Structure
Short Article

Structural Analysis of the G-Box Domain of the Microcephaly Protein CPAP Suggests a Role in Centriole Architecture

Georgios N. Hatzopoulos,[1] Michèle C. Erat,[1] Erin Cutts,[1] Kacper B. Rogala,[1] Leanne M. Slater,[1] Philip J. Stansfeld,[1] and Ioannis Vakonakis[1,*]
[1]Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, UK
*Correspondence: ioannis.vakonakis@bioch.ox.ac.uk
http://dx.doi.org/10.1016/j.str.2013.08.019

SUMMARY

Centrioles are evolutionarily conserved eukaryotic organelles composed of a protein scaffold surrounded by sets of microtubules organized with a 9-fold radial symmetry. CPAP, a centriolar protein essential for microtubule recruitment, features a C-terminal domain of unknown structure, the G-box. A missense mutation in the G-box reduces affinity for the centriolar shuttling protein STIL and causes primary microcephaly. Here, we characterize the molecular architecture of CPAP and determine the G-box structure alone and in complex with a STIL fragment. The G-box comprises a single elongated β sheet capable of forming supramolecular assemblies. Structural and biophysical studies highlight the conserved nature of the CPAP-STIL complex. We propose that CPAP acts as a horizontal "strut" that joins the centriolar scaffold with microtubules, whereas G-box domains form perpendicular connections.

INTRODUCTION

Centrioles are cylindrical organelles with characteristic 9-fold radial symmetry that are present in all eukaryotes, except higher plants (Gönczy, 2012). Centrioles function in two distinct roles, close to the cell membrane as basal bodies from which cilia and flagella emanate or adjacent to the nucleus, where a centriole pair and associated pericentriolar material comprise the centrosome. The latter functions as the microtubule-organizing center in animal cells, directing formation of the mitotic spindle during division and thereby ensuring correct chromosome segregation. Both roles of centrioles are crucial for fundamental cellular processes, including signaling, motility, development, and genome stability. Defects in centriole duplication are linked to multiple diseases, including ciliopathies, male sterility, and primary microcephaly (Nigg and Raff, 2009).

Evolutionarily conserved proteins necessary for centriole assembly have been identified across the eukaryotic taxon (Brito et al., 2012; Gönczy, 2012), and they include CPAP/SAS-4, STIL/Ana2/SAS-5, SAS-6, and Cep135/Bld10p. In most species, SAS-6 proteins form an initial assembly close to the parental centriole, the "cartwheel," with a ring-like hub and radially projecting spokes (Nakazawa et al., 2007). In vitro studies of the *Chlamydomonas reinhardtii* and *Danio rerio* SAS-6 orthologs (Kitagawa et al., 2011b; van Breugel et al., 2011) found that they self-assemble into rings with spokes composed of coiled coils. Cryo-tomographic studies of isolated *Trichonympha sp.* centrioles showed that multiple cartwheels can stack together with periodicity of 8 nm along the long centriolar axis, whereas the spokes of adjacent cartwheels merge at approximately midlength to form coiled-coil bundles every 17 nm (Guichard et al., 2012).

Following cartwheel assembly, CPAP/SAS-4 (also known as CENPJ) (Kohlmaier et al., 2009; Schmidt et al., 2009; Tang et al., 2009) localizes to the site of the growing centriole (Lawo et al., 2012; Mennella et al., 2012). CPAP and its orthologs are endowed with multiple functions: its N terminus interacts with pericentriolar proteins (Gopalakrishnan et al., 2011), including γ-tubulin (Hung et al., 2000), it harbors microtubule and tubulin association epitopes (Cormier et al., 2009; Hsu et al., 2008; Hung et al., 2004), and it is required for microtubule attachment to the initial centriolar scaffold (Pelletier et al., 2006). Interestingly, CPAP overexpression leads to abnormally long centrioles (Kohlmaier et al., 2009; Schmidt et al., 2009; Tang et al., 2009), whereas mutations in CPAP can cause autosomal recessive primary microcephaly, MCPH, a developmental disorder that leads to reduced brain size and associated mental retardation (Thornton and Woods, 2009). Other proteins whose variants lead to MCPH include STIL (Aplan et al., 1991) and Cep135 (Ohta et al., 2002). STIL is an essential centriolar component that interacts with CPAP (Kitagawa et al., 2011a; Tang et al., 2011; Vulprecht et al., 2012), is required for CPAP and SAS-6 (Vulprecht et al., 2012) centriolar localization, and shuttles rapidly between centrioles and the cytoplasm (Vulprecht et al., 2012). Cep135 interacts directly with SAS-6, CPAP, and microtubules (Carvalho-Santos et al., 2012; Lin et al., 2013), whereas Bld10p, the *C. reinhardtii* Cep135 ortholog, localizes to "pinhead" structures that bridge the SAS-6 cartwheel and centriolar microtubule triplets (Hiraki et al., 2007).

Despite the importance of CPAP in centriole assembly, we lack information about the structural features of this protein or the mechanisms that position it in the centriole. Here, we determine the crystallographic structure of the conserved C-terminal domain of CPAP, termed the G-box, alone or in complex with a STIL fragment. The surprising features of the G-box domain,

along with information on the molecular architecture of CPAP allow us to propose a specific role for this protein in the centriole ultrastructure.

RESULTS

Molecular Architecture of CPAP

Human CPAP (HsCPAP) is a 1338 amino acids protein with a glycine-rich C-terminal domain (G-box) (Figure 1A). Previous work showed that separate HsCPAP motifs bind tubulin heterodimers and destabilize microtubules (PN2-3) (Cormier et al., 2009; Hung et al., 2004) or bind and stabilize microtubules (A5N) (Hsu et al., 2008). PN2-3, A5N, and the G-box are the most conserved regions between HsCPAP and its functional orthologs in *D. rerio*, *Drosophila melanogaster*, and *Caenorhabditis elegans* (Gopalakrishnan et al., 2011; Kitagawa et al., 2011a; Tang et al., 2011). All four orthologs also feature a predicted ~150–200 residue coiled coil adjacent to the G-box; in contrast, the three N-terminal coiled coils of HsCPAP are not universally predicted, and they are presumed to be of low stability or unstructured as judged by their short length and biophysical properties (Cormier et al., 2009).

Recombinantly produced HsCPAP coiled-coil domain ($HsCPAP_{CC}$, residues 897–1,056) exhibits circular dichroism (CD) spectra with the double minima characteristic of α-helical proteins (Figure 1B) and a cooperative transition profile upon thermal denaturation (Figure 1C), evidence which support the prediction of an independently folded coiled coil. Size-exclusion chromatography-multiangle light scattering (SEC-MALS) showed that $HsCPAP_{CC}$ is dimeric in solution (Figure 1D). In order to evaluate the relative orientation of the monomeric subunits in the dimer, we performed crosslinking experiments. The sole cysteine in $HsCPAP_{CC}$ is one added during cloning at the protein C terminus; thus, it can form a disulfide with its dimer counterpart only if the coiled coil is parallel. As shown in Figure 1E, $HsCPAP_{cc}$ migrates as a dimer in SDS-PAGE under nonreducing conditions, thereby supporting a parallel orientation (Figure 1I). The length of this ~160 residue coiled coil, if linear, can be calculated as ~24 nm.

Furthermore, we produced the G-box domains of HsCPAP ($HsCPAP_{G\text{-}box}$) and *D. rerio* CPAP ($DrCPAP_{G\text{-}box}$); these are well conserved with 64.5% sequence identity, based on that we expect the two domains to have high structural similarity. Both G-box domains are monomeric under dilute conditions, as determined by SEC-MALS (Figure 1H; Figure S1C available online) and display CD spectra characteristic of β strand proteins (Figures 1F and S1A). However, $DrCPAP_{G\text{-}box}$ shows significantly more cooperativity upon thermal unfolding compared to $HsCPAP_{G\text{-}box}$ (Figures 1G and S1B, respectively). As higher unfolding cooperativity may reflect lower protein plasticity, this property recommended $DrCPAP_{G\text{-}box}$ for crystallographic studies.

Structure of the CPAP G-Box Domain

Next, we determined the structure of the G-box domain. $DrCPAP_{G\text{-}box}$ crystals formed readily and diffracted to 1.44 Å resolution (Table 1). The final model, refined to R/R_{free} of 15.4%/18.0%, consists of one molecule per asymmetric unit and accounts for DrCPAP residues 950–1,106 (Table 1). The first 7 and last 15 residues of $DrCPAP_{G\text{-}box}$ were not observed in the electron density, despite being present in the expressed construct as confirmed with mass spectrometry; therefore, they are presumed disordered in the crystal.

$DrCPAP_{G\text{-}box}$ consists of 17 antiparallel β strands that form a ~75 Å long and ~23 Å wide β sheet, with an ~168° of twist across its length (Figures 2A and 2B). Remarkably, all residues of this structure are solvent exposed; thus, it lacks a well-defined hydrophobic core. A search for similar structures in the Protein Data Bank shows that only OspA, which features a large exposed β sheet (Makabe et al., 2006), is structurally related. However, $DrCPAP_{G\text{-}box}$ does not include globular domains capping the β sheet, in contrast to OspA, yet it remains highly soluble and cooperatively folded (Figure 1G). Analysis of the model suggests that $DrCPAP_{G\text{-}box}$ is stabilized through a number of interstrand ionic interactions (Figure 2C), as well as through clusters of hydrophobic residues that span multiple β strands (Figure 2D). In addition, significant burial of hydrophobic surface area occurs between nonpolar side chains and the aliphatic part of long polar groups, such as arginine and lysine residues. To better understand $DrCPAP_{G\text{-}box}$ stability, we performed a 100 ns atomistic molecular dynamics simulation starting from the crystallographic structure. As shown in Figure S2, $DrCPAP_{G\text{-}box}$ remains folded during the simulation with no signs of β sheet unraveling. The domain length (Figure S2A), the number of main chain hydrogen bonds formed (Figure S2B), and the hydrophobic area buried (Figure S2C) remain essentially unaltered throughout the simulation trajectory (Figure S2D; Movie S1).

$DrCPAP_{G\text{-}box}$ domains associate through a head-to-tail interaction in the crystal along the open ends of β1 and β17, which involves formation of five hydrogen bonds and burial of ~420 $Å^2$ of surface area. This interaction creates a continuous β sheet with ~8 nm periodicity, reminiscent of amyloid fibrils (Figure 2E). To establish whether $DrCPAP_{G\text{-}box}$ populates such supramolecular structures in solution, we performed dynamic light scattering (DLS) measurements. As shown in Figures 2F and 2G, $DrCPAP_{G\text{-}box}$ forms assemblies of increasing hydrodynamic radius as a function of protein concentration. We conclude that $DrCPAP_{G\text{-}box}$ is inherently able to form large assemblies in solution, in a manner likely similar to the fibrils observed in the crystal.

Characterization of the CPAP-STIL Interaction

Previous studies using yeast-two-hybrid and pull-down assays had shown that a HsCPAP fragment spanning the coiled-coil and G-box domains interacts with the central region of human STIL (HsSTIL) (Tang et al., 2011). This interaction is attenuated in the presence of a missense mutation, E1235V, which maps on the G-box and can lead to MCPH (Thornton and Woods, 2009). Using fluorescence polarization, we sublocalized this interaction to $HsCPAP_{G\text{-}box}$ and a short HsSTIL fragment that spans residues 385–433 (Figure 3A). Whereas the wild-type proteins interact with a K_d of 0.55 ± 0.06 μM, $HsCPAP_{G\text{-}box}$ E1235V and E1235K, a more radical substitution that inverts the electrostatic charge, bind the same HsSTIL region with 9- and 75-fold lower affinity, respectively (Figure 3A). Sequence alignments of STIL orthologs, as well as the STIL relatives of *C. elegans* (SAS-5) and *D. melanogaster* (Ana2), showed the presence of a conserved proline-rich motif on HsSTIL residues 404–413 (Figure S3B). A HsSTIL fragment lacking this motif (residues 420–475) showed an ~25-fold decrease in $HsCPAP_{G\text{-}box}$ affinity

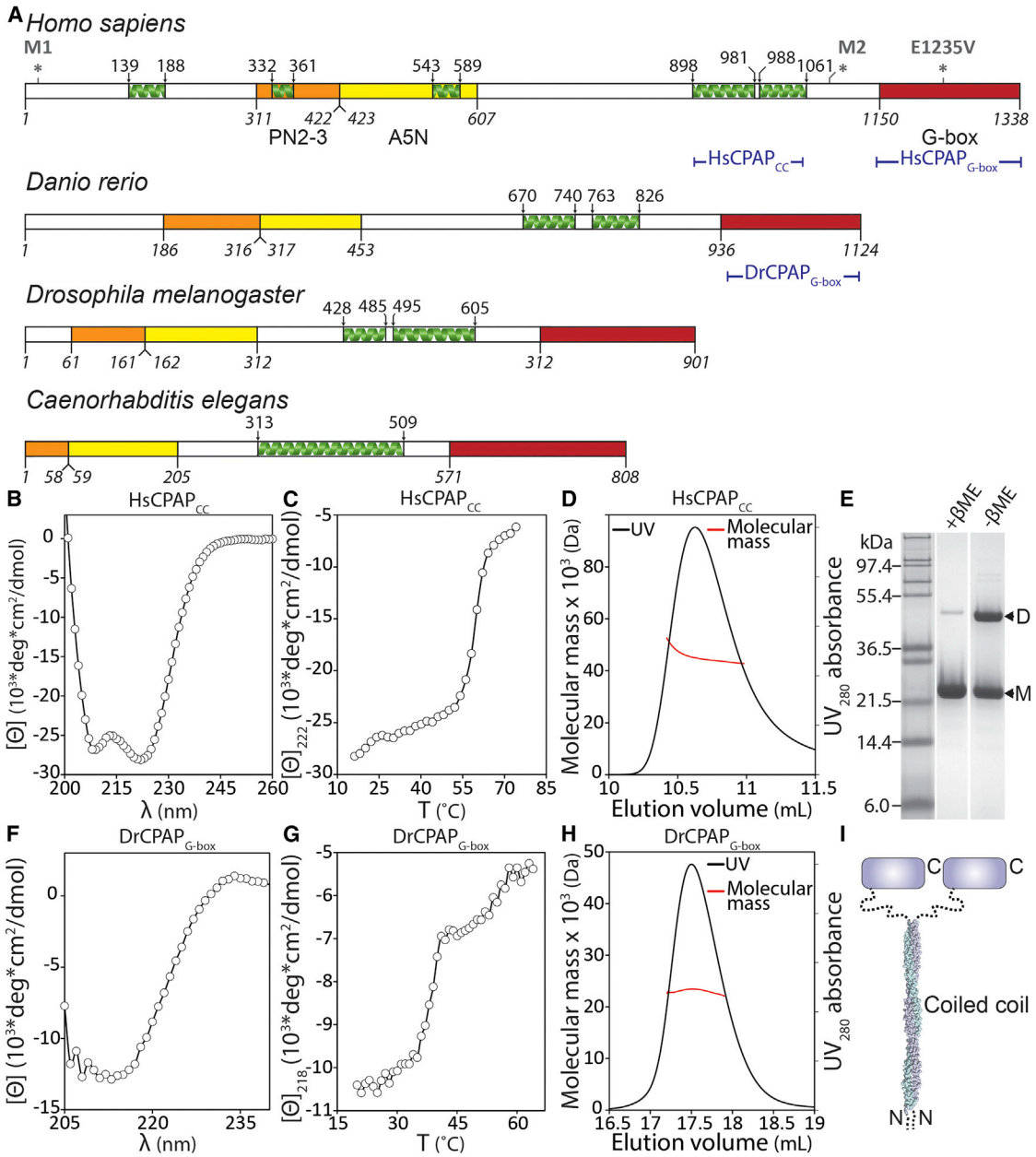

Figure 1. Molecular Architecture of HsCPAP

(A) Schematic representation of CPAP/SAS-4 from human, *D. rerio*, *D. melanogaster*, and *C. elegans*. Orange, yellow, and red boxes correspond to the protein tubulin binding (PN2-3), microtubule binding (A5N), and G-box domains, respectively. The predicted coiled-coil domains are indicated by green helices; the long coiled coil adjacent to the G-box features a short break in some species. Residue boundaries and the extent of our protein constructs are indicated. Asterisks in HsCPAP denote the location of MCPH causing mutations; M1 and M2 correspond to premature stop codons.

(B and C) CD spectrum of HsCPAP$_{CC}$ (B) and change in molar ellipticity (C) at 222 nm showing a cooperative transition upon thermal unfolding of the same construct are shown.

(D) SEC-MALS analysis of HsCPAP$_{CC}$. The UV absorbance profile from SEC (black line) is overlain with the MALS estimation of molecular mass (red line), which agrees well with that of a dimer (39.5 kDa).

(E) SDS-PAGE of HsCPAP$_{CC}$ under reducing (+βMe) and nonreducing (−βMe) conditions. Arrowheads point to protein bands corresponding to monomeric (M) and disulfide-linked dimeric (D) forms.

(F and G) CD spectra (F) and thermal unfolding profile (G) at 218 nm of DrCPAP$_{G-box}$ are shown.

(H) SEC-MALS analysis of DrCPAP$_{G-box}$ as in (D). The estimated molecular mass agrees well with a momoner (20.8 kDa).

(I) Molecular model of CPAP with a parallel coiled coil and a C-terminal domain connected by a flexible linker is shown.

See also Figure S1.

Table 1. Crystallographic Data Collection and Refinement Statistics

Protein	$DrCPAP_{G\text{-}box}$	$DrCPAP_{G\text{-}box}$ (Sm Derivative)	$DrCPAP_{G\text{-}box}$-DrSTIL Residues 398–450	$DrCPAP_{G\text{-}box}$-DrSTIL Residues 414–428
PDB ID code	4LD1	–	4LD3	4LZF
Space group	P1 2_1 1	P1 2_1 1	$P3_2$	C1 2_1 1
Unit cell (Å, °)	a = 41.62, b = 50.25, c = 60.34, β = 108.56	a = 41.74, b = 50.58, c = 60.29, β = 108.97	a = 79.97, b = 79.97, c = 52.27, γ = 120.00	a = 97.58, b = 38.39, c = 70.05, β = 126.68
Beamline	DLS/I02	DLS/I03	DLS/I02	DLS/I02
Wavelength (Å)	0.9795	1.595	0.9795	0.9795
Resolution range (Å)	39.47–1.44	57.02–1.80	28.87–2.44	48.42–1.72
High-resolution shell (Å)	1.49–1.44	1.85–1.80	2.50–2.44	1.76–1.72
R_{Merge}[a]	0.026 (0.504)	0.076 (0.571)	0.051 (0.784)	0.026 (0.656)
R_{Pim}[a]	0.024 (0.453)	0.032 (0.316)	0.029 (0.427)	0.023 (0.365)
Completeness[a] (%)	99.3 (99.9)	90.4 (51.2)	98.0 (97.7)	90.2 (54.2)[b]
Multiplicity[a]	3.7 (3.7)	12.4 (7.7)	5.1 (5.3)	2.9 (1.9)
I/σ(I)[a]	19.2 (2.3)	17.8 (2.9)	12.7 (1.8)	19.5 (2.1)
Phasing				
No. of Sm sites	–	2	–	–
Resolution	–	57.02–1.80	–	–
FOM initial[c]	–	0.21	–	–
FOM DM[d]	–	0.94	–	–
Refinement statistics				
R_{work} (reflections)	0.154 (42,575)	–	0.198 (11,691)	0.190 (19,069)
R_{free} (reflections)	0.180 (2,156)	–	0.214 (613)	0.231 (1,022)
Number of atoms				
Protein atoms[e]	2,688	–	1,362	1,367
Ligands	39	–	21	33
Water	204	–	3	99
Average B factors (Å^2)				
Protein atoms	33.8	–	113.4	40.8
Water	46.2	–	90.5	50.6
Rmsd from ideal values				
Bonds/angles (Å/°)	0.009/1.20	–	0.010/1.08	0.006/1.11
MolProbity statistics[f]				
Ramachandran favored (%)	99.35	–	98.80	98.77
Ramachandran disallowed (%)	0.0	–	0.0	0.0
Clashscore (percentile)	4.04 (95%)	–	3.69 (99%)	4.34 (96%)
MolProbity score (percentile)	1.19 (97%)	–	1.16 (100%)	1.21 (98%)

[a]Values in parentheses correspond to the high-resolution shell.
[b]Reflection data are 99.4% complete to a resolution of 1.99 Å.
[c]From PHASER (McCoy et al., 2007).
[d]From RESOLVE (Terwilliger, 2000).
[e]Atom count includes hydrogens in the P1 2_1 1 model.
[f]Chen et al., 2010.

(Figure 3B). We conclude that the conserved proline-rich motif of HsSTIL is important for binding to $HsCPAP_{G\text{-}box}$; furthermore, as this motif is present in the STIL relatives SAS-5 and Ana2, the case for functional orthology between these molecules is strengthened.

Structure of the $CPAP_{G\text{-}Box}$-STIL Complex

We determined the structure of the CPAP-STIL complex using *D. rerio* STIL (DrSTIL) fragments that span residues 398–450 or 414–428 (equivalent to HsSTIL residues 385–433 and 401–415, respectively). Crystals of the later peptide in complex with $DrCPAP_{G\text{-}box}$ diffracted to 1.72 Å resolution (Table 1), and they provided unambiguous electron density difference for DrSTIL following molecular replacement with $DrCPAP_{G\text{-}box}$ (Figure S3C). This density allowed us to identify specific peptide residues; however, the final model, refined to R/R_{free} of 19.0%/21.3%, shows just eight DrSTIL residues that make significant contact with $DrCPAP_{G\text{-}box}$. In contrast, crystals of $DrCPAP_{G\text{-}box}$ in complex

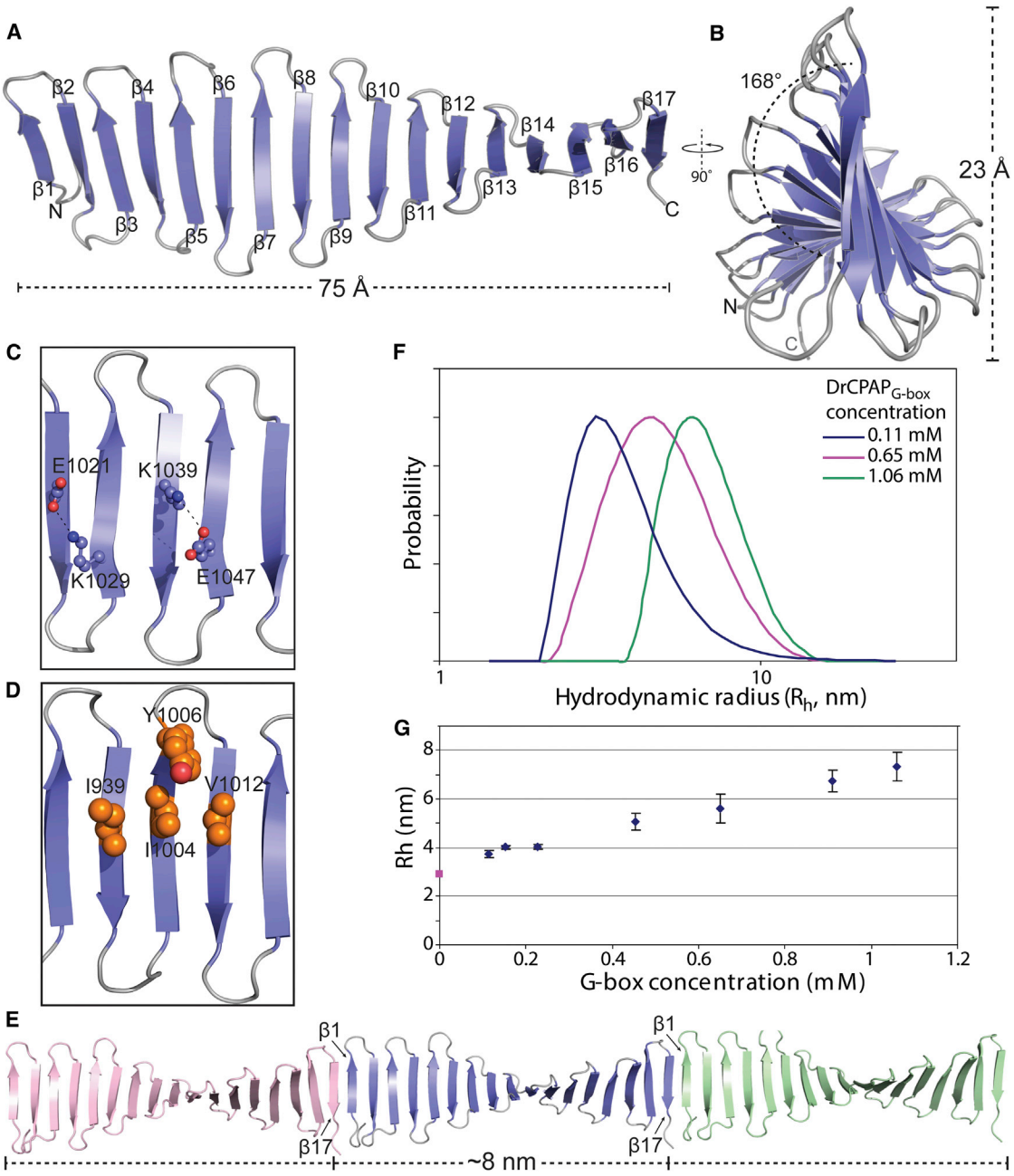

Figure 2. Structural Analysis of DrCPAP$_{G\text{-box}}$

(A and B) Shown is schematic representation of DrCPAP$_{G\text{-box}}$ in two perpendicular views. The secondary structure elements, protein dimensions, and the β sheet twist are shown.

(C and D) Electrostatic (C) and hydrophobic (D) interactions on the surface of the DrCPAP$_{G\text{-box}}$ β sheet are shown. The residues involved are shown in CPK and space-filling representation, respectively.

(E) Schematic representation of the DrCPAP$_{G\text{-box}}$ fibril. Three successive G-box domains are colored pink, blue, and green, with the β1 and β17 strands and the fibril periodicity shown.

(F) Distributions of hydrodynamic radii (R_h) of DrCPAP$_{G\text{-box}}$ assemblies in solution at three different concentrations as measured by DLS. The distributions are broad, with large assemblies populated even at the lowest concentration (blue line); their prevalence increases as a function of protein concentration.

(G) Mean R_h of DrCPAP$_{G\text{-box}}$ assemblies by DLS as a function of protein concentration. R_h at the infinite dilution point (pink square) was calculated from the DrCPAP$_{G\text{-box}}$ structure. Error bars correspond to one SD from triplicate measurements.

See also Figure S2 and Movie S1.

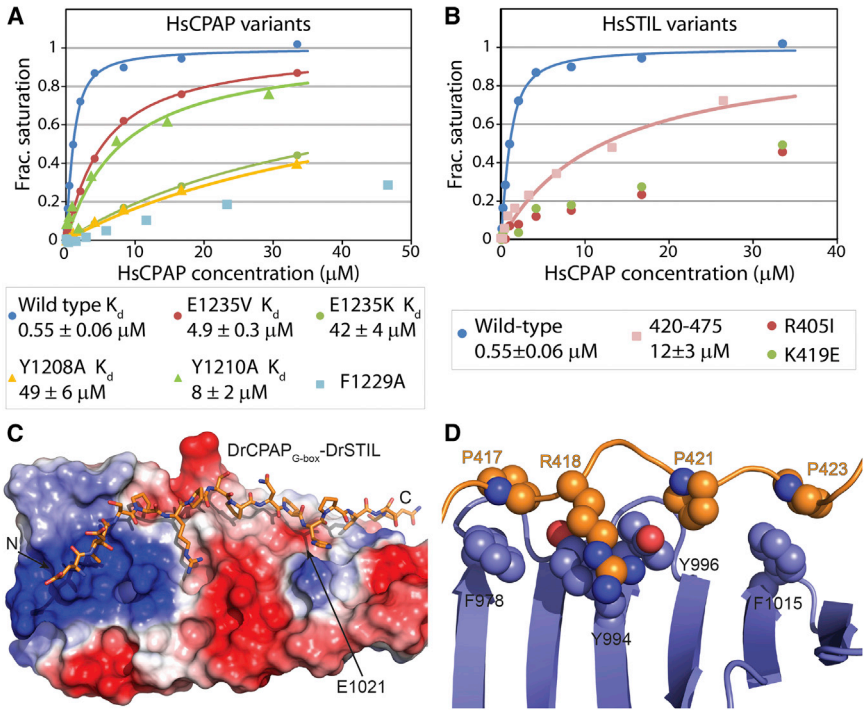

Figure 3. Biophysical and Structural Analysis of the DrCPAP$_{G-box}$-DrSTIL Complex

(A and B) Quantification of the HsCPAP$_{G-box}$-HsSTIL interaction by fluorescence polarization experiments is shown. The binding saturation of fluorescently labeled HsSTIL fragments is plotted against HsCPAP$_{G-box}$ concentration. Solid lines are fits to a 1:1 association model. (A) Shows HsCPAP$_{G-box}$ variants binding to wild-type HsSTIL residues 385–433. (B) Shows wild-type HsCPAP$_{G-box}$ binding to HsSTIL residues 419–475 or HsSTIL residues 385–433 variants. The F1229A variant of HsCPAP, and the R405I and K419E variants of HsSTIL could not be fit satisfactorily.

(C) Electrostatic surface representation of DrCPAP$_{G-box}$ with a DrSTIL fragment as sticks. The DrSTIL N termini and C termini are indicated. The complex shown was obtained using DrSTIL residues 398–450.

(D) Hydrophobic interactions between DrCPAP$_{G-box}$ and DrSTIL. Residues involved in interactions are shown in space-filling representation in blue (DrCPAP$_{G-box}$) or orange (DrSTIL).

See also Figure S3.

with DrSTIL residues 398–450, which diffracted to 2.44 Å resolution (Table 1), yielded more widespread peptide density. Thus, we were able to fit 16 DrSTIL residues (413–428, equivalent to HsSTIL residues 400–415), using the peptide register established from the 1.72 Å resolution complex, and refine the resulting model to R/R_{free} of 19.8%/21.4% (Figures 3C and S3D).

The interface revealed spans approximately half of the G-box β sheet, from the N terminus of that domain to just beyond the MCPH glutamate (E1021 in DrCPAP). DrCPAP$_{G-box}$ remains essentially unaltered upon complex formation, with a C_α root-mean-square deviation (rmsd) of just 0.79 Å compared to DrCPAP$_{G-box}$ alone. The G-box domain still forms β sheet fibrils in the crystal with no differences in their register or spacing. On DrSTIL, the complex involves the proline-rich motif identified above. Complex formation buries ~790 Å^2 of accessible surface area, and it is stabilized by six putative hydrogen bonds; two of these are formed between the E1021 side chain and the DrSTIL H424 backbone. Analysis of the structure suggests that hydrophobic surface burial is important to complex stability, with DrSTIL residues P417, P421, and P423 packing against DrCPAP F978, Y996, and F1015, respectively (Figure 3D). Similarly, DrSTIL R418 is involved in stacking and hydrogen bonding interactions with DrCPAP Y994, Y996, and H1003 (Figure 3D). Amino acid substitutions at these positions reduce complex stability, as shown in fluorescence polarization measurements (Figure 3A), using HsCPAP$_{G-box}$ Y1208A (equivalent to DrCPAP Y994), Y1210A (DrCPAP Y996), and F1229A (DrCPAP F1015). HsSTIL R405I (DrSTIL R418), a relatively conservative substitution that should maintain hydrophobic interactions but perturb the stacking arrangement and hydrogen bonding also results in significant reduction in affinity (Figure 3B).

DrSTIL electron density in the 2.44 Å resolution complex terminates abruptly where crystallographic symmetry-related molecules pack. To determine whether additional DrSTIL residues interact with DrCPAP$_{G-box}$ in solution, we performed nuclear magnetic resonance (NMR) experiments. Titrations of unlabeled DrCPAP$_{G-box}$ to ^{15}N-labeled DrSTIL residues 398–450 yielded the progressive disappearance of amide resonances because of the slow tumbling rate of the resulting complex (Figure S3F). Analysis of NMR resonance intensities suggests that DrCPAP$_{G-box}$ engages DrSTIL residues 413–449 (equivalent to HsSTIL residues 400–436; Figure S3E). To test this suggestion, we substituted HsSTIL residue K419, which is outside the STIL region observed in the crystallographic complex structure yet is conserved or conservatively substituted in all vertebrates. Fluorescence polarization experiments show that HsSTIL K419E significantly reduces affinity for HsCPAP$_{G-box}$ (Figure 3B). Thus, we surmise that in solution the complex likely spans the entire length of the CPAP G-box.

DISCUSSION

Centrioles have fascinated biologists for decades; however, only recently have studies started to elucidate the mechanistic roles of their components. An emerging feature is the structural and functional similarity of their protein components despite sequence divergence (Gönczy, 2012). The CPAP$_{G-box}$-STIL complex is a case in point: STIL and its proposed relatives SAS-5 and Ana2 share little identity making their linkage on the basis of sequence difficult. However, our results suggest that they all contain a conserved CPAP/SAS-4 interaction motif, including universally unaltered residues (Figure S3B). Together with the suggestion of a SAS-6 binding motif in STIL/Ana2/SAS-5 (Hilbert et al., 2013; Qiao et al., 2012), these results support the notion of widespread functional conservation between centriole proteins.

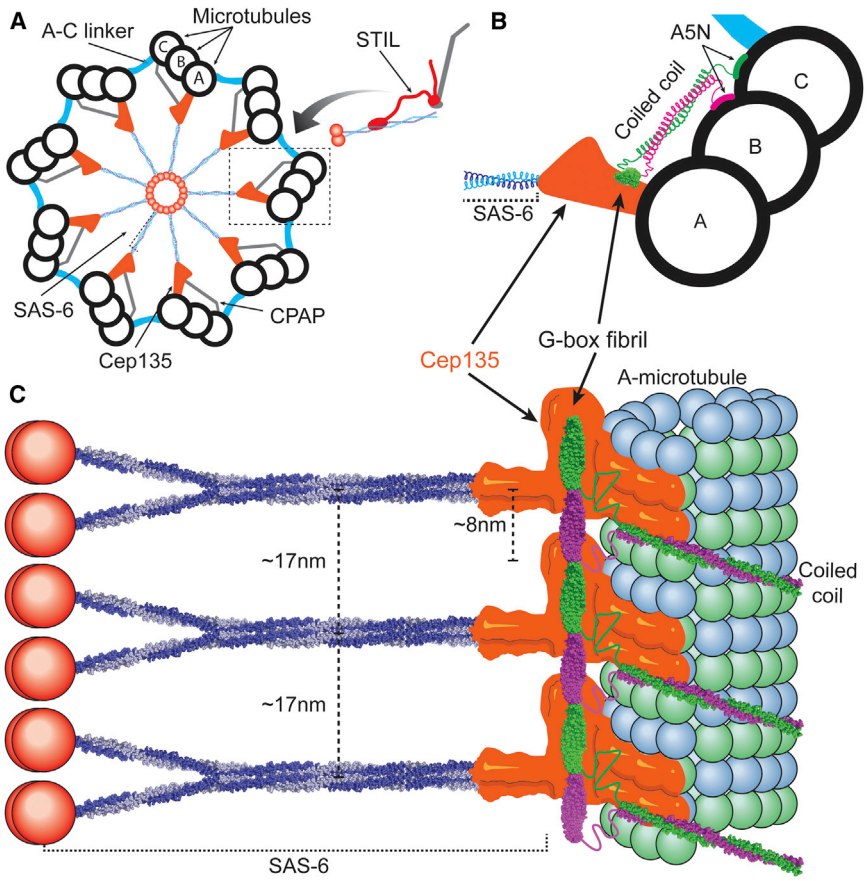

Figure 4. The Role of CPAP/SAS-4 in the Centriole Ultrastructure

(A) Schematic top view of the centriolar assembly, showing the SAS-6 cartwheel, pinheads (orange) connecting the cartwheel to microtubule triplets, and our proposed location of CPAP (gray). STIL may act as transporter of SAS-6 and CPAP to the growing centriole.

(B) View of an area close to microtubules, boxed in (A). The end of the SAS-6 spoke connects to the pinhead, where Cep135 locates. Cep135 interacts with the CPAP C terminus (green), whereas the CPAP coiled coil projects parallel to the microtubule triplet. The microtubule-interacting epitope of CPAP (A5N) can form connections to the B-tubule and C-tubule.

(C) Side view of a centriole section, showing the vertical spacing of SAS-6 spokes as they merge together, the pinhead and A-tubule connection, as well as our proposed fibril of CPAP G-box domains crosslinking cartwheel stacks on the vertical axis.

At the same time, structural studies can place components in the larger centriolar assembly. SAS-6 proteins were shown to self-assemble into cartwheels (Kitagawa et al., 2011b; van Breugel et al., 2011), whereas multiple cartwheels stack together forming a cylindrical scaffold (Guichard et al., 2012). The focus now turns on how microtubules attach to this initial scaffold and how centrioles elongate. CPAP/SAS-4 appears to be implicated in both processes, as its depletion prevents microtubule recruitment (Pelletier et al., 2006), but its overexpression leads to very long centrioles (Kohlmaier et al., 2009; Schmidt et al., 2009). CPAP/SAS-4 is placed close to centriolar microtubules, both by superresolution microscopy studies that tracked the protein N terminus (Mennella et al., 2012) and C terminus (Lawo et al., 2012), and by interaction studies, which showed that the HsCPAP C terminus associates with Cep135 (Lin et al., 2013). Our structural studies offer an intriguing possibility that integrates the aforementioned properties into a mechanistic role for CPAP in the centriole ultrastructure.

A previous cryo-tomographic study of purified *C. reinhardtii* basal bodies showed distinct electron densities in the vicinity of the microtubule wall (Li et al., 2012). We propose that CPAP accounts for an elongated density observed there that lies roughly parallel to the microtubule triplet (Figures 4A and 4B). The CPAP C terminus, which interacts with Cep135 (Lin et al., 2013), would be placed close to the A-tubule as part of the overall "pinhead" assembly that connects cartwheels to microtubules, and where Cep135/Bld10p also locates (Hiraki et al., 2007). The 24-nm-long CPAP coiled coil would then stretch approximately to the beginning of the C-tubule, where the CPAP N terminus could form connections through its microtubule associating segments (Hsu et al., 2008). Thus, we propose that CPAP acts as a horizontal "strut," linking the pinhead with microtubules and reinforcing the overall centriole assembly. It is possible that differences in CPAP length among organisms are related to apparent differences in microtubule tilt with respect to cartwheel spokes; however, further modeling efforts would be required to assess this hypothesis.

G-box domains were shown to form fibrillar structures with ~8 nm periodicity (Figure 2E), which is proportional to the size of the tubulin heterodimer and approximately half of the vertical span between successive cartwheel spokes (Guichard et al., 2012). We propose that the G-box domains of the CPAP dimer bridge that span by associating in a head-to-tail arrangement (Figure 4C); the relatively long linkers connecting the G-box and coiled-coil domains allow for this domain orientation to occur (Figure 1I). G-box domains of successive CPAP dimers would thus link into a continuous fibril in a manner analogous to that observed in the crystal, thereby providing vertical connections and complementing the horizontal strut role outlined above. We believe that the periodicity of the G-box fibril is no coincidence and that it may act as a "molecular ruler" for the initial centriolar scaffold prior to microtubule attachment. Furthermore, excessive CPAP amounts in the cell likely would allow formation of longer G-box fibrils, hence driving the creation of abnormally long centrioles.

What is the role of STIL in this mechanistic picture? Fluorescence recovery after photobleaching assays in human (Vulprecht et al., 2012) and *C. elegans* (Delattre et al., 2004) cells showed that STIL/SAS-5 transits rapidly between centrioles and the cytoplasm; this is in contrast to SAS-6 (Leidel et al., 2005) or

SAS-4 (Leidel and Gönczy, 2003), the fluorescence of which recovers very slowly. This property suggests that STIL does not form permanent structural connections to centrioles as SAS-6 and CPAP do, but rather it may act as a transporter molecule for these components (Figure 4A). The MCPH-causing HsCPAP E1235V mutation reduces CPAP-STIL affinity and thus may reduce the rate of CPAP transport. Although centriole duplication is ultimately little affected in this HsCPAP variant (Kitagawa et al., 2011a), the reduced rate of CPAP accumulation may be the basis of MCPH.

In summary, our findings suggest an intriguing structural role for CPAP in centriole assembly. Observing the unusual means, such as G-box domains, by which nature has overcome mechanistic challenges in building this assembly is fascinating.

EXPERIMENTAL PROCEDURES

Details of the experimental procedures are provided in the Supplemental Experimental Procedures. Briefly, G-box domains were produced by glutathione affinity and gel filtration chromatography, whereas $HsCPAP_{CC}$ and STIL fragments were purified by metal affinity and gel filtration chromatography or were made synthetically. For fluorescence polarization, STIL fragments were derivatized with 5-FAM. Biophysical experiments were performed at 20°C in PBS buffer supplemented with DTT, with the exception of $HsCPAP_{CC}$ crosslinking for which no DTT was added. NMR sequential assignments were performed using triple resonance experiments, and they have been deposited to the BioMagResBank under accession number 19318. Protein crystals were obtained by the sitting drop vapor diffusion technique at 20°C. For Sm-derivatization, crystals were incubated with 10 mM $SmCl_3$ for 1 hr. Phase information for $DrCPAP_{G\text{-}box}$ was obtained using the single-wavelength anomalous diffraction method, whereas $DrCPAP_{G\text{-}box}$-DrSTIL complexes were solved by molecular replacement. Crystallographic data processing and refinement statistics are provided in Table 1. The model and associated data have been deposited in the RCSB Protein Data Bank under the accession numbers 4LD1 ($DrCPAP_{G\text{-}box}$), 4LD3 ($DrCPAP_{G\text{-}box}$-DrSTIL residues 398–450), and 4LZF ($DrCPAP_{G\text{-}box}$-DrSTIL residues 414–428). Details of molecular dynamics simulations are provided in the Supplemental Experimental Procedures.

ACCESSION NUMBERS

The RCSB accession numbers for the structures of $DrCPAP_{G\text{-}box}$ alone, in complex with a DrSTIL residues 398–450 fragment, or in complex with a DrSTIL residues 414–428 fragment are 4LD1, 4LD3, and 4LZF, respectively. The BioMagResBank accession number for the sequence-specific NMR assignments of DrSTIL residues 398–450 is 19318.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, three figures, one movie, and three 3D molecular models and can be found with this article online at http://dx.doi.org/10.1016/j.str.2013.08.019.

ACKNOWLEDGMENTS

We are grateful to David Staunton, Nick Soffe, and Edward Lowe for upkeeping the supporting facilities. We thank Pierre Gönczy, Nicola J. Brown, Michel O. Steinmetz, Paul Guichard, and Virginie Hatchet for insightful discussions and critical reading of our manuscript. We acknowledge the Diamond Light Source for provision of synchrotron radiation facilities. The Wellcome Trust supported the Oxford Biochemistry NMR facility (094872/Z/10/Z), computational resources (092970MA), I.V. (088497/Z/09/Z), and E.C. (PhD studentship). This work was supported by grants from the Biotechnology and Biological Sciences Research Council (BB/J008265/1 to I.V. and BB/1019855/1 to P.J.S.). M.C.E. was supported by a Marie Curie Intra-European Fellowship (235532).

Received: June 27, 2013
Revised: August 19, 2013
Accepted: August 21, 2013
Published: September 26, 2013

REFERENCES

Aplan, P.D., Lombardi, D.P., and Kirsch, I.R. (1991). Structural characterization of SIL, a gene frequently disrupted in T-cell acute lymphoblastic leukemia. Mol. Cell. Biol. *11*, 5462–5469.

Brito, D.A., Gouveia, S.M., and Bettencourt-Dias, M. (2012). Deconstructing the centriole: structure and number control. Curr. Opin. Cell Biol. *24*, 4–13.

Carvalho-Santos, Z., Machado, P., Alvarez-Martins, I., Gouveia, S.M., Jana, S.C., Duarte, P., Amado, T., Branco, P., Freitas, M.C., Silva, S.T., et al. (2012). BLD10/CEP135 is a microtubule-associated protein that controls the formation of the flagellum central microtubule pair. Dev. Cell *23*, 412–424.

Chen, V.B., Arendall, W.B., 3rd, Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S., and Richardson, D.C. (2010). MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr. D Biol. Crystallogr. *66*, 12–21.

Cormier, A., Clément, M.J., Knossow, M., Lachkar, S., Savarin, P., Toma, F., Sobel, A., Gigant, B., and Curmi, P.A. (2009). The PN2-3 domain of centrosomal P4.1-associated protein implements a novel mechanism for tubulin sequestration. J. Biol. Chem. *284*, 6909–6917.

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

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

Gopalakrishnan, J., Mennella, V., Blachon, S., Zhai, B., Smith, A.H., Megraw, T.L., Nicastro, D., Gygi, S.P., Agard, D.A., and Avidor-Reiss, T. (2011). Sas-4 provides a scaffold for cytoplasmic complexes and tethers them in a centrosome. Nat. Commun. *2*, 359.

Guichard, P., Desfosses, A., Maheshwari, A., Hachet, V., Dietrich, C., Brune, A., Ishikawa, T., Sachse, C., and Gönczy, P. (2012). Cartwheel architecture of Trichonympha basal body. Science *337*, 553.

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

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

Hsu, W.B., Hung, L.Y., Tang, C.J., Su, C.L., Chang, Y., and Tang, T.K. (2008). Functional characterization of the microtubule-binding and -destabilizing domains of CPAP and d-SAS-4. Exp. Cell Res. *314*, 2591–2602.

Hung, L.Y., Tang, C.J., and Tang, T.K. (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*, 7813–7825.

Hung, L.Y., Chen, H.L., Chang, C.W., Li, B.R., and Tang, T.K. (2004). Identification of a novel microtubule-destabilizing motif in CPAP that binds to tubulin heterodimers and inhibits microtubule assembly. Mol. Biol. Cell *15*, 2697–2706.

Kitagawa, D., Kohlmaier, G., Keller, D., Strnad, P., Balestra, F.R., Flückiger, I., and Gönczy, P. (2011a). Spindle positioning in human cells relies on proper centriole formation and on the microcephaly proteins CPAP and STIL. J. Cell Sci. *124*, 3884–3893.

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

Kohlmaier, G., Loncarek, J., Meng, X., McEwen, B.F., Mogensen, M.M., Spektor, A., Dynlacht, B.D., Khodjakov, A., and Gönczy, P. (2009). Overly long centrioles and defective cell division upon excess of the SAS-4-related protein CPAP. Curr. Biol. *19*, 1012–1018.

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

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

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

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

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

Makabe, K., McElheny, D., Tereshko, V., Hilyard, A., Gawlak, G., Yan, S., Koide, A., and Koide, S. (2006). Atomic structures of peptide self-assembly mimics. Proc. Natl. Acad. Sci. USA *103*, 17753–17758.

McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and Read, R.J. (2007). Phaser crystallographic software. J. Appl. Crystallogr. *40*, 658–674.

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

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

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

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

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

Qiao, R., Cabral, G., Lettman, M.M., Dammermann, A., and Dong, G. (2012). SAS-6 coiled-coil structure and interaction with SAS-5 suggest a regulatory mechanism in C. elegans centriole assembly. EMBO J. *31*, 4334–4347.

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

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

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

Terwilliger, T.C. (2000). Maximum-likelihood density modification. Acta Crystallogr. D Biol. Crystallogr. *56*, 965–972.

Thornton, G.K., and Woods, C.G. (2009). Primary microcephaly: do all roads lead to Rome? Trends Genet. *25*, 501–510.

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

Vulprecht, J., David, A., Tibelius, A., Castiel, A., Konotop, G., Liu, F., Bestvater, F., Raab, M.S., Zentgraf, H., Izraeli, S., and Krämer, A. (2012). STIL is required for centriole duplication in human cells. J. Cell Sci. *125*, 1353–1362.