryoET (Dobro et al., 2017).

The two major challenges involved in cryoET are even more aggressive radiation damage and the ‘missing wedge’ issue. As each tilt series may contain 120 images, the total dose sometimes exceeds 120 electrons/Å$^2$ for one field of view (40 – 80 electron/Å$^2$ for a cryoEM image), whereas an individual image from such tilt series only has dose as low as 1 electrons/Å$^2$. The high total dose and extremely low individual dose propose a big challenge for tilt series alignment and tomogram reconstruction. In addition, as the microscope stage rarely tilts beyond $+60^\circ$ or $-60^\circ$ due to the thicker ice present at high tilt positions, a wedge-shaped area is constantly missing in Fourier space during tomogram reconstruction (referred as the ‘missing wedge’ problem), which can introduce artefacts in the final tomogram. Despite the challenges, the Briggs lab has recently redefined the conventional limit of cryoET by presenting a 3.9 Å structure of HIV capsid obtained using cryoET and subtomogram averaging (Schur et al., 2016).
Figure D. 5 Data collection and tomogram reconstruction in cryoET, adapted from Steven and Aebi, 2003.

(a) The stage is rotated to successive tilt positions around an axis perpendicular to the electron beam.

(b) At each tilt position, a different projection of the same specimen area is recorded on a CCD camera or direct electron detector.

(c) To reconstruct the 3-D structure of the specimen, each projection is back-projected along its line-of-sight. In this way, the experimental volume of densities is built up. Precise mutual alignment of all the projections prior to back-projecting them is essential to avoid degrading the tomogram, this is often facilitated by gold fiducials that are added during sample preparation stage.
Among all the macromolecular machineries that are studied by cryoEM, helical assemblies represent a class of particularly interesting object. As discussed above, a cryoEM 3D reconstruction requires multiple images corresponding to different views of the object of interest. For helical assemblies such as phage tails, helical viruses and cytoskeletons, multiple views are naturally present. In such assemblies, the orientations of different molecules are defined by helical parameters, which describes the relationship between individual monomers that make up the helical structure. Once the helical parameters are determined, a 3D reconstruction can be generated using helical Fourier inversion method or real-space helical reconstruction method (DeRosier and Moore, 1970; Egelman, 2000). Indeed, the first 3D EM reconstruction was obtained using helical Fourier inversion method to study the bacteriophage T4 tail (Rosier and Klug, 1968).

Helical objects provide a big advantage compared to single-particle objects due to its high symmetry involved, as one segment of such object already contains multiple copies of the same molecule in different orientations defined by helical symmetry. In this thesis and RELION, helical symmetry is simply defined by the rise and the twist (Figure D.6) (He and Scheres, 2017). In addition to the high symmetry, there is ample prior knowledge about the orientations of helical segments. In helical reconstruction, individual segments are extracted along a filament and each segment is the equivalent of one particle in single-particle analysis. Instead of Euler angles $\alpha$, $\beta$ and $\gamma$ in single-particle orientation parameters, rot, tilt and psi angles are used to describe the orientation of each segment (Figure D.7). The rotations along the helical axis of is measured by the ‘rot’ angle, which is strongly linked with the helical symmetry since every asymmetric unit within a helix is identical. In RELION, this is measured by in-plane translational searches within the distance of one helical rise. The tilt angle describes how much the filament is tilted in the z-axis, so that a top view would have a tilt of $0^\circ$ and a bottom view would have a tilt of $180^\circ$. Because the long filaments tend to lie roughly horizontal inside the ice layer, the tilt angles can be fit into a Gaussian distribution that is centred at $90^\circ$. In fact, it is uncommon to capture a top view or a bottom view of a filament unless the sample contains many short stretches of filaments that are shorter than the typical ice thickness of a cryoEM grid. The in-plane rotation of a helical segment is measured by the ‘psi’ angle. During segment picking process, prior knowledge about this psi angle can be obtained by measuring the direction of the filament. Only the polarity of each segment needs to be further determined during refinement, and this is done by a bimodal search of around $0^\circ$ or $180^\circ$ in RELION (He...
and Scheres, 2017). The ample prior knowledge that are available for helical objects, especially about psi and tilt angles, greatly reduces the computational power required during orientation assignment process compared to single-particle analysis. Combined with the high symmetry involved in helical objects, often less number of particles, or helical segments are needed to obtain a near-atomic resolution structure.

Despite the advantage associated with helical assemblies, the most difficult step in solving such structure is often the determination of the helical parameters. Traditionally, helical parameters are derived through the Fourier-Bessel method, which investigates the ‘layer lines’, the signatory symmetrical streaks in the diffraction pattern of a helical object, in order to derive the helicity (Diaz et al., 2010). An alternative approach involves the use of cryoET and subtomogram averaging to obtain an unbiased, reference-free low-resolution model, and the helical parameters can be simply obtained by observing the model. The helical symmetry of BtubABC filament in this thesis was determined using cryoET and subtomogram averaging. Another challenge that is associated with helical reconstruction rises when pseudo symmetry is involved, as demonstrated by 13-pf microtubule and BtubAB filament. The majority of biological filaments are formed by the polymerisation of identical monomers; thus, every lateral interface is identical. However, in some cases heterodimers are involved as building blocks during filament assembly. In such assemblies of heterodimers such as 13-pf microtubule and BtubAB filament, there is one lateral interface that is different from all the rest within the filament (referred as the ‘seam’, see Figure 4.3 in main text for schematic representation) and that particular interface breaks the overall helical symmetry. Such helix is classified as having a pseudo symmetry, and the full symmetry cannot be applied during reconstruction (therefore more particles are needed). In addition, as BtubA and BtubB, and α- and β-tubulins closely resemble each other, it is almost impossible for the software to distinguish the monomer within each heterodimer, resulting a medium-resolution map where BtubA and BtubB, or α- and β-tubulins are averaged. The issue produces profound difficulty in solving microtubule structures to near-atomic resolution. So far, all the near-atomic resolution microtubule and BtubAB structures are solved with an additional protein that are attached at every heterodimer to aid the differentiation between the tubulins, or BtubAB within a heterodimer.
Helical assemblies are generally formed by the polymerisation of identical monomers along the path of a continuous helix. The relationship between individual monomers, represented by the grey spheres here, can be described simply by the ‘rise’ and ‘twist’.
Figure D. 7 Priors on the orientation parameters of a helical segment in RELION, adapted from He and Scheres, 2017.

(a) A top-hat prior on the in-plane translations along the helical axis, and a Gaussian prior on the in-plane translations perpendicular to the helical axis.
(b) A mono-modal Gaussian prior on the tilt angle, which describes the out-of-plane rocking of filaments in the ice layer.
(c) A bi-modal Gaussian prior on the psi angle, which describes the in-plane rotation of the filament.
Appendix E Other experiments

During this project, I have conducted other experiments that are not recorded in this thesis since no positive results were obtained. As BtubAB filament possesses the ‘seam’ problem just like microtubule, a marker protein that only attach to BtubA, or BtubB was needed so that the software can distinguish BtubA from BtubB. Before BtubC was found, several attempts were made by fusing a known domain to either BtubA or BtubB, so that the added domain may function as the ‘marker’ needed to distinguish BtubA from BtubB.

The domains that were tried for fusion were SAM domain (a two-helix bundle), vps28 domain (a four-helix bundle), and monomer superfolder GFP (msGFP). The plasmid containing SAM and vps28 domain were obtained from previous lab member H. Low, and the msGFP plasmid was obtained from previous lab member D. Ghosal. The sequences for each domain is shown below.

<table>
<thead>
<tr>
<th>Domain</th>
<th>DNA sequence</th>
<th>Protein sequence</th>
</tr>
</thead>
</table>
| SAM domain  | CCCAGCCTCGCTAGTTTTTTAACAGGATTG  
GGGCTCCAAACGCCATGATTTCACCT  
CCAAAGTTTACAGTCTATTACCTACCTGCA  
GAACCTGACCATTTGAAGACCTGGGTCTCT  
GAAAATCCCCGAGCATTCGCACTGACCATC  
CTGCGGGGCCCGCTGACCATGACCAAGG  
CTAA | PSLVSFLTGLGAPNAIEYFTS  
QGLQSIYHLQNLTIEDLGALK  
IEPQYRMTIWRGLQDLKQG |
| Vps28 domain| AAAGGGAAATCTGACCGCTACATGCTGAC  
ATTGTATCTTCTCTATCACTGACATGACCA  
AAGTCCAGGCTGATTGCGGGAGCAGTAATGG  
AAACCATGACAGGATGAGGACCACCTCCCA  
CAGATTTTAGGGGACGGGGGAAATGGAGTC  
AGTGTGTTAGAAGCTACAGGATATGCAGTCAG  
CCTCCGATGAAACTGATGATTTCAAGAATGCA  
GAACAAATGCTCTTCAGTGATAGAATCGCTTTGAC  
TAATCATCGTCCATGTCCTCCTCCATCTCAA | KGNLNRCAIADIVSLFTTVMDK  
LRLEIRAMDEIQPDLRELMET  
MNRMSLPPDFEGREKVSQ  
WLQKLSSMASDELDSQVR  
QMLFDLESAYNAFNRFLHS |
| msGFP       | AGTAAAGGGAAAGCTGACATCCTGGTCTGCTCCAGGCTACATGCTGAC  
ATTGTATCGTTCTCTCTATCACTGACATGACCA  
AAGTCCAGGCTGATTGCGGGAGCAGTAATGG  
AAACCATGACAGGATGAGGACCACCTCCCA  
CAGATTTTAGGGGACGGGGGAAATGGAGTC  
AGTGTGTTAGAAGCTACAGGATATGCAGTCAG  
CCTCCGATGAAACTGATGATTTCAAGAATGCA  
GAACAAATGCTCTTCAGTGATAGAATCGCTTTGAC  
TAATCATCGTCCATGTCCTCCTCCATCTCAA | SKGEELFTGVVPILVHDGY  
NGHKFVSRGEGECDATNGK  
LTLLKICCTTGGKLVPWPLT  
TLTYGVQCFSPYDPDMSKHQ  
DDFKSAMPEGYVQERTISFK  
DDGTYKTRAEVKFEGDITLV  
NRIELKIGDFKEDGNNILGHKL  
EYNNPNSHYVYTADKQKNGI  
KANFKIRHNVEDGSVLADH  
YQQMTSIQGPDVPLLDPHEL  
STQSKLSDKPEKRDHMVLLE  
EFVTAAGITHGMDELYK |
SAM, vps28 and msGFP have been test by fusion at the C-terminal (as the C-terminal is facing outwards from the EM structure) with either BtubA or BtubB. Linkers of various length have been introduced between A/B and the domain. Each fusion protein has been expressed, purified and tested for filament assembly ability with the native counterpart. The results were analysed by electron microscopy. Below is a list of the constructs that have been tested and a short summary of the results.

<table>
<thead>
<tr>
<th>Construct</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>pHis17-BtubAB-SAM</td>
<td>B-Sam can be purified, but BtubA and BtubB-SAM don’t assemble into filaments.</td>
</tr>
<tr>
<td>pHis17-BtubAB-vps28</td>
<td>B-vps28 is insoluble.</td>
</tr>
<tr>
<td>pHis17-BtubAB-GSS-SAM</td>
<td>B-GSS-Sam can be purified, but BtubA and BtubB-GSS-SAM don’t assemble into filaments.</td>
</tr>
<tr>
<td>pHis17-BtubAB-GSS-vps28</td>
<td>B-GSS-vps28 is insoluble.</td>
</tr>
<tr>
<td>pHis17-BtubAB-GSSGSS-SAM</td>
<td>B-GSSGSS-Sam can be purified, but BtubA and BtubB-GSSGSS-SAM don’t assemble into filaments.</td>
</tr>
<tr>
<td>pHis17-BtubAB-GSSGSS-vps28</td>
<td>B-GSSGSS-vps28 is insoluble.</td>
</tr>
<tr>
<td>pHis17-BtubAB-msGFP</td>
<td>Protein is soluble but cannot form filaments with BtubA. Results so far suggest that attaching domains on BtubB interferes with filament assembly.</td>
</tr>
<tr>
<td>Construct</td>
<td>Description</td>
</tr>
<tr>
<td>----------------------------</td>
<td>----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>pHis17-BtubA-msGFP-6His</td>
<td>Expresses well, but degrades a bit over the course of purification. BtubA-msGFP-6His polymerise well with BtubB, judging from the results of pelleting assays and fluorescent microscopy. However, the msGFP is too flexible judging by 2D classification results from cryoEM data.</td>
</tr>
<tr>
<td>pHis17-BtubA(1-431)-msGFP-6His</td>
<td>Expresses well, no degradation problem but doesn’t form filament with BtubB.</td>
</tr>
<tr>
<td>pHis17-BtubA(1-436)-msGFP-6His</td>
<td>Expresses well, no degradation problem but doesn’t form filament with BtubB.</td>
</tr>
<tr>
<td>pHis17-BtubA(1-440)-msGFP-6His</td>
<td>Expresses well, no degradation problem and filament observed under EM. However, judging by the 2D classification results from cryoEM data, the msGFP is still too flexible to aid the software to distinguish BtubA from BtubB.</td>
</tr>
</tbody>
</table>