# Integrative approach to analyze the proximal region of the *Trichonympha agilis* centriole

**EPFL**

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Suisse
2018

Ivanovi a Samkovi

# Acknowledgements

I would like to thank many people who were involved directly or indirectly in this project and who helped me to cope with technical difficulties, obstacles or bad moods and enjoyed every big or little achievement with me.

First of all, I thank my supervisor Pierre for giving me the opportunity to work on this challenging project, for his guidance and encouragement to collaborate with many people to enrich my skills.

Then, I thank Virginie and Paul, two amazingly talented scientists who initiated this project, for their guidance and advice at the beginning of my PhD studies. I also thank Paul and Maeva for their enormous help towards the end.

I thank Michel and Natacha for inviting me to their lab at Paul Scherrer Institute and their willingness to teach me core structural biology.

I thank George for his amazing knowledge and experience of structural biology he brought to our lab and his patience to answer my questions even during very busy days. Niccolò, for helping me to solve all types of computational problems and for his mastering of high-speed atomic force microscopy, which I used to observe the behavior of my proteins. Akshari, for just sitting next to me, for her help in processing my data and for our sometimes infinite discussions about the centriolar world. I thank also Jian, the master of native gels; Alex, for our discussions about strange beasts and their beautiful non-canonical centrioles; Radek, for our secret language and bringing me some home-feeling and fun to everyday lab life; Marie, for her small encouragements that improved my days; Tessa, for her positive attitude, and Kerstin, for clear explanations of subjects where I was lost. I cannot forget Tania for our lunch running breaks, Coralie for her encouragement to be eco-friendly and Tatiana for confirmation that "the sky is no limit" if you can build your own aircraft. I thank all of them with all former lab members for making Pierre's lab a nice place to work.

I thank Prof. Petr Leiman for advice in centriole purification, Davide from the Electron Microscopy Facility for doing an excellent job and for his patience in preparing and screening my difficult samples and tomogram acquisition and Sergey for his input in processing my tomograms. I thank also Florence and Aline from the Protein Crystallization Core Facility for help during each step of protein crystallization and Romain from the Proteomic Facility for mastering mass spectrometry.

I want to thank also all members of Michel's lab: Andrea, Sarah, Anil, Ashwari, Sébastien and Marcel for making my stays at PSI agreable and funny; all the AFM people I met during my studies:

# Abstract

The centrosome, the major microtubule organizing center in animal cells, is composed of two orthogonally arranged centrioles and surrounding pericentriolar material. Centrioles are evolutionary conserved organelles characterized by a nine-fold symmetry of microtubules and are built around a cartwheel in their proximal region. The overall architecture at ~ 40 Å resolution of the exceptionally long proximal region of *Trichonympha* spp. basal body has been uncovered previously using cryo-electron tomography (cryo-ET) followed by subtomogram averaging. The resulting 3D map revealed not only a central cartwheel hub that can accommodate rings of the evolutionarily conserved SAS-6 proteins, but also novel features, whose identity is unknown. In this thesis, we used *T. agilis*, which is the only *Trichonympha* sp. with an assembled genome and transcriptomic data, to gain some more insight into the structure, composition and first steps of cartwheel assembly.

With the aim of generating a high-resolution 3D model of the *T. agilis* proximal region, we used cryo-ET and subtomogram averaging to refine the structure of the cartwheel hub and cartwheel inner densities (CID). We uncovered that the CID is polarized, and discovered inter-cartwheel pillars (ICP) between superimposed cartwheel rings. Moreover, we discovered and analysed a new type of the cartwheel of *Teranympha mirabilis* with a novel type of cartwheel central densities (CCD).

We also analyzed the composition of the proximal region of the *T. agilis* centriole using a proteomic approach. We optimized a density gradient centrifugation protocol for isolation of the *T. agilis* rostrum, which harbors ~ 1400 centrioles, and performed protein correlation profiling. In this manner, we have identified 47 candidates of the rostrum proteome, including three new *T. agilis* SAS-6 homologs, which were further characterized.

Furthermore, we addressed the oligomerization properties of human and two *T. agilis* SAS-6 proteins (HsSAS-6, TaSAS-6_1 and TaSAS-6_2) *in vitro*. We found that HsSAS-6 requires an interaction partner in order to be stabilized and form rings and stacks. In contrast, TaSAS-6_1 did not oligomerize in our setup, and TaSAS-6_2 is able to oligomerize and form different types of oligomers, rings included.

In conclusion, the structure and composition of the proximal region of the *T. agilis* centriole uncovered here are anticipated to be of a general significance for studies of the biogenesis and structure of centrioles.

# Résumé

Le centrosome, principal centre organisateur de microtubules dans les cellules animales, comporte deux centrioles structurellement très complexes et arrangés de manière orthogonale ainsi qu'une matrice péricentriolaire. Les centrioles sont des organites conservés au cours de l'évolution et caractérisés par une symétrie radiale d'ordre neuf. Ils sont construits autour d'une structure appelée « roue de voiture » et située dans leur région proximale. L'architecture globale de la région proximale du centriole a été découverte par tomographie cryoélectronique sur des échantillons de *Trichonympha* spp, chez qui cette structure est particulièrement allongée. La résolution de 40 Å ainsi obtenue a permis de découvrir non seulement la structure exceptionnelle de la roue de voiture, où les anneaux de la protéine SAS-6 peuvent être placés, mais aussi de nouvelles densités dont l'identité est inconnue.

Dans cette thèse, nous avons utilisé les centrioles de l'espèce *T. agilis*, seule *Trichonympha* sp., dont le génome est assemblé et les données transcriptomiques disponibles, pour mieux comprendre la structure, la composition et les premières étapes de l'assemblage de la « roue de voiture ».

Premièrement, dans le but de générer un nouveau modèle 3D de la région proximale du centriole qui servira de nouvelle référence, nous nous sommes concentrés sur la détermination de la structure de la partie moyeu de la roue de voiture de *T. agilis*. Nous avons affiné la structure du moyeu de la roue de voiture ainsi que celle des densités visibles en son centre (densités internes ou CID). Cela a permis de découvrir la polarité du CID et de nouvelles densités entre les moyeux. Nous avons aussi identifié et analysé un nouveau type de la roue de voiture de *Teranympha mirabilis* avec différent type des densités internes.

Deuxièmement, nous avons analysé la composition de la région proximale du centriole de *T. agilis*. Nous avons pour cela optimisé le protocole de purification par centrifugation par gradient de densité du rostrum de *T. agilis* qui abrite ~ 1400 centrioles. Nous avons ainsi identifié 47 protéines du rostrum et trouvé trois nouveaux homologues de SAS-6 de *T. agilis* qui ont ensuite été caractérisés.

Troisièmement, nous avons abordé les propriétés d'oligomérisation du SAS-6 humain et de deux homologues de *T. agilis*, TaSAS-6_1 et TaSAS-6_2 *in vitro*. Ces protéines sont impliquées au début de la formation du centriole et forment des roues de voiture de longueur et de stabilité différentes au cours du cycle cellulaire. Nous avons pu déterminer les propriétés d'auto-oligomérisation pour un seul des TaSAS-6 et l'exigence d'un partenaire d'interaction dans la fonction d'oligomérisation de HsSAS-6.

Pour conclure, la meilleure résolution et la détermination de la composition de la région proximale du centriole de *T. agilis* obtenues ici devraient avoir une importance générale pour les études de la formation et de la structure des centrioles.

Mots clés : centriole, *Trichonympha*, tomographie-cryoélectronique, sub-tomogrammes moyennés, protéomique, SAS-6

# Contents

# Chapter 1    **Introduction**

The centrosome is a tiny organelle discovered independently in the same year, more than one century ago, by Walter Flemming (Flemming, 1875) and Édouard van Beneden (van Beneden, 1875). Firstly described as a "polar body" or a "polar corpuscule", the centrosome was named as such by Theodor Boveri 13 years later based on the observation that it acts as an organizing centre of cell division (Fig. 1.1) (Boveri, 1887).
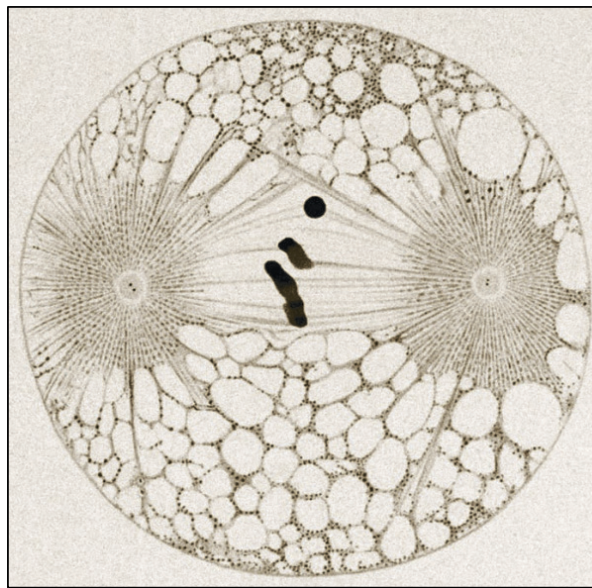


Fig. 1. 1 Centriole pairs within centrosomes. Lithographical illustration of the 1st cell division in the horse parasite *Ascaris megalocephala bivalens* drawn by Theodor Boveri in 1901. The figure shows microtubules arranged in a bipolar spindle with chromosomes in the middle and two centrioles per spindle pole, forming two centrosomes. Note the quality of the illustration and all the cellular details that were already visible more than one century ago (Boveri, 1901).

Van Beneden and Boveri observed independently that the centrosome is a persistent structure, acting as an organizing center around which the whole cellular content is arranged, and that one centrosome is transmitted to each daughter cell after its duplication

(van Beneden and Neyt, 1887; Boveri, 1887). „Das Centrosoma" became a key, even if very small, organelle leading to proper cell division and whose aberrations were postulated already in the early years of centrosome research to be a potential cause of cancer (Boveri, 1914). So, what exactly is the centrosome and what do we know about it?

In the introduction, I will firstly shortly describe the cell division cycle and then portray the centrosome as we know it now. I will describe its structure and its two main components, centrioles and the pericentriolar material. Then, I will focus on centriole biogenesis and will detail the composition of centrosomes and centrioles. I will finish by the description of centrioles in some model organisms.

## 1.1    Cell cycle and cell division

A distinctive feature of living organisms compared to non-living physical entities is their capability of growth and subsequent replication or reproduction to begin a new life cycle. Survival of a multicellular organism does not rely on survival of every single cell within it, but rather on the survival of the whole entity, whereby damaged or very old differentiated cells are replaced by new ones to ensure the continuum of life.

The process through which one so-called mother cell divides and gives rise to two daughter cells is called cell division. We distinguish two types of cell division: mitosis, when a mother cell divides into two daughter cells with an identical DNA content, and meiosis, whereby in a series of two meiotic divisions, a mother cell eventually give rise to four daughter cells with half of the amount of the starting genetic material.
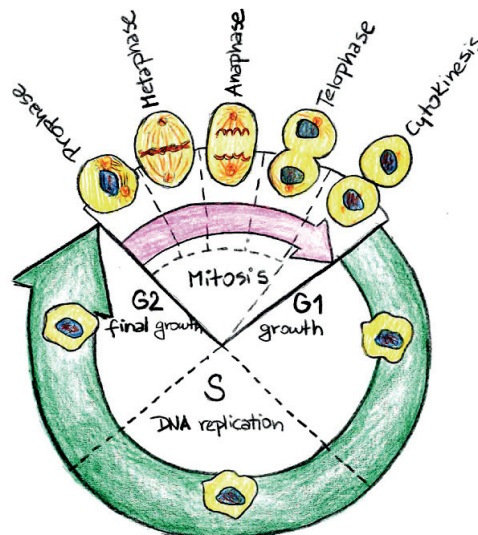


Fig. 1. 2 Simplified representation of the cell cycle with schematic depiction of a cell in individual cell cycle phases. Note that the relative duration of phases is not depicted in real time scale.

The cycle of proliferating cells (Fig. 1.2) is divided into two cytologically distinguishable phases: interphase and mitosis. In interphase, we further distinguish the Gap1 (G1) phase during which the cell grows. When the cell finishes G1 and transits to the next cell cycle phase, the centrosomes start to replicate. The next stage is the synthesis (S) phase when the DNA is replicated, and then the Gap2 (G2) phase when the cell grows further, the two centrosomes separate from one another and the cell prepares for mitosis (M phase). During mitosis, pairs of sister chromatids are pulled by microtubules organized into a bipolar spindle, with one centrosome per spindle pole, to opposite sides of the cell. Mitosis is further divided into four distinct stages, each characterized by a distinct spatial organization of spindle microtubules and chromosomes: prophase, metaphase, anaphase and telophase. After proper separation of the genetic material, the cell physically divides during cytokinesis and gives rise to two daughter cells, each with one centrosome, thus completing the cell cycle.

## 1.2 Overal organization of the centrosome

Structurally, we can describe the centrosome as a non-membraneous organelle that is composed of two major parts: two evolutionary conserved and highly organized barrel shaped centrioles, surrounded by a less structured but very dense network-like scaffold called the pericentriolar matrix or material (PCM), which shapes the centrosome into a spherical organelle.

Let's have a closer look into the amazing structure of canonical centrioles and centrosomes. Firstly, I will describe centrioles and the structural differences between the two centrioles within the centrosome. Secondly, I will discuss the pericentriolar material and its structure.

### 1.2.1 Centriole

The centriole, with its microtubule based cylindrical structure of usually conserved dimensions ~ 250 nm wide and ~ 500 nm long, can be recognized among other cell organelles by conventional or cryo-electron microscopy (Fig. 1.3 A) (reviewed in Bornens, 2012; Winey and O'Toole, 2014). Nine microtubule triplet blades are arranged with radial symmetry into a barrel shape. The triplets are twisted anticlockwise when looked from the proximal end (Fig. 1.3 B), giving the centriole a turbine-like appearance. Only one of the three microtubules in the triplets is a complete one, composed of 13 protofilaments (Tilney et al., 1973). This microtubule is named the A-tubule and is elliptical in cross-section (Guichard et al., 2013). The two others microtubules, called the B- and C-tubule, respectively, are incomplete, being composed of only 10 protofilaments and merging with the wall of the A- or B-tubule, respectively (Guichard et al., 2013; Li et al., 2012; Tilney et al., 1973). The lumen of microtubules is hollow, with the exception of an unknown microtubule inner protein (MIP) density inside the A-tubule (Fig. 1.3 D) (Guichard et al., 2013; Li et al., 2012) and a filament inside the C-tubule present only in the distal part of the tri-

plet (Fig. 1.3 C) (Li et al., 2012). Moreover, a proteinaceous ~ 130 Å long A-C linker connects neighboring microtubule triplets, bridging the A-tubule of one triplet with the C-tubule of the neighboring microtubule triplet (Fig. 1.3 D) (Guichard et al., 2013).

Centrioles can be divided into three distinct regions, based on the presence or absence of given structural components: a proximal region of ~ 100 nm with microtubule triplets bearing a cartwheel (reviewed in Hirono, 2014), a central core with microtubule triplets and specific internal densities (Fig. 1.3 C) (Li et al., 2012), and a distal part where the centriole harbors only A-B-microtubule doublets (reviewed in Winey and O'Toole, 2014). This last region is where distal and subdistal appendages form on the outside of the barrel during centriole maturation (Fig. 1.3 A).
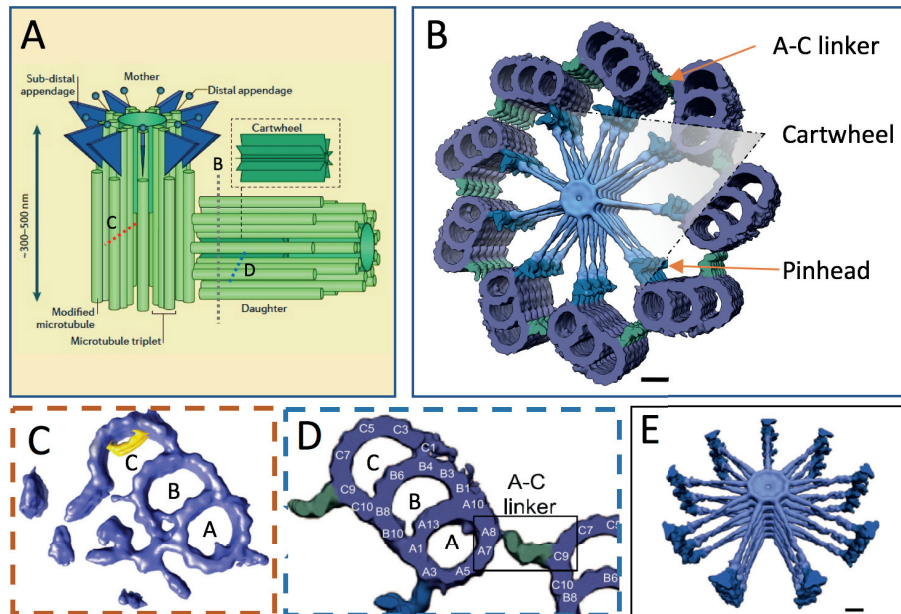


Fig. 1. 3 Centriole structure. **A** Schematic representation of two orthogonaly arranged centrioles with a newly born procentriole harboring the cartwheel at the proximal side. Dashed lines indicates where three cross sections through the centriole were taken to visualize the images shown in B, C and D; with permission from Conduit et al. (2015) **B** Cryo-electron tomography (cryo-ET) map of the proximal region of the *Trichonympha* spp. centriole, seen from the proximal end. Scale bar: 20 nm. **C** Cryo-ET map of the microtubule triplets from the distal part of the *Chlamydomonas reinhardtii* centriole, showing the specific filament inside the C-tubule (in yellow). **D** Magnification of the microtubule triplets from the proximal part of the *Trichonympha* spp. centriole, highlightning the A-C linker between two microtubule triplets. **E** Cartwheel stack. Note double-layers of cartwheel rings. Scale bar: 10 nm. B, D and E adapted from Guichard et al. (2013), C adapted from Li et al. (2012).

The cartwheel is an internal structure of the centriole located in its proximal region (Fig, 1.3 B and E). As its name implies, the cartwheel is composed of a central hub from which

nine radially oriented and symmetrically arranged spokes emerge. The entire cartwheel structure is composed of several layers of cartwheel rings. Based on the data from the elongated *Trichonympha* spp. centriole (Guichard et al., 2012, 2013), two superimposed cartwheel rings merge through their spokes and form double layers that then stack and are connected with microtubule triplets through a Pinhead structure located next to the A-tubule (Fig. 1.3 B). The cartwheel stack exhibits periodicities of 8.5 nm along the centrally located cartwheel hub and of 17 nm close to the peripheral Pinhead (Fig. 1.3 E) (Guichard et al., 2013).

### 1.2.2   Pericentriolar material

The centrosome is a non-membraneous organelle that is shaped mostly by the PCM surrounding the two centrioles. It has clear boundaries as observed by cryo-electron microscopy and super-resolution microscopy techniques (reviewed in Woodruff et al., 2014). The ensemble of PCM proteins was long believed to be only an amorphous unstructured protein mass (Robbins et al., 1968), probably due to the limitations in microscopy techniques. However, recent studies have shown that the pericentriolar material is a complex porous network that can be reconstitued *in vitro* (Woodruff et al., 2015). This network serves notably as a scaffold for microtubule nucleation, as well as for other proteins with various functions throughout the cell cycle. Many PCM components have a discernible localization as shown by superresolution microscopy (Lawo et al., 2012; Mennella et al., 2012).

## 1.3      Centrosome duplication

The centrosome duplication cycle is tightly regulated and coincides with the overall cell cycle, thus ensuring the formation of only one procentriole next to each pre-existing centriole, once per cell cycle (reviewed in Firat-Karalar and Stearns, 2014; Sluder, 2014). Cells in the G1 phase contain one centrosome with two centrioles that are connected by a loose proteinaceous linker (reviewed in Agircan et al., 2014; Nigg and Stearns, 2011). Two important constituents of this linker are rootletin and C-Nap1, which together connect the base of the two centrioles (Mayor et al., 2000). The older mature centriole is called the mother centriole, and the younger one, which has almost reached its final length and maturity, the daughter centriole (Fig. 1.4 A). The mother centriole can be distinguished from the daughter centriole by the presence of distal and subdistal appendages that marks the final maturity (Vorobjev and Chentsov, 1982). As the cell progress through the G1 phase, the two centrioles recruit factors that will be needed for the assembly of new procentrioles; these factors are organized in a toroidal shape around both mother and daughter centriole. Procentriole biogenesis is initiated approximately at the transition between the G1 and the S phases, and continues througout the remainder of the cell cycle. The two centrioles, each with their stably engaged procentriole, remain connected through the linker until late G2 when the linker is severed following

phosphorylation of Nek2A kinase, thus allowing centrosome separation and bipolar spindle formation.
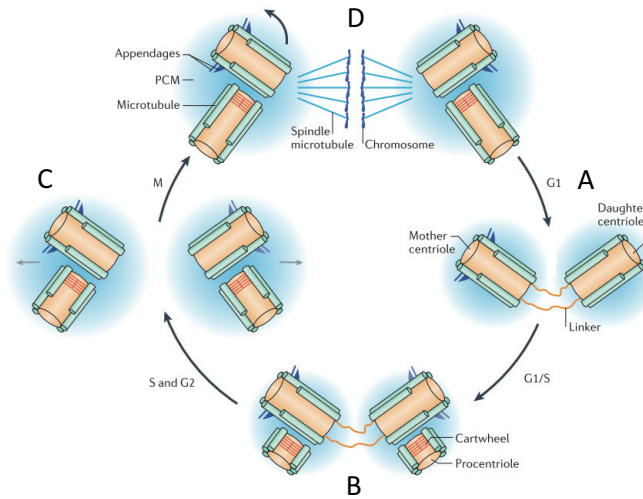


Fig. 1. 4 Schematic representation of the centrosome duplication cycle. **A** Mother and daughter centrioles loosely connected by a proteinaceous linker in the G1 phase. **B** Initiation of centriole duplication at the base of each centriole around the G1/S phase transition and procentriole growth until mitosis. Note orthogonal arrangment between pre-existing centriole and procentriole. The daughter centriole reaches its maturity by the acquisition of distal and subdistal appendages. **C** The two centrosomes with two orthogonally arranged centrioles separate at the onset of mitosis to form the two poles of the mitotic spindle (D). With permission from Gönczy (2015).

At the onset of mitosis, the daughter centriole achieves maturity by the acquisition of distal and subdistal appendages. The end of mitosis is marked by centriole disengagement, whereby the procentrioles loose their orthogonal configuration with respect to the centrioles, but as stated above, remain connected by a linker until the next mitosis.

### 1.3.1 Centriole biogenesis

The assembly of the procentriole is achieved through several steps of a centriole biogenesis pathway whose fundamental components are conserved across evolution (Fig. 1.5) (reviewed in Gönczy, 2012). The proteins identified firstly through studies in *C. elegans* includes recently discovered SAS-7 at the onset of the pathway (Sugioka et al., 2017), SPD-2 (Kemp et al., 2004; Pelletier et al., 2004), the kinase ZYG-1, SAS-6 (Dammermann et al., 2004; Leidel et al., 2005), SAS-5 (Dammermann et al., 2004; Delattre et al., 2004) and SAS-4 (Kirkham et al., 2003; Leidel and Gönczy, 2003)
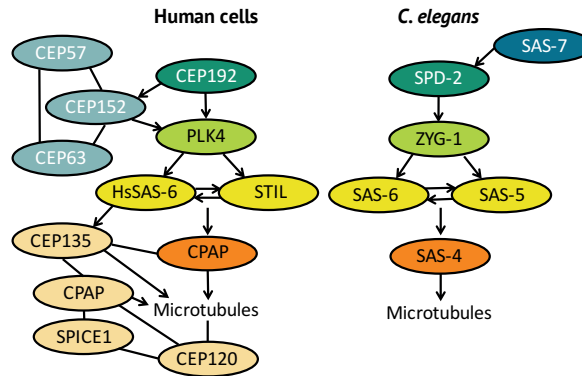
Fig. 1. 5 Core components of the centriole biogenesis pathways conserved across evolution. Schematic representation of known proteins involved in centriole assembly in cycling human cells, as well as *C. elegans* embryos. Proteins homologs or proteins with similar functions are depicted in the same color in the two species. Lines represent physical interactions and arrows the requirement of one component for the continuation of the biogenesis pathway. More information is available in the text. Modified from Gönczy (2012) and Banterle and Gönczy (2017).

Centriole duplication in cycling human cells starts with the recruitment of scaffolding proteins, the partially redundant proteins Cep192 (Gomez-Ferreria et al., 2007; Joukov et al., 2010; Zhu et al., 2008), which is a homolog of *C. elegans* SPD-2, as well as Cep152 (Dzhindzhev et al., 2010; Sonnen et al., 2013). Cep192 and Cep152 form a barrel-like and toroidal structure, respectively, around the centriole, (Lukinavičius et al., 2013a; Sonnen et al., 2012). Both proteins interact with the ZYG-1-related kinase Polo-like-kinase 4 (Plk4) (Kim et al., 2013; Sonnen et al., 2013), thus recruiting it on the torus in preparation for centriole assembly. Additionally, the proteins Cep63 (Brown et al., 2013) and Cep57 (Lukinavičius et al., 2013a) associate with Cep152 and ensure the efficient accumulation of factors needed for procentriole formation.

Plk4 localized on the torus around the proximal end of the pre-existing centriole then focuses onto one spot during the G1/S phase transition, thus probably defining the place from where a procentriole will emerge (Ohta et al., 2014). The Plk4 kinase binds and phosphorylates the SAS-5 relative STIL (Arquint et al., 2015; Kratz et al., 2015; Moyer et al., 2015; Ohta et al., 2014), and this enables phosphorylated STIL to then interact with HsSAS-6 to promote cartwheel assembly (Kratz et al., 2015; Moyer et al., 2015; Ohta et al., 2014). Moreover, Plk4 homodimerizes and autophosphorylates, which leads to its degradation via recruitment to the SCF-Slimb ubiquitin ligase complex member βTrCP (Guderian et al., 2010; Holland et al., 2010; Rogers et al., 2009; Sillibourne et al., 2010). However, the binding of Plk4 with STIL protects the kinase from degradation, which appears to be key to enable Plk4 to promote procentriole assembly (reviewed in Arquint

and Nigg, 2016). The negative auto-regulation of Plk4, the reciprocal positive regulation of STIL and Plk4, as well as oligomerization properties of STIL (David et al., 2016) are likely important to ensure the formation of a single focus around the mother and daughter centriole, where one procentriole each will assemble (Kim et al., 2016; Ohta et al., 2014).

As mentioned earlier, STIL in its active form binds HsSAS-6, and this interaction facilitates the loading of HsSAS-6 onto the focal point of procentriole assembly (Ohta et al., 2014). SAS-6 proteins in general and HsSAS-6 in the case of human cells, are key for imparting the characteristic nine-fold symmetry of the cartwheel and the centriole (reviewed in Gönczy, 2012). HsSAS-6 is present as a homodimer in the cytoplasm, and it is thought that only by being locally concentrated at the centriole it can undergo higher order oligomerization and form the cartwheel (Keller et al., 2014). Cellular levels of HsSAS-6 are tighly regulated during the cell cyle by APC$^{CCdh1}$ and SCF/FBXW5-mediated proteolysis (Puklowski et al., 2011; Strnad et al., 2007), in the latter case modulated by Plk4 (Puklowski et al., 2011). HsSAS-6 higher order oligomerization results in the formation of ring-like structures that stack, thus forming the entire cartwheel structure. SAS-6 proteins are a fundamental component of the centriole and also one of the main subjects of my PhD research project; therefore, the description of SAS-6, its main functions in the centriole and its higher-order oligomerization will be described in more detail in Chapter III.

The HsSAS-6-based cartwheel stack is then connected to growing centriolar microtubules through a structure called the Pinhead (Fig. 1.3 B) (Guichard et al., 2013). The identity of Pinhead proteins has not been uncovered with certainty as of yet, but Cep135 is a promising candidate. Bld10p, the *Chlamydomonas* homolog of Cep135, localizes to the inner wall of the microtule triplets (Matsuura et al., 2004), and various truncations of the Bld10p protein expressed in the background of a *bld10p* null mutant lack the connection between the cartwheel spokes and the microtubule triplets (Hiraki et al., 2007). Moreover, it has been established that Cep135 binds the C-terminal region of HsSAS-6, which forms the peripheral end of cartwheel spokes (Lin et al., 2013b), and also that Cep135 binds and bundles microtubules (Carvalho-Santos et al., 2012; Kraatz et al., 2016). The mechanism controlling the final height of the cartwheel is not known, but different hypotheses were proposed, such as the presence of a molecular ruler, a concentration gradient or structural instability of longer entities (reviewed in Banterle and Gönczy, 2017). Microtubule triplets around the cartwheel, composed of α-/β-tubulin protofilaments, are templated by γ-tubulin (Oakley and Oakley, 1989; Oakley et al., 1990). The formation of triplets is ensured by the presence of the minor tubulin isoforms δ (Dutcher and Trabuco, 1998; O'Toole et al., 2003) and ε (Chang and Stearns, 2000). These isoforms, absent from species which do not form triplets, such as *C. elegans* with its microtubule singlets and *Drosophila* with microtubule doublets (reviewed in Carvalho-Santos et al., 2011; Gupta and

Kitagawa, 2018) act to stabilize proper centriolar structure by maintaining the microtubule triplets (Wang et al., 2017).

The centriole then elongates through the activity of CPAP and Cep120, whose overexpression induce the formation of very long centrioles (Comartin et al., 2013; Kohlmaier et al., 2009; Lin et al., 2013b, 2013c; Schmidt et al., 2009; Tang et al., 2009). CPAP and Cep120 interact with SPICE1, which is also required for centriole formation (Comartin et al., 2013). Furthermore, CPAP slows down microtubule growth and stabilize them (Sharma et al., 2016), whereas CP110 functions as a centriole capping protein, which is removed as the centriole elongates and templates the axoneme (Comartin et al., 2013; Kohlmaier et al., 2009; Schmidt et al., 2009; Spektor et al., 2007).

However, the centriole reaches its final length only in the next cycle. During the subsequent interphase, when assembling its own procentriole, the daughter centriole acquires subdistal and distal appendages, which include the centrosomal proteins Cep164, Odf-2, Cep89, Sclt1 and FBF1, which are all essential for ciliogenesis (Graser et al., 2007; Ishikawa et al., 2005; Tanos et al., 2013)

### 1.3.2 Pericentriolar matrix

Two important components of the PCM in the context of centriole biogenesis are Cep192 and Cep152, the latter surrounding the centriole and forming a scaffold for centriolar protein recruitment (Dzhindzhev et al., 2010; Giansanti et al., 2008; Gomez-Ferreria et al., 2007; Kemp et al., 2004; Pelletier et al., 2004; Varmark et al., 2007). However, other core pericentriolar proteins are also part of the "centromatrix", including CDK5RAP2 or Centrosomin/Cnn in *Drosophila* (Conduit et al., 2010; Lucas and Raff, 2007; Megraw et al., 1999; Vaizel-Ohayon and Schejter, 1999), one of the first identified PCM components, Pericentrin, PCNT (Dictenberg et al., 1998; Martinez-Campos et al., 2004; Tritarelli et al., 2004) and SPD-5 in *C. elegans* (Hamill et al., 2002). When the centriole is formed, an inner pericentriolar layer is formed at its base, composed of Pericentrin (Lawo et al., 2012; Mennella et al., 2012). The PCM is further expanded by the incorporation of PCM components, microtubule interacting proteins and micrutubule nucleators (Decker et al., 2011; Lee and Rhee, 2011). The majority of PCM components are coiled-coil containing proteins (Andersen et al., 2003; Hamill et al., 2002; Pelletier et al., 2004; Tritarelli et al., 2004; Vaizel-Ohayon and Schejter, 1999; Varmark et al., 2007), which suggest their involvement in protein binding and the potential for forming large inter-connected networks. Such an interconnected network is structurally organized in interphase, with the main proteins organized in toroids surrounding the centriole (Fu and Glover, 2012; Lawo et al., 2012; Mennella et al., 2012; Sonnen et al., 2012), and a lack of internal structure when the cell is in metaphase (Lawo et al., 2012).

## 1.4    Roles of centrosomes during the cell cycle

As observed already by Boveri and van Beneden, and as confirmed throughout more than one century of research, the centrosome functions as the major microtubule organizing center (MTOC) in animal cells. Together with chromosome- and microtubule-mediated nucleation pathways, centrosomes function notably to ensure proper mitotic spindle assembly and progress through mitosis (reviewed in Prosser and Pelletier, 2017). Moreover, centrosomes have many other important functions throughout the cell cycle, some of which are linked to their microtubule nucleation activity (reviewed in Arquint et al., 2014; Doxsey et al., 2005; Kellogg et al., 1994; Rieder et al., 2001).

### 1.4.1    Microtubules at centrosomes

The centrosome nucleates microtubules that usually have a regular conserved structure of 13 α-/β-tubulin protofilaments. However, tubulins at higher concentrations can polymerize *in vitro* into microtubules with higher number of protofilaments, whereas when the same amount of tubulin is added onto free centrosome, mostly microtubules with 13 protofilaments are formed (Evans et al., 1985). This fact led to the hypothesis that microtubules are somehow templated at the centrosome (Evans et al., 1985). A few years later, a probable centrosomal microtubule template was discovered (Moritz et al., 1995; Oakley et al., 1990; Vogel et al., 1997) as a multiprotein ring-like structure formed of γ-tubulin (Joshi et al., 1992; Stearns et al., 1991; Zheng et al., 1991) and γ-tubulin complex proteins (GCPs) arranged into a ring or short helix, called the γ-tubulin ring complex (γ-TuRC). Various models of microtubule nucleation from a γTuRC template were proposed (Keating and Borisy, 2000; Moritz et al., 2000), until the 8 Å structure of the major component of γ-TuRC, the γ-tubulin small complex (γ-TuSC) was solved, unraveling thirteen γ-tubulins per turn, matching precisely the 13-fold symmetry of microtubules (Kollman et al., 2010). The nucleation activity of individual γ-TuSC is weak, and the precise mechanism of γ-TuRC activation is still an unsolved issue, although some studies have shown that γ-TuRC activation requires the association with specific activator and adaptors proteins, including pericentriolar protein CDK5RAP2 (Choi et al., 2010) and its priming factor MZT1 (Cota et al., 2017) in human cells.

Furthermore, the conical structure of the γ-TuRC acts as a minus-end capping protein complex to prevent growth and depolymerization of microtubules from the minus-end and thus to control the +/- end orientation of microtubules (Wiese and Zheng, 2000).

Microtubules nucleated by the γ-TuRC in the pericentriolar material are then translocated and anchored by their minus-ends either at the level of subdistal appendages, on the mother centriole (Chrétien et al., 1997; Mogensen, 1999), or at microtubule-anchoring sites localized on the apical side of a cell in the case of epithelial cells (Mogensen et al., 1997). The process of microtubule anchoring is independent of microtubule nucleation, as it involves different proteins than those present in the γ-TuRC, including ninein

(Mogensen et al., 2000), centriolin (Gromley et al., 2005) or the dynein-interacting complex dynactin (Quintyne et al., 1999).

Different processes are linked to microtubule-nucleation from centrosomes. A few of them are discussed in turn below: organization of the interphase microtubule network, organization of the bipolar mitotic spindle, as well as templating of cilia and flagella specifically from the mother centriole. Different nucleation sites at centrosomes or centrioles are involved in these processes. Pericentriolar nucleation sites serve for building the microtubule network and the mitotic spindle, whereas the microtubule based axonemes of cilia or flagella are templated directly from centriolar microtubules.

### 1.4.2 The centrosome and its microtubule network in interphase

In interphase, the position of the centrosome (in G1) or the centrosome pair (after centrosome duplication) is in the cell center (Burakov et al., 2003), usually in close proximity to the nuclear envelope (Burakov and Nadezhdina, 2013). The centrosome is the main microtubule nucleation center of most interphase animal cells and, depending on the cell type, microtubules can stay anchored to the centrosome, as in monocytes (Cassimeris et al., 1986), or be released into the cytoplasm, as in epithelial cells (Müsch, 2004). Microtubule arrays nucleated from centrosomes are important for cell shape, polarity and motility (reviewed in Nigg and Raff, 2009). However, the presence and activity of centrosomes is partially dispensable during interphase, since microtubules can be nucleated and organize a microtubule network also in its absence. Thus, in unperturbed cells, the Golgi apparatus also functions as an MTOC (reviewed in Sanders and Kaverina, 2015). The Golgi apparatus MTOC recruits the same molecular factors that are indispensable for the microtubule nucleation activity at the centrosome, such as γ-TuRC or the γ-TuRC-interacting protein AKAP450 (Rivero et al., 2009). Furthermore, loss of centrosomal microtubule nucleation activity can be compensated by an increase of the Golgi apparatus MTOC activity (O'Rourke et al., 2014), and might ensure the overall survival of acentrosomal p53[-/-] cancer cells treated with a Plk4 inhibitor (Wong et al., 2015). This confirms the partial redundancy of centrosomal and microtubule nucleation activities.

### 1.4.3 Role of the centrosome in spindle assembly and cell division

As the centrosomes approach mitosis, their PCM expands, gaining enough capacity to nucleate the microtubules that will form the bipolar spindle. However, the discovery of two other microtubule-nucleating pathways operating during mitosis, that emanating from chromosomes and that coming from microtubules themselves (reviewed in Meunier and Vernos, 2016), as well as the absence of centrosomes in certain cell types, raises the question of what is the precise function of centrosomes in mitosis. It is now clear that a centrosome and a spindle pole are two distinct physical entities, and that many cell types assemble their spindle without centrosomes, such as higher plant cells (Hashimoto, 2013), oocytes of most animals (Dumont and Desai, 2012) or cells during the first embryonic

divisions of mouse embryos (Courtois et al., 2012). This applies also to cells that normally possess two centrosomes at spindle poles, but in which these centrosomes were physically (Hinchcliffe et al., 2001; Khodjakov et al., 2000) or genetically removed, as in *Chlamydomonas* (Matsuura et al., 2004; Zamora and Marshall, 2005), *Drosophila melanogaster* (Basto et al., 2006), late stages mouse embryos (Bazzi and Anderson, 2014) or cultured chicken cells (Sir et al., 2013). In all these cases, the mitotic spindle is assembled, but the absence of centrosomes causes a delay in spindle assembly, spindle abnormalities and chromosomal instability, demonstrating the importance of centrosomes in cells where they naturally occur. The presence of centrosomes adds efficiency in capturing chromosomal kinetochores and punctuality of cell division timing (Heald et al., 1997; Khodjakov et al., 2000; Sir et al., 2013; Szollosi et al., 1986).

### 1.4.4 Centrosome function in asymmetric cell division

Asymmetric cell division is the ability of mother cells in organisms ranging from bacteria to metazoan to produce two daughter cells with different fates. Either intrinsic cell fate determinants are localized asymmetrically in the mother cell in response to a polarized axis, or extrinsic factors, such as the microenvironment, determine the different fate of daughter cells after division. Studies in one-cell stage *Caenorhabditis elegans* embryos, as well as in *Drosophila* larval neuroblasts and sensory organ precursors (SOPs), provided insights into the mechanisms of intrinsic asymmetric cell division. In *C. elegans*, the paternally inherited centrioles and the resulting centrosome is the initiator of the symmetry breaking mechanism, whereby the position of the centrosome close to the cell cortex determines where anterior-posterior (A-P) embryonic polarity establishment starts, which marks the future posterior of the animal (reviewed in Gönczy, 2008; Gönczy and Rose, 2005). In *Drosophila* male germ line stem cells, the two centrosomes have significantly different sizes, with the older one being bigger and having a more robust microtubule array. This might be the reason why this centrosome is attached sooner to the cell cortex close to the stem cell niche, and thereby determines the stem fate of one of the daughter cells, whereas the other centrosome ends up in a differentiated gonial cell (reviewed in Gonzalez, 2008; Yamashita et al., 2007). In *Drosophila* larval neuroblasts, the apical centrosome nucleates microtubules before the basal centrosome, and thereby also establishes an asymmetry in fate (reviewed in Gönczy, 2008). In vertebrates, for example in mouse neocortex progenitors, centrosome age also determines cell fate, in this case because the centrosome with the older mother centriole remains in the neuronal progenitors, whereas the younger centrosome is inherited preferentially by differentiating cells (Wang et al., 2009), much like in the fly testis.

### 1.4.5 Centriole function in templating the cilia or flagella

The capacity of centrosomes, and more specifically of centrioles, to nucleate microtubules is critical for forming the axoneme, the microtubule based internal scaffold of the cilium

and the flagellum. Firstly, I will describe the function of these two organelles and, secondly, I will shortly describe their formation.

The motile cilium or flagellum is a whip-like structure present mostly on the surface of some unicellular species and in some differentiated cells of multicellular organisms. Motile cilia and flagella are basically the same structures with a similar function, that of moving liquid. In the case of motile cilia, which are present in many copies in specialized cells, such as the epithelial cells lining the trachea and bronchial tubes, ciliary movements are highly synchronous as they move liquid in one direction to clear debris from air passages (Afzelius, 1976). In contrast, the flagella of individual cells, such as the vertebrate sperm flagellum or the flagella of *Chlamydomonas* are utilized simply to swim. The conserved long cylindrical structure of the axoneme in either motile cilia or flagella consists of nine circularly arranged microtubule doublets connected to each other by the motor protein dynein, radial spokes, as well as two central microtubule singlets, together producing a characteristic 9 + 2 arrangement of mirotubules. Mutants of *Chlamydomonas* lacking the two central microtubule singlets are paralyzed, suggesting that they are essential for motility (Adams et al., 1981).

Another, structurally different, type of axoneme is that of the primary cilium, which lacks the inner microtubule singlets, hence yielding a 9 + 0 arrangement. The primary cilium is nonmotile and has a variety of sensory functions, such as mechanosensing, detection of chemicals or of light in photoreceptors, osmolarity, temperature (reviewed in Satir et al., 2010) and also gravity (Moorman and Shorr, 2008). Moreover, the primary cilium is essential for regulation of Hedgehog-ligand-induced signaling, which regulates development and patterning (reviewed in Goetz et al., 2009), and is implicated in growth control throught plateled-derived growth factor α, which resides at the cilium (Schneider et al., 2005). The primary cilium in neurons has functions even in some types of memory, through somatostatin signaling, which is critical fo object recognition memory in mice (Einstein et al., 2010).

When the cell enters the G0 quiescent or resting phase of the cell cycle, the mother centriole with its developed appendages can dock into the plasma membrane, forming a basal body (Vorobjev and Chentsov, 1982) through the appendages present at the distal end of the mother centriole. Microtubule doublets at the distal end of the basal body then function as template for the formation of the axonemal outer doublets. The two inner microtubule singlets, nonexisting in centrioles or basal bodies, are most likely templated by the γ-tubulin complex, which is depleted from the vicinity of basal bodies after axonemal microtubules are formed (McKean et al., 2003; Silflow et al., 1999). Interestingly, CP110, which caps and restricts the length of centrioles in cycling cells (Kleylein-Sohn et al., 2007; Kohlmaier et al., 2009; Schmidt et al., 2009), is removed during cilium formation, which allows the now docked centriole to elongate (Spektor et al., 2007).

The assembly and maintenance of the cilium or flagellum depends on the intraflagellar transport (IFT) protein machinery (reviewed in Taschner and Lorentzen, 2016), which

uses microtubules as double-track railways to transport cargo to the tip of the organelle on the B-tubule and back to the cell on the A-tubule (Stepanek and Pigino, 2016). The specialized region at the very proximal side of the cilium, called the transition zone, functions as a gate controlling entry and exit of proteins from this organelle (reviewed in Gonçalves and Pelletier, 2017).

As we see from the above, the centrioles and basal bodies are related structures. Therefore, for simplicity I will refer to them only as centrioles in the text from now on.

### 1.4.6   Other centrosomal functions

As mentioned above, centrosomes play important, but not essential roles for mitotic spindle assembly and microtubule nucleation in general, and are absolutely required for cilium and flagellum assembly. Do centrosomes have other non-microtubule based functions?

The network-like PCM is an excellent platform to concentrate a variety of signalling molecules, checkpoint proteins and also tumor suppresors into one place throughout the cell cycle (Nigg et al., 2014), making the centrosome a powerful signalling hub.

A non transformed cell arrests in the G1 phase of the cell cycle if centrioles or centrosomes are lacking (Hinchcliffe et al., 2001; Khodjakov and Rieder, 2001; Wong et al., 2015). It is now clear that the dependence of cell cycle progression on the presence of centrosomes reflects stress-related activation of the tumor supressor p53 and the p38 kinase (reviewed in Arquint et al., 2014; Gönczy, 2015), as evidenced by the fact that cells can continue to proliferate even in the absence of centrosomes when p53 is lacking (Wong et al., 2015). Even if centrosome removal does not prevent cells to enter mitosis (Hinchcliffe et al., 2001; Khodjakov and Rieder, 2001), cells require the centrosomal localization of Cdk1 and proteins modulating Cdk1 activity in order to ensure proper progress through mitosis (Krämer et al., 2004). The centrosomal localization of AIR-1 in *C. elegans* embryos furthermore ensures timely regulated mitotic entry and promotes efficient nuclear envelope breakdown (Hachet et al., 2007; Portier et al., 2007). It was also shown that unfertilized starfish oocytes lacking centrosomes are able to progress into meiosis and initiate zygote development after activation by injection of human centrosomes (Picard et al., 1987). Moreover, in *Drosophila* syncytial embryos, the centrosome is involved in responding to damaged or incompletely replicated DNA by its own disruption after accumulation of the checkpoint kinase 2 DmChk2. This leads to defects in spindle assembly and removal of defective nuclei from the embryo (Takada et al., 2003). This short overview of various non-microtubule-based functions gives us an indication that the centrosome functions as a signalling hub throughout the cell cycle and, even if it is dispensable in some cases, its presence provides accuracy to different cellular processes.

## 1.5 Model organisms to study the centriole

Human cells possess canonical centrioles, whose structure, biogenesis and function were described in detail above. Despite the striking conservation of centriole architecture across evolution, we find noteworthy differences already among usual model organisms. Those that will be relevant to aspects of this project are described below.

### 1.5.1 Basal body apparatus of *Chlamydomonas reinhardtii*

Green algae, as well as some lower vascular plants, such as ferns or cycads, possess a variable number of flagella (from 2 to 20,000) at least in one stage of development, which emerge from structurally canonical centrioles (Carvalho-Santos et al., 2010). Among these species, the green alga *Chlamydomonas* is a powerful model organism due to the ease of cultivation in laboratory conditions as well as the availability of tools for molecular genetic studies (Fig. 1.6 A). Furthermore, the simple isolation of nonmotile mutants enables one to select and study genes involved in the structure and function of the flagellum and centrioles (reviewed in Harris, 2001).

In non-dividing *Chlamydomonas* cells, two of the four centrioles of the basal body complex dock into the plasma membrane to template the formation of two motile flagella. All four canonical centrioles keep their cartwheel, which makes *Chlamydomonas* a model suitable to study not only overall centriolar structure, but also the short proximal centriolar region harboring the cartwheel (Fig. 1.6 B). The overall structure of the *Chlamydomonas* centriole is similar to that of the human centriole, with the only difference that the *Chlamydomonas* centriole is slightly longer, going up to ~ 550 nm in *Chlamydomonas* compared to ~ 400-500 nm in human cells.
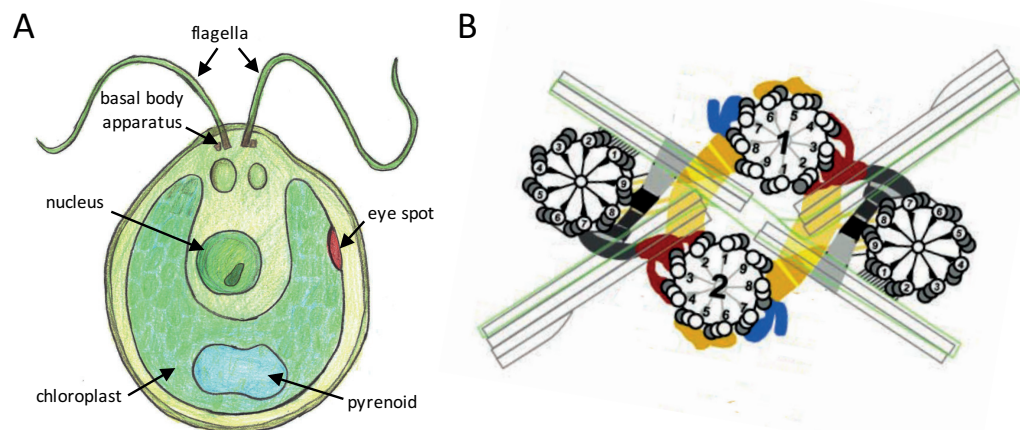


Fig. 1. 6 *Chlamydomonas* and its basal body apparatus. **A** Schematic representation of the unicellular green alga *Chlamydomonas*. **B** Schematic representation of the basal body apparatus with two mother and two daughter centrioles, called probasal bodies (labelled by number 1 and 2). With permission adapted from Geimer and Melkonian (2004).

### 1.5.2 Unusual centriole of *Caenorhabditis elegans*

Small, transparent and easy to cultivate, the nematode *C. elegans* was the first animal with a completely sequenced genome (Fig. 1.7 A) (The *C. elegans* Sequencing Consortium, 1998). The ability to produce mutants through different methods (reviewed in Kutscher and Shaham, 2014), as well as to study loss-of-functions phenotypes by well established RNA interference (reviewed in Conte et al., 2015), or nowadays the use of genome editing techniques based on the clustered regularly interspersed short palindromic repeats (CRISPR)-Cas9 system, render *C. elegans* a powerful model organism.

*C. elegans* possesses a non-canonical microtubule based centriole (Fig. 1.7 B), with nine radially arranged microtubule singlets instead of microtubule triplets. These singlets are connected to an electron dense structures surrounding them termed „paddle wheels", which localize along the entire length of each microtubule singlet (Pelletier et al., 2006; Sugioka et al., 2017). The diameter of the *C. elegans* centriole is smaller compared to human or *Chlamydomonas* centrioles, ~ 110 nm and 175 nm, without and with paddle-wheels, respectively. The length of the centriole is also smaller, being only ~ 200 nm (Fig. 1.7 B).

Another difference comes from the internal structural organization. Here, the canonical cartwheel with nine emanating spokes connecting to the microtubule triplets is replaced by an electron dense central tube connected directly to microtubules, as well as an inner tube (central hub) with a diameter of 18 nm (Sugioka et al., 2017), which likely functions as a cartwheel (Fig. 1.7 B). Structural work on the SAS-6 protein from *C. elegans* revealed that the inner cartwheel tube might be formed by higher order SAS-6 oligomers self-assembled in a spiral arrangement (Hilbert et al., 2013).

Despite the differences with more canonical centrioles, studies in *C. elegans* have had an important impact and revealed that even though centrioles in this species are structurally different, the biogenesis pathway uncovered in worms is conserved across evolution (reviewed in Banterle and Gönczy, 2017; Gönczy, 2012).
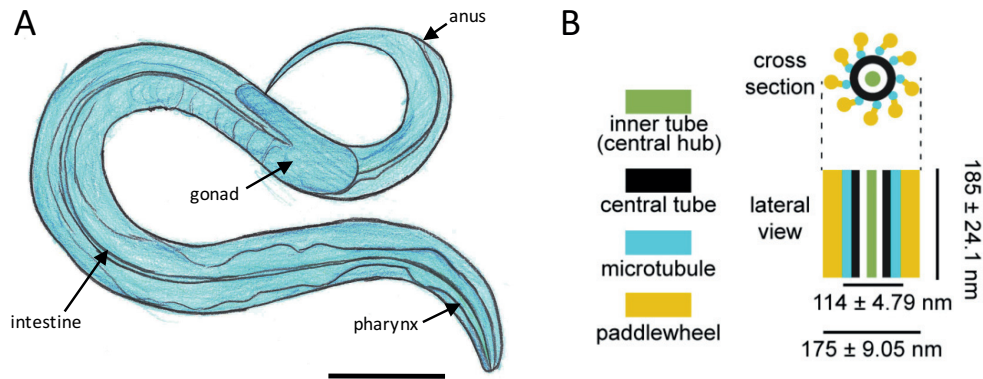
Fig. 1. 7 Centriole of *C. elegans*. **A** Schematic representation of a *C. elegans* hermaphrodite. Note that only the posterior gonad is depicted. Scale bar 0.1 mm. **B** Schematic representation of *C. elegans* centriole. With permission from Sugioka et al. (2017).

### 1.5.3   *Trichonympha* **and its long centriole**

The flagellated protozoan *Trichonympha* spp., with its thousands of centrioles per cell, has attracted attention of centriole biologists for more than 50 years (Fig. 1.8 A) (Gibbons and Grimstone, 1960; Grimstone and Gibbons, 1966). *Trichonympha* spp. is extremely motile in the dense environment of the termite gut, and harbors thousands of flagella templated by very elongated centrioles. The length of *Trichonympha* sp. centrioles spans ~1.5 μm > ~ 5 μm; most interestingly to us, the proximal region harboring the cartwheel is extremely elongated and represents ~ 90% of the entire length of the centriole (Fig. 1.8 B) (Grimstone and Gibbons, 1966; Guichard et al., 2012, 2013). Otherwise, the *Trichonympha* spp. centriole has a canonical organization in cross-section, with nine microtubule triplets going along the long proximal part harboring the cartwheel (Grimstone and Gibbons, 1966; Guichard et al., 2012, 2013).

Even though this termite symbiont is very difficult or perhaps impossible to cultivate in laboratory conditions (Trager, 1934), it became an excellent model organism to study mostly the structure of the proximal region of the centriole (Guichard et al., 2012, 2013).

Even in the absence of direct genetic tools such as mutagenesis or RNAi silencing to study the function of specific genes *in vivo*, the availability of the sequenced and assembled genome of *T. agilis* allowed members of our laboratory to find a *T. agilis* homolog of the important cartwheel protein SAS-6 (Guichard et al., 2013). Armed with the *T. agilis* sequenced genome and a single-cell transcriptomic database, we initiated the project of elucidating the composition and structure of the elongated *T. agilis* centrioles, as well as of SAS-6 assembly, in order to provide novel insights into the biogenesis of centriole. Through our collaboration with Prof. Y. Hongoh, we were able to get japanese termites living in symbiotic relationship with *T. agilis*.
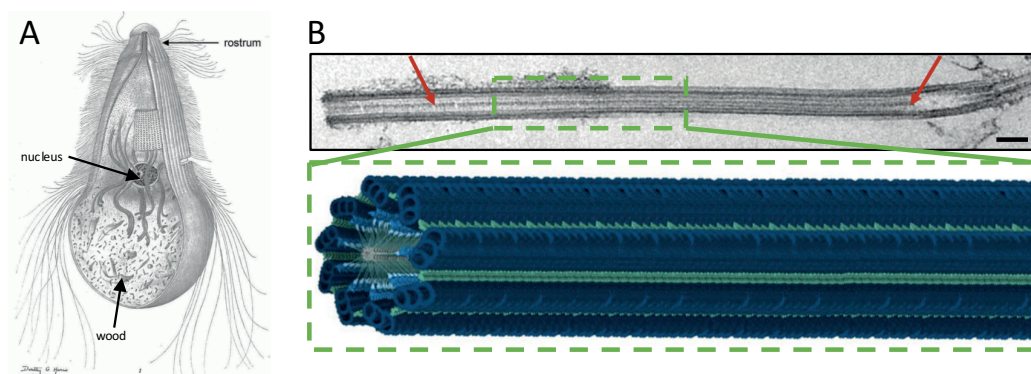
Fig. 1. 8 Elongated proximal region of *Trichonympha* sp. centriole. **A** Schematic drawing of *Trichonympha* sp. Note that the majority of the cell surface is covered by centrioles, with ~ 5.000 to 14.000 centrioles per cell, depending on the species (from Kirby, 1932). **B** Micrograph of the whole centriole (top) showing the presence of the cartwheel along almost the whole length of the centriole (red arrows). The dashed green box depicts the approximate region represented in the cryo-ET density map (bottom). Pictures with permission from Guichard et al. (2015) (top) and Guichard and Gönczy (2016) (bottom).

## 1.6 Aims of the project

As described above, *T. agilis* is an excellent model to study the structure of the proximal region of the centriole. With the availability of the sequenced genome and the transcriptome, we wanted to expand the usefulness of this model system to study not only the structure, but also the composition of the proximal region of the centriole and identify new players potentially involved in the biogenesis and the stability of this amazing structure.

Therefore, the main goals of my PhD studies were to:

1. Obtain a cryo-electron tomography map of the *T. agilis* centriolar proximal region with improved resolution and identify potential species-related differences.
2. Identify proteins of the *T. agilis* centriole using proteomics.
3. Analyze oligomerization properties and behavior of SAS-6 proteins, which initiate procentriole formation, using TaSAS-6_1 and the newly discovered TaSAS-6_2, as well as of HsSAS-6, using biochemistry, electron microscopy and atomic force microscopy.

Many people contributed to the elaboration of these projects; a detailed description of my own contribution, as well as that of other individuals, is provided at the end of each chapter.

# Chapter 2    3D architecture of the proximal region of the *T. agilis* centriole

> *"If you want to understand function, study structure."*
>
> Francis Crick

## 2.1    Background

### 2.1.1    Electron microscopy studies on centrioles

The centriole is one of the smallest organelles in a cell, but also one of the largest protein-based complexes (reviewed in Winey and O'Toole, 2014). The nine-fold symmetrical structure of basal bodies and centrioles was apparent already in the micrographs from the first era of electron microscopy (Fawcett and Porter, 1954; de Harven and Bernhard, 1956). Since then, substantial efforts have been undertaken to uncover the detailed structure of the centriole in its native environment, using conventional electron microscopy (reviewed in Bornens, 2012; Tassin and Bornens, 1999), super-resolution light microscopy (Gartenmann et al., 2017; Hamel et al., 2014; Lau et al., 2012; Wegel et al., 2016) or cryo-electron tomography (Guichard et al., 2010, 2012, 2013; Li et al., 2012). Cryo-electron tomography (cryo-ET) followed by subtomogram averaging has become a powerful visualization technique in the centriole field, because it is particularly suitable for large and internally complex structures, which are difficult or impossible to study by other means so far. The boom of cryo-ET followed by subtomogram averaging brought us into an era of unprecedented insight into the amazing structure of centrioles, more specifically of the *Chlamydomonas* centriole (Li et al., 2012) and of the proximal region of the *Trichonympha* spp. elongated centriole (Guichard et al., 2012, 2013). The resolution attained in these studies is in the range of 30-40 Å, which is significantly better than what was available several years ago (reviewed in Diebolder et al., 2012). The well known α- and β-tubulins subunits, the main constituents of microtubule triplets, can be localized with accuracy into the cryo-ET density map of centriolar microtubule triplets of *Chlamydomonas* (Li et al., 2012). Regarding the other centriolar components, there is only one protein that can be placed into the density map, namely the SAS-6 protein, which resides in the proximal region of the procentriole or basal body, where it forms the cartwheel structure (Guichard et al., 2012, 2013). Unfortunately, the 30-40 Å resolution is not sufficient to place the majority of known important centriolar proteins inside the density maps, even if their crystal structures exist, at the least in part for some of them (reviewed in Banterle and Gönczy, 2017).

### 2.1.2  Cryo-electron tomography and subtomogram averaging

Cryo-electron microscopy is a valuable technique that provides the possibility to study biological molecules and protein complexes in their native form, avoiding dehydration, artefacts coming from slow cubic ice formation or staining needed for conventional electron microscopy (reviewed in Fernandez-Leiro and Scheres, 2016; Murata and Wolf, 2018). However, cryo-EM images are in general very noisy, with a low signal-to-noise ratio. Generally speaking, there are two major techniques enabling one to computationally combine and average hundreds or thousands of cryo-EM images in order to increase the signal-to-noise ratio, to emphasize features of an object and to reconstruct a 3D map: single particle analysis (SPA) and single particle tomography (SPT), also called subtomogram averaging (reviewed in Murata and Wolf, 2018). Below I discuss why SPT is more suitable for the structural determination of the *T. agilis* centriole.

A 3D reconstruction from SPA is calculated from 2D transmission electron micrographs of a purified protein or protein complex in different orientations. The development of direct electron detectors and of analysis software have enabled to reach near-atomic resolution not only for large macromolecular or symmetrical complexes, but also for proteins down to a few kilodaltons in molecular weight (reviewed in Elmlund and Elmlund, 2015). However, in general, the sample requires extraction of the complex from its native environment and purification to concentrate many particles on one grid. Therefore, many biological molecules are excluded from reconstructions using SPA, because their relevant structure exists only in their native environment, and surrounding "contaminating" molecules above or below the structure of interest would hinder the processing of particles by SPA.

Even though the *T. agilis* centriole is symmetrical and keeps its internal structure also after extraction from a cell, it is an elongated and large protein complex of ~ 1.5 – 5 μm present only in one, horizontal, orientation on a grid. Therefore, the second technique, SPT, which takes advantage of acquiring a 3D volume, is more suitable for *T. agilis* centriole reconstruction. In SPT, a sample is rotated around an axis through a set of defined tilts and an image is acquired at each angle. Therefore, one can easily distinguish a structure of interest from other contaminants around the sample, the object of interest can be extracted from the tomogram and whole 3D volumes can be averaged in an iterated manner (Briggs, 2013; Wan and Briggs, 2016).

Cryo-ET followed by subtomogram averaging has indeed provided useful 3D information about the centriole in its native form, with ~ 40 Å resolution (Guichard et al., 2012, 2013; Li et al., 2012). With the use of direct detection devices of the last generation, optimized subtomogram averaging methods (Schur et al., 2013, 2015) and more efficient 3D-contrast transfer function (Turoňová et al., 2017), cryo-ET has the capacity to reach sub-nanometer resolution and even go to the ~ 4 Å resolution, which is at the level of X-ray diffraction.

With this in mind and given the recent acquisition of a new Falcon 3 (FEI) direct electron detector at our electron microscopy facility, we set out to improve the resolution of the 3D architecture of the *T. agilis* centriolar proximal region. Moreover, we aimed to resolve the structure of the *T. agilis* centriole in order to use it as a reference for our next studies. We used the same technique, cryo-ET

with subsequent subtomogram averaging, as was used before for the *Trichonympha* spp. proximal region of centriole (Guichard et al., 2012). Below, I briefly describe the major steps in the workflow of cryo-ET and subtomogram averaging of *T. agilis* centrioles (Fig. 2.1), with the technical details being specified in the Materials and Methods section.

| Sample preparation | Centriole purification |
|---|---|
| Tilt-series acquisition | Tecnai F20 Cryo, 200 kV, Falcon3 direct detection device |
| Tilt-series alignement | Fiducial or patch fiducial-less alignment IMOD software (Kremer et al., 1996) |
| Tomogram reconstruction | IMOD software (Kremer et al., 1996) Parallel processing, GPU-accelerated computing |
| Particle picking / subboxing | Filament model Dynamo software (Castaño-Díez et al., 2012) |
| Pre-alignment | Filament model Dynamo software (Castaño-Díez et al., 2012) |
| Subtomogram alignment and averaging | Dynamo software (Castaño-Díez et al., 2012) GPU-accelerated computing |
| Resolution measurement | FSC measurement (Castaño-Díez et al., 2012) |
| Final structure visualization | UCSF Chimera (Pettersen et al., 2004) |

Fig. 2. 1 Schematic representation of cryo-electron tomography and subsequent subtomogram averaging analysis workflow with specification of protocols, software and hardware used in this study.

## 2.2 Results

### 2.2.1 Sample preparation: *T. agilis* centriole purification

I adapted the previously optimized protocol for *Trichonympha* spp. centriole purification, which yields centrioles attached to flagella in purity sufficient for observation by cryo-EM (see Materials and Methods, Guichard et al., 2015). The flagellum remaining attached to the centriole helped us to unambiguously identify the proximal and distal ends of the centriole. The purified centriole is fragile and can be easily deformed by non-homogeneous ice thickness or the carbon support. Each individual centriole had to fulfil several conditions to be retained for acquisition: (i) it should not be positioned on the carbon lacey support to prevent severe flattening, (ii) it should not be positioned close to the cupper grid, which might shadow the sample in high angle tilts and (iii) if possible, it should be oriented parallel to the tilt axis in order to acquire as much information as possible from the elongated particles.



|  | | Centriole W/H | Hub W/H | Central density | Particles |
|---|---|---|---|---|---|
| *T. agilis* | Tom1 | 1.64 | 1.00 | CID | 122 |
| | Tom2_1 | 1.91 | 1.00 | CID | 39 |
| | Tom2_2 | 1.00 | 1.00 | CID | 83 |
| | Tom2_3 | 1.06 | 1.00 | CID | 91 |
| | Tom3_1 | 1.05 | 0.97 | CID | 86 |
| | Tom3_2 | 1.05 | 1.00 | CID | 110 |
| | Tom4 | 1.14 | 1.00 | CID | 134 |
| | Tom5 | 1.89 | 1.00 | CID | 164 |
| *T. mirabilis* | Tom6 | 1.57 | 1.00 | CCD | 93 |
| | Tom7 | 1.03 | 1.00 | CCD | 120 |
| | Tom8 | 1.69 | 1.00 | CCD | 95 |
| | Tom9 | 1.55 | 1.00 | CCD | 80 |
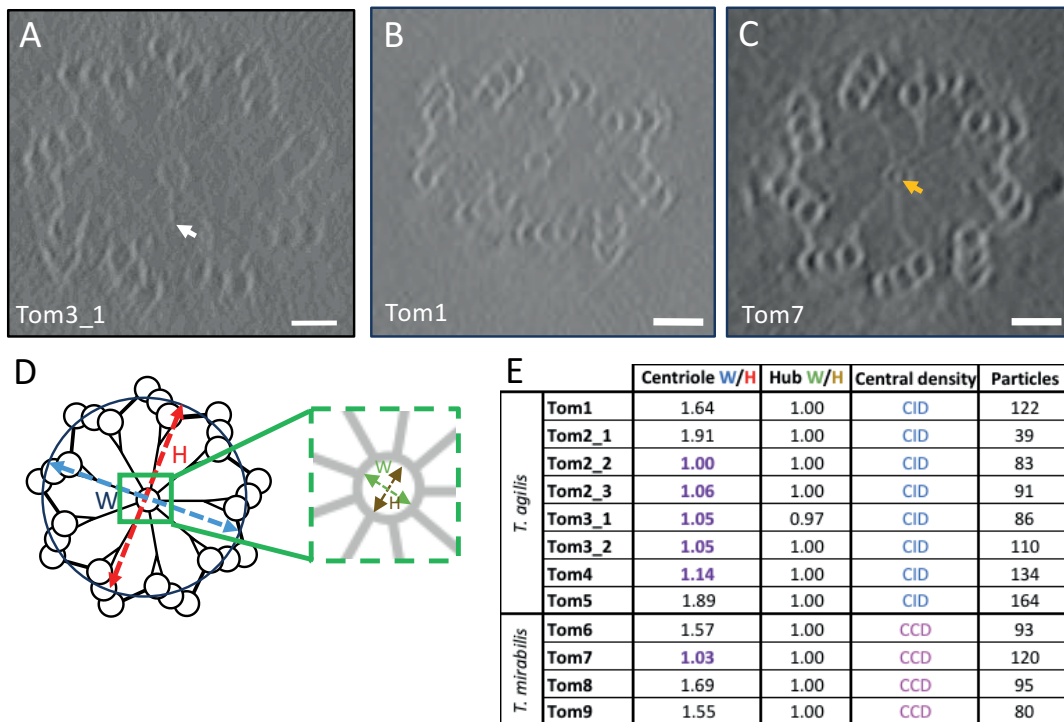
Fig. 2. 2 Centriole distortion and flattening. **A-B** Transverse sections of two tomograms of *T. agilis* centrioles. **A** Almost completely circular centriole, with curved spokes (arrow). Scale bar: 50 nm. **B** Flattened centriole, still taken into consideration for subsequent analysis of the cartwheel hub. Scale bar: 50 nm. **C** Transverse section of *Teranympha mirabilis* centriole. Note electron dense cartwheel central density (CCD, yellow arrow) **D** Circularity measurement of centrioles and their cartwheel hubs taken for this study, Width/Height (W/H). **E** The ratio W/H of centriole around 1 signifies unflattened centriole (in purple), higher ratio shows a flattened centriole. Note that even when a centriole was flattened, the cartwheel hub usually displayed a circular shape. Number of subtomograms taken for averaging is depicted on the right. See text for details.

We opted for acquiring tomograms of centrioles not only in their native, cylindrical shape, but also of centrioles that were a bit flattened, whilst keeping their general organization and position of microtubule triplets (Fig. 2.2 A-C). In this manner, we obtained more tomograms to analyze and process for subtomogram averaging of the central cartwheel region. We found that the cartwheel hub, which appears to be protected by microtubule triplets, stays largely intact after compression, as determined by the measurement of the hub circularity (Fig. 2.2 D and E). The compression of the hub is likely buffered by the radially arranged spokes, which are bent or squeezed during the compression process, whereas the hub remains in a ring-like shape (Fig. 2.2 E). In total, we acquired five suitable tomograms of *T. agilis* centrioles with 8 centrioles, 5 of which were almost unflattened and 3 of which were flattened (Fig. 2.2 E).

The curvature of cartwheel spokes, and the flattening of some peripheral elements in particular, were the reasons why we did not process the cartwheel hub with the entire spokes as previously (Guichard et al., 2013), but instead averaged only the cartwheel hub with ~17 nm of neighboring spoke length.
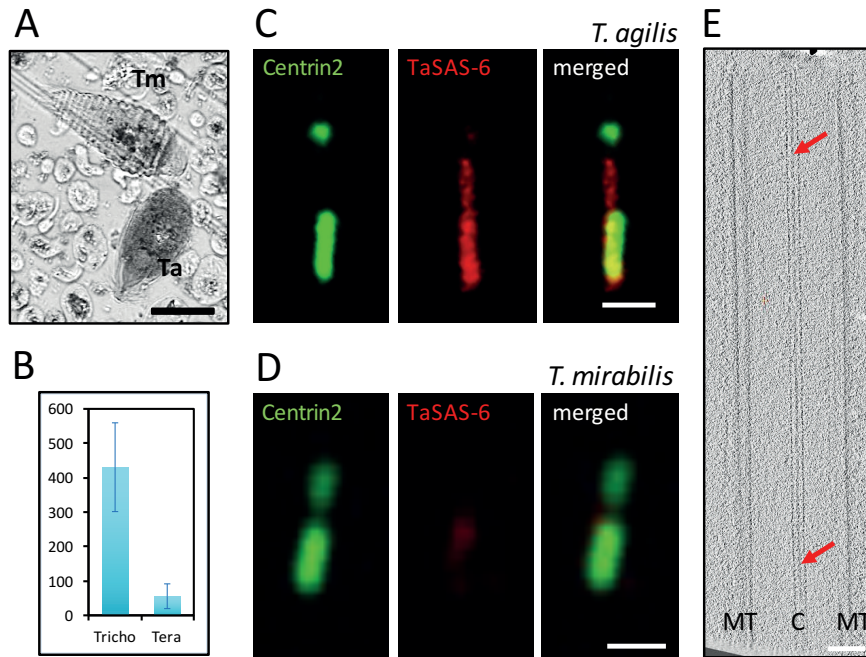


Fig. 2. 3 Cartwheel of *T. mirabilis* **A** Comparison of the two large symbionts present in *R. speratus* gut: *T. mirabilis* (Tm) and *T. agilis* (Ta). Scale bar: 50 μm. **B** Average ratio of the number of the two symbionts present in individual termite guts. Error bars represent standard deviations of the number of large symbionts in one termite, n=70. **C-D** Indirect immunofluorescence staining of purified centrioles from *T. agilis* (C) and *T. mirabilis* (D) stained with Centrin2 and TaSAS-6 antibody. Note the difference in Centrin2 staining between the two species. Proximal side of centrioles is at the bottom. Scale bars: 1 μm. **E** Transverse section through the *T. mirabilis* centriole, with microtubule triples (MT) at the periphery and the cartwheel hub (C) with the presence of CCD at the center. Note the presence of the CCD along the entire analysed region of the centriole (red arrows). Scale bar: 50 nm.

### 2.2.1.1 Centriole of *Teranympha mirabilis*

*T. agilis*, which is the only *Trichonympha* sp. present is the *R. speratus* termite, is not the only large gut symbiont of the latter (Fig. 2.3 A). The second large symbiont is the parabasalid *T. mirabilis* present to a ~ 10x lesser extent on average (Fig. 2.3 B). Even if I tried to optimize the *T. agilis* cell isolation and centriole purification protocol, the centrioles of *T. mirabilis* were sometimes present in the final centriole sample preparation. Interestingly, these centrioles were similarly elongated as are *T. agilis* centrioles (Fig. 2.3 C and D) and harboured a cartwheel along the whole analysed region (Fig. 2.3 E). We found the overall cartwheel and centriole architecture of *T. mirabilis* to be similar to that of *T. agilis* centrioles (Fig. 2.2 C). However, we detected a significant difference between centrioles from the two species: the shape and appearance of the cartwheel inner densities. On the one side, we have the *T. agilis* CID, which is not discernible on raw tomograms, and on the other side, an electron dense elongated structure inside the *T. mirabilis* cartwheel, therefore called Cartwheel Central Densities, or CCD (Fig. 2.3 E). This remarkable difference led us to also analyse the *T. mirabilis* centriole and to investigate the biodiversity of the cartwheel and centriolar structure. We acquired tomograms of four centrioles and processed them similarly as *T. agilis* centrioles (Fig. 2.2 D).

### 2.2.2 Tilt series acquisition

Tomograms are incomplete representations of a 3D structure. Indeed, artefacts and deformation might be introduced in the direction of the "missing wedge" angle (Fig. 2.4), which is a result of purely practical limitations of the tilt series acquisition, reflecting the sample holder maximal tilt or the sample thickness (Schmid and Booth, 2008). Other potential limitations are specimen movement during the tilting process, which might introduce a change in focus and reduce the signal-to-noise ratio due to fractionation of the electron dose that a biological sample can sustain without introducing artefacts. In our case, the tilt series were acquired using a unidirectional tilt scheme (Fig. 2.4), which is the easiest of tilt schemes and that was utilized before with *Trichonympha* spp. (Guichard et al., 2012, 2013). Furthermore, this tilt scheme is the most suitable for the F20 microscope with the side entry holder, where the specimen stage is not stable enough for large motion introduced in the other tilt schemes. Moreover, our collaborator from the electron microscopy facility is most experienced with this acquisition scheme. Even though this tilt scheme might potentially introduce some artefacts towards the end of the acquisition because the electron dose is cumulative in biological samples (Glaeser, 2016), this should not pose a problem in this case because of the symmetrical features present in centrioles and its parallel positioning to the tilt axis.
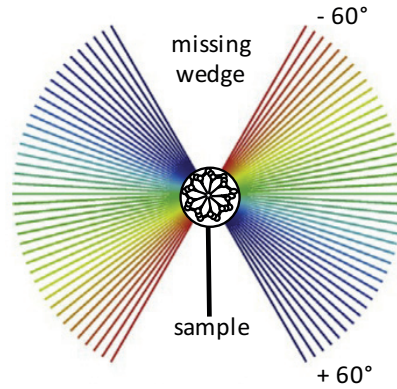
Fig. 2. 4 Unidirectional tilt scheme. Tilt series acquisition from - 60° to + 60° in 2° angle step increments, with depiction of missing wedge and centriole positioning parallel to tilt axis. With permission and adapted from Wan and Briggs (2016).

### 2.2.3 Tilt series alignment and tomogram reconstruction

We used the IMOD package (Kremer et al., 1996; Mastronarde, 1997) to preprocess the data before tomogram reconstruction. Tilt series were aligned using either gold beads or patch "fiducial-less" alignment (Kremer et al., 1996), because some tomograms did not have a sufficiently high concentration of fiducials to align the tilt series with confidence.

### 2.2.4 Particle picking

Manual particle picking is tedious when one needs to pick hundreds or thousands of particles and to define their orientations (Fig. 2.6). The advantage of using an elongated periodical structure, as a *T. agilis* centriole, is the possibility of semi-automated particle picking and cropping using a filament model (Fig. 2.5). This model crops particles alongside a specified axis at regular intervals and is implemented in the Dynamo software (Castaño-Díez et al., 2012). The use of a filament model has also the advantage that it automatically pre-aligns picked particles before iterative alignment and averaging.
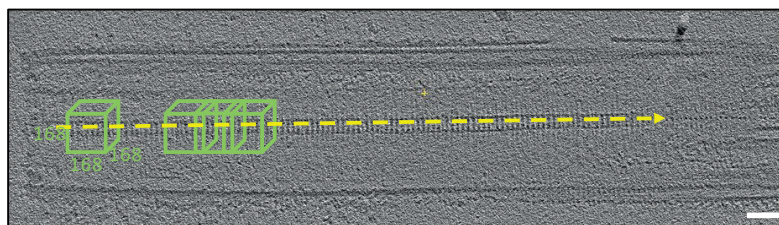


Fig. 2. 5 Particle picking from *T. agilis* centriole cartwheel hub. The filament model is specified in yellow and individual boxes represent individual picked particles that are extracted at regular intervals along the filament model axis. Box size is 168 x 168 x 168 pixels (~ 59 x 59 x 59 nm). Scale bar 50 nm.

### 2.2.5 Subtomogram alignment and averaging

Subtomogram averaging is an iterative process, whereby a set of particles or subtomograms is aligned to a reference. The aligned subtomograms are then averaged with a reference to produce a new reference for the next iteration. This process is iterated until the final reference cannot be improved (Fig. 2.6). Even for the first iteration, a reference is needed. One possibility is to impose an existing reference or template from a structure that was previously determined with lower resolution. However, with its iterative nature, subtomogram averaging might irreversibly bias the averaging and structure determination when imposing an incorrect existing reference. Therefore, one must be very careful in choosing a starting template. In order to be unbiased or to confirm the correctness of the final average obtained with an existing template, a starting reference from a quickly averaged small subset of subtomograms from the dataset or a simple geometrical shape can be used. The final results from the subtomogram averaging analysis, made with or without the existing starting template, should show analogous densities. The alignment of each tomogram is made by angular search whereby a reference is masked and then rotated by a defined set of angles and compared to each individual subtomogram. The final resolution is then assessed by a Fourier shell correlation (FSC), and density maps visualized in UCSF Chimera and coloured based on their local resolution.
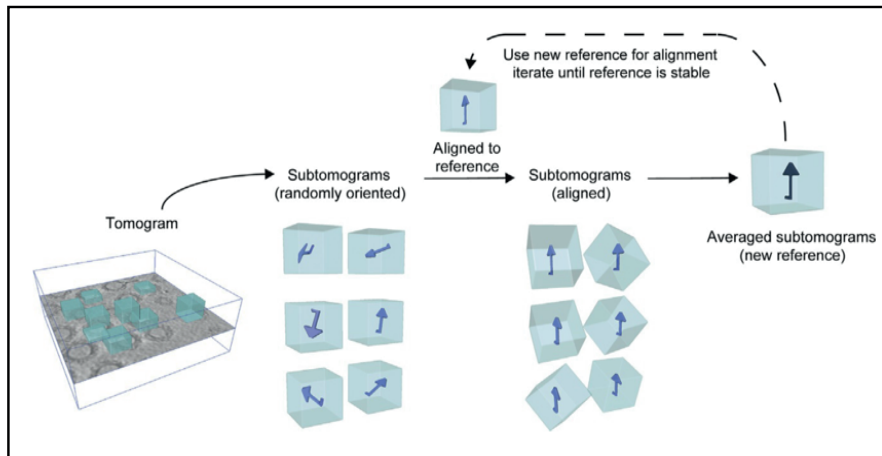


Fig. 2. 6 Schematic representation of subtomogram averaging workflow. Particles are firstly extracted (blue boxes), from individual tomograms. Particles (also called subtomograms) that are randomly oriented or coarsely pre-aligned, depending on the choice of the particle picking method, are aligned to a reference and then averaged together. This average is used as a new reference for the next iteration. This process is repeated several times until the final reference is stable. With permission from Briggs (2013).

### 2.2.6 *T. agilis* cartwheel

Particle extraction for the subtomogram averaging of the *T. agilis* cartwheel hub took place every 8.5 nm, which corresponds to the periodicity of the *Trichonympha* spp. cartwheel hub (Guichard et al.,

2013). Altogether 872 particles were extracted (Fig. 2.2 E). The subtomogram averaging was done with or without the use of pre-existing template of the previously published cartwheel hub map (Guichard et al., 2013), which was adapted in order to fit the dimensions of subvolumes generated here and used in Dynamo. Altogether, four averages were computed, with or without the starting reference and with or without imposing nine-fold symmetrization on the average.

- "Ref-C9" – Imposing an existing template (Guichard et al., 2013) as starting reference and using nine-fold symmetrization
- "Ref-C1" – Imposing an existing template (Guichard et al., 2013) as starting reference, without symmetrization (default in the Dynamo software)
- "Avg-C9" – Using as starting reference an average made from 50 random particles picked from the present dataset and using nine-fold symmetrization
- "Avg-C1" Using as starting reference an average made from 50 random particles picked from the present dataset, without symmetrization

We thus obtained four different averages for the *T. agilis* cartwheel hub and could assess whether imposition of the pre-existing template or symmetrization might have generated artefacts (Fig. 2.7).
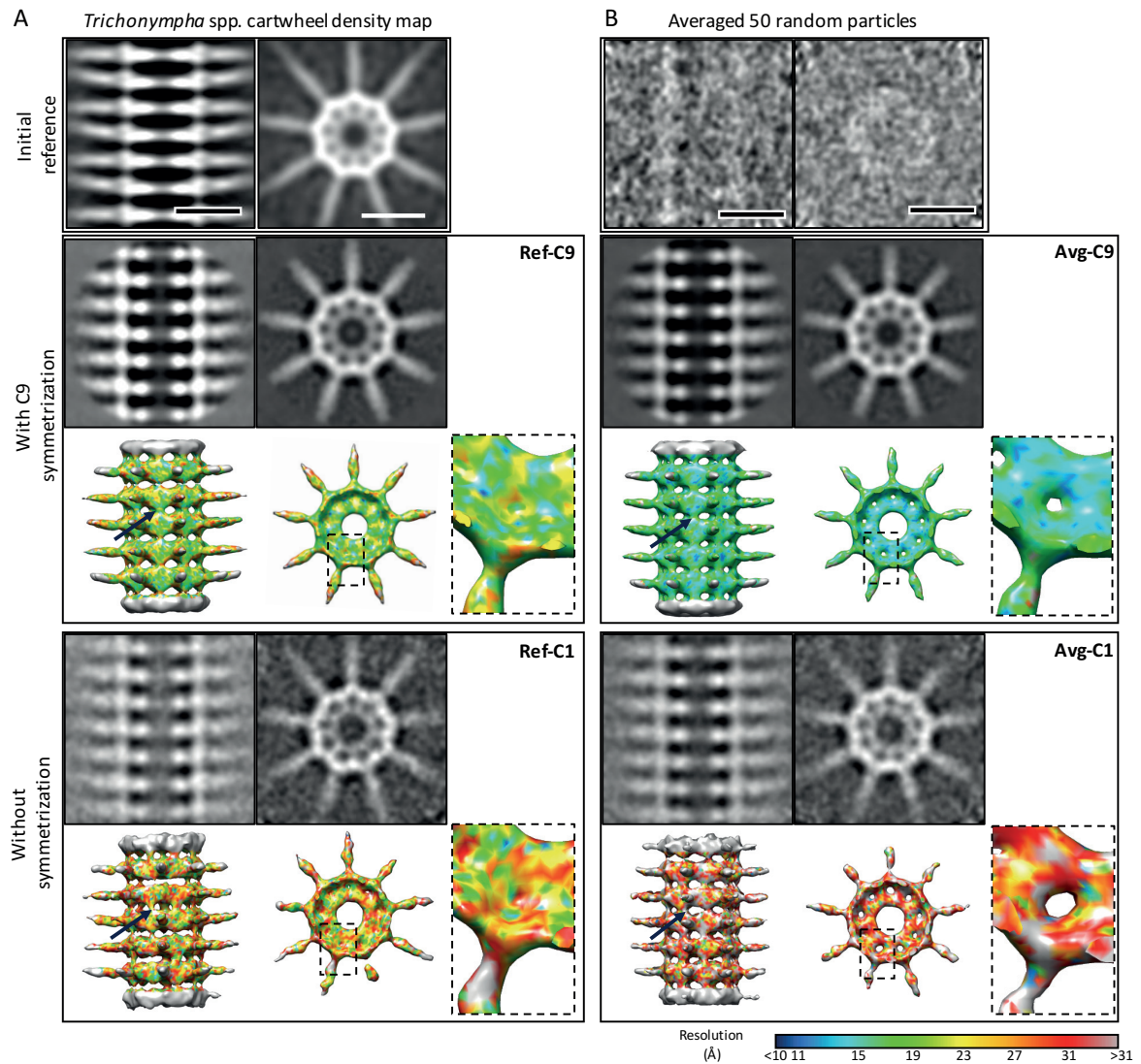
Fig. 2. 7 Symmetrization helps to create a smooth average without introducing artefacts. **A-B** Comparison of four averages calculated with (A) or without (B) imposition of pre-existing reference (Ref) adapted from Guichard et al. (2013) and with (middle panels) or without (bottom panels) imposing nine-fold (C9) symmetrisation. The subtomogram averaging without pre-existing reference was calculated with a low-resolution template averaged from 50 random particles from the *T. agilis* cartwheel hub dataset (Avg). Note the presence of Inter-Cartwheel Pillars (ICPs, arrows). The resulting averages are depicted as tomographic slices and as full density maps, with higher-magnification views of depicted region (dashed line). The electron density maps are shown at the countour level 1.5 σ and depicted using rainbow coloring based on their local resolution, with the scale shown on the bottom in Ångstroms. 2D projections and electron density maps were filtered by a Gaussian filter with value 2 to increase the signal-to-noise ratio and to smoothen the density surface. Noisy particles or small blobs on the density maps disconnected from the main density were removed by the function Hide Dust implemented in UCSF Chimera. The local resolution was calculated using ResMap (Kucukelbir et al., 2014).

### 2.2.6.1 Cartwheel hub

The Ref-C9 density map of *T. agilis* cartwheel hub confirmed previously uncovered features of the cartwheel hub (Guichard et al., 2012, 2013), such as the periodical arrangement of cartwheel rings, with 8.5 nm periodicity (Fig. 2.9 C) and nine radial spokes emanating from it (Fig. 2.7). We confirmed the presence of the CID, which are densities present inside the cartwheel hub and connected to its inner wall at the position where oligomerization between neighboring SAS-6 homodimers occurs (Fig. 2.7) (Guichard et al., 2013). In addition, the new *T. agilis* cartwheel hub map shows novel features, which were not uncovered before, probably because the previous *Trichonympha* spp. map was analyzed at lower resolution (Guichard et al., 2013) or because of their specificity to the *T. agilis* centriole (Fig. 2.9).

### 2.2.6.2 Cartwheel inner density (CID)

All four final electron density maps of the *T. agilis* cartwheel hub confirms the presence of the CID (Fig. 2.7, 2.8), which was uncovered previously as a continuous disc with a hole in the middle at lower threshold density, or as nine triangular inter-connected densities interacting with the cartwheel hub around the SAS-6 N-terminal dimerization interfaces at higher threshold density (Guichard et al., 2013). The new structure allows us to visualize a little more clearly the shape of the CID, even without largely increasing the threshold density (Fig. 2.8 A).
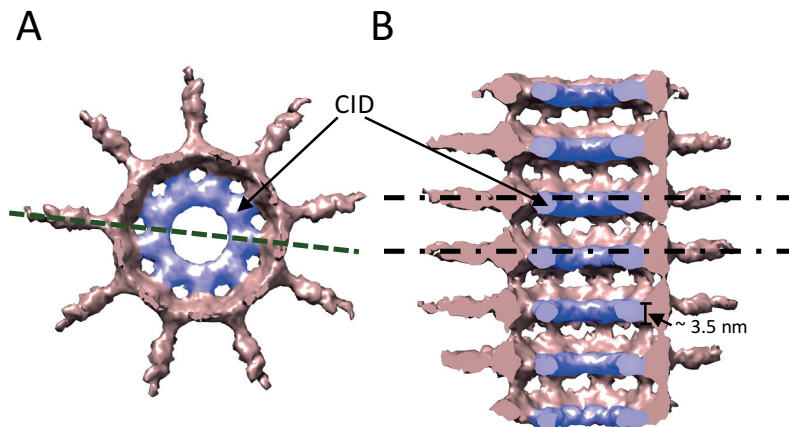


Fig. 2. 8 Asymmetrically localized CID inside cartwheel hub. **A** CID (in purple) viewed from the distal end of the cartwheel hub. Only 1 cartwheel hub layer is shown. Dashed line indicates where the transverse section in (B) was done through the whole stack. **B** Transverse section through the *T. agilis* cartwheel hub, showing that the CID localizes asymmetrically inside the cartwheel hub ring (in purple). Dashed lines are axes going through the centre of the spokes and clearly show that the CID is localized proximally in each ring. Electron density map is visualized at the contour level 1.5 σ, without filtering.

Moreover, given that we extracted all subtomograms in the same orientation along the proximal-distal centriole axis, we were able to determine that the CID is a ~ 3.5 nm thick density localized asymmetrically, on the proximal side of the cartwheel ring (Fig. 2.8 B). The discovery of the polar-

ized localization of the CID is likely to have implications in understanding how the cartwheel forms as a polarized structure along the proximal-distal centriole axis.
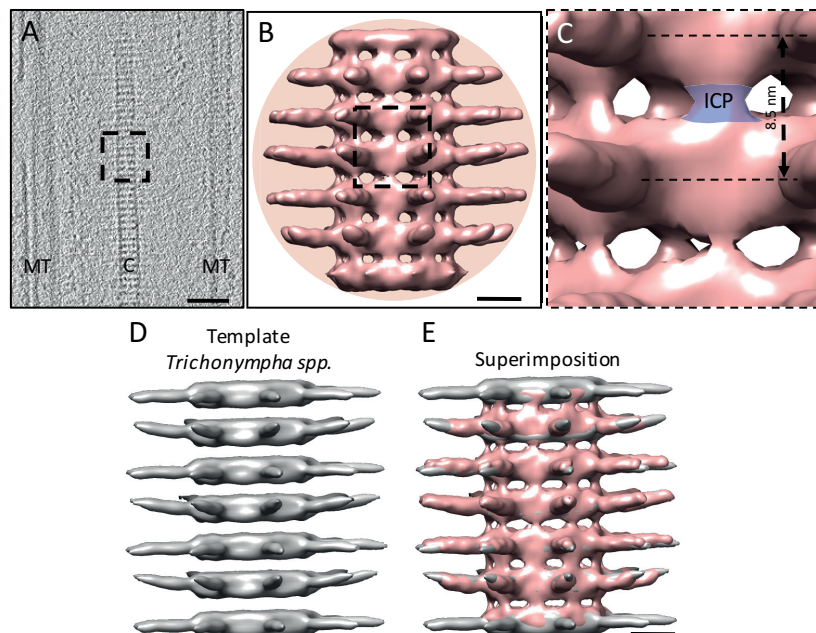


Fig. 2. 9 The *Trichonympha agilis* cartwheel hub reveals new features. **A** Transverse section through a portion of a *T. agilis* centriole tomogram depicting the periodical cartwheel hub in the middle (C) and microtubule triplets (MT) at the periphery. The dashed square represents the size of the box used for subtomogram averaging of the cartwheel hub (168 x 168 x 168 pixels or ~ 59 x 59 x 59 nm). Scale bar: 50 nm. **B** Final Ref-C9 density map of *T. agilis* cartwheel hub. Red disk represents the shape of the mask applied for the subtomogram averaging procedure. Dashed square shows the region magnified in C. **C** Inter-Cartwheel Pillar (ICP) represented in detail. The spacing between rings is 8.5 nm. Note that the thickness of individual ICPs might vary, probably due to their artefactual erasing during the iterative process of subtomogram averaging when the starting reference without the ICP was imposed. **D-E** Comparison of *Trichonympha* spp. density map (Guichard et al., 2013) (D) used as the pre-existing template and superimposed *T. agilis* final Ref-C9 density map (E). The *T. agilis* electron density map was denoised by the function Hide Dust implemented in UCSF Chimera and filtered by a Gaussian filter with value 2 to smoothen the density surface and to be able to compare it to *Trichonympha* spp. map. Scale bars in B and E: 10

### 2.2.6.3   Inter-Cartwheel Pillar (ICP)

Interestingly, the cryo-ET map revealed that the *T. agilis* cartwheel rings seem to be interconnected by a regularly spaced electron density, therefore called Inter-Cartwheel Pillar (ICP) (Fig. 2.9 B and C). These densities become apparent mostly in the averages made without the pre-existing starting reference of the *Trichonympha* spp. cartwheel (Fig. 2.7 B), suggesting that ICPs are either absent in that species or were not resolved properly (Fig. 2.9 D and E) (Guichard et al., 2013). The ICPs are localized at the N-terminal interaction interface of SAS-6 proteins, between spokes, and form continuous cartwheel pillars with the SAS-6 N-terminal domains, possibly to ensure robustness of the long *T. agilis* cartwheel (Fig. 2.9 B and C). The thinning of the ICP from the average made with the

pre-existing starting reference (Fig. 2.7 A) is an indication that this reference might potentially bias the structure determination. However, the ICPs are still apparent at each of the four averages, which confirms their bona fide presence (Fig. 2.7).

### 2.2.6.4 Beyond the cartwheel hub

To confirm the structure of the *Trichonympha* spp. cartwheel spokes and their merging towards the periphery of the centriole, we opted for subtomogram averaging of particles with larger dimensions, ~105 nm. If the periodicity of spokes towards the centriole periphery is truly 17 nm, the averaging of particles picked every 17 nm along the central hub would reinforce the spoke curvature and merging. If the spoke periodicity is different, this experiment would lead to an average with blurry spokes and impossibility to extract information from the periphery of the averaged density map. The visualization of the final average however confirmed the distal spoke periodicity of 17 nm (Fig. 2.10).
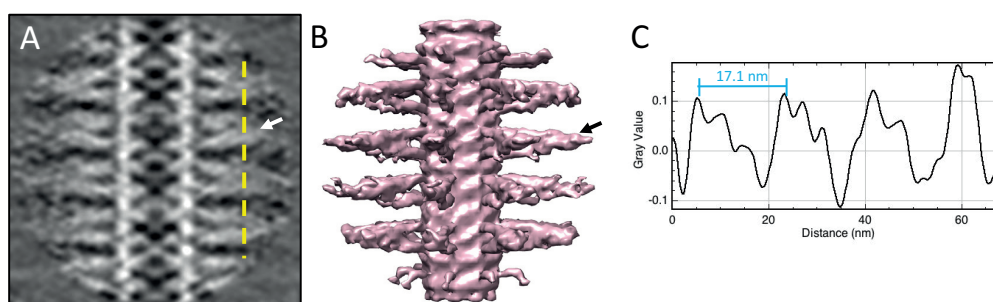


Fig. 2. 10 *T. agilis* cartwheel spokes merge towards the centriole periphery. **A** Transverse section of the final average with the dimension 105 x 105 x 105 nm. Note spokes merging towards the periphery (white arrow). Dashed line represents the section where the plot profile in (C) was made. Scale bar: 10 nm. **B** Electron density map of the averaged cartwheel hub with spokes emanating outside and merging at the periphery (black arrow). Map was filtered with Gaussian filter with value 3 and visualized at the 1.5 σ contour level. Blobs and unspecific noise were removed by the function Hide dust implemented in UCSF Chimera. **C** Plot profile depicting the periodicity of the *T. agilis* cartwheel spokes.

### 2.2.7 *T. mirabilis* cartwheel

### 2.2.7.1 Cartwheel hub with CCD

As stated above, subtomograms picked from the *T. mirabilis* cartwheel hubs with a strong central density were averaged separately, using the pre-existing 3D cartwheel hub density map (Guichard et al., 2013), which was adapted to dimensions of picked particles, as the starting reference, as well as without the external reference. Even when using *Trichonympha* sp. density map as the external reference, which does not contain the CCD, the strong central density of the CCD present in the *T. mirabilis* centriole tomograms was well resolved (Fig. 2.11). The presence of the CCD was confirmed also on electron density maps without imposing a starting reference or symmetrization (Fig. 2.11 B), which indicates that the pre-existing template and/or symmetrization did not introduce a bias into the structure. The CDD appears as a continuous elongated rod with a diameter of ~8 nm (Fig. 2.12 B-D), with no apparent connections to the cartwheel inner wall (Fig. 2.12 A, D and E).
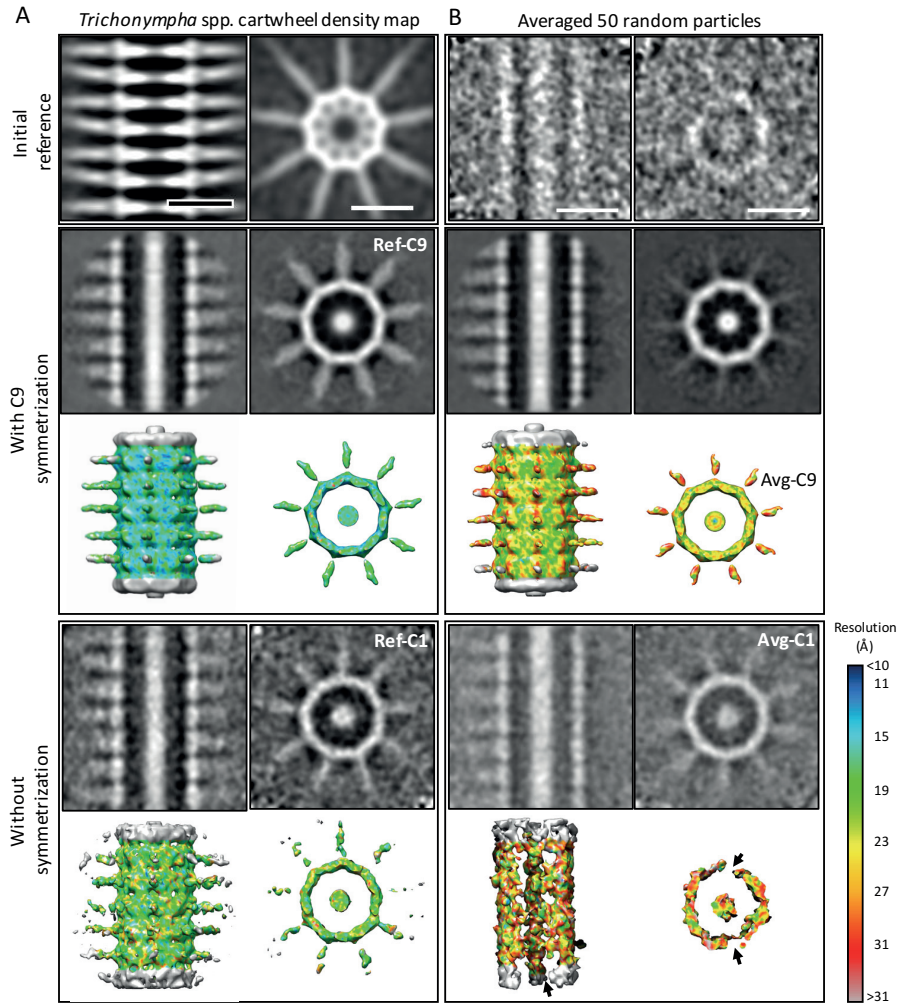
Fig. 2. 11 *T. mirabilis* cartwheel hub with CCD. **A** Two final density maps of the *T. mirabilis* cartwheel hub aligned and averaged using published and adapted *Trichonympha* spp. electron density map as the starting reference (top panel) depicted as a transverse section (left) and top view (right). Scale bars: 20 nm. (Guichard et al., 2013). Transverse section or side view of the cartwheel hub (right) and top view of one cartwheel layer of the average using nine-fold symmetrization (Ref-C9, middle panel) and without symmetrisation (Ref-C1, bottom panel). **B** Two final density maps of the *T. mirabilis* cartwheel hub averaged without an external reference. The reference used for the alignment was made by averaging of 50 random particles from the dataset (top panel, transverse section on the left, topview on the right). Scale bars: 20 nm. Transverse section and side view of the stack (left) and top view of one cartwheel ring layer (right) of the final density maps without starting reference using nine-fold symmetrization (Avg-C9, middle panel) and without symmetrisation (Avg-C1, bottom panel). Note the loss of information from the electron density map Avg-C1, where the absence of an external reference and/or symmetrisation could not compensate for the loss of information from the missing wedge area (black arrows). The resulting averages are depicted as tomographic slices and as full density maps. The electron density maps are shown at the countour level 1.5 σ and depicted using rainbow coloring based on their local resolution, with scale shown on the right in Ångstroms. 2D projections and electron density maps were filtered by a Gaussian filter with value 2 to increase the signal-to-noise ration and to smoothen the density surface, respectively. Noisy particles or small blobs on the density maps disconnected from the main density were removed by the function Hide Dust implemented in UCSF Chimera. The local resolution was calculated using ResMap (Kucukelbir et al., 2014).
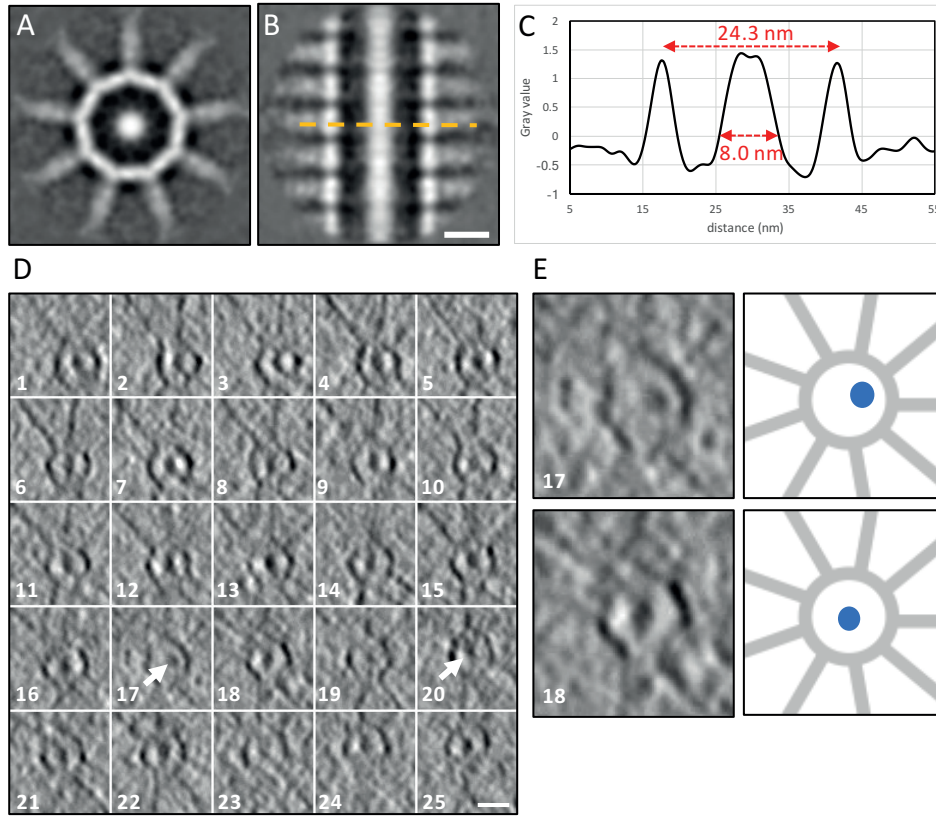
Fig. 2. 12 The CCD is present along the cartwheel hub. **A** Top view of the averaged cartwheel stack with CCD as a central dense dot. **B** Transverse section through the final average of particles with CCD. Note the absence of resolved connection of CCD to the cartwheel hub inner wall. Scale bar: 10 nm. **C** Plot profile representing the diameter of the cartwheel ring with the CCD. **D** 25 low resolution projections of the *T. mirabilis* centriole, corresponding to ~ 30 nm of the cartwheel hub in each case, distributed at equal distances across the ~ 1 μm long proximal region. Note slight movement of the CCD inside the cartwheel hub (white arrows). Scale bar: 20 nm. **E** Magnification of the projections 17 and 18, with corresponding schematic representation. The movement of the CCD is the most apparent in the projection 17 in comparison to the CCD central location in the projection 18.

Furthermore, as observed in all four tomograms with the CCD, the rod is visible along the entire length of this type of cartwheel hub (Fig. 2.12 D). Moreover, we observed a slight movement of the CCD inside the cartwheel hub, along the length of the centriole (Fig. 2.12 D and E), which is compatible with the notion that the CCD harbors connections with the hub that are very thin -and hence not detectable- and flexible.

### 2.2.7.2 *T. mirabilis* cartwheel spokes

Already the preliminary Fourier transform (FFT) analysis of the reconstructed tomograms indicated that the cartwheel hub of *T. mirabilis* shows ~ 8.5 nm periodicity as found previously for *Trichonympha* spp. (Guichard et al., 2012) and confirmed also here for *T. agilis.* Similarities between these species were found also for cartwheel spokes, which display ~ 17 nm periodicity. This perio-

dicity is lost when only the cartwheel hub is analysed by FFT and again present when only spokes were taken into consideration. This indicates that spokes probably merge and cartwheel rings form double layers similarly as observed in *T. agilis* cartwheel (Fig. 1.3 E). The subtomogram averaging of 55 particles, with dimension of ~ 105 nm from one *T. mirabilis* centriole tomogram with the least distorted spokes confirms the spoke merging (Fig. 2.13 D). The FFT analysis was done on three different regions of the *T. mirabilis* cartwheel; the whole cartwheel, where we discovered four periodicities: 4.2 nm, 5.6 nm, 8.3 nm and 16.8 nm (Fig. 2.13 A); the cartwheel hub only with 4.2 nm and 8.4 nm periodicities (Fig. 2.13 B); and the spokes only, where three periodicities of 5.6 nm, 8.5 and 16.5 nm were found (Fig. 2.13 C). The presence of the ~ 8.5 nm periodicity in the cartwheel hub as well as in the spokes likely corresponds to the periodicity of the cartwheel rings with short part of spokes pointing outwards. The spokes then merge and create the periodicity of the ~ 17 nm (measured here as ~ 16.6 nm, Fig. 2.13 A and C). We hypothesize that newly found 4.2 nm periodicity might results from the presence of the ICPs in between cartwheel rings or it is a specific not yet resolved periodicity of the CCD inside the hub. The last periodicity found by the FFT analysis of the cartwheel spokes is the 5.6 nm, which was reported also for *Trichonympha* spp. cartwheel (Guichard et al., 2012). The precise localization of this periodicity remains to be determined, however it might correlate with the presence of the measured ~ 16.6 nm periodicity, as the latter is roughly a triple of the 5.6 nm.
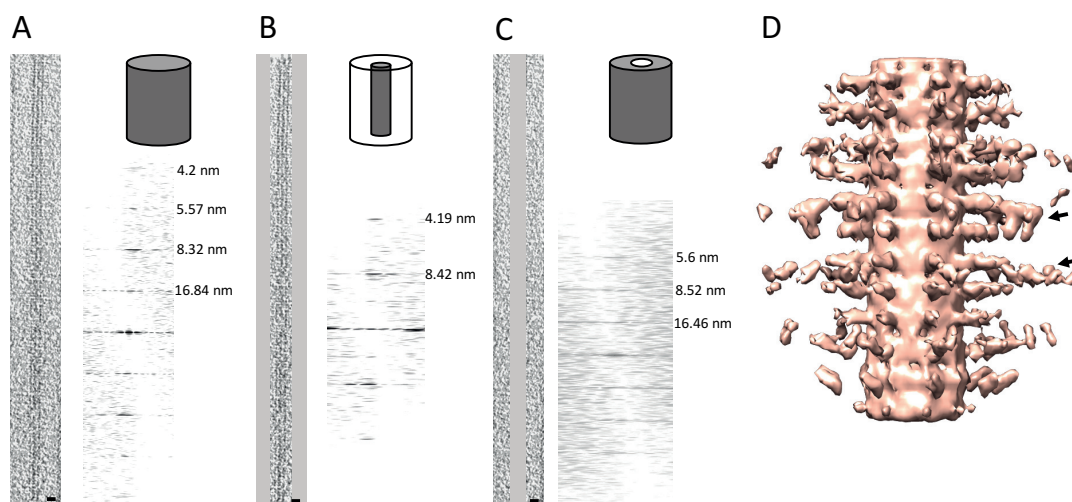


Fig. 2. 13 Cartwheel spokes of *T. mirabilis* centriole display ~ 17 nm periodicity. **A-C** Projections of a cartwheel tomogram (left) with schematic representation of a cartwheel portion used for the FFT analysis (right top) and 2D Fourier transform (right bottom) of depicted area. The FFT analysis of the whole cartwheel (A) revealed four types of periodicities; 16.8 nm, 8.3 nm, 5.7 nm and 4.2 nm, two periodicities of the cartwheel hub only; 8.4 nm and 4.2 nm (B) and three of the cartwheel spokes; 16.5 nm, 8.5 nm and 5.6 nm (C). Scale bars: 15 nm. **D** Electron density map of *T. mirabilis* cartwheel with dimensions of ~ 105x105x105 nm. Note the merging of spokes (black arrows). The density map is visualized at the contour level 1.5 σ, with Gaussian filter 3.0 and denoised by the function Hide Dust implemented in UCSF Chimera.

### 2.2.8 Resolution determination

In order to determine the resolution of the obtained averages of *T. agilis* and *T. mirabilis* cartwheel hubs, we divided each dataset into two groups with the same number of subtomograms, aligned them and averaged independently. The final averages displayed low resolution blobs and noise at outer regions of the box, which are removed by filtering or Hide Dust function (Chimera) for the density map visualization. However, they might bias the overall resolution determination of the final averages towards lower resolution values. Therefore, we calculated the FSC only of the cartwheel hub with ~ 3 nm of spokes. Half-averages were masked by a centered spherical mask of ~ 33 nm (96 pixels) and the Fourier Shell Correlation (FSC) was calculated. Dynamo software was used to plot FSC curve and estimate the approximate resolution at different FSC cutoffs. To confirm the estimated values, we used also the FSC validation server (EMBL-EBI), which determines the resolution as an exact numerical value. However, both softwares likely use different algorithms to calculate FSC, so the final resolution determination might vary, mainly when a result comprises artefactual averaging of noise components (Fig. 2.14 B).

#### 2.2.8.1 Final average resolution of *T. agilis* cartwheel hub density maps

The Avg-C1 and Avg-C9 FSC curves shows larger discrepancies, of ~ 10 Å, between determined resolutions using the two softwares, mainly at the lower FSC cuttoffs, 0.143 and 0.333 for Avg-C1 and Avg-C9, respectively (Fig. 2.14 B). Therefore, we considered the achieved resolution for Ref-C1 and Ref-C9 maps as the final attained resolution. The resolution estimated by Dynamo and calculated by FSC validation server show a difference of only ~ 3 Å, thus were in a good agreement. We obtained ~ 40 Å at the FSC criterion 0.5 and ~ 25 Å using the FSC criterion 0.143 for the Ref-C1 map (Fig. 2.14 A). This is slightly better than what was obtained previously, 34 Å at FSC 0.143 (Guichard et al., 2013); however, because the curve does not drop completely to 0 and oscillate at low FSC values (~0.1-0.15), the resolution at FSC (0.143) must be considered with caution, due to possible artefacts. The symmetrisation improved the final resolution; however, this curve also shows a probable artefactual noise averaging, as visible on right side of the plot (Fig. 2.14 A) (Penczek, 2010). Therefore, only the resolution of ~25 Å at FSC (0.5) is determinant. We see the same trend also in Avg-C1 and Avg-C9 FSC curves (Fig. 2.14 B).

The local resolution of all four *T. agilis* cartwheel hubs (Fig. 2.7) corresponds to values determined by FSC (Fig. 2.14). However, the apparent high resolution of ~ 15 Å of the Avg-C9 map (Fig. 2.7 B) might be probably due to the artefactual noise averaging, as seen in the FSC curve (Fig. 2.14 B). Regardless the average resolution of the density maps, already the direct comparison of the four final averages shows a relative improvement of the final resolution in maps done without the external starting reference, i.e. the CID inside the ring is resolved more in Avg-C1/C9 maps, than in Ref-C1/C9 (Fig. 2.7). This might be due to the bias imposed by the starting reference, where the CID is resolved as continuous disc at the similar threshold value as used here (Guichard et al., 2013).
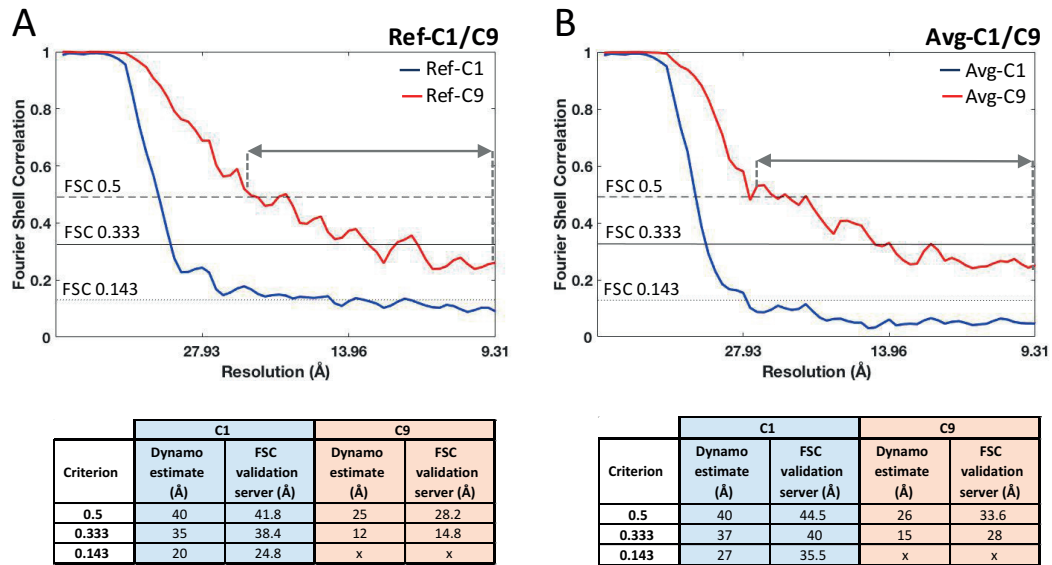
Fig. 2. 14 Resolution determination of *T. agilis* cartwheel hub density maps using Fourier Shell Correlation (FSC). **A** Comparison of FSC curves calculated from the two datasets using *Trichonympha* sp. map (Guichard et al., 2013) as a template for subtomogram alignment and averaging with (C1) and without (C9) ninefold symmetrisation. **B** Comparison of FSC curves calculated from the two datasets without an external template for subtomogram alignment and averaging with (C1) and without (C9) ninefold symmetrisation. Grey double-arrows depict a probable noise averaging artefact, visible on the Ref-C9 and Avg-C9 curves, as a curve portion which does not drop to 0 and oscillate at higher FSC frequencies. The tables show resolutions estimated using Dynamo software and calculated using FSC validation server (EMBL-EBI) at different FSC cutoffs.

## 2.2.8.2 Final average resolution of *T. mirabilis* cartwheel hub density maps

We calculated the FSC also for all four *T. mirabilis* cartwheel hub density maps to determine the overall resolution. On average, we generated C1 maps of ~ 40 Å resolution when using FSC 0.5 criterion and ~ 34 Å when using FSC 0.143 criterion (Fig. 2.15). C9 symmetrization improved the overall resolution as seen from the shift of the red curves to the higher resolution values (Fig. 2.15). However, this again increased artefactual noise averaging at higher resolution and made the resolution determination at lower FSC thresholds (0.143 and 0.333), as well as the local resolution, going down to ~ 15 Å for Ref-C9 (Fig 2.11 A), inaccurate.

**A** Ref-C1/C9 — Ref-C1, Ref-C9

**B** Avg-C1/C9 — Avg-C1, Avg-C9

(Fourier Shell Correlation vs Resolution (Å); FSC 0.5, FSC 0.333, FSC 0.143)

| Criterion | C1 | | C9 | |
|---|---|---|---|---|
| | Dynamo estimate (Å) | FSC validation server (Å) | Dynamo estimate (Å) | FSC validation server (Å) |
| 0.5 | 42 | 40.5 | 31 | 36.1 |
| 0.333 | 39 | 38.8 | 25 | 34.4 |
| 0.143 | 35 | 36.4 | x | 10.5 |

| Criterion | C1 | | C9 | |
|---|---|---|---|---|
| | Dynamo estimate (Å) | FSC validation server (Å) | Dynamo estimate (Å) | FSC validation server (Å) |
| 0.5 | 40 | 40.7 | 28 | 35 |
| 0.333 | 37 | 38.3 | 20 | 24 |
| 0.143 | 32 | 34.6 | x | 8.6 |

Fig. 2. 15 Resolution determination of *T. mirabilis* cartwheel hub density maps using Fourier Shell Correlation (FSC). **A** Comparison of FSC curves calculated from the two datasets using *Trichonympha* spp. map (Guichard et al., 2013) as a template for subtomogram alignment and averaging with (C1) and without (C9) ninefold symmetrisation. **B** Comparison of FSC curves calculated from the two datasets without an external template for subtomogram alignment and averaging with (C1) and without (C9) ninefold symmetrisation. Grey double-arrows depict a probable noise averaging artefact, visible on the Ref-C9 and Avg-C9 curves. The tables show resolutions estimated using Dynamo software and calculated using FSC validation server (EMBL-EBI) at different FSC cutoffs

.

## 2.2.9   SAS-6 as the major constituent of cartwheel ring

As we already know, the major component of the cartwheel ring is the protein SAS-6. It was shown that SAS-6 possesses intrinsic properties to self-oligomerize and form ring-like structures with a diameter of ~22 nm *in vitro* (van Breugel et al., 2011, 2014; Guichard et al., 2017; Hilbert et al., 2016; Kitagawa et al., 2011). Moreover, the modelled *T. agilis* TaSAS-6 could be placed inside the electron density of the cartwheel hub (Guichard et al., 2012). Here, we fitted the same ring into the *T. agilis* cartwheel hub density map, and tried the same also with *T. mirabilis* hub density map. In both cases, we observed that the predicted structure did not fit perfectly inside the cartwheel ring densities (Fig. 2.16). This is because the shape of the cartwheel hub is not a perfect ring, but instead a nonagon (Fig. 2.16 A and D), with a diameter larger than the ~22 nm of the predicted SAS-6 ring, i.e. of ~24.1 or 24.3 nm in the *T. agilis* and *T. mirabilis* cartwheel hub, respectively. This raises the possibility that the conformation of higher order SAS-6 oligomers is different *in situ* than *in vitro*.
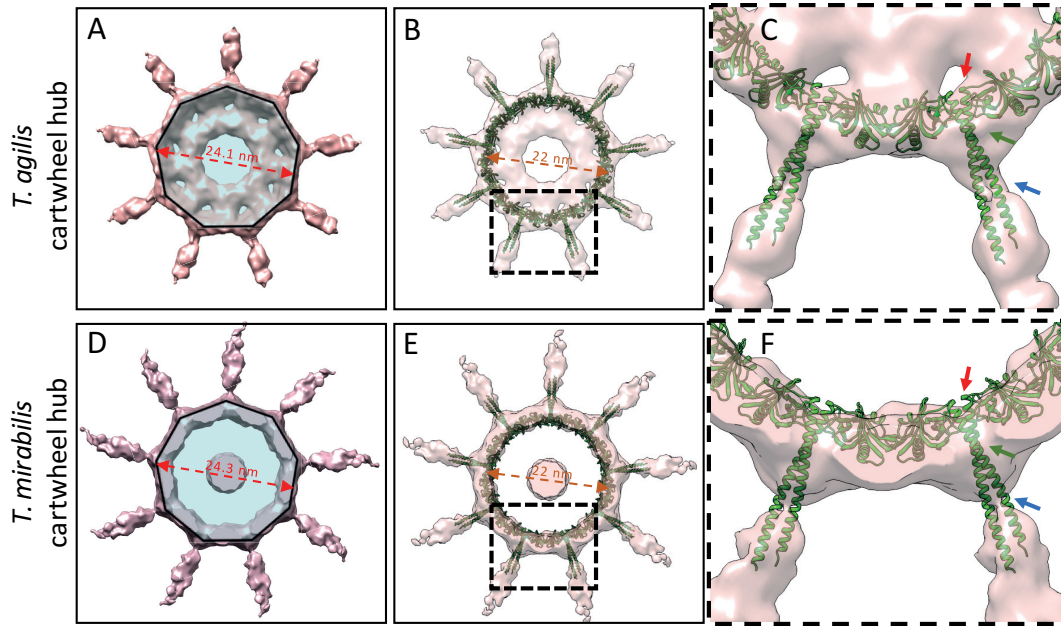
Fig. 2. 16 The nonagonal shape of the cartwheel hub does not fully embed the CrSAS-6 ring. **A** and **D** Top views of cartwheel hubs of *T. agilis* (A) or *T. mirabilis* (D), showing that their shapes are perfect nonagons. **B** and **E** Fitting of the structural model of CrSAS-6 ring (green) with a diameter of ~ 22 nm into the 3D density map of the cartwheel hubs. Dashed boxes represent regions magnified in C and F. **C** and **F** Magnified view of two CrSAS-6 homodimers fitted in the 3D map. The ring-shaped CrSAS-6 model does not fit perfectly into the polygonal shape of the cartwheel hub, since CrSAS-6 molecules protrude outside the density (red and blue arrows). Green arrows point to a CrSAS-6 homodimer region in the 1st coiled-coil turn that probably should be placed where the blue arrow points, according to the narrow 3D densities at that point and the absence of fitting of these densities with the coiled-coil turn.

## 2.3    Discussion

The evolutionarily very conserved centriolar structure and the knowledge that *Trichonympha* spp. harbors centrioles up to ~5 µm in length with a long cartwheel-bearing portion made this genus an amazing model to study the proximal region of the centriole. Guichard et al. (2012, 2013) uncovered the 3D architecture of the *Trichonympha* spp. centriole proximal region at ~ 40 Å resolution using cryo-ET and subtomogram averaging. Another cryo-ET study with *Chlamydomonas* centrioles did not analyze the cartwheel, which was lost during purification (Li et al., 2012). However, in the advent of high-resolution cryo-ET and improved subtomogram averaging, we wanted to go further and also exploit the acquisition of a direct electron detector in our electron microscopy facility to get a higher resolution view not only of the *Trichonympha* centriole. Further, we wanted to compare the centriole structures from different species to find out the extent to which it is conserved. Moreover, as you will read further, we studied the composition of the *T. agilis* centriole and plan to use the new 3D density map as a reference to place identified and annotated proteins into it.

### 2.3.1  *T. agilis* cartwheel ring

Already visible in the first density map solved at ~ 40 Å resolution (Guichard et al., 2013) and now confirmed in this study, the cartwheel hub of *Trichonympha* spp. centrioles displays a perfect nonagonal shape (Fig. 2.16), with a diameter of ~24 nm. The diameter is ~ 2 nm larger than the ninefold symmetrical rings formed by SAS-6 proteins from different species *in vitro*, including *Chlamydomonas* (Guichard et al., 2017; Hilbert et al., 2016; Kitagawa et al., 2011), *Leishmania major* (van Breugel et al., 2014) or human (data in Chapter IV). A diameter of ~ 25 nm was already presented in the first study of *Trichonympha* spp. centrioles (Gibbons and Grimstone, 1960). Careful investigation of the symmetrized cartwheel hub from the published data on *Chlamydomonas* and on *Trichonympha* spp. centrioles confirms that the shape of the cartwheel hub is polygonal and not ring-like (Guichard et al., 2013) (Fig. 2.17).
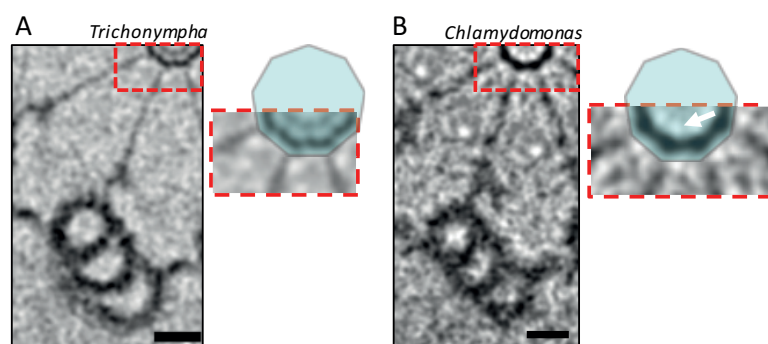


Fig. 2. 17 Nonagonal shape of cartwheel hub in *Trichonympha* and *Chlamydomonas* centriole. **A** and **B** Portion of *Trichonympha* spp. (A) and *Chlamydomonas* (B) centriole with depiction of nonagonal shape of cartwheel hub (in red dashed box, overlaid with blue nonagon). Note density inside the *Chlamydomonas* cartwheel hub (white arrow). Scale bar: 20 nm. With permission, adapted from Guichard et al. (2013).

The modelled ring of CrSAS-6 does not fit inside the polygonal shape and is obviously smaller than the cartwheel hub densities (Fig. 2.16). Up to date, we have not determined the crystal structure of TaSAS-6 or its newly discovered paralogs (see Chapter III). However, the structural prediction of all four TaSAS-6 based on the CrSAS-6 crystal structure shows a large extent of structural homology (Table 3.4 and Fig. 4.3). Therefore, we do not expect a different fitting with the actual structure with any of the four TaSAS-6.

The difference in SAS-6 protein quaternary structure between *in vitro* and *in situ* might result from different folding of SAS-6 proteins when in the presence of an interaction partner. For instance, the CID might regulate the symmetry and diameter of the SAS-6 ring as it interacts with the cartwheel hub at the SAS-6 N-terminal oligomerization interface (Fig. 2.9). However, further analyses and more importantly the identity of the CID are needed to unravel the folding process of the SAS-6 and its interaction with binding partners.

In the case of *T. mirabilis*, the cartwheel hub displays the same shape and diameter. However, the CCD inside the cartwheel hub is not visibly connected to the inner cartwheel hub wall. We hypothesize the presence of some thin unresolved connections of the CCD to the cartwheel, but it remains unclear how a probably thin structure in a symmetry different than ninefold would change the conformation of a SAS-6 ring with a perfect ninefold symmetry.

### 2.3.2  Cartwheel inner densities (CID)

Even without high resolution 3D cryo-ET maps of the cartwheel hub from other species, there are indications that some sort of cartwheel inner densities might be a conserved feature of centrioles. Thus, some weak density resembling the CID are apparent when examining the *Chlamydomonas* centriole and its cartwheel hub (Fig. 2.17 B) (Guichard et al., 2013). Furthermore, more than 30 years ago, some unknown densities, named stellate granules, were described also in the fungus *Phlyctochytrium irregular* (Fig. 2.18) (McNitt, 1974), which would suggest that the presence of some sort of densities inside the cartwheel hub is a rather general feature.
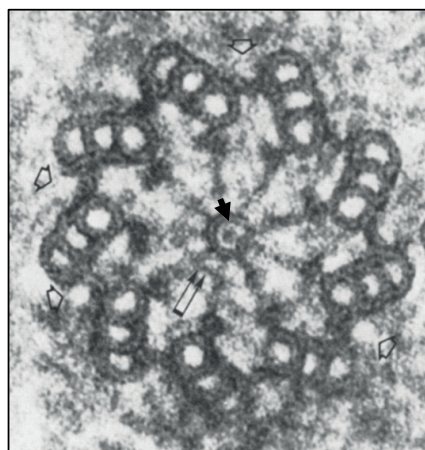


Fig. 2. 18 Cross section of *Phlyctochytrium irregular* centriole. Note the central density inside the cartwheel hub, named the stellate granule (black arrow).

We were able to determine also the polarity of the CID, which is present on the proximal side of each cartwheel ring. Together with the polarized Pinhead and A-C linker (Guichard et al., 2013), which are on the periphery of the centriole, so does the central cartwheel exhibit polarity along the proximal-distal centriole axis, which may have important consequences when considering mechanisms of cartwheel formation. *In vitro* data of CrSAS-6 stacking revealed a stepwise mechanism of preferential incorporation of pairs of SAS-6 rings onto each other (Guichard et al., 2017). Perhaps the CID functions as a seed to promote SAS-6 ring formation with the correct symmetry and diameter before stacking of pairs of SAS-6 rings during centriole assembly. Alternatively, the CID might

have another stabilization function by interacting with the SAS-6 ring internal wall during or perhaps only after cartwheel assembly to stabilize the ring.

The CID addition during or only after SAS-6 ring stacking might be possible to observe in the elongated centriole of *T. agilis*, because this would imply the presence of centrioles with some intermediate forms of cartwheel stack with the CID present only partially along the cartwheel hub. However, the *Trichonympha* centriolar complex is a stable structure and do not degrade when a cell enters mitosis (Kubai, 1973). Therefore, to explore CID addition after the centriole and cartwheel formation, one would need to analyse the *T. agilis* centrioles just after mitosis when centriole formation of approximately half of centrioles occurs (the other half is inherited from the parental cell) (Kubai, 1973). For this to be possible, we would need to know the life cycle and the process of centriolar complex formation in *T. agilis*, and also aquire many tomograms with the entire length of the cartwheel.

### 2.3.3 *T. mirabilis* centriole with cartwheel central densities (CCD)

My work reveals also a second type of cartwheel densities called CCD. The CCD has the shape of an elongated rod present along the whole length (~1 μm) of four *T. mirabilis* cartwheels imaged for cryo-ET analysis. The structure of the CCD is reminiscent of that of the stellate granules found in *P. irregular* (Fig. 2.18) (McNitt, 1974), which might suggest some sort of evolutionary conservation. The striking feature of the CCD is the absence of a visible periodicity. The density map of the CCD appears to be an elongated rod with smooth surface without any apparent connections to the inner wall of the cartwheel hub along the whole structure (Fig. 2.12). We tried to average the CCD by applying the cylindrical mask only on the volume inside the cartwheel hub to uncover possible weak features, which are likely to have been erased from the average of the whole hub, if they do not to follow a nine-fold symmetry. However, we did not obtain a better average or an indication of any connection in doing so (data not shown), even when the particles for subtomogram averaging were picked every 17 nm, which is the periodicity of merged spokes. The CCD might have also a helical confirmation, which should be resolved by applying dedicated helical calculation parameters, which we plan to do in the near future. Nevertheless, the presence of CCD inside the *T. mirabilis* cartwheel hub illustrates the biodiversity of the cartwheel internal structures and might suggest that they are specialized structures with possible different functions in different species.

### 2.3.4 Flexible fitting of SAS-6 ring

Based on the cartwheel ring diameter, as well as on the direct fit of the modelled SAS-6 ring into the cartwheel density map, with the CID, as well as the CCD, we could assess that the SAS-6 ring (Fig. 2.16) is present in a different conformation *in situ* than that expected from *in vitro* experiments. The conformational change might be due to a binding partner, such as the CID or the ICP, which directly binds to TaSAS-6. We noted that when the CCD is present in *T. mirabilis* cartwheel hub, with no apparent connections to the SAS-6 ring, the ICPs appear to be larger (Fig. 2.19 B) than in *T. agilis* cartwheel hubs with the CID (Fig. 2.19 A). This might suggest that the ICP stabilizes the cartwheel stack in the absence of the CID
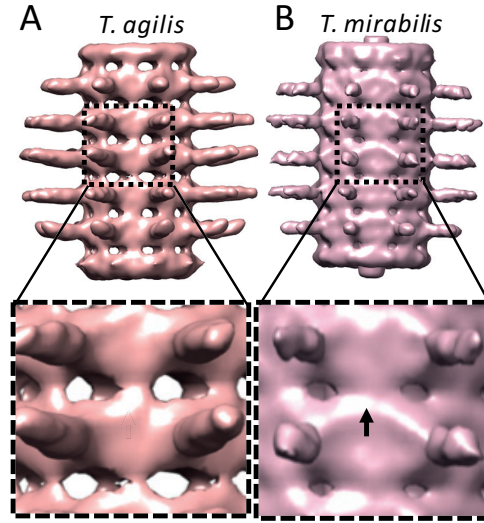
Fig. 2. 19 Direct comparison of the ICPs present in the cartwheel hubs of *T. agilis* with the CID and *T. mirabilis* with CCD. **A** Side view of the *T. agilis* cartwheel hub stack with zoom in on the ICP. **B** Side view of the *T. mirabilis* cartwheel hub stack with zoom in on the ICP. Note that the *T. mirabilis* ICPs are bigger compared to *T. agilis* ICP.

To uncover the conformation of TaSAS-6 *in situ*, we want to proceed with a Molecular Dynamics Flexible Fitting simulation (Trabuco et al., 2009), which is designed to fit the protein structure into a density map using it as a template and by adding adequate external forces to the protein atomic structure.

### 2.3.5 Centriole as a puzzle

The cartwheel hub of *T. agilis* or *T. mirabilis* refined in this study is only a small part of the whole proximal region. We took advantage of the stability of the hub even in flattened centrioles to analyse and average it separately from other centriolar components, such as spokes that can be distorted.

If we continue in this manner to resolve small subvolumes from the same dataset of tomograms, we will probably not be able to merge all the individual portions, because of missing parts, such as the spokes. However, we can place such small reconstructions into the previously uncovered 3D density map from *Trichonympha* spp., if it does not diverge importantly from the new structure. In this manner, this could as a template where we can insert new individual pieces of centriole with improved resolution.

The above procedure with numerous small reconstructions might be beneficial also outside *Trichonympha* centrioles. Canonical centrioles of ~ 500 nm are much shorter than those of *Trichonympha*, and the short ~ 100 nm cartwheel inside the proximal region might be unstable during the centriole purification process, and therefore be slightly distorted in the final tomograms. As a result, it is likely to be challenging to obtain hundreds to thousands of undistorted centrioles in order to get their cartwheel structure. Therefore, one might process small portions of cartwheel or

centriole independently and then assemble them using the available *Trichonympha* spp. template (Guichard et al., 2013).

### 2.3.6 Can we reach better resolution?

We acquired and processed 5 tomograms with 8 *T. agilis* centrioles and 4 tomograms with 4 *T. mirabilis* centrioles. The final average resolution of ~ 25 Å at FSC 0.5 of *T. agilis* cartwheel hub that we reached (Fig. 2.14) is improved compared to 38 Å from the earlier study (Guichard et al., 2013). The overall resolution of *T. mirabilis* is ~ 30 Å at FSC 0.5 (Fig. 2.15). However, we are far from reaching the limit of resolution attainable in our setup. With a pixel size value of 3.491 Å, we should be able to reach up to ~ 7 Å resolution. Several practical reasons, as the electron dose accumulation in the sample or high defocus value during tomogram acquisition, might limit the maximal practically attainable resolution. However, it might be still possible to reach ~ 3x pixel size resolution, ~ 15 Å. There are several ways how we can improve the final resolution, which some of them are discussed below.

Cryo-electron tomography with subsequent subtomogram averaging is a technique based of the averaging of a large number of identical particles to increase the signal-to-noise ratio. However, we worked with less than 900 for *T. agilis* and less than 400 for *T. mirabilis* cartwheel hub, which was enough for the overall structure determination, but might be insufficient to refine the structures at higher resolution. Therefore, one might increase the number of particles taken for subtomogram averaging in order to improve the resolution.

Furthermore, due to sample motion during the tilt-series acquisition and data collection in movie mode on the direct electron detector, the important step of frame alignment should be done during data pre-processing in order to correct for beam-induced, or drift, motion. This is done automatically by the FEI software pipeline during tilt-series acquisition. However, we would like to test a different software, MotionCor2 (Zheng et al., 2017), and assess whether with the use of a different algorithm for motion correction, we can reach a better resolution.

With the technical advances and improvements in the entire field of cryo-EM, including the development of direct detection devices and of software that have revolutionized the field of cryo-electron tomography, one can reach near-atomic resolution, below 4 Å (Turoňová et al., 2017). However, the type of microscope is another important factor when considering what resolution can be reasonably achieved. We observed our sample and acquired tilts series on a FEI Tecnai F20 Cryo microscope with 200 kV field emission gun, which is dedicated to cryo-electron microscopy and tomography of biological specimens requiring low dose imaging. However, our need of thicker samples to prevent centriole flattening would profit from a higher voltage (300 kV) microscope, in which penetration depth is higher, which means less blur on thicker or denser samples and fewer multiple electron scattering. Moreover, higher accelerating voltage causes less damage to a specimen at the same beam intensity, because the number of interaction of electrons with the specimen is lower.

Furthermore, biological samples are damaged by an electron beam. This damage accumulates during exposure and causes changes in molecular structure (reviewed in Glaeser, 2016). Moreover, the higher-resolution information is lost first, so that it is present only in the first few acquired images from the tilt series. This might cause a loss of resolution also in our case. We used the unidirectional tilt scheme (Fig. 2.20), beginning at higher tilt angle (-60º) where our thick sample appears even thicker and the sample might be damaged and/or high resolution information lost until it comes to tilt angles around 0º, where the ice thickness is lowest and could provide the best resolution. Therefore, we might envisage using a bidirectional or a dose-symmetric tilt scheme, where the acquisition starts at 0º, where the ice is thinnest and high-resolution information remains present (reviewed in Wan and Briggs, 2016).
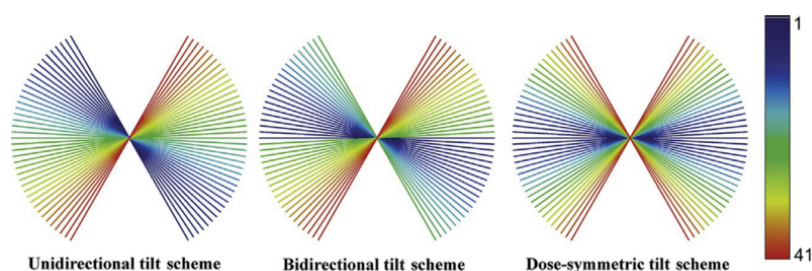


Fig. 2. 20 Tilt schemes used in cryo-ET. Unidirectional, where the acquisition starts at the high tilt angle (here in blue) and continues through 0 until the opposite high tilt angle (on the left). A bidirectional tilt scheme can be used to acquire high resolution structures at the tilt angle 0º at the beginning, where the ice thickness is thinnest (in the middle). The oscillating dose-symmetric tilt scheme ensures that the information with the highest resolution is acquired where the ice thickness is thinnest, and that the distribution of the beam-induced loss of information is spread uniformly (on the right). With permission from Wan and Briggs (2016)

## 2.4    Personal contributions

I prepared the sample for the tomogram acquisitions. I aligned acquired tilt-series and reconstructed all tomograms used for the study. I picked and prepared subtomograms and ran the subtomogram alignment and averaging for all final averages which are showed here, as well as of all half-datasets for the FSC calculation. I determined the resolution using Dynamo software, FSC validation server and ResMap software. I prepared all 2D and 3D representations of final averages. I analysed the *T. mirabilis* centriole raw tomogram by FFT.

This project was done in collaboration with:

- Davide Demurtas from the Interdisciplinary Center for Electron Microscopy and Graham Knott from the Biological Electron Microscopy Facility at EPFL (Lausanne). Davide screened the samples and acquired tomograms.

- Maeva Le Guennec and Paul Guichard from the University of Geneva who helped me to analyse some tomograms and improve tomogram reconstruction and subtomogram averaging. Maeva designed the subtomogram alignment and averaging method, helped me with the overall and local resolution determination and with fiducial-less patch alignment.

- Sergey Nazarov from the University of Basel for his input and discussion.

# Chapter 3     Uncovering new proteins of the proximal centriolar region

*"No one can whistle a symphony.*

*It takes a whole orchestra to play it."*

H. E. Luccock

## 3.1     Background

The centriole, as one of the largest non-membraneous protein-based organelle of the cell, is composed of many proteins. Before the first comprehensive proteomic approaches, all studies of centrosomal or centriolar components were focused on one protein or a set of proteins, and therefore did not uncover the overall composition of centrioles or centrosomes. However, substantial efforts have been made in the last fifteen years to identify an exhaustive list of centriolar and/or centrosomal components in a number of systems. The first proteomic characterization of the human centrosome demonstrated the advantage of protein correlation profiling in the analysis of samples not purified to homogeneity (Andersen et al., 2003). A comparative genomics approach was then used to uncover the centriolar and flagellar proteome of human and *Chlamydomonas,* respectively (Li et al., 2004). The first purely centriolar proteomic study was performed one year later on *Chlamydomonas* centrioles, which do not harbor any appreciable PCM (Keller et al., 2005). Additional studies, including of naked centrioles of human sperm cells (Firat-Karalar et al., 2014), of *Tetrahymena* (Kilburn et al., 2007), or more recently of *Chlamydomonas* again (Hamel et al., 2017), complement the information about centriole composition across evolution. What about specific regions of the centriole? Can one obtain the composition of proteins present in the proximal centriolar region harboring the cartwheel, which is critical at the onset of centriole biogenesis?

The proximal cartwheel-containing region plays an important role in centriole biogenesis and in imparting the evolutionary conserved nine-fold symmetry of centrioles (reviewed in Gönczy, 2012). However, except for a few core structural proteins of the cartwheel and Pinhead, including SAS-6/Bld12p (Dammermann et al., 2004; Leidel et al., 2005; Nakazawa et al., 2007), Cep135/Bld10p (Matsuura et al., 2004; Ohta et al., 2002) (Hiraki et al., 2007; Matsuura et al., 2004), CPAP (Kirkham et al., 2003; Leidel and Gönczy, 2003) and STIL (Tang et al., 2011), alpha and beta-tubulins of microtubule triplets, and some further promising candidates such as POC1 (Keller et al., 2009), we do not have a detailed protein composition of the centriolar proximal region. Due to its short length of ~ 100 nm, which occupies only a small portion of the ~500 nm long centriole, and the impossibility

to extract the proximal region to analyse it separately, candidate proteins from this region might have been under represented in proteomic analyses of centrioles. This is exemplified by the quantitative analysis of the human centrosome, where Bauer et al. (2016) found only 170 and 72 molecules of HsSAS-6 or STIL per centriole, respectively. As a result, such proteins are likely to be overlooked in most proteomic screens, shadowed by other more abundant centriolar proteins that are present throughout the entire centriole. Therefore, one aim of my PhD project has been to uncover proteins of the proximal region of centriole, using the exceptionally long proximal part of the *T. agilis* centriole, with available genomic and single cell transcriptomic data from this species.

## 3.2    Results

### 3.2.1    *Trichonympha* and its host

The *Trichonympha* spp. centriole, with its elongated cartwheel-bearing proximal part, is an interesting object of study in this regard. Indeed, proteins building the cartwheel and additional structures, such as Pinhead or A-C linker, should be enriched in a ~ 4 μm long centriole where ~ 90% of the length is occupied by the cartwheel-bearing region.

Previous work on *Trichonympha* spp. was conducted mostly on the species *T. collaris* and *T. campanula*, which are present in the hindgut of the termite *Zootermopsis nevadensis* (Guichard et al., 2012, 2013). However, the mixture of two species, as well as the absence of available genomic or transcriptomic information, left us unable to utilize these for proteomic analysis.

Therefore, we searched for a *Trichonympha* sp. in which we could obtain genomic or transcriptomic information. We teamed up with Prof. Y. Hongoh, from the Tokyo Institute of Technology, Japan, who works on the symbiotic relationship of the termite *Reticulitermes speratus* with its hindgut symbiont, *T. agilis*, and the latter's bacterial endosymbiots. The Hongoh team sequenced and has assembled the gigantic ~ 20 Gbs genome of *T. agilis* and obtained single-cell transcriptomic data from five cells altogether (Y. Hongoh, personal communications). These two databases were used in our proteomic studies (Table 3.1).

*Trichonympha* spp., its symbiotic relationship with termites, its life cycle and its interesting cellular composition, have been studied for more than 60 years (Gibbons and Grimstone, 1960; Guichard et al., 2012, 2013; Ikeda-Ohtsubo and Brune, 2009; Slaytor et al., 1997). Yet, no proper straightforward cultivation method has been developed to keep this obligate anaerobe alive for more than two weeks. Only *T. sphaerica* was kept outside of the termite gut in an anaerobic setup for two weeks, during which time it only doubled its initial cell number (Trager, 1934). Therefore, the principal method of investigation of *Trichonympha* spp. is based on their careful extraction from the gut of their termite host.

The *T. agilis* host, the termite *Reticulitermes speratus*, is significantly smaller than the *Zootermopsis nevadensis* termite used previously (Guichard et al., 2012, 2013). Consequently, *R. speratus* harbors a significantly lower number of *Trichonympha* cells than *Z. nevadensis*. I could estimate the number

of *Trichonympha* spp. cell in *Z. nevadensis* to be ~ 10,000 compared to ~ 100 - 700 *T. agilis* cells in *R. speratus*, depending on the particular individual termite.

Table 3. 1 Details about the *T. agilis* genome and transcriptome database. The estimated size of the *T. agilis* genome is 10x larger than the genome assembly size, which might suggest either the incompleteness of the assembly or the presence of large genome duplications (Y. Hongoh, personal communication). Single-cell transcriptomic database obtained from five cells altogether contains redundant translated mRNA sequences. Therefore, likely duplicates of protein sequences were grouped based on the percentage of sequence identity and length to obtain an approximative number of unique protein sequences. Sequences with at least 80% sequence and length identity were clustered together (80% cutoff).

| | Estimated genome size | Assembly size | Annotations |
|---|---|---|---|
| Genome | ~ 20 Gbs | ~ 2 Gbs | NA |

| | Raw database protein number | Estimated protein number (80% cutoff) | Annotations |
|---|---|---|---|
| Transcriptome | 143,480 | 14,017 | 0.16% |

Termites in general, just like bees or ants, are eusocial insects that live in highly organized colonies with differentiated castes, including reproductive individuals such as primary and secondary queens and kings. Termite queens are among the longest living insects on Earth (up to 50 years). Non-reproductive active workers build a brood and take care of the reproductive group and of soldiers, which defend the colony and serve as antifungal agents to keep the colony in good conditions (reviewed in Tian and Zhou, 2014).

A laboratory "nest" is in general a piece of wood taken from the termite's natural habitat (Fig. 3.1 A), which contains only workers and soldiers (Fig. 3.1 B and C). Due to the different functions of these two castes, only workers should be removed from the colony in order to leave soldiers and thus keep the colony healthy as long as possible. Many individuals from the worker caste differentiate irreversibly into a reproductive alate stage (Fig. 3.1 C) after several months of separation from their queen in laboratory conditions. During this time, they gradually loose their flagellate symbionts, probably due to the change in their social status, as well as hormonal or environmental change during ecdysis (reviewed in Nalepa, 2017). Therefore, the entire colony must be replaced by a new one once workers change into the alate reproductive stage.

Fig. 3. 1 *Reticulitermes speratus* termites. **A** A termite "nest" kept in the laboratory. **B** Five workers. Scale bar: 5 mm. **C** Three casts present in the laboratory nest; a worker, a mandibulate soldier and an alate reproductive stage. The nymph, the young individual, and the reproductive nymph, the latter being an intermediate stage between worker and alate stage, are not depicted, but are sometimes present depending on the age of colony and time spent in the laboratory. Nymphs are mostly present inside a colony freshly collected from a natural habitat, whereas reproductive nymphs are mostly present when the colony is several months without a queen, when workers turn into alate reproductive stage animals. Female reproductive "primary queen" or "secondary queen" and male reproductive "primary king" are not depicted and not present in the laboratory, because the removal of the queen from its natural habitat would destroy the whole colony (Y. Hongoh, personal communication).

### 3.2.1.1  *T. agilis* inside the gut of *R. speratus*

The termite gut microbiome is a fascinating complex of symbiosis between animal and its microbes, consisting of bacteria, archaea, protists and fungi (reviewed in Brune and Dietrich, 2015). The advantage of using the termite *R. speratus* is that, in contrast to other termite species, this one harbors a single species of the *Trichonympha* genus, *T. agilis* (Fig. 3.2 A and C), which is easy to recognize in the microbiome. The second large protist present in the *R. speratus* gut is the Christmas-tree-like shaped flagellar parabasalid *Teranympha mirabilis* (Koidzumi, 1916), which has a different cell morphology (Cleveland, 1938) (Fig. 3.2 A and B). However, as we observed in Chapter II, both species share the overall morphology of centrioles, which are overly elongated with the presence of the cartwheel hub along the centriole. Moreover, contrary to *T. agilis*, *T. mirabilis* might be sometimes absent from the gut of workers (personal observation), and, when present, usually it is there ~ 10x less frequently than *T. agilis* (Fig. 2.3 B). However, this is variable and dependent on a batch of termites, as seen in the the previous Chapter II, where we worked with a batch of termites with larger amount of *T. mirabilis* in their gut.
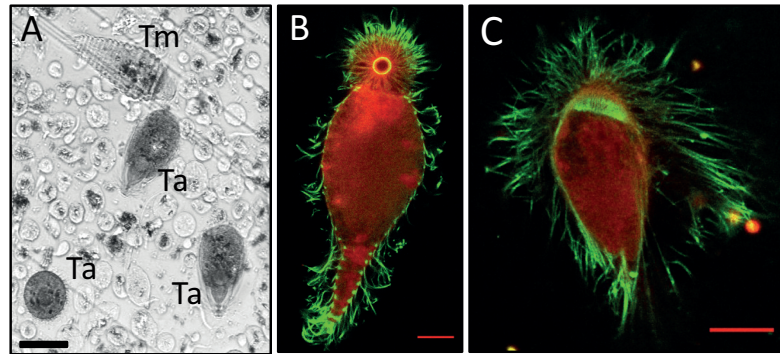
Fig. 3. 2 Large protists inside *R. spe*ratus gut **A** Representative image of the termite gut symbionts, *Trichonympha agilis* (Ta), *Teranympha mirabilis* (Tm) and other smaller species. Scale bar: 50 μm. **B**-**C** immunofluorescence images of isolated cells of the two large symbionts. Flagella stained using antibodies against acetylated tubulin (in green) and membranes by FM4-64 dye (in red). Scale bars: 20 μm. **B** *Teranympha mirabilis*, **C** *Trichonympha agilis*.

### 3.2.2 Rostrum purification for proteomic studies

The protocol for the isolation of *T. agilis* centrioles with their flagella for cryo-ET is well established (see Chapter II and Materials and Methods, Guichard et al., 2012, 2015). However, the presence of long flagella in the resulting suspension, as well as that of contaminants, which are less problematic for cryo-EM studies, because of their easier discrimination on the grid, hinders the use of the resulting sample for proteomic studies.

As mentioned above, *Trichonympha* spp. is an obligate anaerobe living its entire life in the stable environment of the hindgut of its host. Perhaps as a result, stress-related defense mechanisms that shear flagella to reduce the body surface for example in *Chlamydomonas* (reviewed in Quarmby, 2009) appear not to be developed or are silenced in *T. agilis*. Indeed, I tested several methods to shear flagella, including mild mechanical shearing (Lefebvre et al., 1978; Rosenbaum and Child, 1967), dibucaine treatment (Thompson et al., 1974), $Ca^{2+}$- or pH shock treatment (Mitchell et al., 2009; Witman et al., 1972; Zhangs et al., 1991), ethanol treatment (Stevenson and Beane, 2010), sonication or flash-freezing in liquid nitrogen with immediate thawing, but many *T. agilis* cells, in all cases still with flagella, lysed during or just after these treatments. Therefore, I opted instead for mechanical deflagellation with a tissue homogenizer (MagNA lyser), which efficiently sheared flagella from centrioles even in the suspension of lysed cells.
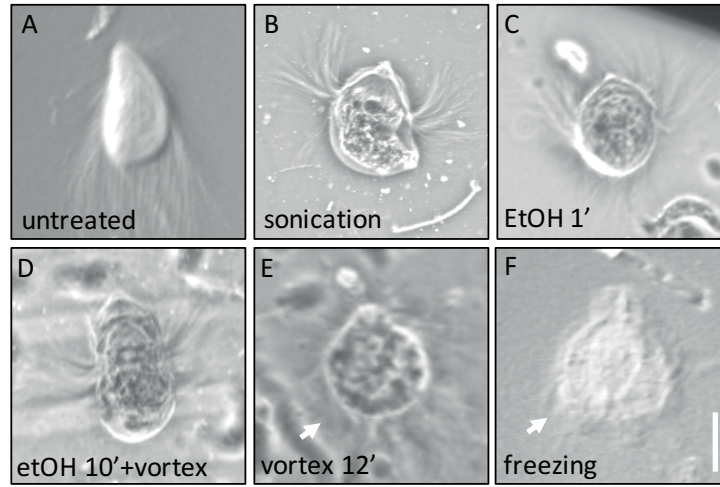
Fig. 3. 3 Gallery of *T. agilis* cells after several trials of deflagellation treatment. **A** Untreated *T. agilis* cell exhibiting long "hairy" flagella around the cell body. **B**-**F** Individual cells keeping their flagella or flagellar fuzzy remnants (arrows) after sonication (B), ethanol treatment or extended ethanol treatment followed by vortex (D), or only high-speed vortex for extended period of time (E) and flash-freezing in liquid nitrogen with immediate thawing (F). Because of the anaerobic nature of the *T. agilis* protist, many cells were lysed during every treatment. Scale bar: 40 μm.

Moreover, the relatively small size of the *R. speratus* gut and the low number of *T. agilis* cells inside required to dissect many termites for an experiment.

Therefore, I set out to optimize the purification of *T. agilis* centrioles in order to i) dissect under the stereoscopic microscope to remove as many contaminants as possible, including gut remnants, wood particles and *T. mirabilis* cells; ii) remove other microbiotic organisms (bacteria, spirochetes and others) in a gentle centrifugation step, iii) shear flagella from centrioles by mechanical treatment with a MagNA lyser tissue homogenizator (Roche) to prevent the co-purification of centrioles and flagella, which would lead to massive detection of flagellar components in proteomic analysis, iv) use OptiPrep density gradient to separate centriole-rich material from remaining cell contaminants.

Due to the elongated shape of the *T. agilis* centriole and the presence of flagellar pieces after the deflagellation step, the established protocols for centriole or centrosome purification that work for other species, including *C. reinhardtii* (Keller et al., 2005) and human centrosomes (Bornens and Moudjou, 1999a; Jakobsen et al., 2013), were not suitable for purification of *T. agilis* centrioles, because pieces of flagella undesirably co-purified with centrioles (data not shown).

However, we observed that despite detergent and mechanical deflagellation treatments, the rostral part of *T. agilis* cell stayed intact (Fig. 3.4). The rostrum is a structure of the shape of a truncated

cone at the cell tip, which serves as a food opening and harbors tightly packed centrioles around the rostral tube in the middle of the structure (Fig. 3.4 C).
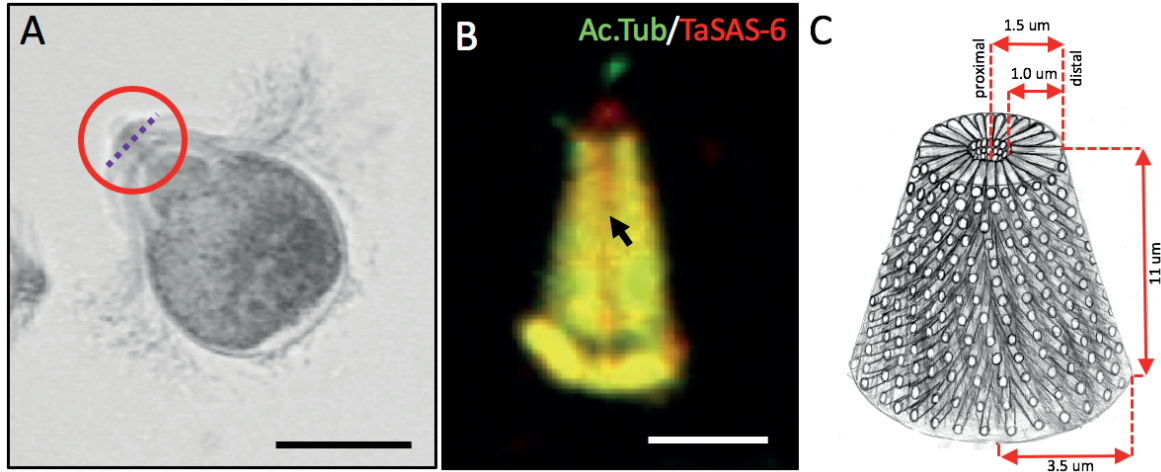


Fig. 3. 4 Rostrum of *T. agilis*. **A** Depiction of *T. agilis* cell with encircled food opening, i.e. the rostrum. Dashed line indicates the cross-section through the rostrum, shown in (D). Scale bar: 40 μm. **B** Deflagellated rostrum stained with anti-acetylated tubulin antibody (green) and anti-TaSAS-6 antibody (red), confirming the presence of a high number of centrioles in the rostrum of *T. agilis*. The rostrum, as a food opening, is a hollow conical frustrum, with the tube in the centre being called the rostral tube (arrow). Scale bar: 10 μm **C** Schematic representation of rostrum and rostrum dimensions. Note the radial arrangement of centrioles as seen on the top layer of the complex. The proximal side of centrioles point to the center of the rostrum, the distal side outwards. Flagella emerging from centriole outwards are not depicted.

### 3.2.2.1   Estimation of centriole number per rostrum

The radial arrangement of centrioles inside the rostrum wall, their compact arrangement close to the rostral tube (Fig. 3.4 C, 3.5.B) (Grimstone and Gibbons, 1966) and the ~ 75-150 μm distance in between each other on the outer wall of the rostrum (Gibbons and Grimstone, 1960) (Fig. 3.5 A and C), as well as the defined dimensions of the rostrum (Fig. 3.4 C) allowed us to estimate the number of centrioles per rostrum. Assuming the shape of the rostrum to be a hollow truncated cone, with a diameter of 7 μm at the base and 3 μm on the top, with a height of 11 μm, the lateral surface of the rostrum where centrioles are localized is ~ 198 μm². Centrioles are separated from each other by a distance of 163 μm horizontally and 75 μm vertically (measured in Fig. 3.5 A). Therefore, we can fit ~ 1430 centrioles into the outer rostrum surface of ~198 μm² (Fig. 3.5 C). To confirm this estimate, we can calculate the number of centrioles by calculating the inner surface of the conical rostral tube, where the centrioles are tightly packed next to each other (Fig. 3.5 B). We estimated that ~ 1380 centrioles might be packed there (Fig. 3.5 B and D), which is in agreement with the previous calculation.
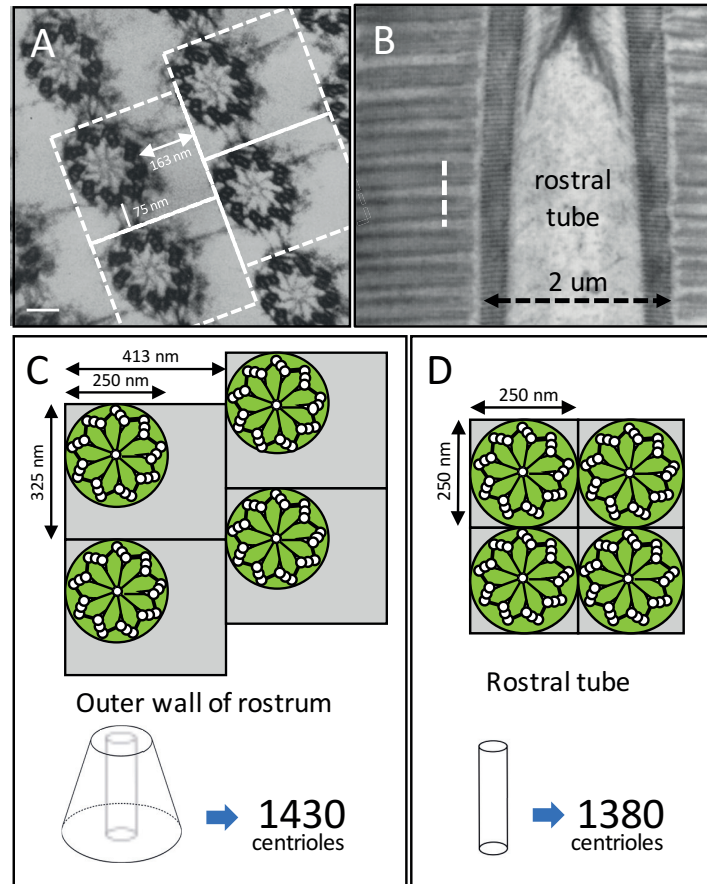
Fig. 3. 5 Two ways to estimate centriole numbers in the rostrum *of T. agilis*. **A** Cross-section through a rostrum with centrioles regularly arranged and interconnected with filamentous structures (Gibbons and Grimstone, 1960). White dashed squares indicate the schematic representation of the centriolar arrangement shown in (C). Scale bar: 50 nm. **B** Transverse section through the rostrum inner tube. The intercentriolar distance is shorter in this region (Grimstone and Gibbons, 1966). The white dashed line indicates where the imaginary cross-section of centrioles was taken to represent the centriolar arrangement shown in (D). **C-D** Schematic representations of centriolar arrangement in the outer wall of rostrum (C) and close to the central tube (D). Green circles represent centrioles in cross-section. Estimated numbers of centrioles are similar.

To uncover *T. agilis* centriolar proteins using a proteomic approach, we established a new protocol for rostrum purification based on density centrifugation gradient using OptiPrep™ (Axis-Shield). This protocol allowed us to separate large and complex rostral structures from the majority of flagellar pieces and other cell contaminants. The detailed rostrum purification protocol is described in Appendix 1.

### 3.2.3   Protein correlation profiling

Based on immunofluorescence analysis of each fraction using the rabbit antibodies against acetylated-tubulin, which stains also flagella, and the mouse antibodies against centrin-2, the number of rostra was quantified, and a qualitative determination of the presence of other components, such as flagella, pieces of flagella or other unspecifically stained contaminants was assessed in order to determine the fractions most enriched and most pure in rostra.

To increase the amount of material in the analysed fractions, we pooled qualitatively similar fractions (Fig 3.6 C). Altogether, 6 fractions from the gradient were processed further (A-F). The detailed procedure of sample preparation for mass-spectrometry based proteomics is described in the Material and Methods section.

As seen in Fig. 3.6 B, rostra were enriched in the 40% OptiPrep density and peaked at the interface between the 40% and 50% OptiPrep layers in this particular experiment. However, the rostrum-enriched fractions harbored also some contaminants as judged by unspecific anti-Centrin2 antibody staining (Fig. 3.6 D). Such unspecific staining was visible in both C and D peak fractions and was highly enriched in the A fraction. In the C and D peak fractions, we could observe also some flagellar pieces (Fig. 3.6 D), stained by antibodies against acetylated-tubulin, but these were mostly enriched in the side fraction E (Fig. 3.6 D). Because of the presence of such contaminants in the rostrum-enriched fractions, and probably of other contaminants that are not detected by the mere use of these two antibodies, we performed mass-spectrometry based proteomic analysis by profiling the relative protein abundance in each fraction A-F.

We analysed the 6 fractions from two independent experiments by nanoscale liquid-chromatography coupled to tandem mass spectrometry (nanoLC-MS/MS) (Table 3.2). However, in experiment 2, the number of proteins assigned was lower, only 523 compared to 810 from the experiment 1, despite using ~ 5x larger samples as starting material. The reason was probably failed sample processing due to the scaling up the sample amount. Therefore, we decided to focus mainky on the results obtained from experiment 1. Nevertheless, results from experiment 2 were analysed similarly to experiment 1, and the hits obtained in that second round are reported further below.

In experiment 1, we identified a total of 810 proteins from more than 14,000 spectra (Table 3.2). Termite contaminants were removed using the *R. speratus* cDNA database (Hayashi et al., 2013). We identified 435 *T. agilis* proteins across six fractions of the gradient. The protein abundance in each fraction was determined by the total spectrum count, as well as by the percentage of the protein sequence covered by the identified peptides, called percent coverage. I defined a fraction enrichment parameter that accounts for both values, EP = total spectrum count * percent coverage * 100.
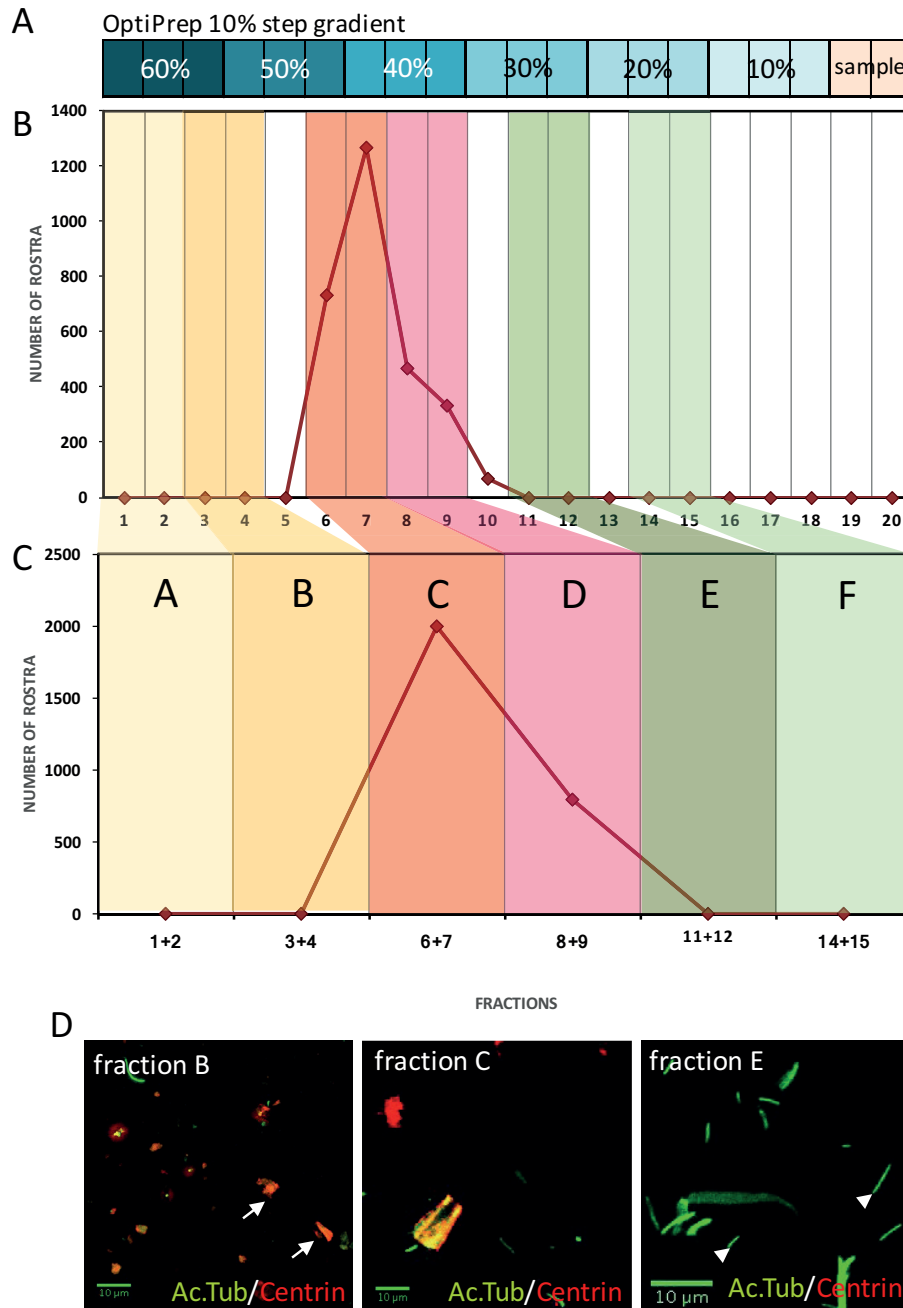
Fig. 3. 6 Rostrum purification by OptiPrep density gradient centrifugation. **A** Schematic representation of the OptiPrep 10% step gradient used for rostrum purification, composed of 6 layers of OptiPrep solution with indicated concentrations. Each layer was fractionated into three fractions, as illustrated. **B** Quantification of rostrum number in each fraction of the density gradient in a given experiment, with a peak in fraction 7. **C** Qualitatively similar fractions assessed by immunofluorescence were pooled together and the 6 resulting fractions, depicted as A-F, were subjected to mass spectrometry analysis. **D** Representative images of three pooled fractions. Note unspecific Centrin2 antibody staining (white arrow) in side fraction B, the rostrum in the peak fraction C, as well as fiber contaminants and flagellar pieces stained with anti-acetylated antibody in side fraction E (white arrowhead).

Table 3. 2 Comparison of two independent proteomic analyses of purified *T. agilis* rostra.

|  | Experiment 1 | Experiment 2 |
|---|---|---|
| number of termites | 70 | 809 |
| number of *T. agilis* cells | 16700 | 249740 |
| number of rostra applied onto gradient | 6300 | ND |
| number of rostra estimated to be in gradient | 2800 | 34500 |
|  |  |  |
| collected MS/MS* | 36380 | 48654 |
| identified MS/MS* | 2181 | 1076 |
| identified MS/MS %* | 5 | 2 |
| peptide sequences identified* | 1154 | 707 |
| peaks repeatedly sequenced* | 1648 | 1067 |
|  |  |  |
| proteins identified in total for *T. agilis* and *R. speratus* | 810 | 523 |
| proteins identified for *T. agilis* | 437 | 310 |
| false discovery rate | 1.0% | 0.7% |

To confirm the enrichment of centriolar proteins in the C and D fractions, as compared to flagellar proteins, we examined the fractionation profiles of known centriolar and flagellar proteins, including TaSAS-6, Centrin2s, as well as α- and β- tubulins. We determined that the centriolar proteins TaSAS-6 and both Centrin2s followed the rostrum profile established based on the immunofluorescence staining of aliquots from each fraction, both in their peptide number and the protein coverage profiles (Fig. 3.7 A-C). Both tubulins, which are present in the centriolar wall and in flagella sheared from rostra had a different fractionation profile. As can be seen in Fig. 3.7 D and E, although tubulins were present in the peak fraction C, an almost two-fold enrichment compared to this was observed in the neighboring fraction D. Overall, we conclude that we can distinguish rostral (or centriolar) proteins from flagellar proteins or other contaminants base on their elution profile.
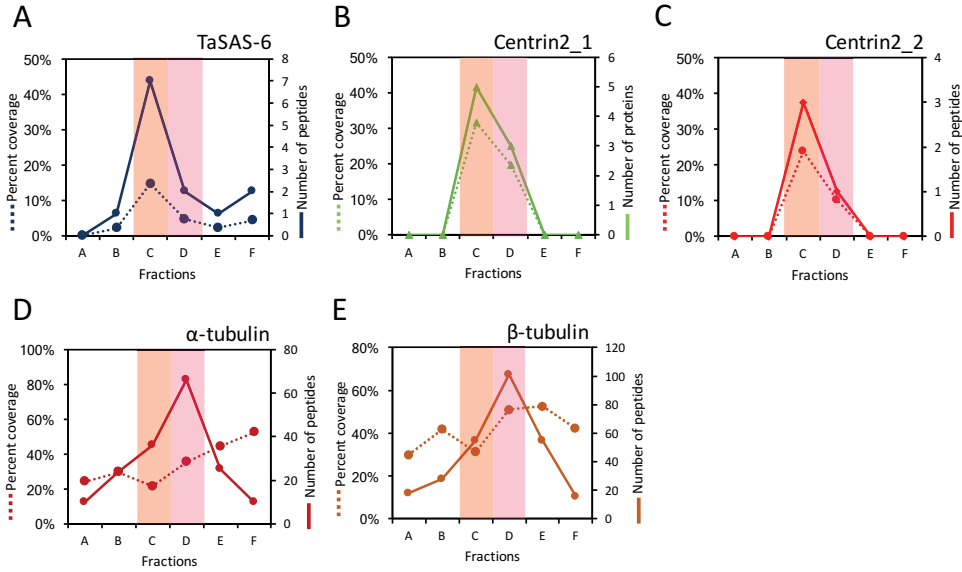
Fig. 3. 7 Fractionation profiles of known centriolar and flagellar components. **A-C** Known core centriolar proteins, TaSAS-6 (A), Centrin2_1 (B) and Centrin2_2 (C), follow the rostral profile with a peak in fraction C and a presence in fraction D. **D-E** α- and (D) β- tubulins (E) are present in the rostrum peak fraction D, but they peak in the neighboring fraction D. Solid lines represents peptide number profiles and dashed line indicates the percentage of the protein covered by the identified peptides.

To identify all proteins that followed the centriolar fractionation profile, we proceeded as illustrated schematically in Fig. 3.8. Firstly, we assessed the abundance of all identified 435 *T. agilis* proteins by the calculated EP value and removed all proteins with 0 or 1 detectable peptide in the fraction C. Secondly, we removed all proteins that were identified with an EP > 0 in the four external side fractions, A, B, E and F, because we expected no centriolar proteins in those fractions. The 33 proteins that met these two criteria were placed in the main hit list. However, we observed an EP value of 8.7 for TaSAS-6-1, due to 2 peptides and ~ 4% percent coverage being present in the side fraction F, even though its overall fractionation profile was consistent with the centriolar profile (see Fig. 3.7 A). A small enrichment for TaSAS-6-1 (EP ~ 2) was seen also in fractions B and E. Overall, this could be due to the degradation of centriolar or rostral structures and the release of individual TaSAS-6 protein or small oligomers thereof. These structures would be expected to have a different density than the whole rostrum, and therefore be localized in different OptiPrep fractions. Therefore, we decided to keep all proteins peaking in fraction C, and with an average EP value < 5 of the four side fractions A, B, E and F into a separate, 14 membered, list (Fig. 3.8). Finally, we got a list of hits with 33 rostral proteins in the main list and 14 on the separate list.
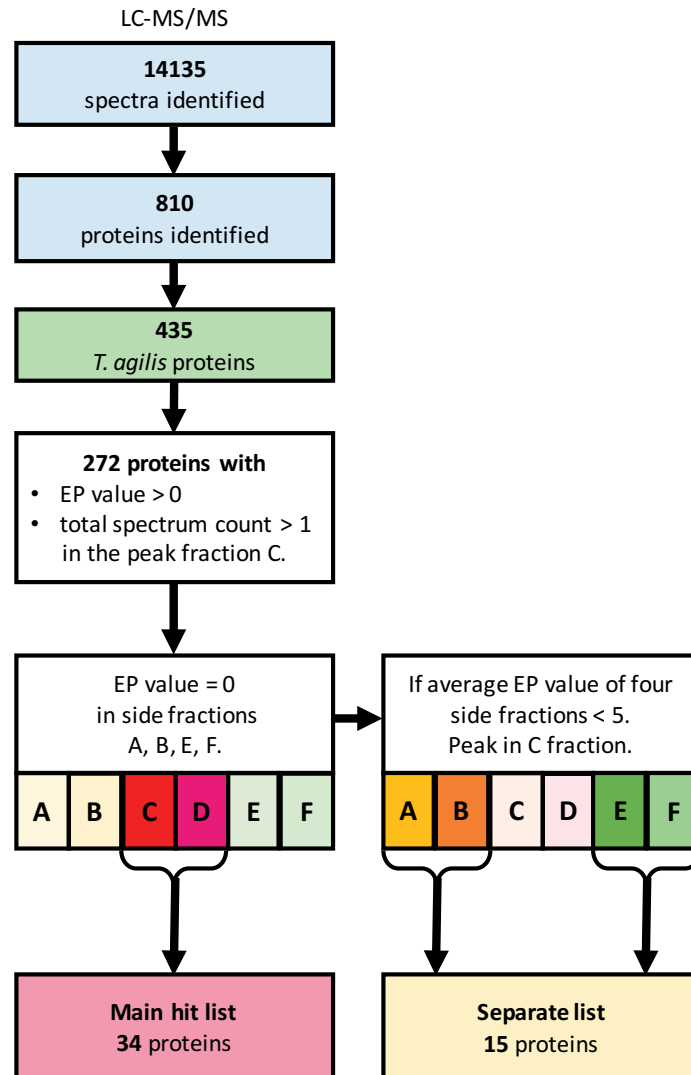
Fig. 3. 8 Schematic representation of subtractive proteomic strategy in identification of *T. agilis* rostrum proteome. Each step was done by manual inspection following the criteria specified in the chart.

### 3.2.4   Proteome of *T. agilis* rostrum

In total, we identified 33 proteins that follow the centriolar profile, as well as 14 proteins that were kept in the separate list, which were present in lower abundance also in one or more side fractions. The identified proteins from the main list had $M_W$ values in the range of ~ 12,000 Da – 135,000 Da, with an average of ~ 48,000 Da. Domain analysis revealed that a high proportion (67%) of these proteins contained coiled-coil domains, which is a characteristic feature of centrosomal or centriolar proteins. Other recognizable domains or motifs were present amongst the proteins from both lists, such as WD40 repeats, EF-hand, FYVE domain or the PISA domain of SAS-6 proteins, some

of which were already described as conserved domains of centriolar and centrosomal components (Andersen et al., 2003; Hamel et al., 2017; Keller et al., 2009; Leidel et al., 2005; Li et al., 2004).

All 47 proteins were annotated for homologous proteins by Basic Local Alignment Search Tool (BLAST) searches against non-redundant protein databases, without species restriction, as well as with the restriction mainly against another parabasalid *Trichomonas vaginalis* (Aurrecoechea et al., 2009), which is the closest species with available, even if not very well annotated, proteome. Based on BLAST analysis, proteins were classified into seven broad categories: homologs of centriolar or centrosomal proteins (1 - Centrosome), proteins present in flagella or in the transition zone (2 – Flagellum), proteins involved in metabolic processes (3 – Metabolism), proteins with known function involved in a variety processes (4 – Other), proteins identified as homologs of non annotated hypothetical proteins of *T. vaginalis* or other protists (5 – Hypothetical), and proteins with no homology found (6 – Unknown) (Fig. 3.9). All proteins from the main group in categories 5 and 6 were labeled as Rostrum Of *T. agilis* (ROT) proteins. The proteins from the separate list in these categories were labeled as probable ROTs (pROT), and all of them were also numbered by their molecular weight (e.g. ROT64).
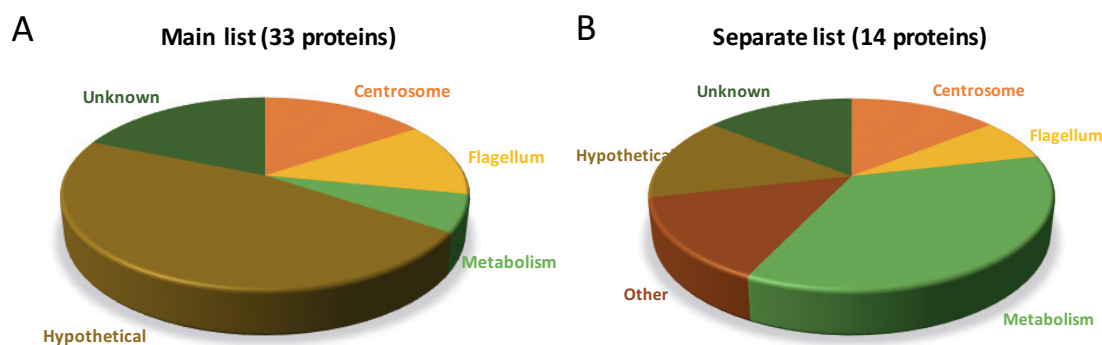


Fig. 3. 9 Proteins of *T. agilis* rostrum **A-B** Based on BLAST analysis, proteins were classified into seven broad categories: (1) Centrosome, proteins known to associate with a centriole and/or centrosome; (2) Flagellum, proteins known to associate with cilium or flagellum; (3) Metabolism, proteins involved in a variety of processes; (4) Other, proteins involved in other processes, not classified in one of previous categories; (5) Hypothetical proteins (Hypothetical), proteins with existing hypothetical homologs in *T. vaginalis*, but not annotated and with unknown function; (6) Unknown, proteins with no homolog found. A Proteins classified in the main list, B Proteins from the separate list.

Table 3. 3 New components of centriole and rostrum of *T. agilis* found in the proteomic analysis. Proteins with identified homology in other species were annotated as their homologs. Homologs of centrosomal proteins are labelled in red. All proteins in the main list without clear homology or with unannotated predicted homolog were named R̲ostrum O̲f *T. agilis* (ROT) or probable ROT (pROT) if they are present in the separate list. Names, molecular weights, fractionation profile based on the EP value and accession number in the single cell transcriptomic database is defined. Color scale: dark red – high EP value, light pink – EP value 0.

**Main list (33 proteins)**

| Protein name | | MW | A | B | C | D | E | F | Accession number in the database |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | EP = Total spectrum count * Percent coverage * 100 | | | | |
| ROT56 | | 56 kDa | 0.0 | 0.0 | 1223.1 | 149.6 | 0.0 | 0.0 | Ta5_comp132978_c0_seq1\|m.12448 |
| FAZ-like | | 32 kDa | 0.0 | 0.0 | 487.2 | 118.0 | 0.0 | 0.0 | Ta6_comp7948_c0_seq1\|m.261 |
| ROT113 | | 113 kDa | 0.0 | 0.0 | 403.2 | 11.0 | 0.0 | 0.0 | Ta1_comp7357_c0_seq1\|m.9686 |
| ROT22 | | 22 kDa | 0.0 | 0.0 | 233.1 | 0.0 | 0.0 | 0.0 | Ta6_comp135577_c0_seq1\|m.299 |
| ROT24 | | 24 kDa | 0.0 | 0.0 | 206.5 | 19.0 | 0.0 | 0.0 | Ta5_comp125696_c0_seq1\|m.2919 |
| TaSAS-6_2 | | 64 kDa | 0.0 | 0.0 | 161.1 | 72.6 | 0.0 | 0.0 | Ta1_comp11369_c0_seq1\|m.8481 (+2) |
| Centrin2_1 | | 21 kDa | 0.0 | 0.0 | 158.5 | 59.1 | 0.0 | 0.0 | Ta1_comp5142_c0_seq1\|m.10195 (+2) |
| ROT82 | | 82 kDa | 0.0 | 0.0 | 136.8 | 16.2 | 0.0 | 0.0 | Ta5_comp7022_c0_seq1\|m.4793 |
| ROT45 | | 45 kDa | 0.0 | 0.0 | 80.5 | 9.2 | 0.0 | 0.0 | Ta1_comp9307_c0_seq1\|m.8956 |
| Centrin2_2 | | 19 kDa | 0.0 | 0.0 | 71.4 | 10.1 | 0.0 | 0.0 | Ta1_comp114472_c0_seq1\|m.1941 (+3) |
| ROT20 | | 20 kDa | 0.0 | 0.0 | 66.6 | 0.0 | 0.0 | 0.0 | Ta6_comp410502_c0_seq1\|m.23950 |
| ROT43 | | 43 kDa | 0.0 | 0.0 | 52.8 | 3.2 | 0.0 | 0.0 | Ta1_comp8169_c0_seq1\|m.12722 |
| ROT12 | | 12 kDa | 0.0 | 0.0 | 45.6 | 12.9 | 0.0 | 0.0 | Ta5_comp2726_c0_seq1\|m.12672 |
| ROT82 | | 82 kDa | 0.0 | 0.0 | 41.5 | 0.0 | 0.0 | 0.0 | Ta6_comp10626_c0_seq1\|m.18684 |
| Dynein light chain, Tctex-type 1 | | 12 kDa | 0.0 | 0.0 | 40.0 | 0.0 | 0.0 | 0.0 | Ta1_comp217365_c0_seq1\|m.12748 (+1) |
| Myosin heavy chain | | 32 kDa | 0.0 | 0.0 | 36.3 | 0.0 | 0.0 | 0.0 | Ta6_comp6140_c0_seq1\|m.14720 |
| ROT135 | | 135 kDa | 0.0 | 0.0 | 34.4 | 9.2 | 0.0 | 0.0 | Ta1_comp11999_c0_seq3\|m.5612 |
| ROT18 | | 18 kDa | 0.0 | 0.0 | 31.8 | 0.0 | 0.0 | 0.0 | Ta1_comp1487_c0_seq1\|m.6748 |
| ROT43b | | 43 kDa | 0.0 | 0.0 | 30.6 | 0.0 | 0.0 | 0.0 | Ta5_comp7472_c1_seq1\|m.3131 |
| ROT27 | | 27 kDa | 0.0 | 0.0 | 28.8 | 0.0 | 0.0 | 0.0 | Ta6_comp185710_c0_seq1\|m.3219 |
| ROT48 | | 48 kDa | 0.0 | 0.0 | 23.2 | 0.0 | 0.0 | 0.0 | Ta5_comp4302_c0_seq3\|m.9644 |
| ROT22b | | 22 kDa | 0.0 | 0.0 | 22.6 | 0.0 | 0.0 | 0.0 | Ta5_comp213707_c0_seq1\|m.9600 |
| Variable flagella number protein | | 38 kDa | 0.0 | 0.0 | 18.6 | 0.0 | 0.0 | 0.0 | Ta6_comp8011_c0_seq1\|m.22688 |
| ROT27b | | 27 kDa | 0.0 | 0.0 | 17.1 | 0.0 | 0.0 | 0.0 | Ta1_comp233286_c0_seq1\|m.24163 |
| ROT36 | | 36 kDa | 0.0 | 0.0 | 16.8 | 5.8 | 0.0 | 0.0 | Ta1_comp9230_c0_seq1\|m.4372 |
| ROT41 | | 41 kDa | 0.0 | 0.0 | 15.3 | 15.3 | 0.0 | 0.0 | Ta5_comp5738_c0_seq1\|m.1819 (+1) |
| POC1 | | 46 kDa | 0.0 | 0.0 | 13.0 | 1.9 | 0.0 | 0.0 | Ta1_comp12628_c0_seq1\|m.23561 (+3) |
| TaSAS-6_3 | | 64 kDa | 0.0 | 0.0 | 9.2 | 7.7 | 0.0 | 0.0 | Ta1_comp11078_c0_seq1\|m.11687 (+2) |
| ROT79 | | 79 kDa | 0.0 | 0.0 | 7.1 | 0.0 | 0.0 | 0.0 | Ta1_comp81220_c0_seq1\|m.2938 (+2) |
| GAP | | 59 kDa | 0.0 | 0.0 | 6.7 | 0.0 | 0.0 | 0.0 | Ta1_comp12464_c0_seq1\|m.5892 (+3) |
| ROT83 | | 83 kDa | 0.0 | 0.0 | 6.2 | 0.0 | 0.0 | 0.0 | Ta1_comp12237_c0_seq3\|m.6477 (+2) |
| CFAP | | 93 kDa | 0.0 | 0.0 | 5.1 | 0.0 | 0.0 | 0.0 | Ta1_comp3574_c0_seq1\|m.22419 |
| ROT84 (Cep135/Bld10p) | | 84 kDa | 0.0 | 0.0 | 5.0 | 3.2 | 0.0 | 0.0 | Ta1_comp10240_c0_seq1\|m.14753 (+1) |

**Separate list (14 proteins)**

| Protein name | | MW | A | B | C | D | E | F | Accession number in the database |
|---|---|---|---|---|---|---|---|---|---|
| actin-related P2/3C subunit 3-like | | 21 kDa | 5.8 | 5.8 | 60.0 | 0.0 | 0.0 | 0.0 | Ta1_comp2719_c0_seq1\|m.20944 (+1) |
| pROT35 | | 35 kDa | 2.9 | 2.9 | 12.8 | 0.0 | 0.0 | 0.0 | Ta1_comp846_c0_seq1\|m.5554 |
| TaSAS-6_1 | | 63 kDa | 0.0 | 2.1 | 102.9 | 9.4 | 2.3 | 8.7 | Ta1_comp9396_c0_seq2\|m.17030 |
| TBC1 domain family member | | 57 kDa | 0.0 | 0.0 | 8.2 | 3.1 | 0.0 | 4.1 | Ta1_comp8737_c0_seq1\|m.17640 |
| 60S acidic ribosomal protein P0 | | 33 kDa | 0.0 | 0.0 | 4.3 | 3.7 | 3.7 | 3.7 | Ta5_comp8550_c0_seq1\|m.7867 |
| CFAP 161 | | 30 kDa | 0.0 | 5.3 | 24.4 | 5.7 | 5.7 | 0.0 | Ta1_comp2882_c0_seq1\|m.8131 (+1) |
| pROT26 | | 26 kDa | 0.0 | 0.0 | 23.2 | 6.3 | 2.7 | 0.0 | Ta1_comp246511_c0_seq1\|m.561 |
| pROT64 | | 64 kDa | 0.0 | 8.7 | 1242.0 | 290.4 | 0.0 | 0.0 | Ta6_comp7948_c0_seq2\|m.19694 |
| SLR-like protein | | 42 kDa | 0.0 | 6.3 | 72.0 | 42.4 | 0.0 | 0.0 | Ta1_comp4902_c0_seq1\|m.5386 |
| peptidyl-prolyl cis-trans isomerase | | 20 kDa | 0.0 | 4.4 | 8.3 | 0.0 | 0.0 | 0.0 | Ta5_comp4697_c0_seq1\|m.13660 |
| pROT45 | | 45 kDa | 0.0 | 2.5 | 25.0 | 22.8 | 0.0 | 0.0 | Ta1_comp11404_c0_seq1\|m.9951 (+5) |
| TPR domain protein | | 44 kDa | 0.0 | 2.1 | 3.1 | 0.0 | 0.0 | 0.0 | Ta1_comp12595_c0_seq1\|m.3863 (+4) |
| pROT83 (Cep135/Bld10p) | | 83 kDa | 0.0 | 1.8 | 6.1 | 1.5 | 0.0 | 0.0 | Ta6_comp11857_c0_seq1\|m.19826 |
| peroxidase | | 96 kDa | 0.0 | 1.6 | 13.0 | 0.9 | 0.0 | 0.0 | Ta6_comp11683_c0_seq2\|m.24262 (+3) |

Legend: centrosome | flagellum | metabolism | other | hypothetical | unknown

### 3.2.5 Proteome of *T. agilis* rostrum – experiment 2

As specified above, we performed a second protein correlation profiling experiment of purified *T. agilis* rostra. We proceeded in a similar way to prepare samples for subsequent proteomic analysis; however, we then pooled qualitatively similar peak fractions and neighboring side fractions in order

to increase the final sample concentration in individual samples. Therefore, the final gradient consisted of only three fractions; A', B' and C', with B' as the peak rostral fraction and A' and C' as side fractions. For this experiment, I dissected ~ 10x more termites and obtained ~ 10x more rostra in the gradient (Table 3.2). Each of three final fractions was split in two to analyse each fraction twice in order to have a duplicate of this experiment, giving rise to fractions with ~ 5x larger protein amount.

Table 3. 4 *T. agilis* centriole and rostrum proteins found in experiment 2. Proteins with identified homology in other species were annotated as their homologs. Homologs of centrosomal proteins are labelled in red. Proteins identified in experiment 1 are labelled with a bullet. Note that few centriolar proteins labelled in red were also identified in experiment 1. All proteins without clear homology or with unannotated predicted homolog were named probable ROT (pROT) regardless their presence in the main or separate list. Names, molecular weights, fractionation profile based on the EP value from both gradients and accession number in the single cell transcriptomic database is defined. A gradient that determined the presence of a protein in the hit list is labelled in shadows of pink and red; the EP values from the second non-determining gradient are depicted in grey. Color scale: dark red – high EP value, light pink – EP value 0.

**Main list (21 proteins)**

| Protein name and group | MW | | Gradient 1 A' | Gradient 1 B' | Gradient 1 C' | Gradient 2 A' | Gradient 2 B' | Gradient 2 C' | Accession number in the database |
|---|---|---|---|---|---|---|---|---|---|
| •ROT27 | 27 kDa | | 0 | 47.1 | 0 | 56.4 | 61.5 | 0 | Ta6_comp185710_c0_seq1\|m.3219 |
| •Myosin heavy chain | 32 kDa | | 0 | 37.5 | 0 | 0 | 15.38 | 0 | Ta6_comp6140_c0_seq1\|m.14720 |
| pROT20 | 20 kDa | | 0 | 26 | 0 | 0 | 7.57 | 0 | Ta5_comp79342_c0_seq1\|m.17434 |
| •ROT82 | 82 kDa | | 0 | 22.64 | 0 | 2.35 | 58.02 | 2.35 | Ta6_comp10626_c0_seq1\|m.18684 |
| pROT30 | 30 kDa | | 0 | 21.6 | 0 | 5.02 | 21.6 | 0 | Ta6_comp4718_c0_seq1\|m.9296 |
| •ROT45b | 45 kDa | | 0 | 11.22 | 0 | 0 | 2.81 | 0 | Ta1_comp9307_c0_seq1\|m.8956 |
| pROT48 | 48 kDa | | 0 | 10.1 | 0 | 0 | 2.88 | 0 | Ta6_comp12104_c0_seq4\|m.11519 |
| •ROT48 | 49 kDa | | 0 | 9.9 | 0 | 2.59 | 28.32 | 0 | Ta5_comp4302_c0_seq2\|m.4811 |
| pROT62 | 62 kDa | | 0 | 8.64 | 0 | 2.07 | 2.07 | 0 | Ta1_comp139393_c0_seq1\|m.8515 |
| pROT19 | 19 kDa | | 0 | 5.39 | 0 | 0 | 21.6 | 0 | Ta1_comp268808_c0_seq1\|m.9430 (+1) |
| Metacaspase-like cysteine peptidase | 29 kDa | | 0 | 3.49 | 0 | 24 | 0 | 0 | Ta1_comp7556_c0_seq1\|m.20992 (+2) |
| pROT67 | 67 kDa | | 0 | 2.84 | 0 | 0 | 7.36 | 0 | Ta1_comp3452_c0_seq1\|m.6547 (+1) |
| Restin homolog isoform | 84 kDa | | 0 | 1.68 | 0 | 1.26 | 28.56 | 0 | Ta5_comp6388_c0_seq1\|m.9396 |
| •pROT26 | 26 kDa | | 31.2 | 48.3 | 0 | 0 | 98.4 | 0 | Ta1_comp246511_c0_seq1\|m.561 |
| CAMK family protein kinase | 31 kDa | | 0 | 26.43 | 5.75 | 0 | 29.88 | 0 | Ta1_comp8710_c0_seq1\|m.7841 |
| •Centrin2 | 21 kDa | | 6.56 | 6.56 | 24 | 0 | 21.8 | 0 | Ta5_comp7079_c0_seq2\|m.9670 |
| •ROT27b | 27 kDa | | 5.41 | 0 | 0 | 0 | 20.8 | 0 | Ta1_comp233286_c0_seq1\|m.24163 |
| POC1b | 35 kDa | | 0 | 0 | 0 | 0 | 15.98 | 0 | Ta1_comp8987_c0_seq1\|m.25077 (+1) |
| pROT51 | 51 kDa | | 3.39 | 12.18 | 3.39 | 0 | 12.18 | 0 | Ta6_comp637_c0_seq1\|m.10414 |
| pROT68 | 68 kDa | | 0 | 0 | 0 | 0 | 7.74 | 0 | Ta5_comp72897_c0_seq1\|m.13644 |
| ROT84b (Cep135/Bld10p) | 84 kDa | | 0 | 1.76 | 1.49 | 0 | 6.52 | 0 | Ta1_comp12505_c0_seq3\|m.2751 |

Header note: Total spectrum count * Percent coverage * 100

**Separate list (14 proteins)**

| Protein name and group | MW | | Gradient 1 A' | Gradient 1 B' | Gradient 1 C' | Gradient 2 A' | Gradient 2 B' | Gradient 2 C' | Accession number in the database |
|---|---|---|---|---|---|---|---|---|---|
| •TBC1 domain family member | 57 kDa | | 3.09 | 22.26 | 0 | 3.09 | 3.09 | 0 | Ta1_comp8737_c0_seq1\|m.17640 |
| POC1c | 46 kDa | | 5.07 | 31.8 | 0 | 3.62 | 33.3 | 0 | Ta5_comp9767_c0_seq3\|m.13948 (+1) |
| pROT30 | 30 kDa | | 0 | 21.6 | 0 | 5.02 | 21.6 | 0 | Ta6_comp4718_c0_seq1\|m.9296 |
| •TaSAS-6_3 | 64 kDa | | 0 | 20.97 | 2.57 | 9.2 | 9.56 | 0 | Ta1_comp11078_c0_seq1\|m.11687 (+2) |
| •ROT135 | 135 kDa | | 1.29 | 72 | 0 | 10.8 | 28.3 | 0 | Ta1_comp11999_c0_seq3\|m.5612 |
| •ROT41 | 41 kDa | | 0 | 20.34 | 3.95 | 15.82 | 37.2 | 0 | Ta5_comp5738_c0_seq1\|m.1819 |
| pROT30 | 30 kDa | | 0 | 36.9 | 4.6 | 0 | 0 | 0 | Ta1_comp8019_c0_seq1\|m.16289 |
| pROT59 | 59 kDa | | 2.71 | 34.8 | 0 | 0 | 15 | 15.84 | Ta1_comp6358_c0_seq1\|m.6909 |
| •ROT43 | 43 kDa | | 3.23 | 3.49 | 0 | 12.9 | 19.35 | 0 | Ta1_comp8169_c0_seq1\|m.12722 |
| •TaSAS-6_1 | 63 kDa | | 7.9 | 10.92 | 1.69 | 21.48 | 39.16 | 0 | Ta1_comp9396_c0_seq2\|m.17030 (+1) |
| •ROT82 | 82 kDa | | 0 | 22.64 | 0 | 2.35 | 58.02 | 2.35 | Ta6_comp10626_c0_seq1\|m.18684 |
| •ROT84 (Cep135/Bld10p) | 84 kDa | | 1.65 | 9.62 | 1.65 | 1.65 | 12.39 | 0 | Ta1_comp10240_c0_seq1\|m.14753 (+1) |
| Coiled-coil containing protein 61 | 36 kDa | | 5.05 | 15.78 | 5.05 | 2.84 | 45.6 | 0 | Ta1_comp9020_c0_seq2\|m.3216 |
| •ROT36 | 36 kDa | | 19.42 | 19.42 | 5.83 | 5.83 | 36.9 | 0 | Ta1_comp9230_c0_seq1\|m.4372 |

Legend: centrosome | flagellum | metabolism | other | hypothetical | unknown

Despite the careful processing of all six samples, the final list of identified proteins was shorter than in experiment 1, with only 523 *T. agilis* proteins from both three-fraction gradients. The resulting

list was analysed as in experiment 1, keeping in the main hit list all proteins with > 1 peptide and EP > 0 in B' and EP = 0 in A' and C' fractions. Proteins peaking in the B' fraction and with the average EP < 5 of two side fractions were kept in the separate list. Proteins following the rostral profile only in one of the two gradients were kept in the final list of hits (Table. 3.4). As can be seen, 40% of the proteins that would have been retained in gradient 1 would not have been retained in gradient 2, and reciprocally. Such lack of reproducility is likely due to the low number of peptides detected in either case. Even we found several new centriolar homologs, all proteins with unidentified homology were called pROT regardless of their placement in the main or separate hit list. Several centriolar proteins that had been identified in experiment 1, such as TaSAS-6_2, ROT83 (Cep135/Bld10p) or one of Centrin2s, were not identified in this analysis. However, some new centriolar homologs were found in experiment 2, such as the third Cep135/Bld10p homolog, called ROT84b, or two other homologs of POC1. The presence of several paralogs of a protein could reflect not only a trend for duplication of centriolar proteins in this species harboring thousands of centrioles, but also a possible specialization of centrioles across the *T. agilis* rostrum. The precise function and localization of individual paralogs remain to be determined.

## 3.3    Discussion

### 3.3.1   Components of the *T. agilis* rostrum

Even though we analyzed the whole rostrum and not individual centrioles, the proteomic study of *T. agilis* rostrum provided a first insight into the probable composition of the elongated proximal region of *T. agilis* centriole. This might be of general significance in uncovering the structural proteins involved in the biogenesis of the centriole and establishing the conserved nine-fold symmetry.

Several homologs of centriolar proteins were identified (Table 3.3 and Table 3.4), including the previously uncovered Centrin2s and TaSAS-6. Moreover, three *T. agilis* Cep135/Bld10p homologs were identified based on sequence similarity in the conserved first 200 residues of the N-terminal domain and the presence of an elongated coiled-coil region, one of which in experiment 2. These proteins are smaller than Cep135 or Bld10p, of only ~ 83/84 kDa, and are therefore tentatively referred to as ROT83, ROT84 and ROT84b, respectively.

Several predicted flagellar proteins were uncovered in this analysis. Considering that we should have subtracted pure flagellar components from the hit list, this might indicate either that these components are enriched close to the rostrum, around or in the transition zone that is probably still present, or else the presence of flagellar proteins in the centriole.

As shown in Fig. 3.9, the identified proteins could be divided into several categories, including potential homologs of proteins involved in metabolic processes, which might have been co-separated with rostrum during the density gradient centrifugation. These proteins might have also an additional function inside the centriolar apparatus and potentially we can expect some other unidentified proteins, classified now as ROTs, to be involved in different non-centriolar functions and be present as contaminants.

Approximately two thirds of the 47 proteins either have no clear homolog in other organisms, or are homologs of components annotated as hypothetical proteins. This might have several explications; i) evolutionary distance, as *T. agilis* is distant from any other organisms with well annotated and available databases (Fig. 3.10), ii) poor annotation of the huge *Trichomonas vaginalis* proteome, which is the closest species with available genomic and proteomic data (Aurrecoechea et al., 2009; Carlton et al., 2007) and iii) truly novel proteins found only in the *Trichonympha* elongated centriole. The latter class of proteins might be important to form and stabilize the elongated cartwheel or the rostrum structure, and thus be absent in other species that have short or transient cartwheels.
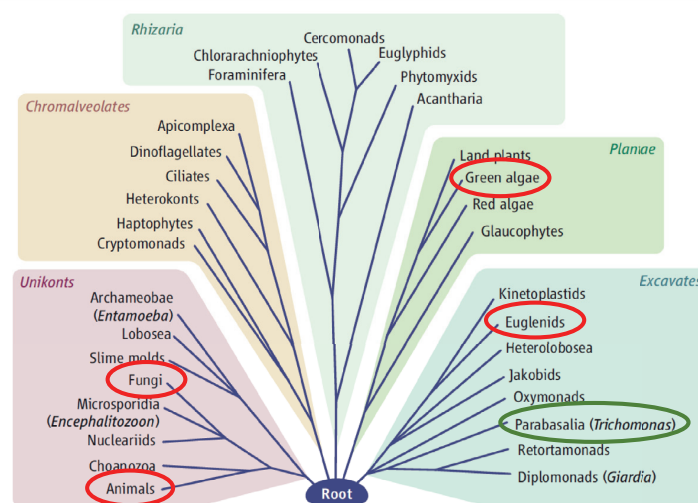


Fig. 3. 10 Hypothetical evolutionary tree of eukaryotes. *T. agilis* as well as *Trichomonas vaginalis* belong to the class *Parabasalia*, which is evolutionary distant from other species with a well annotated proteome, for example human, *Mus musculus*, *C. elegans* (Animalia), *Saccharomyces cerevisiae* (Fungi), *Chlamydomonas* (Green algae) or *Leishmania major* (Euglenida). Note that the branching order is unresolved, adapted from Keeling (2007).

TaSAS-6, a genuine centriolar protein, was assigned to the separate list due to the presence of peptides in the side fractions A and F. However, overall, this protein follows the rostrum fractionation profile (see Fig. 3.7). Together with one Cep135/Bld10p homolog, the placement of TaSAS-6 in the separate list was an indication to manually analyze all the proteins peaking in fraction C.

Interestingly, we have identified two new TaSAS-6 proteins amongst the 47 retained overall, based on their secondary structure prediction, presence of the PISA domain and analogous molecular weight. Therefore, we decided to analyze further these two proteins, hereafter referred to as TaSAS-6_2 and TaSAS-6_3.

### 3.3.2 *T. agilis* homologs of SAS-6

#### 3.3.2.1 TaSAS-6_2

Probably due to the incompleteness of the *T. agilis* genome sequence or of its subsequent assembly, only one TaSAS-6, called now TaSAS-6_1, was identified initially using tBLASTn search with the *T. vaginalis* SAS-6 sequence (Guichard et al., 2013). In this study, we identified TaSAS-6_2 as a paralog of TaSAS-6_1. TaSAS-6_2 followed a similar peptide number fractionation profile as TaSAS-6_1 and passed stringent tests regarding the total spectral count enrichment in the peak C fraction (number of spectra in fraction C > fraction D), as well as side fractions (0). Sequence-wise, TaSAS-6_2 shares 25% identity and 46% similarity with TaSAS-6_1. The sequence identity and similarity is uniformly spread throughout the sequence (Fig. 3.11 A). Coiled-coiled prediction of the whole protein and structural prediction of ~ 250 residues from the N-terminus confirm the structural conservation of TaSAS-6_2 with its paralog TaSAS-6_1, as well as with SAS-6 proteins from other species (Fig. 3.11 B-D).

The polyclonal antibody previously raised against TaSAS-6_1 stained the entire length of the proximal region of the *T. agilis* centriole (Fig. 3.11 E, Guichard et al., 2013) and is present in all centrioles stained by Centrin2 (personal observation). To analyse TaSAS-6_1 and TaSAS-6_2 in more detail, and determine their precise localization along or in the *T. agilis* centriole, we designed and produced specific anti-peptides antibodies raised against the most divergent regions of both paralogs in the N-terminal region (Fig. 3.11 A) (Eurogentec).
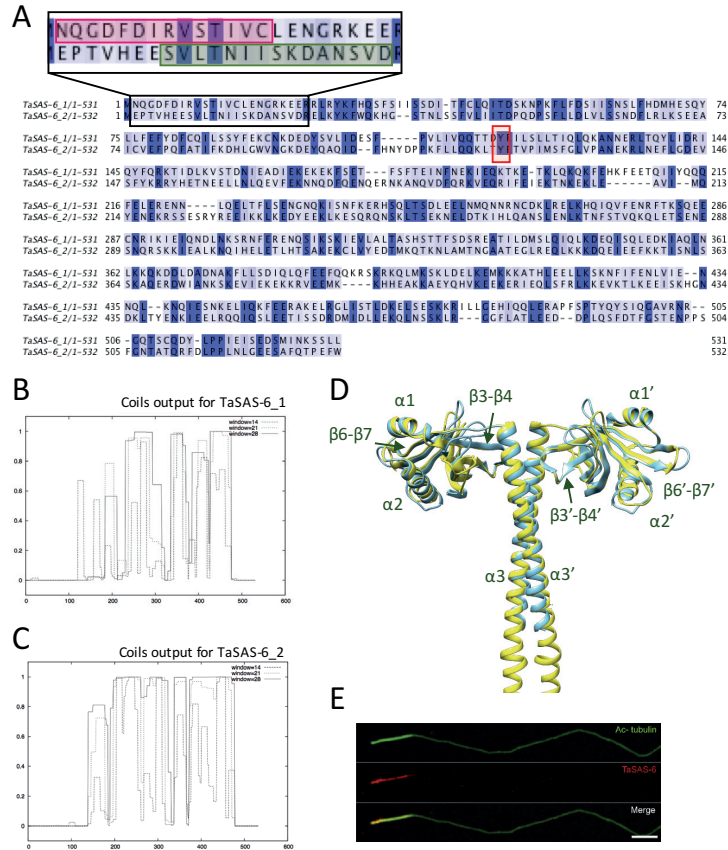
Fig. 3. 11 TaSAS-6_2. **A** Whole sequence alignment of TaSAS-6_1 and TaSAS-6_2 showing substantial identity (25%) and similarity (46%) along the entire sequence. The alignment was done in Jalview2.10 using the Mafft alignment, coloured by % identity colour scheme. Note the presence of an unusual Tyr120 (Tyr118 in TaSAS-6_1) as the hydrophobic residue in the position that normally mediates higher order oligomerization of SAS-6 proteins. Black box magnification depicts the region used for the production of specific anti-peptide antibodies, rabbit anti-TaSAS-6_1 (in red) and rat anti-TaSAS-6_2 (in green). **B-C** Coiled-coil prediction for TaSAS-6_1 (B) and TaSAS-6_2 (C). **D** Superimposition of structural predictions of ~ 250 amino acids from the N-terminus of both paralogs. Structural predictions were made in Phyre[2] (Kelley et al., 2015). **E** The antibody raised against TaSAS-6_1 stains the whole length of the *T. agilis* centriolar proximal region. With permission from Guichard et al. (2013). Scale bar: 2 μm.

The anti-peptide antibodies purified from the sera were tested by western blot analysis to assay their specificity (Fig. 3.12 A), which was confirmed, and then used in indirect immunofluorescence analysis of purified *T. agilis* centrioles. We found that TaSAS-6_1 and TaSAS-6_2 colocalize throughout the length of the *T. agilis* centriolar proximal region (Fig. 3.12 B). This might be explained either by their possible hetero-oligomerization inside a cartwheel ring or else the alternation of rings containing solely TaSAS-6_1 or solely TaSAS-6_2 inside the long cartwheel.
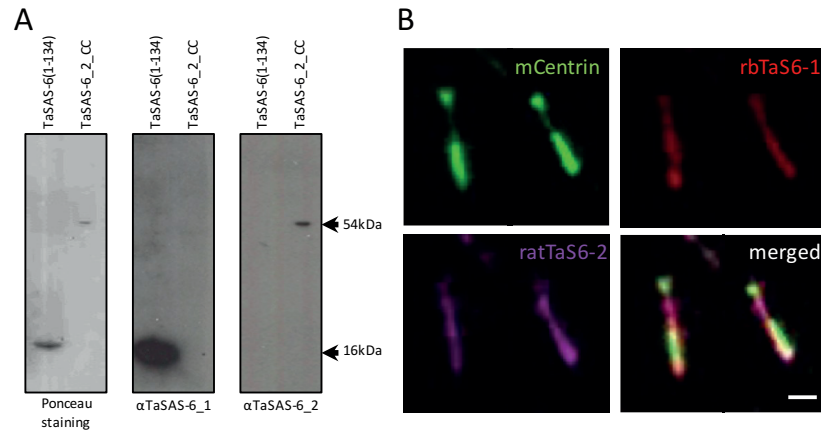
Fig. 3. 12 Anti-TaSAS-6_1 and anti-TaSAS-6_2 specific anti-peptide antibodies. **A** Western blot of purified TaSAS-6_1 N-ter with $M_W$ of ~ 16 kDa and TaSAS-6_2_CC (with the whole coiled-coil region) with $M_W$ of ~ 54 kDa. The antibodies do not cross-react. **B** Indirect immunofluorescence staining of individual centriole from *T. agilis* lysate stained with Centrin2 antibody (in green) to distinguish centrioles in the lysate and TaSAS-6_1 (in red) and TaSAS-6_2 (in purple). Note the colocalization of both paralogs of TaSAS-6 along the cartwheel bearing portion of centrioles. Proximal side of centrioles is at the bottom.

### 3.3.2.2 TaSAS-6_3
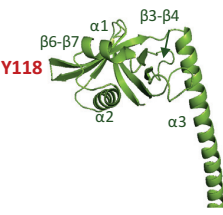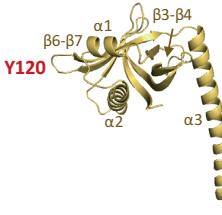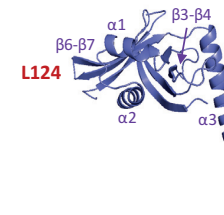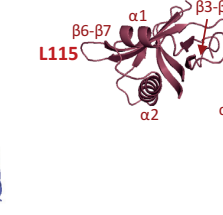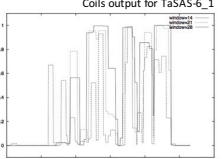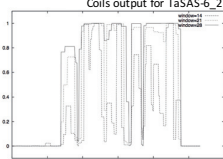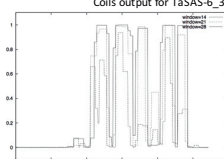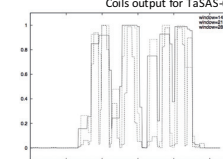
A third paralog of TaSAS-6 was identified as a protein only with two peptides in the peak fraction C and with the same peptide number in the neighboring D fraction. However, based on the absence of its peptides in all four side fractions A, B, E and F, TaSAS-6_3 was nevertheless identified as a hit from the main list (Table 3.3). The sequence and structural analysis confirmed its homology to other SAS-6 (Table 3.4). However, the only obvious difference of TaSAS-6_3 in comparison with other two paralogs is that this protein harbors another unusual hydrophobic residue, Leu124, in the position that normally mediates higher order oligomerization of SAS-6 proteins.

### 3.3.2.3 TaSAS-6_4

The presence of the second and third TaSAS-6 in the rostral proteome led me to analyse the whole redundant transcriptomic database to search for possible additional paralogs of TaSAS-6. Therefore, I used a BLAST search of all three known TaSAS-6 sequences against the *T. agilis* transcriptomic database using Geneious software. In doing so, I found a short sequence of a protein very similar to the three other TaSAS-6s. To find out the length and predicted structure of the corresponding entire protein, this short amino acid sequence of TaSAS-6_4 was Blasted against the whole *T. agilis* genome assembly. The rest of the sequence was found very easily in this manner, because even if the genome of *T. agilis* is gigantic, introns inside genes are rare (Y. Hongoh, personal communication). I analysed also TaSAS-6_4, and found that it shares the length, ~ 25% of overall sequence identity and structural homology with the three other TaSAS-6 paralogs (Table 3.4). However, TaSAS-6_4

was not identified at all in the proteomic analysis and, therefore, we can only speculate if this fourth TaSAS-6 is present at centrioles. Note also that the incomplete sequence of TaSAS-6_4 indicates that the transcriptomic database is incomplete.

Table 3. 5 Comparison of four TaSAS-6 paralogs. All four proteins are similar in several aspects, including their length, molecular weight, structural prediction of the conserved N-terminal domain and coiled-coil. Note the presence of Tyrosine in TaSAS-6_1 and TaSAS-6_2, and of Leucine in the other two paralogs at the position that normally mediates higher order oligomerization of SAS-6 proteins.



| | TaSAS-6_1 | TaSAS-6_2 | TaSAS-6_3 | TaSAS-6_4 |
|---|---|---|---|---|
| protein length | 531 residues | 532 residues | 544 residues | 529 residues |
| $M_W$ | 63.1 kDa | 62.6 kDa | 64.4 kDa | 63.1 kDa |

### 3.3.3 Proteome of *T. agilis* centriole

Here, I presented the first approach to assess the composition of the centriolar proximal region using the elongated *T. agilis* centrioles with an extraordinarily long cartwheel bearing region. Difficulties encountered during optimization of *T. agilis* centriole purification led me to establish a protocol for rostrum purification (Appendix 1), therefore establishing the first proteome of *T. agilis* rostrum. The total number of flagella, and thus of centrioles, in a *T. agilis* cell is estimated to be ~ 5.000 (Y. Hongoh, personal communication), whilst a rostrum harbors ~ 1400 of them (Fig. 3.5). Furthermore, the rostrum is a compact and resistant structure differing in shape, size and density from other cellular components, making it an appropriate structure to purify. Because we did not achieve rostrum purification to homogeneity, we performed mass spectrometry based protein correlation profiling, which was proven to be successful previously for the analysis of centrosomal components from human cells (Andersen et al., 2003; Jakobsen et al., 2013).

The number of identified proteins in all gradient fractions was in the same range in the two experiments, even if we analysed a ~ 5x larger sample the second time (Table 3.2). This might be due to a failure in scaling up the purification and incompleteness of OptiPrep removal, which might hinder peptide detection (personal observation, data not shown). However, more detailed analysis of experiment 2 results confirms the elution and identification of ~ 37% of rostral proteins identified in ex-

periment 1 (Fig. 3.13), including two TaSAS-6s or one homolog of Cep135/Bld10p. Moreover, we identified the third homolog of Cep135/Bld10p, named ROT84b and two other homologs of POC1 protein. The presence of several paralogs of a protein would indicate not only the trend of the centriole protein duplication in this species with a complex of thousands of centrioles around the cell, but also a possible specialization of centrioles with different functions in different location in the *T. agilis* cell. The precise function and localization of individual paralogs remain to be determined.
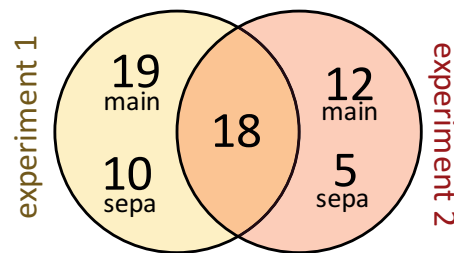


Fig. 3. 13 Comparison of numbers of identified T. agilis rostrum proteins from two independent experiments. Number of proteins identified in the main (main) or separate (sepa) list is depicted. Because some proteins from the main list in experiment 1 were identified as the separate list proteins in experiment 2, and vice versa, no distinction is made for overlapping proteins.

Even though more than 30.000-40.000 spectra were detected in total in both experiments, only 2-5% of them were identified. This might be due to the incompleteness of the transcriptomic database and the genome assembly. We could observe such incompleteness in the case of TaSAS-6_4, which was not present as a full sequence in the database. Furthermore, the available single cell transcriptomic database most certainly does not cover all stages of the *T. agilis* life cycle, and some mRNA sequences of centriolar proteins, might be absent from the single-cell transcriptomic database, because they are present only when the whole centriolar complex is formed. The stage of the *T. agilis* cell cycle at which the transcriptome was studied was not determined. Moreover, *T. agilis* lives in symbiosis with its bacterial endosymbionts (Ikeda-Ohtsubo and Brune, 2009), so it is possible that some fraction of the detected spectra were from peptides of *T. agilis* internal bacterial symbionts, which might be tested by including small endosymbionts proteomes or translated genome of some sequenced species into our analyses.

Therefore, to establish a better identification success rate, we would need to work with a full *T. agilis* proteome, which is not existent for now, and to improve the purification method before and also after cell lysis to discard all possible contaminants from other species.

Furthermore, the availability of centriolar antibodies and the large size of the rostrum (7 μm x 7 μm x 11 μm) suggests that the *T. agilis* rostrum could be isolated from the cell lysate by flow cytometry (Gauthier et al., 2008), which is likely to improve the purity of the final sample.

Of the all proteins identified in the rostral proteome from both experiments, only few annotated centriolar proteins were identified that have been extensively studied in different model organisms. This includes three homologs of POC1 (Keller et al., 2009), ROT83, ROT84 and ROT84b as the potential homologs of Cep135/Bld10p (Bayless et al., 2012; Kraatz et al., 2016; Lin et al., 2013a; Roque et al., 2012), as well as three TaSAS-6 paralogues (van Breugel et al., 2011, 2014; Dammermann et al., 2004; Hilbert et al., 2013; Kitagawa et al., 2009, 2011; Leidel et al., 2005; Nakazawa et al., 2007). Altogether, we identified only 11 centriolar protein homologs. This might seem in contradiction with the fact that the centriole is an evolutionary conserved organelle. However, this might be due to the evolutionary distance with other species, such that only some proteins with important functions (e.g. establishing nine-fold symmetry, connecting the cartwheel with microtubule triplets and inter-triplet connections) are well conserved. Another reason may be our focus on the proximal region. Proteins from only cartwheel bearing portion of the centriole, which are more likely to be overlooked when analysing the whole centriole with a short cartwheel, are likely to be enriched in our analysis, and might be found in more detailed analyses of centrioles from other species in the future. Furthermore, the presence of the elongated cartwheel-bearing portion also suggests that *T. agilis* centriole possess also other proteins required for cartwheel stability, as for example the ICPs, which might be specific to this elongated centriole.

Indirect immunofluorescence staining was chosen to test the newly discovered TaSAS-6 paralogs: in this manner, TaSAS-6_2 was confirmed to be located along the whole cartwheel together with TaSAS-6_1. There are several hypotheses regarding how the two proteins might be organized inside one long cartwheel as depicted in Fig. 3.14 for two of them. Antibodies for TaSAS-6_3 and TaSAS-6_4 have been already produced and will be tested in the near future.

The presence of myosin and actin-related proteins or of proteins connected to metabolic functions, might indicate that some of the identified proteins have non-centriolar functions or localization. Obviously, all proteins detected in this list are enriched in the rostrum. This may be because they are present for instance in membrane-associated domains, which might be absent from all other fractions, as such material would likely have been pelleted before loading onto the OptiPrep gradient. However, we did not find transmembrane proteins in our list of candidates, nor a clear signature of other specific groups of proteins..

We did not find homologs in other organisms for a large number of rostrum proteins. In the future, one can nevertheless test some of them by indirect immunofluorescence staining to confirm their presence in the rostrum, and to determine their exact localization in individual centrioles, in order to derive functional predictions in the centriole or possibly find candidates of new densities, such as CCD or ICP.
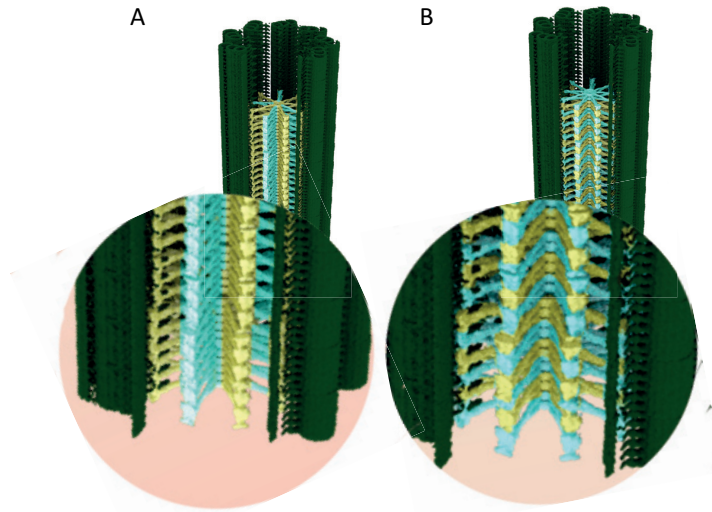
Fig. 3. 14 Two different speculative models regarding how two TaSAS-6s might be placed inside the long cartwheel. **A** Two TaSAS-6 paralogs interact with each other through their N-terminal domain or through coiled-coil domains (not depicted) within a given ring, and thus form rings with heterogeneous TaSAS-6 populations. **B** Intercalation of two types of homogenous rings (not depicted) or double layers each formed by one of the two paralogs of TaSAS-6.

For now, we do not have reliable genomic tools to study protein function in symbiotic protists of termites. Therefore, further characterization of ROTs without available homologs in other species is likely to prove difficult. Nevertheless, the interest in termites as animal pests led to the development of RNA interference via voluntary feeding as a potential pest control method (Zhou et al., 2008), as well as a tool to conduct functional genomic research in such species (reviewed in Huvenne and Smagghe, 2010). This method was further developed and extended to target termite symbionts (Itakura et al., 2009; Liu et al., 2017). Even though this method was not tested on *R. speratus* or other termite hosts of *Trichonympha* spp., and would thus need to be established, this might provide a tool to study the function of individual *T. agilis* proteins *in vivo*.

## 3.4    Personal contributions

I optimized the protocol for the *T. agilis* rostrum purification (see Appendix I) and prepared the samples for the proteomic analysis. Then, I analysed the resulting dataset and performed the bioinformatics analysis of the hit list proteins, including BLAST and PSI-BLAST search, domain prediction and 3D structure prediction for some of them. I analysed more carefully TaSAS-6 paralogs and designed peptides for the specific antibody production. I tested the specificity for the antibodies against TaSAS-6_1 and TaSAS-6_2.

This project was done in collaboration with:

- Yuichi Hongoh from Tokyo Institute of Technology (Tokyo, Japan) who kindly provided *R. speratus* termites, *T. agilis* single cell transcriptomic database and the assembly of the *T. agilis* genome.
- Romain Hamelin and Marc Moniatte from Proteomics Core Facility at EPFL (Lausanne, Switzerland), who performed the mass spectrometry based proteomic analysis of purified *T. agilis* rostra.

# Chapter 4    SAS-6 proteins and cartwheel assembly

*"Do not disturb my circles."*

*"Noli turbare circulos meos."*

Archimedes

## 4.1    Background

Centrioles are composed of hundreds of proteins, as uncovered by several proteomic studies of so-called "naked" centrioles, i.e. centrioles devoid of substantial PCM, of *Chlamydomonas* (Hamel et al., 2017; Keller et al., 2009) or mammalian sperm cells (Firat-Karalar et al., 2014). Several such proteins are conserved across evolution and the SAS-6 protein is one of them (Fig. 4.1 A) (Carvalho-Santos et al., 2010; Hodges et al., 2010).

Just like the conserved canonical structure of centrioles and of the proximally localized cartwheel, SAS-6 homologs share not only the primary structure of the protein, but also the secondary, tertiary and in many species with a canonical cartwheel also the quaternary structure of the protein.

SAS-6 proteins harbor an N-terminal globular domain, a long coiled-coil domain and an unstructured C-terminal region. SAS-6 can dimerize through the interaction of coiled-coil moieties and then oligomerize through the interaction of N-terminal domains. All important details about the sequence, structure, dimerization and oligomerization properties are described below.

### 4.1.1    Primary structure

Sequence-wise, SAS-6 homologs share substantial similarity within a conserved region in the SAS-6 N-terminal domain, called the PISA domain (Present In SAS-6, Fig. 4.2 A). However, the N-terminal domain itself diverged more throughout the evolution, than the rest of the SAS-6 protein sequence (Fig. 4.1 A and B). The PISA domain is a ~ 50 amino acids long sequence characterized by several conserved hydrophic residues, including Ile62 in the human homolog of SAS-6, HsSAS-6 (Fig. 4.2 A), which is suggested to be required for the proper folding and/or function of HsSAS-6 (Khan et al., 2014). A critical residue in the N-terminal part of SAS-6 proteins is a given hydrophobic residue, which is Phenylalanine in the majority of species, such as Phe131 HsSAS-6, or Phe145 in *Chlamydomonas* homolog of SAS-6, called Bld12p or CrSAS-6. This residue mediates the higher order oligomerization of homodimers of SAS-6 proteins, whereby the hydrophobic residue at this

position docks into a hydrophobic pocket of the neighboring N-terminal part. There are some exceptions to having a Phenylalanine at this location, as for example Isoleucine in *C. elegans*, Tyrosine in two TaSAS-6 homologs and Leucine in the other two (Fig. 4.2 B).

A
**Whole SAS-6 sequences**



B
**SAS-6 N-terminal domains**
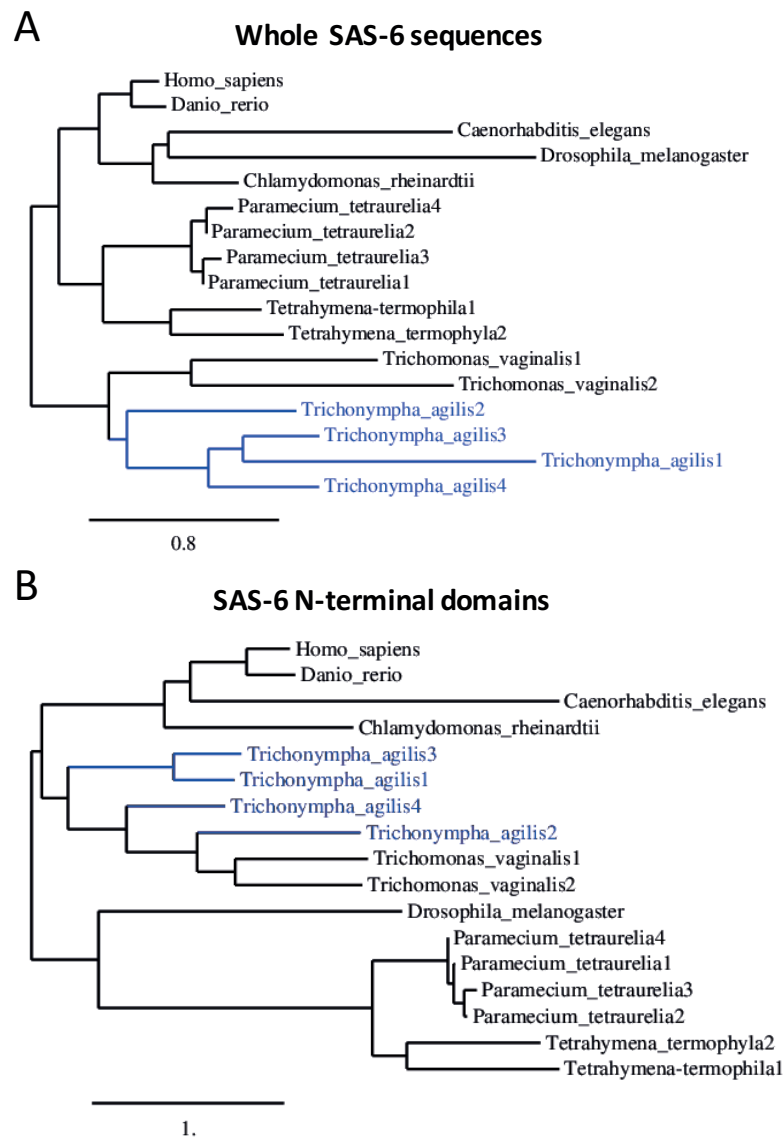


Fig. 4. 1 SAS-6 protein phylogenetic trees. **A-B** Phylogenetic tree of the entire SAS-6 protein sequences (A) and of their N-terminal domains (B). The trees were constructed from MUSCLE alignment (Edgar, 2004), based on PhyML phylogeny (Guindon and Gascuel, 2003), and rendered in TreeDyn (Anisimova and Gascuel, 2006; Chevenet et al., 2006; Dereeper et al., 2010) in Phylogeny.fr software (Dereeper et al., 2008). Gblocks curation was done only for the phylogenetic tree in A (Castresana, 2000). Scales represent the length of branch with an amount of genetic change in nucleotide substitutions per site.
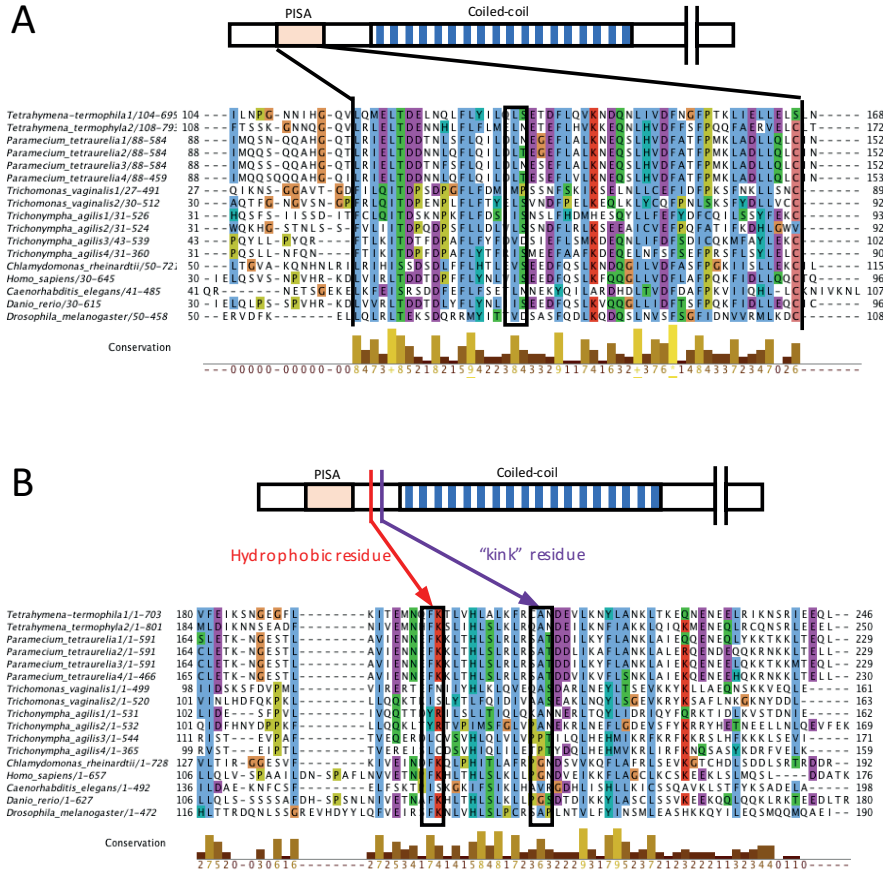
Fig. 4. 2 Conserved residues in N-terminal region of SAS-6. **A** PISA domain of 17 homologs of SAS-6 from nine species. SAS-6 sequences were aligned in Jalview2.10 using the Mafft alignment and coloured by Clustal X colour scheme. Note the conservation of residues. Black rectangle depicts the conservation of the hydrophobic residue at the position Ile62 in HsSAS-6 that is important for function and/or protein folding (Khan et al., 2014). **B** Conservation of amino acids around the hydrophobic residue responsible for N-N dimerization (black rectangle on the left) and the Val167 residue in *C. elegans*, which is at the position of a small amino acid residue in the majority of SAS-6 homolog; Val167 has been suggested to be responsible for a spiral conformation of SAS-6 in nematodes (Hilbert et al., 2013). The alignment of the entire sequences can be found in Appendix 2.

### 4.1.2 Secondary and tertiary structure

Up to now, several crystal structures of SAS-6 protein fragments from different species were solved: *C. elegans* (Kitagawa et al., 2011), *Chlamydomonas* (Kitagawa et al., 2011), zebrafish (van Breugel et al., 2011), *Drosophila* (Cottee et al., 2015), and *Leishmania major* (van Breugel et al., 2014). These crystal structures revealed that SAS-6 proteins have a conserved tertiary structure. Thus, the N-terminal part of the protein comprises ~ 150 amino acids and is formed from a β-barrel, two α - helices, and flexible loops with various lengths (Fig. 4.3 B). Then the protein elongates with a long α-helix and after a short "neck" region, the α-helix dimerizes with an equivalent of another SAS-6 protein, with which it forms a parallel coiled-coil. In most species, the coiled-coil region contains ~ 40

heptad repeats and is thus predicted to be ~ 45 nm long. Past the coiled-coil region, SAS-6 proteins are predicted to be unstructured in the entire C-terminal region, which varies in length and sequence amongst species (Fig. 4.3 A).
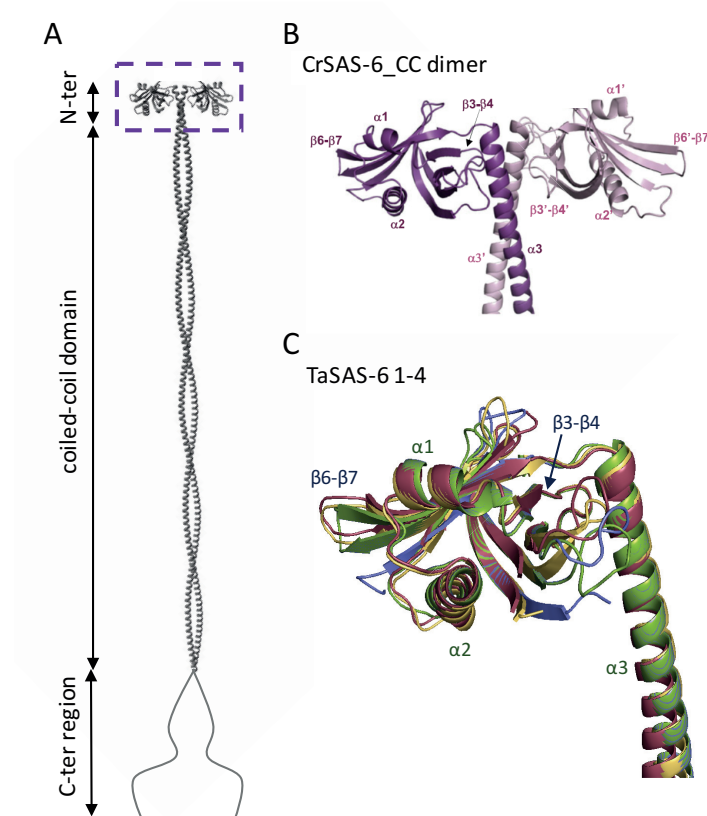


Fig. 4. 3 SAS-6 structure. **A** Structural prediction of the whole HsSAS-6 dimer with the representation of the globular N-terminal domain (purple box), the long coiled-coil domain and the unstructured C-terminal region. **B** Crystal structure of CrSAS-6_CC dimer (Kitagawa et al., 2011). **C** Superimposition of the predicted structures of four TaSAS-6 homologs, TaSAS-6_1 in green, TaSAS-6_2 in yellow, TaSAS-6_3 in purple and TaSAS-6_4 in pink. Structural predictions were made in Phyre[2] (Kelley et al., 2015).

### 4.1.3   Quaternary structure

From what we know from previous *in vitro* studies (van Breugel et al., 2011, 2014; Guichard et al., 2017; Hilbert et al., 2013, 2016; Kitagawa et al., 2011; Pfreundschuh et al., 2014), it is clear that SAS-6 proteins have an intrinsic capacity to form unique ring-like structures with mostly nine-fold radial symmetry. The globular N-terminal parts of neighboring SAS-6 homodimers interact and eventually form the cartwheel hub, from which nine radial spokes composed of the coiled-coil domains of the protein emanate (Fig. 4.4). These spokes are probably connected through the C-terminal region of SAS-6 proteins to the Pinhead bridging the cartwheel and the microtubule triplets (Guichard et al., 2013). Moreover, two juxtaposed SAS-6 rings form a double layer *in situ* (Guichard et al., 2012) as well as *in vitro* (Guichard et al., 2017), probably through an interaction mediated by their C-

terminal halves. This results in a stack with a periodicity of ~8.5 nm and ~17 nm at the cartwheel hub and at the peripheral end of spokes, respectively (Fig. 4.5). Structural information about the 3D conformation of the C-terminal merged parts of SAS-6 and its connection to the Pinhead is not available.
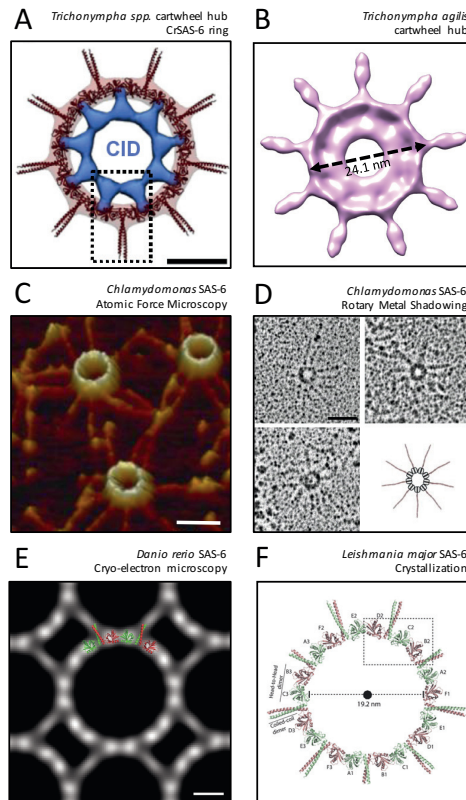


Fig. 4. 4 Intrinsic property of SAS-6 proteins to oligomerize and form ring-like structures. **A** CrSAS-6 ring modelled *in silico* fitted into the 3D density map of *Trichonympha* spp. cartwheel hub (Guichard et al., 2013). Scale bar: 10 nm. **B** 3D architecture of *T. agilis* cartwheel from this study (see Chapter II) **C-D** CrSAS-6 rings formed on Mica surface and observed using atomic force microscopy (Pfreundschuh et al., 2014). Scale bar: 25 nm (C) and analyzed using rotary metal shadowing (Kitagawa et al., 2011). Scale bar: 50 nm (D). **E** Cryo-electron micrograph of thin crystal of DrSAS-6 (van Breugel et al., 2011). Scale bar: 60 Å. Note changed symmetry to eight-fold due to crystal packing. **F** LmSAS-6 crystallized into rings with nine-fold symmetry (van Breugel et al., 2014).
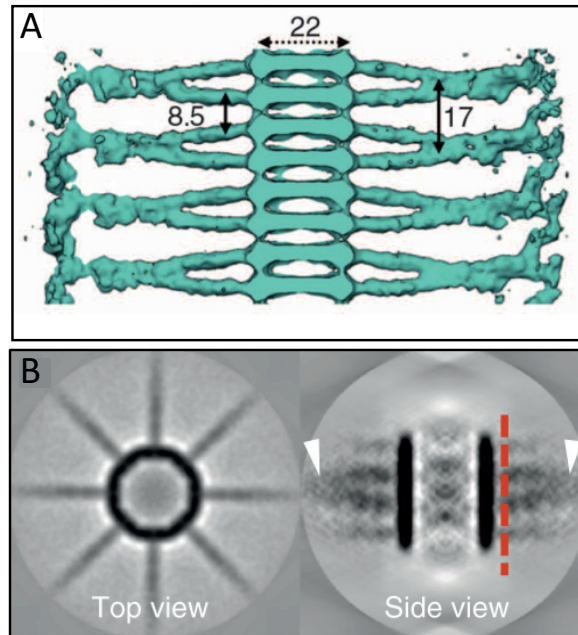
Fig. 4. 5 SAS-6 proteins form double-layers and stack with 8.5 nm periodicity. **A** Cryo-ET map of *Trichonympha* spp. cartwheel (Guichard et al., 2012) showing the internal organization and periodicity of the cartwheel. **B** CrSAS-6 protein can oligomerize into eight-fold or nine-fold (not shown) symmetrical ring and stack in double-layers *in vitro* with ~ 7.5 nm periodicity (Guichard et al., 2017).

## 4.2    Results

As described above, several homologs of SAS-6 proteins, with the whole or part of the coiled-coil domain sufficient for homodimerization, can oligomerize and in ideal cases also form ring-like structures. The aim of the third part of my project was to study the oligomerization properties of TaSAS-6 orthologs and of HsSAS-6 using biochemistry, microscopy and structural methods. HsSAS-6 localizes to the ~100 nm long cartwheel of the human centriole (Guichard et al., 2010; Keller et al., 2014; Lukinavičius et al., 2013b), where it is present from the G1/S transition until metaphase (Kleylein-Sohn et al., 2007; Strnad et al., 2007) and remains absent in G1/G0, including when the centriole docks below the plasma membrane to template axoneme formation. By contrast, SAS-6 proteins from *T. agilis* are structural components of an extremely long centriole, which harbors an elongated and stable cartwheel, including when the centriole is docked into the plasma membrane, possibly to reinforce the centriole so as to sustain external forces from the flagellum.

### 4.2.1    Negative staining TEM for SAS-6 ring visualization

Negative staining for transmission electron microscopy (TEM) is a powerful technique to visualize biological samples, such as different viruses or large protein complexes (reviewe in De Carlo and Harris, 2011). Even with the broad range of available ring formation assays, such as AFM (Hilbert et al., 2016; Pfreundschuh et al., 2014), rotary metal shadowing (Kitagawa et al., 2011) or cryo-EM of

crystals (van Breugel et al., 2011; Guichard et al., 2017), none of them is ideally suited as a screening method due to long sample preparation time. Therefore, I aimed to optimize a protocol for SAS-6 ring formation using negative staining TEM, which is a simple and rapid method requiring little amount of sample, and for which the grid can be stored for a long time. Of course, as with any other technique where a biological sample is finally air-dried, also negative staining TEM can produce some artefacts. However, we felt it was nevertheless a suitable method for initial screening of conditions allowing SAS-6 oligomerization.

During the optimization process of negative staining TEM, I used *Chlamydomonas* SAS-6 without the unstructured C-terminal part (Fig. 4.6 A), which is a CrSAS-6 construct with established *in vitro* ring formation capabilities (Guichard et al., 2017; Hilbert et al., 2016; Kitagawa et al., 2011; Pfreundschuh et al., 2014) to screen for a suitable buffer and salt concentration. Three buffers (Tris pH 7.5, HEPES pH 8.0, K-PIPES pH 7.0) and 3 different salt concentrations (0, 150 and 300 mM) were tested. I found the most suitable buffer to be K-Pipes, pH 7.0, without salt addition, which is used in human or *Chlamydomonas* centriole purification protocols (Bornens and Moudjou, 1999b; Hamel et al., 2017; Jakobsen et al., 2013). Then, I screened for a suitable incubation time of the sample on the grid to provide enough time for the protein to oligomerize, while preventing overcrowding of the grid. Sparse rings were found even with a 2 min incubation time (Fig. 4.6 C) and crowding was observed when incubating longer than 15 min (Fig. 4.6 B). A detailed protocol is described in the Material and Methods section. The irregularities in the ring shape (circular, elliptical, compressed) and in position of spokes (Fig. 4.6 B and C) are probably due to the conformation freedom of individual monomers in the neck region between the N-terminal and coiled-coil domain, as discussed for the TaSAS-6_2 (see below and Discussion), or due to the staining and drying artefacts.
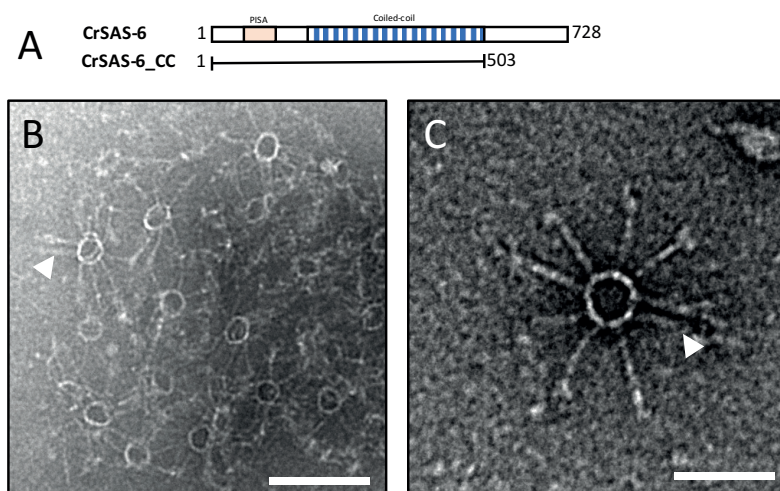


Fig. 4. 6 CrSAS-6 rings visualized by negative staining – TEM **A** Positive control CrSAS-6 without unstructured C-terminal region was used for optimization process, CrSAS-6_CC. **B-C** Electron micrographs of clustered (B) or individual rings (C). White arrowheads point to spokes at an angle smaller than 40º that shows the flexibility of CrSAS-6 protein conformation. Scale bars are 100 nm and 50 nm, for B and C, respectively.

### 4.2.2 TaSAS-6 oligomerization properties

We then set out to analyze the oligomerization properties of two TaSAS-6 homologs: TaSAS-6_1 which was identified previously (Guichard et al., 2013), as well as TaSAS-6_2, which was identified in the proteomic analysis of the *T. agilis* centriole (see Chapter III).

#### 4.2.2.1 TaSAS-6_1

The predicted globular N-terminal domain of TaSAS-6_1 (1-132) could be purified and concentrated only to 2 mg/ml (125 μM), above which the protein started to precipitate. We have tested the dimerization properties of the N-terminal domain by multi-angle light scattering analysis (MALS) and saw that even at the highest possible soluble concentration of 2 mg/ml, the protein was still present as a monomer in solution (Fig. 4.7 B). Moreover, our collaborator from Prof. Michel Steinmetz group, Natacha Olieric, tried to measure the $K_d$ of the N-terminal domain using isothermal titration calorimetry (ITC), but could not detect an interaction. We analyzed also the behavior of a longer construct containing the N-terminal domain plus 6 heptat repeats of the coiled-coil region (TaSAS-6_1 6HR) by negative staining TEM, but could observe only aggregates instead of rings (Fig. 4.7 A and C). Moreover, we tried to optimize a protocol for the purification of TaSAS-6_1 with the whole coiled-coil domain. However, this construct, as well as other constructs with only a portion of the coiled-coil, aggregated into inclusion bodies already during bacterial protein expression and could be extracted in very low quantities only using denaturing methods. Therefore, further optimization of the expression and purification method will be needed in the future to analyze TaSAS-6_1.
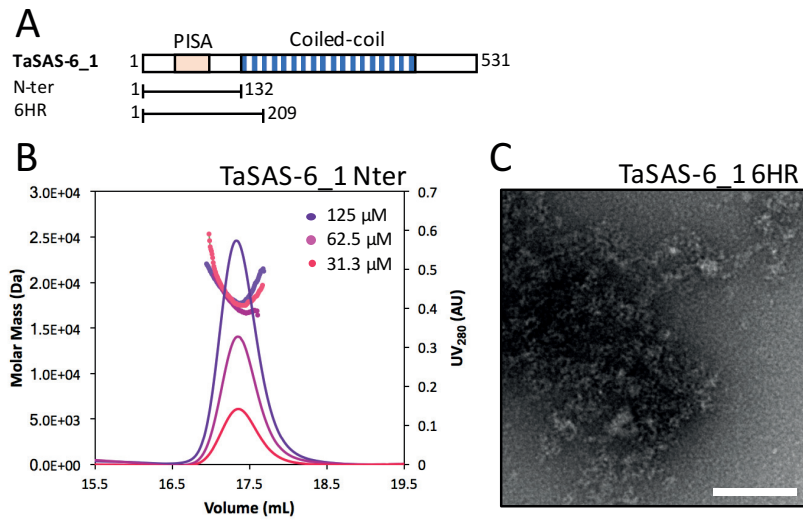


Fig. 4. 7 TaSAS-6_1 is present as a monomer in solution. **A** Two constructs that were used to study TaSAS-6_1, N-terminal domain (N-ter) and N-ter with 6 heptat repeats (HR) of the coiled-coil region. **B** Multi-angle light scattering analysis of TaSAS-6_N-ter at three different concentrations showing the presence of only monomers in solution, with an estimated $M_W \sim 17$ kDa, which corresponds to the molecular weight of a TaSAS-6_1 N-ter monomer. Size-exclusion chromatography elution profiles are shown as absorbance at $UV_{280}$ and overlaid with the estimation of the protein molecular weight assessed by multi-angle light scattering. **C** Example of an aggregate observed by negative staining-TEM of TaSAS-6_1 6HR. Scale bar: 100 nm.

#### 4.2.2.2 TaSAS-6_2

In contrast to TaSAS-6_1, the expression and purification of TaSAS-6_2 was optimized quickly and we obtained proteins in high purity and concentrations. TaSAS-6_2_CC behaved as an oligomer when applied on native PAGE gel and ran as a single band at ~ 300 kDa (Fig. 4.8 A). This might correspond to a higher order-oligomer, but because coiled-coil proteins run in native electrophoresis differently than globular proteins, this might correspond also to a dimer. The protein was also applied on a Mica surface to study its behavior at a high local surface concentration using the AFM. In doing so, we observed that TaSAS-6_2_CC is capable of forming ring-like structures at high surface density on Mica surface, when injection the solution of ~ 50 nM concentration (Fig. 4.8 B and C). However, when we decreased the solution concentration, to ~ 18 nM, with the aim of improving the resolution of the AFM topographs, we found that the protein forms straight oligomers through its N-terminal domain interaction, with spokes (coiled-coil dimers) oriented outwards to both sides (Fig. 4.8 D and E). This unexpected behavior may reflect flexibility of the TaSAS-6_2 neck region, between the globular N-terminal domain and the coiled-coil, or perhaps the formation of a helix, as seen for example in *C. elegans* SAS-6 (Hilbert et al., 2013).
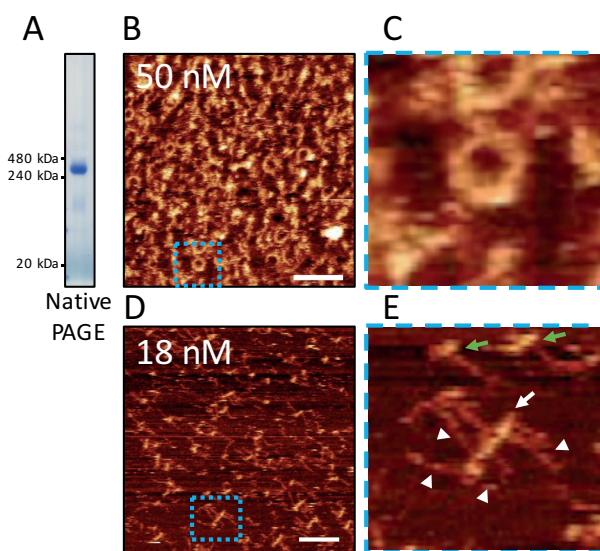


Fig. 4. 8 TaSAS-6_2_CC (1-476) forms ring-like structures and rod-like oligomers *in vitro*. **A** Native PAGE gel of purified TaSAS-6_2 with the whole coiled-coil domain (CC) shows the presence of probable higher-order oligomers in solution. **B-C** Mica surface covered with the TaSAS-6_2_CC oligomers and rings at a solution concentration of 50 nM. The dashed square represents the field magnified in C. Note that the crowding effect on the Mica surface lowers the resolution of the image, such that the spokes are not visible. Scale bar: 60 nm. **D-E** Injection of TaSAS-6_2_CC protein with a lower concentration resulted in the loss of rings, but reveals the capability of the protein to oligomerize through its N-terminal domain. White arrow depicts the N-terminal domains assembled into straight oligomer, which shows up with higher intensity than the coiled-coil dimers. Note spokes of individual dimers (white arrowheads) on both sides of the straight oligomer. Green arrows show the presence of dimers with one spoke. The dashed square represents the region magnified in E. Scale bar: 50 nm. AFM topographs taken by N. Banterle.

### 4.2.3 TaSAS-6_1 and TaSAS-6_2 hetero-oligomerize

In order to determine a possible interaction between TaSAS-6_1 and TaSAS-6_2, we mixed the two proteins with different lengths of their coiled-coil domains to distinguish them and used high-speed AFM to observe their behaviour on the Mica surface (Fig. 4.9 A and B). Preliminary observations indicate that the two paralogs might interact and form hetero-oligomers (Fig. 4.9 B and C).
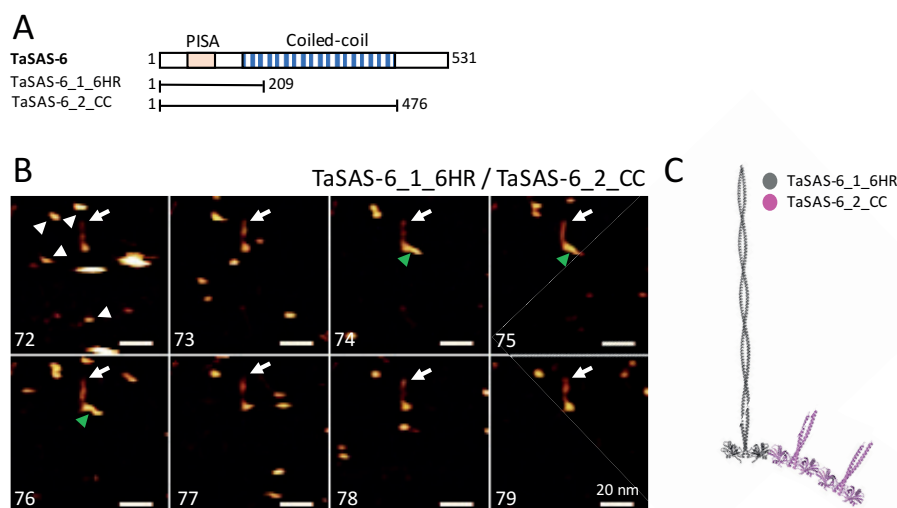


Fig. 4. 9 TaSAS-6_1 and TaSAS-6_2 interact *in vitro*. **A** Schematic representation of two paralogs of the TaSAS-6 used for the experiment with different lengths of their coiled-coil domains. **B** Montage of eight frames from a high-speed AFM acquisition of the Mica surface with a mixture of TaSAS-6_1_6HR and TaSAS-6_2_CC proteins in 1:1 ratio, with a solution concentration of ~ 55 nM for each. Each time frame corresponds to 2.56 s. Arrows point TaSAS-6_2_CC with the elongated coiled-coil domain. White arrowheads point the TaSAS-6_1_6HR with very short coiled-coil region, not visible on the topograph. Green arrowheads point to the high intensity rod-like structure extending from the coiled-coil of the TaSAS-6_2_CC, which might indicate the binding of one or two TaSAS-6_1_6HR dimers to the N-terminal domain of the TaSAS-6_2. Scale bar: 20 nm. Experiment done by N. Banterle. **C** Probable schematic model of the oligomer formed in B in the frame 75.

### 4.2.4 Human homolog of SAS-6

HsSAS-6 is well studied in the cellular context, however many questions about its oligomerization and ring-formation properties, as well as its crystal structure, are still unresolved. HsSAS-6 has a predicted canonical SAS-6 protein tertiary structure (Fig. 4.10 B), and harbors Phe131 as the hydrophobic residue that docks into the hydrophobic pocket in order to mediate oligomerization through the N-terminal domain. The protein dimerizes through the coiled-coil moiety and is present primarily as a homodimer in the cytoplasm of human cells, oligomerizing only when at centrioles (Keller et al., 2014). We suppose that the entire coiled-coil moieties of HsSAS-6 dimerize with a $K_d$ similar to the $K_d$ of ~ 1 μM of the CrSAS-6 coiled-coil (Kitagawa et al., 2011). The N-terminal part

of HsSAS-6 (1-164) dimerizes with a $K_d$ of ~ 50 μM (van Breugel et al., 2011) which is comparable to that of ring-assembly competent CrSAS-6 (1-159) (Kitagawa et al., 2011).
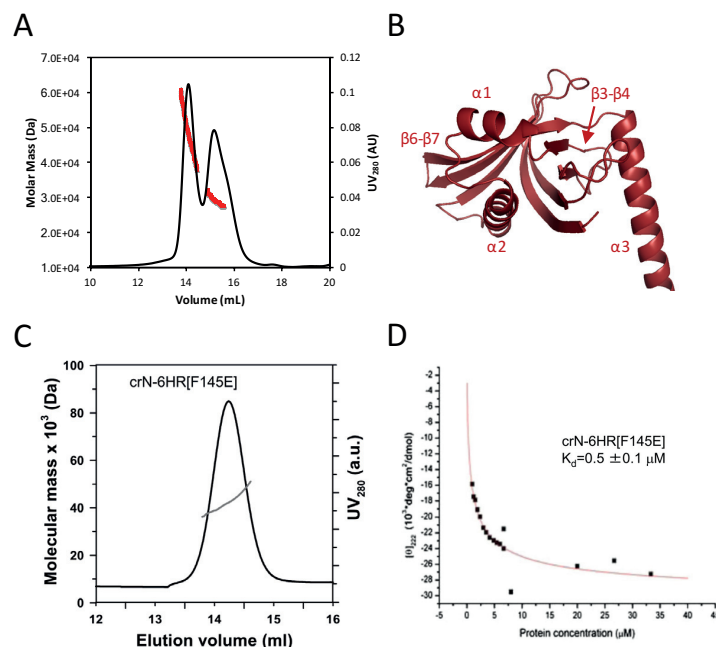


Fig. 4. 10 Comparison of HsSAS-6_7HR and CrSAS-6_6HR dimerization efficiency. **A** MALS analysis of HsSAS-6_7HR construct at 57 μM. Size exclusion chromatography profile is shown in black as absorbance at $UV_{280}$ and is overlaid by multi-angle light scattering estimation of molecular weight in red. HsSAS-6_7HR protein elutes in two fractions, with estimated molecular weight of 49.0 kDa (on the left) and of 29.2 kDa (on the right), which suggests the presence of both dimer and monomer of HsSAS-6_7HR with a molecular weight of 24.5 kDa. **B** Structural prediction of the conserved HsSAS-6 N-terminal domain with short coiled-coil domain. α-helices and β-strands are depicted. **C** MALS analysis of CrSAS-6_6HR[F145E] dimerization deficient mutant which is present only as a homodimer in solution, with molecular weight of 48.3 kDa corresponding to a stable homodimer (Kitagawa et al., 2011). **D** Dilution series of CrSAS-6_6HR[F145E] monitored by circular dichroism at 222 nm. $K_d$ is estimated based on the monomer-dimer model fitted into the data (Kitagawa et al., 2011).

We estimated also the dissociation constant of HsSAS-6_7HR based on multi-angle light scattering measurement, where we observed that the protein exists in equilibrium as a dimer and as a monomer at a concentration of ~ 50 μM (Fig. 4.10 A). This is in striking contrast with the $K_d$ of ~ 0.5 μM of the equivalent *Chlamydomonas* CrSAS-6_6HR protein (Fig. 4.10 C and D) (Kitagawa et al., 2011), and might explain the different behavior of these two constructs *in vitro*. The $K_d$s of HsSAS-6 N-terminal domain only and of its short coiled-coil domain-containing construct (HsSAS-6 7HR) are however in good agreement with each other, and we assumed that the protein should oligomerize when in sufficient concentration. We confirmed that the protein is present as a dimer in concentrations higher than its $K_d$ from the size exclusion chromatography – multi-angle light scattering (SEC-

MALS) experiments (Fig. 4.11 A). The dimerization is more likely through the HsSAS-6 6HR coiled-coil domain, because the N-terminal dimerization is disrupted in SEC-MALS as seen for HsSAS-6_(1-170)_cortexillin chimeric construct, which exists only as a dimer in solution, presumably through its coiled-coil domain (Fig. 4.11 B). Therefore, we decided to use also the shorter 7HR construct for oligomerization and ring-formation assays. However, the use of longer constructs with longer coiled-coil version should stabilize the coiled-coil interaction and reduce its $K_d$, as well as the use of chimeric constructs, such as HsSAS-6_(1-170)_cortexilin.
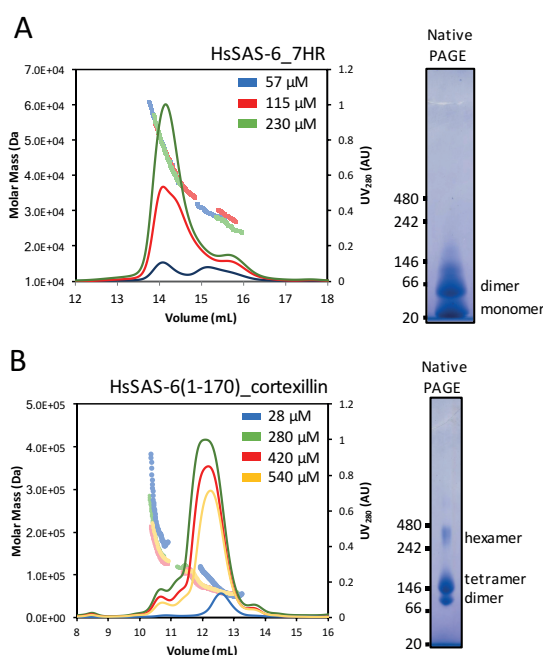


Fig. 4. 11 Stability of two HsSAS-6 constructs in solution at different concentrations. **A-B** MALS analysis of HsSAS-6_7HR (A) and HsSAS-8(1-170)_cortexillin (B) at different concentrations. Size-exclusion chromatography elution profiles assessed by $UV_{280}$ and overlaid by the multi-angle light scattering estimation of the protein or the protein complex molecular weight. A HsSAS-6_7HR is present as a dimer and as a monomer in solution. Native PAGE gel of native protein confirms the presence of dimers and monomers (on the right). B Chimeric construct HsSAS-6(1-170)_cortexillin is present as a dimer in solution even at very low concentration (28 μM). Note the shift of the peak to the left at increasing concentrations suggesting the formation of probable oligomers falling appart during migration through the column resin. The presence of several oligomer forms is visible also by Native PAGE gel (on the right).

Various wild-type and chimeric constructs of HsSAS-6 (Fig. 4.12 A) were subjected to AFM analysis in order to probe their oligomerization status. We tested different lengths of the wild-type protein, as well as a chimeric construct HsSAS-6_(1-170)-cortexillin, in which the coiled-coil domain is substituted by that of cortexillin. The latter is a well studied coiled-coil protein of 19 nm in length pre-

sent only as a dimer in solution (Fig. 4.11 B) (Burkhard et al., 2000; Faix et al., 1996), which was used here to enhance and ensure HsSAS-6 protein dimerization through the coiled-coil domain in solution (Fig. 4.11 B).
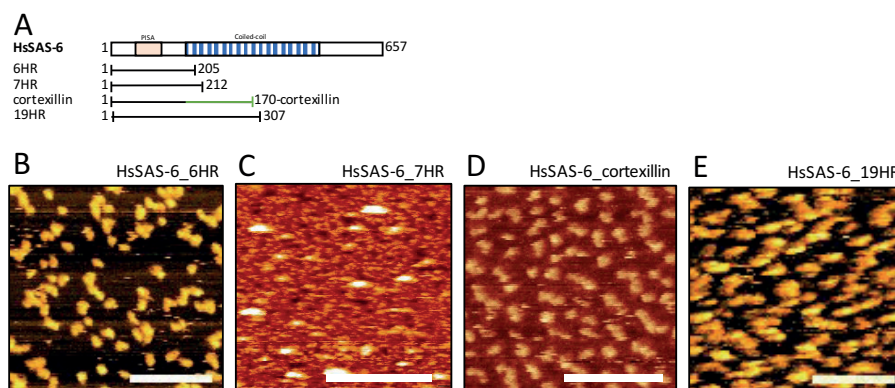


Fig. 4. 12 HsSAS-6 oligomerization trials on Mica surface observed by AFM. **A** HsSAS-6 protein constructs used for the analyses. **B**-**E** HsSAS-6 aggregates or is not folded properly on the Mica surface. AFM analysis was done in collaboration with Moritz Pfreudschuh (ETHZ) (B and E) and with Pascal Odermatt (EPFL) (C and D). Scale bars: 100 nm for B and E and 60 nm for C and D.

Even though we tested different constructs, a range of protein concentrations and various buffers, the proteins did not undergo higher-order oligomerization or did not adsorb on the Mica surface (Fig. 4.12 B-E). Therefore, I decided to screen conditions for HsSAS-6 oligomerization and ring formation using negative staining-TEM, as discussed above. I used a construct with the whole coiled-coil domain to ensure dimer formation (Fig. 4.13 A) and tested several buffers with different ionic strength, as well as a few salt concentrations at three different temperatures (Fig. 4.13 B). The concentration of protein in solution was ~ 60-120 nM, which is typically used for CrSAS-6 ring formation studies (Pfreundschuh et al., 2014), even if it is significantly lower than the $K_d$ of the HsSAS-6 N-terminal domains. CrSAS-6 oligomerizes and forms rings even at this low concentration (Hilbert et al., 2016; Pfreundschuh et al., 2014), due to the increase in local protein concentration on the Mica surface or the grid. I indeed observed grid crowding at concentrations higher than 120 nM. Unfortunately, however, conditions for HsSAS-6_CC oligomerization were not found (Fig. 4.13 B and C).
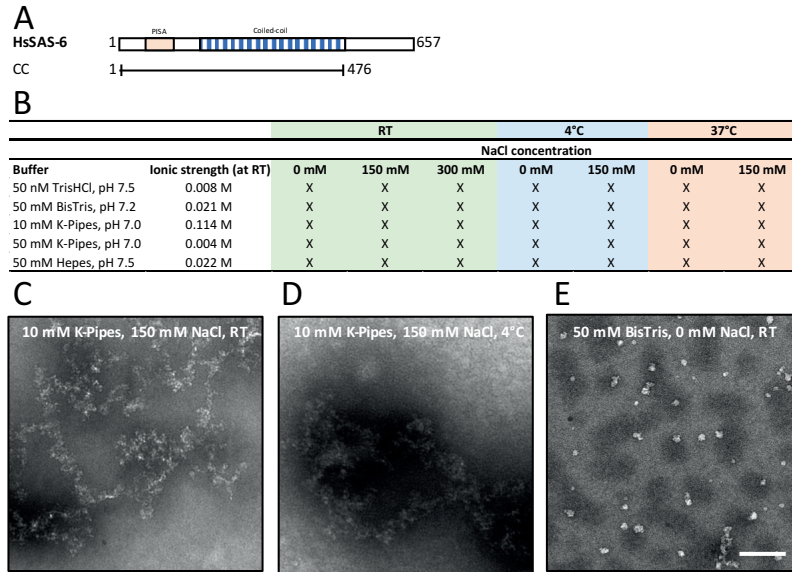
Fig. 4. 13 Inability of HsSAS-6 with the whole coiled-coil domain to oligomerize *in vitro*. **A** HsSAS-6 construct used for screening. **B** HsSAS-6_CC was tested at a range of buffers, salt concentrations and temperatures, as indicated. "X" represents aggregates or no visible structure formation. **C-E** Three examples of aggregates (C and D) and misfolded or collapsed protein that we judge from the fact that the elongated coiled-coils are not visible (E). Scale bar: 50 nm.

The inability of HsSAS-6 to form rings *in vitro* might have several explanations. First, even though we tested many conditions, the screening was not exhaustive, which may explain why we did not find a buffer suitable for protein stability and oligomerization properties. Second, the fact that the protein concentration is far below its $K_d$ might be a critical point. We know that HsSAS-6 is present as a dimer in the cytoplasm of human cells and oligomerizes only when present at very high local concentration at centrosomes (Keller et al., 2014). Therefore, the negative staining TEM or the AFM method might not be suitable for the observation of HsSAS-6 oligomerization due to the crowding of the support at concentrations > 120 nM, which might be needed to promote oligomerization of the human homolog and might be higher than the concentration necessary for CrSAS-6 ring formation. Third, the protein might be unstable or have a high degree of flexibility without an interaction partner. We know several interaction partners of HsSAS-6 in the cell, including STIL and Cep135, and therefore we can hypothesize that HsSAS-6 cannot form rings properly in their absence.

### 4.2.5 HsSAS-6 specific monobodies

The oligomerization properties of SAS-6 are critical for the initiation of cartwheel formation and subsequent procentriole growth (Leidel et al., 2005; Nakazawa et al., 2007; Strnad et al., 2007). As the results from the AFM and EM studies indicate, HsSAS-6 is likely to be misfolded or to aggregate *in vitro*. As mentioned above, one possibility is that HsSAS-6 oligomerization is regulated by an interaction partner, which may stabilize the conformation of HsSAS-6 and thus lead to proper dis-

play of the hydrophobic pocket and the hydrophobic residue to allow oligomerization through N-terminal domain interactions. Therefore, we decided to develop synthetic binding monobody proteins against the HsSAS-6 N-terminal domain or the first heptad repeat of its coiled-coil domain to stabilize the protein. It has been already shown in various instances that such binding can stabilize a protein of interest and restrict its conformation states, and thus improve the protein crystal lattice formation and provide better results in crystallization trials (Koide, 2009; Reckel et al., 2017; Stockbridge et al., 2015). Therefore, we aimed to raise monobodies against HsSAS-6 as a tool, not only to stabilize HsSAS-6 protein conformational status for ring-formation assays, but also for crystallization screening and as a tool to disrupt HsSAS-6 oligomerization *in vivo* in future studies.

The monobody is a small antibody mimic (Fig. 4.14 A), which uses the tenth fibronectin type III domain of human fibronection (FN3) as a scaffold (Koide et al., 1998, 2012). FN3 is a small, monomeric beta-sandwich protein (Baron et al., 1991), resembling immunoglobulins, and which is involved in specific molecular recognition. Monobodies are small monomeric proteins (~ 10 kDa), for which libraries containing many diversified positions on surface loops and some β-sheets have been generated (Koide et al., 2012), and which are used for screening (Batori et al., 2002; Koide et al., 2012). Monobodies raised against recombinantly expressed and purified biotinylated HsSAS-6_6HR were generated using phage and yeast display from the so-called *side-and-loop* library (Koide et al., 2012). Monobodies with high affinity were identified after two rounds of yeast display (Table 4.1).

Table 4. 1 Selected monobody clones for HsSAS-6. Eight binders were generated in total, 4 of which have the same amino acid sequence, as shown below.

| Name | Clone | Amino acid sequence |
|------|-------|---------------------|
| | | 10    20    30    40    50    60    70    80    90 |
| *library* | side-and-loop | GSVSSVPTKLEVVAATPTSLLISWDAPAVTVOUYOITYGETG(X$_{5-6}$)QTFTVPGSKSTATISGLKPGVDYTITVYA(X$_{7-13}$)--------SPISINYRT |
| MB01 | F3-I | GSVSSVPTKLEVVAATPTSLLISWDAPAVTVVYYLITYGETGYGPPVQTFTVPGSKSTATISGLKPGVDYTITVYAYMGNDYWDSYSPY-SPISINYRT |
| MB02 | F3-II | GSVSSVPTKLEVVAATPTSLLISWDAPAVTVVYYLITYGETGYGPPVQTFTVPGSKSTATISGLKPGVDYTITVYAYMGNDYWDSYSPY-SPISINYRT |
| MB03 | F3-III | GSVSSVPTKLEVVAATPTSLLISWDAPAVTVVYYLITYGETGYGPPVQTFTVPGSKSTATISGLKPGVDYTITVYAYMGNDYWDSYSPY-SPISINYRT |
| MB04 | F3-V | GSVSSVPTKLEVVAATPTSLLISWDAPAVTVVYYLITYGETGYGPPVQTFTVPGSKSTATISGLKPGVDYTITVYAYMGNDYWDSYSPY-SPISINYRT |
| MB05 | F3-VI | GSVSSVPTKLEVVAATPTSLLISWDAPAVTVIFYVITYGETGGNSPVQEFTVPGSKSTATISGLKPGVDYTITVYAKFDGSYW-----Y-SPISINYRT |
| MB06 | F4-I | GSVSSFPTKLEVVAATPTSLLISWDAPAVTVFYVITYGETGYASYVQEFAVPGSKSTATISGLKPGVDYTITVYAWVRHKGW-----S-SPISINYRT |
| MB07 | F4-V | GSVSSVPTKLEVVAATPTSLLISWDAPAVTVDFYVITYGETGW-PGYQEFEVPGSKSTATISGLKPGVDYTITVYAWHY--IY-----YDSPISINYRT |
| MB08 | F4-VI | GSVSSVPTKLEVVAATPTSLLISWDAPAVTVDFYVITYGETGW-PGYQEFEVPGSKSTATISGLKPGVDYTITVYAGFTDQYY-----YGSPISINYRT |

From 5 unique monobodies, we chose the monobody 01 (MB01), as it exhibited a high binding affinity, with a $K_d$ of ~150 nM, as measured in the yeast-display format (Fig. 4.14 B). This $K_d$ value was in a good agreement with the one determined using isothermal titration calorimetry (ITC) with purified monobody MB01 and HsSAS-6_6HR (Fig. 4.14 C).
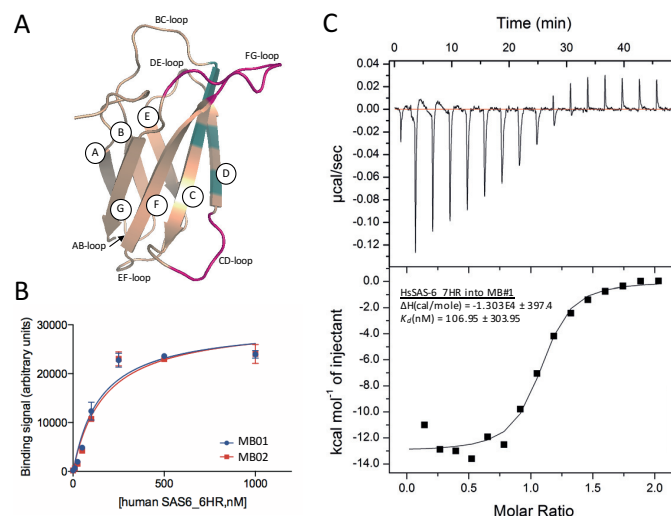
Fig. 4. 14 Monobody MB01 binds HsSAS-6 with nanomolar affinity. **A** Structural prediction of MB01 with its conserved fibronectin type III domain scaffold; β-strands are labelled A-F. Variegated loops are coloured in pink and side chains in blue. Structural prediction was made using Phyre[2] (Kelley et al., 2015) **B** Binding measurement by yeast surface display. The mean fluorescence intensities of yeast cells displaying MB01 and MB02 are plotted as a function of HsSAS-6_6HR concentration. The standard deviations from curve fitting of the 1:1 binding model are shown as error bars. Experiment done by T. Kükenhorner. **C** ITC measurement of MB01 with HsSAS-6_6HR. The calorimetric titration was performed at 25°C. The top panel shows raw heat signal of the ITC experiment and the integrated calorimetric data of each peak are at the bottom. The best fit of the data based on 1:1 binding model is represented as a continuous line.

The complex of the MB01 monobody with HsSAS-6_7HR used also for the ring-formation assays and later for the crystallization screening was co-purified in order to promote the interaction and remove all unbound proteins. We found that after at least 24 h incubation at 4°C at the high concentration of ~ 150 mM at 1:1 ratio (Fig. 4.15 A), higher-order oligomers formed, as assessed by native gel protein electrophoresis (Fig. 4.15 B). Based on this result, we could estimate the highest molecular weight of the complex formed by the monobody MB01 and HsSAS-6_7HR protein to be ~ 720 kDa. This would correspond to the oligomer formed by 18 HsSAS-6_7HR molecules bound to its monobodies, with $M_W$ of 27 kDA and 10 kDa, respectively. The short coiled-coil region should not change significantly the running behaviour of the protein as it is the case of the elongated coiled-coils. Furthermore, a smear in the smaller molecular weight region (Fig. 4.15 B) suggests that a range of smaller oligomers may be present in solution.
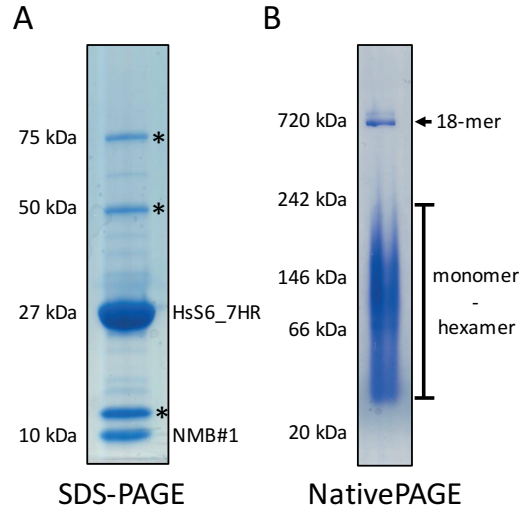
Fig. 4. 15 HsSAS-6 oligomerizes when bound to its specific monobody MB01. **A** SDS-PAGE gel of the complex of HsSAS-6_7HR and the monobody MB01. Asterisks indicate contaminants inside the solution. **B** Blue NativePAGE gel shows higher-order oligomer formation of the HsSAS-6_7HR with MB01. Probable mixture of various oligomer forms, with as indicated as a smear on the gel and possible 18-mer, which would correspond to a ring formation. Monomer comprises a complex of HsSAS-6_7HR and MB01 of ~ 37 kDa

Therefore, we analysed this sample also by cryo-EM (Fig. 4.16 A-D) and negative staining TEM (Fig. 4.16 E-H). We found that HsSAS-6 7HR can form ring-like structures in the presence of MB01. The observed rings have the SAS-6 specific nine-fold symmetry (Fig. 4.17 A, B, D and E) and a diameter similar to those of SAS-6 nine-fold rings from other species observed *in vitro* (van Breugel et al., 2014; Hilbert et al., 2016; Kitagawa et al., 2011) (Fig. 4.17 C and F). Strikingly, such rings can also stack into long tube-like structures, resembling the cartwheel. However, the periodicity of ~ 4-5 nm, as assessed by 2D Fourier transform (Fig. 4.16 C and G) or by the Plot Profile function implemented in Fiji (Schindelin et al., 2012) across the stack (Fig. 4.16 D and H), does not correspond to the ~ 8.5 nm cartwheel periodicity observed *in vivo* (Guichard et al., 2012, 2013). This might be due to the absence of the HsSAS-6 spokes, which are likely to be involved in the establishment of the 8.5 nm periodicity, when connecting to the microtubule triplets through the Pinhead (Guichard et al., 2013).
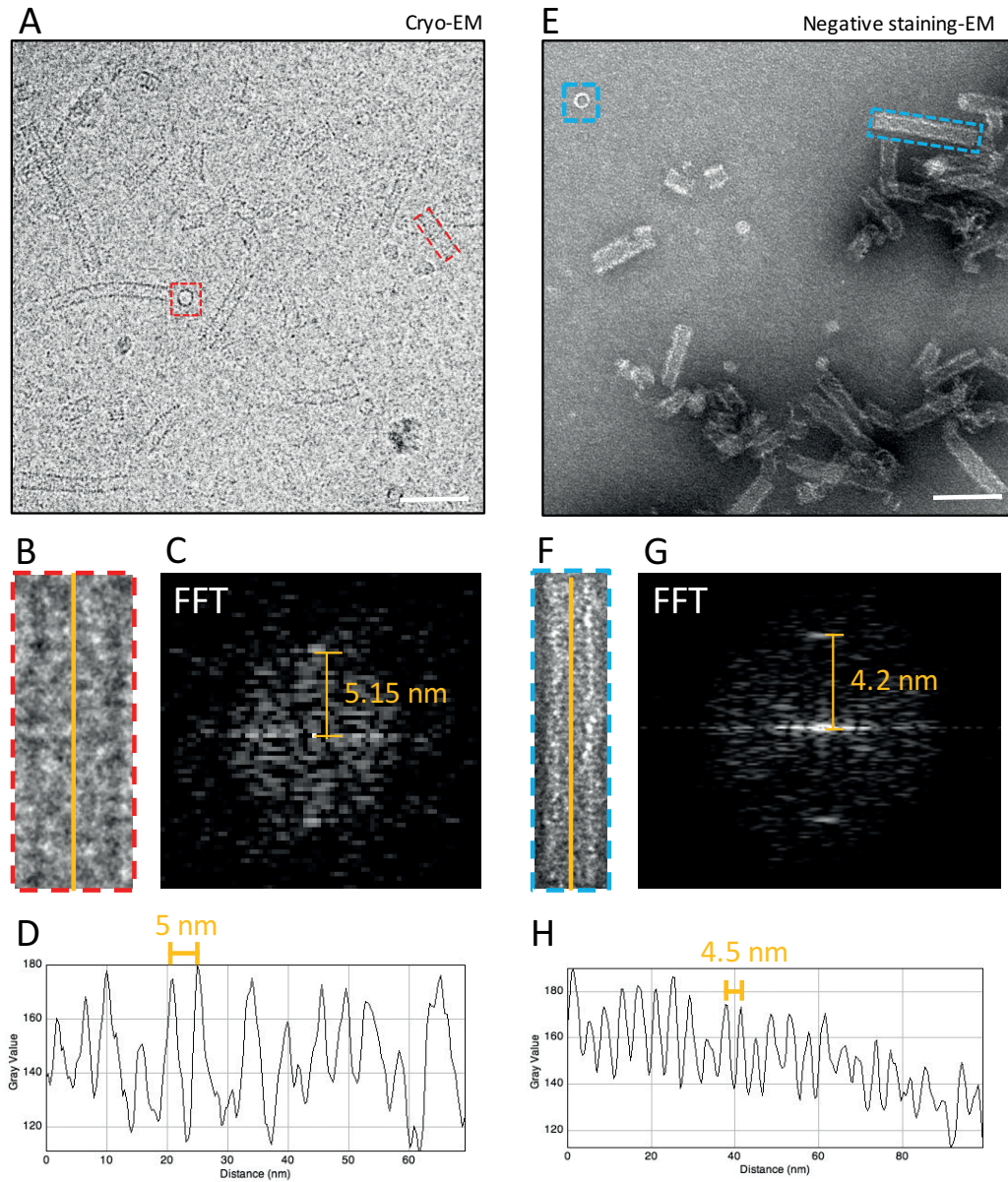
Fig. 4. 16 HsSAS-6_7HR forms cartwheel-like stacks when bound to its monobody MB01. **A** and **E** Micrographs from cryo-EM (A) or from negative staining-TEM (E) showing the rings and elongated tubular structures of the complex HsSAS-6_7HR and MB01. Red and blue dashed rectangles represent two stacks, which are enlarged in B and F, respectively. **B-F** One HsSAS-6_7HR/MB01 stack taken into consideration for periodicity determination using Fourier transform (FFT, in **C** and **G**, respectively) or using the Plot Profile function (D and H, respectively) implemented in Fiji (Schindelin et al., 2012).
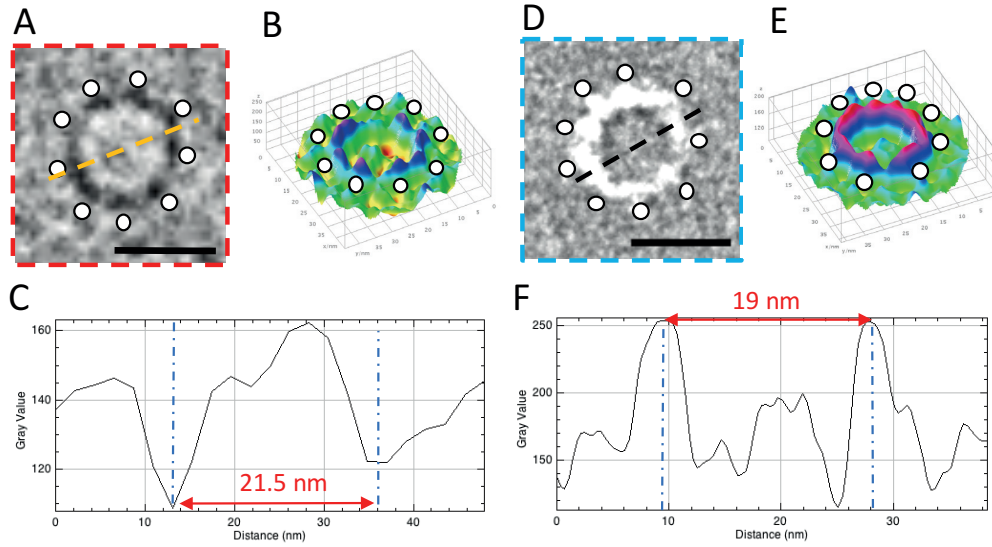
Fig. 4. 17 Determination of HsSAS-6_7HR ring symmetry and diameter. **A** Top view of a stack seen in the cryo-EM micrograph, with the depiction of HsSAS-6_7HR dimers (white circles). The dashed line represents the section, where the plot profile (C) was determined. Scale bar: 20 nm. **B** 3D visualization of a ring (A) by surface plot showing peaks at the positions of HsSAS-6_7HR. **C** Plot profile across the ring, showing the diameter of 21.5 nm. **D** A short stack of individual HsSAS-6 rings seen on the grid from the negative staining method, with the depiction of HsSAS-6_7HR dimers (white circles). The dashed line represents the section, where the plot profile (F) was calculated. Scale bar: 20 nm. Micrograph of a top view from cryo-EM. **E** 3D visualization of the ring (D) by surface plot showing peaks at the positions of HsSAS-6_7HR dimers in the ring. **F** Plot profile across the ring, showing the diameter of 19 nm. Surface plot and Plot profile are functions implemented in Fiji (Schindelin et al., 2012).

Furthermore, we used the HsSAS-6_7HR-MB01 complex for crystallization screening, as discussed further down.

### 4.2.6 Crystallization trials

In order to determine the variations in protein structure between known crystals structures of *Chlamydomonas* (Kitagawa et al., 2011), *C. elegans* (Kitagawa et al., 2011), zebrafish (van Breugel et al., 2011) or *Leishmania major* (van Breugel et al., 2014) and that of HsSAS-6 and the *T. agilis* homologs, we aimed to crystallize HsSAS-6, TaSAS-6_1 and TaSAS-6_2.
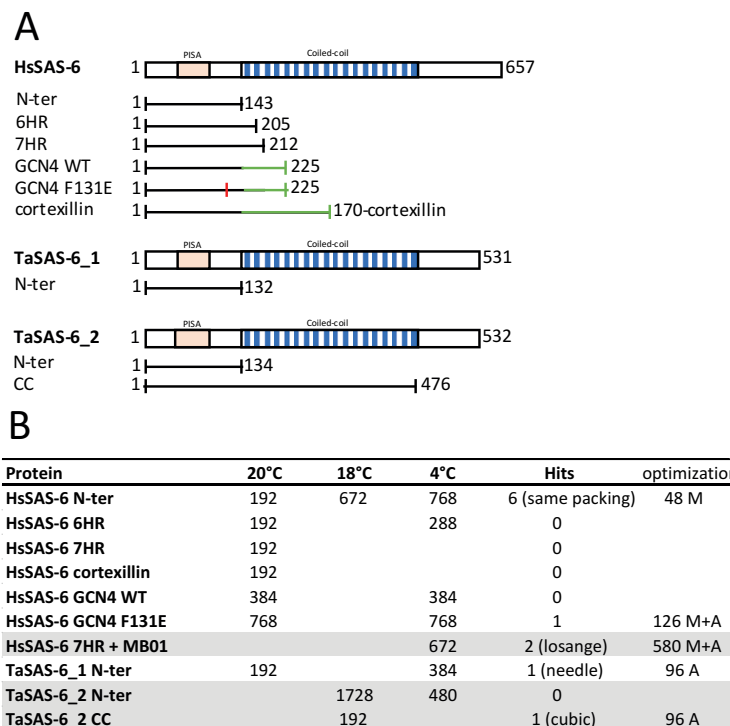
Fig. 4. 18 HsSAS-6 and TaSAS-6s crystallization screening. **A** Constructs of the three SAS-6 proteins used for the crystallization screenings. 6 constructs of HsSAS-6 were used: N-terminal domain (Nter), N-ter domain with 6 or 7 heptat repeats from the coiled-coil domain (6HR and 7HR, respectively), chimeric constructs with the coiled-coil proteins GCN4 and cortexilin. One N-ter construct was used for TaSAS-6_1 and two constructs for TaSAS-6_2: N-ter and N-ter with the whole coiled-coil domain. **B** Crystallization screenings with representation of the number of conditions at three different temperatures used for each construct. Number of hits with the shape of crystals, as well as the number of optimization conditions of the initial hits, set up automatically (A) in 96-well plate or manually (M), differing in the size of the final drop, 400 nl and 2 μl, respectively.
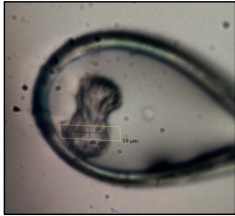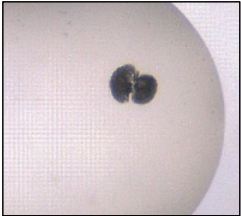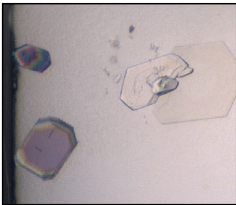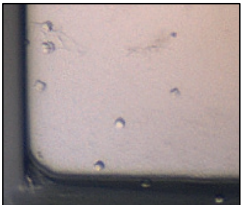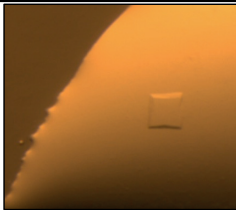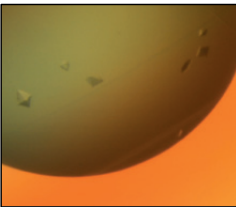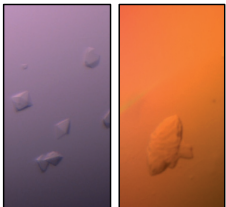
As suggested by our oligomerization studies of HsSAS-6 *in vitro* (Fig. 4.11 and 4.12), HsSAS-6 is likely unstable and adopts a range of conformational states, which was hindering our crystallization screenings. We found a few conditions where two HsSAS-6 constructs, HsSAS-6_N-ter and oligomerization deficient mutant HsSAS-6_GCN4_F131E, crystallized and diffracted at 15 Å and 3.5 Å, respectively. Even though diffraction at 3.5 Å resolution could be sufficient to solve the atomic structure, we were unable to index the diffraction pattern and determine the crystal geometry. This was probably due to an uncommon crystallization, where the HsSAS-6_GCN4_F131E crystallized into a small and thin 2D crystal seeded on a side of a large crystal of unspecific molecule from the buffer. Therefore, we could not obtain comprehensive data to solve the structure. Therefore, we tried to optimize these constructs further (Table 4.2). HsSAS-6 N-ter crystals displayed the same crystal packing of twisted tightly packed needles, impossible to separate into individual needles, also

after optimization. The final resolution we reached in this case was only ~ 8 Å, which is insufficient to solve the protein structure. The HsSAS-6_GCN4_F131E was subjected to further optimization as well, but we could not reproduce the specific crystal seeding described above. The crystallization took much longer, going from 3 days to ~ 2 weeks, and the final crystals were small and not diffracting well, unfortunately (Table 4.2).

The possibility to use the HsSAS-6 specific monobodies to reduce its conformational freedom led us to do a crystallization screening also with the complex of the protein HsSAS-6_7HR with the monobody MB01. This protein complex co-crystallized in the form of small octahedron crystals only after 3 days and the hit was reproducible. However, these crystals diffracted only at ~ 12 Å resolution. Thereafter, we set up large optimization screens with different protein concentrations, crystal seeding and temperatures, and obtained two new crystal shapes. Unfortunately, both of them did not diffract at all to proceed further.

Furthermore, we tried to crystallize one construct of TaSAS-6_1, its N-terminal domain, and two constructs of TaSAS-6_2, the N-terminal domain only and the N-terminal domain with the whole coiled-coil region. We were able to concentrate the TaSAS-6_1 N-ter only to ~ 125 μM (2 mg/ml), as higher concentrations led to protein precipitation. This is at the lower limit of useful protein concentration for crystallization screening (Dessau and Modis, 2011). We initially obtained one hit with crystals in a needle shape, but could not reproduce it (Table 4.2). For TaSAS-6_2, we obtained only one crystal hit for what must have been a degradation product of TaSAS-6_2 CC, as revealed by X-ray diffraction experiment and assessed by the indexing of the diffraction data, which showed that the unit cell of the crystal is too small to harbor the whole TaSAS-6_CC construct. We screened the shorter version of TaSAS-6_2 at different concentrations, ranging from 4-20 mg/ml, at two different temperatures, but the protein did not crystallize at all.

Table 4. 2 Details about the crystallization hits. The name of the construct with the used concentration, as well as the composition of the mother liquor, are listed. The crystals directly from screens and the crystals from the optimization trials are shown with the resolution at which they diffracted.

| Protein | Conditions | Crystal / resolution | Optimized crystals |
|---|---|---|---|
| **HsSAS-6 N-ter** at 6 mg/ml | Morpheus, C2: 30% w/v EDO-PEG8000 1 M MB1: imidazole, MES, pH 6.5 10% NPS: 0.03 M |  **15-20 Å** |  **8 Å** |
| **HsSAS-6_GCN4_F131E** at 20 mg/ml | JSCG IV, B4 0.1 M Bicine 65% v/v MPD pH 8.5 |  **3.5 Å** |  **worse** |
| **TaSAS-6_1 N-ter** at 2 mg/ml | PACT, B5 0.1 MIB 25% PEG1500 ph 8.0 |  **diffraction not tested** | |
| **TaSAS-6_2 CC** at 5 mg/ml | JSCG+, C4 0.1 M HEPES 10% w/v PEG6000 ph 7.0 |  **12 Å (degradation)** | |
| **HsSAS-6_7HR with MB01 1:1** at 5.5 mg/ml | JSCG+, H1 0.3 M Magnesium formate dihydrate 0.1 M BIS-Tris pH 5.5 |  **12 Å** |  **not diffracting** |

## 4.3     Discussion

In this study, we performed oligomerization tests of two *T. agilis* homologs, TaSAS-6_1, which was identified previously (Guichard et al., 2013), and TaSAS-6_2, which was identified in our proteomic analysis as the second TaSAS-6 paralog. Because the expression and behaviour of various constructs

of these two proteins varied significantly, and because TaSAS-6_1 was tested mostly at the beginning of the project, whereas TaSAS-6_2 was studied mostly towards the end, after we identified it from the proteomic data, we did not us the same approaches in the two cases. We could not find evidence of higher order oligomerization of TaSAS-6_1 *in vitro* at the highest concentration of 2 mg/ml we could reach before the protein starts to precipitate, which corresponds to ~ 125 μm. This was true even using the negative staining method with the TaSAS-6_1 construct with the short coiled-coil region. This is in striking contrast to TaSAS-6_2, which exhibits dimerization through the N-terminal domains and form rings on the Mica surface at a concentration of 50 nM. However, these rings are probably formed only at very high local concentration during protein absorption on the Mica surface. When the protein is diluted ~ 3x to 18 nM, the rings are not formed. Perhaps TaSAS-6_2 can form rings only at very high local concentration during centriole formation, as was shown for HsSAS-6. Indeed, in human cells, HsSAS-6 is present as a homodimer in the cytoplasm and oligomerizes only when at centrioles (Keller et al., 2014). Furthermore, even in lower concentration, probably below that needed for ring formation, TaSAS-6_2 still forms higher-order oligomers of another sort, i.e. elongated rod-like structures with spokes emerging outwards in all directions (Fig. 4.8 D and E). The property of the protein to form rings as well as straight oligomers might suggest that the conformation of the protein is very flexible, possibly in the neck region, from residue 135 to 160, where the α-helix starts to dimerize and form the parallel coiled-coil. Another possibility is that TaSAS-6_2 forms helical oligomers as was already observed for *C. elegans* SAS-6 (Hilbert et al., 2013). We were unable to assess the helicity of these oligomers in the AFM topographs. Analysis by negative staining, as done before for *C. elegans* SAS-6 (Hilbert et al., 2013) might answer this question.

Moreover, we found that *T. agilis* contains four paralogs of TaSAS-6 altogether. What is their respective function in the centriole? Up to now, we were able to assess the localization of the first two. TaSAS-6_1 and TaSAS-6_2 colocalize along the whole proximal cartwheel-bearing part of the *T. agilis* centriole.

However, to assess whether they interact inside one ring or whether instead the cartwheel is formed by intercalating layers of different TaSAS-6s, as discussed in Chapter III (see Fig. 3.13), one would need to use one of the available super-resolution microscopy techniques, which can localize centriolar proteins with a 4-5 nm precision (Gartenmann et al., 2017).

In the future, a probable interaction between TaSAS-6_1 and TaSAS-6_2 should be confirmed using a range of biochemical methods, such as size-exclusion chromatography, native protein electrophoresis, co-immunoprecipitation, as well as by more profound AFM imaging or negative staining, if one uses the protein constructs with different coiled-coil lengths as we tested in Fig. 4.9.

### 4.3.1 HsSAS-6 ring formation

The large screening of various HsSAS-6 constructs, conditions and techniques allows us to speculate that the HsSAS-6 protein is either unstable *in vitro* or might adopt a range of conformational states that would hinder its higher-order oligomerization properties. This would suggest that an interaction partner is needed in order to stabilize the protein tertiary structure and/or to expose the inter-

action surfaces, the hydrophobic pocket on one HsSAS-6 molecule and the F131 residue that docks into it on the other moiety to promote higher order oligomerization.

In collaboration with Tim Kükenshöner from the Hantschel laboratory, we developed specific monobodies against the HsSAS-6_6HR protein. The monobody MB01 promoted not only HsSAS-6 oligomerization, but also HsSAS-6 ring stacking, which implies that MB01 stabilized the protein and/or reduced the range of its conformational freedom, thus promoting HsSAS-6 oligomerization. By determining the structure of the HsSAS-6/MB01 complex, we would be able to localize the binding site of the MB01 on HsSAS-6, and by comparison with known HsSAS-6 binders to uncover a possible interaction partner responsible for HsSAS-6 stabilization and cartwheel formation.

### 4.3.2 Further studies on monobodies

We found a specific monobody MB01 with high affinity to its target HsSAS-6_6HR or 7HR, with a $K_d$ of ~ 107 nM. MB01 significantly improved the oligomerization properties of HsSAS-6, probably by stabilizing its conformation. One would need to test other monobodies that were found to be specific to HsSAS-6 (Table 4.1) or to screen for new ones in order to use them as tools to disrupt oligomerization and ring formation of HsSAS-6. By contrast, MB01 promotes HsSAS-6 oligomerization. Because the majority of the monobodies raised against CrSAS-6_6HR construct bind the coiled-coil domain (G. Hatzopoulos, personal communication), we can hypothesize that the binding site of MB01 is also somewhere at the level of the coiled-coil domain close to or interacting also with the N-terminal region, on a site where it does not inhibit the N-terminal domain interaction.

### 4.3.3 Crystallization trials

Many constructs of HsSAS-6 or TaSAS-6_1 and _2 were used for solving the crystal structures of these three proteins. We obtained several hits, which were not reproducible, or in some cases the optimization did not yield a crystal with a diffraction sufficient to solve the structure. However, we screened mostly constructs that were designed to be used in the oligomerization assays, meaning proteins with variably long coiled-coild domains, which might be too flexible to pack into a crystal. In the future, one might conduct more rigorous crystallization screening in order to design and produce proteins with small changes in the length of the SAS-6 neck region or the coiled-coil domain to find a better construct suitable for crystal packing.

## 4.4    Personal contributions

I cloned several SAS-6 constructs used in the study (see Table A.1), optimized protein expression and purification of all proteins, monobodies included, that were used for the oligomerization and crystallization studies (see Table A.2). I performed protein characterization using biochemical methods, including SEC-MALS, isothermal titration calorimetry (ITC) and native PAGE. I optimized the negative staining method for SAS-6 ring visualization and acquired all negative staining-TEM images. I performed the HsSAS06_CC oligomerization screening, crystallization screening and crystal growth optimization of all used proteins. I analysed TaSAS-6s and HsSAS-6 and made structural predictions. I processed and analysed all negative staining and cryo-TEM images.

This project was done in collaboration with:

- Natacha Olieric, Andrea Prota and Michel O. Steinmetz from Paul Scherrer Institute (Villigen). Natacha Olieric supervised the protein purification and crystallization trials of several protein constructs and together with Andrea Prota supervised the collection of diffraction data.
- Moritz Pfreudschuh and Daniel O. Müller from ETHZ (Basel), Pascal Odermatt, Adrien P. Nievergelt and Georg Fantner from EPFL (Lausanne) and Niccolò Banterle from EPFL (Lausanne) who performed the traditional and high-speed AFM topographs acquisition of the proteins used for oligomerization studies.
- Tim Kükenshorner and Oliver Hantschel from EPFL (Lausanne) who performed the monobody screening.
- Aline Reynaud and Florence Pojer from the Protein Crystallography Core Facility at EPFL (Lausanne) who helped with the crystallization screening settings. Florence Pojer directed collection of diffraction data.

# Chapter 5 **Conclusions**

My PhD thesis project was based on an uncommon model organism to study centriolar structure, *Trichonympha* sp. We aimed to extend its utilisation from the model used purely for structural studies of the centriole to one with an available centriole proteome, with the goal in mind that individual centriolar proteins would be analyzed by biochemical and crystallization methods to then be placed back into the cryo-EM map.

*Trichonympha* spp. harbors thousands of extremely long centrioles, up to 5 μm in length, with an elongated proximal region harboring the cartwheel. Therefore, it is an ideal model to study the structure and the composition of the proximal centriolar region. To our knowledge, only one species of the *Trichonympha* genus, *T. agilis*, has a sequenced and assembled genome and available transcriptomic data enabling us to elucidate the centriolar proteome.

All the results and possible future directions were already discussed at the end of each results chapter. Therefore, here I will only summarize the findings again and link them to other studies in order to get a more complete view of the state of knowledge regarding the proximal region of the centriole.

In this study, firstly, we aimed to determine the structure of the *T. agilis* centriolar proximal region in order to compare it with the previously published map (Guichard et al., 2013) using cryo-electron microscopy and subtomogram averaging. Due to centriole flattening, we analysed and averaged only the cartwheel hub. We improved the resolution up to ~ 25 Å, when using ninefold symmetrization. Moreover, we discovered that the second symbiont of *R. speratus* termite, namely *T. mirabilis*, harbors elongated centrioles with cartwheel along the entire analysed regions. However, the *T. mirabilis* cartwheel differs in the presence of CCD instead of CID. The CCD is in the shape of an elongated rod present along the whole length of the proximal end of the centriole. We resolved the *T. mirabilis* cartwheel hub at ~ 30 Å resolution. Various methods and tips how to improve the final resolution of the cartwheel hubs of both species are discussed in the Chapter II.

Secondly, we aimed to elucidate the proteome of the centriole proximal region using *T. agilis* centrioles. I optimized the protocol for isolation and purification of the centriole-rich rostra of the *T. agilis* cells using density gradient centrifugation, and performed a mass-spectrometry based protein correlation profiling of the fractions from the gradient. We identified 40 proteins, whose fractionation profiles correlated with the rostrum profile and another 20, whose profiles were slightly different, but still peaked in the rostrum peak fractions. We identified several homologs of known centriolar proteins in other species. Furthermore, we uncovered and characterized two paralogs of TaSAS-6, and found a fourth one by an additional search of the transcriptome and genome data bases.

Thirdly, we tested the oligomerization properties of TaSAS-6_1 and TaSAS-6_2, as well as of the human homolog HsSAS-6, and thus discovered an intrinsic property of the TaSAS-6_2 protein to form ring *in vitro* as it was observed for other species (van Breugel et al., 2011, 2014; Hilbert et al., 2016; Kitagawa et al., 2011; Pfreundschuh et al., 2014). Moreover, we found that TaSAS-6_2 can oligomerize into tree-like oligomers with a straight trunk at lower protein concentration. Through the analysis of the HsSAS-6 oligomerization properties and the protein stabilization by the specific monobodies, we found that the HsSAS-6 can form rings and also to stack.

## 5.1    SAS-6 in the cartwheel hub

We defined the shape of the cartwheel hub *in situ* to be a perfect nonagon with a slightly enlarged diameter when compared to CrSAS-6 rings observed *in vitro* or modelled *in silico* (van Breugel et al., 2011; Guichard et al., 2012, 2013, 2017; Hilbert et al., 2016; Kitagawa et al., 2011), with ~ 24.2 nm for the cartwheel hub and 22 nm for the CrSAS-6 ring *in vitro*. Consequently, the modelled CrSAS-6 ring does not fit the 3D density map perfectly (Fig. 2.13). The range of SAS-6 ring symmetries formed *in vitro*, as assessed mainly on CrSAS-6 rings, and ranging from 8 to 10 (Hilbert et al., 2016), indicates some conformational freedom to adopt slightly different conformations inside the ring. Furthermore, we found that TaSAS-6_2 oligomerizes and forms rings, as well as other higher order oligomeric structures. The rod-like oligomers can adopt an elongated straight conformation with coiled-coil dimers, or spokes, emanating towards the outside. This indicates that the TaSAS-6_2 protein itself can adopt variable conformations and might form nine-folded rings with a changed conformation and diameter, as suggested by the 3D density map.

As seen from the cryo-EM micrographs of *Chlamydomonas* centrioles (Guichard et al., 2013), the shape of the cartwheel ring is likely to be enlarged also in this species. Therefore, the CrSAS-6 ring is likely to be also present in a slightly altered conformation *in situ*. This would be another indication that even though SAS-6 can form rings mainly with nine-fold symmetry, other players are involved to ensure the exact symmetric arrangement of centrioles (reviewed in Gönczy, 2012).

We hypothesize that the protein(s) helping to establish the precise nine-fold symmetry in the case of *T. agilis* might reside either inside the cartwheel, for instance be the CID, or more externally to the cartwheel, for instance be the A-C linker, which might position the microtubule triplets in a specific orientation with respect to one another and form barrels with nine-fold symmetry, which would allow to accommodate SAS-6 rings only with nine-fold symmetry. One of the interesting candidates for the A-C linker is the centriolar protein POC1, whose localization was described on the centriolar wall (Keller et al., 2009; Meehl et al., 2016; Pearson et al., 2009) and whose removal results in the partial loss of microtubule triplets. (Meehl et al., 2016). A homolog of POC1 was also identified also in the proteomic analysis of *T. agilis* here.

## 5.2    Building a centriole model

Conventional light microscopy is a powerful technique that helped discover the existence of centrosomes and centrioles over a century ago. The size of most centrioles at ~250 nm x 500 nm is just above the diffraction limit of light. Therefore, the discovery and improvements of different super-resolution microscopy techniques (reviewed in Sydor et al., 2015) are allowing the field to localize individual proteins inside the centriole (Gartenmann et al., 2017; Hamel et al., 2014; Sonnen et al., 2012; Wegel et al., 2016), and thus build a "colored wooden ring baby toy" model of the centriole, where individual molecules are spatially organized to give rise to a molecular map (Fig. C.1). With the availability of the cryo-ET maps of the cartwheel and the whole proximal region of the centriole, one would want to correlate all this information with superresolution microscopy studies to generate a full integrative structural and molecular centriolar model. The results from the proteomic analysis of the *T. agilis* centriole should include mostly components present in the proximal region, which, when localized to specific regions by superresolution techniques in the future, might also be localized into an improved sub-nanometer cryo-ET map even without available crystal structures.
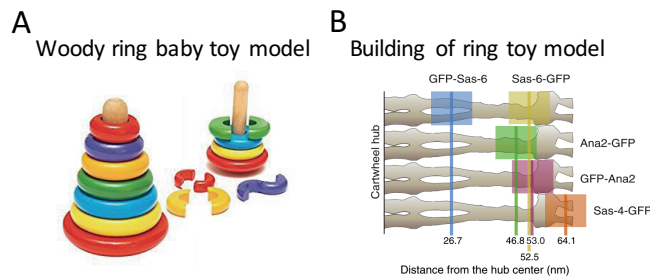


Fig. C. 1 Depiction of a ring toy model of centriole. **A** Centriolar components labelled and visualized by super-resolution microscopy are often seen as a ring or torus surrounding the cylindrically shaped centriole. Therefore, the centriolar ring toy model might look like a baby toy. **B** Schematic representation of the localization of a few centriolar proteins at nanometer resolution and depicted as rings, from (Gartenmann et al., 2017)

## 5.3    Concluding remarks

The centriole, as the tiny dot that was detected a long time ago during mitosis using light microscopy, has led to the foundation of an entire field and attracted many scientists around the world, who have gathered substantial knowledge about its composition, structure and functions in the cell.

Technical improvement in electron microscopy in the last years allowed one to uncover new and interesting structural features of the centriole and its cartwheel, and we are still trying to improve the resolution to better understand how centriolar proteins are organized inside.

We took advantage of the evolutionary conservation on the one hand and the amazing diversity of centriolar shape and length and centriole number per cell on the other hand to get more insight into the composition of the proximal region of the centriole.

Centrioles from different species might have a conserved overall structure, but also from this study we see that they differ in the numbers of centriolar protein paralogs present inside the centriole, as well as by their requirement during the first step in SAS-6 ring formation and cartwheel assembly.

With this project, we established the base and resources to study the long centriole of the *T. agilis* and its elongated proximal cartwheel-bearing region, which is formed at the onset of the centriole assembly. Therefore, this might have appliations in research about the cartwheel and centriole bio-genesis.

# Materials and Methods

## 5.4 Molecular Biology

### 5.4.1 Cloning of HsSAS-6 and TaSAS-6s

DNA encoding fragments of wild-type, mutant or chimeric constructs of HsSAS-6 (Uniprot ID O6UVJO), TaSAS-6_1 (Uniprot ID R4WPE9), as well as of the newly identified TaSAS-6_2 (un-published), were cloned using a Positive Selection method (Olieric et al., 2010) into the pET-based bacterial expression vector PSTCm1 encoding for N-terminal 6xHis-tag, the PSTCm8 vector encoding for N-terminal thioredoxin, and 6xHis-tag, a modified pGEX6p-2 vector encoding for N-terminal GST-tag and pET15b vector encoding for 6xHis-tag (Table 5.1).

Table A. 1 List of HsSAS-6, TaSAS-6_1 and TaSAS-6_2 constructs used in this study

| Expression vector | N-terminal tag | Construct | Simplified name | By |
|---|---|---|---|---|
| PSTCm1 | 6xHis | HsSAS-6 (1-212) | 7HR | Natacha Olieric |
| | | HsSAS-6 (1-170)-cortexillin | | Natacha Olieric |
| | | HsSAS-6 (1-170)-GCN4 WT | | Veronika Nemcikova |
| | | HsSAS-6 (1-170)-GCN4 F131E | | Natacha Olieric |
| | | TaSAS-6 (1-134) | N-ter | Natacha Olieric |
| PSTCm8 | 6xHis, thioredoxin | HsSAS-6 (1-205) | 6HR | Natacha Olieric |
| pGEX-6P-2 | GST | HsSAS-6 (1-484) | CC | Virginie Hamel |
| | | HsSAS-6 (1-370) | 19HR | Veronika Nemcikova |
| | | TaSAS-6_1 (1-207) | 6HR | Veronika Nemcikova |
| | | TaSAS-6_2 (1-145) | N-ter | Veronika Nemcikova |
| | | TaSAS-6_2 (1-476) | CC | Veronika Nemcikova |
| pET15b | 6xHis | HsSAS-6 (1-143) | N-ter | Ioannis Vakonnasis |

### 5.4.2 Protein expression and purification of 6xHis-tagged proteins

Recombinant protein expression was performed in the *Escherichia coli* strains BL21 gold (DE3) or RosettaTM (DE3)pLysS (Novagen) in Luria-Bertani (LB) medium. Protein expression was induced at 20°C for CrSAS-6_CC (Kitagawa et al., 2011), HsSAS-6_N-ter, HsSAS-6_6HR, HsSAS-6_1-170-cortexillin, HsSAS-6_1-170-GCN4, HsSAS-6_1-170-(F131E)-GCN4, TaSAS-6_Nter, and at 18°C for HsSAS-6_7HR by the addition of 0.5 mM IPTG for 16h. Proteins were purified using immobilized metal-affinity chromatography (IMAC) on HisTrap HP $Ni^{2+}$-Sepharose columns (GE Healthcare) at 4°C according to the manufacturer's protocol. Cleavage of the hexahistidine (6xhis) tag was performed during dialysis against thrombin cleavage buffer (20 mM Tris-HCl, pH 7.5 supplemented

with 150 mM NaCl and 2.5 mM CaCl$_2$) for 16h at 4°C using 2 units of human thrombin (Sigma) per milligram of recombinant protein. Cleaved proteins were reapplied to IMAC in order to separate the cleaved proteins from the 6xHis-tag and still tagged proteins, in the event of incomplete cleavage. Cleaved constructs were concentrated and gel filtrated either on a SEC HiLoad Superdex 200 16/60 column (GE Healthcare) or SEC HiLoad Seperdex 75 16/60 (GE Healthcare), which were equilibrated in 20 mM Tris-HCl, pH 7.5 supplemented with 150 mM NaCl and 2 mM DTT. The proteins were then concentrated and the purity of the proteins was assessed by SDS-PAGE (Table A.1). Protein concentration was estimated by UV at 280 nm. Proteins were immediately used for crystallization screenings and ring formation assays or flash frozen in liquid nitrogen and stored at -80°C until used.

### 5.4.3 Proten expression and purification of GST-tagged proteins

Recombinant protein expression was performed in the *E. coli* strains BL21 gold (DE3) or RosettaTM (DE3)pLysS (Novagen) in Luria-Bertani (LB) medium. Protein expression was induced at 18°C by the addition of 0.5 mM IPTG for 16h. Cell pellets resuspended in lysis buffer containing 50 mM Tris-HCl, pH 7.5, 500 mM NaCl, 5 mM EDTA, 1 mM DTT, 1:1000 protease inhibitor cocktail (Sigma) with the addition of lysozyme 15 mg/ml were lysed by sonication. The lysates were incubated with Glutathione sepharose beads (GE healthcare). The beads were then washed 5x with wash buffer: 50 mM Tris-HCl, pH 7.5, 800 mM NaCl, 5 mM EDTA, 1 mM DTT, and then additionally three times with decreasing NaCl concentration: 500 mM – 300 mM – 150 mM. GST-tags were cleaved on beads overnight with Prescission protease (GE Healthcare or homemade) in a cleavage buffer containing 20 mM Tris-HCl, pH 7.5, 150 mM NaCl, 5 mM EDTA, 1 mM DTT, and then eluted from beads. Additional washes of beads with cleavage buffer or with buffer with 300-800 mM NaCl concentrations are required to elute all cleaved protein from beads. Proteins were then concentrated and loaded on size exclusion chromatography either on SEC HiLoad Superdex 200 or Superdex 75 (GE Healthcare) in 20 mM Tris-HCl, pH 7.5, 150 mM NaCl, 1 mM DTT. The proteins were then concentrated and the purity of the proteins was assessed by SDS-PAGE (Table A.2). Protein concentration was estimated by UV at 280 nm. Proteins were immediately used for crystallization screenings and ring formation assays or flash frozen in the liquid nitrogen and stored at -80°C until used.

Table A. 2 Assessment of the purity of proteins used in this study and their stock concentrations.



| HsSAS-6 (1-212) | HsSAS-6 (1-170)-cortexillin | HsSAS-6 (1-170)-GCN4 WT | HsSAS-6 (1-170)-GCN4 F131E |
|---|---|---|---|
| 25 kDa => 6mg/ml | 35 kDa / 25 kDa => 20.5 mg/ml | 25 kDa / 20 kDa => 6 mg/ml | 25 kDa / 20 kDa => 41.5 mg/ml |
| **HsSAS-6 (1-143)** | **HsSAS-6 (1-205)** | **HsSAS-6 (1-484)** | **HsSAS-6 (1-370)** |
| 20 kDa / 15 kDa => 7 mg/ml | 25 kDa / 20kDa => 7.8 mg/ml | 63 kDa / 48 kDa => 0.35 mg/ml | 48 kDa / 35 kDa => 0.19 mg/ml |
| **TaSAS-6_2 (1-145)** | **TaSAS-6_2 (1-476)** | **TaSAS-6 (1-134)** | **CrSAS-6 (1-503)** |
| 15 kDa => 22 mg/ml | 63 kDa / 48 kDa => 10 mg/ml | 15 kDa => 2 mg/ml | 63 kDa / 48 kDa => 3 mg/ml |

## 5.4.4   Monobody screening

### 5.4.4.1   Cloning

HsSAS-6_6HR (1-205) was cloned into the BamHI and XhoI restriction sites of the pHBT vector (modified pET vector) containing 6xHis-Tag, tobacco etch virus (TEV) protease cleavage site and an Avi-tag for biotinylation.

### 5.4.4.2   Biotinylation of Avi-tagged HsSAS-6_6HR

Biotinylation was performed *in vivo* by co-transformation of *E. coli* BL21 gold (DE3) cells with a BirACm plasmid and pHBT_HsSAS-6_6HR. When the $OD_{600}$ of the culture reached 0.6, a final Biotin concentration of 50 μM was added and protein expression was induced at 18°C by the addition of 0.5 mM IPTG for 16h.

### 5.4.4.3   Monobody selection

The method for phage and yeast display library sorting has been described elsewhere (Koide et al., 2012), as well as the "side-and-loop" library of monobodies, which was used (Koide et al., 2012).

Briefly, yeast libraries were generated after four rounds of phage display, where the selected loops were amplified and the enriched monobody sequences were transformed into yeast. Binding of the yeast cells displaying the monobodies directed against the biotinylated HsSAS-6_6HR was assessed using Fluorescence-activated cell sorting (FACS) by their positive signal emission when bound. Final clones were isolated, sequenced and cloned into a pHBT plasmid. Biotinylated HsSAS-6_6HR with a concentration ranging from 10 nM to 1 µM, in Tris-buffered saline (TBS) with 0.1% BSA, was used in the yeast binding assay (Wojcik et al., 2016). The Protein was then incubated with yeast cells and anti-V5 antibody (mouse) at room temperature for 30 min. After two washes and incubation for 30 min at room temperature with streptavidin-DyLight650 and a FITC-coupled anti-mouse IgG, the sample was analysed on a Gallios (Beckman Coulter) or BD Accuri C6 (BD Bioscience) flow cytometer. Data were fitted to a 1:1 binding model and $K_d$ values were determined using Prism (Graphpad).

## 5.5    Biophysical methods

### 5.5.1    Isothermal titration calorimetry (ITC)

ITC experiments were performed at 25°C on a MicroCal iTC200 (GE) instrument. The proteins were dialyzed against 25 mM Hepes (pH 7.5) for 16 h and degassed. The HsSAS-6_6HR protein in the syringe was loaded for step-wise injections into MB01 solution with 10x lower concentration. The MicroCal doftware was used to integrate resulting heats and fit them with the one-step association model provided by the software package.

### 5.5.2    Size-exclusion chromatography Multi-angle light scatterning (SEC-MALS)

SEC-MALS was performed at 20°C in 25 mM TrisHCl, pH 7.5, 150 mM NaCl supplemented with 2mM DTT using an S-200 analytical size exclusion chromatography column. The column was connected in-line to mini-DAWN TREOS light scattering and Optilab T-rEX refractive index detectors (Wyatt Technology). Protein samples with concentrations of 0.25 - 20 mg/ml were used in the experiments.

## 5.6    Centriole and rostrum purifications

### 5.6.1    Centriole purification

Colonies of *Reticulitermes speratus* termites were collected from a forest in Japan (delivered by Y. Hongoh) and kept in the dark. Centrioles were purified by adapting a previously published method (Guichard et al., 2012, 2015). Shortly, the whole guts of cold-paralyzed termites were dissected and opened into dissection buffer (ice-cold 10 mM K-Pipes, pH 7.0). *T. agilis* cells were separated from the rest of gut contents by pipetting the visible contaminant cells and debris and by subsequent centrifugation for 20 s at 100xg in a tabletop centrifuge. The cells were then resuspended in the dissec-

tion buffer supplemented with 0.5% NP-40. After incubation for 1h on ice with occasional mixing, cell debris were separated by centrifugation for 3 min at 500xg at 4°C in a tabletop centrifuge and centrioles were then pelleted by centrifugation for 5 min at 1000xg at 4°C in a tabletop centrifuge. The last step was repeated with the dissection buffer to remove residual detergent from the sample.

## 5.7    Proteomic analysis

### 5.7.1    Identification of proteins by NanoLC-MS/MS

Each fraction was diluted in denaturing solution (sodium dodecyl sulfate [SDS] 2%, dithiothreitol [DTT] 50 nM) and incubated at 37°C for 1 h.  Samples with denatured proteins were concentrated on a 10 kDa filter (Amicon, Merck) and washed several times during the concentration procedure with the denaturing solution. The samples were loaded on a 10% SDS-polyacrylamide gel (Biorad) to remove iodixanol (OptiPrep) from the protein mixture and ran until the entire sample entered the gel. Note that samples containing iodixanol might run differently and need more time to enter the gel.

Each SDS-PAGE gel lane was sliced and proteins were digested in-gel using trypsin. Briefly, samples were reduced in 10 mM dithioerythritol (DTE), alkylated in 55 mM iodoacetamide (IAA), and individual pieces of the gel were dried. Digestion was performed overnight at 37°C using modified Mass Spectrometry grade trypsin. Desalting of peptides was done on C18 StageTips (Rappsilber et al., 2007) and then the peptides were dried down by vacuum centrifugation prior to Mass Spectrometric analysis. For LC MS/MS analysis, peptides were resuspended and separated by reverse-phase chromatography on a Dionex Ultimate 3000 RSLC nanoUPLC system in-line connected with an Orbitrap Q Exactive Mass-Spectrometer (Thermo Fischer Scientific). Database search was performed using Mascot (Matrix Science, Boston, USA), MS-Amanda (Dorfer et al., 2014) and SEQUEST in Proteome Discoverer v.1.4. against a concatenated database consisting of *R. speratus* database (Hayashi et al., 2013) and single-cell transcriptomic database of *T.agilis* databases (gift from Y. Hongoh). Further data processing and validation was performed in Scaffold 4.8.4. (Proteome Software, Portland, USA).

## 5.8    Antibodies and indirect immunofluorescence

### 5.8.1    Custom antibodies

Antibodies against TaSAS-6_1 and TaSAS-6_2 were raised against short specific peptides: 14 amino acids sequence, NQGDFDIRVSTIVC, for TaSAS-6_1 and 16 amino acids sequence, CSVLTNIISKDANSVD, for TaSAS-6_2, and injected into a rabbit and rat, respectively (Eurogentec). Antibodies were subsequently strip-purified against the corresponding recombinant proteins, TaSAS-6_1_N-ter and TaSAS-6_2_CC, and eluted with glycine pH 2.5 and neutralized with Tris-HCl pH 8.0.

### 5.8.2 Immunofluorescence

For the observation of large symbionts living in the termite gut, *T. agilis* and *T. mirabilis* cells freshly dissected from the gut were left to adhere on microscope slides treated with 0.1% poly-L-lysine, then fixed in -20°C methanol for 3-5 min, washed with 1xPBS and then incubated 1 h at room temperature in 1x PBS, 0.5% Tween-20 and 1% bovine serum albumin with primary mouse monoclonal antiacetylated tubulin antibody 6-11B-1 (1:1000, SigmaAldrich), then washed for 5 min in 1xPBS, and incubated 1 h at room temperature with anti-mouse Alexa 488 secondary antibodies (Life technologies) and the membrane dye FM™4-64 (1:1000, Thermo Fischer).

5 μl aliquots of OptiPrep gradient fractions were centrifuged at 10,000xg (JS13.1, Beckman Coulter) onto a 12 mm a coverslip in a 15 ml Corex tube with a home-made adaptor and concentrator (Fig A.1). Coverslips with samples were then fixed in -20°C methanol for 3-5 min, washed with 1xPBS and then incubated 1 h at room temperature in 1x PBS, 0.5% Tween-20 and 1% bovine serum albumin with primary antibodies, then washed for 5 min in 1x PBS, and incubated 1 h at room temperature with secondary antibodies. Primary antibodies were rabbit anti-TaSAS-6_1 (1:300), rat anti-TaSAS-6_2 (1:100), mouse monoclonal anti-acetylated tubulin antibody 6-11B-1 (1:1000, SigmaAldrich), mouse anti-centrin-2 (1/1000, 20H5) and rabbit anti-acetylated tubulin (1:800, Abcam). Secondary antibodies were goat anti-rabbit coupled to Alexa 568, goat anti-mouse coupled to Alexa 488 and goat anti-rat coupled to Cy5 (all 1:1000, Molecular Probes). Counting was done on an epifluorescence microscope, with a 100x oil objective and imaging performed on a confocal fluorescence microscope LSM700 (Zeiss), with a 63x oil objective.
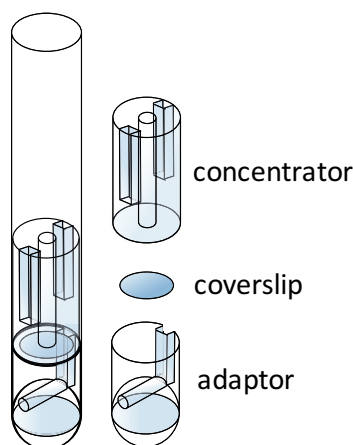


Fig. A. 1 Adaptor and concentrator to perform immunostaining. Home-made adaptors for Corex tubes to pellet samples in the centre of the coverslip. 2 ml of the buffer is pipetted onto the adapters to fill the empty spaces until the upper level of the concentrator. 5-10 μl of the sample is then resuspended inside the tunnel that delimits the surface of the coverslip to be covered by the sample. Both adaptors have a notch to facilitate the removal of the coverslip after centrifugation using a needle hook.

## 5.9 Microscopy methods and image-processing

### 5.9.1 Atomic force microscopy

Different proteins used for oligomerization assays were diluted in adsorption buffer (20 mM Tris-HCl, pH 7.2, 300 mM KCl) to reach a final concentration of 60-120 nM. 30 μl of the sample was adsorbed for 15-20 min onto a freshly cleaved Mica surface. The samples were subsequently rinsed three times with imaging buffer (20 mM Tris-HCl, pH 7.2, 150 mM KCl) and characterized by AFM in imaging buffer at room temperature (26 °C) as described previously (Pfreundschuh et al., 2014). Briefly, the AFM (Multimode 8 equipped with a Nanoscope V controller, Bruker) operated in the 'PeakForceTapping' mode was equipped with a 120 μm piezoelectric scanner. Rectangular Si3N4 cantilevers having a nominal spring constant of ~0.04–0.08 N m$^{-1}$ and a resonance frequency in water of ~35 kHz were chosen (Biolever Mini, Olympus). The recording of AFM topographs (frame size 0.5–3 μm) were done by applying a vertical oscillation frequency of 2 kHz and amplitude of 25 nm to the cantilever, an imaging force of 45–65 pN, a scanning frequency of 0.3–0.8 line s$^{-1}$. The recorded topographs were $512 \times 512$ or $1{,}024 \times 1{,}024$ pixels in size.

### 5.9.2 Electron microscopy – negative staining

Protein samples used for oligomerization studies were diluted in 10 mM K-Pipes, pH 7.0 to a final concentration of 2-30 μg/ml and incubated on glow-discharged carbon-coated 400 mesh cupper grids (Canemco Inc.) for 2-15 min. Excess protein was washed off with 10 mM K-Pipes, pH 7.0. The sample was then stained with 2% uranyl acetate for 40 s, blotted with a filter paper (Whatman n°1) to remove excess stain and dried on air (Fig. A.2). Electron micrographs were taken in a Tecnai Spirit (FEI) operated at 120 kV equipped with an Eagle CCD camera (FEI).



Fig. A. 2 Procedure of the negative staining method. The sample is applied on a hydrophobic surface (e.g. a piece of parafilm) and incubated with a glow-discharged carbon-coated cupper grid. After sample excess removal and washing, the grid is applied on a drop of 2% uranyl acetate. Excess stain is then removed by a piece of filter paper and the grid is air-dried.

### 5.9.3 Cryo-electron microscopy

The protein sample (HsSAS-6_7HR with monobody MB01), after 48 h incubation at 4°C, was applied on a holey carbon grid (Agar scientific), blotted with filter paper (Whatman 1) and vitrified in liquid ethane using a home-made plunger. The grid was then transferred into a cryo-specimen holder (626, Gatan) at cryogenic liquid nitrogen temperature and inserted into a Tecnai F20 microscope (FEI) operating at 200 kV. Observations were made using 29,000x nominal magnification. Digital images were acquired with the direct electron detector Falcon 3 (FEI).

### 5.9.4 Cryo-electron tomography

Isolated *T. agilis* centrioles with flagella were applied on holey carbon grid (Agar scientific), blotted with filter paper (Whatman 1) and vitrified in liquid ethane using a home-made or an automatic plunging apparatus (Vitrobot, FEI). The grid was then transferred into a cryo-specimen holder (626, Gatan) at cryogenic liquid nitrogen temperature and inserted into a Tecnai F20 microscope (FEI) operating at 200 kV.

Tilt-series were collected automatically with the continuous tilt scheme from -60° to +60° using the FEI tomography software, using 29,000x nominal magnification, pixel size 3.491 Å, a defocus range from -3 μm to -5 μm and a 20 eV energy window. The maximum estimated total dose for a single tilt series was 40 e-/$Å^2$. Digital images acquired with the direct electron detector Falcon 3 (FEI) 4098 x 4098 pixels were automatically drift corrected.

Tomograms were aligned using 15 nm colloidal gold beads as fiducial markers and reconstructed by R-weighted back projection using IMOD (IMOD package, Boulder, Colorado, Kremer et al., 1996). A total of 7 tomograms with 10 centrioles without extreme flattening, with a width/height ratio < 2.0 (Fig. 2.2D) were acquired, reconstructed and used for sub-tomogram averaging.

### 5.9.5 Subtomogram averaging

Using Dynamo software (Castaño-Díez et al., 2012), subtomograms of dimensions 168 x 168 x 168 or 300 x 300 x 300 pixels, with a pixel size of 3.491 Å, were boxed out and extracted using a filament model with a fixed starting and final ends at the proximal and distal ends of a centriole, respectively. The position of the filament model defined the center of each extracted subtomogram and its approximate orientation, so that the cartwheel hub is parallel to the z-axis. A total of 872 subtomograms, with a spacing of 8.5 nm, were extracted for averaging of the cartwheel hub of *T. agilis* with the CID, and 388 and 194 subtomograms, with a spacing of 8.5 and 17 nm, respectively, were extracted for averaging of the cartwheel hub of *T. mirabilis* with the CCD. Particle alignment was done in an iterative manner, so that a starting reference was rotated around a specified range of angles and compared to each particle by the calculation of the locally normalized cross-correlation. The final particle orientation is defined by the maximum cross-correlation value between the particle and the reference and by the set of alignment parameters, comprising shifts vectors and three Euler angles. The aligned particles are then averaged and the resulting average is used as a new starting reference during the next iteration. The size of the previously published map of the cartwheel (EMD-2329) (Guichard et al., 2013) was cropped to the dimensions of the subtomograms and used as a starting reference.

After several iterations of alignment and averaging, we obtained final maps of both cartwheel hubs with or without imposing a starting reference. The maps were denoised by Gaussian filter (value 1-3) and used to produce images with ImageJ and the 3D representation using UCSF Chimera.

The overall resolution of the final averages was determined using Fourier Shell Correlation calculation in Dynamo and FSC validation server (EMBL-EBI). The local resolution was calculated using ResMap software implemented as a webtool of Scipion software (Kucukelbir et al., 2014).

# Appendix 1

**Protocol for *T. agilis* rostrum isolation**

**Materials and Reagents**

Materials and reagents needed for each step of rostrum isolation and subsequent analysis are listed below. Usual material for molecular biology is also needed (pipettes, tips, Eppendorf tubes, …)

***T. agilis* cell lysate preparation**

- *Reticulitermes speratus* termites
- Ice bucket
- Pre-cooled petri dish
- Tweezers DUMONT n°5 (Dumont, ref 0101-5-PO)
- Precooled K-PIPES 10 mM, pH 7.0
- Watch glass
- Stereoscopic microscope
- Tabletop centrifuge at 4°C
- 10% NP-40 (IGEPAL)
- 2-ml sample tubes filled with silicon beads (Roche)
- MagNA lyser (Roche)
- Benchtop centrifuge at 4°C

**Density gradient separation and immunofluorescence**

- K-PIPES 20 mM, pH 7.0
- Open-top polyallomer or polypropylene 36 ml tubes (Beckmann Coulter)
- Concentrated OptiPrep solution (Axis-Shield)
- Ultracentrifuge
- Swinging rotor SW32.Ti
- Needle
- Clamp and clamp holder
- Liquid nitrogen

**Immunofluorescence**

- 15-ml Kontes™ KIMAX™ reusable high strength open centrifuge tubes (Fisher Scientific, ref 09-500-34) – "Corex tubes"

- Adapters for coverslips (Fig. 5.1)
- 12-mm round glass coverslips
- Concentrators (Fig 5.1)
- JS-13 (Avanti) or HB6 (Beckman) centrifuge rotors
- 16-mm-diameter adapter rubber to put the Corex tubes in the rotor (Fisher Scientific, ref 05-569-11)
- Tweezers DUMONT n°5 (Dumont, ref 0101-5-PO)
- Curved-tip needle
- 12-well plates (used in cell culture)
- Methanol 100%, cooled at -20°C
- PBS 1x: Phosphate Buffered Saline
- PBST: PBS 1x, Tween20 0.05%, Bovine Serum Albumin (BSA) 1%
- Parafilm
- Mouse centrin-2 antibody 20H5 (Merck Millipore, ref 04-1624) 1:1000
- Rabbit anti-acetylated tubulin antibody (Abcam) 1:800
- Secondary antibodies anti-mouse Alexa 568 (Life technologies, ref A11004) 1:1000
- Secondary antibodies anti-rabbit Alexa 488 (Life technologies, ref A11034) 1:1000
- Mounting medium: 4% n-Propyl-Gallate; 1× PBS; 90% glycerol
- Nail polish (transparent)
- Glass slide
- Fluorescence microscope (epifluorescence, confocal, 63× oil objective, etc.)

## *T. agilis* rostrum purification

### *T. agilis* cell lysate preparation

- Collect 80-100 *R. speratus* termites and put them in the petri dish precooled on ice.
- Termites are anesthetized by cold. Be careful that the whole petri dish is in constant contact with ice, because these termites are less sensitive to cold treatment than *Zootermopsis nevadensis* and might recover if temperature rises (personal observations).
- The dissection step is well described in Guichard et al. (2015).
- Place 10-15 guts at once into a watch glass with 0.5 ml K-PIPES, pH 7.0, open them one by one under the dissection microscope and remove with tweezers or by pipetting the empty gut tissue, pieces of wood or other visible contaminants from suspension. It is not recommended to dissect more termites at once, because they undergo cell lysis in aerobic conditions after ~ 10-15 min, which might have a negative impact on the subsequent rostrum isolation steps.
- Aspirate the suspension of cells in the buffer and centrifuge 20 s at 100xg in the tabletop centrifuge to pellet large protozoan symbionts, including *T. agilis*. Smaller symbionts remain in the supernatant and are removed by careful pipetting of the supernatant.
- Resuspend the pellet in 0.5 ml of K-PIPES 10 mM, pH 7.0 supplemented with NP-40 0.5% and incubate for 1 h with occasional mixing by inverting the tube up and down.
- To determine the number of cells and assess the potential loss of cells during the wash, take a 5 μl aliquot of the suspension and compare the number of *T. agilis* cells after the dissection and after the wash.

- Repeat with other termites
- Pipette 200 μl of the cell lysate into tubes with silicone beads and deflagellate rostra using MagNA lyser, for 10 s at 6,000 rpm.
- Remove the liquid and wash twice with K-PIPES 10 mM, pH 7.0. Collect all washes.
- Centrifuge the cell lysate suspension from all rounds of dissection 3 min, at 500xg, in the benchtop centrifuge
- Remove the supernatant. The pellet is enriched in rostra and contains also large cell debris.

**Rostrum purification by density gradient centrifugation**
- Prepare 4 ml of each of OptiPrep dilutions with K-PIPES 20 mM, pH 7.0: 10%, 20%, 30%, 40%, 50% and 60%.
- Carefully pipet 3 ml of each, starting from the highest concentration into a polyallomer centrifuge tube and prepare a discontinuous gradient. Be careful not to mix different layers.
- Resuspend the cell lysate enriched in rostra in 7 ml of K-PIPES 10 mM, pH 7.0 and load onto the OptiPrep gradient.
- Prepare a balance, i.e. the same tube with the same amount of the OptiPrep solution with buffer to balance the rotor of the ultracentrifuge
- Centrifuge in the ultracentrifuge Centrikon T-1080 at 25,000 rpm, at 4°C for 2 h.
- Carefully remove the tube and punch a hole in the middle of the bottom curvature and fractionate into 20 1.0 ml fractions.
- Mix well and take 5 ul aliquot from each fraction for subsequent immunofluorescence analysis.
- Flash freeze individual fractions in liquid nitrogen and store at -20°C or -80°C until the next analysis.

# Appendix 2

## Full sequence alignment of 17 homologs of SAS-6 from nine species.

SAS-6 sequences of nine usual model organisms used for centriolar studies. Five species with one copy of SAS-6 protein and four species where a gene duplication occured with two to four SAS-6 protein sequences, including *T. agilis*. Sequences were aligned in Jalview2.10 using the Mafft alignment, coloured by Clustal X colour scheme. Phylogenetic tree of the SAS-6 sequence is depicted in Fig. 4.1A.

Multiple sequence alignment (positions 386–589)

| Sequence | Start | Alignment | End |
|---|---|---|---|
| Tetrahymena-termophila1/1-703 | 386 | TELTVKIESLEKQIIDKDQLISS--SKSLNSNNEDQKKMLEKELNEKNQSIDKLEKKLEQCKVEIEKGNDVIDKIREALEKKKKQKQKEAI | 475 |
| Tetrahymena_termophyla2/1-801 | 390 | TELSFVNQGMTKQIEDKENLLVN--NKDLVENFNNQKDLLKQQCDEYRNQILKLEQKLDLCSEEINKGNKIIEKLQEELQKSKNKIKMKNQM | 479 |
| Paramecium_tetraurelia1/1-591 | 369 | IELRIQKEGLEKILRDKEDFLNN--KQQLVDNEKKQNVILEEQTNNQKKQIDKLEQRIAVMSDEVNKGNQIIEKLENELSKQKEKIKLKNAV | 458 |
| Paramecium_tetraurelia2/1-591 | 369 | IELRIQKEGFEKLIHDKEEYLNN--KQQLVDNEKKQNSILEEQVANQKKQIEKLETRISVMSEEVNKGNQIIEKLENELSKQKDKIKLKNTV | 458 |
| Paramecium_tetraurelia3/1-591 | 369 | IELRIQKEGLEKLLHDKEDFLNN--KQQLVDNEKKQNVILEEQTNNQKKQIEKLEQRIAVMSDEVNKGNQIIEKLENELSKQKEKNKLKNTV | 458 |
| Paramecium_tetraurelia4/1-466 | 370 | IELRIQKEGFEKLIHDKEEYLNN--KQQLVENEKKQNSILEEQVINQKKQIEKLESRISVMSEEVNKGNQIIEKLENELSKQKEKIKLKNTV | 459 |
| Trichomonas_vaginalis1/1-499 | 308 | SSLMNENSQLKQQIEKNSQQINQ--YVEQIRKILSESSQKDSIIEEMNHKISEYQEQLSQFNMIKVKSKEVIEKYLFQNNELKAKYSELYND | 397 |
| Trichomonas_vaginalis2/1-520 | 311 | ETLKEKITLLEKTVEHDKDVQDS--QAKIIEGLQADVAKKDEENAQIRARLSENEDKARQHDWIAAKSKQVIEKFMAEAKELNLQIAQLKSQ | 400 |
| Trichonympha_agilis1/1-531 | 309 | QSIKSKIEVLALTASHSTTFSDS--REATILDMSLQIQLKDEQISQLEDKIAQLNLKKQKDDLDADNAKFLLSDIQLQFEEFQQKRSKRKQL | 398 |
| Trichonympha_agilis2/1-532 | 311 | TSAKEKCLVYEDTMKQTKNLAMT--NGAATEGLREQLKKKDQEIEEFKKTISNLSSKAQERDWIANKSKEVIEKEKKRVEEMKKHHEAKKAE | 400 |
| Trichonympha_agilis3/1-544 | 318 | IKLESKYEV--LTQNSDDKFSQAFRKESTVEELLEEYKIKDARVSNLQEQLQQIILSSKQVAWETEHVNQQRQEIEILFDTLKEKEITQKQR | 407 |
| Trichonympha_agilis4/1-365 | 306 | LSLQGQYSI--HTNHNENNFSVALKKESTFEELIECQIKDKKLSTLQEEIQGFLLSSKQ------------------------------- | 363 |
| Chlamydomonas_rheinardtii/1-728 | 332 | NEAKAKVLALDEKAQSQGDVIEQ--QRGRLRDMEAALRQTEQRCADLRDTLASAEGRAKEAQAEVLKGNQAIEKLTNDLRLAKEKTKRKAAI | 421 |
| Homo_sapiens/1-657 | 318 | NQLQTKVAVLEQEIKDKDQLVLR--TKEAFDTIQEQKVVLEENGEKNQVQLGKLEATIKSLSAELLKANEIIKKLQGDLKTLMGKLKLKNTV | 407 |
| Caenorhabditis_elegans/1-492 | 257 | ADLKQDTESLQKQLEENQEELEI--VGNMLREEQGKVDQLQKRNVAHQKEIGKLRAELGTAQRNLEKADQLLKRNSQQ------------- | 332 |
| Danio_rerio/1-627 | 322 | NQLQTRVAVLEQEIKDKDQLVLR--TKEVLEATQQQKNSVEGNAESKQLQISKLESTVKSLSEELIKANGIIKKLQADLKALLGKIKVKNSV | 411 |
| Drosophila_melanogaster/1-472 | 322 | EDLKQHTVELQEHIQKGNK-----EKANIAAELEAEKKILHTK----------RQALEMASEEISKANQIIVKQSQELLNLKKTIAWRTEV | 397 |

Conservation
3 4 7 3 2 2 3 3 3 1 0 2 2 5 1 3 3 3 5 0 0 0 0 - - 2 2 3 3 4 2 3 3 2 2 2 3 2 2 5 4 3 3 0 0 0 0 1 0 1 2 0 0 0 2 2 3 3 3 2 0 0 0 0 1 0 1 1 1 0 0 0 0 2 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0

Quality

Consensus
IELKIKKE+LEKTI+DKE++LNN++K+QLVEN+++QKKILEQ+EN++KQISKLEQ++KQMSEEVNKGNQIIEKLENELSKLKEKIKLKN+V

Occupancy

---

| Sequence | Start | Alignment | End |
|---|---|---|---|
| Tetrahymena-termophila1/1-703 | 476 | IL---QQEENIQNLQNSIDTLNRQIIEIQREKDNSEIKVKELQTANSDLKTKIQESQKVLEQNQSMIQYLNNQLNNTRPN----------- | 552 |
| Tetrahymena_termophyla2/1-801 | 480 | VL---KQEQFLEEIRQKNDELGNEILRKEREIINNKDLKIKEQETQINDYKEKLQESCKTIESSQSMIQYLNKQLNEGGIK--------- | 556 |
| Paramecium_tetraurelia1/1-591 | 459 | VL---QQEQTVQQLQDANDQSSKQINE---QKDSLELKYKDQENTNLDLKNKLAESQKLLESNNQMIQYLNKCLNETKVQ---------- | 532 |
| Paramecium_tetraurelia2/1-591 | 459 | FL---QQEQTVQQLQDANDQNHKQLNE---QKDSLEIRLKDQENMNLDLKNKLAESQKLLESNNQMIQYLNKCLNETKVQ---------- | 532 |
| Paramecium_tetraurelia3/1-591 | 459 | FL---QQEQTVQQLQDANDQNNKQINE---QKDSLELRVKDQENTNLDLKNKLAESQKLLESNNQMIQYLNKCLNETKVQ---------- | 532 |
| Paramecium_tetraurelia4/1-466 | 460 | FL--QQECK--------------------------------------------------------------------------------- | 466 |
| Trichomonas_vaginalis1/1-499 | 398 | YQ---KDEEYINQLK-------QSLEKGSEILNQYQSQIQELKEKNDELSLEVSDLHNNLQVSESTIKTLEIQLEQK------------ | 464 |
| Trichomonas_vaginalis2/1-520 | 401 | LD---ERTQQMTEMQVENAHAEERIKALNDTIDRENAEMKELRTRNDALQKQIAEMKDKSATDEITIQHLNKRLNEHEL---------- | 476 |
| Trichonympha_agilis1/1-531 | 399 | MK---SKLDELKEMKKKATHLEELLKSKNFIFENLVIENNQLKNQIESNKELIQKFEERAKELRGLISTLDKELSESKKR-------- | 475 |
| Trichonympha_agilis2/1-532 | 401 | YQHVKEEKERIEQLS---FRLKKEVKTLKEEISKHGNDKLTYENKIEELRQQIQSLEETISSDRDMIDLLEKQLNSSKLR-------- | 477 |
| Trichonympha_agilis3/1-544 | 408 | LD---TKFEFQHLTLRSAELEDEIQQITSQIDAIRAENSALQVSTQKQKYFLKKYEEEAFQHQKLVRFIENEVKDSRKK-------- | 484 |
| Trichonympha_agilis4/1-365 | 364 | ------------------------------------------NS------------------------------------------- | 365 |
| Chlamydomonas_rheinardtii/1-728 | 422 | TI---RQEEELQEREQSLANATRDVGVLGQQAESLRKDVASLHSENDSLRSKLEDSKQQLQSNEQMIRWLNQQVTEAQLHASGATVAGSRYA | 510 |
| Homo_sapiens/1-657 | 408 | TI---QQEKLLAEKEEKLQKEQKELQDVGQSLRIKEQEVCKLQEQLEATVKKLEESKQLLKNNEKLITWLNKLNENQLVRKQDVLGPST-- | 494 |
| Caenorhabditis_elegans/1-492 | 333 | ---QNQQSLDMRK--LGELEADLKEKDSMVESLTETIGILRKELENEKLKAAENMDSFEKLSMENENLKEKIAHYRAQ--------RFS | 408 |
| Danio_rerio/1-627 | 412 | TV---PQEKILQETSDKLQRQQRELQDTQQRLSLKEEEAAKLKEQLEATVQKLDESREVLKTNENVITWLNKQLNENQLSRKQETVAMFE- | 498 |
| Drosophila_melanogaster/1-472 | 398 | AL---QQEKAVQAKESLLSLRENELREARITIEKLR----------------EEIPQQLQSMRNFAQGLEQKYSKQILI---------- | 457 |

Conservation
10---0100001000---00000001000---011000010000010000000000010100100000010010210010010000

Quality

Consensus
FLHVKQQEQ+LQELQDK+D+LEKELKE++QQIDSLE+EVKELENQNEDLKNKLAESQKLLESNEQMIQYLNKQLNETKLQRKQ+TVA+SR++

Occupancy

---

| Sequence | Start | Alignment | End |
|---|---|---|---|
| Tetrahymena-termophila1/1-703 | 553 | -IPTGLS---STLAGQTSSAYRPNYAATS----------------TISHT-----KYG-TY------------------AN---- | 589 |
| Tetrahymena_termophyla2/1-801 | 557 | -T-SMIS---TTAKPYFSSNYESKYATTN-------------YASNTAGITRKY--SYLDLNTESS---QQNANGVTTPYSLAKDSQTAG | 623 |
| Paramecium_tetraurelia1/1-591 | 533 | -APAQMPG-FSTLKSQTFTKY-------TS-------------PI------------------------------- | 555 |
| Paramecium_tetraurelia2/1-591 | 533 | -APAQMPG-ITSLKSQTFAKY-------TS-------------PI------------------------------- | 555 |
| Paramecium_tetraurelia3/1-591 | 533 | -APAQMPG-FSTLKSQTFTKY-------TS-------------PI------------------------------- | 555 |
| Paramecium_tetraurelia4/1-466 | | -------------------------------------------------------------------------- | |
| Trichomonas_vaginalis1/1-499 | 465 | ------------------------------------------------SSAFN------------------------- | 469 |
| Trichomonas_vaginalis2/1-520 | 477 | ------------------------------ENP-------PCDG----FSSSFKGGSQNKK------------- | 496 |
| Trichonympha_agilis1/1-531 | 476 | ------------ILLGEHIQQ------------LERA------P------FSPTYQ-YSI----------- | 498 |
| Trichonympha_agilis2/1-532 | 478 | ----------GGFLATLEEDDPLQSFDTF--GSTENP------PSFGN-TATQRFD---------- | 514 |
| Trichonympha_agilis3/1-544 | 485 | ----------ILYASQIQN----------IERTG----IIP------FTKSLS-SSMNET----------- | 513 |
| Trichonympha_agilis4/1-365 | | -------------------------------------------------------------------------- | |
| Chlamydomonas_rheinardtii/1-728 | 511 | FRPSA----------ALPGSMPPP----SVAATPAVGLLAAPGPSYAYKPGGT-----TSRYA-STLGVTGLSSGLSGAGATSVGAAV------AG | 580 |
| Homo_sapiens/1-657 | 495 | -TPPAHSS-SNTIRSGISPNL---NVVD----GRLTYPTCGIGY-PVSSAFAFQNTFP-HSIS----------AKNTSHP------- | 553 |
| Caenorhabditis_elegans/1-492 | 409 | PAPSGLPGLQTGLTNRLTPSF----------KPVLGP----HTPYGAN--LNSRTYP------------- | 448 |
| Danio_rerio/1-627 | 499 | -TPAA-------ALRSAAVPH----------NMAF-PITS-TINSKYP-LALSCV-------SSGSRSVLTS----- | 542 |
| Drosophila_melanogaster/1-472 | 458 | ----------------------------------------LKERLA------------------------------ | 463 |

Conservation
-----------------------------------------------------0---------------------------

Quality

Consensus
+APA+MPGLF+TLKSQT+PKY++N+A+TSAVG+LE+PG+++AY+PIS+++TFTSRYPGSSL++T++SSGLS++GANSV+TP+SLAKDSQTAG

Occupancy

Multiple sequence alignment

**Block 1 (positions ~590–707)**

| Sequence | Start | Alignment | End |
|---|---|---|---|
| Tetrahymena-termophila1/1-703 | 590 | ---NFGGEDRY--TQDHPRSHTPLSYSQKNNNPLQYTNLYSQ-------------SLNQGATYHLTKSP------LTKDFDYDN------ | 649 |
| Tetrahymena_termophyla2/1-801 | 624 | LNPTFGGTTNYNTVQNKP-SYTPTSQS-TNPNDLVFTSKQSNYNNYSAI-----NTQNQSSTINYNNNSSN-YLPNLSKQLSSSNNPNNYSD | 707 |
| Paramecium_tetraurelia1/1-591 | 556 | -------------------------PLQEDRSFR----------NQSVTY----------------------------- | 570 |
| Paramecium_tetraurelia2/1-591 | 556 | -------------------------PLQEDRSFR----------NQSVTY----------------------------- | 570 |
| Paramecium_tetraurelia3/1-591 | 556 | -------------------------PLQEDRSFR----------NQSVTY----------------------------- | 570 |
| Paramecium_tetraurelia4/1-466 | | --- | |
| Trichomonas_vaginalis1/1-499 | 470 | ---------------------------------------------PQLTSFGL--------- | 477 |
| Trichomonas_vaginalis2/1-520 | 497 | ---------------------------------------------LLPSLQPIA---------- | 505 |
| Trichonympha_agilis1/1-531 | 499 | -------------------------------QGAVRNRGQTSCQDYLPPI------------- | 517 |
| Trichonympha_agilis2/1-532 | 515 | ---------------------------------------------LPPL-------------- | 518 |
| Trichonympha_agilis3/1-544 | 514 | -----------------------------QKKTTRRNGPGG---YLPAL------------- | 530 |
| Trichonympha_agilis4/1-365 | | --- | |
| Chlamydomonas_rheinardtii/1-728 | 581 | GSYGAGGDSSYRPGVGGAYASSGGAGG--------------FGGSVGGGGQVGGSGSAIGAATGAGS-ALGTYRSSYAASTHP----- | 648 |
| Homo_sapiens/1-657 | 554 | ---GSGTKVQFNLQFTKPNASL----------GDV---------------QSGATISMPC--- | 585 |
| Caenorhabditis_elegans/1-492 | 449 | ---------------------------F--------RDNTTLNFQNSTI--ATPHA----- | 467 |
| Danio_rerio/1-627 | 543 | ---SNGPKVQFNPMSVKPSAA--------------EV-----------------SPAAFSQP----- | 570 |
| Drosophila_melanogaster/1-472 | | --- | |

Conservation

Quality

Consensus

+++GFGGKVQYNP+Q+KP+ASTP+S+SQ+N+NPLQEDRSFR+YNNF+++VGGGG+V+NQSVTYN+GN+S++DYLPPLT+SGAAS++P+NYSD

Occupancy

**Block 2 (positions ~650–788)**

| Sequence | Start | Alignment | End |
|---|---|---|---|
| Tetrahymena-termophila1/1-703 | 650 | ----EEKGQENHTNLNTKYTATLG--------GQRN--SNSFSVNNTISAVGEGN-----------IV | 692 |
| Tetrahymena_termophyla2/1-801 | 708 | LVSYNPNSTNNNNNISQQSFTQSGQSQNLQKYEKISKEEE--NGDDQDF-SQKYLFNNKNSNINDQKKNNFFQEDQADEPQR------VT | 788 |
| Paramecium_tetraurelia1/1-591 | 571 | --------QQNLNNVSQISNATI--------QNKFNF----------- | 591 |
| Paramecium_tetraurelia2/1-591 | 571 | --------QQNQNNMSQISNATI--------SNKFNF----------- | 591 |
| Paramecium_tetraurelia3/1-591 | 571 | --------QQNQNNVSQISNQTI--------QNKFNF----------- | 591 |
| Paramecium_tetraurelia4/1-466 | | --- | |
| Trichomonas_vaginalis1/1-499 | 478 | --------KTNQYGLNNDSSVRG-----------QLWSSLN--- | 499 |
| Trichomonas_vaginalis2/1-520 | 506 | ---------DVNGAASV------------WDSTAFF--- | 520 |
| Trichonympha_agilis1/1-531 | 518 | --------EISEDSMINKS---------SLL--- | 531 |
| Trichonympha_agilis2/1-532 | 519 | ---------NLGEESAFQTP--------EFW--- | 532 |
| Trichonympha_agilis3/1-544 | 531 | ---------EITDDQYFATP--------NLF--- | 544 |
| Trichonympha_agilis4/1-365 | | --- | |
| Chlamydomonas_rheinardtii/1-728 | 649 | ------SPGGVTAGMGSLTTP--AVNRYAAGSYSSV--------SSSLSFGGGLGAARAVDAT--------PAA------- | 699 |
| Homo_sapiens/1-657 | 586 | ----STDKENGENVGLES-------KYLKKREDSIPLRGLSQNLFSNSDHQRDGTLGALHTSSK----------PTALPSAS---- | 646 |
| Caenorhabditis_elegans/1-492 | 468 | ----FRFNSQLIADETGSVTNTP--------PAQR------- | 492 |
| Danio_rerio/1-627 | 571 | --------ANKENSEPVGLDS----------KYFERRDDSIPLRGL-----------LPSMHLNREV---------PKPLNTAAAKAT | 620 |
| Drosophila_melanogaster/1-472 | 464 | ----IPTGKENRR------------ | 472 |

Conservation

-------------------011000-------------------

Quality

Consensus

LVS++P++QENQNNVSQDSNATTPQSQ+++KY+K+++DSIPLRGL+QNLWSSNKFNFN+TL+A+H+++K+NFFQEDQADEPAAL++A+AK+T

Occupancy

**Block 3 (positions ~693–801)**

| Sequence | Start | Alignment | End |
|---|---|---|---|
| Tetrahymena-termophila1/1-703 | 693 | PIKYRDPTK-NQ----------------- | 703 |
| Tetrahymena_termophyla2/1-801 | 789 | PIKYRQPKNPNQN----------------- | 801 |
| Paramecium_tetraurelia1/1-591 | | ----------------------------- | |
| Paramecium_tetraurelia2/1-591 | | ----------------------------- | |
| Paramecium_tetraurelia3/1-591 | | ----------------------------- | |
| Paramecium_tetraurelia4/1-466 | | ----------------------------- | |
| Trichomonas_vaginalis1/1-499 | | ----------------------------- | |
| Trichomonas_vaginalis2/1-520 | | ----------------------------- | |
| Trichonympha_agilis1/1-531 | | ----------------------------- | |
| Trichonympha_agilis2/1-532 | | ----------------------------- | |
| Trichonympha_agilis3/1-544 | | ----------------------------- | |
| Trichonympha_agilis4/1-365 | | ----------------------------- | |
| Chlamydomonas_rheinardtii/1-728 | 700 | -SKYGVG-FSNAAMGVVGTVGGAGGTDAQRA | 728 |
| Homo_sapiens/1-657 | 647 | -SAYFPGQLPNS--------- | 657 |
| Caenorhabditis_elegans/1-492 | | ----------------------------- | |
| Danio_rerio/1-627 | 621 | PSAFFPG-------------- | 627 |
| Drosophila_melanogaster/1-472 | | ----------------------------- | |

Conservation

Quality

Consensus

PSKY+PG++PNQ+MGVVGTVGGAGGTDAQRA

Occupancy

# Bibliography

Adams, G.M.W., Huang, B., Piperno, G., and Luck, D.J.L. (1981). Central-pair microtubular complex of Chlamydomonas flagella: Polypeptide composition as revealed by analysis of mutants. J. Cell Biol. *91*, 69–76.

Afzelius, B.A. (1976). A human syndrome caused by immotile cilia. Science (80-. ). *193*, 317–319.

Agircan, F.G., Schiebel, E., and Mardin, B.R. (2014). Separate to operate: control of centrosome positioning and separation. Philos. Trans. R. Soc. B Biol. Sci. *369*, 2–10.

Andersen, J.S., Wilkinson, C.J., Mayor, T., Mortensen, P., Nigg, E.A., and Mann, M. (2003). Proteomic characterization of the human centrosome by protein correlation profiling. Nature *426*, 570–574.

Anisimova, M., and Gascuel, O. (2006). Approximate likelihood ratio test for branchs: A fast, accurate and powerful alternative. Syst. Biol. *55*, 539–552.

Arquint, C., and Nigg, E.A. (2016). The PLK4–STIL–SAS-6 module at the core of centriole duplication. Biochem. Soc. Trans. *44*, 1253–1263.

Arquint, C., Gabryjonczyk, A.-M., and Nigg, E.A. (2014). Centrosomes as signalling centres. Philos. Trans. R. Soc. B *369*, 1–12.

Arquint, C., Gabryjonczyk, A.M., Imseng, S., Böhm, R., Sauer, E., Hiller, S., Nigg, E.A., and Maier, T. (2015). STIL binding to Polo-box 3 of PLK4 regulates centriole duplication. Elife *4*, 1–22.

Aurrecoechea, C., Brestelli, J., Brunk, B.P., Carlton, J.M., Dommer, J., Fischer, S., Gajria, B., Gao, X., Gingle, A., Harb, G.G.O.S., et al. (2009). GiardiaDB and TrichDB: integrated genomic resources for the eukaryotic protist pathogens Giardia lamblia and Trichomonas vaginalis. Nucleic Acids Res. *37*, D526–D530.

Banterle, N., and Gönczy, P. (2017). Centriole Biogenesis : From Identifying the Characters to Understanding the Plot. Annu. Rev. Cell Dev. Biol. *33*, 23–49.

Baron, M., Norman, D.G., and Campbell, J.D. (1991). Protein modules. Trends Biochem. Sci. *16*, 13–17.

Basto, R., Lau, J., Vinogradova, T., Gardiol, A., Woods, C.G., Khodjakov, A., and Raff, J.W. (2006). Flies without Centrioles. Cell *125*, 1375–1386.

Batori, V., Koide, A., and Koide, S. (2002). Exploring the potential of the monobody scaffold: effects of loop elongation on the stability of a fibronectin type III domain. Protein Eng. *15*, 1015–1020.

Bauer, M., Cubizolles, F., Schmidt, A., and Nigg, E.A. (2016). Quantitative Analysis Of Human Centrosome Architecture By Targeted Proteomics And Fluorescence Imaging. EMBO J. *35*, 2152–2166.

Bayless, B.A., Giddings, T.H., Winey, M., and Pearson, C.G. (2012). Bld10/Cep135 stabilizes basal

bodies to resist cilia-generated forces. Mol. Biol. Cell *23*, 4820–4832.

Bazzi, H., and Anderson, K. V. (2014). Acentriolar mitosis activates a p53-dependent apoptosis pathway in the mouse embryo. Proc. Natl. Acad. Sci. *111*, E1491–E1500.

van Beneden, E. (1875). Recherches sur les Dicyemides, survivant actuels d'un embranchement des Mesozoaires. Bull. Acad. R. Med. Belg. *41*, 1160–1205.

van Beneden, E., and Neyt, A. (1887). Nouvelle recherches sur la fécondation et la division mitosique chez l'Ascaride mégalocéphale. Bull. Acad. R. Belgique *3*, 215–295.

Bornens, M. (2012). The centrosome in cells and organisms. Science (80-. ). *335*, 422–426.

Bornens, M., and Moudjou, M. (1999a). Studying the composition and function of centrosomes in vertebrates. Methods Cell Biol. *61*, 13–34.

Bornens, M., and Moudjou, M. (1999b). Studying the composition and function of centrosomes in vertebrates. Methods Cell Biol. *61*, 13–34.

Boveri, T. (1887). Ueber den Antheil des Spermatozoon an der Teilung des Eies. Sitzungsber. Ges. Morph. Physiol. München *3*, 157–164.

Boveri, T. (1901). Zellen-Studien: Heft 4, Ueber die natur der centrosomen. Jenaische Zeitschr. Naturwiss *35*, 1–220.

Boveri, T. (1914). Zur Frage der Entstehung maligner Tumoren. G. Fischer.

van Breugel, M., Hirono, M., Andreeva, A., Yanagisawa, H., Yamaguchi, S., Nakazawa, Y., Morgner, N., Petrovich, M., Ebong, I.-O., Robinson, C. V, et al. (2011). Structures of SAS-6 suggest its organization in centrioles. Science (80-. ). *331*, 1196–1199.

van Breugel, M., Wilcken, R., McLaughlin, S.H., Rutherford, T.J., and Johnson, C.M. (2014). Structure of the SAS-6 cartwheel hub from Leishmania major. Elife *3*, 1–16.

Briggs, J.A.G. (2013). Structural biology in situ — the potential of subtomogram averaging. Curr. Opin. Struct. Biol. *23*, 261–267.

Brown, N.J., Marjanović, M., Lüders, J., Stracker, T.H., and Costanzo, V. (2013). Cep63 and Cep152 Cooperate to Ensure Centriole Duplication. PLoS One *8*.

Brune, A., and Dietrich, C. (2015). The Gut Microbiota of Termites: Digesting the Diversity in the Light of Ecology and Evolution. Annu. Rev. Microbiol. *69*, 145–166.

Burakov, A. V., and Nadezhdina, E.S. (2013). Association of nucleus and centrosome: Magnet or velcro? Cell Biol. Int. *37*, 95–104.

Burakov, A., Nadezhdina, E., Slepchenko, B., and Rodionov, V. (2003). Centrosome positioning in interphase cells. J. Cell Biol. *162*, 963–969.

Burkhard, P., Kammerer, R.A., Steinmetz, M.O., Bourenkov, G.P., and Aebi, U. (2000). The coiled-coil trigger site of the rod domain of cortexillin I unveils a distinct network of interhelical and intrahelical salt bridges. Structure *8*, 223–230.

De Carlo, S., and Harris, J.R. (2011). Negative staining and cryo-negative staining of macromolecules and viruses for TEM. Micron *42*, 117–131.

Carlton, J.M., Hirt, R.P., Silva, J.C., Delcher, A.L., Schatz, M., Zhao, Q., Wortman, J.R., Bidwell, S.L., Alsmark, U.C.M., Besteiro, S., et al. (2007). Draft Genome Sequence of the Sexually Transmitted

Pathogen Trichomonas vaginalis. Science *315*, 207–212.

Carvalho-Santos, Z., Machado, P., Branco, P., Tavares-Cadete, F., Rodrigues-Martins, A., Pereira-Leal, J.B., and Bettencourt-Dias, M. (2010). Stepwise evolution of the centriole-assembly pathway. J. Cell Sci. *123*, 1414–1426.

Carvalho-Santos, Z., Azimzadeh, J., Pereira-Leal, J.B., and Bettencourt-Dias, M. (2011). Tracing the origins of centrioles, cilia, and flagella. J. Cell Biol. *194*, 165–175.

Carvalho-Santos, Z., Machado, P., Alvarez-Martins, I., Gouveia, S.M., Jana, S.C., Duarte, P., Amado, T., Branco, P., Freitas, M.C., Silva, S.T.N., et al. (2012). BLD10/CEP135 Is a Microtubule-Associated Protein that Controls the Formation of the Flagellum Central Microtubule Pair. Dev. Cell *23*, 412–424.

Cassimeris, L.U., Wadsworth, P., and Salmon, E.D. (1986). Dynamics of microtubule depolymerisation in monocytes. J. Cell Biol. *102*, 2023–2032.

Castaño-Díez, D., Kudryashev, M., Arheit, M., and Stahlberg, H. (2012). Dynamo: A flexible, user-friendly development tool for subtomogram averaging of cryo-EM data in high-performance computing environments. J. Struct. Biol. *178*, 139–151.

Castresana, J. (2000). Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol. Biol. Evol. *17*, 540–552.

Chang, P., and Stearns, T. (2000). Delta-tubulin and epsilon-tubulin: two new human centrosomal tubulins reveal new aspects of centrosome structure and function. Nat. Cell Biol. *2*, 30–35.

Chevenet, F., Brun, C., Banuls, A.L., Jacq, B., and Chisten, R. (2006). TreeDyn: towards dynamic graphics and annotations for analyses of trees. BMC Bioinformatics *10*, 439.

Choi, Y.-K., Liu, P., Sze, S.K., Dai, C., and Qi, R.Z. (2010). CDK5RAP2 stimulates microtubule nucleation by the γ-tubulin ring complex. J. Cell Biol. *191*, 1089–1095.

Chrétien, D., Buendia, B., Fuller, S.D., and Karsenti, E. (1997). Reconstruction of the centrosome cycle from cryoelectron micrographs. J. Struct. Biol. *120*, 117–133.

Cleveland, L.R. (1938). Morphology and mitosis of Teranympha. Arch. Protistenk. *91*, 441–451.

Comartin, D., Gupta, G.D., Fussner, E., Coyaud, É., Hasegan, M., Archinti, M., Cheung, S.W.T., Pinchev, D., Lawo, S., Raught, B., et al. (2013). CEP120 and SPICE1 cooperate with CPAP in centriole elongation. Curr. Biol. *23*, 1360–1366.

Conduit, P.T., Brunk, K., Dobbelaere, J., Dix, C.I., Lucas, E.P., and Raff, J.W. (2010). Centrioles regulate centrosome size by controlling the rate of Cnn incorporation into the PCM. Curr. Biol. *20*, 2178–2186.

Conduit, P.T., Wainman, A., and Raff, J.W. (2015). Centrosome function and assembly in animal cells. Nat. Rev. Mol. Cell Biol. *16*, 611–624.

Conte, D., MacNeil, L.T., Walhout, A.J.M., and Mello, C.C. (2015). RNA Interference in *Caenorhabditis elegans*. Curr. Protoc. Mol. Biol. 26.3.1-26.3.30.

Cota, R.R., Teixidó-Travesa, N., Ezquerra, A., Eibes, S., Lacasa, C., Roig, J., and Lüders, J. (2017). MZT1 regulates microtubule nucleation by linking γTuRC assembly to adapter-mediated targeting and activation. J. Cell Sci. *130*, 406–419.

Cottee, M. a, Muschalik, N., Wong, Y.L., Johnson, C.M., Johnson, S., Andreeva, A., Oegema, K., Lea,

S.M., Raff, J.W., and van Breugel, M. (2013). Crystal structures of the CPAP/STIL complex reveal its role in centriole assembly and human microcephaly. Elife *2*, e01071.

Cottee, M.A., Muschalik, N., Johnson, S., Leveson, J., Raff, J.W., and Lea, S.M. (2015). The homo-oligomerisation of both Sas-6 and Ana2 is required for efficient centriole assembly in flies. Elife *10*, 1–25.

Courtois, A., Schuh, M., Ellenberg, J., and Hiiragi, T. (2012). The transition from meiotic to mitotic spindle assembly is gradual during early mammalian development. J. Cell Biol. *198*, 357–370.

Dammermann, A., Müller-Reichert, T., Pelletier, L., Habermann, B., Desai, A., and Oegema, K. (2004). Centriole assembly requires both centriolar and pericentriolar material proteins. Dev. Cell *7*, 815–829.

David, A., Amartely, H., Rabinowicz, N., Shamir, M., Friedler, A., and Izraeli, S. (2016). Molecular basis of the STIL coiled coil oligomerization explains its requirement for de-novo formation of centrosomes in mammalian cells. Sci. Rep. *6*, 1–10.

Decker, M., Jaensch, S., Pozniakovsky, A., Zinke, A., Connell, K.F.O., Zachariae, W., Myers, E., and Hyman, A.A. (2011). Limiting Amounts of Centrosome Material Set Centrosome Size in C . elegans Embryos. Curr. Biol. *21*, 1259–1267.

Delattre, M., Leidel, S., Wani, K., Baumer, K., Bamat, J., Schnabel, H., Feichtinger, R., Schnabel, R., and Gönczy, P. (2004). Centriolar SAS-5 is required for centrosome duplication in C. elegans. Nat. Cell Biol. *6*, 656–664.

Dereeper, A., Guignon, V., Blanc, G., Audic, S., Buffet, S., Chevenet, F., Dufayard, J.F., Guindon, S., Lefort, V., Lescot, M., et al. (2008). Phylogeny.fr: robust phylogenetic analysis for the non-specialist. Nucleic Acids Res. *36*, 465–469.

Dereeper, A., Audic, S., Claverie, J.M., and Blanc, G. (2010). BLAST-EXPLORER helps you building datasets for phylogenetic analysis. BMC Evol. Biol. *10*.

Dessau, M.A., and Modis, Y. (2011). Protein Crystallization for X-ray Crystallography. J. Vis. Exp. *9*, 1–6.

Dictenberg, J.B., Zimmerman, W., Sparks, C.A., Young, A., Vidair, C., Zheng, Y., Carrington, W., Fay, F.S., and Doxsey, S.J. (1998). Pericentrin and γ-tubulin form a protein complex and are organized into a novel lattice at the centrosome. J. Cell Biol. *141*, 163–174.

Diebolder, C.A., Koster, A.J., and Koning, R.I. (2012). Pushing the resolution limits in cryo electron tomography of biological structures. J. Microsc. *248*, 1–5.

Dorfer, V., Pichler, P., Stranzl, T., Stadlmann, J., Taus, T., Winkler, S., and Mechtler, K. (2014). MS Amanda, a Universal Identi fi cation Algorithm Optimized for High Accuracy Tandem Mass Spectra. J. Proteome Res. *13*, 3679–3684.

Doxsey, S., McCollum, D., and Theurkauf, W. (2005). Centrosomes in Cellular Regulation. Annu. Rev. Cell Dev. Biol. *21*, 411–434.

Dumont, J., and Desai, A. (2012). Acentrosomal spindle assembly and chromosome segregation during oocyte meiosis. Trends Cell Biol. *22*, 241–249.

Dutcher, S.K., and Trabuco, E.C. (1998). The UNI3 gene is required for assembly of basel bodies of Chlamydomonas and encodes d-tubulin, a new member of the tubulin superfamily. Mol. Biol. Cell *9*, 1293–1308.

Dzhindzhev, N.S., Yu, Q.D., Weiskopf, K., Tzolovsky, G., Cunha-Ferreira, I., Riparbelli, M., Rodrigues-Martins, A., Bettencourt-Dias, M., Callaini, G., and Glover, D.M. (2010). Asterless is a scaffold for the onset of centriole assembly. Nature *467*, 714–718.

Edgar, R.C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. *32*, 1792–1797.

Einstein, E.B., Patterson, C.A., Hon, B.J., Regan, K.A., Reddi, J., Melnikoff, D.E., Mateer, M.J., Schulz, S., Johnson, B.N., and Tallent, M.K. (2010). Somatostatin Signaling in Neuronal Cilia Is Critical for Object Recognition Memory. J. Neurosci. *30*, 4306–4314.

Elmlund, D., and Elmlund, H. (2015). Cryogenic Electron Microscopy and Single-Particle Analysis. Annu. Rev. Biochem. *84*, 499–517.

Evans, L., Mitchison, T., and Kirschner, M. (1985). Influence of the centrosome on the structure of nucleated microtubules. J. Cell Biol. *100*, 1185–1191.

Faix, J., Steinmetz, M., Boves, H., Kammerer, R.A., Lottspeich, F., Mintert, U., Murphy, J., Stock, A., Aebi, U., and Gerisch, G. (1996). Cortexillins, major determinants of cell shape and size, are actin-bundling proteins with a parallel coiled-coil tail. Cell *86*, 631–642.

Fawcett, D.W., and Porter, K.R. (1954). A study of the fine structure of ciliated epithelia. J. Morphol. *94*, 221–281.

Fernandez-Leiro, R., and Scheres, S.H.W. (2016). Unravelling biological macromolecules with cryo-electron microscopy. Nature *537*, 339–346.

Firat-Karalar, E.N., and Stearns, T. (2014). The centriole duplication cycle. Philos. Trans. R. Soc. Lond. B. Biol. Sci. *369*, 1–10.

Firat-Karalar, E.N., Sante, J., Elliott, S., and Stearns, T. (2014). Proteomic Analysis Of Mammalian Sperm Cells Identifies New Components Of The Centrosome. J. Cell Sci. *127*, 4128–4133.

Flemming, W. (1875). Studien in der Entwicklungsgeschichte der Najaden. Sitzungsberichte Der K. Akad. d. Wissensch, K. K. Staatsdruckerei.

Fu, J., and Glover, D.M. (2012). Structured illumination of the interface between centriole and peri-centriolar material. Open Biol. *2*, 120104–120104.

Gartenmann, L., Wainman, A., Qurashi, M., Kaufmann, R., Schubert, S., Raff, J.W., and Dobbie, I.M. (2017). A combined 3D-SIM/SMLM approach allows centriole proteins to be localized with a precision of ~4–5 nm. Curr. Biol. *27*, R1054–R1055.

Gauthier, D.L., Sobota, J.A., Ferraro, F., Mains, R.E., and Lazure, C. (2008). Flow cytometry-assisted purification and proteomic analysis of the corticotropes dense-core secretory granules. Proteomics *8*, 3848–3861.

Geimer, S., and Melkonian, M. (2004). The ultrastructure of the Chlamydomonas reinhardtii basal apparatus: identification of an early marker of radial asymmetry inherent in the basal body. J. Cell Sci. *117*, 2663–2674.

Giansanti, M.G., Bucciarelli, E., Bonaccorsi, S., and Gatti, M. (2008). Drosophila SPD-2 Is an Essential Centriole Component Required for PCM Recruitment and Astral-Microtubule Nucleation. Curr. Biol. *18*, 303–309.

Gibbons, I.R., and Grimstone, A. V. (1960). On flagellar structure in certain flagellates. J. Biophys.

Biochem. Cytol. *7*, 697–716.

Glaeser, R.M. (2016). Specimen Behavior in the Electron Beam. In The Resolution Revolution: Recent Advances In CryoEM, (Elsevier Inc.), pp. 19–50.

Goetz, S.C., Ocbina, P.J.R., and Anderson, K. V. (2009). The primary cilium as a Hedgehog signal transduction machine. (Elsevier).

Gomez-Ferreria, M.A., Rath, U., Buster, D.W., Chanda, S.K., Caldwell, J.S., Rines, D.R., and Sharp, D.J. (2007). Human Cep192 Is Required for Mitotic Centrosome and Spindle Assembly. Curr. Biol. *17*, 1960–1966.

Gonçalves, J., and Pelletier, L. (2017). The Ciliary Transition Zone: Finding the Pieces and Assembling the Gate. Mol. Cells *40*, 243–253.

Gönczy, P. (2008). Mechanisms of asymmetric cell division: Flies and worms pave the way. Nat. Rev. Mol. Cell Biol. *9*, 355–366.

Gönczy, P. (2012). Towards a molecular architecture of centriole assembly. Nat. Rev. Mol. Cell Biol. *13*, 425–435.

Gönczy, P. (2015). Centrosomes and cancer: revisiting a long-standing relationship. Nat Rev Cancer *15*, 639–652.

Gönczy, P., and Rose, L.S. (2005). Asymmetric cell division and axis formation in the embryo. In WormBook, Ed. The C. Elegans Research Community, pp. 1–20.

Gonzalez, C. (2008). Centrosome function during stem cell division: the devil is in the details. Curr. Opin. Cell Biol. *20*, 694–698.

Graser, S., Stierhof, Y.D., Lavoie, S.B., Gassner, O.S., Lamla, S., Le Clech, M., and Nigg, E.A. (2007). Cep164, a novel centriole appendage protein required for primary cilium formation. J. Cell Biol. *179*, 321–330.

Grimstone, A. V., and Gibbons, I.R. (1966). The fine structure of the centriolar apparatus and associated structures in the complex flagellates Trichonympha and Pseudotrichonympha. Philos. Trans. R. Soc. London - Ser. B Biol. Sci. *250*, 215–242.

Gromley, A., Yeaman, C., Rosa, J., Redick, S., Chen, C.T., Mirabelle, S., Guha, M., Sillibourne, J., and Doxsey, S.J. (2005). Centriolin anchoring of exocyst and SNARE complexes at the midbody is required for secretory-vesicle-mediated abscission. Cell *123*, 75–87.

Guderian, G., Westendorf, J., Uldschmid, A., and Nigg, E.A. (2010). Plk4 trans-autophosphorylation regulates centriole number by controlling TrCP-mediated degradation. J. Cell Sci. *123*, 2163–2169.

Guichard, P., and Gönczy, P. (2016). Basal body structure in Trichonympha. Cilia *5*, 9–11.

Guichard, P., Chrétien, D., Marco, S., and Tassin, A.-M. (2010). Procentriole assembly revealed by cryo-electron tomography. EMBO J. *29*, 1565–1572.

Guichard, P., Desfosses, A., Maheshwari, A., Hachet, V., Dietrich, C., Brune, A., Ishikawa, T., Sachse, C., and Gönczy, P. (2012). Cartwheel Architecture of Trichonympha Basal Body. Science (80-. ). *337*, 553.

Guichard, P., Hachet, V., Majubu, N., Neves, A., Demurtas, D., Olieric, N., Fluckiger, I., Yamada, A., Kihara, K., Nishida, Y., et al. (2013). Native architecture of the centriole proximal region reveals features underlying its 9-fold radial symmetry. Curr. Biol. *23*, 1620–1628.

Guichard, P., Hamel, V., Neves, A., and Gönczy, P. (2015). Isolation, cryotomography, and three-dimensional reconstruction of centrioles. Methods Cell Biol. *129*, 191–209.

Guichard, P., Hamel, V., Guennec, M. Le, Banterle, N., Iacovache, I., Nemcikova, V., Fluckiger, I., Goldie, K.N., Stahlberg, H., Levy, D., et al. (2017). Cell-free reconstitution reveals centriole cartwheel assembly mechanisms. Nat. Commun. *8*, 1–9.

Guindon, S., and Gascuel, O. (2003). A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst. Biol. *52*, 696–704.

Gupta, A., and Kitagawa, D. (2018). Ultrastructural diversity between centrioles of eukaryotes. J. Biochem. *mvy031*.

Hachet, V., Canard, C., and Gönczy, P. (2007). Centrosomes Promote Timely Mitotic Entry in C. elegans Embryos. Dev. Cell *12*, 531–541.

Hamel, V., Guichard, P., Fournier, M., Guiet, R., Flückiger, I., Seitz, A., and Gönczy, P. (2014). Correlative multicolor 3D SIM and STORM microscopy. Biomed. Opt. Express *5*, 3326.

Hamel, V., Steib, E., Hamelin, R., Armand, F., Borgers, S., Flückiger, I., Busso, C., Olieric, N., Sorzano, C.O.S., Steinmetz, M.O., et al. (2017). Identification of Chlamydomonas Central Core Centriolar Proteins Reveals a Role for Human WDR90 in Ciliogenesis. Curr. Biol. *27*, 2486–2498.e6.

Hamill, D.R., Severson, A.F., Carter, J.C., and Bowerman, B. (2002). Centrosome maturation and mitotic spindle assembly in C. elegans require SPD-5, a protein with multiple coiled-coil domains. Dev. Cell *3*, 673–684.

Harris, E.H. (2001). Chlamydomonas as a model organism. Annu. Rev. Plant Physiol. Plant Mol. Biol. *52*, 363–406.

de Harven, E., and Bernhard, W. (1956). Etude au microscope electronique de l'ultrastructure du centriole chez les vertebres. Zeitschrift Fiir Zellforsch. *45*, 378–398.

Hashimoto, T. (2013). A ring for all: γ-tubulin-containing nucleation complexes in acentrosomal plant microtubule arrays. Curr. Opin. Plant Biol. *16*, 698–703.

Hayashi, Y., Shigenobu, S., Watanabe, D., Toga, K., Saiki, R., Shimada, K., Bourguignon, T., Lo, N., Hojo, M., Maekawa, K., et al. (2013). Construction and Characterization of Normalized cDNA Libraries by 454 Pyrosequencing and Estimation of DNA Methylation Levels in Three Distantly Related Termite Species. PLoS One *8*, 1–12.

Heald, R., Tournebize, R., Habermann, A., Karsenti, E., and Hyman, A. (1997). Spindle assembly in Xenopus egg extracts: respective roles of centrosomes and microtubule self-organization. J. Cell Biol. *138*, 615–628.

Hilbert, M., Erat, M.C., Hachet, V., Guichard, P., Blank, I.D., Flückiger, I., Slater, L., Lowe, E.D., Hatzopoulos, G.N., Steinmetz, M.O., et al. (2013). Caenorhabditis elegans centriolar protein SAS-6 forms a spiral that is consistent with imparting a ninefold symmetry. Proc. Natl. Acad. Sci. U. S. A. *110*, 11373–11378.

Hilbert, M., Noga, A., Frey, D., Hamel, V., Guichard, P., Kraatz, S.H.W., Pfreundschuh, M., Hosner, S., Flückiger, I., Jaussi, R., et al. (2016). SAS-6 engineering reveals interdependence between cartwheel and microtubules in determining centriole architecture. Nat. Cell Biol. *18*, 393–403.

Hinchcliffe, E.H., Miller, F.J., Cham, M., Khodjakov, A., and Greenfield, S. (2001). Requirement of a Centrosomal Activity for Cell Cycle Progression Through G 1 into S Phase. Science (80-. ). *291*,

1547–1551.

Hiraki, M., Nakazawa, Y., Kamiya, R., and Hirono, M. (2007). Bld10p constitutes the cartwheel-spoke tip and stabilizes the 9-fold symmetry of the centriole. Curr. Biol. *17*, 1778–1783.

Hirono, M. (2014). Cartwheel assembly. Philos. Trans. R. Soc. Lond. B. Biol. Sci. *369*.

Hodges, M.E., Scheumann, N., Wickstead, B., Langdale, J. a, and Gull, K. (2010). Reconstructing the evolutionary history of the centriole from protein components. J. Cell Sci. *123*, 1407–1413.

Holland, A.J., Lan, W., Niessen, S., Hoover, H., and Cleveland, D.W. (2010). Polo-like kinase 4 kinase activity limits centrosome overduplication by autoregulating its own stability. J. Cell Biol. *188*, 191–198.

Huvenne, H., and Smagghe, G. (2010). Mechanisms of dsRNA uptake in insects and potential of RNAi for pest control: A review. J. Insect Physiol. *56*, 227–235.

Ikeda-Ohtsubo, W., and Brune, A. (2009). Cospeciation of termite gut flagellates and their bacterial endosymbionts: Trichonympha species and "Candidatus Endomicrobium trichonymphae." Mol. Ecol. *18*, 332–342.

Ishikawa, H., Kubo, A., Tsukita, S., and Tsukita, S. (2005). Odf2-deficient mother centrioles lack distal/subdistal appendages and the ability to generate primary cilia. Nat. Cell Biol. *7*, 517–524.

Itakura, S., Murayama, S., Kamata, Y., Tanaka, H., and Enoki, A. (2009). RNA Interference in Symbiotic Protists of the Termite Coptotermes formosanus (Isoptera: Rhinotermitidae) Through Ingestion of siRNA by the Host Termite. Sociobiology *54*.

Jakobsen, L., Schrøder, J.M., Larsen, K.M., Lundberg, E., and Andersen, J.S. (2013). Centrosome isolation and analysis by mass spectrometry-based proteomics (Elsevier Inc.).

Joshi, H.C., Palacios, M.J., McNamara, L., and Cleveland, D.W. (1992). γ-Tubulin is a centrosomal protein required for cell cycle-dependent microtubule nucleation. Nature *356*, 80–83.

Joukov, V., De Nicolo, A., Rodriguez, A., Walter, J.C., and Livingston, D.M. (2010). Centrosomal protein of 192 kDa (Cep192) promotes centrosome-driven spindle assembly by engaging in organelle-specific Aurora A activation. Proc. Natl. Acad. Sci. *107*, 21022–21027.

Keating, T.J., and Borisy, G.G. (2000). Immunostructural evidence for the template mechanism of microtubule nucleation. Nat. Cell Biol. *2*, 352–357.

Keeling, P.J. (2007). Deep Questions in the Tree of Life. Science (80-. ). *317*, 1875–1876.

Keller, D., Orpinell, M., Olivier, N., Wachsmuth, M., Mahen, R., Wyss, R., Hachet, V., Ellenberg, J., Manley, S., and Gönczy, P. (2014). Mechanisms of HsSAS-6 assembly promoting centriole formation in human cells. J. Cell Biol. *204*, 697–712.

Keller, L.C., Romijn, E.P., Zamora, I, Yates, J.R., and Marshall, W.F. (2005). Proteomic analysis of isolated Chlamydomonas centrioles reveals orthologs of ciliary-disease genes. Curr. Biol. *15*, 1090–1098.

Keller, L.C., Geimer, S., Romijn, E., Yates, J., Zamora, I., and Marshall, W.F. (2009). Molecular Architecture of the centriole Proteome: The Conserved WD40 Domain Protein POC1 Is Required for Centriole Duplication and Length Control. Mol. Biol. Cell *20*, 1150–1166.

Kelley, L.A., Mezulis, S., Yates, C.M., Wass, M.N., and Sternberg, M.J.E. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. Nat. Protoc. *10*, 845–858.

Kellogg, D.R., Moritz, M., and Alberts, B.M. (1994). The Centrosome and Cellular Organization. Annu. Rev. Biochem. *63*, 639–674.

Kemp, C.A., Kopish, K.R., Zipperlen, P., Ahringer, J., and O'Connell, K.F. (2004). Centrosome maturation and duplication in C. elegans require the coiled-coil protein SPD-2. Dev. Cell *6*, 511–523.

Khan, M.A., Rupp, V.M., Orpinell, M., Hussain, M.S., Altmüller, J., Steinmetz, M.O., Enzinger, C., Thiele, H., Höhne, W., Nürnberg, G., et al. (2014). A missense mutation in the PISA domain of HsSAS-6 causes autosomal recessive primary microcephaly in a large consanguineous pakistani family. Hum. Mol. Genet. *23*, 5940–5949.

Khodjakov, A., and Rieder, C.L. (2001). Centrosomes enhance the fidelity of cytokinesis in vertebrates and are required for cell cycle progression. J. Cell Biol. *153*, 237–242.

Khodjakov, A., Cole, R.W., Oakley, B.R., and Rieder, C.L. (2000). Centrosome-independent mitotic spindle formation in vertebrates. Curr. Biol. *10*, 59–67.

Kilburn, C.L., Pearson, C.G., Romijn, E.P., Meehl, J.B., Giddings, T.H., Culver, B.P., Yates, J.R., and Winey, M. (2007). New Tetrahymena basal body protein components identify basal body domain structure. J. Cell Biol. *178*, 905–912.

Kim, M., O'Rourke, B.P., Soni, R.K., Jallepalli, P. V., Hendrickson, R.C., and Tsou, M.F.B. (2016). Promotion and Suppression of Centriole Duplication Are Catalytically Coupled through PLK4 to Ensure Centriole Homeostasis. Cell Rep. *16*, 1195–1203.

Kim, T.-S., Park, J.-E., Shukla, A., Choi, S., Murugan, R.N., Lee, J.H., Ahn, M., Rhee, K., Bang, J.K., Kim, B.Y., et al. (2013). Hierarchical recruitment of Plk4 and regulation of centriole biogenesis by two centrosomal scaffolds, Cep192 and Cep152. Proc. Natl. Acad. Sci. *110*, E4849–E4857.

Kirby, H. (1932). Flagellates of the genus Trichonympha in termites. Univ. Calif. Publ. Zool. *37*, 349–476.

Kirkham, M., Müller-Reichert, T., Oegema, K., Grill, S., and Hyman, A.A. (2003). SAS-4 is a C. elegans centriolar protein that controls centrosome size. Cell *112*, 575–587.

Kitagawa, D., Busso, C., Flückiger, I., and Gönczy, P. (2009). Phosphorylation of SAS-6 by ZYG-1 is critical for centriole formation in C. elegans embryos. Dev. Cell *17*, 900–907.

Kitagawa, D., Vakonakis, I., Olieric, N., Hilbert, M., Keller, D., Olieric, V., Bortfeld, M., Erat, M.C., Flückiger, I., Gönczy, P., et al. (2011). Structural basis of the 9-fold symmetry of centrioles. Cell *144*, 364–375.

Kleylein-Sohn, J., Westendorf, J., Le Clech, M., Habedanck, R., Stierhof, Y.D., and Nigg, E.A. (2007). Plk4-Induced Centriole Biogenesis in Human Cells. Dev. Cell *13*, 190–202.

Kohlmaier, G., Lončarek, J., Meng, X., McEwen, B.F., Mogensen, M.M., Spektor, A., Dynlacht, B.D., Khodjakov, A., and Gönczy, P. (2009). Overly Long Centrioles and Defective Cell Division upon Excess of the SAS-4-Related Protein CPAP. Curr. Biol. *19*, 1012–1018.

Koide, S. (2009). Engineering of recombinant crystallization chaperones. Curr. Opin. Struct. Biol. *19*, 449–457.

Koide, A., Bailey, C.W., Huang, X., and Koide, S. (1998). The fibronectin type III domain as a scaffold for novel binding proteins. J. Mol. Biol. *284*, 1141–1151.

Koide, A., Wojcik, J., Gilbreth, R.N., Hoey, R.J., and Koide, S. (2012). Teaching an old scaffold new

tricks: Monobodies constructed using alternative surfaces of the FN3 scaffold. J. Mol. Biol. *415*, 393–405.

Kollman, J.M., Polka, J.K., Zelter, A., Davis, T.N., and Agard, D.A. (2010). Microtubule nucleating γTuSC assembles structures with 13-fold microtubule-like symmetry. Nature *466*, 879–882.

Kraatz, S., Guichard, P., Obbineni, J.M., Olieric, N., Hatzopoulos, G.N., Hilbert, M., Sen, I., Missimer, J., Gönczy, P., and Steinmetz, M.O. (2016). The Human Centriolar Protein CEP135 Contains a Two-Stranded Coiled-Coil Domain Critical for Microtubule Binding. Structure *24*, 1358–1371.

Krämer, A., Mailand, N., Lukas, C., Syljuåsen, R.G., Wilkinson, C.J., Nigg, E.A., Bartek, J., and Lukas, J. (2004). Centrosome-associated Chk1 prevents premature activation of cyclin-B-Cdk1 kinase. Nat. Cell Biol. *6*, 884–891.

Kratz, A.-S., Barenz, F., Richter, K.T., and Hoffmann, I. (2015). Plk4-dependent phosphorylation of STIL is required for centriole duplication. Biol. Open *4*, 370–377.

Kremer, J.R., Mastronarde, D.N., and McIntosh, J.R. (1996). Computer visualization of Three-Dimensional Image Data Using IMOD. J. Struct. Biol. *116*, 71–76.

Kubai, D.F. (1973). Unorthodox mitosis in Trichonympha agilis: kinetochore differentiation and chromosome movement. J. Cell Sci. *13*, 511–552.

Kucukelbir, A., Sigworth, F.J., and Tagare, H.D. (2014). Quantifying the Local Resolution of Cryo-EM Density Maps. Nat. Methods *11*, 63–65.

Kutscher, L.M., and Shaham, S. (2014). Forward and reverse mutagenesis in C. elegans. In WormBook, Ed. The C. Elegans Research Community, pp. 1–30.

Lau, L., Lee, Y.L., Sahl, S.J., Stearns, T., and Moerner, W.E. (2012). STED microscopy with optimized labeling density reveals 9-fold arrangement of a centriole protein. Biophys. J. *102*, 2926–2935.

Lawo, S., Hasegan, M., Gupta, G.D., and Pelletier, L. (2012). Subdiffraction imaging of centrosomes reveals higher-order organizational features of pericentriolar material. Nat. Cell Biol. *14*, 1148–1158.

Lee, K., and Rhee, K. (2011). PLK1 phosphorylation of pericentrin initiates centrosome maturation at the onset of mitosis. *195*, 1093–1101.

Lefebvre, P.A., Nordstrom, S.A., Moulder, J.E., and Rosenbaum, J.L. (1978). Flagellar elongation and shortening in Chlamydomonas IV. Effects of Flagellar Detachment, Regeneration, and Resorption on the Induction of Flagellar Protein Synthesis. J. Cell Biol. 8–27.

Leidel, S., and Gönczy, P. (2003). SAS-4 is essential for centrosome duplication in C. elegans and is recruited to daughter centrioles once per cell cycle. Dev. Cell *4*, 431–439.

Leidel, S., Delattre, M., Cerutti, L., Baumer, K., and Gönczy, P. (2005). SAS-6 defines a protein family required for centrosome duplication in C. elegans and in human cells. Nat. Cell Biol. *7*, 115–125.

Li, J.B., Gerdes, J.M., Haycraft, C.J., Fan, Y., Teslovich, T.M., May-Simera, H., Li, H., Blacque, O.E., Li, L., Leitch, C.C., et al. (2004). Comparative Genomics Identifies a Flagellar and Basal Body Proteome that Includes the BBS5 Human Disease Gene. Cell *117*, 541–552.

Li, S., Fernandez, J.-J., Marshall, W.F., and Agard, D. a (2012). Three-dimensional structure of basal body triplet revealed by electron cryo-tomography. EMBO J. *31*, 552–562.

Lin, Y.-C., Chang, C.-W., Hsu, W.-B., Tang, C.-J.C., Lin, Y.-N., Chou, E.-J., Wu, C.-T., and Tang,

T.K. (2013a). Human microcephaly protein CEP135 binds to hSAS-6 and CPAP, and is required for centriole assembly. EMBO J. *32*, 1141–1154.

Lin, Y.C., Chang, C.W., Hsu, W. Bin, Tang, C.J.C., Lin, Y.N., Chou, E.J., Wu, C.T., and Tang, T.K. (2013b). Human microcephaly protein CEP135 binds to hSAS-6 and CPAP, and is required for centriole assembly. EMBO J. *32*, 1141–1154.

Lin, Y.N., Wu, C.T., Lin, Y.C., Hsu, W. Bin, Tang, C.J.C., Chang, C.W., and Tang, T.K. (2013c). CEP120 interacts with CPAP and positively regulates centriole elongation. J. Cell Biol. *202*, 211–219.

Liu, X.J., Xie, L., Liu, N., Zhan, S., Zhou, X.G., and Wang, Q. (2017). RNA interference unveils the importance of Pseudotrichonympha grassii cellobiohydrolase, a protozoan exoglucanase, in termite cellulose degradation. Insect Mol. Biol. *26*, 233–242.

Lucas, E.P., and Raff, J.W. (2007). Maintaining the proper connection between the centrioles and the pericentriolar matrix requires Drosophila Centrosomin. J. Cell Biol. *178*, 725–732.

Lukinavičius, G., Lavogina, D., Orpinell, M., Umezawa, K., Reymond, L., Garin, N., Gönczy, P., and Johnsson, K. (2013a). Selective chemical crosslinking reveals a Cep57-Cep63-Cep152 centrosomal complex. Curr. Biol. *23*, 265–270.

Lukinavičius, G., Lavogina, D., Orpinell, M., Umezawa, K., Reymond, L., Garin, N., Gönczy, P., and Johnsson, K. (2013b). Selective chemical crosslinking reveals a Cep57-Cep63-Cep152 centrosomal complex. Curr. Biol. *23*, 265–270.

Martinez-Campos, M., Basto, R., Baker, J., Kernan, M., and Raff, J.W. (2004). The Drosophila pericentrin-like protein is essential for cilia/flagella function, but appears to be dispensable for mitosis. J. Cell Biol. *165*, 673–683.

Mastronarde, D.N. (1997). Dual-axis tomography: An approach with alignment methods that preserve resolution. J. Struct. Biol. *120*, 343–352.

Matsuura, K., Lefebvre, P. a, Kamiya, R., and Hirono, M. (2004). Bld10p, a novel protein essential for basal body assembly in Chlamydomonas: localization to the cartwheel, the first ninefold symmetrical structure appearing during assembly. J. Cell Biol. *165*, 663–671.

Mayor, T., Stierhof, Y.D., Tanaka, K., Fry, A.M., and Nigg, E.A. (2000). The centrosomal protein C-Nap1 is required for cell cycle-regulated centrosome cohesion. J. Cell Biol. *151*, 837–846.

McKean, P.G., Baines, A., Vaughan, S., and Gull, K. (2003). γ-Tubulin functions in the nucleation of a discrete subset of microtubules in the eukaryotic flagellum. Curr. Biol. *13*, 598–602.

McNitt, R. (1974). Centriole ultrastructure and its possible role in microtubule formation in an aquatic fungus. Protoplasma *80*, 91–108.

Meehl, J.B., Bayless, B.A., Giddings, T.H., Pearson, C.G., and Winey, M. (2016). Tetrahymena Poc1 ensures proper intertriplet microtubule linkages to maintain basal body integrity. Mol. Biol. Cell *27*, 2394–2403.

Megraw, T.L., Li, K., Kao, L.R., and Kaufman, T.C. (1999). The centrosomin protein is required for centrosome assembly and function during cleavage in Drosophila. Development *126*, 2829–2839.

Mennella, V., Keszthelyi, B., Mcdonald, K.L., Chhun, B., Kan, F., Rogers, G.C., Huang, B., and Agard, D.A. (2012). Subdiffraction-resolution fluorescence microscopy reveals a domain of the centrosome critical for pericentriolar material organization. Nat. Cell Biol. *14*, 1159–1168.

Meunier, S., and Vernos, I. (2016). Acentrosomal Microtubule Assembly in Mitosis: The Where, When, and How. Trends Cell Biol. *26*, 80–87.

Mitchell, K.A.P., Szabo, G., and Otero, A.D.S. (2009). Methods for the isolation of sensory and primary cilia - an overview. Methods Cell Biol. *94*, 87–101.

Mogensen, M. (1999). Microtubule release and capture in epithelial cells. Biol. Cell *91*, 331–341.

Mogensen, M.M., Mackie, J.B., Doxsey, S.J., Stearns, T., and Tucker, J.B. (1997). Centrosomal deployment of γ-tubulin and pericentrin: Evidence for a microtubule-nucleating domain and a minus-end docking domain in certain mouse epithelial cells. Cell Motil. Cytoskeleton *36*, 276–290.

Mogensen, M.M., Malik, A., Piel, M., Bouckson-Castaing, V., and Bornens, M. (2000). Microtubule minus-end anchorage at centrosomal and non-centrosomal sites: the role of ninein. J. Cell Sci. *113*, 3013–3023.

Moorman, S.J., and Shorr, A.Z. (2008). The primary cilium as a gravitational force transducer and a regulator of transcriptional noise. Dev. Dyn. *237*, 1955–1959.

Moritz, M., Braunfeld, M.B., Sedat, J.W., Alberts, B., and Agard, D.A. (1995). Microtubule nucleation by γ-tubulin-containing rings in the centrosome. Nature *378*, 638–640.

Moritz, M., Braunfeld, M.B., Guénebaut, V., Heuser, J., and Agard, D.A. (2000). Structure of the γ-tubulin ring complex: A template for microtubule nucleation. Nat. Cell Biol. *2*, 365–370.

Moyer, T.C., Clutario, K.M., Lambrus, B.G., Daggubati, V., and Holland, A.J. (2015). Binding of STIL to Plk4 activates kinase activity to promote centriole assembly. J. Cell Biol. *209*, 863–878.

Murata, K., and Wolf, M. (2018). Cryo-electron microscopy for structural analysis of dynamic biological macromolecules. Biochim. Biophys. Acta - Gen. Subj. *1862*, 324–334.

Müsch, A. (2004). Microtubule organization and function in epithelial cells. Traffic *5*, 1–9.

Nakazawa, Y., Hiraki, M., Kamiya, R., and Hirono, M. (2007). SAS-6 is a cartwheel protein that establishes the 9-fold symmetry of the centriole. Curr. Biol. *17*, 2169–2174.

Nalepa, C.A. (2017). What Kills the Hindgut Flagellates of Lower Termites During the Host Molting Cycle ? Microorganisms 1–28.

Nigg, E. a, and Raff, J.W. (2009). Centrioles, centrosomes, and cilia in health and disease. Cell *139*, 663–678.

Nigg, E.A., and Stearns, T. (2011). The centrosome cycle: Centriole biogenesis, duplication and inherent asymmetries. Nat. Cell Biol. *13*, 1154–1160.

Nigg, E.A., Čajánek, L., and Arquint, C. (2014). The centrosome duplication cycle in health and disease. FEBS Lett. *588*, 2366–2372.

O'Rourke, B.P., Gomez-Ferreria, M.A., Berk, R.H., Hackl, A.M.U., Nicholas, M.P., O'Rourke, S.C., Pelletier, L., and Sharp, D.J. (2014). Cep192 controls the balance of centrosome and non-centrosomal microtubules during interphase. PLoS One *9*.

O'Toole, E.T., Giddings, T.H., McIntosh, J.R., and Dutcher, S.K. (2003). Three-dimensional organization of basal bodies from wild-type and δ-tubulin deletion strains of Chlamydomonas reinhardtii. Mol. Biol. Cell *14*, 2999–3012.

Oakley, C.E., and Oakley, B.R. (1989). Identification of γ-tubulin, a new member of the tubulin

superfamily encoded by mipA gene of Aspergillus nidulans. Nature *338*, 662–664.

Oakley, B.R., Oakley, C.E., Yoon, Y., and Jung, M.K. (1990). γ-tubulin is a component of the spindle pole body that is essential for microtubule function in Aspergillus nidulans. Cell *61*, 1289–1301.

Ohta, M., Ashikawa, T., Nozaki, Y., Kozuka-Hata, H., Goto, H., Inagaki, M., Oyama, M., and Kitagawa, D. (2014). Direct interaction of Plk4 with STIL ensures formation of a single procentriole per parental centriole. Nat. Commun. *5*, 5267.

Ohta, T., Essner, R., Ryu, J.H., Palazzo, R.E., Uetake, Y., and Kuriyama, R. (2002). Characterization of Cep135, a novel coiled-coil centrosomal protein involved in microtubule organization in mammalian cells. J. Cell Biol. *156*, 87–99.

Olieric, N., Kuchen, M., Wagen, S., Sauter, M., Crone, S., Edmondson, S., Frey, D., Ostermeier, C., Steinmetz, M.O., and Jaussi, R. (2010). Automated seamless DNA co-transformation cloning with direct expression vectors applying positive or negative insert selection. BMC Biotechnol. *10*, 56.

Pearson, C.G., Osborn, D.P.S., Giddings, T.H., Beales, P.L., and Winey, M. (2009). Basal body stability and ciliogenesis requires the conserved component Poc1. J. Cell Biol. *187*, 905–920.

Pelletier, L., Ozlu, N., Hannak, E., Cowan, C., Habermann, B., Ruer, M., Muller-Reichert, T., and Hyman, A.A. (2004). The Caenorhabditis elegans centrosomal protein SPD-2 is required for both pericentriolar material recruitment and centriole duplication. Curr. Biol. *14*, 863–873.

Pelletier, L., O'Toole, E., Schwager, A., Hyman, A. a, and Müller-Reichert, T. (2006). Centriole assembly in Caenorhabditis elegans. Nature *444*, 619–623.

Penczek, P.A. (2010). Resolution measures in molecular electron microscopy. Methods Enzymol. *482*, 73–100.

Pfreundschuh, M., Alsteens, D., Hilbert, M., Steinmetz, M.O., and Müller, D.J. (2014). Localizing chemical groups while imaging single native proteins by high-resolution AFM. Nano Lett. *14*, 2957–2964.

Picard, A., Karsenti, E., Dabauvalle, M.C., and Doree, M. (1987). Release of mature starfish oocytes from interphase arrest by microinjection of human centrosomes. Nature *327*, 170–172.

Portier, N., Audhya, A., Maddox, P.S., Green, R.A., Dammermann, A., Desai, A., and Oegema, K. (2007). A Microtubule-Independent Role for Centrosomes and Aurora A in Nuclear Envelope Breakdown. Dev. Cell *12*, 515–529.

Prosser, S.L., and Pelletier, L. (2017). Mitotic spindle assembly in animal cells: A fine balancing act. Nat. Rev. Mol. Cell Biol. *18*, 187–201.

Puklowski, A., Homsi, Y., Keller, D., May, M., Chauhan, S., Kossatz, U., Grünwald, V., Kubicka, S., Pich, A., Manns, M.P., et al. (2011). The SCF-FBXW5 E3-ubiquitin ligase is regulated by PLK4 and targets HsSAS-6 to control centrosome duplication. Nat. Cell Biol. *13*, 1004–1009.

Quarmby, L.M. (2009). Deflagellation. In The Chlamydomonas Sourcebook: Cell Motility and Behaviour, G. Witman, ed. pp. 43–70.

Quintyne, N.J., Gill, S.R., Eckley, D.M., Crego, C.L., Compton, D.A., and Schroer, T.A. (1999). Dynactin is required for microtubule anchoring at centrosomes. J. Cell Biol. *147*, 321–334.

Rappsilber, J., Mann, M., and Ishihama, Y. (2007). Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. Nat. Protoc. *2*, 1896–1906.

Reckel, S., Gehin, C., Tardivon, D., Georgeon, S., Kükenshöner, T., Löhr, F., Koide, A., Buchner, L., Panjkovich, A., Reynaud, A., et al. (2017). Structural and functional dissection of the DH and PH domains of oncogenic Bcr-Abl tyrosine kinase. Nat. Commun. *8*, 1–14.

Rieder, C.L., Faruki, S., and Khodjakov, A. (2001). The centrosome in vertebrates: More than a microtubule-organizing center. Trends Cell Biol. *11*, 413–419.

Rivero, S., Cardenas, J., Bornens, M., and Rios, R.M. (2009). Microtubule nucleation at the cis-side of the golgi apparatus requires AKAP450 and GM130. EMBO J. *28*, 1016–1028.

Robbins, E., Jentzsch, G., and Micali, A. (1968). The centriole cycle in synchronized HeLa cells. J. Cell Biol. *36*, 329–339.

Rogers, G.C., Rusan, N.M., Roberts, D.M., Peifer, M., and Rogers, S.L. (2009). The SCF Slimb ubiquitin ligase regulates Plk4/Sak levels to block centriole reduplication. J. Cell Biol. *184*, 225–229.

Roque, H., Wainman, A., Richens, J., Kozyrska, K., Franz, A., and Raff, J.W. (2012). Drosophila Cep135/Bld10 maintains proper centriole structure but is dispensable for cartwheel formation. J. Cell Sci. *125*, 5881–5886.

Rosenbaum, J.L., and Child, F.M. (1967). Flagellar regeneration in protozoan flagellates. J. Cell Biol. *34*, 345–364.

Sanders, A.A.W.M., and Kaverina, I. (2015). Nucleation and dynamics of Golgi-derived microtubules. Front. Neurosci. *9*, 1–7.

Satir, P., Pedersen, L.B., and Christensen, S.T. (2010). The primary cilium at a glance. J. Cell Sci. *123*, 499–503.

Schindelin, J., Arganda-Carreras, I., and Frise, E. (2012). Fiji: an open-source platform for biological-image analysis. Nat. Methods *9*, 676–682.

Schmid, M.F., and Booth, C.R. (2008). Methods for aligning and for averaging 3D volumes with missing data. J. Struct. Biol. *161*, 243–248.

Schmidt, T.I., Kleylein-Sohn, J., Westendorf, J., Le Clech, M., Lavoie, S.B., Stierhof, Y.D., and Nigg, E.A. (2009). Control of Centriole Length by CPAP and CP110. Curr. Biol. *19*, 1005–1011.

Schneider, L., Clement, C.A., Teilmann, S.C., Pazour, G.J., Hoffmann, E.K., Satir, P., and Christensen, S.T. (2005). PDGFRαα signaling is regulated through the primary cilium in fibroblasts. Curr. Biol. *15*, 1861–1866.

Schur, F.K.M., Hagen, W.J.H., De Marco, A., and Briggs, J.A.G. (2013). Determination of protein structure at 8.5Å resolution using cryo-electron tomography and sub-tomogram averaging. J. Struct. Biol. *184*, 394–400.

Schur, F.K.M., Hagen, W.J.H., Rumlová, M., Ruml, T., Müller, B., Kraüsslich, H.G., and Briggs, J.A.G. (2015). Structure of the immature HIV-1 capsid in intact virus particles at 8.8 Å resolution. Nature *517*, 505–508.

Sharma, A., Aher, A., Dynes, N.J., Frey, D., Katrukha, E.A., Jaussi, R., Grigoriev, I., Croisier, M., Kammerer, R.A., Akhmanova, A., et al. (2016). Centriolar CPAP/SAS-4 Imparts Slow Processive Microtubule Growth. Dev. Cell *37*, 362–376.

Silflow, C.D., Liu, B., LaVoie, M., Richardson, E.A., and Palevitz, B.A. (1999). γ-tubulin in Chlamydomonas: characterization of the gene and localization of the gene product in cells. Cell

Motil. Cytoskeleton *42*, 285–297.

Sillibourne, J.E., Tack, F., Vloemans, N., Boeckx, A., Thambirajah, S., Bonnet, P., Ramaekers, F.C.S., Bornens, M., and Grand-Perret, T. (2010). Autophosphorylation of Polo-like kinase 4 and its role in centriole duplication. Mol. Biol. Celliology Cell *21*, 547–561.

Sir, J.H., Pütz, M., Daly, O., Morrison, C.G., Dunning, M., Kilmartin, J. V., and Gergely, F. (2013). Loss of centrioles causes chromosomal instability in vertebrate somatic cells. J. Cell Biol. *203*, 747–756.

Slaytor, M., Sugimoto, a, Azuma, J.-I., Murashima, K., and Inoue, T. (1997). Cellulose and Xylan Utilisation in the Lower Termite Reticulitermes speratus. J. Insect Physiol. *43*, 235–242.

Sluder, G. (2014). One to only two: a short history of the centrosome and its. Philos. Trans. R. Soc. B *369*, 1–8.

Sonnen, K.F., Schermelleh, L., Leonhardt, H., and Nigg, E.A. (2012). 3D-structured illumination microscopy provides novel insight into architecture of human centrosomes. Biol. Open *1*, 965–976.

Sonnen, K.F., Gabryjonczyk, A.-M., Anselm, E., Stierhof, Y.-D., and Nigg, E.A. (2013). Human Cep192 and Cep152 cooperate in Plk4 recruitment and centriole duplication. J. Cell Sci. *126*, 3223–3233.

Spektor, A., Tsang, W.Y., Khoo, D., and Dynlacht, B.D. (2007). Cep97 and CP110 Suppress a Cilia Assembly Program. Cell *130*, 678–690.

Stearns, T., Evans, L., and Kirschner, M. (1991). γ-Tubulin is a highly conserved component of the centrosome. Cell *65*, 625–636.

Stepanek, L., and Pigino, G. (2016). Microtubule doublets are double-track railways for intraflagellar transport trains. Science (80-. ). *352*, 721–724.

Stevenson, C.G., and Beane, W.S. (2010). A low percent ethanol method for immobilizing planarians. PLoS One *5*.

Stockbridge, R.B., Kolmakova-Partensky, L., Shane, T., Koide, A., Koide, S., Miller, C., and Newstead, S. (2015). Crystal structures of a double-barrelled fluoride ion channel. Nature *525*, 548–551.

Strnad, P., Leidel, S., Vinogradova, T., Euteneuer, U., Khodjakov, A., and Gönczy, P. (2007). Regulated HsSAS-6 levels ensure formation of a single procentriole per centriole during the centrosome duplication cycle. Dev. Cell *13*, 203–213.

Sugioka, K., Hamill, D.R., Lowry, J.B., McNeely, M.E., Enrick, M., Richter, A.C., Kiebler, L.E., Priess, J.R., and Bowerman, B. (2017). Centriolar SAS-7 acts upstream of SPD-2 to regulate centriole assembly and pericentriolar material formation. Elife *6*, 1–25.

Sydor, A.M., Czymmek, K.J., Puchner, E.M., and Mennella, V. (2015). Super-Resolution Microscopy: From Single Molecules to Supramolecular Assemblies. Trends Cell Biol. *25*, 730–748.

Szollosi, A., Ris, H., Szollosi, D., and Debec, A. (1986). A centriole-free Drosophila cell line. A high voltage EM study. Eur. J. Cell Biol. *40*, 100–104.

Takada, S., Kelkar, A., and Theurkauf, W.E. (2003). Drosophila checkpoint kinase 2 couples centrosome function and spindle assembly to genomic integrity. Cell *113*, 87–99.

Tang, C.J.C., Fu, R.H., Wu, K.S., Hsu, W. Bin, and Tang, T.K. (2009). CPAP is a cell-cycle regulated protein that controls centriole length. Nat. Cell Biol. *11*, 825–831.

Tang, C.J.C., Lin, S.Y., Hsu, W. Bin, Lin, Y.N., Wu, C.T., Lin, Y.C., Chang, C.W., Wu, K.S., and Tang, T.K. (2011). The human microcephaly protein STIL interacts with CPAP and is required for procentriole formation. EMBO J. *30*, 4790–4804.

Tanos, B.E., Yang, H.J., Soni, R., Wang, W.J., Macaluso, F.P., Asara, J.M., and Tsou, M.F.B. (2013). Centriole distal appendages promote membrane docking, leading to cilia initiation. Genes Dev. *27*, 163–168.

Taschner, M., and Lorentzen, E. (2016). The intraflagellar transport machinery. Cold Spring Harb. Perspect. Biol. *8*, 1–20.

Tassin, A.-M., and Bornens, M. (1999). Centrosome structure and microtubule nucleation in animal cells. Biol. Cell *91*, 343–354.

The C. elegans Sequencing Consortium (1998). Genome sequence of the nematode C. elegans: a platform for investigating biology. Science (80-. ). *282*, 2012–2018.

Thompson, G.A., Baugh, L.C., and Walker, L.F. (1974). Nonlethal deciliation of Tetrahymena by local anesthetic and its utility as a tool for studying cilia regeneration. J. Cell Biol. *61*, 253–257.

Tian, L., and Zhou, X. (2014). The soldiers in societies: Defense, regulation, and evolution. Int. J. Biol. Sci. *10*, 296–308.

Tilney, L.G., Bryan, J., Bush, D.J., Fujiwara, K., Mooseker, M.S., Murphy, D.B., and Snyder, D.H. (1973). Microtubules: Evidence for 13 Protofilaments. J. Cell Biol. *59*, 267–275.

Trabuco, L.G., Villa, E., Schreiner, E., Harrison, C.B., and Schulten, K. (2009). Molecular dynamics flexible fitting: A practical guide to combine cryo-electron microscopy and X-ray crystallography. Methods *49*, 174–180.

Trager, W. (1934). The cultivation of a cellulose-digesting flagellate, Trichomonas termopsidis, and of certain other termite protozoa. Biol. Bull. *66*, 182–190.

Tritarelli, A., Oricchio, E., Ciciarello, M., Mangiacasale, R., Palena, A., Lavia, P., Soddu, S., and Cundari, E. (2004). p53 Localization at Centrosomes during Mitosis and Postmitotic Checkpoint Are ATM-dependent and Require Serine 15 Phosphorylation. Mol. Biol. Cell *15*, 3751–3737.

Turoňová, B., Schur, F.K.M., Wan, W., and Briggs, J.A.G. (2017). Efficient 3D-CTF correction for cryo-electron tomography using NovaCTF improves subtomogram averaging resolution to 3.4 Å. J. Struct. Biol. *199*, 187–195.

Vaizel-Ohayon, D., and Schejter, E.D. (1999). Mutations in centrosomin reveal requirements for centrosomal function during early Drosophila embryogenesis. Curr. Biol. *9*, 889–898.

Varmark, H., Llamazares, S., Rebollo, E., Lange, B., Reina, J., Schwarz, H., and Gonzalez, C. (2007). Asterless Is a Centriolar Protein Required for Centrosome Function and Embryo Development in Drosophila. Curr. Biol. *17*, 1735–1745.

Vogel, J.M., Stearns, T., Rieder, C.L., and Palazzo, R.E. (1997). Centrosomes isolated from Spisula solidissima oocytes contain rings and an unusual stoichiometric ratio of α/β tubulin. J. Cell Biol. *137*, 193–202.

Vorobjev, I.A., and Chentsov, Y.S. (1982). Centrioles in the cell cycle. I. Epithelial cells. J. Cell Biol. *98*, 938–949.

Wan, W., and Briggs, J.A.G. (2016a). Cryo-Electron Tomography and Subtomogram Averaging. *579*,

329–367.

Wan, W., and Briggs, J.A.G. (2016b). Cryo-Electron Tomography and Subtomogram Averaging. In The Resolution Revolution: Recent Advances In CryoEM, (Elsevier Inc.), pp. 329–367.

Wang, J.T., Kong, D., Hoerner, C.R., Loncarek, J., and Stearns, T. (2017). Centriole triplet microtubules are required for stable centriole formation and inheritance in human cells. Elife 6, 1–17.

Wang, X., Tsai, J.W., Imai, J.H., Lian, W.N., Vallee, R.B., and Shi, S.H. (2009). Asymmetric centrosome inheritance maintains neural progenitors in the neocortex. Nature 461, 947–955.

Wegel, E., Göhler, A., Lagerholm, B.C., Wainman, A., Uphoff, S., Kaufmann, R., and Dobbie, I.M. (2016). Imaging cellular structures in super-resolution with SIM, STED and Localisation Microscopy: A practical comparison. Sci. Rep. 6, 1–13.

Wiese, C., and Zheng, Y. (2000). A new function for the γ-tubulin ring complex as a microtubule minus-end cap. Nat. Cell Biol. 2, 358–364.

Winey, M., and O'Toole, E. (2014). Centriole structure. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 369, 20130457-.

Witman, G.B., Carlson, K., Berliner, J., and Rosenbaum, J.L. (1972). Chlamydomonas flagella: I. isolation and electrophoretic analysis of microtubules, matrix, membranes, and mastigonemes. J. Cell Biol. 54, 507–539.

Wojcik, J., Lamontanara, A.J., Grabe, G., Koide, A., Akin, L., Gerig, B., Hantschel, O., and Koide, S. (2016). Allosteric inhibition of Bcr-Abl kinase by high affinity monobody inhibitors directed to the Src homology 2 (SH2)-kinase interface. J. Biol. Chem. 291, 8836–8847.

Wong, Y.L., Anzola, J. V., Davis, R.L., Yoon, M., Motamedi, A., Kroll, A., Seo, C.P., Hsia, J.E., Kim, S.K., Mitchell, J.W., et al. (2015). Reversible centriole depletion with an inhibitor of Polo-like kinase 4. Science (80-. ). 348, 1155–1160.

Woodruff, J.B., Wueseke, O., and Hyman, A.A. (2014). Pericentriolar material structure and dynamics. Philos. Trans. R. Soc. B Biol. Sci. 369, 20130459–20130459.

Woodruff, J.B., Wueseke, O., Viscardi, V., Mahamid, J., Ochoa, S.D., Bunkenborg, J., Widlund, P.O., Pozniakovsky, A., Zanin, E., Bahmanyar, S., et al. (2015). Regulated assembly of supramolecular centrosome scaffold in vitro. Science (80-. ). 348, 808–812.

Yamashita, Y.M., Mahowald, A.P., Perlin, J.R., and Fuller, M.T. (2007). Asymmetric inheritance of mother versus daughter centrosome in stem cell division. Science (80-. ). 315, 518–521.

Zamora, I., and Marshall, W.F. (2005). A mutation in the centriole-associated protein centrin causes genomic instability via increased chromosome loss in Chlamydomonas reinhardtii. BMC Biol. 3, 1–7.

Zhangs, Y., Ross, E.M., and Snell, W.J. (1991). ATP-dependent regulation of flagellar adenylylcyclase in gametes of Chlamydomonas reinhardti. J. Biol. Chem. 266, 22954–22959.

Zheng, S., Palovcak, E., Armache, J.-P., Cheng, Y., and Agard, D. (2017). MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. Nat. Methods 14, 331–332.

Zheng, Y., Jung, M.K., and Oakley, B.R. (1991). γ-Tubulin is present in Drosophila melanogaster and homo sapiens and is associated with the centrosome. Cell 65, 817–823.

Zhou, X., Wheeler, M.M., Oi, F.M., and Scharf, M.E. (2008). RNA interference in the termite Reticulitermes flavipes through ingestion of double-stranded RNA. Insect Biochem. Mol. Biol. *38*, 805–815.

Zhu, F., Lawo, S., Bird, A., Pinchev, D., Ralph, A., Richter, C., Müller-Reichert, T., Kittler, R., Hyman, A.A., and Pelletier, L. (2008). The Mammalian SPD-2 Ortholog Cep192 Regulates Centrosome Biogenesis. Curr. Biol. *18*, 136–141.

# CURRICULUM VITAE

## PERSONAL INFORMATION

|  |  |
|---|---|
| Name | Veronika Nemčíková Villímová |
| Nationality | Slovak |
| Place and date of birth | Malacky (Slovakia), 10. 11. 1987 |
| Present address | Pue Pierre de Savoie 7, 1680 Romont |

## EDUCATION

| 2013 - present | École Polytechnique Fédérale de Lausanne (EPFL), Switzerland |
|---|---|
|  | Ph.D. in Molecular Life Sciences |
| 2011 - 2013 | École Normale Supérieure de Lyon / Université Claude Bernard Lyon 1, France |
|  | M. Sc. in Biosciences |
| 2008 - 2011 | Pavol Jozef Šafárik University in Košice, Slovakia |
|  | B. Sc. in Biology |

## PROFESSIONAL EXPERIENCE

| 2013 - present | Swiss Institute for Experimental Cancer Research, École Polytechnique Fédérale de Lausanne, Switzerland |
|---|---|
|  | Laboratory of Prof. Pierre Gönczy |
| • PhD research project | Study of the centriolar structure using biochemistry, cryo-electron microscopy, and mass spectrometry on *Trichonympha* centrioles |
| January 2012 – Mai 2012, February 2013 – June 2013 | Laboratory of plant reproduction and development, École Normale Supérieure de Lyon, France |
| • Research internships | Participated in studies of transcriptional interactions of genes controlling carpel development by surface plasmon resonance analysis and its applications in "evo-devo" research |
| September 2012 – December 2012 | Conservatoire et Jardin Botaniques, Geneva, Switzerland |
|  | Laboratory of anatomy and micromorphology |
| • Research internship | Studied micromorphology and systematics in the Neotropical palm tribes |

Leopoldinieae and Manicarieae (Arecoideae), using histology techniques and scanning electron microscopy

## PUBLICATIONS

Guichard P., Hamel V., Le Guennec M., Banterle N., Iacovache I, Nemčíková V., Flückiger I., Goldie K. N., Stahlberg H., Lévy D., Zuber B. & Gönczy P. Cell-free reconstitution reveals centriole cartwheel assembly mechanisms. Nat. Commun. 8, 14813 doi: 10.1038/ncomms14813 (2017).

Villímová, V. and Stauffer, F. W. (2013). Floral Structure in the Neotropical Tribes Leopoldinieae and Manicarieae (Arecaceae: Arecoideae). Palms. 57, 181-193.

## SKILLS

*Languages:*    Proficient in English and French, mother tongue is Slovak

*Computer:*    Microsoft Office, Internet

Graphic and image processing softwares

## INTERESTS

- road cycling, running, swimming and mountain hiking
- travelling