Building a ninefold symmetrical barrel: structural dissections of centriole assembly

Gang Dong

Max F. Perutz Laboratories, Medical University of Vienna, Vienna 1030, Austria

Centrioles are short microtubule-based organelles with a conserved ninefold symmetry. They are essential for both centrosome formation and cilium biogenesis in most eukaryotes. A core set of five centriolar proteins has been identified and their sequential recruitment to procentrioles has been established. However, structures at atomic resolution for most of the centriolar components were scarce, and the underlying molecular mechanisms for centriole assembly had been a mystery—until recently. In this review, I briefly summarize recent advancements in high-resolution structural characterization of the core centriolar components and discuss perspectives in the field.

1. Introduction

Centrioles are cylindrical cellular structures present in almost all eukaryotic lineages. They are 0.1–0.5 \( \mu \text{m} \) long and 0.1–0.2 \( \mu \text{m} \) in diameter, and are usually composed of nine microtubule triplets at their outer wall [1,2]. Exceptions to this organization are found in \textit{Drosophila} embryos (nine doublet microtubules), and in \textit{Caenorhabditis elegans} sperm cells and early embryos (nine singlets). Centrioles usually exist in pairs, oriented perpendicularly to one another, and are surrounded by the pericentriolar material (PCM) to form the centrosome [3,4]. The PCM has a layered structure made of fibres and matrix proteins, and is responsible for microtubule nucleation and anchoring [5–8]. As such, the centrosome is the major microtubule organizing centre of animal cells. Centrosome duplication is controlled by centriole replication, which in most animal cells is tightly coordinated with the mitotic cycle. Miscoordination of the two processes often leads to alterations in centrosome number and/or structure that result in chromosome missegregation and severe consequences such as cancerous growth of the cell [9–12]. Besides nucleating centrosomes, centrioles also play a pivotal role in the biogenesis and operation of the cilium, an antenna-like structure projecting out from the cell surface [13]. Cilia are present on virtually every human cell type, and defects in their assembly or function lead to a plethora of human disorders, which are collectively called the ciliopathies [14,15].

Despite cell biologists’ great interest in the centriole and extensive cellular and microscopic studies on this organelle for several decades, the molecular composition of the centriole was not fully determined until recently. Because of its favourable features for laboratory study and the availability of powerful genetic and genomic tools, the nematode \textit{C. elegans} has become an important model organism in studying centrosome duplication and centriole assembly. Studies in several research groups have uncovered five core \textit{C. elegans} centriolar proteins: the polo-like kinase ZYG-1 [16], and the four coiled-coil-containing proteins, SPD-2 [17,18], SAS-4 [19,20], SAS-5 [21] and SAS-6 [22,23]. All of these proteins contain both structured domains and disordered segments and are conserved in all ciliated cells.

Centriole duplication is a multistep process, with each of the five centriolar proteins being recruited in a hierarchical manner [24,25]. In \textit{C. elegans}, centriole assembly is initiated by the recruitment of SPD-2 to the proximal side of the
mother centriole. The kinase ZYG-1 is then incorporated into the SPD-2 scaffold, and is required for the subsequent recruitment of the SAS-5/SAS-6 complex. SAS-4 is then recruited to add an outer tube around the central tube formed by SAS-5 and SAS-6. Finally, nine singlet microtubules are assembled onto the outer tube to generate a daughter centriole (figure 1).

In other organisms, including *Drosophila* and humans, additional essential proteins besides the five core components are also recruited during centriole assembly (figure 1).

Recent years have seen a surge of structural dissections of the centriole, with an increasing number of high-resolution structures reported for most of its core components. Together, these have greatly advanced our understanding of centriole duplication at the atomic level. Nevertheless, a full molecular view of centriolar assembly is a mission that is not yet completed. This review briefly summarizes recent progress in the structural characterization of centriole assembly, which promises a challenging but rewarding future.

2. SPD-2/Cep192 and Asterless/Cep152

Centriole duplication in animals is initiated by recruiting the centriolar receptors SPD-2 and/or Asterless to the vicinity of the mother centriole. These receptors then recruit polo-like kinase 4 (Plk4) and ZYG-1, depending on which organism is under consideration (figure 2a). SPD-2 is a coiled-coil-containing protein localizing to both the centriole and the PCM in *C. elegans* [17,18]. It has an N-terminal acidic region that binds to the cryptic polo-box (CPB) domain of ZYG-1 (figure 2b) [26]. The *Drosophila* orthologue of SPD-2 lacks the acidic region for Plk4 binding, and thus is dispensable for centriole duplication [27,28] (figure 2b). The mammalian SPD-2 orthologue Cep192 also bears an acidic region that interacts with the Plk4 CPB domain, but it is located a bit further downstream in the sequence (figure 2b).

Asterless was originally identified in *Drosophila* as a protein essential for aster formation during male meiosis [29]. The protein was later revealed to be a constitutive centriolar component involved in the initiation of centriole duplication [30–33]. Cep152, the mammalian orthologue of Asterless, is also involved in both centriole duplication and PCM recruitment [34–36]. Both *Drosophila* Asterless and human Cep152 bind to the CPB domain of Plk4 via their N-terminal acidic region [32,35]. Recent studies demonstrated that Plk4 additionally binds to the C-terminal region of Asterless, which serves to stabilize Plk4 during mitosis [37].

Although it is now well established how SPD-2 and Asterless interact with their kinase partners ZYG-1 and Plk4, they have been structurally less characterized. Bioinformatics analysis suggests that all SPD-2/Cep192 orthologues contain a mostly unstructured N-terminal region with a short coiled-coil domain and a long C-terminal β-strand-rich region (figure 2b). By contrast, all Asterless/Cep152 orthologues contain mostly helical structures that are predicted to form coiled-coils, with no β-strand content (figure 2c). *Caenorhabditis elegans* and mammalian SPD-2, and *Drosophila* and mammalian Asterless, also contain a highly acidic region towards their N-termini, which interacts directly and tightly with their kinase partners [26]. This tight interaction creates something of a puzzle however, as the receptors and the kinases do not seem to form a robust complex in the cell until the receptors are targeted to mother centrioles. Recent studies by multiple imaging/biophysical techniques have shown that *C. elegans* SPD-2 is monomeric in the cytoplasm [38]. It was postulated that the acidic region for kinase binding might be concealed via its interaction with another (probably basic) part of the same protein; receptors targeted to the nascent
centriole might then release such auto-inhibition to facilitate interaction with the downstream kinases [26].

Another unresolved issue about the centriolar receptors (SPD-2 and Asterless) in centriole assembly is how they themselves are recruited to the mother centriole. One possibility is that Plk1, another critical polo-like kinase linking centriole biogenesis to cell-cycle progression [39–41], might phosphorylate SPD-2/Asterless to promote their targeting and the initiation of centriole assembly. In human cells, both SPD-2 and Asterless bind Plk4 but in a mutually exclusive manner to temporally and spatially regulate Plk4 recruitment. To determine the high-resolution structures of Asterless and Cep152 might prove challenging because of their high coiled-coil content (figure 2c). Reconstitution of a complex of the coiled-coil with an interaction partner might help to make its structural studies amenable. The MitoCheck consortium has reported that at least nine other proteins interact with Cep152 in mammalian cells [43].

3. Plk4/ZYG-1

In metazoans, the serine/threonine kinase Plk4 controls daughter centriole assembly and couples centriole duplication with cell-cycle progression [44,45]. As the master regulators of centriole assembly, Plk4 and ZYG-1 are essential for the recruitment of downstream proteins to the procentriole [24,25,46]. In *C. elegans*, ZYG-1 directly binds SAS-6 in order to recruit it and its interaction partner SAS-5 to the mother centriole [47]. In *Drosophila* and mammals, Plk4 phosphorylates Ana2/STIL to trigger the direct interaction between Ana2/STIL and SAS-6, which is essential for their recruitment during centriole assembly [48–51]. Plk4 is the most divergent member of the polo-like kinase family, and *C. elegans* ZYG-1 is one of the most divergent of Plk4 orthologues [52]. Despite the high divergence in their primary
sequences, ZYG-1 and Plk4 share a similar structural arrangement consisting of three structural domains: the N-terminal kinase domain, the central CPB domain and the C-terminal single polo-box (PB3) domain (figure 3a).

The kinase domain of Plk4 adopts a similar conformation to that of Plk1, with the nucleotide binding site located in the cleft between two lobe-like modules (figure 3b). The central and C-terminal regions of Plk4 and ZYG-1 are the least similar in terms of protein sequence comparison. It has been debated whether ZYG-1 is a bona fide Plk4 or a structurally distinct functional orthologue [44]. To address this, it was critical to compare their respective CPB and PB3 domains to find out how similar they really are in three-dimensional structure.

Initially, a crystal structure of the Drosophila Plk4 CPB was determined, based on which a side-by-side dimeric conformation was proposed [53]. However, this orientation of the Plk4 CPBs was later found to be a crystallographic artefact [26,54]. Other more recent structural studies on the ZYG-1 and Plk4 CPBs revealed that they in fact form a similar Z-shaped dimer, with a basic patch across the dimeric interface for binding by SPD-2 and/or Asterless [26,55]. Interestingly, despite the high similarity in their overall conformation, there are subtle structural variations in the CPBs of ZYG-1 and Plk4 that confer their selective binding to the respective receptors [26]. The CPB of Drosophila Plk4 bears a flat and broad basic patch (figure 3c), whereas the basic patch on the ZYG-1 CPB is narrow and elongated because of the expanded angle between the arms (figure 3d). A crystal structure of the human Plk4 CPB revealed that it is very similar to that of the Drosophila Plk4 CPB [55]. The structure of the human Plk4 CPB in complex with either SPD-2 (figure 3e) or Asterless (figure 3f) showed that the two receptors are bound with distinct modes, suggesting an intricate spatio-temporal regulation of Plk4 recruitment [55].

A crystal structure of the C-terminal PB3 of mouse Plk4 showed that it forms a dimeric structure distinct from the canonical polo-box (figure 3g) [56]. Recent studies suggested that the PB3 of human Plk4 plays an important regulatory role by relieving auto-inhibition of the kinase domain [57]. Similarly, the C-terminal PB3 in ZYG-1 has also been shown to play a regulatory role in its function. ZYG-1 deletion mutations lacking its PB3 make the kinase behave differently in mitosis and meiosis [58]. It was proposed that the ZYG-1 C-terminus may possess two distinct regulatory roles: one for inhibiting the kinase activity and the other for differential regulation of ZYG-1 recruitment in mitosis and meiosis [58]. However, the underlying mechanism for how the PB3 regulates ZYG-1 activity remains elusive. A crystal structure of the mouse Plk4 PB3 dimer revealed a deep interfacial pocket between the two molecules, which was proposed to be a potential ligand-binding site [56]. However, no ligands binding to this pocket of the Plk4 PB3 have yet been identified. Recent studies further demonstrated that the human Plk4 PB3 does not dimerize [57]. Therefore, it remains unclear whether...
the PB3 of Plk4 in different organisms indeed all forms a dimer and, if so, whether the ZYG-1 PB3 shares a similar conformation to it. Determination of high-resolution structures for the PB3 domains of ZYG-1 and various Plk4s would also help investigate how PB3 mechanistically regulates the function of Plk4 and ZYG-1 in vivo.

4. SAS-5/Ana2/STIL and SAS-6

SAS-5 and SAS-6 are co-dependent for their centriolar localization, which is essential for the subsequent recruitment of SAS-4 [23]. SAS-5 is functionally orthologous to Drosophila Ana2 and vertebrate STIL, with a 90-residue STIL/Ana2 (STAN) motif at the C-terminus [59]. Reflecting evolutionary history, this motif is more conserved among vertebrates and flies (greater than 30% identity) than with the corresponding region in C. elegans SAS-5 (12% identity). Likewise, the binding modes of Ana2 and SAS-5 to their respective partners seem to be different. While the binding of SAS-5 to SAS-6 is mediated via synergetic hydrophobic and electrostatic interactions between the C-terminal 15 residues of SAS-5 (which is predicted to form a short α-helix) and a central region of the SAS-6 coiled-coil [60], such an interaction seems to be absent in Drosophila and mammals. This is despite the fact that both Ana2 and STIL are also predicted to possess a short helix towards their C-termini and that two of the four residues at the C-terminus of SAS-5 for SAS-6 binding are also conserved in this helix of Ana2 and STIL. Instead, the interaction between Ana2/STIL and SAS-6 was recently revealed to be mediated by the Plk4-dependent phosphorylation of a number of conserved serine/threonine residues in the STAN motif of Ana2/STIL [48–51]. ZYG-1 was initially reported to phosphorylate SAS-6 at serine 123 residue [61]. However, later studies with mutations of 38 potential phosphorylation sites in C. elegans SAS-6 (including serine123) suggest that SAS-6 is unlikely to be the phosphorylation target of ZYG-1 [47]. It remains to be determined whether ZYG-1 phosphorylates SAS-5 in a similar manner to that of Plk4 on Ana2/STIL.

Studies using various biophysical methods suggested that the central region of SAS-5 forms a tetramer, which may function as a cross-linker to strengthen the higher-order structure formed by SAS-6 [62]. Detailed molecular mechanisms underlying SAS-5 oligomerization have been revealed very recently. Crystal structures show that SAS-5 contains two independent oligomeric domains: a coiled-coil region spanning residues 125–180 and an Implico domain consisting of residues 210–265 (figure 4a) [63]. While the Implico motif forms a stable dimer via the intertwining of the zigzagged triple-helix structures, the coiled-coil can exist as either a dimer or trimer depending on its concentration. Together, these properties allow SAS-5 to form a higher-order assembly that may create a seeding point for SAS-6 oligomerization [63]. In the meantime, a crystal structure of the central coiled-coil domain of Ana2 has been reported, which reveals that it forms a parallel, symmetrical 4-helix bundle (figure 4b) [64]. This tetrameric form, together with its N-terminal SAS-4 binding site and C-terminal phosphorylation-dependent SAS-6-binding STAN motif, was believed to actively contribute to centriole assembly by interconnecting neighbouring layers of cartwheels formed by SAS-6 [64]. While auto-assembly of both SAS-5 and Ana2 is essential for centriole biogenesis, their different oligomerization modes might reflect the morphological variations of the respective SAS-6 oligomers that have been observed (see below).

SAS-6 is one of the most thoroughly characterized centriolar proteins at a structural level. It is a modular protein with an N-terminal head domain followed by a long coiled-coil and an intrinsically disordered C-terminal tail (figure 5a). Numerous crystal structures of the head domain of SAS-6 from various organisms have been reported, which show a conserved intermolecular interaction between the head domains [65–68]. Furthermore, a crystal structure of the C. elegans SAS-6 coiled-coil region shows that two SAS-6 molecules form a parallel dimer via a strong interaction between their coiled-coils [60]. Together, these studies suggest that 18 SAS-6 molecules can form nine parallel homodimers that then assemble into a cartwheel-like structure. The cartwheel is stabilized through the interactions between the neighbouring head domains in the closed ring (figure 5b–d). This working model finally provides a molecular explanation for the ninefold cartwheel structure originally identified by electron microscopy [69].

While the cartwheel formed by SAS-6 is believed to be important for dictating the ninefold symmetry of centrioles, the interaction between the SAS-6 head domains is relatively weak, with a Kd of only 60–100 µM. This argues against the autonomous assembly of the cartwheel by SAS-6 alone. In the cell, this issue might be solved via the cross-linking action of SAS-5/Ana2/STIL as described above. Interestingly, it was shown recently that in human cells cartwheel assembly occurs in the proximal lumen of the mother centriole [70]. This assembly process is initiated by the recruitment of SAS-6 dimers to disengaged centrioles, and mediated by the interaction between the SAS-6 C-terminal tail and SAS-4 located at the luminal wall [70]. However, this seems to be species-specific and might be employed only by cycling cells with strictly controlled centriole number, as de novo assembly of centrioles has been observed in many other species [71–73].

To form a cylindrical centriole, the SAS-6-mediated cartwheel structure apparently has to be assembled longitudinally in the lumen of the centriole. One way to achieve this is by stacking multiple rings on top of one another, as observed in the flagellate Trichonympha (figure 5c, left) [74]. This assembly mechanism might also occur in other organisms [75,76]. The other means of SAS-6 assembly was found in C. elegans where SAS-6 forms a spiral-like structure, and two such spirals wind around each other to form the central tube of the centriole (figure 5c, middle) [65]. It is postulated that SAS-6 might also assemble in a third way, in which SAS-6 forms a single tight spiral formed by the slightly tilted hinge region between the head group and the coiled-coil (figure 5c, right) [77]. However, the single spiral model has not been seen in any real centrioles so far and its existence remains undetermined.

5. SAS-4/CPAP

SAS-4/CPAP is the last core centriolar protein to be recruited. It is located at the outer wall of the centriolar barrel and plays an essential role for subsequent recruitment of microtubules [24,25]. SAS-4/CPAP was originally identified in mammals by yeast two-hybrid screens [78]. Homologues of SAS-4/CPAP were also later identified in Drosophila and C. elegans [19,20,79]. All SAS-4/CPAP proteins share a similar overall
domain arrangement, with a mostly disordered N-terminus, a central coiled-coil and a globular C-terminal T complex protein 10 (TCP) domain (figure 6a). Crystal structures of the TCP domain of SAS-4 proteins from zebrafish and Drosophila showed that it forms a long β-meander consisting of about 20 β-strands, which interacts directly with a conserved proline-rich region in STIL/Ana2—the consensus sequence of this region is PR/C2/C2P/C2P (figure 6b) [80, 81]. Interestingly, the seemingly stable TCP structure does not contain a defined hydrophobic core or flanking globular domains on either end.

While the interaction mode between the SAS-4 TCP domain and STIL/Ana2 is conserved in Drosophila and vertebrates, the interaction between C. elegans SAS-4 and SAS-5 may have diverged significantly. For instance, although the TCP domain is required for C. elegans SAS-4 to interact with the N-terminal part of SAS-5, by itself it is not sufficient [80]. Furthermore, SAS-5 does not contain the consensus TCP-binding motif (PR××P×P) present in STIL/Ana2 [80].

The function of SAS-4/CPAP in recruiting microtubules during centriole assembly has been confirmed. However, the mechanism is still unclear. One hypothesis is that the TCP domain of Drosophila CPAP forms monomers in solution, but upon incorporation into procentrioles the TCP domain would bind to STIL towards the interior of the centriole whereas the N-terminal part of SAS-4/CPAP interacts with Asterless/Cep152 and microtubules at the outer wall [80]. Another is that the TCP β-sheet is capable of forming large assemblies via a head-to-tail interaction. This was observed in the crystal structure and was postulated to help CPAP to form continuous connections along the long axis of the centriolar outer wall. Such connections, together with the central coiled-coil of the protein, would then link the pinhead with the microtubules to stabilize the cartwheel stacks in the centriole [81]. Structural and functional analysis of Drosophila SAS-4 TCP domain by another group further confirmed the role of SAS-4 as a mediator to tether PCM complexes to the centriole to facilitate centrosome assembly.

Figure 4. Domain arrangements and crystal structures of SAS-5 and Ana2. (a) Domain organization and crystal structures of SAS-5. The central coiled-coil may form a dimer or trimer depending on the protein concentration. The Implico domain forms an intertwined homodimer. The C-terminal 15 residues of SAS-5 (shown in red) are predicted to form an α-helix that binds directly to a central region of the SAS-6 coiled-coil. (b) Domain organization of Drosophila Ana2 and crystal structure of its parallel, symmetrical 4-helix bundle. Plk4 phosphorylates four conserved serine residues in the STAN motif to promote the strong interaction between Ana2 and SAS-6.
6. Other proteins involved in centriole assembly

Besides the five core centriolar components, a number of other proteins participating in centriole duplication have recently been identified. Bld10/Cep135 is an approximately 170 kDa protein originally found in *Chlamydomonas* and essential for cartwheel formation [83]. It consists mostly of coiled-coils and was later found to be a major spoke-tip component, serving to bridge the cartwheel to the triplet microtubules [84]. A homologue of Bld10/Cep135 was subsequently identified in *Drosophila*, and was reported to play a role in the formation of both the central tube in mature primary spermatocytes [85] and the central pair of microtubules in the flagellar axoneme [86,87]. However, no Bld10/Cep135 homologue has been identified in nematodes.

The localization of Bld10/Cep135 to procentrioles depends on both Cep120 and SPICE1 [88]. Cep120 is an approximately 100 kDa protein participating in centriole assembly by asymmetrically localizing to the daughter centriole [89]. The protein contains an N-terminal C2 domain, a central region for SAS-4/CPAP binding and a C-terminal long coiled-coil for dimerization [90]. It was reported that Cep120 interacts with SPICE1, another large coiled-coil-containing protein, and the two proteins are co-dependent for their localization to procentrioles [88]. Cep120 and SPICE1 are believed to work cooperatively with SAS-4/CPAP to regulate centriole elongation [88,90].

Comparative genomics and proteomics studies of isolated basal bodies from *Chlamydomonas reinhardtii* have identified 18 proteins (Poc 1–Poc 18) proposed to be centriolar constituents [91]. Poc1 is one of the best-characterized Poc proteins. It contains an N-terminal WD40 domain and a C-terminal coiled-coil, and like Bld10/Cep135, is hypothesized to stabilize microtubules, and thus maintains centriole assembly and integrity [92–95].
In animal cells, a cap structure at the distal end of centrioles is formed to suppress centriole-templated ciliogenesis. The cap complex consists of CP110 and Cep97, two large proteins consisting mostly of coiled-coils [96]. Additionally, several other proteins have been identified to be more specifically involved in assembly of an appendage structure at the distal end of mother centrioles. These proteins include Ccdc41/Cep83, Cep89/Cep123, Sclt1, Fbf1 and Cep164 [97–99]. The distal appendage serves to anchor the mother centriole onto the plasma membrane during ciliogenesis.

7. Conclusion and perspectives

The recent progress in structural studies on centriolar components, with each of the five core components having been at least partially characterized, has greatly advanced our understanding of centriole assembly at the high-resolution level. However, further studies will be necessary to precisely determine their contributions, both physically and temporally, during centriole assembly, and to uncover their interplay with other newly identified constitutive or accessory centriolar proteins and each other. The field faces several challenges in this process including (i) to understand how centriole assembly is initiated, (ii) to uncover how centriole duplication is regulated by both kinases and phosphatases, (iii) to determine the high-resolution structures and the functional roles for those coiled-coil-containing centriolar proteins, and (iv) to precisely assign the relative positions and the interaction network of different centriolar proteins and their structural/functional roles.

Recent studies using various techniques have identified many new centriolar components. It is likely that more centriole-associated proteins will be reported in the future.
However, unlike the five core centriolar proteins, at least some of them are organism- or cell-specific. For example, although Bld10/Cep135 is crucial for cartwheel formation in green algae [83,84], Paramecium tetraurelia [100] and mammalian cells [46,101], its homologue in Drosophila is dispensable for cartwheel formation [85,87,102], and no homologue of Bld10/Cep135 seems to exist in nematodes. Another example is CP110 and Cep97, both of which exist only in animals, and are needed for centriole duplication in S-phase-arrested human cells to control centriole length [46,103–105]. Phylogenetic analysis suggests that both CP110 and Cep97 might have been recruited to the centriole-assemble pathway in animals during evolution to coordinate centriole assembly with ciliogenesis and/or cytokinesis [86]. Therefore, it is necessary to carefully examine which of these newly identified centriolar proteins are structural constituents of the centriole and which are accessory proteins, and what their structural and functional roles are in centriole assembly.

It is foreseeable that many more structural characterizations on centriole assembly will be reported in the coming years, which, together with continuous advancements in functional and in vivo studies, will provide a more refined view of one of the most important and fascinating cellular structures that has captivated cell biologists for over a century.

Note added in proof

Arquint et al. [106] just reported high-resolution structures of the PB3 of human Plk4, both alone and in complex with STIL. Their studies demonstrate that the human plk4 PB3, unlike the previously reported dimeric mouse Plk4 PB3 (PDB code: 1MBY), folds into a monomeric canonical polo-box and binds to the central coiled-coil of STIL via a novel polo-box—peptide interaction mode mimicking coiled-coil formation.

Competing interests. The author declares that there is no conflict of interest.

Funding. This research was supported by institutional support from the Max F. Perutz Laboratories and grant P23440-B20 from the Austrian Science Fund (FWF).

Acknowledgements. The author is grateful to Brooke Morriswood (University of Würzburg) for critical reading of the manuscript.

References


89. Mahjoub MR, Xie Z, Stearns T. 2010 Cep120 is asymmetrically localized to the daughter centriole and is essential for centriole duplication. J. Cell Biol. 191, 331 – 346. (doi:10.1083/jcb.201003009)


