DISSERTATION

Titel der Dissertation

Part 1: Cloning and Expression of Intraflagellar Transport Complex A

Part 2: The SAS-6 Coiled Coil Structure and Its Specific Interaction with SAS-5 Suggest a Regulatory Mechanism in Centriole Assembly

Verfasserin

Mag. Renping Qiao

angestrebter akademischer Grad

Doktorin der Naturwissenschaften (Dr.rer.nat.)

Wien, September 2012

Acknowledgments

I thank my supervisor **Gang Dong**, who gave me the opportunity to work on these interesting, challenging and competitive projects. Gang taught me side by side when I started my PhD, gave me the first hand experience which I benefit all the time. He provided me with sufficient resources, generous materials and collaborations for my projects. He is always up for the discussions and questions, pushing the projects forward. Gang's enthusiasm to science and hardworking spirit set a best example for me. Most importantly, I would like to thank Gang for his consistent encouragement, patience, criticism and trust, which makes the Ph.D study an exciting and valuable chapter of my life.

I thank my Ph.D committee members **Graham Warren** and **Tim Clausen**. Graham gave me valuable opinions and input to my project and spent time checking our manuscript, even when he was very busy with the direction of the institute. I really appreciate it. I want to thank Tim for very honest and sharp scientific discussion during each of my Ph.D committee meetings and generously sharing the synchrotron beam time and the instruments with us. It brought me a lot of convenience for my research.

I thank my lovely colleague **Clara Pleban**, who has accompanied me throughout my Ph.D study. I thank her for her awesomeness, her support, her comfort, her humor and kindness. She contributed a lot of time and energy to our lab maintenance before the technician came, which made our work much more efficient. She was always there when I was frustrated and down. Our jokes, funny talking, daily chitchat made my everyday life in the lab much easier and more pleasant. I always enjoy our time together both during and after work. I feel very lucky to have her around.

I thank **Keni Vidilaseris**. He is constantly calm and always finds time to help others whenever being asked. I know sometimes it is probably overwhelming for him when all three girls are talking and laughing, but he is always so nice and has never complained. I thank Keni for his good advice on the scientific books. I really enjoyed the experience that we practiced structure determination together.

I thank **Ekaterina Shimanovskaya** for her help on SAS-5 constructs and EM images recording. Katya gave me her comfort and nice girly talk when I needed her company. I thank Katya for her passion and dynamic character and sharing various interesting topics with me.

Table of contents

Abbreviations

Å	Angstroem
Amp (or Ap)	Ampicillin
b-ME	β-mercaptoethanol
Bac	Bacmid
BBS	Bardet-Biedl syndrome
BBSome	A complex of Bardet–Biedl syndrome (BBS) proteins
C. elegans (or *Ce*)	*Caenorhabditis elegans*
C-terminal	Carboxy terminal
CCD	Coiled coil domain
Chl	Chloramphenicol
CPB	Cryptic polo-box
C. reinhardtii (or *Cr*)	*Chlamydomonas reinhardtii*
CTD	C-terminal domain
cv	Column volume
D. rerio (or *Dr*)	Danio *rerio*
Da	Dalton
DLS	Dynamic light scattering
Dm	*Drosophila melanogaster*
ddH$_2$O	Double-distilled water
DTT	Dithiothreitol
E.coli	*Escherichia coli*
eYFP	Enhanced yellow fluorescent protein
FGF	Fibroblast growth factor
G0	Resting *phase*
G1	Gap phase 1
G2	Gap phase 2
HF cells	High Five insect cell line that originated from the ovarian cells of the cabbage looper, *Trichoplusia ni*
Hh	Hedgehog signaling

His-tag	Hexahistidine-tag
IPTG	Isopropyl-β-D-thiogalactopyranoside
IFT	Intraflagellar transport
IFT-A	Intraflagellar transport complex A
IFT-B	Intraflagellar transport complex B
ITC	Isothermal titration calorimetry
IgG	Immunoglobulin G
IPTG	Isopropyl-β-D-thiogalactopyranoside
JBTS	Joubert syndrome
Kan	Kanamycin
K_d	*Dissociation constant*
LB	*Lysogeny broth*
MBP	Maltose binding protein
MKS	Meckel-Grüber syndrome
MTOC	Microtubule-organizing center
MT(s)	Microtubule(s)
Ni-NTA	Nickel nitrile triacetate
N-terminal	Amino terminal
OFDS	Orofaciodigital syndrome
PBD(s)	Polo-box domain(s)
PCM	Pericentriolar material
PCP	Planar cell polarity
PDGF	Platelet-derived growth factor
PKD	Polycystic kidney disease
PLK1	Polo-like kinase 1
PLK4	Polo-like kinase 4
PP2A	Protein phosphatase 2A
RNAi	RNA interference
S phase	Synthesis phase
SAS-5	Spindle assemble abnormal 5
SAS-6	Spindle assemble abnormal 6

SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
Sf9 cells	*Spodoptera frugiperda* ovary cells
SLS	Static light scattering
Spec	Spectinomycin
Tb (or T. brucei)	*Trypanosoma brucei*
TPR	Tetratricopeptide repeat
v/v	Volume per volume
WAA	Degenerate TPR-like repeats
WD repeats	Stretch of ~40 amino acids usually ending with Trp-Asp
wt	wild-type
w/w	Mass per mass
X-gal	5-bromo-4-chloro-indolyl-β-D- galactopyranoside

Amino acids were abbreviated according to the standard one or three letter Nomenclatures

PART 1: CLONING AND EXPRESSION OF INTRAFLAGELLAR TRANSPORT COMPLEX A

1. Abstract

Cilia and flagella are specialized organelles present in most eukayotic cells and consist of a membrane-sheathed axoneme and about 700 associated proteins. These organelles play important roles in cell motility and/or signal transduction and have recently been associated with a plethora of human disorders. One of my PhD projects focused on the cloning and probably structural studies of a few proteins and their complexes that are essential for ciliogenesis.

Devoid of ribosomes and membrane-bound vesicles, cilia and flagella are assembled and maintained by intraflagellar transport (IFT), which is a bidirectional transport process along the microtubules of the axoneme. I started my PhD study trying to clone and over-express the IFT complex A (IFT-A), which is responsible for the retrograde transport in cilia. The six IFT-A genes were successfully cloned into three different MultiBac transfer vectors specifically designed for multi-gene amplification. These constructs were fused using Donor-Acceptor *in vitro* Cre-*loxP* recombination to generate two multi-gene expression constructs, which later were integrated into two MultiBac plasmids using Tn7 transposition for baculovirus production in insect cells. Expressions of all six IFT-A proteins in both Sf9 and Hi-Five insect cells were confirmed by anti-6×His western blot. However, the yield was low and purification by Ni-NTA proved not very successful. Future plans are both to increase the yield of co-expressed proteins and to refine purification strategies.

2. Zusammenfassung

Zilien oder Flagellen sind hochspezialisierte Organellen, die auf der Mehrzahl eukaryotischer Zellen zu finden sind und aus einem von der Plasmamembran umschlossenen Axonem bestehen. Diese Zellfortsätze spielen eine wichtige Rolle bei der zellulären Fortbewegung sowie in verschiedenen Signalkaskaden und konnten bereits mit diversen Erkrankungen in Verbindung gebracht werden. Eines meiner PhD Projekte beschäftigte sich mit der Klonierung und Aufreinigung von sechs verschiedenen Proteinen, die eine wichtige Rolle in funktionierenden Zilien spielen.

Essenziell für den Aufbau und die Instandhaltung von Zilien ist der sogenannte Intraflagellare Transport (IFT), ein bi-direktionaler Transport von ciliären Bestandteilen entlang des mikrotubulären Axonems. Dieser Transport wird von zwei großen Proteinkomplexen, IFT Komplex A und IFT Komplex B, durchgeführt.

Zu Beginn meines PhD-Studiums versuchte ich den IFT Komplex A (IFT-A), der für den retrograden Transport verantwortlich ist, zu klonieren und in einem eukaryotischen Expressionssystem zu überexprimieren. Die sechs verschiedenen Untereinheiten von IFT-A aus *Trypanosoma brucei* wurden in drei MultiBac Transfervektoren kloniert, welche speziell für die Amplifikation von mutliplen Genen geeignet sind. Die erfolgreich generierten Konstrukte wurden via einer Donor-Akzeptor *in vitro* Cre-loxP Rekombination fusioniert um zwei multi-Gen Expressionsvektoren zu generieren. Diese wurden darauffolgend mittels Tn7 Transposition in zwei MultiBac Plasmide integriert und für die Baculovirus-Produktion in Insektenzellen verwendet.

Die Expression aller sechs IFT-A Proteine in sowohl Sf9 als auch Hi-Five Inseketenzellen konnte mittels anti-6xHis Western Blot nachgewiesen werden. Leider konnten nur geringe Mengen des Komplexes isoliert werden und die Aufreinigung via Affinitätschromatographie war nicht zufriedenstellend. Zukünftige Studien werden sich auf eine Steigerung des Expressionslevels der rekombinanten Proteine und auf die Verbesserung der Aufreinigungstrategie fokussieren.

3. Introduction

3.1 Cilia

Cilia and flagella are specialized organelles present in most eukaryotic cells with the exception of fungi and higher plants. There are two types of cilia: motile and non-motile (primary) cilia. Both types of cilia regulate an excessive number of cell signaling and cell development processes (Christensen et al. 2007; Eggenschwiler and Anderson 2007).

Motile cilia and flagella are typically assembled by nine peripheral doublet microtubules surrounding two central single microtubules (central pair), hence designated '9+2' cilia (**Figure 1**). The axoneme of 9+2 cilia also contains accessory structures including inner and outer dynein arms, radial spokes and nexin links, which are involved in motility. There are exceptions of motile cilia that are devoid of the central microtubules, for example, motile 9+0 cilia on the nodal cells of developing mammalian embryos (Hirokawa et al. 2006).

Most non-motile cilia contain 9+0 axonemes, which lack the central microtubules and the dynein arms (**Figure 1**). The 9+0 axonemes of the motile nodal cilia lack the central microtubules but have dynein arms.

Despite of structural variations in different classes of cilia, all cilia and flagella share similarities in the basic structure, which includes a basal body, a transition zone, a microtubule-based axoneme and the ciliary membrane (**Figure 1**). The ciliary axoneme projects out from the cell surface and is surrounded by a bilayer lipid membrane that is continuous with the plasma membrane of the cell body but contains a different subset of membrane receptors and ion channels. The region that separates the two membrane compartments at the ciliary base, called 'ciliary necklace', is connected by fibers to the transition zone of the basal body (**Figure 1**) (Gilula and Satir 1972). Different classes of cilia have various tissue-specific locations, which accordingly carry out different tissue-specific functions.
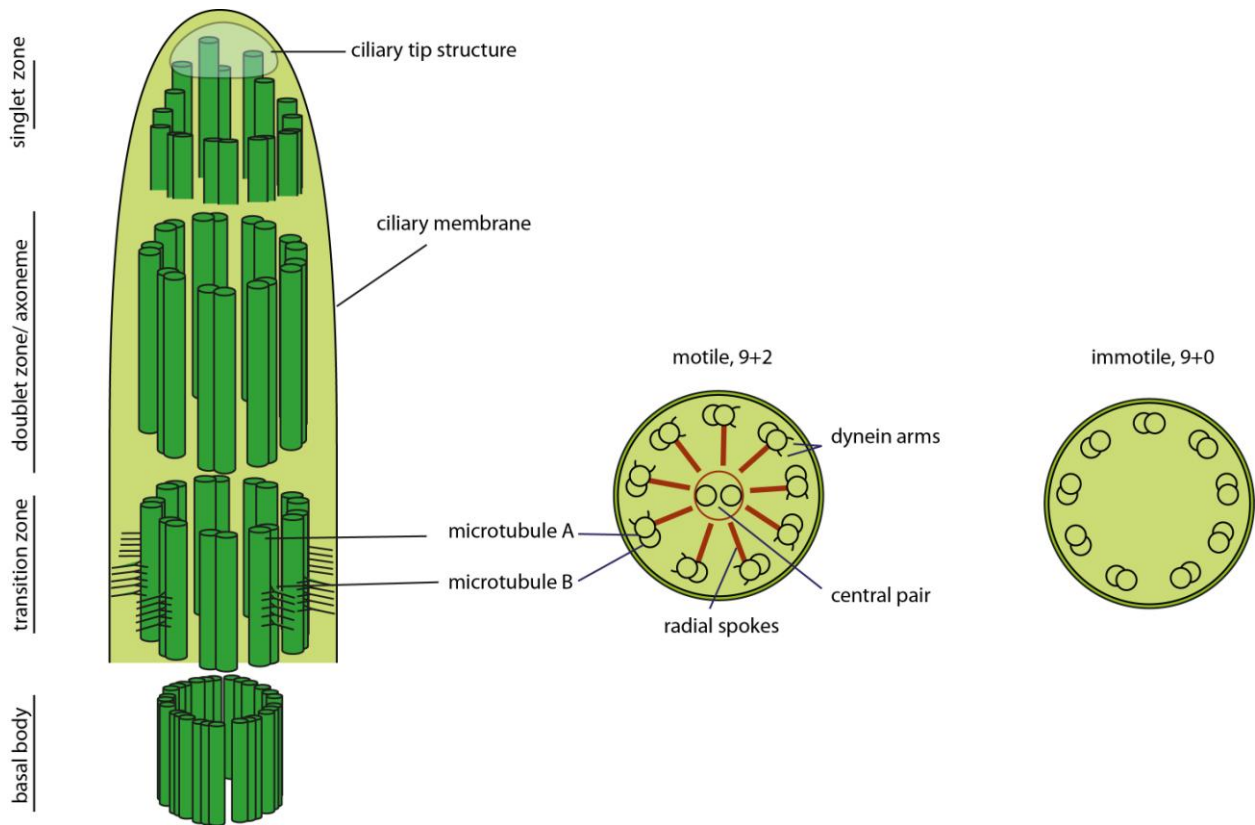
Figure 1 Ciliary structure. Cilia are composed of the basal body, the transition zone, the axoneme, the singlet zone, and the ciliary membrane. With a few exceptions, most cilia are divided into two groups: primary immotile cilia and motile cilia. Immotile cilia are characterized by 9+0 architecture, which has nine outer doublets, but no central microtubules, whereas motile cilia share a common 9+2 scaffold, which has outer nine doublets and two central microtubules (Czarnecki and Shah 2012).

3.1.1. Functions of Cilia

Primary cilia (non-motile cilia) and motile cilia have different specific functions. Primary cilia are mainly involved in sensing environmental cues, as cellular antennae that receive signals from the periphery (Berbari et al. 2009; Marshall and Nonaka 2006). Mutational analyses of individual cilia proteins have exhibited that cilia are involved in processing of both activator and repressor signal-transducing proteins (Goetz and Anderson 2010; Murdoch and Copp 2010). The cell signaling that primary cilia are associated with, include the Hedgehog (Hh), Wnt (canonical and non-canonical), platelet-derived growth factor (PDGF) and fibroblast growth factor (FGF) signaling pathways.

The motile functions of cilia are generated by coordinated activation and inactivation of the dynein motor proteins within the inner and outer dynein arms along the length of the axoneme (**Figure 1**). The motility of cilia is required in numerous biological processes. For example, the motility of cilia allows them to generate an extra-embryonic fluid flow (nodal flow) at the embryonic nod, which is important for the establishment of left-right (L-R) patterning during development (McGrath and Brueckner 2003; Tabin and Vogan 2003). Motile cilia of respiratory epithelial cells play an important role in lung clearance. Cerebrospinal fluid circulation and ependymal flow requires the motility of cilia (Ibanez-Tallon et al. 2004). Flagellar motility is essential for male and female fertility. Recently, motile cilia have been suggested to have sensory functions as well (Bloodgood 2010).

3.1.2. Ciliopathies

An expending number of cilia-related diseases have been identified, which are commonly referred to ciliophathies. Ciliophathies are related with developmental defects affecting the central nervous system, the skeleton or other organ systems.

Polycystic kidney disease (PKD or PCKD, also known as polycystic kidney syndrome) is a very common inherited cilia-related disease. The phenotype of PKD is frequently characterized by loss of glomerular filtration, variable ages of reaching end-stage renal disease, hypertension, renal cysts and subarachnoid hemorrhage from intracranial 'berry' aneurysm (Cano et al. 2004; Igarashi and Somlo 2002; Kiser et al. 2004; Oh and Katsanis 2012). Kidney epithelial cells with polycystin defects proliferate excessively, fail to differentiate fully and form cysts that can eventually destroy the kidney in about half of affected individuals. Scanning electron microscopy showed that PKD mice had abnormal cilia in their kidneys (Pazour et al. 2000).

Defects in cilia also lead to another human disorder called Bardet-Biedl syndrome (BBS), which is a rare genetic disorder characterized by renal cysts, polydactyly, mental retardation, obesity, hepatic fibrosis, hypertension, hypogonadotropism, anosmia, decrease in peripheral sensation and neocortical and hippocampal volume loss (Pan et al. 2005). BBS is genetically heterogeneous. Up to now, 16 loci have been identified in humans. Proteins encoded by the BBS loci localize primarily to the ciliary basal body and axoneme (Zaghloul and Katsanis 2009); many of them are components of the so-called BBSome, a protein complex involved in cilia targeting (Jin and

Nachury 2009; Jin et al. 2010). Mutations of a number of BBS loci in mouse had caused defects in cilium assembly and function, in addition to the phenotypes of BBS (Norris and Grimes 2012).

Other ciliopathies include Meckel-Grüber syndrome (MKS), which is one of the most severe human ciliopathies and consistently results in neonatal lethality; orofaciodigital syndrome (OFDS), which encompasses oral, facial and digital defects; Joubert syndrome (JBTS), which is an autosomal recessive disorder that comprises mental retardation, abnormal breathing, atypical eye movements and ataxia in association with agenesis of the cerebellar vermis (Sattar and Gleeson 2011).

3.2.3. Ciliogenesis

Formation of cilia is tightly coupled with the cell cycle. During G1 stage, the mother centriole-associated Golgi-derived vesicles fuse with the plasma membrane, forming a membrane sheath, which later covers the newly emerging ciliary axoneme. The distal appendages of mother centrioles were thought to be responsible for docking of mother centriole/basal body to the apical cell membrane. Next, microtubules nucleated by the basal body begin to grow, protruding beneath the membrane. The distal part of the basal body is called transition zone, where the nine outer MT doublets begin to emerge. The outer doublets only assemble at the distal end of the cilia. However, cilia and flagella are not able to synthesize de novo proteins. Therefore, assembly and maintenance of cilia require the transport of axonemal precursors from the cytoplasm to the growing ciliary tip, which is carried out by IFT. Exception, however, does exist in some species, such as *Drosophila* sperm cells, which assemble axonemes in the cytoplasm in an IFT-independent manner (Han et al. 2003).

When cilia stop to elongate, they still remain dynamic. New materials for cilia assembly are still imported to the tip, whereas the turn-over products are constantly transferred back to the base of the cilium, where there is a large pool of IFT components. This dynamic metabolism is re-quired for cilia to maintain their steady-state.

3.2. IFT

As mentioned above, since cilia/flagella lack the machinery for synthesizing proteins and since the axoneme assembly exclusively happens at the tip of cilia/flagella (Johnson and Rosenbaum

1992), transport of axonemal components is required. The cell has solved this logistical problem by means of IFT, a bi-directional movement of large protein complex along the axoneme. The IFT particles move to the distal tip of the flagellum (anterograde) by kinesin II at the speed of ~2 µm/s whereas back to the cell body (retrograde) by cytoplasmic dynein 2 at the speed of ~3.5 µm/s (Kozminski et al. 1995; Kozminski et al. 1993; Porter et al. 1999).
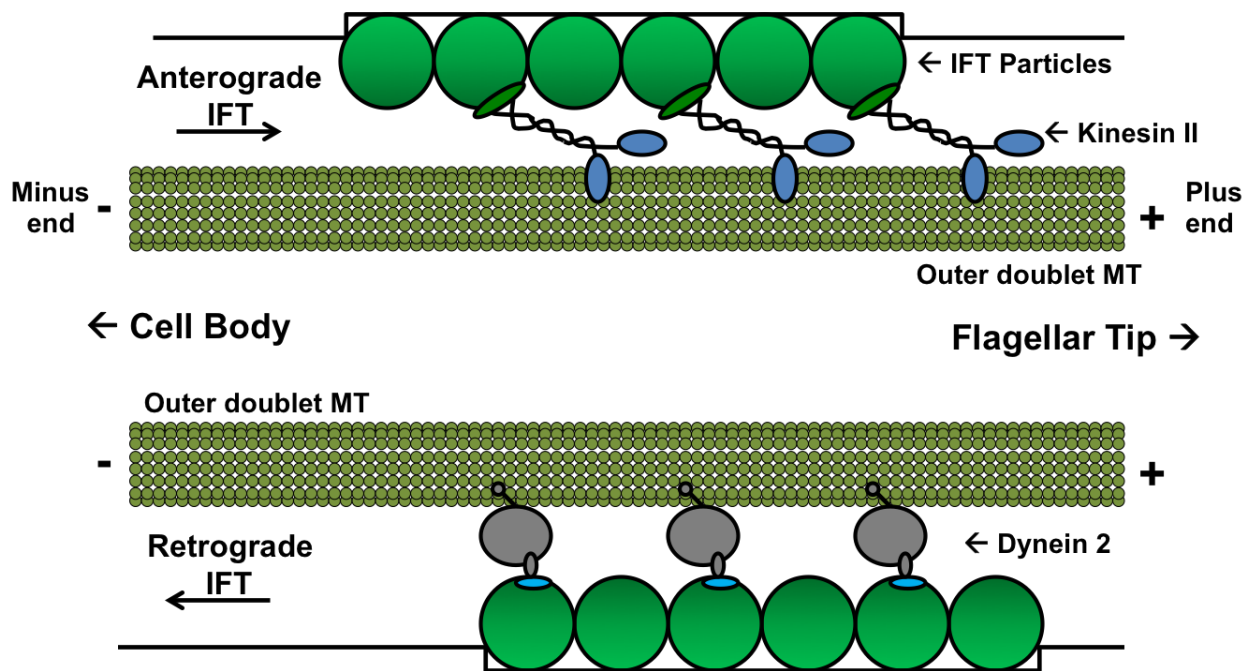
Figure 2 Schematic representation of intraflagellar transport (IFT). Anterograde IFT particles, which are motivated by Kinesin II along B tubules of outer doublet microtubules, transfer axoneme precursors from the base of the cillium out to the distal tip by heterotrimeric kinesin-II. Retrograde IFT particles, which are mediated by cytoplas-mic dynein 2, transfer the turn-over products from ciliary distal tip back to the ciliary base (Cole 2003).

IFT particles are composed of at least 18 different proteins, which form two sub-complexes, IFT-A and IFT-B (Taschner et al. 2012). IFT-A and IFT-B function in different aspects of the IFT process. IFT-A is specifically involved in anterograde and IFT-B is involved in retrograde transport. Mutational analysis studies have shown that mutations of IFT-B subunits cause cilia dramatical-ly shortened or absent (Brazelton et al. 2001; Hou et al. 2007), analogous with kinesin II muta-tions (Matsuura et al. 2002). IFT-A mutants, however, lead to malformed cilia with prominent bulges containing accumulations of IFT proteins (Iomini et al. 2009; Piperno et al. 1998), which is similar to what was observed in cytoplasmic dynein 2 mutations (Pazour et al. 1999; Porter et al. 1999; Schafer et al. 2003).

IFT proteins are highly conserved in ciliated organisms from green algae to mammals (Rosenbaum and Witman 2002). Proteins of the IFT complexes in *Chlamydomonas* were named after their apparent molecular weight on SDS-PAGE as IFT43, 121, 122, 139, 140, and 144 for IFT-A subunits and IFT20, 22, 25, 27, 46, 52, 54, 57, 70, 74, 80, 81, 88 and 172 for IFT-B subunits. Mutational analyses indicated that a subset of IFT-A proteins (IFT144/140/122) form a stable 12S subcomplex, referred to the IFT-A core (Behal et al. 2012; Taschner et al. 2012) (**Figure 3**). Analogously, biochemical studies suggested that the IFT-B complex can be stripped of weakly associated subunits to form a nine-protein salt-stable subcomplex (IFT22/25/27/46/52/70/74/81/88), which is referred as the IFT-B core (Bhogaraju et al. 2011; Lucker et al. 2005; Lucker et al. 2010; Taschner et al. 2012) (**Figure 3**). At present, high-resolution structures are only available for the IFT25/IFT27 complex (Bhogaraju et al. 2011) and cytoplasmic dynein motor domain (Kon et al. 2012) (**Figure 3**). Pigino *et al.* have published electron tomographic reconstruction of IFT trains using subtomographic averaging, which rep-resents a major step forward in structural understanding of IFT particles (Pigino et al. 2009) (**Figure 3**).
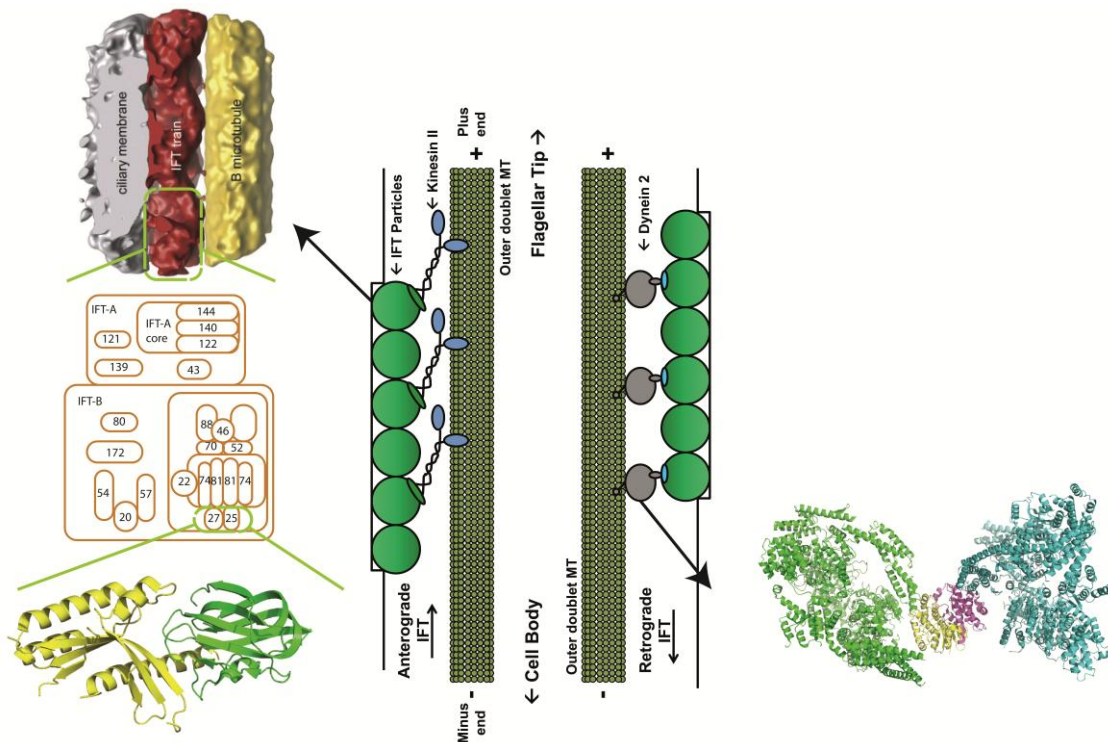
 Figure 3 High-resolution structures of dynein, electron tomographic reconstruction of IFT particles, and high-resolution crystal structure of IFT complex 25/27. (Mizuno et al. 2012)

3.2.1. IFT-A

Bioinformatic analysis suggests that, except for IFT43, all other components of the IFT-A complex are large proteins with molecular weights more than 120 kDa. According to secondary structure prediction, IFT144, IFT140, IFT 122 and IFT121 have a notably similar domain organization with an N-terminal WD-repeat and a C-terminal α-helical tetratricopeptide repeats (TPR) motif or degenerate TPR-like repeats (WAA). WD repeats, also known as WD40, are degenerate repeats of 40 or more amino acids that often contain conserved tryptophan and aspartic acid residues (Smith et al. 1999; Yu et al. 2000). WD-repeats are known to form β-propellers. The TPR motif consists of highly degenerate 34 amino acid repeats where a minimum of three tandem repeats are thought to form a binding domain that cradles a short, specific amino acid sequence, often the C-terminal part of a binding partner (Das et al. 1998; Lamb et al. 1995; Scheufler et al. 2000). Two IFT-A proteins without β-propellers are IFT139, which appears to contain only TPR repeats, and IFT43, which probably does not contain any motif (**Figure 4**). The predicted domain arrangement reflects that the IFT-A complex likely serves as a structural platform that binds to multiple proteins, such as motor proteins and the IFT-B components.
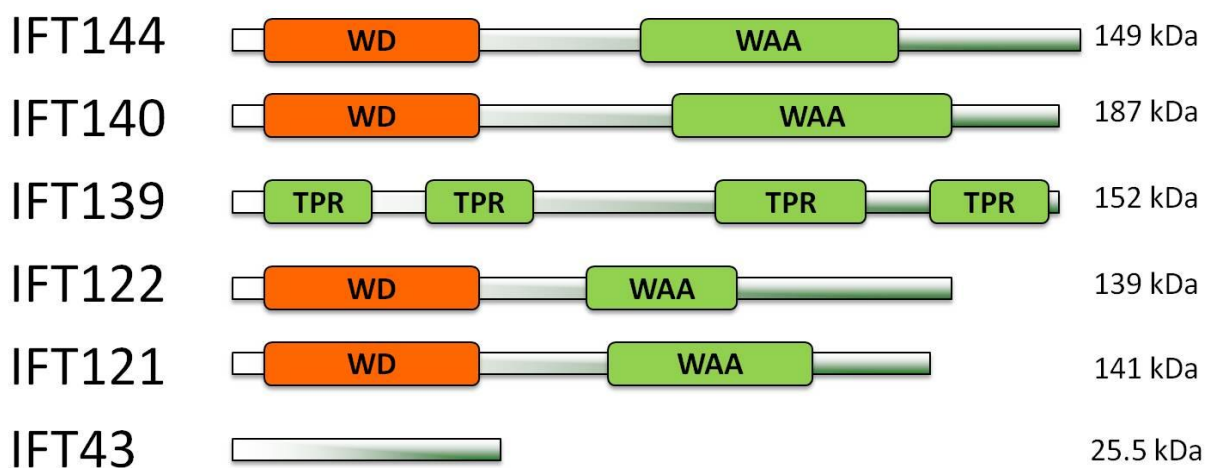
Figure 4 IFT-A particle components and domain predictions. IFT proteins were purified from *Chlamydomonas* and identified at the level of the gene. Putative protein-protein binding motifs have been identified through sequence analysis. WD: WD repeats; TPR: tetratricopeptide repeats; WAA: degenerate TPR-like repeats.

3.2.1.1. IFT144

IFT144 is predicted to have two 7-blade β-propellers (each blade having 4 β-strands), with the highest confidence in the first β-propeller. *Chalmydomanas* IFT144 mutant causes decreased velocities of retrograde IFT particles, reduced concentration of IFT-A proteins in flagella and the formation of a characteristic flagellar 'bulge' containing IFT-B proteins at the permissive temperature (Iomini et al. 2009; Lin et al. 2003). Homologous proteins of IFT144 are WDR19 in humans and DYF-2 in *C. elegans* (Efimenko et al. 2006; Lin et al. 2003). In agreement with the function of *Cr*IFT144, *dyf-2* mutant similarly caused the accumulation of IFT-B proteins in cilia. Mutations in human IFT144 lead to both Sensenbrenner and Jeune Syndromes with skeletal anomalies and renal insufficiency phenotypes (Bredrup et al. 2011).

3.2.1.2. IFT140

Similar to IFT144, IFT140 has one highly plausible β-propeller followed by a number of potential β-propellers. However, literature about IFT140 is limited.

3.2.1.3. IFT139

Different from other components of IFT-A, IFT139 does not contain any WD repeats, but instead have TPR repeats throughout the whole sequence (**Figure 4**). TPR repeats are believed to often act as a scaffold which transiently interacts with two or more proteins. Secondary structure analysis suggests that TPR repeats may allow IFT139 to simultaneously bind to as many as five proteins. Potential binding partners of IFT139 are IFT-A and IFT-B proteins, as well as axonemal precursors. *Cr*IFT139 mutants cause similar phenotypes to those from IFT144 mutants (Iomini et al. 2009; Piperno et al. 1998). IFT139 is suggested to be essential for recycling IFT complex A from the flagellum back to the cell body (Williamson et al. 2011). Analogously, mutations in mammalian orthologs result in defective retrograde IFT (Goetz and Anderson 2010; Tran et al. 2008).

3.2.1.4. IFT122 and IFT121

IFT122 and IFT121 show a remarkable similarity in domain organization with predicted β-propellers formed by WD repeats at the N-terminus and degenerate TPR repeats (WAA) at the C-terminus, respectively. IFT122 is not required for ciliogenesis in *Tetrahymena*. The null mutant, however, caused IFT-B proteins to accumulate at ciliary tips (Tsao and Gorovsky 2008b). IFT122 orthologs in *C. elegans*, humans and other mammals lead to severe defects in retrograde IFT (Bell et al. 2006; Cortellino et al. 2009; Qin et al. 2001; Qin et al. 2011). Mutant IFTA-1, the ortholog of IFT121 in *C. elegans*, causes chemosensory and dye-filling defects, abnormal cilia structure, and the accumulation of IFT-B proteins in cilia (Blacque et al. 2006). Mutant WDR35, the human ortholog of IFT121, leads to multiple ciliophathies, such as Sensenbrenner Syndrome (Mill et al. 2011).

3.2.1.5. IFT43

IFT43, the smallest component of IFT-A, does not contain any clearly distinguishable motif. Initially, IFT43 was not included in the IFT complex, probably due to different purification procedure (Piperno et al. 1998). This was corrected later when it was found out that defective IFT43 in human cells resulted in a characteristic malfunction in retrograde IFT, as anticipated for dysfunctional IFT-A components (Arts et al. 2011).

3.2.2. Subunit interactions of IFT-A

Behal *et al* have carried out extensive studies on the interactions between IFT-A subunits. A ~12 S subcomplex, referred as the IFT-A core, was fractionated during co-sedimentation of whole cell extract from *ift121* cells, which contains IFT122, IFT140 and IFT144 (Behal et al. 2012). The yeast-based two-hybrid analyses showed that IFT43 and the IFT121 C-terminal part interact with each other. It was further confirmed by the observation that IFT43 co-eluted with the MBP-tagged IFT121 when they were heterologously expressed in bacteria. The interaction between IFT43 and IFT122 was also observed. Mutational analysis indicates that the presence of IFT139 required IFT121, whereas the presence of IFT139 relied on IFT122.

3.2.3. IFT-B

IFT-B is biochemically divided into a salt-stable core of nine subunits and several peripheral subunits. Sufficient evidences have shown that knock-out of any components of IFT-B almost completely abolishes the assembly of cilia, suggesting that IFT-B is essential for anterograde transport of axonemal precursors to the ciliary tip (Absalon et al. 2008; Beales et al. 2007; Kobayashi et al. 2007; Li et al. 2008; Omori et al. 2008; Tran et al. 2008; Tsao and Gorovsky 2008a).

3.2.3.1. The IFT-B core proteins

It was shown that there is an absolute request of IFT88 for flagellar assembly in *Chlamydomona*s and as well as for ciliogenesis in vertebrates (Pazour et al. 2002; Pazour et al. 2000). IFT70 has a conserved role among green algae, nematodes and zebrafish. Deletion of IFT70 leads to abnormal and destabilized axonemes (Fan et al. 2010). IFT74 and IFT81 both contain long predicted coiled coil region. It was initially suggested that there exists a tetrameric IFT74-81-81-74 complex (Lucker et al. 2005). According to Taschner and colleagues, however, an IFT88/81/74/70/52/46/27/25 octamer can form in their reconstruction and purification experiments, suggesting that the formation of a $(IFT81)_2/(IFT74)2$ heterotetramer is not a condition for complex assembly (Taschner et al. 2011). Mutation of IFT52 had an analogous phenotype as IFT88 mutant, which causes a failure in flagellum assembly. IFT52 directly interacts with IFT46, IFT70 and IFT88 as well as with the IFT81/74/27/25 heterotetramer (Taschner et al. 2011). In agreement with the observation that IFT46 interacts with IFT52, IFT70 and IFT88 within the IFT-B core, western blot results revealed that when IFT46 was mutated, multiple components of IFT-B core level dramatically decreased. These results indicate that IFT46 is required for the stability of the IFT-B core (Fan et al. 2010; Lucker et al. 2010; Taschner et al. 2011). Additionally, IFT46 mutant also caused the axonemal outer dynein arms to be missing, implying that IFT46 might have a secondary role of recognizing the outer dynein arms (Hou et al. 2007). IFT27 and IFT25 form a stable subcomplex, whose structure has been solved at a high resolution (**Figure 3**) (Bhogaraju et al. 2011). Resent data have shown that the IFT27/IFT25 complex acts both during IFT complex assembly and at IFT initiation (Bhogaraju et al. 2011; Wang et al. 2009). IFT22 was the last identified component of IFT-B (Cole et al. 1998; Lucker et al. 2005), whose function awaits further characterization.

3.2.3.2. The peripheral IFT-B proteins

IFT172, IFT80, IFT57, IFT54 and IFT20 are not parts of the IFT-B core complex (**Figure3**), since they weakly associate with IFT-B core and dissociate at 300 mM slat concentration (Lucker et al. 2005). Nevertheless, these peripheral IFT-B proteins play important roles in IFT regulation and carrier versicle targeting from Golgi network to the clilium (Taschner et al. 2012).

4. Aim of the study

Whereas the complexes IFT A/B have been studied for nearly two decades, little is known about their architecture and assembly. The lack of high-resolution structural information on these complexes has now become a limiting step in gaining an understanding of their function at the molecular level.

This project mainly focused on cloning and expression of the IFT-A complex in a eukaryotic expression system, which, if successful, will enable us to carry out structural studies of the complex to elucidate the assembly mechanisms of the IFT-A complex for cargo transport within the cilium at the atomic level. The strategy we used for producing the IFT-A complex is to co-express all six components of IFT-A using the baculovirus-insect cell expression system that has been established and used successfully for other multimeric protein complexes (Bieniossek et al. 2008; Nie et al. 2009; Palomares et al. 2004). Following confirmation of overexpression of these proteins, purification by affinity and size exclusion chromatography would be carried out. Purified samples will be subjected to electron microscopy analysis and further crystallization trials. While successful X-ray crystallography will allow us to visualize these proteins and their complexes at the atomic level, large particles could also be examined by cryo-electron microscopy to provide a medium resolution but still useful information about the assembly of the complex.

5. Results

5.1. Cloning of IFT-A complex proteins into multiBac vectors

To establish the multigene expression system for the IFT-A complex, we took advantage of the established expression plasmids classified as either 'Acceptor' or 'Donor', which can be fused *in vitro* to generate multigene constructs (Fitzgerald et al. 2006) (**Figure 5**). A *loxP* sequence present on all vectors allows *in vitro* fusion of one or two Donors (pUCDM and pSPL) with the Acceptor (pFL) by Cre recombinase (**Figure 5**). Acceptor plasmids contain standard replication origins, while Donor plasmids carry an R6Kg conditional origin, rendering their propagation dependent on hosts expressing the *pir* gene (see Materials and Methods). Transformation of *in vitro* fusion reaction mixture of Donors and Acceptors into *pir*⁻ bacterial strains would eliminate non-fused Donors, while the unique antibiotic-resistance markers present on the Donor (s) and the Acceptor together allow for selection of the desired Donor-Acceptor fusions.

In this study, IFT-A complex proteins were divided into two groups of constructs during cloning, due to the large gene sizes. The first group contains *T. brucei* IFT43, IFT121, IFT122, IFT140. One Donor (pUCDM) carries *T. brucei* IFT43 and *T. brucei* IFT121, and the other Doner (pSPL) carries *T. brucei* IFT122. They were fused with the Acceptor pFL carrying IFT140 using *in vitro* recombination reaction by the Cre recombinase (**Figure 5a**). The second group contains *T. brucei* IFT43, IFT139, IFT144, IFT140. Donor pUCDM carrying *T. brucei* IFT43 and IFT139 and Doner pSPL carrying *T. brucei* IFT144 were fused with Acceptor pFL carrying IFT140 using *in vitro* recombination (**Figure 5b**).

Donor-Acceptor fusions enter the MultiBac genome by Tn7 transposition in DH10Multi-Bac cells (Berger et al. 2004). Successful Multibac bacmids are ready to transfect insect cells to generate bacoluvirus for protein expression.
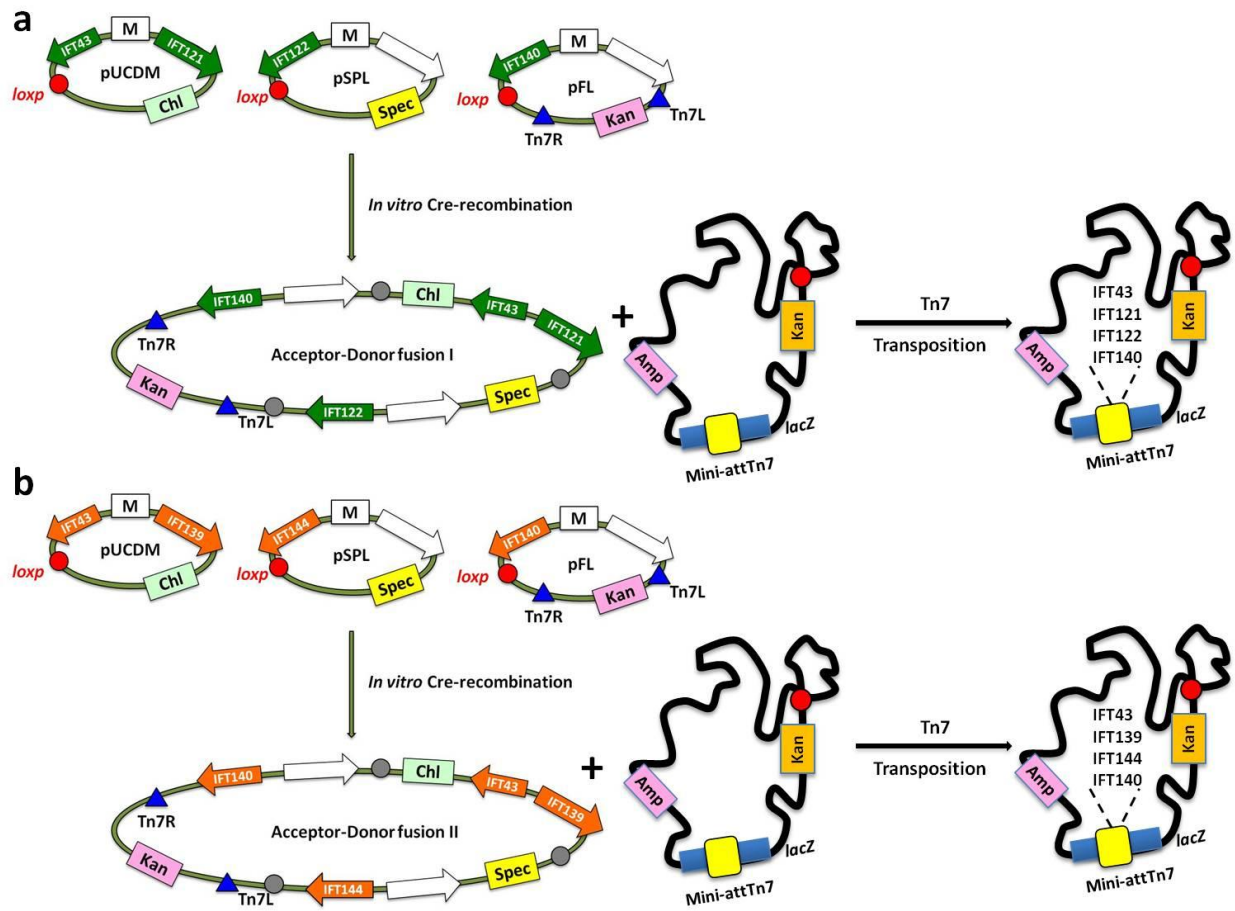
Figure 5 *In vitro* recombination of Acceptors and Donors carrying target genes. The Acceptor plasmid pFL and two Donor plasmids pUCDM and pSPL are fused to an Acceptor-Donor-Donor triple fusion plasmid by a single Cre-loxP reaction. Recombinant genes encode for **(a)** *T. brucei* IFT43, IFT121, IFT122 and IFT140, **(b)** *T. brucei* IFT43, IFT139, IFT144 and IFT140. Antibiotic markers (colored boxes), expression cassettes (open arrows), loxP (red and grey circles) and Tn7 transposition sequences (blue triangles). Acceptor-Donor fusions access the Multi-Bac plasmid through Tn7 transposition.

5.2. IFT-A complex proteins were expressed in insect cell system

Baculovirus were generated by transfecting the Sf9 or High-Five insect cells with multibac bacmids. Overexpression of heterologous proteins using the third generation of viruses (V3) was analyzed. Each IFT-A complex component has a C-terminal 6×His tag, which allowed us to use anti-His Western Blot to analyze the expression of the IFT-A proteins. Western Blot results revealed that IFT43, IFT121, IFT122 and IFT140 were successfully expressed in both Sf9 and High-Five cells transfected with virus A; so did for IFT43, IFT139, IFT144 and IFT140 in the Sf9

and High-Five cells transfected with virus B. The expression levels of two viruses were comparable to each other.

IFT43	IFT121	IFT122	IFT139	IFT140	IFT144
25.5 KDa	141 KDa	139 KDa	152 KDa	187 KDa	149 KDa
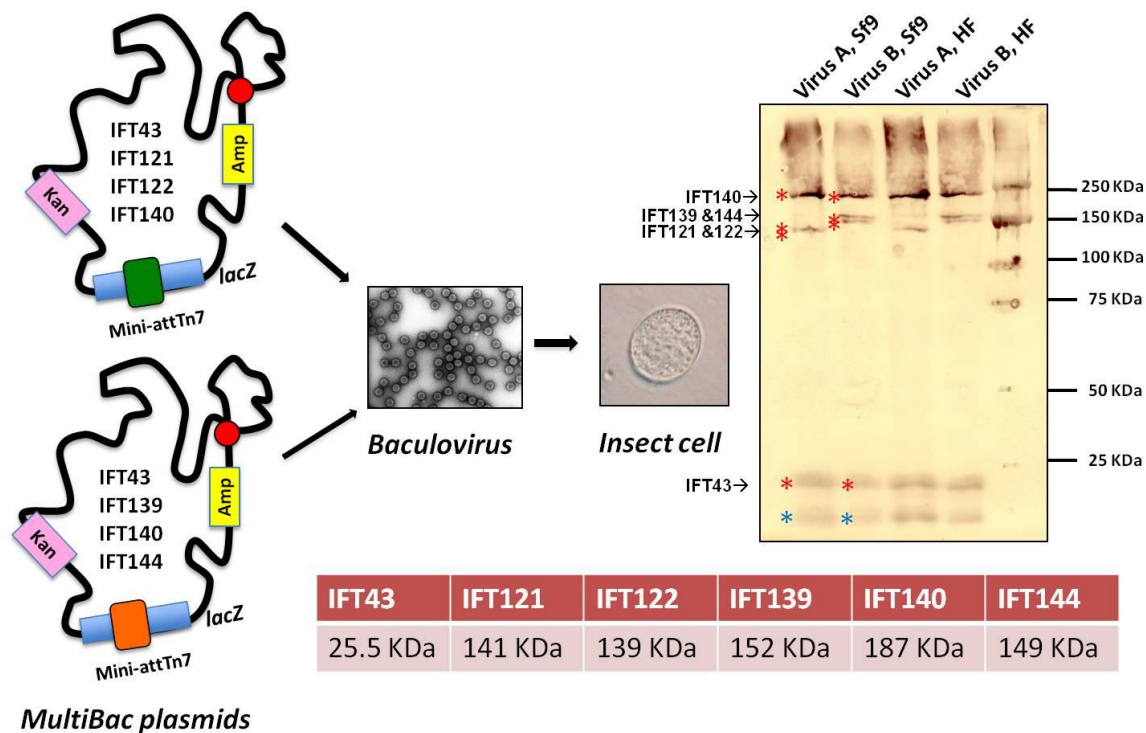
Figure 6 Expression of IFT-A compelx proteins in the baculovirus/insect cell system. Multigene constructs shown in Figure 5 were assembled in the MultiBac plasmid and overexpressed proteins were analyzed by anti-His Western blot. Each IFT-A protein contains a 6×His tag at its C-terminus. IFT-A proteins (red asterisks) and a degradation band (blue asterisks) were identified in cells by western blot with a mouse anti-His antibody and a secondary goat antibody coupled to alkaline phosphatase for identification of target proteins using color development.

Most of the protein bands were visualized on the Western blot except that *Tb*IFT121 and IFT122 bands were too close to each other on the 12% SDS-PAGE to be distinguished apart. A degradation band with a smaller molecular weight than IFT43 was also constantly detected on the Western blot, which we suspect to be a N-terminal degraded product of either IFT43 or IFT140 (**Figure 6**, blue asterisks). To assess whether the expression level of the IFT-A proteins are sufficient for protein production for further structural studies, I carried out Ni-NTA purification tests. The results suggested that the soluble proteins had very low abundance and multiple degradation bands appeared, indicating that some IFT-A proteins are not stable under current purification condition. Taken together, I have established the multiBac expression system for the *T. brucei* IFT-A complex, however, the expression needs to be further optimized so as to

increase the expression yield, and different purification buffers/strategies need to be tested to maintain the stability of the proteins.

6. Discussion

Using the multiBac expression system, I have established the protocol for co-expressing the 800-kDa IFT-A complex consisting of 6 protein components. Considering the large protein size, the complex was divided into two subgroups for expression, and expressed proteins in both groups have been identified with anti-His Western blots. This result should be further confirmed by Mass Spectrometry, especially for *Tb*IFT121 and IFT122.

The expression system that we used especially favors the expression of eukaryotic protein complex by taking advantage of insect cell culture, which provides necessary post-translational modifications to target proteins that might stabilize and/or solubilize the proteins. However, the expression level appears not high enough to proceed with further structural studies. In the future, expression should be further optimized to identify conditions for a higher yield. To monitor the expression level and to optimize the best harvesting time point, a donor carrying eYFP (enhanced yellow florescence protein) could be inserted into the bacmid via Cre-*loxP* recombination before the Tn7 transpostion takes place (Fitzgerald et al. 2007). By evaluating the florescent level of eYFP, the expression level of the IFT-A complex could be analyzed to decide when expression reaches the peak point and insect cells should be harvested. DH10EMBacY can be used for this purpose, which has YFP gene on the bacmid backbone already (described on Dr. Berger lab homepage http://www.embl.fr/multibac/multiexpression_technologies/multibac/MultiBac_manual.pdf).

To constitute the IFT-A complex *in vivo*, the proteins could be expressed via co-transfecting insect cells using the two viruses carrying the two subgroups of IFT-A components. This would ensure all interactions among different components to occur *in vivo*. For purification, the protein complex can be isolated first via Ni-NTA, which is followed by size exclusion chromatography to remove extra copies of certain components to obtain only the complex assembled in the correct stoichiometry. The complex should be further confirmed by mass spectroscopy and/or Western blot. Crystallization trials could be applied to the purified IFT-A complex to obtain high-resolution structural information.

In case that the whole complex could not be crystallized, the smaller IFT-A core complex instead should be tried for structural and/or biochemical studies. It has been reported that the

IFT-A core (IFT122, IFT144 and IFT140) could be obtained using high-salt stripping (Behal et al. 2012; Taschner et al. 2012). Compared with the IFT-A complex, the IFT-A core is much smaller in size, around 470 kDa, with the assumption that all subunits are single copies. Therefore, it might be easier to handle during purification and structural studies.

Similarly, the IFT-B complex and its core could also be expressed using this approach. Successful purification of both IFT-A and IFT-B complexes would offer an opportunity for combinatorial analyses of interactions between the two protein complexes, which still remain mysterious at present. Furthermore, future characterization of the interactions between IFT complexes and their accessory proteins, such as chaperones, specific kinases and phosphatases, would open another door for biochemical studies of ciliogenesis

7. Methods and materials

7.1. Constructs used for insect cell expression

The multigene baculoviral constructs which were designed for this study are listed in **Table 1**.

C-terminal tag	Protein	Type	vector
6×His tag	*Tb*IFT43+IFT121	Full length, wild type	pUCDM
6×His tag	*Tb*IFT43+IFT139	Full length, wild type	pUCDM
6×His tag	*Tb*IFT122	Full length, wild type	pSPL
6×His tag	*Tb*IFT144	Full length, wild type	pSPL
6×His tag	*Tb*IFT140	Full length, wild type	pFL

Table 1 All multigene baculoviral constructs used in this study

7.2. Generation of multigene expression cassettes

All IFT-A component constructs described in this study were expressed using the multiBac expression system (Fitzgerald et al. 2006). The acceptor plasmid pFL is a derivative of pFBDM (Berger et al. 2004) with an additional *loxP* sequence. pFL contains a high copy number ColE1 replicon and an ampicillin (Ap)-resistance marker. The donor plasmids are pUCDM (Berger et al. 2004) and pSPL. pUCDM and pSPL are identical except that the resistance markers are chloramphenicol and spectinomycin, respectively. All vectors contain two expression cassettes and a multiplication module that allows for adding additional cassettes. All vector sequences and further details about expression cassettes, recombination elements, and restriction maps can be found on the website of Dr. Berger lab homepage (http://www.embl.fr/multibac/multiexpression_technologies/multibac/index.html). Insert PCR amplifications were carried out with Phusion high-fidelity DNA polymerase (Finzymes). To create multigene expression cassettes, two *in vitro* Cre-*loxP* reactions were carried out. Donor plasmids pUCDM and pSPL were fused via *loxP* with the acceptor plasmid pFL by *in vitro* Cre-fusion reaction, which was carried out as recommended by the manufacturer (NEB). The first reaction generated the fused plasmid combining pUCDM containing *Tb*IFT43 and *Tb*IFT121, pSPL containing *Tb*IFT122 and pFL containing *Tb*IFT140. The second reaction generated the fused plasmid combining pUCDM containing *Tb*IFT43 and *Tb*IFT139, pSPL containing *Tb*IFT144 and pFL containing *Tb*IFT140. 5 microliters of the Cre recombination reaction were used to

transform competent Top10 *E.coli* cells and selection for growth was carried out by plating the transformed cells on LB agar plates containing Chloramphenicol, Spectinomycin and Ampicillin resistance markers. The plasmid DNA amplified from single colonies was verified by restriction digestion and further by DNA sequencing.

7.3. Generation of recombinant bacmids carrying the genes of interest

Multibac bacmids were generated by chemical transformation of the acceptor-donor fused plasmids into the DH10Bac *E. coli* strain. Transformed cells were incubated in a shaker incubator for at least 8 hours at 37°C to allow transposition of the donor-acceptor fusion expression construct into the lacZ gene of the bacmid. Afterwards, cells were plated onto a LB-Agar plate containing kanamycin (50 µg/ml), gentamycin (10 µg/ml), tetracycline (10 µg/ml), X-gal (500 µg/ml) and IPTG (0.5 mM) to enable blue-white selection of the clones containing the genes of interest (white colonies). White colonies were re-streaked and re-evaluated to confirm the result. Positive clones were amplified in an over-night cultures containing 50 µg/ml kanamycin, 10 µg/ml gentamycin and 10 µg/ml tetracycline in 5 ml LB-medium using a shaker incubator (200 rpm) at 37°C.

The recombinant bacmid DNA was purified with a modified version of the Qiagen Miniprep protocol. Cells were resuspended in 250 µl of Qiagen buffer P1, mixed with 250 µl of Qiagen buffer P2 and inverted the tubes 10-15 times to lyse the cells. 420 µl of Qiagen buffer P3 was added to precipitates cell debris. Lysate was cleared by centrifugation at 13,000×rpm for 20 minutes at room temperature. The supernatant was transferred into a clean tube, supplemented with 800 µl of ice cold isopropanol and thoroughly mixed. Precipitated bacmid DNA was pelleted by centrifugation at 13,000×rpm for 40 minutes at 4°C. The supernatant was decanted carefully. The bacmid DNA pellet was washed with 1 ml of 70% (v/v) ethanol. Ethanol in the supernatant was discarded after centrifugation at 13,000×rpm for 5 minutes at 4°C. The DNA pellets were resuspended in 50 µl of steril ddH$_2$O.

7.4. Generation of recombinant baculoviruses and protein expression in insect cells

To transfect insect cells, approximately 1.0×10^6 adherently grown *S.frugiperda* (Sf9) cells or High-Five cells (Invitrogen) per well grown in a 6-well plate (Nunc) were washed with 2 ml of Sf-900 II serum-free medium. To generate recombinant baculovirus, the transfection mix contain-

ing 6 µl of resuspended bacmid DNA and 6 µl of Cellfectin II Reagent (Invitrogen) in 200 µl of Sf-900 II serum-free medium (Gibco) was incubated for 30 minutes in a laminar flow cabinet. About 200 µl of transfection mix and 800 µl of Sf-900 II serum-free medium were added to one well and incubated for 4 hours at 27°C. After transfection, insect cells were washed once with 2 ml of Sf-900 II serum-free medium and then incubated in 3 ml of Sf-900 II serum-free medium for 72 hours. Baculoviruses start to be released into the medium 3 days post transfection and transfected cells appear larger in size and tend to detach.

To generate the 1st virus amplification (V1), 1 ml of baculovirus containing supernatant derived from bacmid transfection (V0) was added to semi-confluent Sf9 cells (~2X10^6 cells/ml) in a T-80 flask containing 25 ml of Grace's insect medium supplemented with 9% L-glutamine (Sigma-Aldrich), 100 U/ml penicillin-streptomycin solution (Sigma-Aldrich) and 5% heat inactivated fetal bovine serum (Gibco). After 5-6 days of incubation at 27°C, baculoviruses containing supernatant were filtered using 0.22 µm Steriflips (Millipore) and were stored at 4°C to protect it from light. To generate the 2nd virus amplification (V2), 5 ml of V1 was added to semi-confluent Sf9 cells (~2X10^6 cells/ml) or HF cells in a T-175 flask containing 50 ml of Grace's insect medium supplemented with 9% L-glutamine (Sigma-Aldrich), 100 U/ml penicillin-streptomycin solution (Sigma-Aldrich) and 1% heat inactivated fetal bovine serum (Gibco). Usually, the 3rd virus amplification (V3) was used for protein expression. For large-scale protein production, 600-800 ml of 0.8-1.0x10^6 Sf9 suspension cells or High-Five cells grown in Grace's insect medium supplemented with 9% L-glutamine (Sigma-Aldrich), 100 U/ml penicillin-streptomycin solution (Sigma-Aldrich) and 5% heat inactivated fetal bovine serum (Gibco) were inoculated with 15-20 ml of V3 viruses and incubated in a shaker incubator (100×rpm) at 27°C for 60-72 hours. After protein expression, insect cells were spun down at 1,500×rpm for 30 minutes at 4°C. Cell pellets were resuspended in lysis buffer containing 20 mM Tris-HCl (pH8), 300 mM NaCl, 0.5% Tween 20, 10% glycerol, 20 mM Imidazole pH 8, supplemented with 1 mM phenylmethanesulfonylfluoride (PMSF). The cells were broken using emulsiflex-C3 homogenizer (Avestin). The centrifugation at 13,000×rpm for 40 minutes at 4°C was carried out to obtain the supernatant containing soluble proteins.

7.5. Anti-His Western blot to detect expressed IFT-A complex proteins

IFT-A complex proteins were identified in cell lysates by western blot using a mouse antibody against 6×His tag (Calbiochem) and a secondary anti-Mouse IgG (H+L) antibody conjugated with alkaline phosphatase. The proteins were visualized by adding the Sigma FAST BCIP/NBT Tablet (Sigma-Aldrich).

PART 2: THE SAS-6 COILED COIL STRUCTURE AND ITS SPECIFIC INTERACTION WITH SAS-5 SUGGEST A REGULATORY MECHANISM IN CENTRIOLE ASSEMBLY

1. Abstract

The centriole is a conserved microtubule-based organelle essential for both centrosome formation and cilium biogenesis. Five centriolar proteins have been identified in *Caenorhabditis elegans* and their homologues in other species have also been reported. Two of them, SAS-5 and SAS-6, physically interact with each other and are codependent for their targeting to procentrioles. However, it remains unclear how these two proteins interact at the molecular level and why the *C. elegans* centriole has a unique central tube that is absent in non-nematode centrioles. Here, I demonstrate that the SAS-5 C-terminal domain (CTD, residues 390-404) specifically binds to the central region (residues 275-289) of the SAS-6 coiled coil. To further investigate their interaction, I have solved the crystal structure of the SAS-6 coiled coil domain (CCD, residues 248-410) and established that the association of the SAS-6 CCD and the SAS-5 CTD is mediated by synergistic hydrophobic and electrostatic interactions. The crystal structure also shows a periodic charge pattern along the SAS-6 CCD which, in the absence of SAS-5, gives rise to an anti-parallel tetramer of its CCD. Electron microscopy studies of the SAS-5/SAS-6 complex suggest that the central tube of *C. elegans* centrioles is formed by SAS-5 circularly arranged on SAS-6; SAS-5 alone forms aggregates. We further show that mutations of key residues within the CCD disrupt SAS-6 recruitment and function in centriole assembly *in vivo*. Overall our findings establish the molecular basis of the specific interaction between SAS-5 and SAS-6, and suggest that both proteins individually adopt a self-associated conformation that is disrupted upon the formation of the hetero-complex to facilitate the correct assembly of the nine-fold symmetric centriole.

2. Zusammenfassung

Das Zentriol ist ein konserviertes, mikrotubuläres Organell, welches für die Bildung von Zentrosomen und die Biogenese von Zilien verantwortlich ist. Fünf verschiedene Proteine, die für die Duplikation von Zentrosomen notwendig sind, wurden zuerst in *Caenorhabditis elegans* identifiziert, und ihre Homologe konnten später auch in anderen Spezies nachgewiesen werden.

Zwei dieser Proteine, SAS-5 und SAS-6, interagieren direkt miteinander und sind voneinander abhängig um zu den Prozentriolen lokalisieren zu können. Wie diese beiden Proteine miteinander interagieren und warum das Zentriol in *C. elegans* eine einzigartige tubuläre Zentralstruktur besitzt, welche nur in Nematoden zu finden ist, konnte bisher noch nicht geklärt werden.

In folgender Studie zeige ich, dass die C-terminale Domäne von SAS-5 (CTD, Aminosäuren 390-404) spezifisch in der zentralen Region der SAS-6 Coiled-coil Domäne (Aminosäuren 275-289) bindet. Um ihre Interaktion genauer zu untersuchen habe ich die Kristallstruktur der Coiled-coil Domäne von SAS-6 (CCD, Aminosäuren 248-410) gelöst und gezeigt, dass die Bindung von SAS-6 CCD und SAS-5 CTD auf hydrophoben und elektrostatischen Wechselwirkungen basiert. Weiters besitzt die Kristallstruktur von SAS-6 CCD eine periodische Ladungsverteilung entlang des Coiled-coils, wodurch das Protein in der Abwesenheit von SAS-5, ein antiparalleles Tetramer bildet. Elektonenmikroskopische Studien des SAS-5/SAS-6 Komplexes implizieren, dass SAS-5 zirkulär an SAS-6 bindet und so die zentrale Röhrenstruktur der *C.elegans* Zentriolen bildet. Interessanterweise formt SAS-5, dass ohne SAS-6 aufgereinigt wird, Aggregate.

Zusammengenommen zeigen die vorliegenden Ergebnisse die molekulare Grundlage der spezifischen SAS-5/SAS-6 Interaktion, und legen nahe, dass beide Proteine in der Abwesenheit des jeweils anderen selbst-assoziierte Konformationen einnehmen. Diese werden erst durch die Bildung des Heterokomplexes SAS-5/SAS-6 aufgelöst, so dass ein korrekter Aufbau des Zentriols ermöglicht wird.

3. Introduction

Centrioles are microtubule based, cylindrical shaped organelles that exist in most animal eukrayotic cells, although they were lost in yeast and vascular plants (Carvalho-Santos et al. 2011) (**Figure 1**). The centriole is usually a cylinder-like structure composed of nine microtubule triplets. It is ~0.5 μm long and ~0.2 μm wide and has appendages at the distal end upon maturation. There are variations of this structure, in which triplets are substituted by singlet or doublets or there are no appendages. Centrioles have two distinct functions that involve the biogenesis of two different organelles – centrosomes and cilia (Marshall 2007). The centrosome is the microtubule-organizing center of animal cells. It regulates cell motility, polarity and adhesion, and facilitates the organization of spindle poles during mitosis (Bettencourt-Dias and Glover 2007). Abnormalities of centrosomes have been firmly linked to a variety of cancers, mainly due to genomic instability caused by irregular cell division. In addition, during ciliogenesis the mother centriole anchors on the plasma membrane to become the 'basal body', which serves as a template to nucleate cilia or flagella. Cilia and flagella in turn play an important role in physiology, cell development and diseases.

3.1. Centriole structure

The centrosome, which is comprised of two centrioles and the surrounding pericentriolar material (PCM), is the major microtubule-organizing center (MTOC) in animal cells and regulates the nucleation and spatial organization of microtubules (MT). The two orthogonally arranged centrioles are surrounded by the PCM to form the centrosome. They are linked together by the interconnecting fiber until disengagement at the exit from mitosis. The centriole pair displays structural and functional asymmetry due to the generational difference between each member of the pair: the old fully mature mother centriole is distinguished by two sets of nine appendages at its distal end while the young immature daughter centriole, assembled during the previous cell cycle, is about 80% the length of the mother centriole (**Figure 2**) (Bornens 2012). In most ciliated cells, centrioles are composed of a cartwheel structure and nine triplets of MTs (Marshall 2001; Preble et al. 2000). In each triplet, the internal tubule is termed the A-tubule, followed by the B-tubule and C-tubule. The C-tubule does not extend to the distal end of the centriole.
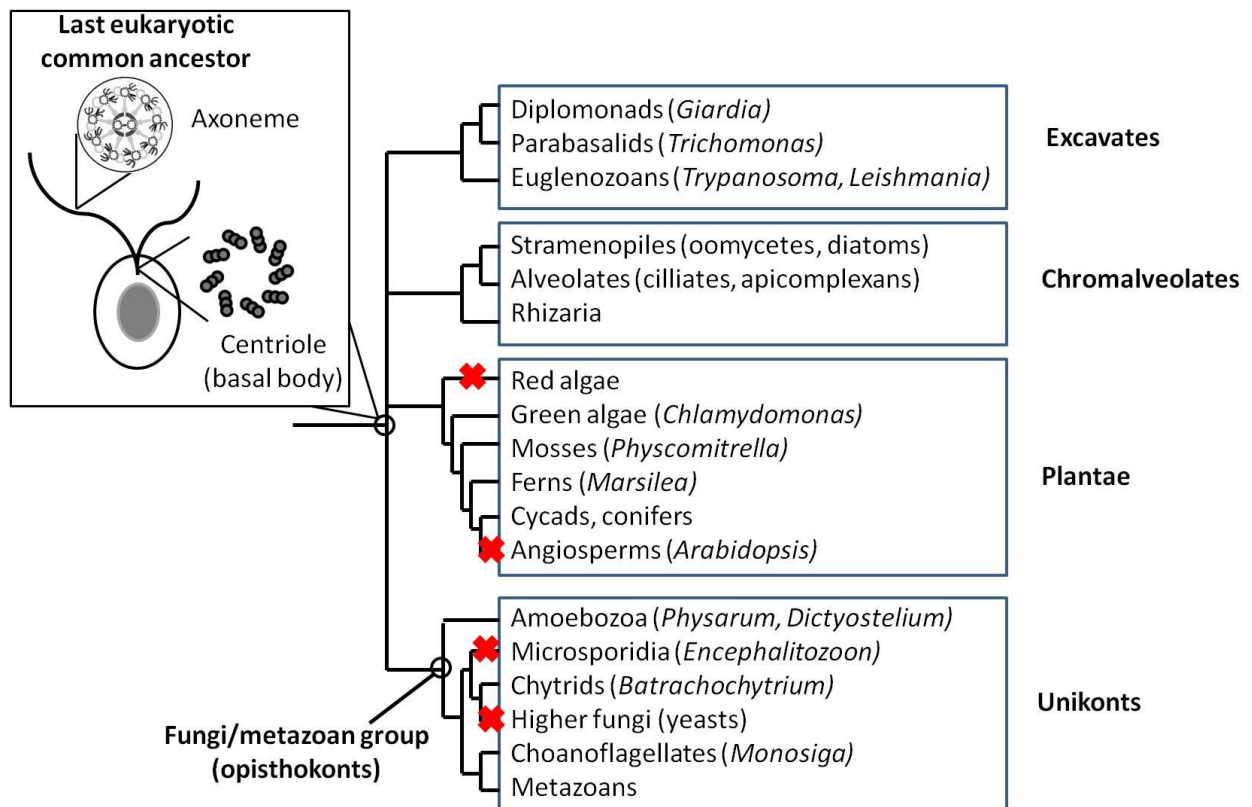
Figure 1 Conservation of the centriole/basal body structure throughout the organisms. The feathered microtubule core structure of cilia and flagella are present in most of eukaryotes. Centrioles and axonemes were lost in some species, such as angiosperms and higher fungi, which are marked by a red cross. *Adapted from* (Azimzadeh and Marshall 2010).

However, there are variations on centriole structure. For instance, centrioles may have 9 MT triplets, doublets or singlets, depending on species. It is 200-500 nm in length and 100-200 nm in diameter (**Figure 3**). Examples are the nine doublet MTs in cenrtrioles of *Drosophila melanogaster* embryos and the nine singlets in centrioles of *C. elegans* sperm cells and early embryos (Delattre and Gonczy 2004). Additionally, nematode centrioles presents a double-layered central tube, which is in contrast to the cartwheel structure seen in non-nematode species (Pelletier et al. 2006).
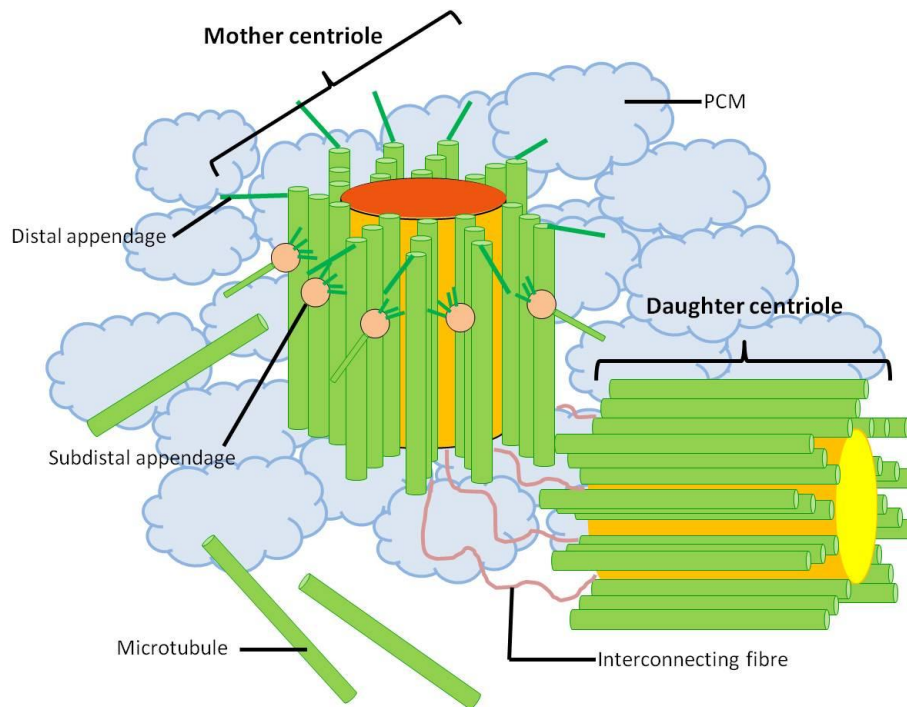
Figure 2 Schematic view of a centrosome containing mother and daughter centrioles. Adapted from (Bettencourt-Dias and Glover 2007)**.**

Despite the morphological differences, centrioles from different species share common origins. During the last decade, a number of conserved centriolar proteins have been identified (Azimzadeh and Marshall 2010; Carvalho-Santos et al. 2011). Functional genomics in nematodes and flies has identified a small set of conserved proteins required for the initiation of centriole/basal body assembly and for centrosome reproduction (Bornens 2012). The biochemical composition of the centrosome explains its structural complexity: more than 100 proteins are localized either in the centrioles or in the centrosomal matrix (Kobayashi and Dynlacht 2011; Nigg and Raff 2009), among which several are disease gene products. For example, all genes involved in microcephaly syndromes, which have been identified so far, encode centrosomal proteins (Megraw et al. 2011).
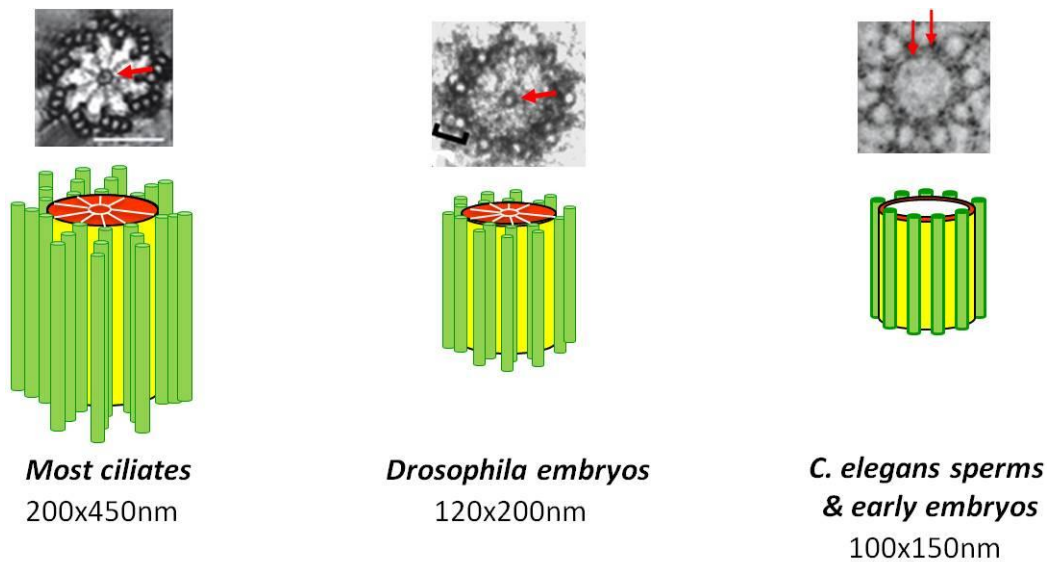
Figure 3 Different centriole structures of different species. The canonical centriole has 9 MT triplets and is 450 nm in length and 200 nm in diameter; for *Drosophila* melanogaster embryos, the centriole has nine doublet MTs and is 200 nm long and 120 nm wide; for *Caenorhabditis elegans* embryos, the centriole has singlet MTs and is 150 nm in length and 100 nm in diameter.

3.2. Centrioles functions

3.2.1. Centrosome as the microtubule-organizing center

As mentioned before, the centrosome acts as the main microtubule-nucleating organelle in animal cells and plays a critical role in mitotic spindle orientation and in genome stability (Bornens 2012). The most evident function is that the centrosome serves as the MTOC. During cell division, the centrosome shapes the bipolar mitotic spindle to ensure faithful chromosome segregation (Marshall 2009). At the prophase stage, centrosomes are associated with the nuclear membrane. Then the nuclear membrane breaks down so that centrosomes nucleated MTs can interact with the chromosomes to build up the mitotic spindle. Interestingly, however, centrioles are not required for the progression of mitosis. When the centrioles are irradiated by a laser, mitosis proceeds normally with a morphologically normal spindle. Moreover, development of the fruit fly *Drosophila* is largely normal when centrioles are absent due to a mutation in a gene required for their duplication (Basto et al. 2006). In the absence of the centrioles the MTs of the spindle are focused by motors, allowing the formation of a bipolar spindle. Many cells can completely undergo interphase without centrioles (Rieder et al. 2001). Unlike centrioles, centrosomes are required for survival of the organism. Acentrosomal

31

cells lack radial arrays of astral MTs. They are also defective in spindle positioning and inability to establish a central localization site in cytokinesis. The function of centrosomes in this context is hypothesized to ensure the fidelity of cell division because it greatly increases the efficacy.

3.2.2. Centriole as a template for cilia and flagella

Another distinct function of centrioles is to 'seed' cilia and flagella. Cilia and flagella are projections from the cell surface, which either enable movement of the cell itself or facilitate the movement or sensing of substances around cells. The cilium is composed of a MT-based core structure called the axoneme, which is surrounded by the ciliary membrane that is continuous with the plasma membrane. The axoneme is constructed from nine parallel doublet MTs known as outer doublets, which elongate from the basal body (Ishikawa and Marshall 2011). The axoneme of primary cilia typically has a ring of nine outer MT doublets (called a 9+0 axoneme), and the axoneme of a motile cilium has two central MT singlets in addition to the nine outer doublets (called a 9+2 axoneme) (**Figure 4**). In vertebrate cells, only the mature mother centriole can nucleate primary cilia. The appendages that are only present on mother centriole seem important. In mouse cells, knockout ODF2 (an appendage marker) leads to the mother centriole missing appendages and thus no cilia are formed (Ishikawa et al. 2005).

The majority of eukaryotic cells form cilia at certain point during their life cycle. Ciliogenesis begins when cells exit the cell cycle and enters a quiescent (G0 phase) and/or differentiated state, with the centrosome being translocated from the periphery of the nucleus to the plasma membrane. There, the cilium emanates from the basal body, which is converted from the mother centriole. The mature basal body is anchored to the plasma membrane and serves as the template for the outgrowth of the ciliary axoneme. For cells to exit G0 to re-enter the cell cycle, cilia have to be resorbed to convert the basal bodies back to centrioles to re-form centrosomes. Importantly, while centrioles are not strictly required for mitosis, they are indispensable for ciliogenesis (Pedersen and Rosenbaum 2008).

Abnormalities of cilia or flagella have now been firmly linked to a variety of human diseases, which are collectively called ciliopathies, including McKusick-Kaufmann syndrome, Joubert syndrome, Meckel-Gruber syndrome, Bardet-Biedl syndrome, Nephronophthisis, and Leber congenital amaurosis (Rachel et al. 2012).
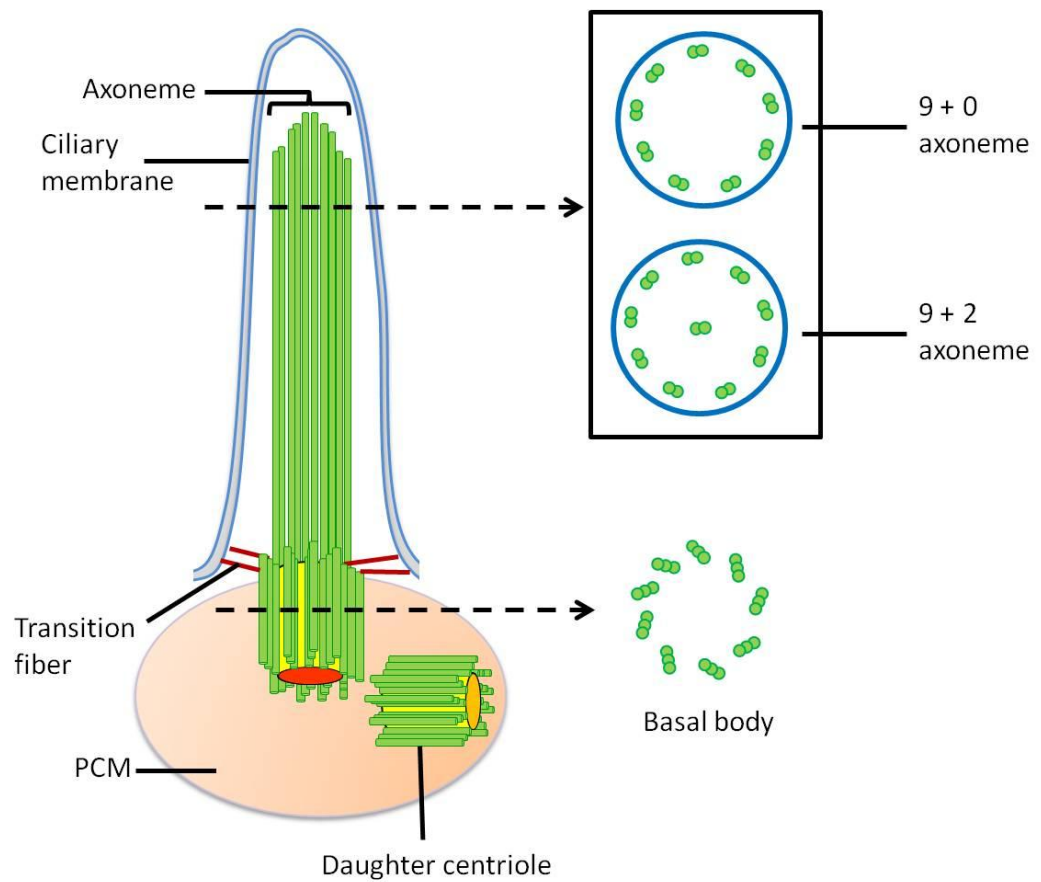
Figure 4 Schematic of the cilium structure. Primary cilia are non-motile, as their axonemal microtubules have only 9 microtubule pairs (9 + 0), while motile cilia have an additional pair of central microtubules (9 + 2) (Saudou 2012).

3.3. Centrosome cycle

There are four consecutive steps characterized for centrosome cycles: centriole disengagement, nucleation of the daughter centrioles, elongation of the procentrioles, and separation of the centrosomes (Kuriyama and Borisy 1981) (**Figure 5**). To maintain the constant centriole numbers in proliferation cells, it requires two types of controls. The first one is cell-cycle control. Every centriole must duplicate once and only once in every cell cycle. And the second is to ensure that only one new centriole is formed during centriole duplication. As summarized in **Figure 5**, centriole duplication begins at the G1 to S transition of the cell cycle with the formation of one procentriole adjacent to each pre-existing parental centriole.
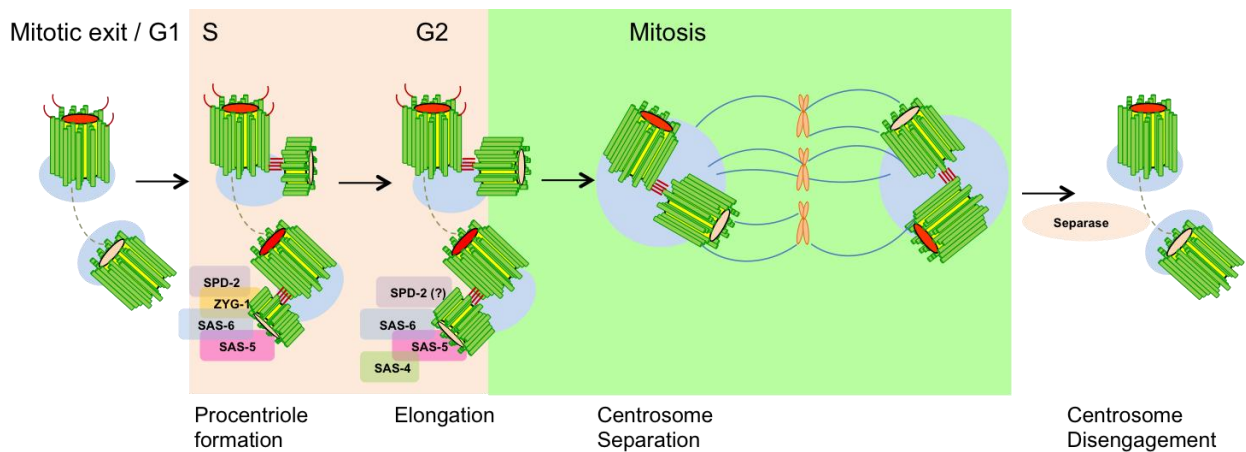
Figure 5 The centriole duplication cycle in *C. elegans* is coupled with the chromosome replication and segregation cycle. In *C. elegans*, SPD-2 and ZYG-1 recruitment initializes the procentriole formation in S phase, which is followed with the recruitment of SAS-6 and SAS-5. In G2, ZYG-1 diminishes and SAS-6 and SAS-5 form an elongate the central tube. SAS-4 incorporates into the newly forming procentriole to recruits MTs to complete centriole assembly. The daughter centriole reaches full elongation and maturation. When a cell exits mitosis, sister chromatid separation and centriole disengagement are triggered by the separase. Figures are redrawn based on (Debec et al. 2010).

The Polo-like kinase 4 (PLK4) initiates this event by phosphorylation its substrates, hence triggers the centriole duplication. In *C. elegans*, SAS-6 has been identified as a substrate of ZYG-1 (zygotic defective: emryotic lethal 1), which is a functional homolog of PLK4 (Kitagawa et al. 2009). However, it is not clear yet whether SAS-6 is also phosphorylated by PLK4 in other species. The importance of PLK4 during the centriole duplication has been underlined by the observation that overexperssion of PLK4 induces multiple newly formed centrioles adjacent to the mother centriole; conversely, deletion of PLK4 causes the reduction of centriole numbers (Habedanck et al. 2005).

The initiation of procentrioles is morphologically different in *C. elegans* compared with other species, which is a tubular structure rather than a cartwheel. But the proteins involved in procentriole initiation are conserved and limited in numbers. SAS-6 protein has been identified recently to be critical for the emergence of procentrioles (Kitagawa et al. 2011b; van Breugel et al. 2011). In *Chalamydomonas*, SAS-6 protein is localized at the cartwheel structure of centriole; when SAS-6 is null (*bld12*), the centriole lacks the central hub of the cartwheel structure (Nakazawa et al. 2007). Likewise, in *Drosophila*, SAS-6 is important for centriole assembly but itself cannot form the 9-fold symmetry. It is a tetramer by itself, but it can form the tubules

highly reminiscent of the cartwheel hub by co-experssion with its binding partner ANA2 (Gopalakrishnan et al. 2010; Rodrigues-Martins et al. 2007; Stevens et al. 2010b). Functional homologs of ANA2 are STIL in humans and zebrafish and SAS-5 in *C. elegans*, but they are structurally distinct based on bioinformatic analysis. Recent studies suggest that there are other proteins besides SAS-6 being crucial for centriole assembly initiation, one of which is strongly suggested to be SAS-5 (Delattre et al. 2004). Therefore, it would be very important to investigate the interaction between SAS-6 and SAS-5 so as to understand the mechanism of the initiation of the centriole assembly.

When centriole formation is initiated, the procentriole starts to elongate throughout S and G2 phase. Several proteins are required during procentriole elongation, which include SAS-4 (CPAP or CENPJ in humans) (Dammermann et al. 2008; Tang et al. 2009), POC5 (Azimzadeh et al. 2009), OFD1 (Singla et al. 2010), and CP110 (Schmidt et al. 2009). Disengagement of the centriole pair requires PLK1 function, a mitotic kinase (Tsou et al. 2009), as well as separase, the protease responsible for sister chromatid separation during metaphase-to-anaphase transition (Uhlmann et al. 2000). Available evidence suggests that PLK1 and separase are only active during mitosis. This fact favors the explanation for the coupling of centriole duplication to traverse of the cell cycle. It also ensures that the disengagement of the centriole pair is coordinated with chromatid separation and in the meantime prevents the formation of multiple centrosomes.

3.4. Centriole biogenesis and regulation in *C. elegans*

In metazoans, there is a dual parental contribution to forming the centrosome in the wild-type one-cell-stage *C. elegans* embryo. The sperm provides a pair of centrioles stripped of PCM components, whereas the oocyte is devoid of centrioles but provides an abundant source of PCM proteins. After fertilization, the parentally contributed centrioles recruit PCM components to form the centrosome in the zygote. This single centrosome undergoes duplication during S phase, yielding two centrosomes, each containing one mother and one daughter centriole, which assemble a bipolar spindle during mitosis (Leidel and Gonczy 2005).
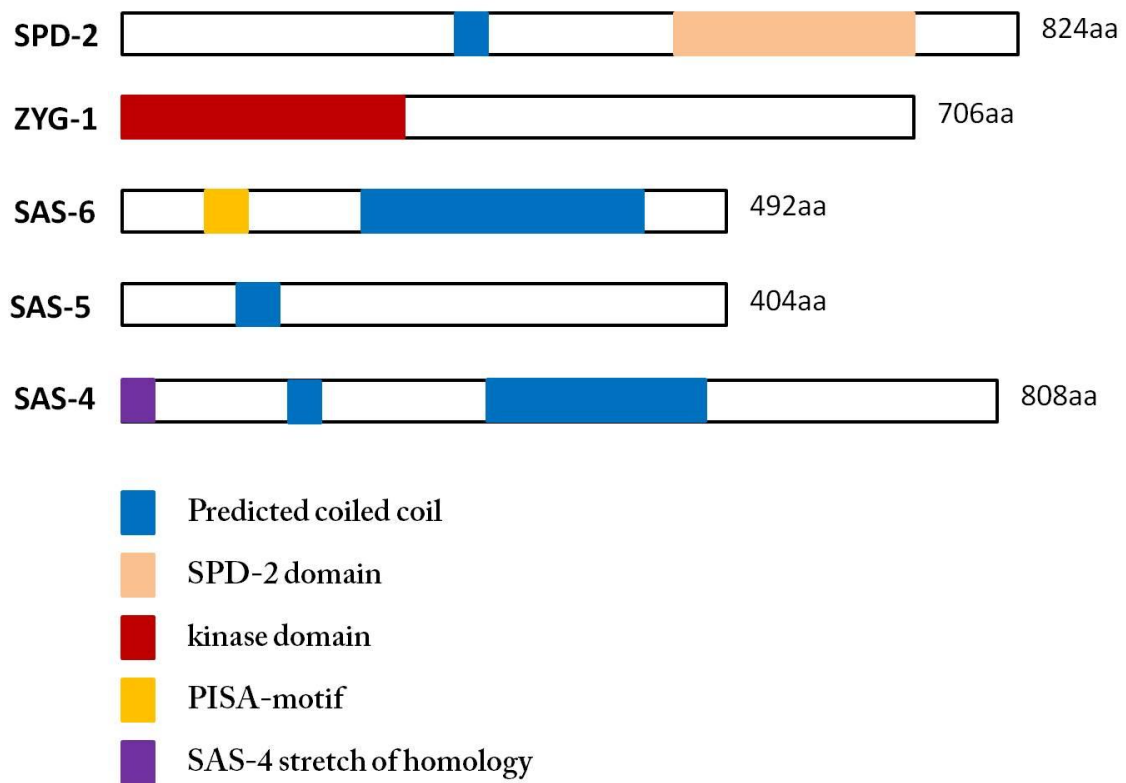
Figure 6 Structural schematic representation of proteins required for daughter centriole formation in *C. elegans* embryos. Predicted domains and motifs are represented by colored boxes. Adapted from (Leidel and Gonczy 2005).

Studies in a number of research groups have uncovered five *C. elegans* centriolar proteins, namely the kinase ZYG-1 (O'Connell et al. 2001), and the four coiled coil-containing proteins, SPD-2 (Kemp et al. 2004; Pelletier et al. 2004), SAS-4 (Kirkham et al. 2003; Leidel and Gonczy 2003), SAS-5 (Delattre et al. 2004), and SAS-6 (Dammermann et al. 2004; Leidel et al. 2005) (**Figure 6**). Homologues of these proteins have been identified in flies and humans (Andersen et al. 2003; Basto et al. 2006; Bettencourt-Dias et al. 2005; Habedanck et al. 2005; Hung et al. 2000; Leidel et al. 2005; Stevens et al. 2010a; Tang et al. 2011) (**Table 1**) .

Centrosomal proteins and their localization and interaction partners

Centrosomal proteins	Homologs	Interaction partners	Centrosomal localization		
			PCM, mother, daughter procentriole	Proximal distal along	Region
PLK4/Sak	Plk4 (Dm), Zyg1 (Ce)	Cep152, CPAP, Sas6, FBXW5, β-TrCP/Slimb	M, D, Pr	P, A	Outer walls, lumen
Cep152	Asl (Dm), Cep 152 (Dr)	CPAP, Plk4	M, D, PCM	P	Outer walls
Cep192	Spd2 (Ce, Dm)		M, D, PCM	A	Outer walls
hSas6	Sas6 (Ce, Dm, Dr, Tt, Pm), CrSas6/Bld12p (Cr)	Sas5, Ana2, Zyg1	Pr	P	Cartwheel (spokes, hub)
Cep135	Bld10 (Dm, Pm), Bld10p (Cr)	C-Nap1	M, D, Pr	P	Cartwheel, lumen, outer walls
Centrin	Cdc31, (Sc, Cp), CEN2/3 (Pt), CEN1 (Tt), VFL2 (Cr)	hPoc5, CP110	M, D, Pr	Di	Lumen
Cep120	Cep120(Mm), Uni2 (Cr)	Ninein, Cep164, Cep290	M, D, Pr	A	Outwalls
CPAP	Sas4 (Ce, Dm)	γ-tubulin, αβ-tubulin, Cep152, Plk2, Plk4	M, D, Pr	P, A, Di	Cartwheel, lumen, walls
γ-tubulin	γ-tubulin (Dm, Pt), Gtu1 (Tt), Tbg1 (Ce), Tubg (Mm), Tug (Cr)	CDK5RAP2 CPAP, Cep170	M, D, Pr, PCM	P	Lumen
Centrobin	CG5690 (Dm), Cntrob (Mm), Nud1p (Sc), Cdc11p (Sp)	αβ-tubulin, Plk1, NEK2	D, Pr	P	Lumen, outer walls
CP110	CP110 (Dm)	Centrin, Cep97, Kif24, Cep76, Cep290	M, D, Pr	Di	Cap
Cep97	Cep97 (Dm)	CP110, Cep76, Kif24	M, D, Pr	Di	Cap
Cep76	Cep97 (Xl, Dr)	CP110, Cep97, Kif24	M, D, Pr	Di	
ε-tubulin	Bld2 (Cr), Tube1 (Dr, Mm, Xl)	EB1	M, D, Pr	Di, A	Subdistal appendages, outer walls
δ-tubulin	Tubd (Dr, Mm, Xl) Uni3 (Cr), δPT1 (Pt)		M, D, Pr	P	outer walls
hPoc5	Poc5 (Cr, Pm)	Centrin	M, D	Di, A	Lumen

hPoc1	Poc1 (Dm, Cr, Tt, Pm)	αβ-tubulin	M, D, Pr	Di, A	Lumen, walls
Ofd1	Ofd1 (Mm), BUG11 (Cr)	γ-tubulin	M, D, Pr	Di, A	Lumen
Ofd2	Ofd2 (Mm)	Ninein, Trichoplein	M	Di	All appendages
Cep164	XP_929307 (Mm), NP_611787 (Dm), XP_697015 (Dr)		M	Di	Distal appendages
		γ-tubulin, Odf2			
Ninein	Nin (Dr, Mm)	Trichoplein, EB1	M	Di	Subdistal appendages
EB1	Mal3 (Sp), Bim1 (Sc)	CDK5RAP2, FOP, Cep290, Cep170	M	Di	Subdistal appendages
Cep170	Cep170 (Mm)	Plk1, EB1	M	Di	Subdistal appendages
CAP350	Cep350 (Mm)	FOP	M, D	A	
FOP	Fgfr1op (Mm)	CAP350, EB1	M, D	A	Outer walls
Kif24	Kif24 (Mm)	CP110, Cep97	M	Di	Subdistal appendages
CDK5RAP2	Cnn (Dm), CDK5RAP2 (Mm)	Cdc20, PCNT, γ-tubulin, EB1	M, PCM	Di	Outer walls, append-ages
C-Nap1	Cep250 (Mm)	Cep135	M, D	P	Linker
Plk1	Cdc5 (Sc), Plk1 (Dr, Mm), Plk1/2 (Ce), Plo1 (Sp), Polo (Dm)	Cep170	PCM		
Plk2	Plk2 (Dm, Mm), Plk2b (Dr)	CPAP	M, D		
β-TrCP	Slimb (Dm), β-TrCP (Mm)	SKP1, Plk4	Centrosome		
Cul1	Cul1 (Ce, Mm, Sp), Cul1a (Dr)	SKP1, SKP2, PPP1CA	M	Di	
Stil	Ana2 (Dm), Sas5 (Ce)	Sas6 (Ce, Dm)	M, D, BB	P, Di	

Table 1 The homologs, interaction partners, and centriolar localization details of human centriole proteins. Abbreviations in 'Centrosomal localization' in the table: M, mother centriole; D, daughter centriole; Pr, procentriole; P, proximal region; Di, distal region; A, along the centriole. Abbreviations in 'Homologs' in the table: Ce, Caenorhabditis elegans; Cr, Chlamydomonas reinhardtii; Dm, Drosophila melanogaster; Dr, Danio rerio; Mm, Mus musculus; Pt, Paramecium tetraurelia; Sc, Schizosaccharomyces cerevisiae; Sp, Schizosaccharomyces pombe; Tt, Tetrahymena thermophila; Xl, Xenopus laevis. Adapted from (Brito et al. 2012)

RNAi and mating-based assays in *C. elegans* have shown that centriole duplication is a multistep process, with the five centriolar proteins being recruited in a sequential manner (Delattre et al. 2006; Pelletier et al. 2006). Firstly, SPD-2 is brought close to the mother centriole. The kinase ZYG-1, which is required for the subsequent recruitment of the SAS-5/SAS-6 complex, is then

incorporated into the nascent daughter centriole. SAS-5 and SAS-6 together form the initial central tube and following their arrival ZYG-1 dissociates. Subsequently, SAS-4 is recruited to build an outer wall of the central tube. Finally, nine singlet MTs are assembled around the central tube to generate a daughter centriole that is identical to the mother (**Figure 7**).
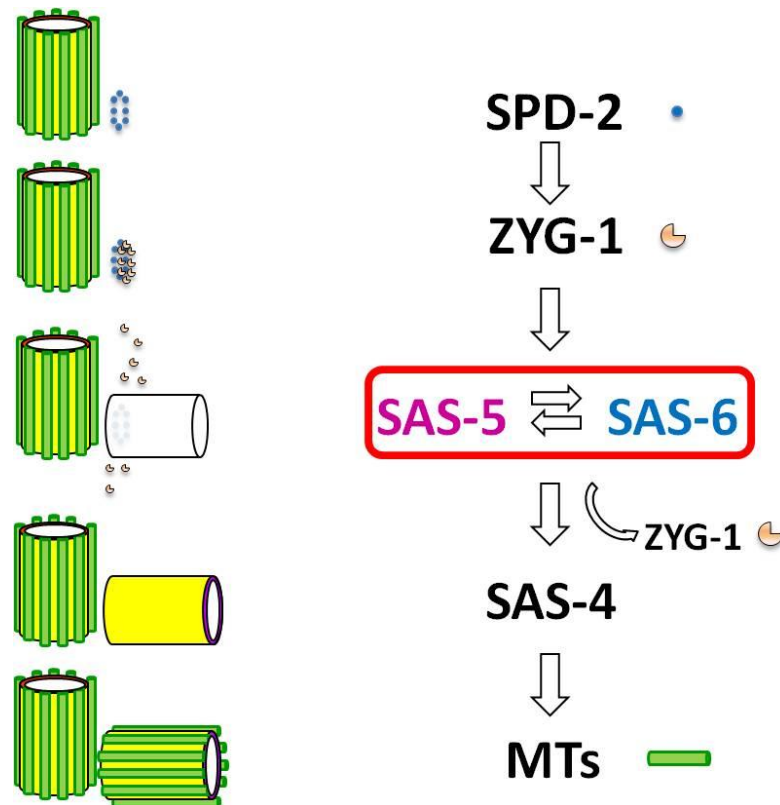
Figure 7 Schematic representation of daughter centriole assembly during the first mitotic division in *C. elegans*. SPD-2 and ZYG-1 are recruited during meiosis, before daughter centriole assembly. SPD-2 is required to recruit ZYG-1 to the mother centriole. Both SPD-2 and ZYG-1 are required for the recruitment of the SAS proteins, coincident with the formation of the daughter centriole central tube. SAS-5 and SAS-6 are required for SAS-4 recruitment, which in turn recruits MTs. My work has focused on the SAS-5/SAS-6 complex.

3.4.1. SPD-2

SPD-2, spindle defective protein-2, is a component of the PCM. It has been found to be the first loaded and required for the centriolar localization for other proteins (Kemp et al. 2004; Pelletier et al. 2004). Intriguingly, SPD-2 family members are only found in the genomes of Unikonts, a branch of the eukaryotic tree comprising the Fungi/Metazoan group as well as Amoebozoa, such as the model organism *Dictyostelium discoideum* (Pelletier et al. 2004). The

results of the database searches with the sequence of the *C. briggase* ortholog of SPD-2 identified one human and one *D. melanogaster* homologous protein. There are six putative CDK phosphorylation sites conserved in nematodes (**Figure 8**). Among these species, the homologous region is restricted to a 200 amino acid domain, called SPD-2 domain, which is located at the C-terminal part of the protein in nematodes. The similarity between *C. elegans* SPD-2 domain and its human and *Drosophila* orthologs are 15%/36% and 14%/34%, respectively, suggesting that the function of the SPD-2 domain might be conserved. In addition, like *C. elegans* SPD-2, its orthologs also comprise a large coiled coil domain preceding the SPD-2 domain (**Figure 8**).
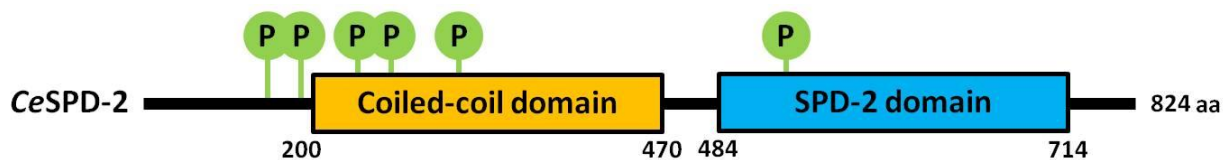
Figure 8 Domain structure of *C. elegans* SPD-2. SPD-2 has 824 amino acids, containing a coiled coil domain (aa 200–470, orange box) and a SPD-2 domain (aa 484–714, blue box) defined by similarity to its orthologs in other species. There are six conserved CDK phosphorylation sites at aa positions 143, 171, 220, 233, 259 and 545 in *C. elegans* SPD-2. The putative CDK phosphorylation sites are circled in green. *Ce*, *C. elegans*. Adapted from (Pelletier et al. 2004).

Previous studies on SPD-2 orthologs in human and *Drosophila* indicate that the primary function of SPD-2 is to recruit PCM proteins. In *C. elegans* embryos, SPD-2 localizes to the PCM and the centrioles, suggesting that SPD-2 is also a centrosomal component (Pelletier et al. 2004). It genetically interacts with the kinase ZYG-1 (Kemp et al. 2004). The human ortholog of SPD-2, called Cep192, is a major regulator of PCM recruitment, centrosome maturation and centriole duplication (Zhu et al. 2008). Interestingly, its ortholog in *Drosophila* seems dispensable for centriole duplication (Dix and Raff 2007). Taken together, *C. elegans* SPD-2 shares homology with human and *Drosophila* centrosome proteins and involves in both PCM recruitment and centriole duplication.

3.4.2. ZYG-1

ZYG-1 is a Ser/Thr kinase. As mentioned above, centriole assembly is regulated by the Polo-like kinase Plk4/SAK, and ZYG-1 in *C. elegans* is suggested to be the functional ortholog of PLK4 (Giddings et al. 2010; Kitagawa et al. 2009; Leidel and Gonczy 2005). PLK4 and SAK are necessary for centriole duplication in human and *Drosophila* cells, respectively (Bettencourt-Dias et al. 2005). Over-expression of PLK4/SAK4 induces formation of multiple new centrioles in these cells (Kleylein-Sohn et al. 2007; Basto et al. 2008; Peel et al. 2007). In *C. elegans*, ZYG-1 transiently localizes to centrosomes. The two parental ZYG-1 alleles play different roles. Paternal ZYG-1 regulates duplication and bipolar spindle assembly while maternal ZYG-1 regulates these processes thereafter. It has been reported that there are several temperature-sensitive mutations which affect ZYG-1 regulatory function, such as *oj7* and *b1* (O'Connell et al. 2001). Interestingly, it has been shown that truncations of ZYG-1 blocked centrosome duplication in mitotic cycle but leaded to centrosome amplification in the meiotic cycle. The mechanism of ZYG-1 regulation in centriole assembly still remains elusive. Nevertheless, the phosphorylation of centriole conserved protein SAS-6 at serine 123 *in vitro* has been demonstrated recently. Moreover, such a phosphorylation is shown to be required for the maintenance of SAS-6 at the newly formed centriole (Kitagawa et al. 2009).
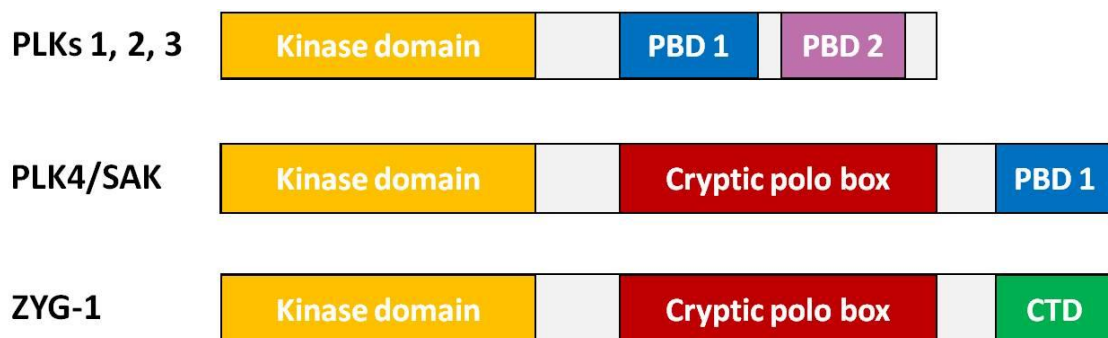
Figure 9 Structure Schematic representation of Polo-like kinases and their functional homologous proteins. Predicted domains and motifs are represented with colored boxes. PBD: Polo-box domain, CTD: C-terminal domain.

Structurally, ZYG-1 shares certain similarity with its functional homologs (**Figure 9**). There are four types of Polo-like kinases, PLKs 1-4. All of them share a conserved N-terminal Ser/Thr protein kinase domain. PLKs 1-3 are also similar in their C-terminal part by the presence of two highly conserved polo-box domains (PBDs). However, PLK4 differs from PLKs 1-3 significantly in

its C-terminal region. Instead of two PBDs, PLK4 has a central cryptic polo-box (CPB) domain and a C-terminal PBD. It has been shown that both the CPB and the PBD in PLK4 form a dimer (Leung et al. 2002). ZYG-1 is the putative homolog of mammalian PLK4 in *C. elegans*. Bioinformatical analysis indicates that ZYG-1 has three domains: an N-terminal kinase domain, a central CPB, and a globular C-terminal domain (CTD) (**Figure 9**). The CTD of ZYG-1 is necessary for targeting to centrosomes and is important for differentiating mitotic and meiotic centriole duplication (Peters et al. 2010).

3.4.3. SAS-6 and SAS-5

Recently, crystal structures of the N-terminal head group of SAS-6 from several organisms have been determined, which suggested that SAS-6 could self-associate *in vitro* into assemblies akin to the central hub of the cartwheel via the head domain (**Figure 10**) (Kitagawa et al. 2011b; van Breugel et al. 2011).

Figure 10 Crystal structure of *C. elegans* SAS-6 N-terminal domain dimer (PDB code: 3PYI). I154 forms a hydrophobic plug inserting into a hydrophobic socket formed by a second SAS-6 N-terminal domain, as shown in stick presentation in details (Kitagawa et al. 2011a).

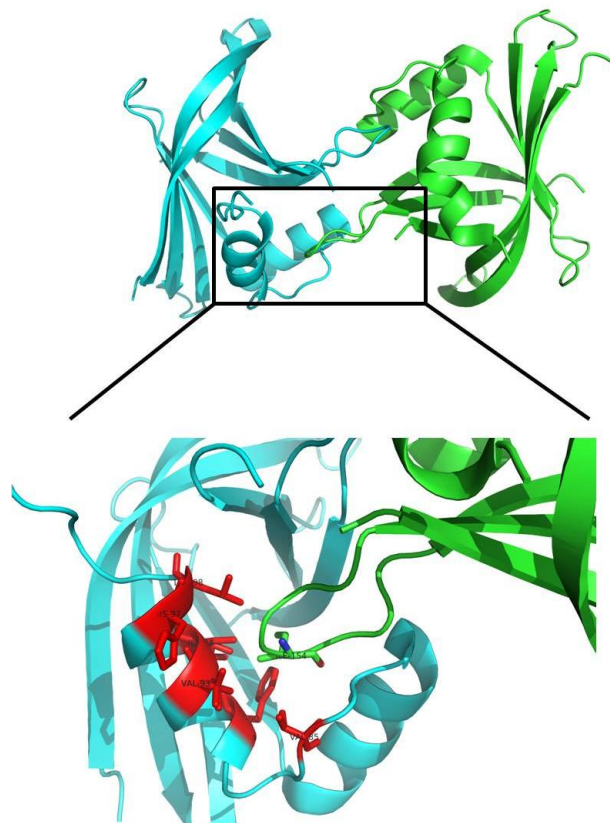This head domain interaction takes place mainly through a residue I154, which forms a hydrophobic plug that inserts into a socket located toward the end of the PISA motif of the neighboring molecule (**Figure 10**). Both I154 and the residues shaping the hydrophobic cavity are well conserved among SAS-6 orthologs, suggesting functional relevance. However, whether SAS-6 alone could faithfully drive the formation of the strict nine-fold symmetry of centrioles or whether it merely has a propensity to oligomerize into circular structures but the strict nine-fold symmetry is enforced by interactions with other factors is still a matter of some debates (**Figure 11**) (Cottee et al. 2011).
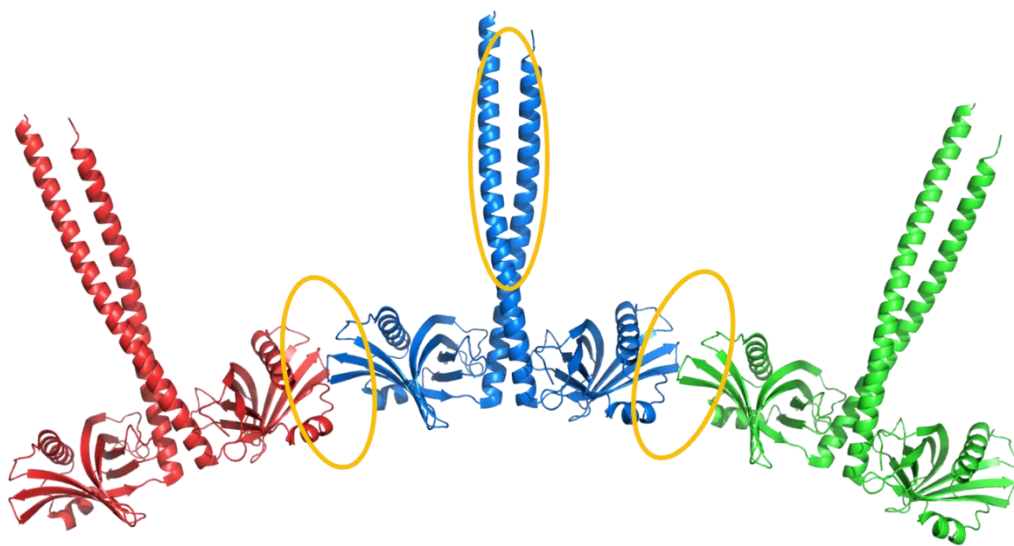
Figure 11 The interactions of CrSAS-6 take place via the coiled coil interaction and N-terminal head interaction, which is through a hydrophobic plug. The coiled coil domain initializes the SAS-6 dimerization and the head domain interaction leads to SAS-6 oligomerization, forming a ring structure. Adapted from (Cottee et al. 2011)

Indeed, modelings based on the available crystal structures show the potential for SAS-6 to adopt variant oligomeric conformations: spirals of different orientations or a flat ring that contains eight or nine dimers (**Figure 12**). For example, *Cr*SAS-6 head domain crystal structure supports both a flat ring oligomer structure and a left-handed spiral, whereas the *D. rerio* structures can be modeled into either a right-handed spiral or an almost flat ring with approximately eight-fold symmetry (**Figure 12**). Therefore, the head domain of SAS-6 is seemingly not sufficient to answer the question how the oligomer is arranged.
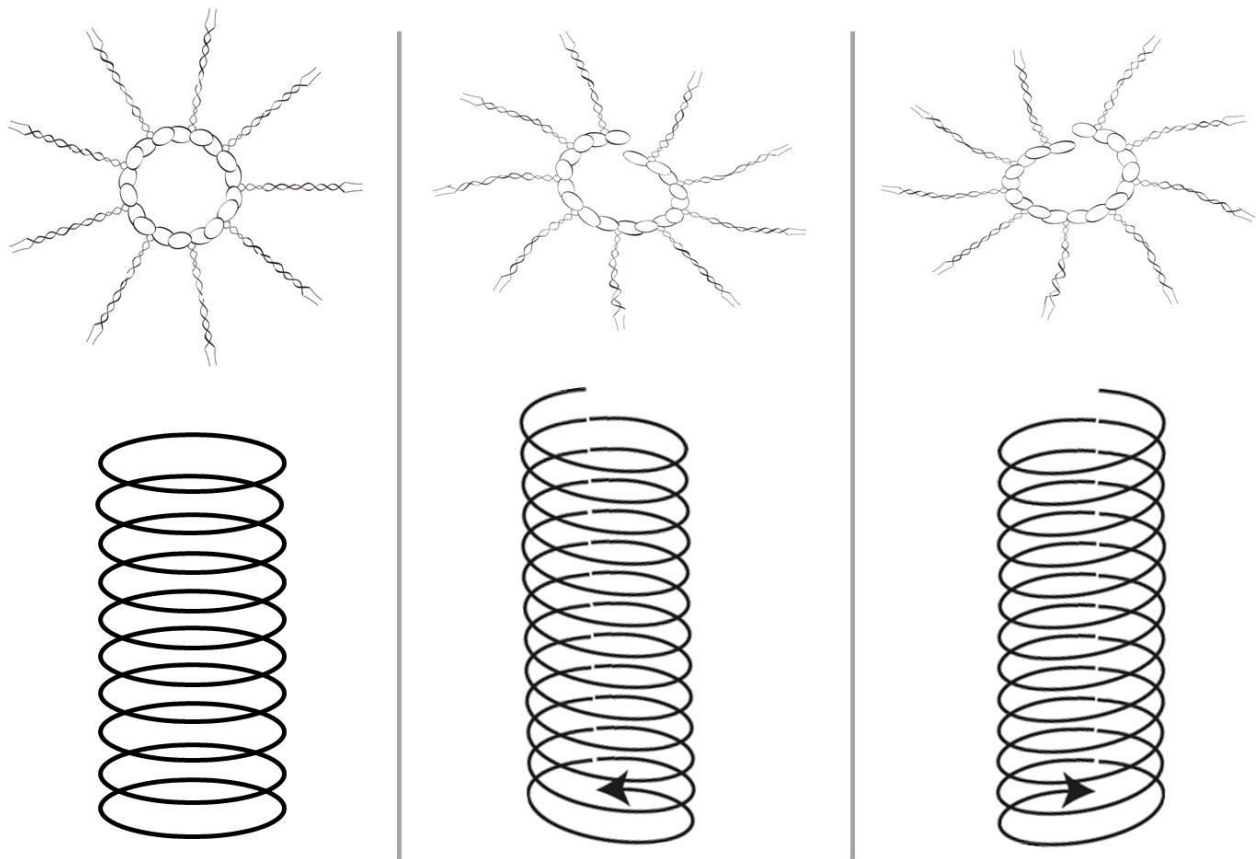
Figure 12 Possible SAS-6 oligomer models generated using different SAS-6 N-terminal head domain structures from *Chlamydomonas* and Zebra fish. Three alternative models are shown. SAS-6 N-terminal crystal structure from *Chlmaydomonas* could form either a flat ring structure (on left side) with ten dimer per circle or a left-handed spiral structure (in the middle) with ten dimer per helical turn. Similarly, SAS-6 N-terminal crystal structure from Zebra fish could generate a flat ring (on left side) with eight fold symmetry or a right-handed spiral structure with eight or nice dimers per helical turn (Cottee et al. 2011).

Even if we assume that SAS-6 is sufficient to establish the nine-fold symmetry, the spoke dimers would have to associate rigidly with a precise geometry to generate a 40° angle. Can it be restrained by SAS-6 head domains alone? Although the binding affinity between the head groups of *C. elegans* SAS-6 is comparable to that of *C. reinhardtii*, *H. sapiens*, and *D. rerio SAS-6*, recombinant *C. elegans* SAS-6 alone does not form a cartwheel-like structure similar to that of non-nematode SAS-6 proteins (Kitagawa et al. 2011b; Pelletier et al. 2006; van Breugel et al. 2011). Notably, the head group interaction of different SAS-6 proteins (dissociation constant, K_d ~ 60 - 110 µM) is relatively weak and has been thought unlikely to be the driving force for forming the nine-fold symmetry (Cottee et al. 2011).

Additionally, the oligomeric structure of recombinant *D. melanogaster* SAS-6 (*Dm*SAS-6) is different from the *in vivo* structure of centrioles. It was characterized as a stable tetramer by electron microscopy (Gopalakrishnan et al. 2010). Furthermore, velocity sedimentation analyses indicate that the endogenous *Dm*SAS-6 is in significantly denser fractions than recombinant *Dm*SAS-6, suggesting *Dm*SAS-6 probably needs the presence of additional factors that are required to enhance the structural assembly (Gopalakrishnan et al. 2010). Taken together, these data suggest that faithful duplication of the strict nine-fold symmetric centrioles likely requires other symmetry-ensuring factors.

It was shown previously that, although over-expression of *Dm*SAS-6 alone resulted in an irregular tube-like structure, co-expression of *Dm*SAS-6 with Ana2, the putative *Drosophila* ortholog of SAS-5, generated a highly ordered tubular structure, the SAS tubule. This structure looked the same as the *in vivo* cartwheel structure, suggesting that Ana2 assists SAS-6 in *Drosophila* centriole assembly (Stevens et al. 2010b). Similarly, STIL, the vertebrate homologue of SAS-5, regulates centrosome integrity (Castiel et al. 2011), and depletion of either SAS-6 or STIL made the other protein fail to target to the procentriole, implying that SAS-6 and STIL in vertebrates are also mutually dependent for centriolar localization (Arquint et al. 2012; Tang et al. 2011; Vulprecht et al. 2012). Earlier experiments carried out in worms revealed that SAS-5 and SAS-6 physically interact with each other for their codependent centriolar localization and that centriole duplication failed in embryos with a *sas-5*-mutant that fails to interact with SAS-6, indicating that SAS-5 works synergistically with SAS-6 in *C. elegans* centriole assembly (Delattre et al. 2004; Leidel et al. 2005). Altogether, these findings suggest that the SAS-5/Ana2/STIL family of proteins is likely the extra factor needed for SAS-6 to generate the nine-fold symmetry of centrioles.

In *C. elegans*, SAS-5 is a dynamic protein that shuttles between centrioles and cytoplasm throughout the cell cycle, whereas SAS-6 is stably located in centrioles. It has been found that SAS-5 centriolar localization is regulated by ZYG-1. In double allele mutants of *SAS-5*, the cells had monopolar spindle assembly and failed to divide, suggesting that SAS-5 presence is crucial for daughter centriole formation (Delattre et al. 2004). Bioinformatic analysis suggests that SAS-5 consists of a coiled coil domain, and its orthologs in the related nematode species can be readily identified by NCBI BLAST.
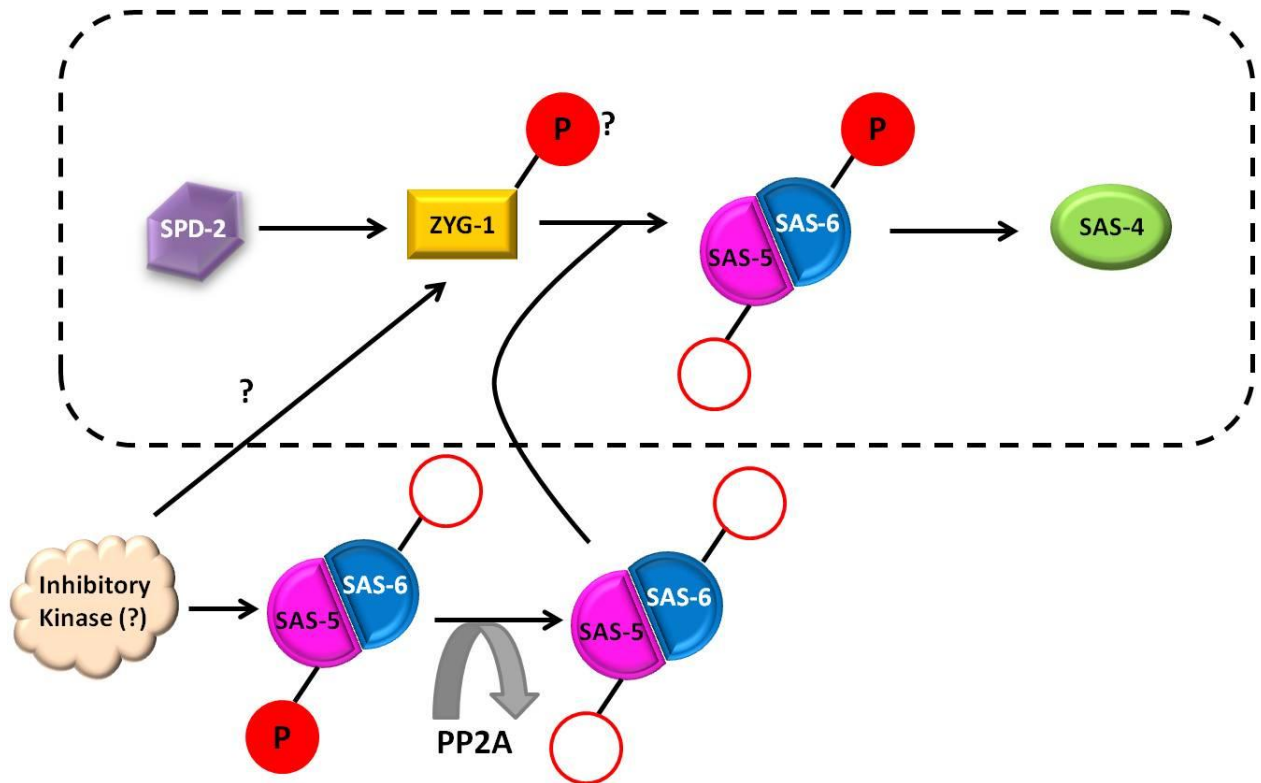
Figure 13 Role of PP2A in the centriole assembly pathway Dephosphorylation of SAS-5 by PP2A assists its localization to centrioles. However, a putative kinase, phosphorylating SAS-5 is still not known. It is suggested PP2A might also act on ZYG-1 to stabilize it. It seems that recruitment of SAS-5/SAS-6 to centrioles needs both the phosphorylation of SAS-6 by ZYG-1 and dephosphorylation of SAS-5 by PP2A (Megraw 2011).

Recently, it has been reported that ZYG-1 and SAS-5 level is regulated by protein phosphatase 2A (PP2A) (**Figure 13**). In PP2A subunit SUR-6 deleted embryos, the levels of both ZYG-1 and SAS-5 significantly decreased and SAS-6, the recruitment of which depends on ZYG-1 and SAS-5, failed to take on at the nascent centriole (Song et al. 2011). However, PP2A phosphatase does not antagonize ZYG-1 kinase. PP2A dephosphorylates SAS-5 while ZYG-1 phosphorylates SAS-6. It has been reported that ZYG-1 is required for SAS-6 maintenance rather than its recruitment, by serine phosphorylation at position 123 (Kitagawa et al. 2009). ZYG-1 might act on SAS-5/SAS-6 after PP2A-SUR6, which does not localize to centrioles (**Figure 13**). Moreover, it has been shown that in PP2A-deleted embryos, SAS-5 is undetectable in centrioles, indicating that PP2A might dephosphorylate SAS-5 before its localizing at centrioles (Megraw 2011).

It was previously reported that SAS-5 binds to the SAS-6 coiled coil and that SAS-6 fails to interact with the *sas-5* (*t2079*) mutant, which corresponds to a C-terminal single residue mutation of

SAS-5, R397C (Leidel et al. 2005). Live imaging studies of centriole duplication in *C. elegans* indicate that the SAS-5/SAS-6-based central tube is the first observable structure in procentrioles, which grows wider upon the recruitment of SAS-4 (Pelletier et al. 2006). However, it remains elusive what the molecular mechanism of SAS-5 and SAS-6 interaction is and how the central tube of *C. elegans* centrioles is formed.

3.4.4. SAS-4

SAS-4 (spindle assembly defective 4) has been identified as a component necessary for proper spindle assembly in *C. elegans* (Leidel and Gonczy 2003; Salisbury 2003a). Its importance is underlined by the fact that in *sas-4* RNAi treated cytoplasm, the pair of centrioles from sperm individually get mature but fail to duplicate, which leads to the blastomeres of the two-cell embryos inheriting a centrosome with only a single centriole (**Figure 14a**). SAS-4 is an 808-residue protein with a calculated molecular weight of 92 kDa. It contains two coiled coil domains, which are common motifs for mediating protein-protein interactions. Such a character is shared by the majority of centrosome related proteins (Salisbury 2003b). BLAST result of SAS-4 shows that it contains a short N-terminal motif that spans about 70 amino acid residues and is conserved among different species (**Figure 14b**). In *C. elegans*, SAS-4 localizes to the centrosome during the cell cycle and determines the centrosome size (Kirkham et al. 2003). CPAP, the homolog of SAS-4 in human, interacts with γ-tubulin complex (Tang et al. 2009) and regulates the daughter centriole length (Tang et al. 2009). The functional similarity of SAS-4 between species suggests SAS-4 is a centriole assembly organizer.
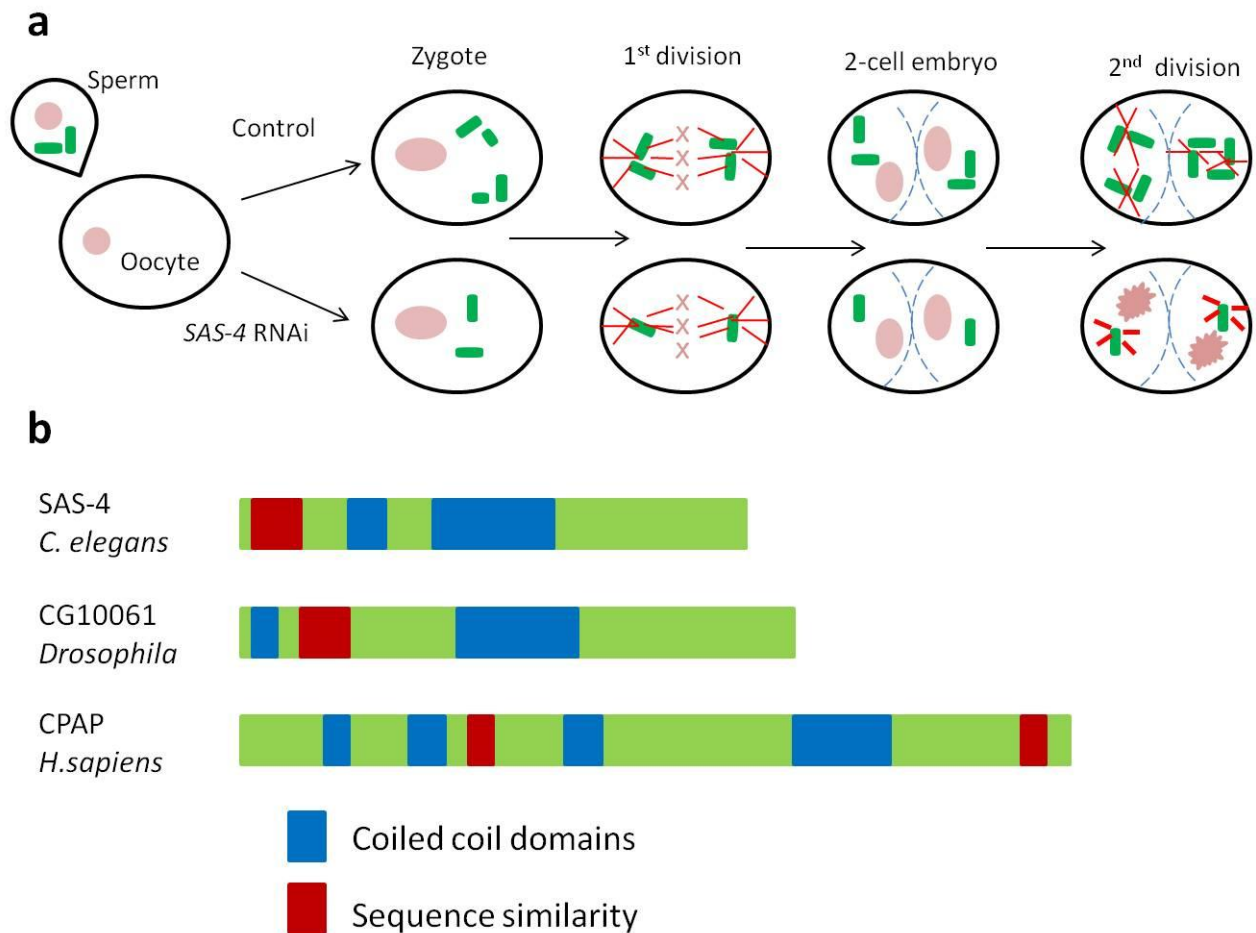
Figure 14 Schematic representations of SAS-4 affecting centrosome dynamics in *C. elegans* cell cycle and structural comparison with its orthologs. **(a)** In contrast to wide type worms (upper row), sas-4 RNAi treated worms (lower row) show that the sperm contributed centrioles are not able to duplicate. Although centrioles separate to form bipolar during the first cell division, each spindle pole only inherits one centriole, which in turn leads to the failure of bipolar spindle assembly in the second division cycle. **(b)** Structural features of *C. elegans* SAS-4 (808 amino acids) and its orthologs ORF CG10061 (891 amino acids) in *Drosophila melanogaster* and CPAP (1338 amino acids) in *Homo sapiens*. Predicted domains and motifs are represented with colored boxes. (Salisbury 2003a).

4. Aim of the study

The centriole is a conserved nine fold symmetric MT-based organelle essential for both centrosome formation and cilium biogenesis. Five centriolar proteins initially identified in *C. elegans* are essential for centriole assembly, and their related proteins serve analogous functions in other organisms. Among them, SAS-6 plays a crucial role of assembling the nine-fold radial symmetry of centriole across evolution. SAS-5 physically interacts with SAS-6 and may act as a symmetry-ensuring factor in centriole assembly.

I propose to carry out structural studies and biochemical characterization on SAS-6 and SAS-5 in hopes of gaining insight into how SAS-6 and SAS-5 interact at the molecular level and how the central tube of the procentriole forms. To achieve these goals, high-resolution structure information of SAS-6, SAS-5, and the SAS-6/SAS-5 complex needs to be obtained using X-ray crystallography, which will allow us to visualize the single proteins and/or the SAS-5/SAS-6 complex and to understand how the complex contribute to the nine fold symmetry. In case that structure determination of the SAS-6/SAS-5 complex is unsuccessful, the complex will be examined using electron microscopy. Mutational analysis and biochemical and biophysical studies will as well be used to reveal how SAS-5 and SAS-6 interact via short motifs. The ultimate goal is to provide a structural insight into how SAS-6 and SAS-5 cooperate in centriole assembly and what their roles are in defining the nine-fold symmetry.

5. Results

5.1. The C-terminal domain of SAS-5 is necessary and sufficient for interaction with SAS-6

Previous studies using yeast two-hybrid assays showed that the interaction between SAS-6 and SAS-5 was undetectable when using SAS-5 corresponding to the *sas-5*(*t2079*) mutant allele (Leidel et al. 2005). Since this natural mutation (R397→C) is located close to the C-terminus of SAS-5 and as the last 15 residues of SAS-5 are predicted to form an alpha helix, we wondered whether this C-terminal helix alone is sufficient for binding SAS-6.
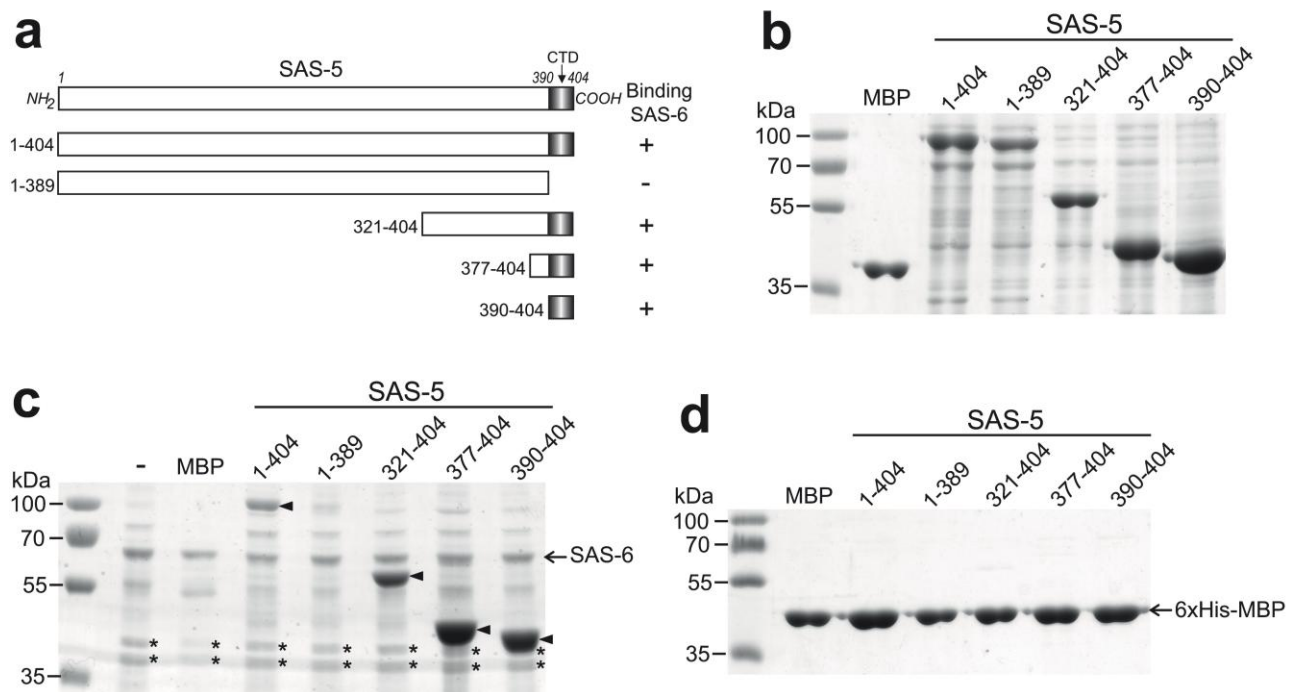
Figure 15 The C-terminal domain of SAS-5 interacts with SAS-6. **(a)** Deletion constructs of SAS-5 used for *in vitro* binding assays with SAS-6. CTD, carboxy-terminal domain. Numbers indicate amino acid positions and ranges. The right column shows the summary of the binding results in (c). **(b)** Purified maltose binding protein (MBP) or soluble fractions of SAS-5 proteins for *in vitro* pull-down assays. **(c)** *In vitro* pull-down results of SAS-5 proteins using Ni-NTA bound SAS-6 as the bait. MBP is used as a negative control for tag-dependent binding. SAS-5 proteins specifically pulled down by SAS-6 are indicated by arrow heads. Marked by asterisks are the two degradation products of SAS-6. The upper and lower bands are confirmed by N-terminal amino acid analyses to be sequences starting at residues 225 and 239, respectively. **(d)** Mock experiments to check the binding of SAS-5 proteins to Ni-NTA resin loaded with 6×His-tagged MBP. No nonspecific interaction to the resin was detected.

To test this, we carried out *in vitro* pull-down experiments using five truncations of SAS-5 in addition to the full-length protein (**Figure 15a**). In order to increase yield and to better visualize the smaller fragments, we added maltose-binding protein (MBP, M.W. ~42kDa) as a fusion tag to the N-termini of all six constructs (**Figure 15b**). Our *in vitro* pull-down results showed that the SAS-5 CTD (residues 390-404) is both necessary and sufficient to bind SAS-6 (**Figure 15c**). In a control experiment, we used MBP-loaded Ni-NTA beads to pull down SAS-5, which shows no bound SAS-5 (**Figure 15d**), indicating that the determined interaction between SAS-5 and SAS-6 is specific. Notably, during purification of SAS-6 we consistently observed two degraded fragments on SDS-PAGE gels (**Figure 15c**, asterisks). Using N-terminal amino acid analysis, we found that the two fragments correspond to sequences starting at residues 225 and 239, respectively. It indicates that this neck region of SAS-6 (in reference to the head group and the coiled coil tail), approximately spanning residues 220-240, is flexible and prone to proteolysis.

5.2. The C-terminal domain of SAS-5 binds specifically to the central part of the SAS-6 coiled coil

It was previously reported that SAS-5 binds specifically to the coiled coil domain of SAS-6 (residues 180-415) (Boxem et al. 2008; Leidel et al. 2005). The results in **Figure 15** show that only the last 15 residues of the SAS-5 CTD are required for SAS-6 interaction. As this CTD is much smaller in size compared to the SAS-6 coiled coil, we anticipated that only a small segment of the SAS-6 coiled coil would be involved in their interaction. To locate the SAS-5 binding site, we generated six truncations of SAS-6 (**Figure 16a**), which were all expressed and soluble (**Figure 16b**). *In vitro* binding assays were carried out. MBP-tagged SAS-5 CTD preloaded on the amylose beads was used as the bait to pull down SAS-6 proteins. The results visualized on the SDS-PAGE showed that the SAS-5 CTD specifically bound to the central region of the SAS-6 coiled coil, spanning residues 248-303 (**Figure 16c**).

To exclude the possibility that MBP tag caused the unspecific binding during the pull-down assay, a control assay was carried out. MBP alone was used as the bait to pull down SAS-6 and no significant binding was detected (**Figure 16d**). Taken together, the pull-down assays showed that the interaction between the SAS-5 CTD and the SAS-6 coiled coil is specific.

Figure 16 The SAS-5 CTD binds specifically to the central part of the SAS-6 coiled coil. **(a)** Truncation constructs of SAS-6 used for *in vitro* binding assays with the SAS-5 CTD (residues 390-404). Numbers in front of the constructs indicate amino acid ranges. The right column shows the summary of the binding results in (c). **(b)** Soluble fractions of SAS-6 proteins used in the *in vitro* pull-down assays. Arrowheads indicate the target proteins. **(c)** *In vitro* pull-down results of SAS-6 proteins using amylose beads preloaded with the MBP-tagged SAS-5 CTD as the bait. Filled arrowheads indicate SAS-6 proteins pulled down by SAS-5. An empty arrowhead indicates the MBP-dependent nonspecific binding of the construct containing residues 1-218 of SAS-6, which is comparable to what is seen in the control experiment in (d). Marked by asterisks are the degradation products of SAS-6. **(d)** Control experiments to check the nonspecific binding of SAS-6 proteins to the MBP tag. Except for SAS-6 (residues 1-218), which is indicated by an empty arrowhead, no other significant binding was detected.

In a reciprocal binding experiment, we used Ni-NTA bound SAS-6 constructs to pull down the MBP-tagged SAS-5 CTD, which further confirms that SAS-5 binds to the same region of the SAS-6 coiled coil (**Figure 17**).
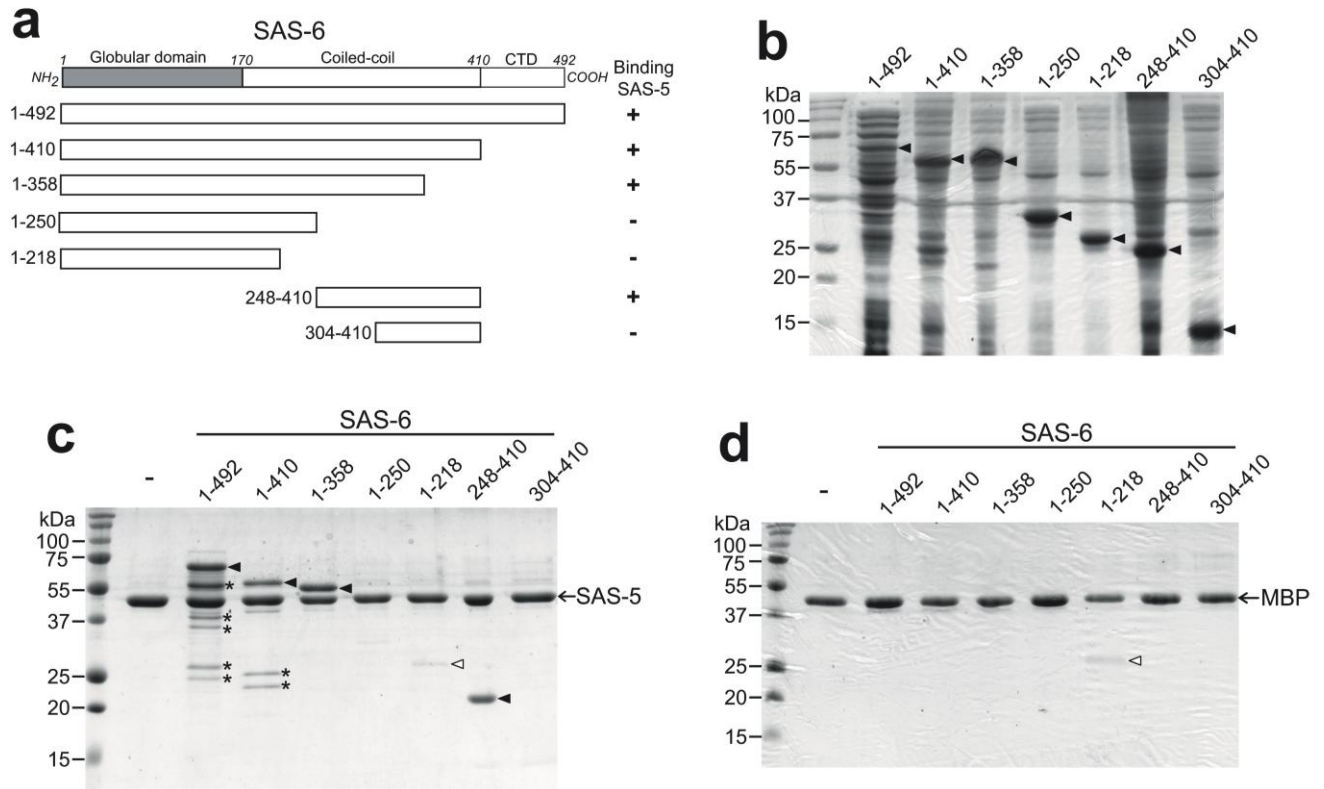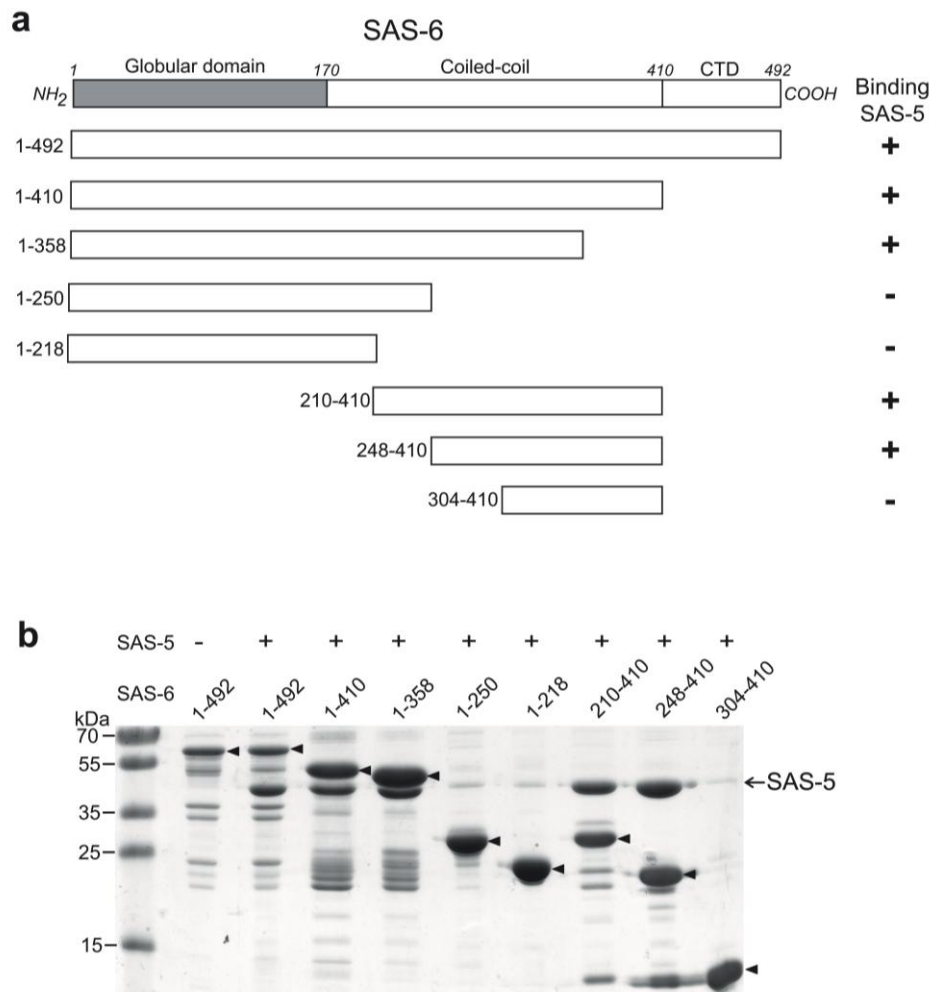
Figure 17 The SAS-5 CTD binds specifically to the central part of the SAS-6 coiled coil. **(a)** Truncation constructs of SAS-6 used for in vitro binding assays with the SAS-5 CTD (residues 390-404). Numbers in front of the constructs indicate amino acid ranges. The right column shows the summary of the binding results in (b). **(b)** In vitro pull-down results of MBP-tagged SAS-5 CTD using Ni-NTA bound SAS-6 of various lengths as the baits. Arrowheads indicate SAS-6 proteins.

5.3. Crystal structure of the SAS-6 coiled coil domain reveals an electrostatic periodicity along the coiled coil

To investigate the interaction between SAS-5 and SAS-6 at the molecular level, we determined the crystal structure of the SAS-6 CCD (residues 248-410). This contains the SAS-5 binding site mapped above. The structure was determined to 3.3 Å resolution (**Table 2** and **Figure 18a**).

Table 2 Data collection and refinement statistics

Data collection	
Space group	P6$_1$
Cell dimensions	
a, *b*, *c* (Å)	140.29, 140.29,
Resolution (Å)	39-3.3 (3.48-3.30)
No. unique reflections	12775 (1851)
I / σ*I*	10.4 (1.9)
Overall completeness (%)	99.9 (100.0)
Overall redundancy	7.2 (7.4)
Anomalous completeness	99.5 (99.7)
Anomalous redundancy	3.7 (3.7)
Resolution (Å)	39-3.3
R$_{work}$ / *R*$_{free}$ (%)	25.8/29.9
No. atoms	
Protein	2578
R.m.s. deviations	
Bond lengths (Å)	0.009
Bond angles (°)	1.30

*Values in parentheses are for highest-resolution shell.

In the structure, each of the two chains is folded into a continuous alpha helix spanning residues 250-405. The two helices form a parallel coiled coil extending to 230 Å in length. Interestingly, an electrostatic surface plot indicated that the SAS-6 CCD exhibits a periodically charged pattern along the helices - the segments spanning residues 250-293 and 343-390 are predominantly negatively charged, and residues 294-342 and 391-407 are mainly positively charged (**Figure 18b**).
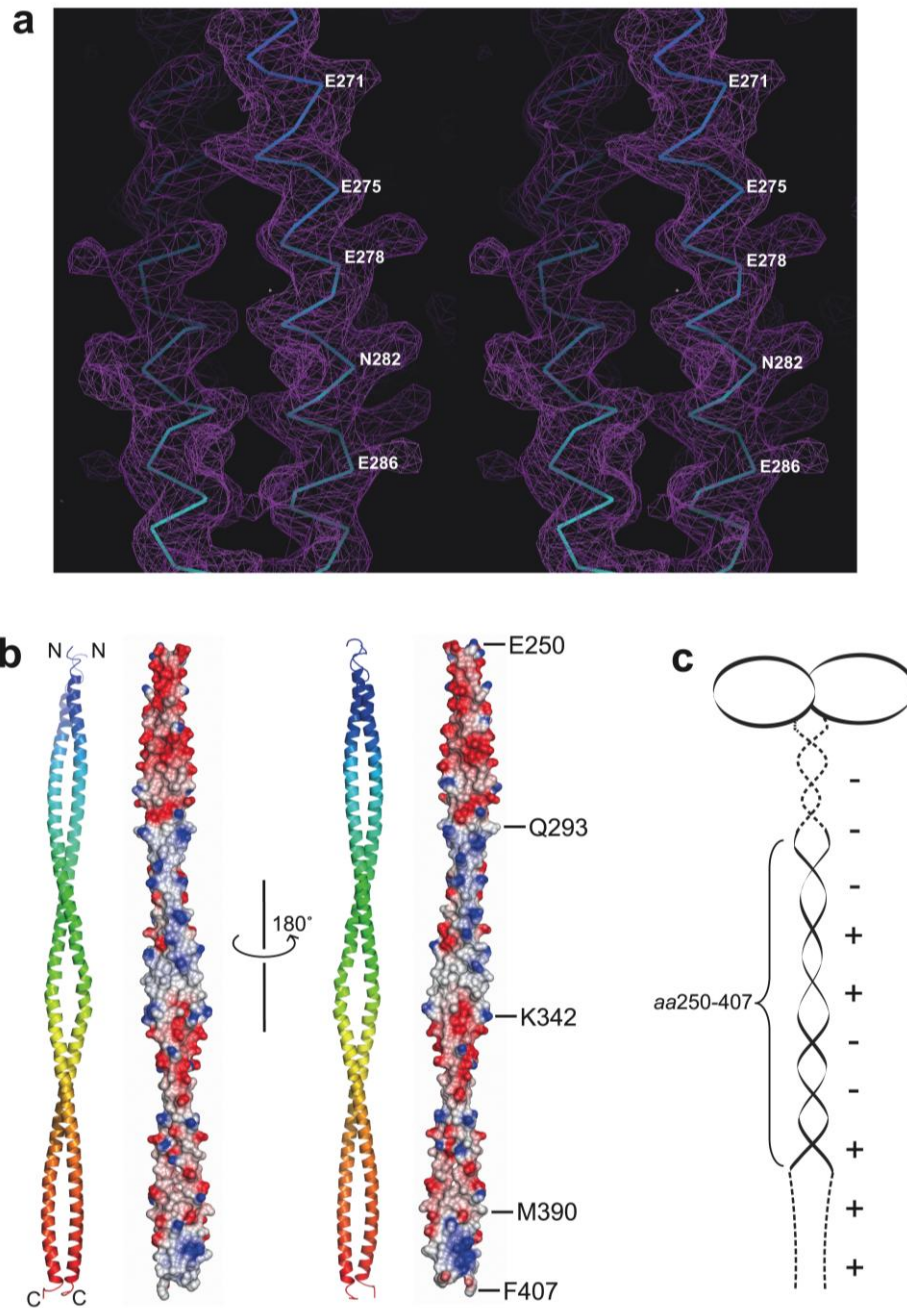
Figure 18 Crystal structure of the SAS-6 coiled coil domain. **(a)** Stereo view of a representative portion of the $2F_o - F_c$ experimental electron density map (contoured at 2.0 σ) covering residues 268-290. For clarity, only the main chains of the final model are shown. **(b)** Ribbon diagram and electrostatic surface plot of the SAS-6 CCD structure. Residues at the boundaries of differently charged segments are indicated. **(c)** Schematic representation of the SAS-6 dimer. Dashed lines indicate the regions lacking a known structure. Positive and negative charges along the coiled coil and in the C-terminal domain are depicted as "+" and "-", respectively.

Sequence alignment of SAS-6 proteins from three different *Caenorhabditis* species indicated that the surface of the coiled coil region preceding the CCD, spanning residues 220-250, is also

negatively charged (**Figure 19**), whereas the C-terminal part of SAS-6 (residues 410-492) contains five conserved lysine/arginine residues and is positively charged (**Figure 20**).
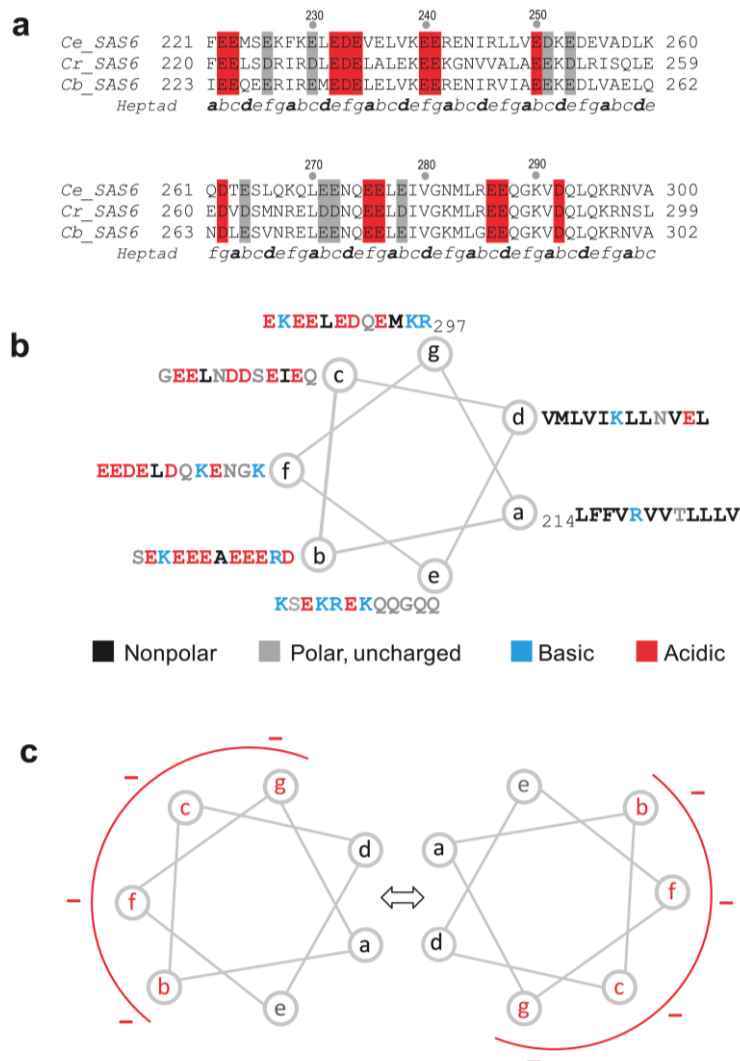
Figure 19 The "neck" region of the *C. elegans* SAS-6 coiled coil(residues 221-300) is predominantly negatively charged. **(a)** Sequence alignment of the N-terminal part of the SAS-6 coiled coil from three *Caenorhabditis* species. Ce, *Caenorhabditis elegans*; Cr, *C. remanei*; Cb, *C. briggsae*. Identical negatively-charged residues (D or E) are highlighted in red; conserved ones are highlighted in grey. The heptad registers of the residues are shown beneath the aligned sequences. Registers of residues 250-300 are extracted from the crystal structure and those for the preceding residues are derived from coiled coil predictions (http://www.ch.embnet.org/software/COILS_form.html). **(b)** Helical wheel diagram looking down the helix axis from the N- to the C-terminus of the SAS-6 coiled coil (residues 221-300). Heptad positions are labeled from a to g. Positions a and d are occupied predominantly by nonpolar hydrophobic residues whereas other positions are mostly polar or charged residues. **(c)** Helical wheel representation of the parallel homodimer of the SAS-6 coiled coil (residues 221-300). Inter-helical hydrophobic interactions are denoted as a wide arrow. Negatively charged surfaces formed by residues at positions b, c, f, and g are shown by red curves.

Previously published crystal structures of SAS-6 proteins have shown that the head group to-gether with the N-terminal part of the coiled coil domain of SAS-6 forms a dimer (Kitagawa et al. 2011b; van Breugel et al. 2011). Taken together, we conclude that *C. elegans* SAS-6 folds into a tadpole-like structure with an alternating charge distribution along its coiled coil tail (**Figure 18**).
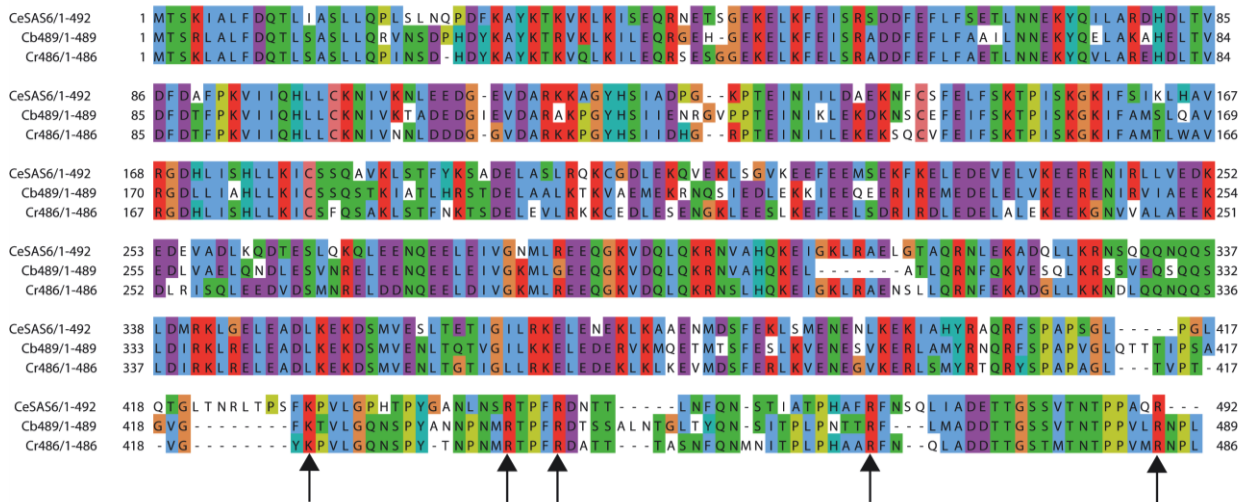
Figure 20 Sequence alignment of SAS-6 proteins from three *Caenorhabditis* species. Ce, *Caenorhabditis elegans*; Cr, *C. remanei*; Cb, *C. briggsae*. The five conserved positive residues in the C-terminal domain are indicated by arrows.

5.4. Crystallization of the SAS-6-CCD/SAS-5-CTD complex

To understand the mechanism of SAS-6 and SAS-5 interaction, different length of truncations of the SAS-6 CCD and the MBP fusion SAS-5 CTD (residues 390-404) have been expressed in *E. coli* and purified by Ni-NTA and amylose affinity chromatography, respectively (**Figure 21a**). After mixing the SAS-6 CCD with excess amount of the SAS-5 CTD, the complex was purified on a size exclusion column (**Figure 21b**). Crystallization trials have been carried out for purified complex-es of SAS-6 (residues 192-410) and the SAS-5 CTD, SAS-6 (residues 210-410) and the SAS-5 CTD, SAS-6 (residues 248-410) and the SAS-5 CTD. Thin needle-shaped crystals grew from multiple conditions at both 22°C and 4°C. Crystals finally reached the maximum size of 0.5 × 0.02 × 0.01 mm after optimization (**Figure 21c**). The best crystals of SAS-6 (residues 248-410) and the SAS-5 CTD were gained by the hanging drop vapor diffusion method with the reservoir of 8% PEG4000, 0.1 M MES pH6.5, 0.2 M $MgCl_2$. Unfortunately, these crystals only diffracted to ~7 Å. The other two types of crystals of SAS-6 (192-410) and SAS-5 CTD, and SAS-6 (residues 210-410) and SAS-5 CTD diffracted more poorly, to ~10 Å.
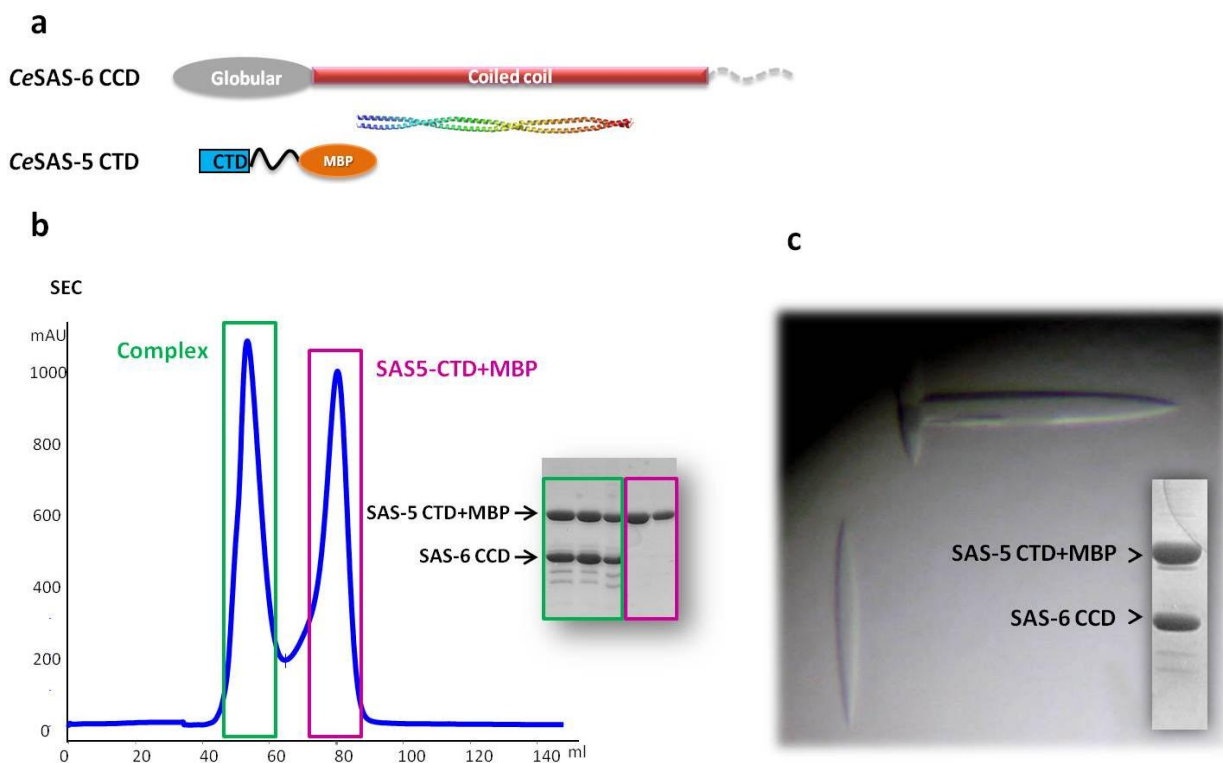
Figure 21 Purification and crystallization of the SAS-6-CCD/SAS-5-CTD complex. SAS-6 CCD, SAS-6 coiled coil domain (residues 248-410); SAS-5 CTD, SAS-5 C-terminal domain (residues 390-404). **(a)** Constructs of the SAS-6 CCD used for co-purification with the SAS-5 CTD. **(b)** Size exclusion column (SEC) chromatography of mixture of the SAS-6 CCD and an excess amount of the SAS-5 CTD. Mixed proteins were applied to Superdex 200 16/60. SDS-PAGE confirmed that the first peak was the complex of both proteins. **(c)** Needled-shaped crystals were obtained at 8% PEG4000, 0.1 M MES pH6.5, 0.2 M MgCl$_2$, in a hanging droplet after optimization. About ten crystals were picked, thoroughly washed in reservoir solution, and applied on an SDS-PAGE to confirm that the crystals are the protein complex instead of a single protein.

5.5. Association of SAS-5 and SAS-6 is based on synergistic hydrophobic and electrostatic interactions

Due to the limitation of the low resolution diffraction data of the crystal of the SAS-6 CCD and the SAS-5 CTD, biochemical analyses instead were carried out to further locate the specific binding site of SAS-5 on SAS-6. We generated multiple structure-based deletion constructs of SAS-6 in the coiled coil region covering residues 248-303 that is essential for their interaction (**Figure 17**). To avoid disrupting or distorting the coiled coil structure, each deletion removed n×7 residues (n = 2, 3, or 4) to maintain the register of the heptad repeats of the coiled coil (**Figure 22a**).

58
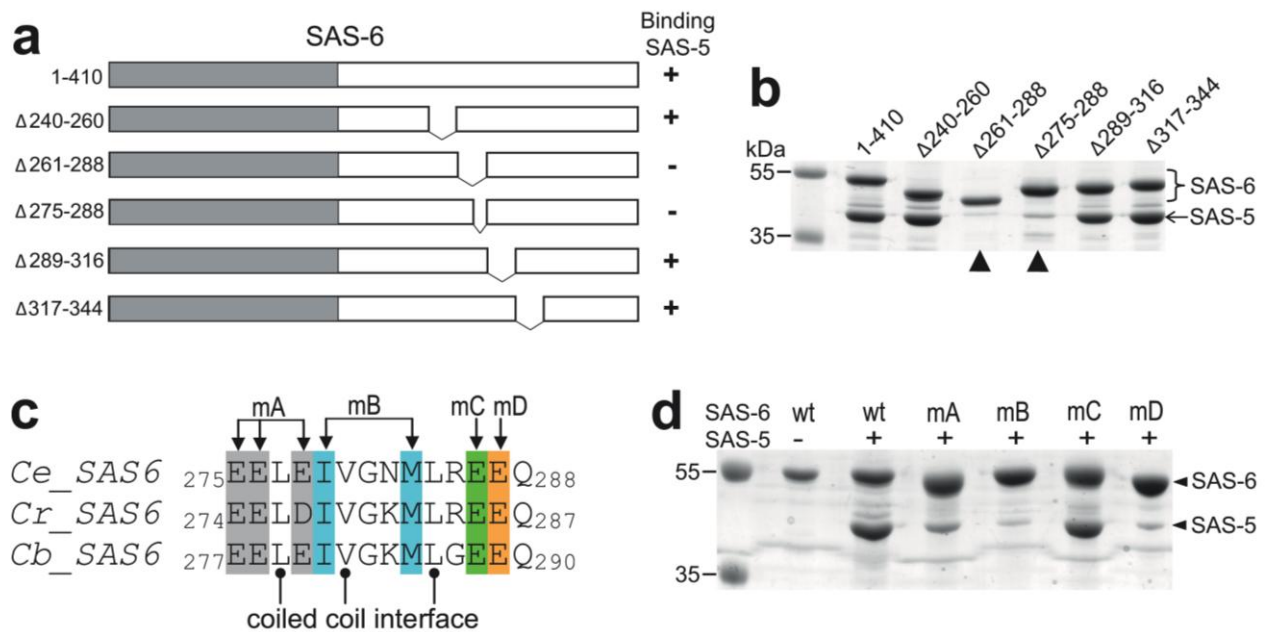
Figure 22 SAS-5 binds to a short region in the SAS-6 CCD, residues 275-288. **(a)** Schematic of SAS-6 deletion constructs. The right column summarizes the interaction results in (b). **(b)** *In vitro* pull-down results of SAS-5 using Ni-NTA bound SAS-6 as the bait. The two deletions of SAS-6 that failed to pull down SAS-5 are indicated by arrowheads. **(c)** Sequence alignment of the SAS-5 binding site from three *Caenorhabditis species. Ce, Caenorhabditis elegans; Cr, C. remanei; Cb, C. briggsae.* The three hydrophobic residues on the coiled coil interface are indicated beneath the sequences. Mutations of the four groups of conserved, solvent-exposed residues (to alanines) are highlighted in different colors. **(d)** Coomassie stained SDS-PAGE gel showing the result of *in vitro* pull-down of the SAS-5 CTD by wild-type (wt) and the four mutations of SAS-6. All mutations except for mC failed to interact with SAS-5.

We also generated a deletion outside of this region, spanning residues 317 - 344, as a control to show that a partial deletion of the SAS-6 coiled coilded did not affect its folding or the binding ability of the neighboring region to SAS-5. Using *in vitro* pull-down assays, we determined that the region containing residues 275-288 of SAS-6 is essential for SAS-5 binding (**Figure 22a and b**). To determine which individual residues are directly involved in the interaction, we generated four structure-based mutations of all residues in this 14-residue segment that are conserved among *Caenorhabditis* species but are not a part of the coiled coil interface (**Figure 22c**, mA, mB, mC & mD). Results of the *in vitro* pull-down experiments indicated that mutations of either the central hydrophobic residues (mB: I279A+M283A) or the flanking negatively charged residues (mA: E275A+E276A+E278A, mD: E287A) nearly completely abolished the SAS-5/SAS-6 interaction. In contrast, mutation of another charged residue in the same region (mC: E286A)

seemed not to affect the interaction (**Figure 22d**). The influence of the mutations on the SAS-5/SAS-6 interaction was further measured by isothermal titration calorimetry (*ITC*) assays. These indicated that each of the three mutations (mA, mB and mD) completely abolished the interaction between SAS-5 and SAS-6, whereas the mutation mC only slightly reduced the binding affinity (**Figure 23**). Given that all these mutated residues are solvent exposed, as seen in the crystal structure, this suggests that the interaction between SAS-5 and SAS-6 is structure-based rather than a nonspecific electrostatic interaction.

To test whether SAS-6 mutations in vivo also disrupt the interaction with SAS-5, in turn defect centriole assembly, we generated constructs carrying each of the three sets of mutations which disrupted interaction with SAS-5 *in vitro* (mA, mB and mD) and used these for *C. elegans* transformation. We collaborated with Dr. Alex Dammermann, who obtained several independent strains of mA and mD with identical behavior. However the strain carrying an integration of mB was not able to be obtained. This could be explained by technical limitations or reflect differences in the ability of *C. elegans* to tolerate the two classes of mutations (mA, mD charge vs. mB hydrophobicity). Interestingly, SAS-6 mutants mA and mD both localized to centrioles, however, after deleting the endogenous protein by RNAi, the embryos showed the characteristic monopolar second division phenotype, in turn neither mutant could successfully assemble centriole. The results from mating experiment (Dammermann et al. 2008; Dammermann et al. 2004; Kirkham et al. 2003) to specifically assess SAS-6 recruitment showed that SAS-6 recruitment was nearly completely abolished for both mutants when endogenous SAS6 was depleted, similar to what was seen with wild-type SAS-6 following depletion of SAS-5. Consistent with *in vitro* data, the in vivo results suggest that the residues mutated in mA and mD are indeed critical for SAS-6 recruitment and function in centriole assembly, and the severity of the phenotype mirrors that of depleting SAS-5 itself.
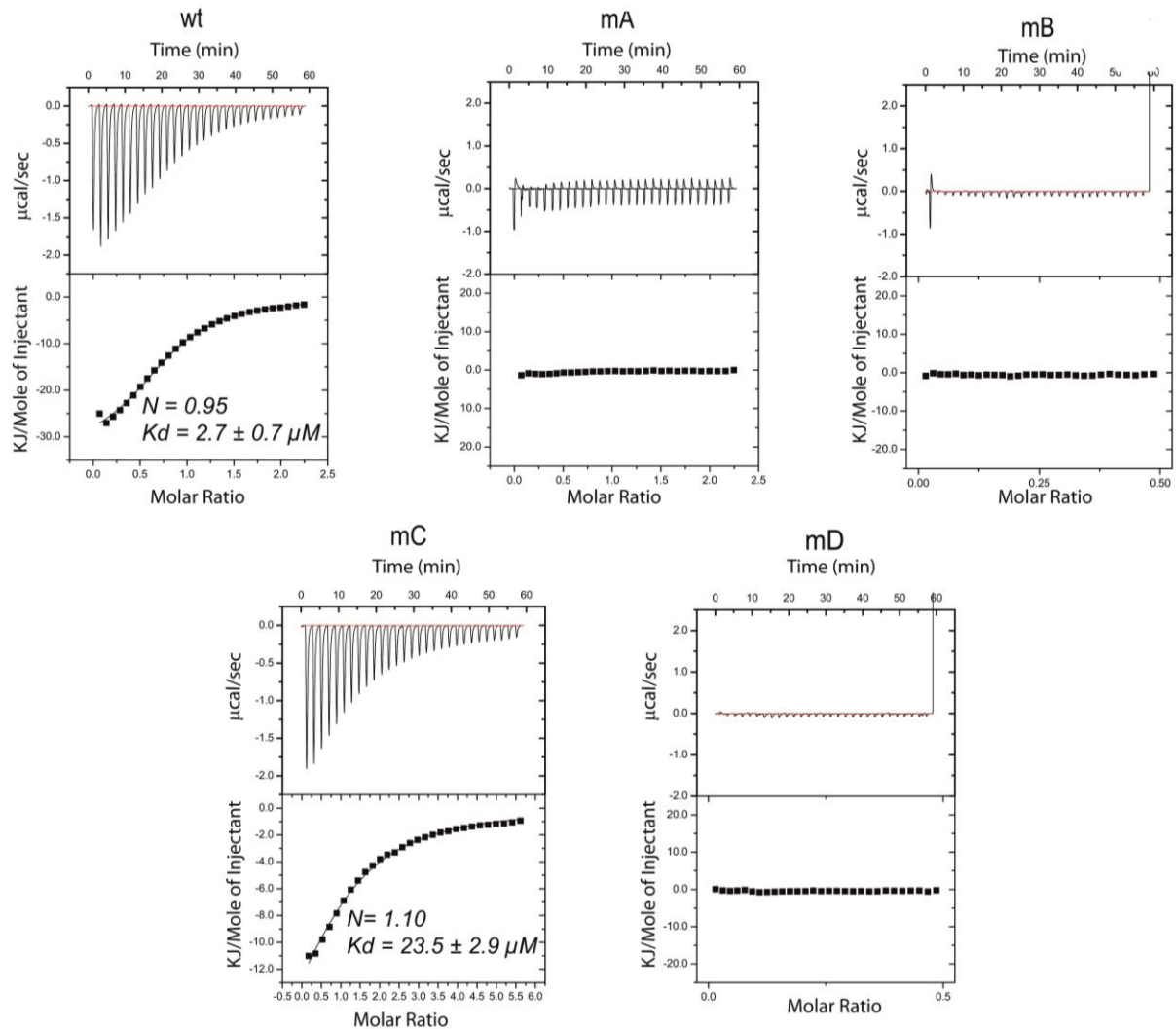
Figure 23 ITC titration and fit curves for wild-type (wt) and the four mutations of SAS-6 interacting with the SAS-5 CTD. Mutations mA, mB, and mD completely abolished the interaction between SAS-6 and SAS-5, whereas the mutation mC only mildly reduced the binding affinity. For wild-type (wt) and mC, "N" represents the molar ratio between the monomeric SAS-5 CTD and the dimeric SAS-6 CCD.

We showed above that the SAS-5 CTD (residues 390-404) is responsible for interacting with SAS-6 (**Figure 15**). To further identify which residues in this region are directly involved in the interaction with SAS-6, we generated eleven mutations in the SAS-5 CTD that substituted each non-alanine residue with an alanine, except for the residue R397, which was replaced by a cysteine as in the previously reported *sas-5*(*t2079*) mutant (Leidel et al. 2005) (**Figure 24a**). All SAS-5 CTD mutations were fused to the C-terminus of MBP to facilitate the visualization of the proteins on SDS-PAGE gels. We then carried out *in vitro* binding assays using Ni-NTA bound 6×His-SAS-6 (residues 1-410) to pull down SAS-5. While most of the mutations did not affect the

amount of SAS-5 pulled down, four of them (M4/I396A, M5/R397C, M8/Y400A, and M10/R403A) drastically reduced the interaction between SAS-6 and SAS-5 (**Figure 24b**). To quantify the influence of the mutations on the interaction, we again carried out ITC experiments using individually purified proteins. None of the four mutations that could not be pulled down by SAS-6 showed a measurable K_d (**Figure 24c**). In contrast the other seven mutations of SAS-5 showed no or only slight reduction of binding affinity toward SAS-6 (data not shown).
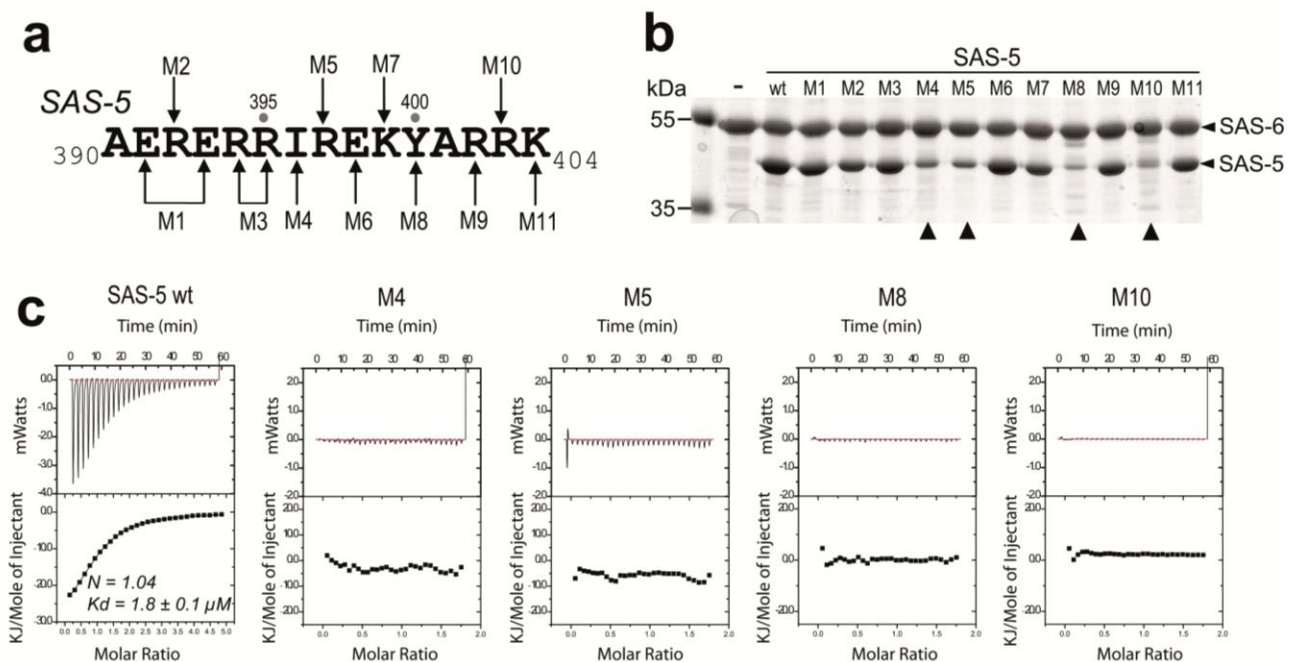
Figure 24 Four of the 15 residues in the SAS-5 CTD contribute to its specific interaction with SAS-6. **(a)** Sequence of the SAS-5 CTD. Mutations (to Alanines) are indicated as M1 - M11. **(b)** Coomassie stained SDS-PAGE gel showing the results of *in vitro* pull-down of wild-type or mutants of the SAS-5 CTD by SAS-6. The four mutations that show a drastic decrease of binding to SAS-5 are indicated by arrowheads. **(c)** ITC titration and fit curves for wild-type and four mutations of the SAS-5 CTD interacting with the SAS-6 CCD. No measurable K_d was detected for all these mutations.

Based on the pull-down and ITC results, a docking trial was carried out by ClusPro 2.0 (Kozakov et al. 2010) using the crystal structure of the SAS-6 CCD as the receptor and a theoretical helical model of the SAS-5 CTD as the ligand. Multiple predicted interaction models were generated (**Figure 25**). In a representative docked model (**Figure 26**), the helix of the SAS-5 CTD was placed nearly perpendicular onto the SAS-6 coiled coil. This arrangement allowed both the hydrophobic interactions between the central two pairs of non-polar residues (SAS-6: I279/M283 vs. SAS-

6: I396/Y400) and the electrostatic interactions between the flanking oppositely charged residues (SAS-6: E275/E276/E278 vs. SAS-5: R403).
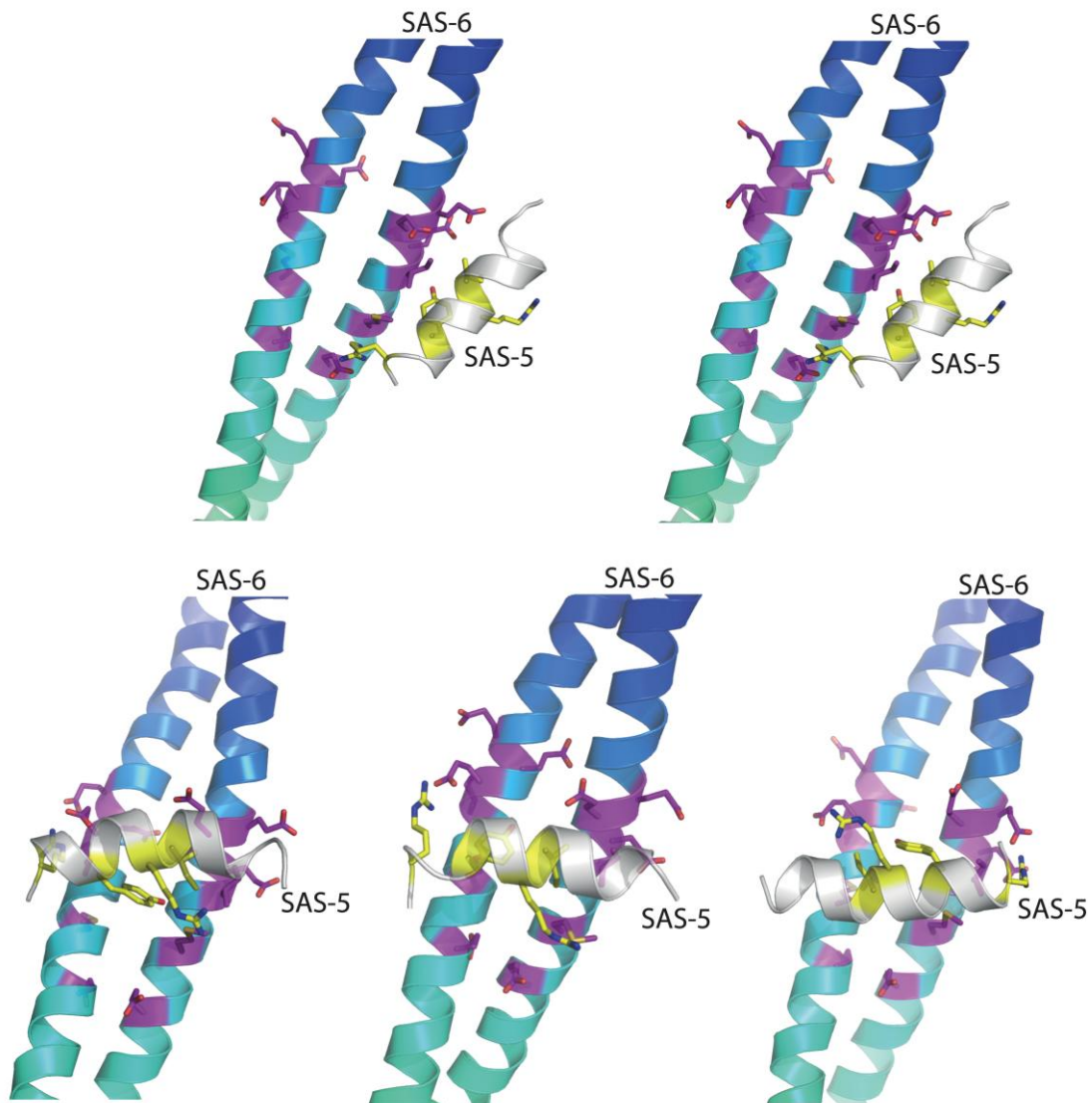
Figure 25 Five alternative interacting models of SAS-5 and SAS-6 predicted by the ClusPro 2.0 docking server (http://cluspro.bu.edu/). Side chains of the residues involved in the interaction, which were mapped by our mutagenesis studies, are shown in sticks with color schemes of magenta in SAS-6 and yellow in SAS-5.

However, the distance between E287 of SAS-6 and R397 of SAS-5 seems too far for establishing an electrostatic interaction. One possibility is that the side chains of these two charged residues form salt bridges with the backbone of the opposite molecule. An alternative is that the SAS-6 coiled coil might bend at the interaction site to maximize the intermolecular contacts, which has been seen in other interactions between a helix and a coiled coil (Sibanda et al. 2001). Given the nearly symmetric arrangement of the SAS-6 coiled coil and the stoichiometry of the

complex (SAS-6 dimer : SAS-5 = 1:1. See the ITC results for "wt" in **Figure 23** and **Figure 24c**), it suggests that SAS-5 only binds to one side of the SAS-6 coiled coil, which implies that binding of SAS-5 to one site either occludes the other site or disrupts the structural symmetry by inducing local conformational changes of the SAS-6 coiled coil.

In summary, we have identified the residues on both SAS-5 and SAS-6 that are directly involved in the interaction between these two centriolar proteins. These conclusions, together with the crystal structure of the SAS-6 CCD and the predicted helical structure of the SAS-5 CTD, allowed us to generate a docking model for the interaction between the two proteins (**Figure 26**).
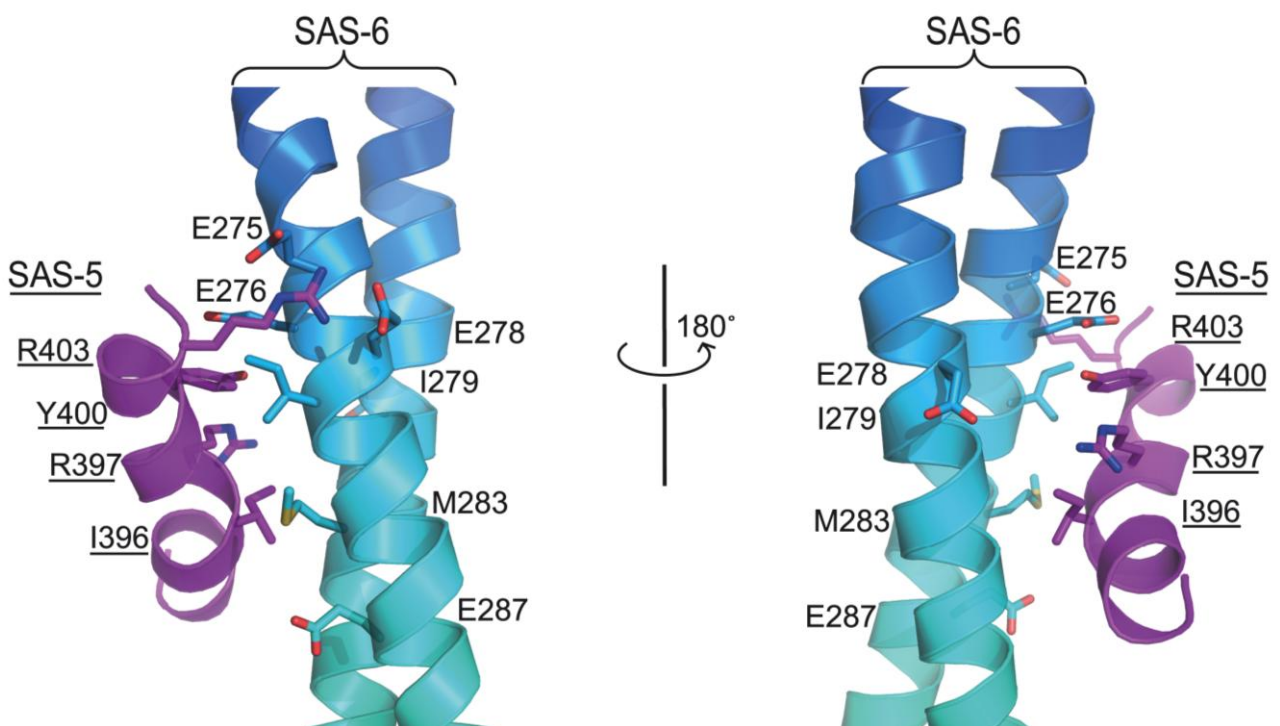
Figure 26 Docking the SAS-5 CTD to its binding site on the SAS-6 CCD. Side chains of the residues that participate in the interaction are shown and labeled.

5.6. SAS-6 molecules form an anti-parallel tetramer through the electrostatic interactions of their coiled coil domains

Rotary metal shadowing electron microscopy studies of purified recombinant SAS-6 protein showed that although many of the observed particles are tadpole-like structures as reported previously (Kitagawa et al. 2011b), a significant fraction of them (~20%) are dumbbell-like structures with a central rod measuring 35-45 nm in length (**Figure 27a**). As shown in **Figure 18c**,

there is an alternating charge distribution along the SAS-6 CCD, which we suspected may direct further self-association of SAS-6 dimers along their coiled coils to form a tetramer. To test this hypothesis, we carried out a dilution ITC experiment that has been used successfully for analyzing the dissociation equilibrium of other proteins (Lovatt et al. 1996). In this assay, a series of small aliquots of concentrated SAS-6 CCD were injected into a large volume of buffer, which generated a sequence of endothermic heat pulses, characteristic of molecular dissociation (**Figure 27b, top**). It has been reported that the K_d of the SAS-6 coiled coil dimer is ~0.9 μM (Kitagawa et al. 2011b). The fit dissociation curve had a K_{diss} of 56.2 ± 7.6 μM and ΔH_{diss} of 2.84 ± 0.05 kcal/mole, which suggests a dimer-tetramer equilibrium (**Figure 27b, bottom**). The tetrameric association of the SAS-6 CCD is apparently much weaker than the coiled coil dimer and thus may not always survive the grid preparation for rotary metal shadowing as the coiled coil dimer does. This explains why the dumbbell-like tetramer structure was not as frequently observed as the tadpole-like structure of the SAS-6 dimer.

To better understand the self-association of the SAS-6 CCD, we subjected our solved crystal structure of the SAS-6 CCD to the ClusPro 2.0 protein-protein docking server (Kozakov et al. 2010). The docking results suggested an anti-parallel interaction of the coiled coils, with the opposite charges complementing each other in each segment (**Figure 27c**). The fully extended central rod of the anti-parallel SAS-6 tetramer based on this docking model is calculated to be ~45 nm long, which is in agreement with the length of the rod in the dumbbell-like structure seen in the electron micrographs (**Figure 27a**). The length variation of the central rods in the dumbbell-like structures is likely due to the flexible region at the N-terminal part of the coiled coil as shown above (**Figure 15c**, asterisks). These data altogether suggest that *C. elegans* SAS-6 forms a dumbbell-like tetramer through the anti-parallel association of the coiled coils. Notably, the SAS-5 binding sites on the SAS-6 CCD are obscured in the anti-parallel SAS-6 tetramer (**Figure 27c**, boxes).
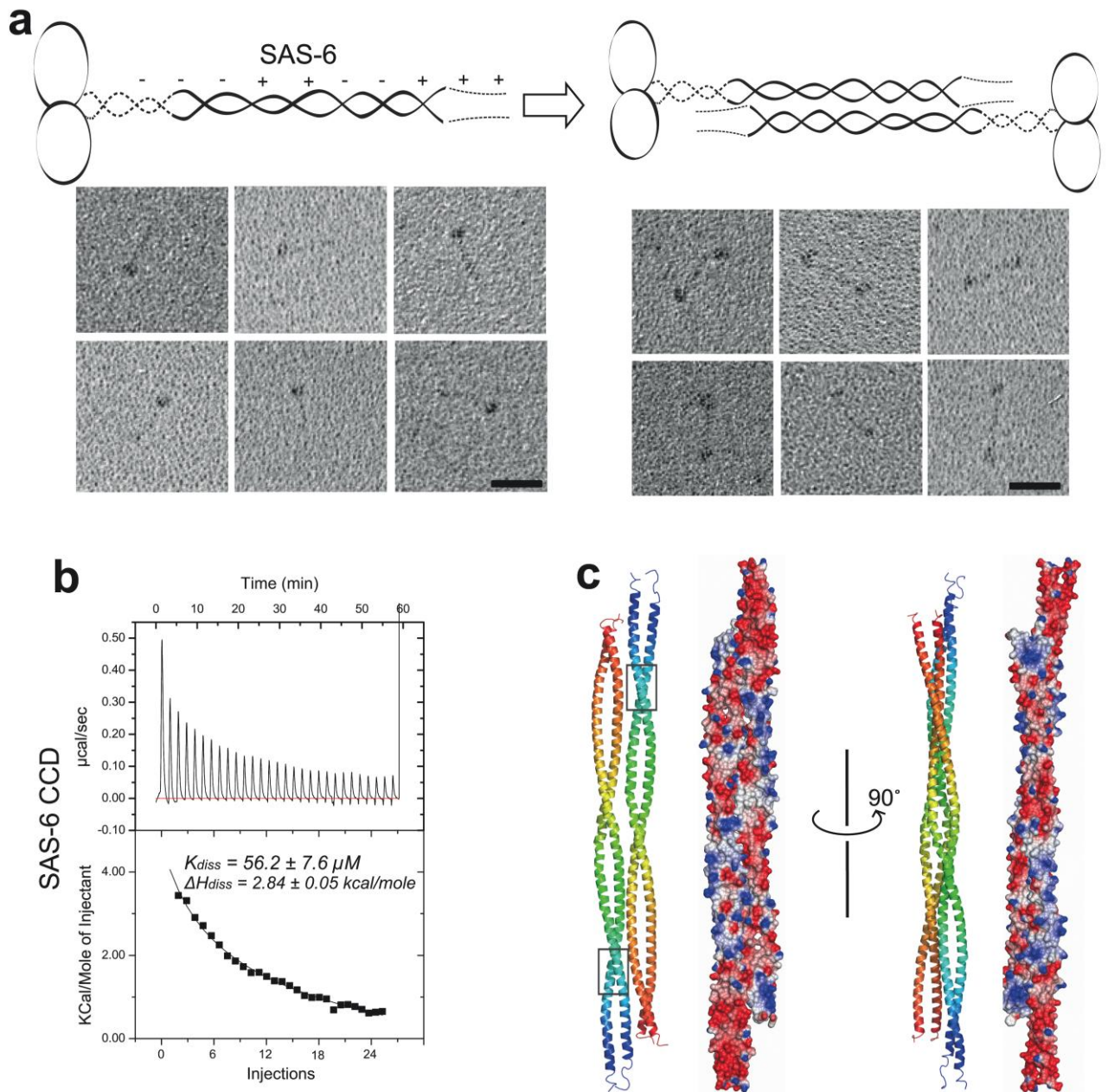
Figure 27 *C. elegans* SAS-6 forms an anti-parallel tetramer via the electrostatic interaction of the coiled coils. **(a)** Schematic model and the rotary metal shadowing electron micrographs of recombinant SAS-6. Scale bars, 30 nm. **(b)** Experimental and integrated dilution ITC curves for the SAS-6 CCD (residues 248-410). **(c)** Docking of the SAS-6 CCD self-association by the automated protein docking program ClusPro 2.0 (Kozakov et al. 2010). Both ribbon diagrams and electrostatic surface plots are shown. Ribbons are rainbow colored from blue (N-terminus) to red (C-terminus). In the surface plots, positive and negative charges are shown in blue and red, respectively. Boxed are the SAS-5 binding sites on the SAS-6 CCD.

5.7. Binding of SAS-5 both disrupts the tetrameric association of the SAS-6 CCD and promotes the formation of a ring-like structure resembling the central tube of *C. elegans* centrioles

We found that SAS-5 binds specifically to a segment of the SAS-6 CCD that is part of the periodic charge region being obscured in the anti-parallel tetramer (**Figure 27c**, boxes). Moreover, the interaction between SAS-5 and SAS-6 (K_d ~ 2 µM) is more than one order of magnitude stronger than the self-association of the SAS-6 CCD tetramer (K_d ~ 56 µM). Therefore, binding of SAS-5 may disrupt SAS-6 self-association. To test this, we carried out a dilution ITC assay similar to that for the SAS-6 CCD alone but with 1.5 × fold (molar ratio) of the SAS-5 CTD supplemented to both the injections and the buffer. There would be two possibilities for the results. One possible result is that the presence of the SAS-5 CTD in the buffer prevents the dissociation of the SAS-5/SAS-6 complex and no self-dissociation of SAS-6 is observed. The other possibility is that the anti-parallel self-association of SAS-6 is not affected by SAS-5 binding and we would still observe the dissociation of the tetramer of the SAS-6 CCD to dimers. As shown in **Figure 28a**, no endothermic heat pulses were observed, which was in contrast to the strong dissociation signal for the SAS-6 CCD (**Figure 27b**). This suggests that the binding of SAS-5 might disrupt the tetrametric association of the SAS-6 CCD. This interpretation was further supported by an additional experiment to exclude the possibility that the binding of SAS-5 stabilizes the SAS-6 CCD tetramer. As shown in **Figure 28b**, analysis of the complex by static light scattering (SLS) indicated that the complex is a hetero-trimer (MW ~80kDa; MWs of the SAS-6 CCD monomer and the MBP-SAS-5 CTD are 19kDa and 44kDa, respectively). Therefore, disruption of the tetrametric association of the SAS-6 CCD by binding of SAS-5 accounts for the loss of endothermic heat pulses in the ITC.

Moreover, examination using dynamic light scattering (DLS) also indicates that the SAS-5 CTD shifts the equilibrium between two species of the SAS-6 CCD into one species when mixing the two proteins in a stoichiometric 1:1 ratio (**Figure 29**). Therefore, binding of SAS-5 disrupts the tetrameric association of the SAS-6 CCD.
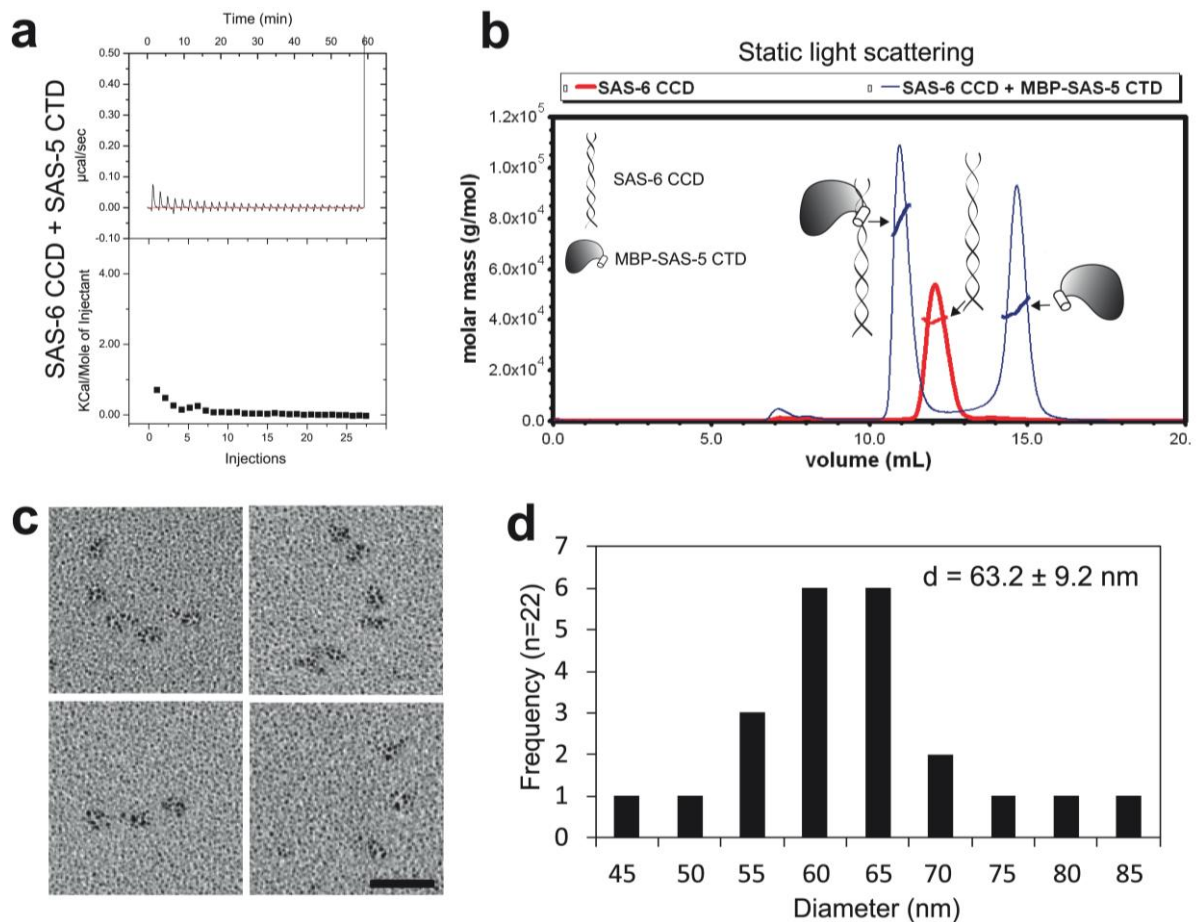
Figure 28 The SAS-5/SAS-6 complex forms curved structures with an extrapolated diameter similar to the central tube of *C. elegans* centrioles. **(a)** Experimental and integrated dilution ITC curves for SAS-6 CCD + SAS-5 CTD. **(b)** SLS analysis of the complex of the SAS-6 CCD and the MBP-SAS-5 CTD. The SAS-6 CCD by itself forms a dimer (MW ~38kDa), whereas mixing it with the MBP-SAS-5 CTD (molar ratio = 1:1.5) gave rise to a hetero-trimer (MW ~80kDa). MW of the MBP-SAS-5 CTD is 44kDa. **(c)** Rotary metal shadowing electron micrographs of the full-length SAS-5-/SAS-6 complex. Scale bar, 30 nm. **(d)** Histogram representation of mean diameters of the rings measured from circle structures. The majority of the circle structures have a diameter of 60-65 nm, which is in good agreement with that of the central tube of *C. elegans* centrioles.

It was previously reported that SAS-5 and SAS-6 are mutually dependent for their centrosome localization and removal of either one results in failure of centriole duplication in *C. elegans* (Pelletier et al. 2006). Observation of the tadpole- and dumbbell-like structures rather than the cartwheel structure of recombinant *C. elegans* SAS-6 (**Figure 27a**) suggests that the nine-fold symmetry formation of *C. elegans* centrioles needs additional symmetry-ensuring factor to associate with SAS-6. In this case, SAS-5 may take an important part.
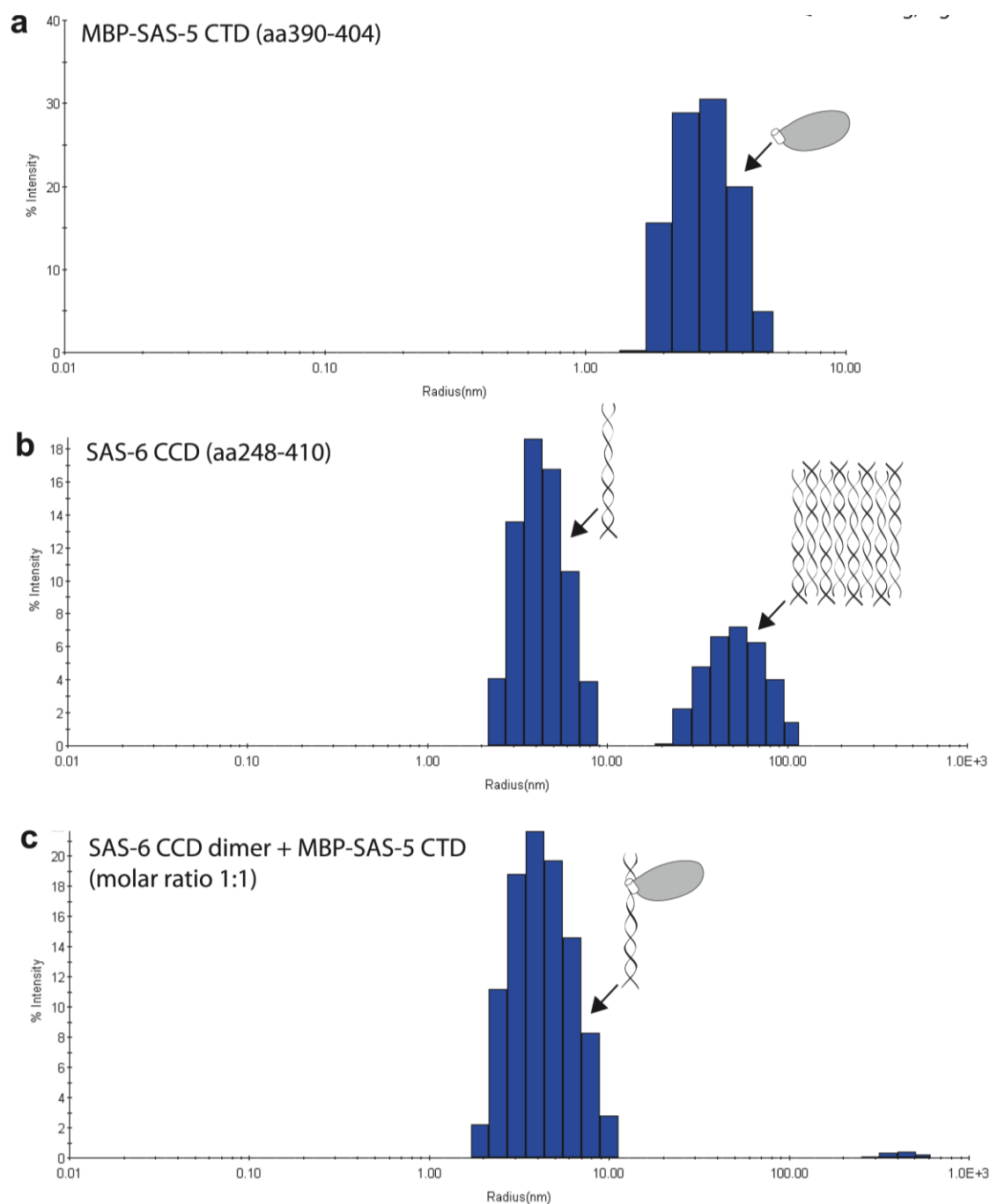
Figure 29 DLS-determined size distributions of SAS-5 CTD, SAS-6 CCD, and their complex. **(a)** MBP-SAS-5 CTD (residues 390-404, 20μM) shows a monodisperse signal. **(b)** SAS-6 CCD (residues 248-410, 15μM) shows two peaks. The majority of the mass (~99.9%) is in the 4-nm peak, which is the dimer of the SAS-6 CCD, whereas the 50-nm peak only represents 0.1% of the total mass and is likely the oligomeric form of the CCD, which, without the obstruction of the head group, could form a large assembly. **(c)** Mixture of the SAS-6 CCD dimer and the SAS-5 CTD (1:1, 15μM each) shows a single peak, which suggests that binding of SAS-5 prevents the oligomerization of the SAS-6 CCD.

To find this out, we used rotary metal shadowing electron microscopy to examine the purified complex of recombinant SAS-6 and MBP-tagged SAS-5. Interestingly, we repeatedly observed curved structures (n=22) with particles of the same size arranged with a similar distance in between (**Figure 28c**). These particles are 2-3 times larger than the SAS-6 head group shown in **Figure 27a** and are likely SAS-5 molecules. The mean diameter of the fit rings of these structures is 63.2 ± 9.2 nm (**Figure 28d**), which is in good agreement with the 60-nm diameter of the central tube of *C. elegans* centrioles (Pelletier et al. 2006).
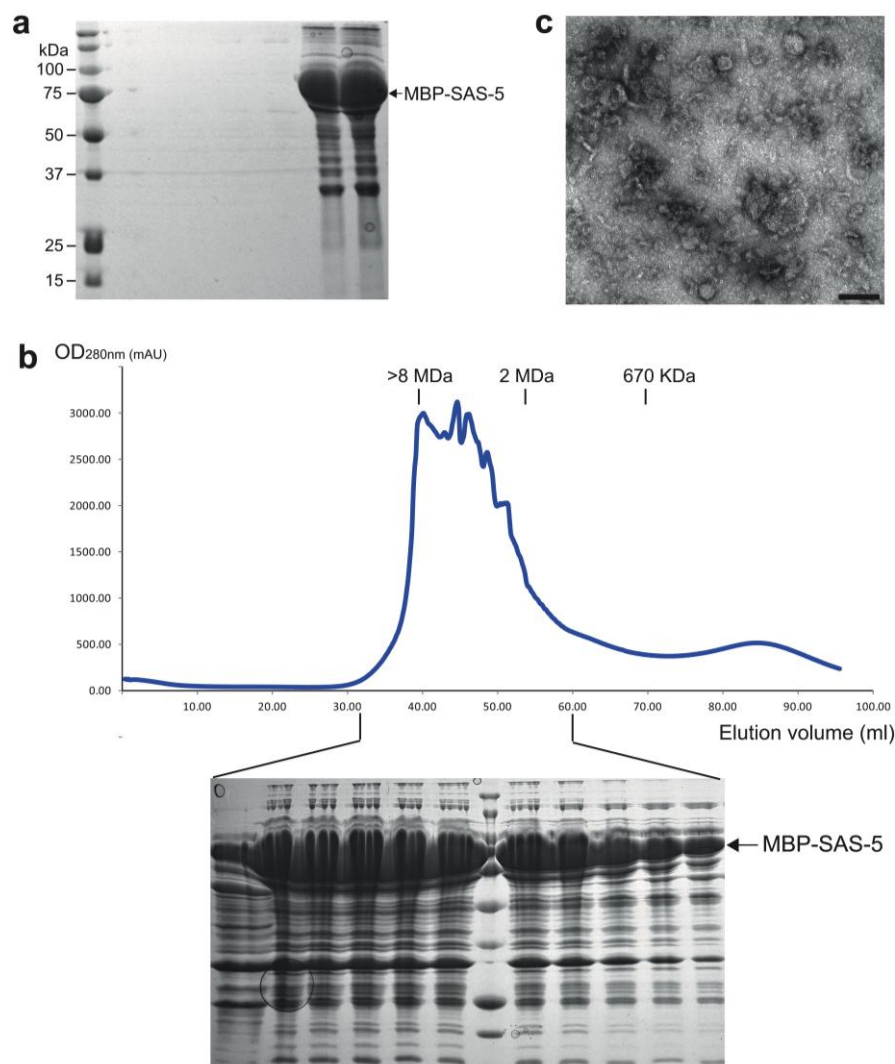
Figure 30 SAS-5 on its own forms aggregates. **(a)** Partially purified recombinant MBP-tagged SAS-5 using ammonium sulfate precipitation. **(b)** Size exclusion chromatography of MBP-SAS-5 on a Superdex-400 (16/60) column. SAS-5 was eluted at factions near the void volume of the column. **(c)** Negative staining electron micrograph of MBP-SAS-5. Large irregular aggregates can be seen. Scale bar, 100 nm.

To find out whether SAS-5 alone can form such structures, we tried to purify the MBP-SAS-5 with a C-terminal 6×His-tag. However, we could not purify the protein using Ni-NTA resin (data not shown), suggesting that the His-tag is inaccessible either due to being shielded by a neighboring structure or being buried in aggregates. Using size exclusion chromatography and negative staining electron microscopy we found that SAS-5 forms large aggregates (**Figure 30**).

Interestingly, the C-terminus of SAS-5 became accessible when mixing with SAS-6 as demonstrated by the success in the pull-down of SAS-5 by Ni-NTA bound SAS-6 (**Figure 15c**), and the complex shows semicircle and arc-like structures (**Figure 29c**). This indicates that binding to SAS-6 indeed releases SAS-5 from its aggregates.

Overall, our data suggest that while SAS-6 and SAS-5 individually form a self-associated conformation, together they can assemble into a highly ordered structure resembling the central tube of the *C. elegans* centrioles.

6. Discussion

Accumulating data indicate that SAS-5 and its functional orthologs, Ana2 in flies and STIL in vertebrates, work cooperatively with SAS-6 in centriole formation (Arquint et al. 2012; Stevens et al. 2010a; Stevens et al. 2010b; Tang et al. 2011; Vulprecht et al. 2012). To find out how SAS-5 assists SAS-6 in centriole assembly, we first need to know how the two proteins interact. Here, we demonstrate that the short SAS-5 CTD specifically interacts with a narrow segment of the SAS-6 coiled coil. We have also solved the crystal structure of the SAS-6 CCD that contains the binding site of SAS-5 and further used structure-based mutagenesis studies to identify the residues on both proteins that are directly involved in their interaction. Interestingly, we found that the interaction is mediated by synergistic hydrophobic and electrostatic interactions of multiple residues on either protein. Single residue mutation analyses showed that mutating any one of these residues completely abolished the interaction. We further showed that the recombinant SAS-5/SAS-6 complex could form semicircular or arc-like structures. How can one put this into the context of centriole duplication?

Unlikely the clearly visible cartwheel structure in non-nematode centrioles, centriole duplication in *C. elegans* begins with a 60-nm central tube dependent on SAS-5/SAS-6. This central tube grows wider and longer at the pronuclear migration stage when SAS-4 is recruited, implying that SAS-4 may contribute to the tube expansion (Pelletier et al. 2006). Further examination of the effect of SAS-4 depletion by RNAi showed that daughter centriole central tubes failed to expand in width whereas elongation occurred normally, suggesting that SAS-4 contributes only to the broadening of the central tube (Pelletier et al. 2006). Additionally, the increase in the width of the central tube coincides with the emergence of an outer wall and hook-like structures around it at positions where MT assembly occurs (Pelletier et al. 2006). In earlier studies, SAS-4 had been revealed to localize to the outer wall of centrioles and have a ring-like distribution around the centriole (Kirkham et al. 2003). Consistently, SAS-4 homologues in flies and humans also localize to the outer wall of centrioles and are essential for recruiting MTs and pericentriolar materials (Gopalakrishnan et al. 2011; Kohlmaier et al. 2009; Tang et al. 2009).
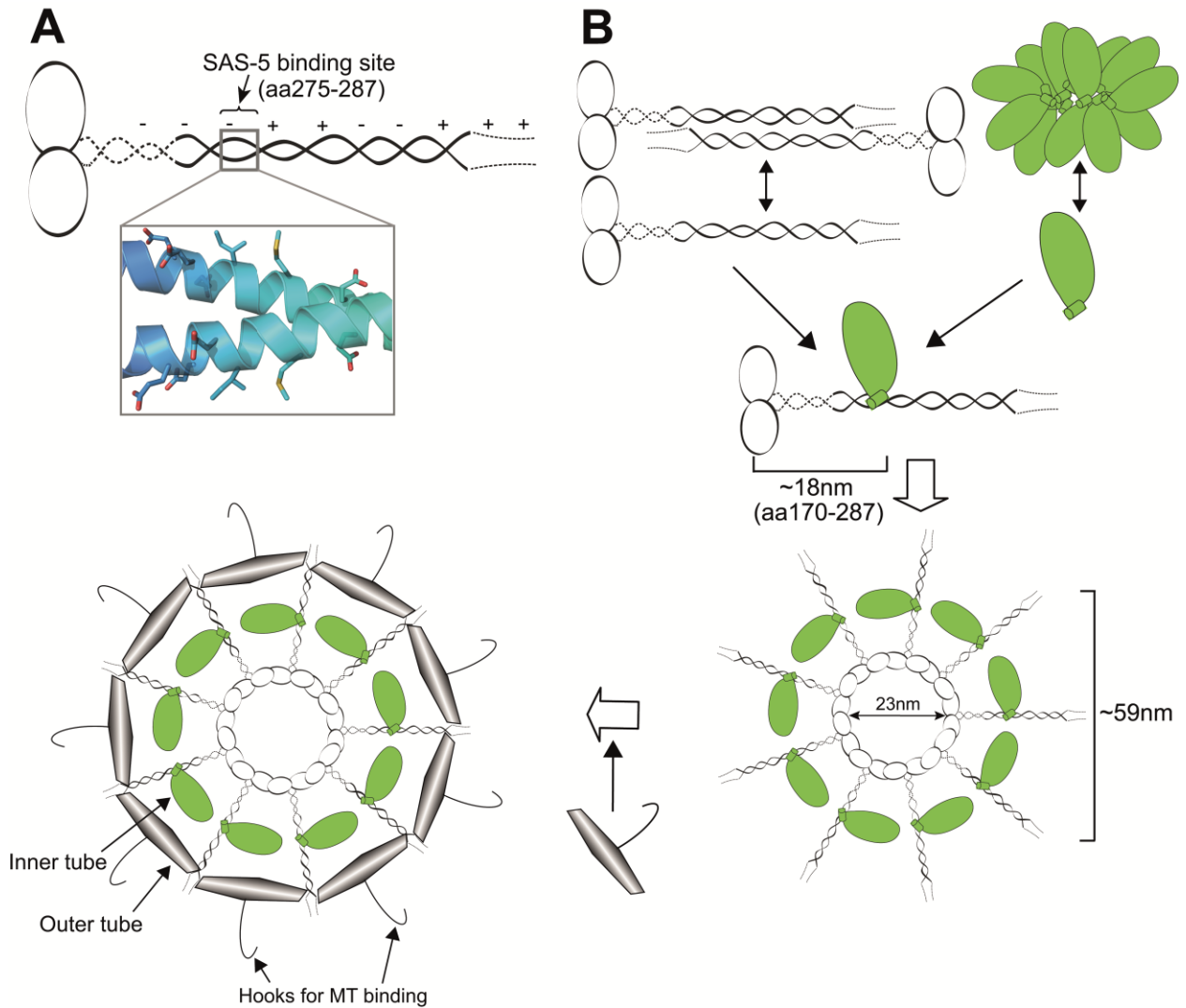
Figure 31 A working model for centriole assembly in *C. elegans*. **(a)** Schematic of SAS-6 dimer and the nearly symmetric arrangement of the residues involved in SAS-5 binding. **(b)** Hypothetical mechanism of recruitment of SAS-5, SAS-6, and SAS-4. Interaction of SAS-5 and SAS-6 releases SAS-5 from the aggregates and opens up the anti-parallel tetramer of SAS-6, which subsequently assembles into a nine-fold cartwheel with SAS-5 arranged as a ring. The diameter of the ring is calculated to be about 59 nm, which is in good agreement with the diameter of the emerging central tube of *C. elegans* centrioles (60 nm). The assembled structure of SAS-5 and SAS-6 may direct the loading of SAS-4, which subsequently generates a wider tube with outward projecting hooks for recruiting MTs as reported previously (Pelletier et al. 2006).

In this study, we show that SAS-6 forms a self-associated tetramer whereas SAS-5 aggregates. Crystal structure of the SAS-6 CCD reveals a periodic charge distribution with the SAS-5 binding site in the center of the coiled coil (**Figure 31a**). We also discovered that binding of the SAS-5 CTD to the SAS-6 coiled coil both releases SAS-5 from its aggregates and prevents the

tetrameric association of SAS-6. This would allow the efficient interaction between the SAS-6 head groups. Our electron microscopy studies demonstrate that the recombinant SAS-5/SAS-6 complex assembles into arc-like structures with an average diameter of 63.2 ± 9.2 nm for the corresponding rings (**Figure 29**). We therefore believe that the emerging 60-nm wide central tube in the procentriole is formed by circularly arranged SAS-5 molecules bound onto the coiled coils of SAS-6 (**Figure 31b**). SAS-5 and SAS-6 together assemble into an unstable tubular structure, whereas loading of SAS-4 stabilizes this tube (now the inner wall) by generating an outer wall with protruding hook-like appendages that serve to recruit the nine singlet MTs (**Figure 31b**).

Apart from disrupting the tetrameric association of SAS-6, SAS-5 may also play a more active role in centriole assembly by inducing conformational changes in the SAS-6 coiled coil. The crystal structure of the SAS-6 CCD indicates that each SAS-6 coiled coil contains two almost symmetrically arranged SAS-5 binding sites (**Figure 31a**). However, results of both ITC and SLS experiments indicate that only one SAS-5 molecule could bind to the SAS-6 dimer (**Figure 28b**). How might this occur? The crystal structure of the *C. elegans* SAS-6 head group shows that it has a similar fold to that of the Xrcc4 family of DNA repair proteins (Junop et al. 2000; Kitagawa et al. 2011b). DNA ligase IV binds to a short segment of the coiled coil domain of Xrcc4. Interestingly, the interaction occurs in a similar manner to that between SAS-5 and SAS-6, i.e., a helix from one protein binds to the central part of the coiled coil of the other. The crystal structure of the Xrcc4-DNA ligase IV complex shows that binding of a short helix of DNA ligase IV induces a bending of the Xrcc4 coiled coil (**Figure 32**) (Sibanda et al. 2001).

In analogy to the DNA ligase IV-induced bending of the Xrcc4 coiled coil, we predict that binding of SAS-5 to the SAS-6 CCD may generate a kink of the SAS-6 coiled coil. It has been proposed that SAS-6 may assemble into a spiral-like structure with the coiled coils pointing outwards to form the spokes (Cottee et al. 2011). If that is the case, the SAS-5 induced structural change on the SAS-6 coiled coil may facilitate the spiral assembly of the oligomeric complex of SAS-6 (**Figure 33**).

Notably, SAS-5 was shown to self-associate in a reported domain-based interactome network in *C. elegans* (Boxem et al. 2008). The self-association of SAS-5 could bridge gaps between SAS-6 coil coils and/or neighboring layers of the spiral, which may strengthen the curved structure

and accelerate the elongation of the central tube. These altogether may serve to enforce the nine-fold symmetry of centrioles and to regulate centriole duplication *in vivo*.
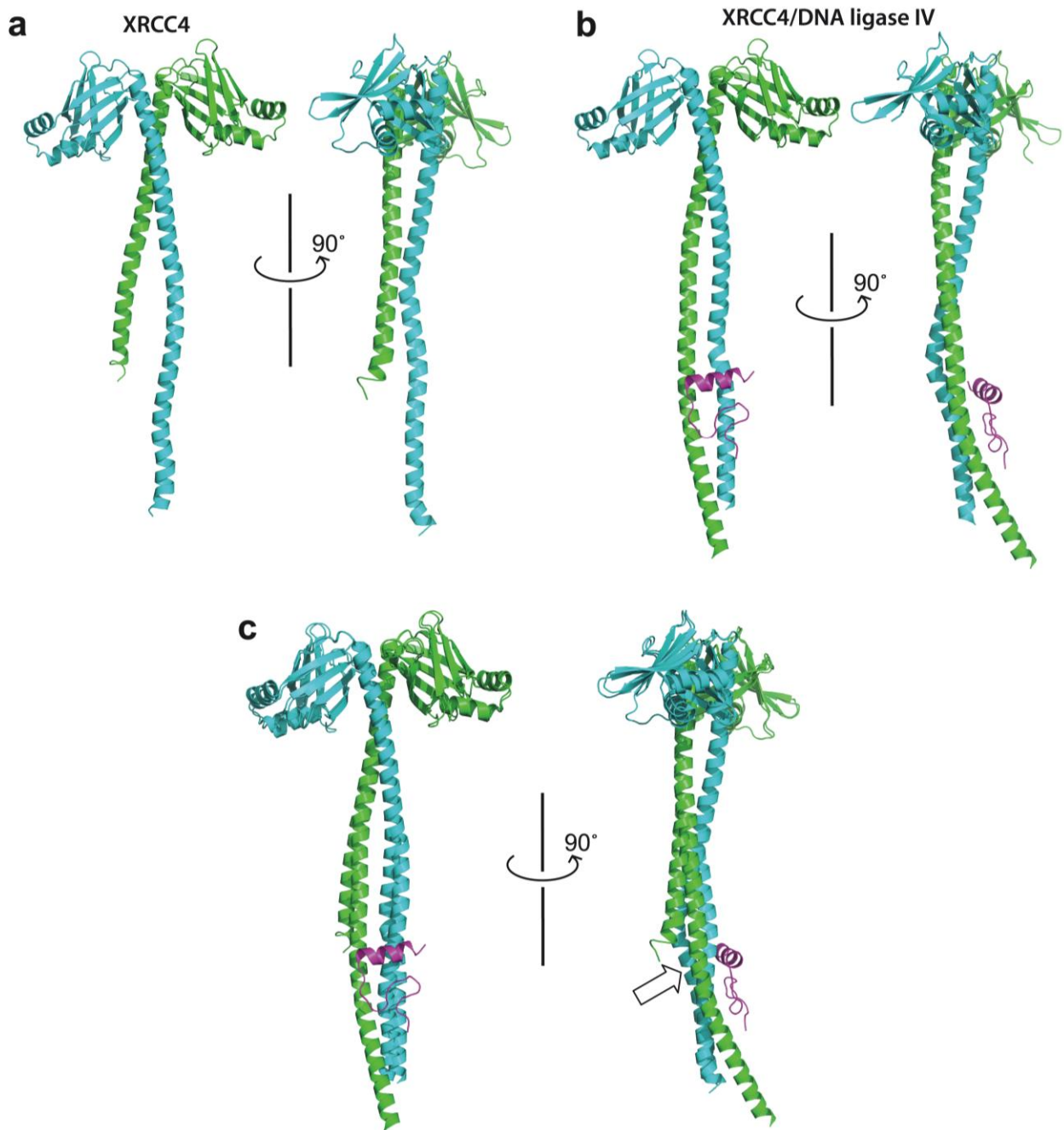
Figure 32 Binding of a helix of DNA ligase IV induces bending of the XRCC4 coiled coil. **(a)** Crystal structure of XRCC4 (pdb code: 1FU1). **(b)** Crystal structure of XRCC4/DNA ligase IV complex (pdb code: 1IK9). **(c)** Superposition of the two structures shows clearly that binding of the short helix of DNA ligase IV induces a kink (indicated by an arrow) of the coiled coil of XRCC4.
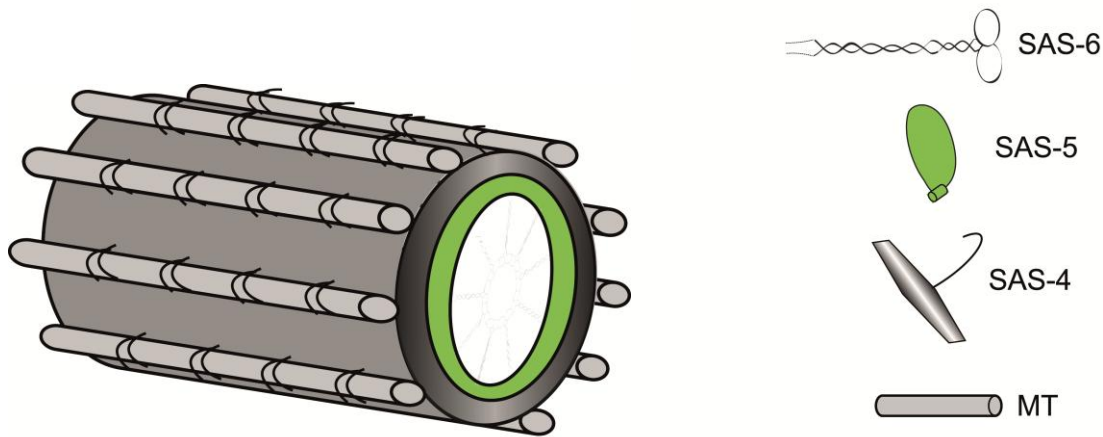
Figure 33 Assembly of *C. elegans* centrioles. Binding of SAS-5 may bend the SAS-6 coiled coil, serving to facilitate the spiral assembly of the SAS-5/SAS-6 complex. This spiral unit would expand the long axis through added SAS-5/SAS-6 proteins to form the initial central tube of the procentriole (SAS-6 is omitted for clarity). Afterwards, SAS-4 would be recruited to generate the outer wall of the central tube. The hook-like structures along the length of the outer tube recruit singlet microtubules to generate the daughter centriole that is identical to the mother.

Since ZYG-1 is required for the recruitment of SAS-5/SAS-6, it is likely that ZYG-1 plays a direct or indirect structural role in the assembly of the central tube. It was reported previously that ZYG-1 phosphorylates SAS-6 and this phosphorylation is crucial for centriole duplication *in vivo* (Kitagawa et al. 2009). The phosphorylation site, serine 123, is located in a long flexible loop (disordered in the crystal structure) next to the dimerization interface of the SAS-6 head group (Kitagawa et al. 2011b). It is conceivable that the phosphorylation by ZYG-1 might strengthen the head group interaction of SAS-6 by providing an electrostatic interaction between the phospho-group and a positively charged surface patch on the opposite molecule (**Figure 34**). The phosphorylation-dependent stable head group interaction may facilitate the SAS-5-induced spiral formation of the complex. This hypothesis is consistent with the observation that ZYG-1-dependent phosphorylation of SAS-6 is needed for both central tube formation and maintenance of SAS-6 at the central tube (Kitagawa et al. 2009). This would explain why we only observed arc-like structures of the SAS-5/SAS-6 complex *in vitro* but not closed rings or spirals. In the future, it should be checked whether adding ZYG-1 to the recombinant SAS5/SAS-6 complex or co-expressing the three proteins could stimulate the formation of ring- or spiral-like structures.
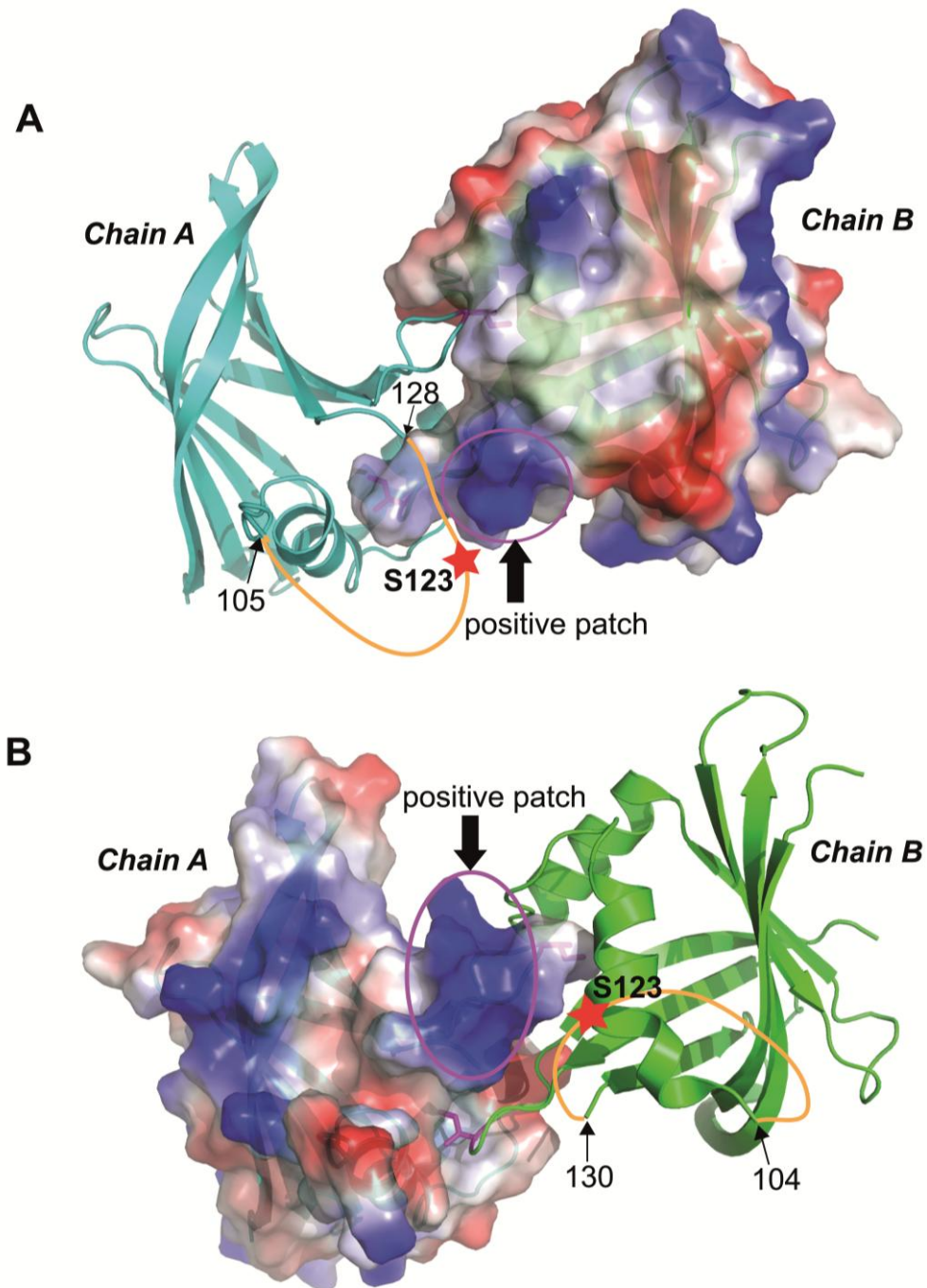
Figure 34 ZYG-1 mediated phosphorylation of S123 on SAS-6 may strengthen the interaction between the head groups of SAS-6. The residue S123 is located in the _exible loop (gold) adjacent to the dimerization interface. It is spatially close to a positively charged surface patch (magenta circles) on the neighboring molecule. It is conceivable that phosphorylation of S123 by ZYG-1 may stabilize the head group interaction of SAS-6 by providing an electrostatic interaction between the phospho-group and the positive surface patch on the opposite molecule. **(A)** View of thespatial orientation of S123 on chain A and a positive patch on chain B. **(B)** The same view as in (A) but showing the spatial orientation of S123 on chain B and a positive patch on chain A.

It was shown previously that SAS-5 failed to localize to centrioles in a mutant *sas-5 (t2033)* corresponding to a single amino acid substitution (R397C) in SAS-5 (Delattre et al. 2004). This substitution disrupts SAS-5 and SAS-6 interaction as demonstrated in a yeast two-hybrid assay (Leidel et al. 2005). Using *in vitro* pull-down and ITC assays, here we showed that the R397C substitution in SAS-5 completely abolishes its interaction with SAS-6. Thus, SAS-5 seems to shuttle to procentrioles through the specific interaction of its CTD with SAS-6.

Given that ZYG-1 dependent recruitment of SAS-6 failed when SAS-5 was depleted (Delattre et al. 2006; Pelletier et al. 2006), SAS-5 likely helps SAS-6 to target the procentriole by forming with SAS-6 a specific recognition site for ZYG-1. Indeed, the combination of conserved charges on the SAS-5 CTD and its binding site on the SAS-6 CCD may provide a unique recognition site for ZYG-1 binding (**Figure 31a**). Notably, although the SAS-5 CTD is sufficient to bind SAS-6, the N-terminal part of SAS-5 seems to ensure the fidelity of the interaction. As shown in **Figure 15c**, while full-length SAS-5 was pulled down stoichiometrically relative to SAS-6, considerably more protein was pulled down for all the three N-terminal truncations of SAS-5. This disproportionate interaction occurs only when using full-length SAS-6 but not for SAS-6 lacking the C-terminal disordered tail (residues 411-492) (see "wt" lanes of **Figure 22d** and **Figure 24b**). Therefore, we conclude that the N-terminal domain of SAS-5 (residues 1 - 389) prevents the nonspecific interaction between the SAS-5 CTD and the unstructured SAS-6 C-terminal tail. Consistently, most of the N-terminal part of SAS-5 is conserved in the three *Caenorhabditis species* (**Figure 35a**), *and* there is a region containing ~20 highly conserved residues in several more divergent SAS-5 homologues (**Figure 35b**). These conserved residues may confer on SAS-5 the ability to regulate its interaction with SAS-6.

How does our finding of SAS-5/SAS-6 interaction relate to the mechanisms of centriole formation in other organisms? Homologues of both *C. elegans* SAS-5 and SAS-6 have been identified in flies and vertebrates, which are *Dm*SAS-6/hSAS-6 and Ana2/STIL, respectively. Crystal structures of several SAS-6 proteins show that the head group of SAS-6 has a conserved fold that mediates the intermolecular interaction in SAS-6 oligomeric assembly, implying that the mechanism of centriole biogenesis may be conserved through evolution. While structural segmentation of SAS-6 family of proteins is easy to define, domain arrangements of SAS-5/Ana2/STIL are very vague because of the lack of distinct motif structures. A ~90 residue region toward the C-terminus of the SAS-5 family of proteins, the STAN motif, was suggested to

be important for their function (Stevens et al. 2010a). However, while the STAN motif is modestly conserved between Ana2 and STIL (31% sequence identity), it is very divergent in SAS-5. Notably, in the STAN motif alignment, none of the four residues in the SAS-5 CTD that participate in its interaction with SAS-6 is conserved in Ana2 or STIL (Stevens et al. 2010a).
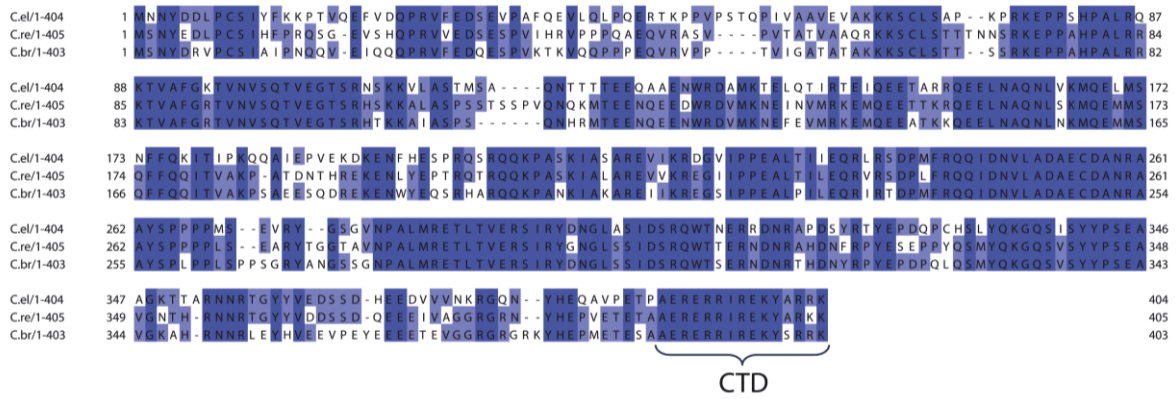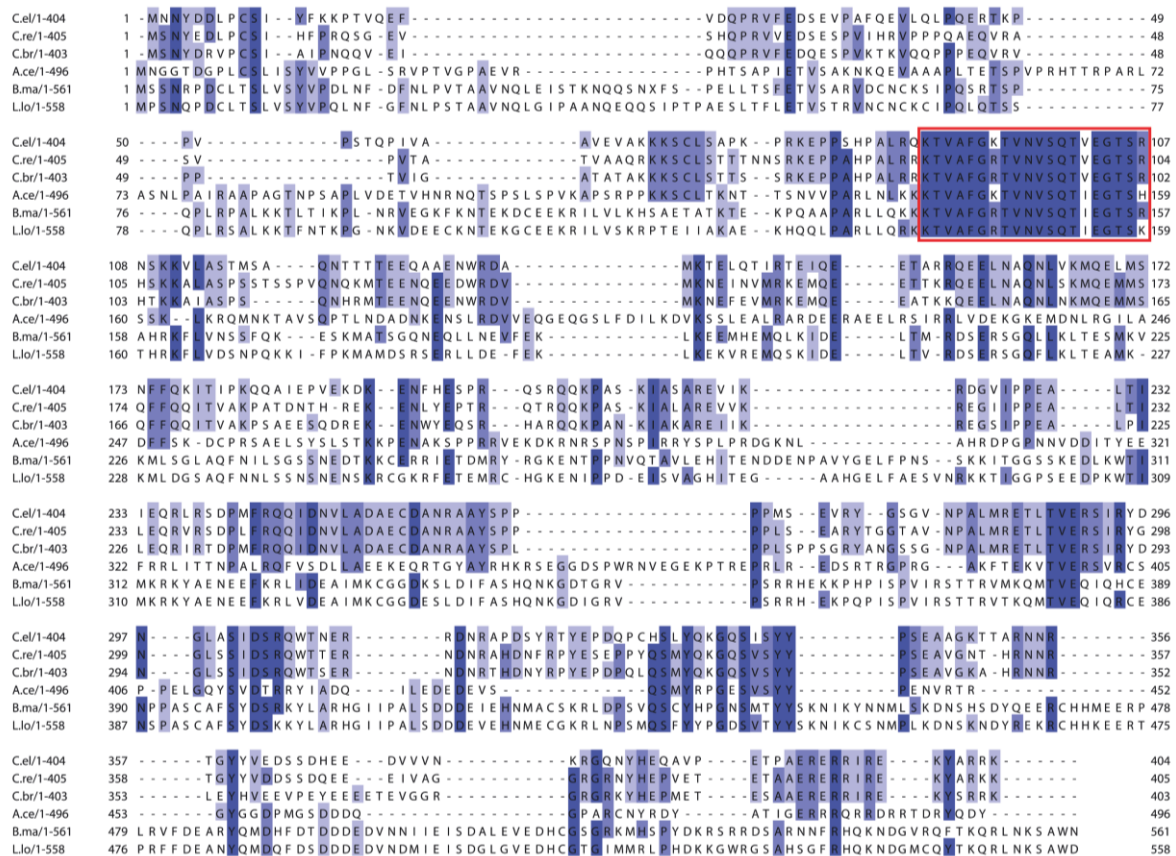
Figure 35 Primary sequence alignment of SAS-5 homologues in different nematode species. **(a)** Sequence alignment of SAS-5 from three Caenorhabditis species. C.el, *Caenorhabditis elegans*; C.re, *C. remanei*; C.br, *C. briggsae*. **(b)** Sequence alignment of six more divergent SAS-5 proteins. The highly conserved region is boxed. A.ce, *Ancylostomaceylanicum*; B.me, *Brugia malayi*; L.lo, *Loa loa*.

Moreover, unlike the helical structure of the SAS-5 CTD, the corresponding regions of Ana2 and STIL in this alignment are predicted to be disordered by the PSIPRED protein structure prediction server (http://bioinf.cs.ucl.ac.uk/psipred/). Most recently, a second conserved motif called TIM was identified at the extreme C-terminus of STIL/Ana2 (Arquint et al. 2012). This motif is predicted to form a helix as the SAS-5 CTD. More interestingly, we found that the TIM motif contains two of the four conserved residues that we found in the SAS-5 CTD essential for its interaction with SAS-6 (**Figure 36a**); the other two residues are conserved only in fly or vertebrate species (**Figure 36b**). It is likely that the STAN motif might have a specific function in flies and vertebrates, whereas the TIM motif participates in or regulates the interaction of Ana2/STIL with SAS-6. The different combinations of the four SAS-6-interacting residues in Ana2 and STIL may reflect the uniqueness of their interaction with SAS-6 in flies and vertebrates, respectively. Notably, two recent studies indicate there is no strong physical interaction between STIL and SAS-6 (Arquint et al. 2012; Vulprecht et al. 2012), while robust interaction between Ana2 and *Dm*SAS-6 could be detected in an immunoprecipitation experiment (Stevens et al. 2010a). The other unique feature of *Dm*SAS-6 is that it forms a parallel tetramer (Gopalakrishnan et al. 2010), and the binding site of Ana2 has been mapped onto a region spanning the head group and the N-terminal part of the coiled coil of *Dm*SAS-6 (Stevens et al. 2010a). It will be important to investigate how interaction with SAS-6 and centriolar localization are affected using Ana2/STIL with these potential binding sites mutated.

Intriguingly, in the rotary metal shadowing micrographs of recombinant SAS-6 or SAS-5/SAS-6 we did not observe the 23-nm central hub as that formed by *C. reinhardtii* and *D. rerio* SAS-6 (Kitagawa et al. 2011b; van Breugel et al. 2011), which is notably consistent with the missing cartwheel structure in *C. elegans* centrioles *in vivo.* We found that although the majority of the *C. elegans* SAS-6 coiled coils well folded as seen in the crystal structure, the N-terminal region (approximately residues 220-240) of the SAS-6 coiled coil seems flexible and sensitive to proteolysis (Fig. 1c, asterisks). It needs to be investigated whether the flexibility of this region accounts for the invisible hub structure in *C. elegans* centrioles.

a

```
Cel_SAS-5   391ERERRIREKYARRK404
Dme_Ana2    408LDLENIRNQPKLL420
Hsa_STIL    1276LDVKRLRQLPKLF1288
```

b

SAS-5
```
C.el   373VNKRGQNYHEQAVPETPAERERRIREKYARRK404
C.re   374AGGRGRNYHEPVETETAYAERKKIREKYARKK405
C.br   372GRGRGRKYHEPMETESAAERERRIREKYSRRK403
```

Ana2
```
D.me   389EPVCPQSPMAAT---APSPHIQLDLENIRNQPKLL420
D.er   385GQVTPQSPMAAT---ASSPHIQLDLENIRNQPKLL416
D.se   388EPVCPQSPMAAT---ASSPHIQLDLENIRNQPKLL419
D.an   397DQVGPQSPVTA----SSSPQIQLDLENIRNQPKLL427
D.wi   420---SPGSPQQQA---EP----MLDLENIRNQPKLL444
D.vi   410PQAVHKIPRSGSPAAAQHQEPMLDLDNIRNQPKLL444
D.gr   391PQGIPKMPQQ-----KVCQEQMLDLDNIRNQPKLL420
D.mo   400QHLISSPQ-------ARSQEPMLDLDNIRNQPKLL427
```

STIL
```
H.sa   1253FPESLQPS--ETLKQMNSMNSVGTFLDVKRLRQLPKLF1288
N.le   1253FPESLQPS--ETLKQMNSMNSVGTFLDVKHLRQLPKLF1288
M.fa   1256FPESLQPS--ETLKQMNSMNSVGTFLDVKHLRQLPKLF1291
A.me   1244LPEGLKP---ETLKQMNSMNSVGTFLDVKRLRQLPKLF1278
C.lu   1242FPESLKP---ETLKQMSSMNSVGTFLDVKQLRQLPKLF1276
T.gu    964FHGNLHTPTLEMLNQSNSINSVGTILDVKQLRQLPKLF1001
G.ga   1266FTGNLHAPGLGTLSQSNSVNSVGTILDVKQLRQLPKLF1303
C.po   1252FPEGLQPS--ETLKQMNSMDSVGPFLDVKRLRQLPKLF1287
L.af   1250FSESLQPS--ETLKQMNSMNSVGTFLDVKRLRQLPKLF1285
M.ga   1266FAGKLHAPALGTLNQSNSVNSVGTILDVKQLRQLPKLF1303
C.gr   1296FPGPLQPP--ETLKQMSSMHSVGTFLDVNRLRQLPKLF1331
B.ta   1253FPESLKPS--ETLKQMNSMNSVGTFLDVKRLRQLPKLF1288
```

Figure 36 Comparison of the SAS-5 CTD to the TIM motifs at the extreme C-terminus of Ana2 and STIL. **(a)** The TIM motifs of Ana2 and STIL contain two of the four residues in the SAS-5 CTD that are essential for SAS-6 interaction. *C. el*, *Caenorhabditis elegans*; *D. me*, *Drosophila melanogaster*; *H. sa*, *Homo sapiens*. **(b)** These regions are all predicted to form a α-helix, which is preceded by a disordered loop. Two of the four conserved residues in the SAS-5 CTD essential for SAS-6 interaction are also conserved in flies and vertebrates (arrows), whereas the other two residues are only conserved in flies or vertebrates species (arrowheads).

In summary, our findings uncover the specific interaction between SAS-5 and SAS-6 and provide an explanation for the unique central tube structure in *C. elegans* centrioles. The data further confirm a role for SAS-5 in assisting SAS-6 to determine the 9-fold symmetry of centrioles and suggest a possible mechanism of the regulation. Our results also provide hints for SAS-6 and Ana2/STIL interaction in other organisms and may have general relevance for future studies.

7. Methods and materials

7.1. Cloning, protein expression and purification

Constructs made in this study for expression and purification from *E. coli* cells
are summarized in **Table 2**.

Name	Expression vectors	Cloning ranges	Lengths
*Ce*SAS-6_1/1H	pET29a,	1-492	492
*Ce*SAS-6_1/2	pET15b, HM15b	1,410	410
*Ce*SAS-6_1/2 Δ240-260	pET15b	1-410 (240-260 deleted)	389
*Ce*SAS-6_1/2 Δ261-288	pET15b	1-410 (261-288deleted)	382
*Ce*SAS-6_1/2 Δ275-288	pET15b	1-410 (275-288 deleted)	396
*Ce*SAS-6_1/2 Δ289-316	pET15b	1-410 (289-316 deleted)	382
*Ce*SAS-6_1/2 Δ317-344	pET15b	1-410 (317-344 deleted)	382
*Ce*SAS-6_1/2 mA*	pET15b	1-410	410
*Ce*SAS-6_1/2 mB*	pET15b	1-410	410
*Ce*SAS-6_1/2 mC*	pET15b	1-410	410
*Ce*SAS-6_1/2 mD*	pET15b	1-410	410
*Ce*SAS-6_1/2dEK11[&]	pET15b	1-410 (11a.a deleted)	399
*Ce*SAS-6_1/3	pET15b	1-116	116
*Ce*SAS-6_2/2	pET15b	119-410	292
*Ce*SAS-6_1/4	pET15b	1-358	358
*Ce*SAS-6_1/5	pET15b,	1-303	303
*Ce*SAS-6_1/6	pET15b	1-250	250
*Ce*SAS-6_1/7	pET15b	1-192	192
*Ce*SAS-6_1/7c	pET15b, KiM5α	1-218	218
*Ce*SAS-6_1/7b	pET15b (no-His), pET15b	1-203	203
*Ce*SAS-6_3/2	pET15b, pET29a	192-410	219
*Ce*SAS-6_4/2	pET15b	230-410	181
*Ce*SAS-6_5/2	pET15b	248-410	163
*Ce*SAS-6_6/2	pET15b	304-410	107
*Ce*SAS-6_6/1H	pET29a, pCDF	304-492	189
*Ce*SAS-6_7/1H	pET29a, pCDF	407-492	85

*Ce*SAS-6_5/1H	pET29a	248-492	245
*Ce*SAS-6_1/1HS123D	pET15b, pET29a	1-492	492
*Ce*SAS-6_1/5S123D	pET15b	1-303	303
*Ce*SAS-6_8/8	pET15b	210-290	80
*Ce*SAS-6_8/2	pET15b, pET29a	210-410	200
*Ce*SAS-6_1/9	pET15b, KiM5α	1-170	170
*Ce*SAS-6_1/9 dRKK[%]	pET15b	1-170	167
*Ce*SAS-6_2b/2	pET15b, pET29a	171-410	240
*Ce*SAS-6_1/9I154E	pET15b	1-170	170
*Ce*SAS-6_7b/1H	pETGB	410-492	83
*Ce*SAS-6_4b/2	pET29a	239-410	172
*Ce*SAS-6_4b/4	pET29a	239-358	120
*Ce*SAS-5_1/1	pET15b, KiM5α	1-404	404
*Ce*SAS-5_1/3	KiM5α	1-389	389
*Ce*SAS-5_1/2b	KiM5α	1-320	320
*Ce*SAS-5_1/2	KiM5α	1-260	260
*Ce*SAS-5_2/1	KiM5α	261-404	143
*Ce*SAS-5_2b/1	KiM5α	321-404	84
*Ce*SAS-5_3/1	KiM5α	377-404	28
*Ce*SAS-5_3/1M1-11[§]	KiM5α	377-404	28

*: mA-E275A+E276A+E278A, mB-I279A+M283A, mC-E286A, mD-E287A

&: dEK11-deletion from E108 to K118

%: dRKK-deletion from R116 to K118

§: M1-E391A+E393A, M2-R392A, M3-R394A+R395A, M4-I396A, M5-R397A, M6-E398A, M7-K399A, M8-Y400A, M9-R402A, M10-R403A, M11-K404A

Table 2 Bacterial expression constructs designed for structural and biochemical studies. pET29a has no tag; HM15b has an N-terminal 6×His tag, which can be cut off by the protease thrombin, followed by MBP and cloning site; pET15b has an N-terminal 6×His tag (cleavable by thrombin); KiM5α has a non-cleavable N-terminal MBP tag.

Sequences encoding full-length *C. elegans* SAS-6 (residues 1-492) and SAS-5 (residues 1-404) were amplified by PCR from cDNA and cloned respectively into pET-29a (Novagen) and a custom vector KiM5α that adds an N-terminal MBP tag to the target protein. Truncations of SAS-6

were cloned into pET-15b (Novagen), which provides an N-terminal 6×His tag cleavable by thrombin. Truncations of SAS-5 were cloned in a similar manner to full-length SAS-5. Deletions and point mutations were generated by the QuickChange Kit (Stratagene) and confirmed by DNA sequencing.

All recombinant proteins were expressed in *Escherichia coli* BL21 (DE3) cells. The cells were grown at 37°C. At an OD_{600} of 0.6-0.8, the cells were cold shocked on ice for 10 min and then shifted to 18°C. Protein induction was done overnight with 0.5 mM of isopropyl-beta-D-thiogalactopyranoside (IPTG). The cells were harvested and resuspended in cold lysis buffer (20 mM Tris-HCl (pH 8), 300 mM NaCl, 20 mM imidazole, and 5% glycerol). The cells were broken by the EmulsiFlex-C3 homogenizer (Avestin) and the lysate was cleared by centrifugation at 30,000 × g for 30 min. The supernatant was filtered through a 0.4 μm filter and loaded onto a Ni-HiTrap column (GE Healthcare) pre-equilibrated in the same lysis buffer. The column was washed with 5 × column volume (cv) of lysis buffer, and bound protein was eluted by a liner gradient concentration of imidazole (20 - 500 mM, 10 × cv) in the lysis buffer. The N-terminal 6×His tag was removed by incubation with 2% (w/w) of thrombin overnight at 4°C. The protein was concentrated and further purified with a Superdex-200 16/60 column (GE Healthcare) pre-equilibrated with 20 mM Tris-HCl (pH 8), 50 mM NaCl and 5% glycerol. The protein was concentrated to 10 mg/ml, divided into aliquots and stored at -80°C.

Selenomethionine(SeMet)-substituted SAS-6 CCD (residues 248-410) for crystallization was expressed using M9 minimal medium supplemented with all amino acids (2 mg/ml) except for methionine. Prior to induction, L-SeMet was added to 80 mg/l, and additional threonine, lysine, phenylalanine, leucine, isoleucine, and valine were added to inhibit the methionine biosynthetic pathway (Doublie 1997). The SeMet-protein was purified as described above, except for the addition of 15 mM β-mercaptoethanol (b-ME) for Ni-HiTrap purification and 10 mM dithiothreitol (DTT) for gel filtration.

7.2. Crystallization and data collection

SAS-6 CCD (residues 248-410) was crystallized at 4°C by the hanging drop method against a reservoir solution containing 0.1 M tri-sodium citrate (pH 5.6), 10% (w/v) PEG 4000 and 10% (v/v) isopropanol. Rod-shaped crystals appeared in two days and reached the maximal size of ~0.03 × 0.03 × 0.5 mm after one week. The crystals belong to space group $P6_1$ (a = b = 140.29 Å, c =

74.67 Å). For harvesting, crystals were soaked in the same reservoir solution augmented with increasing concentrations of glycerol (final concentration 20% [v/v]), loop mounted, and flash frozen in liquid nitrogen. Diffraction data to 3.3Å resolution was collected at the beamline ID23-1 at the European Synchrotron Radiation Facility (ESRF). A complete and highly redundant data set at the anomalous peak of Se (λ = 0.9792 Å) was collected.

7.3. Structure determination and model docking

Data were integrated using iMosflm (Battye et al. 2011) and scaled using the program SCALA (Evans 2006). Selenium sites were located and experimental maps were calculated using AutoSol in the software suite Phenix (Terwilliger et al. 2009). Models were built using the program COOT (Emsley and Cowtan 2004), and refinement carried out using CNS (Brunger et al. 1998) to final R_{work} of 0.258 and R_{free} of 0.299.

For modeling of the SAS-6 coiled coil tetramer and the SAS-6/SAS-5 complex, we submitted our solved crystal structure of the SAS-6 CCD and a theoretical helical model the SAS-5 CTD to the web-based ClusPro 2.0 docking server (http://cluspro.bu.edu/), which filters docked conformations with good surface and charge complementarity and ranks them based on their clustering properties. The docking was carried out with default parameters.

7.4. Pull-down assays

Small aliquots (50 µl of beads) of 6×His-tagged full-length or truncated SAS-6 proteins bound to Ni-NTA beads (QIAGEN) were used to pull down MBP-tagged SAS-5 protein from crude cell lysate. Afterwards, the beads were washed using 5 × cv of lysis buffer supplemented with 0.1% Triton X-100 to remove contaminants. After boiling for 2 min in 1×SDS loading buffer, the proteins were separated on an SDS-PAGE gel and stained with Coomassie Brilliant Blue G250 (Sigma-Aldrich). In a reciprocal binding experiment, we loaded MBP-tagged SAS-5 CTD onto amylose beads and then used these beads to pull down SAS-6 proteins. Subsequent wash and examination were carried out in the same way as the Ni-NTA pull-down. As a negative control to show that SAS-5 proteins did not nonspecifically bind to Ni-NTA beads and SAS-6 did not bind to MBP and/or the amylose beads, mock experiments were carried out, in which we used Ni-NTA bound 6×His-tagged MBP to pull down SAS-5 or MBP alone on amylose beads to pull down SAS-6.

7.5. Isothermal titration calorimetry

The purified SAS-6 CCD and the synthetic SAS-5 CTD (residues 390-404) or purified recombinant SAS-5 CTD fused to MBP were dialyzed overnight against a buffer containing 20 mM Tris-HCl (pH 8.0) and 50 mM NaCl. Protein and peptide concentrations were determined by ND-1000 spectrophotometer (PEQlab). ITC experiments were carried out at 25°C using a VP-ITC Microcal calorimeter (MicroCal, GE healthcare). The cell contained 1.4 ml of SAS-6 CCD dimer (10 µM, wild type or mutants), which was titrated with 29 injections × 10 µl of the SAS-5 CTD at a concentration of 105 µM. The ITC data were analyzed using the program Origin version 7.0 provided by MicroCal. One-site binding model was used to fit the integrated data to calculate the stoichiometry and binding constants.

For dilution ITC experiments, 300 µM of the SAS-6 CCD (dialyzed overnight) alone or in mixture with 450 µM of the SAS-5 CTD was injected into 1.4 ml of the same buffer in the temperature-controlled cell at 25°C. A total of 29 × 10 µl injections were carried out. Dissociation constants (K_d) were calculated by integrating and fitting the endothermic heat pulses to a dimer-tetramer dissociation model using Orign Version 7.0 (MicroCal).

7.6. Static light scattering (SLS)

The SLS studies were carried out on an instrument from the Wyatt Technology Corp. The liquid chromatography equipment consists of a HPLC system (Agilent Technologies) connected in series with a triple-angle laser light scattering detector (miniDAWN TREOS), a UV detector at 280 nm (Agilent technologies) and a refractive index detector (RI-101, Shodex). 100 µl of protein samples (4 mg/ml for the SAS-6 CCD and 6 mg/ml for the MBP-SAS-5 CTD) were eluted from a Superdex 200 10/300 GL column (GE healthcare) at a flow rate of 0.5 ml per min. Data analysis was carried out using the Astra software (Wyatt technology).

7.7. Electron microscopy

Purified full-length SAS-6, either alone or in complex with MBP-tagged SAS-5, was prepared at 0.05-0.1 mg/ml in 100 mM ammonium bicarbonate (pH 7.5), 30% (v/v) glycerol. The samples were sprayed onto freshly cleaved mica chips. After drying in a Bal-Tec MED020 high vacuum coater (Leica Microsystems) for at least 6 h, the chips were rotary shadowed with 0.7 nm platinum/carbon at an elevation angle of 4 degree for SAS-6 and 7 degree for SAS-5/SAS-6 complex

and with carbon at a tilted angle of 45 degree. Electron micrographs were taken on an FEI Morgagni 268D transmission electron microscope operated at 80 kV equipped with a 11 mega-pixel CCD camera. Images were examined and analyzed using ImageJ (http://imagej.nih.gov/ij/).

8. References

Absalon S, Blisnick T, Kohl L, Toutirais G, Dore G, Julkowska D, Tavenet A, and Bastin P. 2008. Intraflagellar transport and functional analysis of genes required for flagellum formation in trypanosomes. Mol Biol Cell 19(3):929-944.

Andersen JS, Wilkinson CJ, Mayor T, Mortensen P, Nigg EA, and Mann M. 2003. Proteomic characterization of the human centrosome by protein correlation profiling. Nature 426(6966):570-574.

Arquint C, Sonnen KF, Stierhof YD, and Nigg EA. 2012. Cell-cycle-regulated expression of STIL controls centriole number in human cells. J Cell Sci.

Arts HH, Bongers EM, Mans DA, van Beersum SE, Oud MM, Bolat E, Spruijt L, Cornelissen EA, Schuurs-Hoeijmakers JH, de Leeuw N et al. . 2011. C14ORF179 encoding IFT43 is mutated in Sensenbrenner syndrome. J Med Genet 48(6):390-395.

Azimzadeh J, Hergert P, Delouvee A, Euteneuer U, Formstecher E, Khodjakov A, and Bornens M. 2009. hPOC5 is a centrin-binding protein required for assembly of full-length centrioles. J Cell Biol 185(1):101-114.

Azimzadeh J, and Marshall WF. 2010. Building the centriole. Curr Biol 20(18):R816-825.

Basto R, Brunk K, Vinadogrova T, Peel N, Franz A, Khodjakov A, and Raff JW. 2008. Centrosome amplification can initiate tumorigenesis in flies. Cell 133(6):1032-1042.

Basto R, Lau J, Vinogradova T, Gardiol A, Woods CG, Khodjakov A, and Raff JW. 2006. Flies without centrioles. Cell 125(7):1375-1386.

Battye TG, Kontogiannis L, Johnson O, Powell HR, and Leslie AG. 2011. iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. Acta Crystallogr D Biol Crystallogr 67(Pt 4):271-281.

Beales PL, Bland E, Tobin JL, Bacchelli C, Tuysuz B, Hill J, Rix S, Pearson CG, Kai M, Hartley J et al. . 2007. IFT80, which encodes a conserved intraflagellar transport protein, is mutated in Jeune asphyxiating thoracic dystrophy. Nat Genet 39(6):727-729.

Behal RH, Miller MS, Qin H, Lucker BF, Jones A, and Cole DG. 2012. Subunit interactions and organization of the Chlamydomonas reinhardtii intraflagellar transport complex A proteins. J Biol Chem 287(15):11689-11703.

Bell LR, Stone S, Yochem J, Shaw JE, and Herman RK. 2006. The molecular identities of the Caenorhabditis elegans intraflagellar transport genes dyf-6, daf-10 and osm-1. Genetics 173(3):1275-1286.

Berbari NF, O'Connor AK, Haycraft CJ, and Yoder BK. 2009. The primary cilium as a complex signaling center. Curr Biol 19(13):R526-535.

Berger I, Fitzgerald DJ, and Richmond TJ. 2004. Baculovirus expression system for heterologous multiprotein complexes. Nat Biotechnol 22(12):1583-1587.

Bettencourt-Dias M, and Glover DM. 2007. Centrosome biogenesis and function: centrosomics brings new understanding. Nat Rev Mol Cell Biol 8(6):451-463.

Bettencourt-Dias M, Rodrigues-Martins A, Carpenter L, Riparbelli M, Lehmann L, Gatt MK, Carmo N, Balloux F, Callaini G, and Glover DM. 2005. SAK/PLK4 is required for centriole duplication and flagella development. Curr Biol 15(24):2199-2207.

Bhogaraju S, Taschner M, Morawetz M, Basquin C, and Lorentzen E. 2011. Crystal structure of the intraflagellar transport complex 25/27. EMBO J 30(10):1907-1918.

Bieniossek C, Richmond TJ, and Berger I. 2008. MultiBac: multigene baculovirus-based eukaryotic protein complex production. Curr Protoc Protein Sci Chapter 5:Unit 5 20.

Blacque OE, Li C, Inglis PN, Esmail MA, Ou G, Mah AK, Baillie DL, Scholey JM, and Leroux MR. 2006. The WD repeat-containing protein IFTA-1 is required for retrograde intraflagellar transport. Mol Biol Cell 17(12):5053-5062.

Bloodgood RA. 2010. Sensory reception is an attribute of both primary cilia and motile cilia. J Cell Sci 123(Pt 4):505-509.

Bornens M. 2012. The centrosome in cells and organisms. Science 335(6067):422-426.

Boxem M, Maliga Z, Klitgord N, Li N, Lemmens I, Mana M, de Lichtervelde L, Mul JD, van de Peut D, Devos M et al. . 2008. A protein domain-based interactome network for C. elegans early embryogenesis. Cell 134(3):534-545.

Bredrup C, Saunier S, Oud MM, Fiskerstrand T, Hoischen A, Brackman D, Leh SM, Midtbo M, Filhol E, Bole-Feysot C et al. . 2011. Ciliopathies with skeletal anomalies and renal insufficiency due to mutations in the IFT-A gene WDR19. Am J Hum Genet 89(5):634-643.

Brito DA, Gouveia SM, and Bettencourt-Dias M. 2012. Deconstructing the centriole: structure and number control. Curr Opin Cell Biol 24(1):4-13.

Brunger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS et al. . 1998. Crystallography & NMR system: A new software suite for macromolecular structure determination. Acta Crystallogr D Biol Crystallogr 54(Pt 5):905-921.

Cano DA, Murcia NS, Pazour GJ, and Hebrok M. 2004. Orpk mouse model of polycystic kidney disease reveals essential role of primary cilia in pancreatic tissue organization. Development 131(14):3457-3467.

Carvalho-Santos Z, Azimzadeh J, Pereira-Leal JB, and Bettencourt-Dias M. 2011. Evolution: Tracing the origins of centrioles, cilia, and flagella. J Cell Biol 194(2):165-175.

Castiel A, Danieli MM, David A, Moshkovitz S, Aplan PD, Kirsch IR, Brandeis M, Kramer A, and Izraeli S. 2011. The Stil protein regulates centrosome integrity and mitosis through suppression of Chfr. J Cell Sci 124(Pt 4):532-539.

Christensen ST, Pedersen LB, Schneider L, and Satir P. 2007. Sensory cilia and integration of signal transduction in human health and disease. Traffic 8(2):97-109.

Cole DG. 2003. The intraflagellar transport machinery of Chlamydomonas reinhardtii. Traffic 4(7):435-442.

Cole DG, Diener DR, Himelblau AL, Beech PL, Fuster JC, and Rosenbaum JL. 1998. Chlamydomonas kinesin-II-dependent intraflagellar transport (IFT): IFT particles contain proteins required for ciliary assembly in Caenorhabditis elegans sensory neurons. J Cell Biol 141(4):993-1008.

Cortellino S, Wang C, Wang B, Bassi MR, Caretti E, Champeval D, Calmont A, Jarnik M, Burch J, Zaret KS et al. . 2009. Defective ciliogenesis, embryonic lethality and severe impairment of the Sonic Hedgehog pathway caused by inactivation of the mouse complex A intraflagellar transport gene Ift122/Wdr10, partially overlapping with the DNA repair gene Med1/Mbd4. Dev Biol 325(1):225-237.

Cottee MA, Raff JW, Lea SM, and Roque H. 2011. SAS-6 oligomerization: the key to the centriole? Nat Chem Biol 7(10):650-653.

Czarnecki PG, and Shah JV. 2012. The ciliary transition zone: from morphology and molecules to medicine. Trends Cell Biol 22(4):201-210.

Dammermann A, Maddox PS, Desai A, and Oegema K. 2008. SAS-4 is recruited to a dynamic structure in newly forming centrioles that is stabilized by the gamma-tubulin-mediated addition of centriolar microtubules. J Cell Biol 180(4):771-785.

Dammermann A, Muller-Reichert T, Pelletier L, Habermann B, Desai A, and Oegema K. 2004. Centriole assembly requires both centriolar and pericentriolar material proteins. Dev Cell 7(6):815-829.

Das AK, Cohen PW, and Barford D. 1998. The structure of the tetratricopeptide repeats of protein phosphatase 5: implications for TPR-mediated protein-protein interactions. EMBO J 17(5):1192-1199.

Debec A, Sullivan W, and Bettencourt-Dias M. 2010. Centrioles: active players or passengers during mitosis? Cell Mol Life Sci 67(13):2173-2194.

Delattre M, Canard C, and Gonczy P. 2006. Sequential protein recruitment in C. elegans centriole formation. Curr Biol 16(18):1844-1849.

Delattre M, and Gonczy P. 2004. The arithmetic of centrosome biogenesis. J Cell Sci 117(Pt 9):1619-1630.

Delattre M, Leidel S, Wani K, Baumer K, Bamat J, Schnabel H, Feichtinger R, Schnabel R, and Gonczy P. 2004. Centriolar SAS-5 is required for centrosome duplication in C. elegans. Nat Cell Biol 6(7):656-664.

Dix CI, and Raff JW. 2007. Drosophila Spd-2 recruits PCM to the sperm centriole, but is dispensable for centriole duplication. Curr Biol 17(20):1759-1764.

Doublie S. 1997. Preparation of selenomethionyl proteins for phase determination. Methods Enzymol 276:523-530.

Efimenko E, Blacque OE, Ou G, Haycraft CJ, Yoder BK, Scholey JM, Leroux MR, and Swoboda P. 2006. Caenorhabditis elegans DYF-2, an orthologue of human WDR19, is a component of the intraflagellar transport machinery in sensory cilia. Mol Biol Cell 17(11):4801-4811.

Eggenschwiler JT, and Anderson KV. 2007. Cilia and developmental signaling. Annu Rev Cell Dev Biol 23:345-373.

Emsley P, and Cowtan K. 2004. Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr 60(Pt 12 Pt 1):2126-2132.

Evans P. 2006. Scaling and assessment of data quality. Acta Crystallogr D Biol Crystallogr 62(Pt 1):72-82.

Fan ZC, Behal RH, Geimer S, Wang Z, Williamson SM, Zhang H, Cole DG, and Qin H. 2010. Chlamydomonas IFT70/CrDYF-1 is a core component of IFT particle complex B and is required for flagellar assembly. Mol Biol Cell 21(15):2696-2706.

Fitzgerald DJ, Berger P, Schaffitzel C, Yamada K, Richmond TJ, and Berger I. 2006. Protein complex expression by using multigene baculoviral vectors. Nat Methods 3(12):1021-1032.

Fitzgerald DJ, Schaffitzel C, Berger P, Wellinger R, Bieniossek C, Richmond TJ, and Berger I. 2007. Multiprotein expression strategy for structural biology of eukaryotic complexes. Structure 15(3):275-279.

Gilula NB, and Satir P. 1972. The ciliary necklace. A ciliary membrane specialization. J Cell Biol 53(2):494-509.

Giddings TH, Jr., Meehl JB, Pearson CG, and Winey M. 2010. Electron tomography and immuno-labeling of Tetrahymena thermophila basal bodies. Methods Cell Biol 96:117-141.

Goetz SC, and Anderson KV. 2010. The primary cilium: a signalling centre during vertebrate development. Nat Rev Genet 11(5):331-344.

Gopalakrishnan J, Guichard P, Smith AH, Schwarz H, Agard DA, Marco S, and Avidor-Reiss T. 2010. Self-assembling SAS-6 multimer is a core centriole building block. J Biol Chem 285(12):8759-8770.

Gopalakrishnan J, Mennella V, Blachon S, Zhai B, Smith AH, Megraw TL, Nicastro D, Gygi SP, Agard DA, and Avidor-Reiss T. 2011. Sas-4 provides a scaffold for cytoplasmic complexes and tethers them in a centrosome. Nat Commun 2:359.

Habedanck R, Stierhof YD, Wilkinson CJ, and Nigg EA. 2005. The Polo kinase Plk4 functions in centriole duplication. Nat Cell Biol 7(11):1140-1146.

Han YG, Kwok BH, and Kernan MJ. 2003. Intraflagellar transport is required in Drosophila to differentiate sensory cilia but not sperm. Curr Biol 13(19):1679-1686.

Hirokawa N, Tanaka Y, Okada Y, and Takeda S. 2006. Nodal flow and the generation of left-right asymmetry. Cell 125(1):33-45.

Hou Y, Qin H, Follit JA, Pazour GJ, Rosenbaum JL, and Witman GB. 2007. Functional analysis of an individual IFT protein: IFT46 is required for transport of outer dynein arms into flagella. J Cell Biol 176(5):653-665.

Hung LY, Tang CJ, and Tang TK. 2000. Protein 4.1 R-135 interacts with a novel centrosomal protein (CPAP) which is associated with the gamma-tubulin complex. Mol Cell Biol 20(20):7813-7825.

Ibanez-Tallon I, Pagenstecher A, Fliegauf M, Olbrich H, Kispert A, Ketelsen UP, North A, Heintz N, and Omran H. 2004. Dysfunction of axonemal dynein heavy chain Mdnah5 inhibits

ependymal flow and reveals a novel mechanism for hydrocephalus formation. Hum Mol Genet 13(18):2133-2141.

Igarashi P, and Somlo S. 2002. Genetics and pathogenesis of polycystic kidney disease. J Am Soc Nephrol 13(9):2384-2398.

Iomini C, Li L, Esparza JM, and Dutcher SK. 2009. Retrograde intraflagellar transport mutants identify complex A proteins with multiple genetic interactions in Chlamydomonas reinhardtii. Genetics 183(3):885-896.

Ishikawa H, Kubo A, and Tsukita S. 2005. Odf2-deficient mother centrioles lack distal/subdistal appendages and the ability to generate primary cilia. Nat Cell Biol 7(5):517-524.

Ishikawa H, and Marshall WF. 2011. Ciliogenesis: building the cell's antenna. Nat Rev Mol Cell Biol 12(4):222-234.

Jin H, and Nachury MV. 2009. The BBSome. Curr Biol 19(12):R472-473.

Jin H, White SR, Shida T, Schulz S, Aguiar M, Gygi SP, Bazan JF, and Nachury MV. 2010. The conserved Bardet-Biedl syndrome proteins assemble a coat that traffics membrane proteins to cilia. Cell 141(7):1208-1219.

Johnson KA, and Rosenbaum JL. 1992. Polarity of flagellar assembly in Chlamydomonas. J Cell Biol 119(6):1605-1611.

Junop MS, Modesti M, Guarne A, Ghirlando R, Gellert M, and Yang W. 2000. Crystal structure of the Xrcc4 DNA repair protein and implications for end joining. EMBO J 19(22):5962-5970.

Kemp CA, Kopish KR, Zipperlen P, Ahringer J, and O'Connell KF. 2004. Centrosome maturation and duplication in C. elegans require the coiled coil protein SPD-2. Dev Cell 6(4):511-523.

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

Kiser RL, Wolf MT, Martin JL, Zalewski I, Attanasio M, Hildebrandt F, and Klemmer P. 2004. Medullary cystic kidney disease type 1 in a large Native-American kindred. Am J Kidney Dis 44(4):611-617.

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

Kitagawa D, Kohlmaier G, Keller D, Strnad P, Balestra FR, Fluckiger I, and Gonczy P. 2011a. Spindle positioning in human cells relies on proper centriole formation and on the microcephaly proteins CPAP and STIL. J Cell Sci 124(Pt 22):3884-3893.

Kitagawa D, Vakonakis I, Olieric N, Hilbert M, Keller D, Olieric V, Bortfeld M, Erat MC, Fluckiger I, Gonczy P et al. . 2011b. Structural basis of the 9-fold symmetry of centrioles. Cell 144(3):364-375.

Kleylein-Sohn J, Westendorf J, Le Clech M, Habedanck R, Stierhof YD, and Nigg EA. 2007. Plk4-induced centriole biogenesis in human cells. Dev Cell 13(2):190-202.

Kobayashi T, and Dynlacht BD. 2011. Regulating the transition from centriole to basal body. J Cell Biol 193(3):435-444.

Kobayashi T, Gengyo-Ando K, Ishihara T, Katsura I, and Mitani S. 2007. IFT-81 and IFT-74 are required for intraflagellar transport in C. elegans. Genes Cells 12(5):593-602.

Kohlmaier G, Loncarek J, Meng X, McEwen BF, Mogensen MM, Spektor A, Dynlacht BD, Khodjakov A, and Gonczy P. 2009. Overly long centrioles and defective cell division upon excess of the SAS-4-related protein CPAP. Curr Biol 19(12):1012-1018.

Kon T, Oyama T, Shimo-Kon R, Imamula K, Shima T, Sutoh K, and Kurisu G. 2012. The 2.8 A crystal structure of the dynein motor domain. Nature 484(7394):345-350.

Kozakov D, Hall DR, Beglov D, Brenke R, Comeau SR, Shen Y, Li K, Zheng J, Vakili P, Paschalidis I et al. . 2010. Achieving reliability and high accuracy in automated protein docking: ClusPro, PIPER, SDU, and stability analysis in CAPRI rounds 13-19. Proteins 78(15):3124-3130.

Kozminski KG, Beech PL, and Rosenbaum JL. 1995. The Chlamydomonas kinesin-like protein FLA10 is involved in motility associated with the flagellar membrane. J Cell Biol 131(6 Pt 1):1517-1527.

Kozminski KG, Johnson KA, Forscher P, and Rosenbaum JL. 1993. A motility in the eukaryotic flagellum unrelated to flagellar beating. Proc Natl Acad Sci U S A 90(12):5519-5523.

Kuriyama R, and Borisy GG. 1981. Centriole cycle in Chinese hamster ovary cells as determined by whole-mount electron microscopy. J Cell Biol 91(3 Pt 1):814-821.

Lamb JR, Tugendreich S, and Hieter P. 1995. Tetratrico peptide repeat interactions: to TPR or not to TPR? Trends Biochem Sci 20(7):257-259.

Leidel S, Delattre M, Cerutti L, Baumer K, and Gonczy P. 2005. SAS-6 defines a protein family required for centrosome duplication in C. elegans and in human cells. Nat Cell Biol 7(2):115-125.

Leidel S, and Gonczy 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(3):431-439.

Leidel S, and Gonczy P. 2005. Centrosome duplication and nematodes: recent insights from an old relationship. Dev Cell 9(3):317-325.

Leung GC, Hudson JW, Kozarova A, Davidson A, Dennis JW, and Sicheri F. 2002. The Sak polo-box comprises a structural domain sufficient for mitotic subcellular localization. Nat Struct Biol 9(10):719-724.

Li C, Inglis PN, Leitch CC, Efimenko E, Zaghloul NA, Mok CA, Davis EE, Bialas NJ, Healey MP, Heon E et al. . 2008. An essential role for DYF-11/MIP-T3 in assembling functional intraflagellar transport complexes. PLoS Genet 4(3):e1000044.

Lin B, White JT, Utleg AG, Wang S, Ferguson C, True LD, Vessella R, Hood L, and Nelson PS. 2003. Isolation and characterization of human and mouse WDR19,a novel WD-repeat protein exhibiting androgen-regulated expression in prostate epithelium. Genomics 82(3):331-342.

Lovatt M, Cooper A, and Camilleri P. 1996. Energetics of cyclodextrin-induced dissociation of insulin. Eur Biophys J 24(5):354-357.

Lucker BF, Behal RH, Qin H, Siron LC, Taggart WD, Rosenbaum JL, and Cole DG. 2005. Characterization of the intraflagellar transport complex B core: direct interaction of the IFT81 and IFT74/72 subunits. J Biol Chem 280(30):27688-27696.

Lucker BF, Miller MS, Dziedzic SA, Blackmarr PT, and Cole DG. 2010. Direct interactions of intraflagellar transport complex B proteins IFT88, IFT52, and IFT46. J Biol Chem 285(28):21508-21518.

Marshall WF. 2001. Centrioles take center stage. Curr Biol 11(12):R487-496.

Marshall WF. 2007. What is the function of centrioles? J Cell Biochem 100(4):916-922.

Marshall WF. 2009. Centriole evolution. Curr Opin Cell Biol 21(1):14-19.

Marshall WF, and Nonaka S. 2006. Cilia: tuning in to the cell's antenna. Curr Biol 16(15):R604-614.

McGrath J, and Brueckner M. 2003. Cilia are at the heart of vertebrate left-right asymmetry. Curr Opin Genet Dev 13(4):385-392.

Megraw T. 2011. PP2A targets SAS-5 in centriole assembly. Dev Cell 20(4):416-417.

Megraw TL, Sharkey JT, and Nowakowski RS. 2011. Cdk5rap2 exposes the centrosomal root of microcephaly syndromes. Trends Cell Biol 21(8):470-480.

Mill P, Lockhart PJ, Fitzpatrick E, Mountford HS, Hall EA, Reijns MA, Keighren M, Bahlo M, Bromhead CJ, Budd P et al. . 2011. Human and mouse mutations in WDR35 cause short-rib polydactyly syndromes due to abnormal ciliogenesis. Am J Hum Genet 88(4):508-515.

Mizuno N, Taschner M, Engel BD, and Lorentzen E. 2012. Structural Studies of Ciliary Components. J Mol Biol.

Murdoch JN, and Copp AJ. 2010. The relationship between sonic Hedgehog signaling, cilia, and neural tube defects. Birth Defects Res A Clin Mol Teratol 88(8):633-652.

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(24):2169-2174.

Nie Y, Viola C, Bieniossek C, Trowitzsch S, Vijay-Achandran LS, Chaillet M, Garzoni F, and Berger I. 2009. Getting a grip on complexes. Curr Genomics 10(8):558-572.

Nigg EA, and Raff JW. 2009. Centrioles, centrosomes, and cilia in health and disease. Cell 139(4):663-678.

Norris DP, and Grimes DT. 2012. Mouse models of ciliopathies: the state of the art. Dis Model Mech 5(3):299-312.

O'Connell KF, Caron C, Kopish KR, Hurd DD, Kemphues KJ, Li Y, and White JG. 2001. The C. elegans zyg-1 gene encodes a regulator of centrosome duplication with distinct maternal and paternal roles in the embryo. Cell 105(4):547-558.

Oh EC, and Katsanis N. 2012. Cilia in vertebrate development and disease. Development 139(3):443-448.

Omori Y, Zhao C, Saras A, Mukhopadhyay S, Kim W, Furukawa T, Sengupta P, Veraksa A, and Malicki J. 2008. Elipsa is an early determinant of ciliogenesis that links the IFT particle to membrane-associated small GTPase Rab8. Nat Cell Biol 10(4):437-444.

Palomares LA, Estrada-Mondaca S, and Ramirez OT. 2004. Production of recombinant proteins: challenges and solutions. Methods Mol Biol 267:15-52.

Pan J, Wang Q, and Snell WJ. 2005. Cilium-generated signaling and cilia-related disorders. Lab Invest 85(4):452-463.

Pazour GJ, Baker SA, Deane JA, Cole DG, Dickert BL, Rosenbaum JL, Witman GB, and Besharse JC. 2002. The intraflagellar transport protein, IFT88, is essential for vertebrate photoreceptor assembly and maintenance. J Cell Biol 157(1):103-113.

Pazour GJ, Dickert BL, Vucica Y, Seeley ES, Rosenbaum JL, Witman GB, and Cole DG. 2000. Chlamydomonas IFT88 and its mouse homologue, polycystic kidney disease gene tg737, are required for assembly of cilia and flagella. J Cell Biol 151(3):709-718.

Pazour GJ, Dickert BL, and Witman GB. 1999. The DHC1b (DHC2) isoform of cytoplasmic dynein is required for flagellar assembly. J Cell Biol 144(3):473-481.

Pedersen LB, and Rosenbaum JL. 2008. Intraflagellar transport (IFT) role in ciliary assembly, resorption and signalling. Curr Top Dev Biol 85:23-61.

Peel N, Stevens NR, Basto R, and Raff JW. 2007. Overexpressing centriole-replication proteins in vivo induces centriole overduplication and de novo formation. Curr Biol 17(10):834-843.

Pelletier L, O'Toole E, Schwager A, Hyman AA, and Muller-Reichert T. 2006. Centriole assembly in Caenorhabditis elegans. Nature 444(7119):619-623.

Pelletier L, Ozlu N, Hannak E, Cowan C, Habermann B, Ruer M, Muller-Reichert T, and Hyman AA. 2004. The Caenorhabditis elegans centrosomal protein SPD-2 is required for both pericentriolar material recruitment and centriole duplication. Curr Biol 14(10):863-873.

Peters N, Perez DE, Song MH, Liu Y, Muller-Reichert T, Caron C, Kemphues KJ, and O'Connell KF. 2010. Control of mitotic and meiotic centriole duplication by the Plk4-related kinase ZYG-1. J Cell Sci 123(Pt 5):795-805.

Pigino G, Geimer S, Lanzavecchia S, Paccagnini E, Cantele F, Diener DR, Rosenbaum JL, and Lupetti P. 2009. Electron-tomographic analysis of intraflagellar transport particle trains in situ. J Cell Biol 187(1):135-148.

Piperno G, Siuda E, Henderson S, Segil M, Vaananen H, and Sassaroli M. 1998. Distinct mutants of retrograde intraflagellar transport (IFT) share similar morphological and molecular defects. J Cell Biol 143(6):1591-1601.

Porter ME, Bower R, Knott JA, Byrd P, and Dentler W. 1999. Cytoplasmic dynein heavy chain 1b is required for flagellar assembly in Chlamydomonas. Mol Biol Cell 10(3):693-712.

Preble AM, Giddings TM, Jr., and Dutcher SK. 2000. Basal bodies and centrioles: their function and structure. Curr Top Dev Biol 49:207-233.

Qin H, Rosenbaum JL, and Barr MM. 2001. An autosomal recessive polycystic kidney disease gene homolog is involved in intraflagellar transport in C. elegans ciliated sensory neurons. Curr Biol 11(6):457-461.

Qin J, Lin Y, Norman RX, Ko HW, and Eggenschwiler JT. 2011. Intraflagellar transport protein 122 antagonizes Sonic Hedgehog signaling and controls ciliary localization of pathway components. Proc Natl Acad Sci U S A 108(4):1456-1461.

Rachel RA, May-Simera HL, Veleri S, Gotoh N, Choi BY, Murga-Zamalloa C, McIntyre JC, Marek J, Lopez I, Hackett AN et al. . 2012. Combining Cep290 and Mkks ciliopathy alleles in mice rescues sensory defects and restores ciliogenesis. J Clin Invest 122(4):1233-1245.

Rieder CL, Faruki S, and Khodjakov A. 2001. The centrosome in vertebrates: more than a microtubule-organizing center. Trends Cell Biol 11(10):413-419.

Rodrigues-Martins A, Bettencourt-Dias M, Riparbelli M, Ferreira C, Ferreira I, Callaini G, and Glover DM. 2007. DSAS-6 organizes a tube-like centriole precursor, and its absence suggests modularity in centriole assembly. Curr Biol 17(17):1465-1472.

Rosenbaum JL, and Witman GB. 2002. Intraflagellar transport. Nat Rev Mol Cell Biol 3(11):813-825.

Salisbury JL. 2003a. Centrosome size is controlled by centriolar SAS-4. Trends Cell Biol 13(7):340-343.

Salisbury JL. 2003b. Centrosomes: coiled coils organize the cell center. Curr Biol 13(3):R88-90.

Sattar S, and Gleeson JG. 2011. The ciliopathies in neuronal development: a clinical approach to investigation of Joubert syndrome and Joubert syndrome-related disorders. Dev Med Child Neurol 53(9):793-798.

Saudou F. 2012. A "so cilia" network: cilia proteins start "social" networking. J Clin Invest 122(4):1198-1201.

Scheufler C, Brinker A, Bourenkov G, Pegoraro S, Moroder L, Bartunik H, Hartl FU, and Moarefi I. 2000. Structure of TPR domain-peptide complexes: critical elements in the assembly of the Hsp70-Hsp90 multichaperone machine. Cell 101(2):199-210.

Schmidt TI, Kleylein-Sohn J, Westendorf J, Le Clech M, Lavoie SB, Stierhof YD, and Nigg EA. 2009. Control of centriole length by CPAP and CP110. Curr Biol 19(12):1005-1011.

Sibanda BL, Critchlow SE, Begun J, Pei XY, Jackson SP, Blundell TL, and Pellegrini L. 2001. Crystal structure of an Xrcc4-DNA ligase IV complex. Nat Struct Biol 8(12):1015-1019.

Singla V, Romaguera-Ros M, Garcia-Verdugo JM, and Reiter JF. 2010. Ofd1, a human disease gene, regulates the length and distal structure of centrioles. Dev Cell 18(3):410-424.

Smith TF, Gaitatzes C, Saxena K, and Neer EJ. 1999. The WD repeat: a common architecture for diverse functions. Trends Biochem Sci 24(5):181-185.

Song MH, Liu Y, Anderson DE, Jahng WJ, and O'Connell KF. 2011. Protein phosphatase 2A-SUR-6/B55 regulates centriole duplication in C. elegans by controlling the levels of centriole assembly factors. Dev Cell 20(4):563-571.

Stevens NR, Dobbelaere J, Brunk K, Franz A, and Raff JW. 2010a. Drosophila Ana2 is a conserved centriole duplication factor. J Cell Biol 188(3):313-323.

Stevens NR, Roque H, and Raff JW. 2010b. DSas-6 and Ana2 coassemble into tubules to promote centriole duplication and engagement. Dev Cell 19(6):913-919.

Tabin CJ, and Vogan KJ. 2003. A two-cilia model for vertebrate left-right axis specification. Genes Dev 17(1):1-6.

Taschner M, Bhogaraju S, and Lorentzen E. 2012. Architecture and function of IFT complex proteins in ciliogenesis. Differentiation 83(2):S12-22.

Taschner M, Bhogaraju S, Vetter M, Morawetz M, and Lorentzen E. 2011. Biochemical mapping of interactions within the intraflagellar transport (IFT) B core complex: IFT52 binds directly to four other IFT-B subunits. J Biol Chem 286(30):26344-26352.

Tang CJ, Fu RH, Wu KS, Hsu WB, and Tang TK. 2009. CPAP is a cell-cycle regulated protein that controls centriole length. Nat Cell Biol 11(7):825-831.

Tang CJ, Lin SY, Hsu WB, Lin YN, Wu CT, Lin YC, Chang CW, Wu KS, and Tang TK. 2011. The human microcephaly protein STIL interacts with CPAP and is required for procentriole formation. EMBO J 30(23):4790-4804.

Terwilliger TC, Adams PD, Read RJ, McCoy AJ, Moriarty NW, Grosse-Kunstleve RW, Afonine PV, Zwart PH, and Hung LW. 2009. Decision-making in structure solution using Bayesian estimates of map quality: the PHENIX AutoSol wizard. Acta Crystallogr D Biol Crystallogr 65(Pt 6):582-601.

Tran PV, Haycraft CJ, Besschetnova TY, Turbe-Doan A, Stottmann RW, Herron BJ, Chesebro AL, Qiu H, Scherz PJ, Shah JV et al. . 2008. THM1 negatively modulates mouse sonic hedgehog signal transduction and affects retrograde intraflagellar transport in cilia. Nat Genet 40(4):403-410.

Tsao CC, and Gorovsky MA. 2008a. Different effects of Tetrahymena IFT172 domains on anterograde and retrograde intraflagellar transport. Mol Biol Cell 19(4):1450-1461.

Tsao CC, and Gorovsky MA. 2008b. Tetrahymena IFT122A is not essential for cilia assembly but plays a role in returning IFT proteins from the ciliary tip to the cell body. J Cell Sci 121(Pt 4):428-436.

Tsou MF, Wang WJ, George KA, Uryu K, Stearns T, and Jallepalli PV. 2009. Polo kinase and separase regulate the mitotic licensing of centriole duplication in human cells. Dev Cell 17(3):344-354.

Uhlmann F, Wernic D, Poupart MA, Koonin EV, and Nasmyth K. 2000. Cleavage of cohesin by the CD clan protease separin triggers anaphase in yeast. Cell 103(3):375-386.

van Breugel M, Hirono M, Andreeva A, Yanagisawa HA, Yamaguchi S, Nakazawa Y, Morgner N, Petrovich M, Ebong IO, Robinson CV et al. . 2011. Structures of SAS-6 suggest its organization in centrioles. Science 331(6021):1196-1199.

Vulprecht J, David A, Tibelius A, Castiel A, Konotop G, Liu F, Bestvater F, Raab MS, Zentgraf H, Izraeli S et al. . 2012. STIL is required for centriole duplication in human cells. J Cell Sci.

Wang Z, Fan ZC, Williamson SM, and Qin H. 2009. Intraflagellar transport (IFT) protein IFT25 is a phosphoprotein component of IFT complex B and physically interacts with IFT27 in Chlamydomonas. PLoS One 4(5):e5384.

Williamson SM, Silva DA, Richey E, and Qin H. 2011. Probing the role of IFT particle complex A and B in flagellar entry and exit of IFT-dynein in Chlamydomonas. Protoplasma.

Yu L, Gaitatzes C, Neer E, and Smith TF. 2000. Thirty-plus functional families from a single motif. Protein Sci 9(12):2470-2476.

Zaghloul NA, and Katsanis N. 2009. Mechanistic insights into Bardet-Biedl syndrome, a model ciliopathy. J Clin Invest 119(3):428-437.

Zhu F, Lawo S, Bird A, Pinchev D, Ralph A, Richter C, Muller-Reichert T, Kittler R, Hyman AA, and Pelletier L. 2008. The mammalian SPD-2 ortholog Cep192 regulates centrosome biogenesis. Curr Biol 18(2):136-141.

9. Appendix: *CURRICULUM VITAE*

Renping Qiao

Dr. Bohrgasse 9/3

1030 Vienna

Autria

Tel: +43 1 4277 61626

Email: renping.qiao@univie.ac.at

Research Interest

My main interest is to understand mechanism of protein/protein interactions involved in conserved cell signaling pathway and macromolecular complex assembly, mostly by X-Ray crystallography but also using other biophysical and biochemical techniques such as electron microscopy (EM), isothermal calorimetry (ITC) and immunoprecipitation.

Academic Background

2009-present Max F. Perutz Laboratories, Vienna Biocenter, Austria

Ph.D student in the laboratory of Dr. Gang Dong

2008-2009 CeMM-Center for Molecular Medicine of the Austrian Academy of Science, Vienna, Austria

Guest student in the laboratory of Dr. Denise P. Barlow

2005-2007 Medicinal Chemistry, Peking University, Beijing, P. R. China

Master of science, Thesis Advisor: Prof. Dr. Zhengjun Yang

2001-2004 Pharmaceutical Science, Peking University, Beijing, P. R. China

Bachelor of science

Research Experience

2009-present Structural Characterization of SAS-5 and SAS-6 in Centriole Assembly

Advisor: Dr. Gang Dong, Max F. Perutz Laboratories, Vienna, Austria

2008-2009 Analysis of DMRs and Macro ncRNAs in imprinted gene clusters

Advisor: Dr. Denise P. Barlow, CeMM-Center for Molecular Medicine of the Austrian Academy of Science, Vienna, Austria

2005-2007 Studies on Amino-isonucleoside Modified siRNAs: Stability and Silencing Activity

Advisor: Prof. Dr. Zhenjun Yang, Peking University, Beijing, P. R. China

2004-2005 One-step synthesis of novel tricyclic isomeric azidonucleosides

Advisor: Prof. Dr. Zhenjun Yang, Peking University, Beijing, P. R. China

Honors and Awards

2012 Invited speaker in 2012 the Annual Meeting of the American Crystallographic Association, Boston, US

2012 Travel funding from the IUCr Inter-Regional Bursary Fund

2010 Protein Production Platform (P-CUBE) training grant for the MultiBac-EMBL Platform
    in Grenoble, France

2007 Outstanding graduate, Triple-A student of Peking University, Peking University

2006 Excellent Medical Student First Order Scholarship, Triple-A student of Peking University, Peking University

2005 Excellent Medical Student Secondary Order Scholarship, Peking University

2004 Excellent Medical Student 'Wu-Si' Scholarship, Peking University

2003 Excellent Medical Student Superfine Scholarship, Triple-A student of Peking University, Peking University

2002 Triple-A student of Peking University, Peking University

Publications

Renping Qiao and Gang Dong. The SAS-6 coiled coil structure and its specific interaction with SAS-5 suggest a mutual regulation of the two proteins in centriole assembly. *Under revision.* 2012

Huang Huang, **Renping Qiao**, Tong Zhang, Youxian Li, Junmei Hong, Fangfang Lai, Zhenjun Yang, Lihe Zhang, Quan Du and Zicai Liang. Profiling of mismatch discrimination in RNAi enabled rational design of allele-specific siRNAs. *Nucleic Acids Res.,* 2009, 37(22), 7560-9.

Junmei Hong, Na Wei, Alistair Chalk, Jue Wang, Yutong Song, Fan Yi, **Renping Qiao** and Quan Du. Focusing on RISC assembly in mammalian cells. *Biochem Biophys Res Commun*., 2008, 368(3), 703-8.

Zong-Sheng Li, **Renping Qiao (co-first author)**, Quan Du, Zhen-Jun Yang, Liang-Ren Zhang, Pei-Zhuo Zhang, Zi-Cai Liang, Li-He Zhang. Studies on Amino-isonucleoside Modified siRNAs: Stability and Silencing Activity. *Bioconjugate Chemistry*, 2007, 18(4), 1017-1024.

Zong-Sheng Li, **Renping Qiao**, Zhen-Jun Yang, Liang-Ren Zhang and Li-He Zhang. One-step synthesis of novel tricyclic isomeric azidonucleosides. *Tetrahedron: Asymmetry*, 2006, 17, 1056–1061.

Academic Activities

2012 the Annual Meeting of the American Crystallographic Association, Boston, US

2012 LRZ Life-Science Molecular Modeling Workshop, Munich, Germany

2011 Advances in Structural Biology Computing Symposium, online workshop

2009 Protein Production Platform (P-CUBE) project, Grenoble, France

2008 13th Annual Meeting of the RNA Society, Berlin, Germany

2007 the 5th National Conference on Chemical Biology, Kunming, China

2006 ICISTS-KAIST 2006, International Conference for the Integration of Science and Technology into Society, Seoul, Korea

2005 The Lee Shiu Socio-Economic Project 2005, Hong Kong