

Separation of function MCPH mutations in CPAP/ CENPJ affect centriole number and length

3.1 Abstract

Centrosomes formed by a pair of centrioles and pericentriolar material are essential organelles for microtubule (MT) organization in most animal cells. The MT triplets that form the cylindrical ultrastructure of centrioles are more robust and slow growing with a definite size when compared to the cytoplasmic MTs. The molecular capping centriole protein, CPAP (centrosomal P4.1-associated protein also known as CENPJ) is a cell cycle-regulated protein involved in regulating centrosome duplication and elongation. It utilizes conserved N-terminal PN2-3 and A5N domains to act as a molecular cap to the growing MTs. The C-terminal region of CPAP has a conserved glycine-rich G-box/TCP-10 domain which is involved in interaction with the centriole inner cartwheel protein STIL. In the neurodevelopmental disorder, autosomal recessive primary microcephaly (MCPH), mutations in the C-terminal region of CPAP have been identified. Two missense point mutations, i.e., E1235V (MCPH6) and D1196N, identified in different MCPH patients, map to the G-box domain of CPAP. However, the role of the G-box domain in CPAP functioning and its relevance in the development of MCPH is not very well understood.

Our study identified that even though the two MCPH missense mutations fall in the C-terminal domain of CPAP they cause distinct effects on protein functioning at centrioles. The E1235V (EV) mutation prevents CPAP centriole recruitment and causes overly long centrioles. The D1196N (DN) mutation causes an increase in centriole numbers without affecting centriole localization. Moreover, both mutations prevent binding to the centriole cartwheel protein STIL, which controls centriole duplication. It interacts with other cartwheel proteins, such as SAS6, and helps to stabilize and organize the cartwheel

architecture. It is also involved in recruiting other proteins required for centriole elongation, such as CPAP that participates in the regulation of MT polymerization at the growing ends of the centriole [Tang et al. 2011]. The work highlights the involvement of an alternate CEP152-dependent route for CPAP centriole localization. Molecular dynamics simulations suggest that EV leads to the destabilization of CPAP, while DN stabilizes it. Collectively, our work reports that the CPAP region outside the MT interacting domains influences centriole number and length, which translates to spindle defects and reduced cell viability. This study provides new insights into the molecular causes of primary microcephaly.

3.2 Introduction

Centrosomes are considered to be the major microtubule organizing centers (MTOCs) in the majority of dividing animal cells. In non-dividing cells, they transition to form the core of basal bodies, which extend to form the axoneme of cilia and flagella. They lack membrane boundaries and concentrate several proteins and regulatory molecules which provide a platform for efficient protein interactions. Hence, centrosomes play crucial roles in cell signaling, motility and cell division [Vertii et al. 2016]. It consists of a pair of MT-based cylindrical centrioles embedded in a proteinaceous matrix known as pericentriolar material (PCM). In mammals, a mature centrosome is approximately 250 nm in diameter and 500 nm in length. Centrioles exhibit a remarkable 9-fold radially symmetric organization of MT triplets which are stable and slow-growing [Gupta and Kitagawa 2018; Vorobjev and YuS 1982; Alieva and Uzbekov 2016].

Evolutionarily conserved centrosome proteins which are involved in regulating centriole duplication during the S-phase of the cell cycle were identified in worms, flies and humans [Delattre et al. 2004; Leidel et al. 2005; Leidel and Gönczy 2003; Kemp et al. 2004; Pelletier et al. 2006]. In humans, the formation of new centrioles (procentrioles) involve proximally located mother centriole proteins, CEP192 (functional orthologue of SPD-2 in worm) and CEP152 (functional orthologue of Asterless in flies) which initiate the recruitment of centrosome specific-kinase, PLK4 (functional orthologue of ZYG1 in worms and SAK in flies) adjacent to the mother centriole [Sonnen et al. 2013; O'Connell et al. 2001; Delattre et al. 2006; Bettencourt-Dias et al. 2005; Habedanck et al. 2005]. PLK4 then recruits STIL (SCL-TAL1 interrupting locus; SIL; functional orthologue is SAS5 in worms and ANA2 in flies) protein and phosphorylates the STAN domain of STIL [Dzhindzhev et al. 2014]. This promotes the recruitment of central cartwheel protein SAS6 which oligomerizes to form the nine-fold symmetry of MTs [Nakazawa et al. 2007] at the centrioles. The conserved proline-rich N-

terminal of STIL, CR2 region is involved in interaction with the C-terminal G-Box domain of CPAP protein [Cottee et al. 2013]. STIL-CPAP interaction helps to recruit CPAP to the outer region of the cartwheel for procentriole formation. The C-terminal of SAS6 interacts with the cartwheel pinhead protein CEP135, which in turn directly interacts with CPAP and MTs [Lin et al. 2013]. SAS4 is the worm orthologue of CPAP, which regulates centrosome size regulation by controlling PCM level [Kirkham et al. 2003]. In cycling human cells, over-expression of CPAP leads to the formation of elongated centrioles leading to abnormal spindle organization and defects in cell division [Kohlmaier et al. 2009; Schmidt et al. 2009; Tang et al. 2009].

Human CPAP was first identified in a study using yeast two-hybrid assay to screen interacting partner of γ -tubulin ring complex [Hung et al. 2000]. It consists of three major conserved domains including N-terminal MT destabilizing PN2-3 domain (311-422 residues) [Hung et al. 2004; Cormier et al. 2009], followed by MT stabilizing A5N domain (423-607 residue) [Hsu et al. 2008] and C-terminal glycine-rich G-box/T-complex protein 10 (TCP-10) domain (1150-1338 residues) [Zheng et al. 2014] (Fig.3.1).

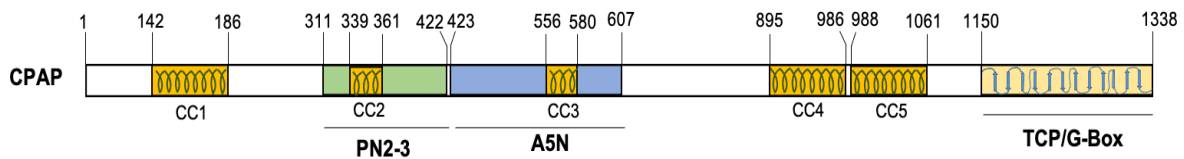


Figure 3.1: Schematic representation of major domains in the human CPAP protein- PN2-3, A5N, TCP-10/G-box domain and five CC coiled-coil regions.

Structural study demonstrates that CPAP MT stabilizing/destabilizing activity defines centriole size by actively engaging its PN2-3 domain with the β -tubulin site involved in the longitudinal interaction of tubulin dimers [Sharma et al. 2016]. CPAP binding to growing plus ends causes a peculiar slow and processive growth of centriole MTs [Sharma et al. 2016]. The A5N domain engagement ensures MT stability. CPAP C-terminal region (CP3, 895-1338 residues), includes the TCP-10 domain involved in direct interaction with STIL, and, the coiled-coil CC4 and CC5 regions responsible for CPAP localization and homo-dimerization, respectively [Lin et al. 2013; Zhao et al. 2010]. CPAP is regulated by PLK2 phosphorylation at S589 and S595 residues in a cell cycle stage-specific manner, such that its phosphorylation increases at the G₁/S transition and decreases during the mitosis exit phase. Phosphorylated CPAP localizes at the procentrioles during the centrosome cycle [Chang et al. 2010]. Additionally, its centriolar localization depends on its binding to centriole

lumen proteins PPP1R35 [Sydor et al. 2018], CEP152 [Cizmecioglu et al. 2010], STIL, and, Centrobin [Gudi et al. 2014]. STIL acts as a scaffold protein aiding the recruitment and localization of CPAP [Cottee et al. 2013]. Centrobin is also required for preventing the proteasome-mediated degradation of CPAP [Gudi et al. 2015]. CPAP recruits CEP120 and SPICE1 to grow centrioles to control centriole elongation [Comartin et al. 2013]. *In vitro* studies show that centriole protein CPAP caps the growing end of the MTs leading to their slow growth and enhanced stability [Gomes Pereira et al. 2021]. CPAP also interacts with free tubulin subunits to possibly regulate their availability for centriole elongation. Other proteins such as CEP135 and CEP295 interact with centriole MTs to promote growth whereas proteins at the distal end of centriole such as CP110 and CEP97 suppress centriole MT growth [Comartin et al. 2013; Spektor et al. 2007; Franz et al. 2013; Delgehyr et al. 2012]. However, mechanistic insights into the regulation of these proteins involved in maintaining correct centriole length remain to be explored. Knockdown of CPAP results in an increased population of mitotic cells with abnormal spindle organization [Tang et al. 2009; Cho et al. 2006; McIntyre et al. 2012; Kitagawa et al. 2011]. CPAP plays a crucial role in cerebral cortex development and neurogenesis regulation [Ding et al. 2019] as its knockdown in mouse embryos lead to extended mitosis and p53-dependent cell death in neural precursor cells [Garcez et al. 2015; Bazzi and Anderson 2014]. Accordingly, mutations in *CPAP* are mapped in patients suffering from neurodevelopmental defects like seckel syndrome and primary microcephaly as shown in Table 3.1.

Table 3.1: CPAP mutations reported in MCPH and SCKL

Disease	DNA Nucleotide change	Mutation Type	Protein amino acid change	Reference
Primary Microcephaly	c.3704 A>T	Substitution	p.Glu1235Val (E1235V)	[Bond et al. 2005]
	c.17delC	Deletion	p.Ser7LeufsTer4 (Thr6fsTer3)	[Bond et al. 2005]
	c.3243_3246delTCAG	Deletion	p.Ser1081ArgfsTer8	[Gul et al. 2006]
	c.3302-1G>C	Substitution	1073_1122del + 1073_1159del (double splicing)	[Al-Dosari et al. 2010]

(MCPH)			defect)	
	c.2462C>T	Substitution	p.Thr821Met	[Darvish et al. 2010]
	c.18delC	Deletion	p.Ser7ProfsTer2	[Sajid Hussain et al. 2013]
	c.3586G>A	Substitution	p.Asp1196Asn (D1196N)	[Makhdoom et al. 2021]
Seckel Syndrome (SCKL)	c.130A>T	Substitution	-	[Cuccurullo et al. 2022]
	IVS11-1G>C	Substitution	-	[Al-Dosari et al. 2010]

Homozygous deletion mutations in *CPAP* (17delC and c.3,243–3,246delTCAG) have been reported in primary microcephaly patients which results in protein truncation and single missense mutation [Bond et al. 2005; Gul et al. 2006]. *CPAP* G-box harboring the missense E1235V (EV) mutation was identified in microcephaly (MCPH6) patients [Bond et al. 2005]. The mutation reduces the direct interaction of *CPAP* with *STIL* protein [Tang et al. 2009]. Additionally, its over-expression leads to abnormally long centrioles, however, their number remains unaffected. Consequently, defective centriole results in erroneous spindle organization and positioning, which could translate to erroneous and precocious transitioning from symmetric to asymmetric division axis in the dividing neural progenitor cells of the developing brain. This affects the stem cell population and brain development in microcephaly patients [Kitagawa et al. 2011]. Structural work of *Danio rerio* *CPAP* G-box reveals its fibrillary form with approximately 8 nM periodicity, which corresponds to the comparable length of tubulin heterodimer and half size of cartwheel spoke found in *Trypanosoma* basal body [Hatzopoulos et al. 2013; Guichard et al. 2012]. These observations suggest that the G-box of *CPAP* is crucial for cartwheel assembly to regulate centriole number and possibly also involved in establishing vertical connections with the stacks of cartwheel to control centriole size. Specific mutations separating these two functions of *CPAP* would be a powerful tool to investigate mechanisms dictating this cross-talk between the centriole size and length. Recently identified microcephaly D1196N (DN) mutation occurs in the same *CPAP* G-box and its role in *CPAP* function has not yet been elucidated

[Makhdoom et al. 2021]. CPAP MCPH patients show mild to severe symptoms across different families [Leal et al. 2003; Makhdoom et al. 2021]. The patients with EV variation showed mental retardation of moderate severity with normal motor development, whereas the families harboring the DN variation showed additional phenotypes like slopping forehead, articulation difficulties, short stature, clinodactyly of toes, mild ID and drooling.

In this study, we simultaneously characterized two different CPAP MCPH mutations, EV and DN to elucidate the role of CPAP in controlling centriole length and number. Interestingly, even though both the mutations occur in CPAP G-box and are closely spaced, they showed different centriole phenotypes. The CPAP EV caused an increase in centriole length while the DN mutant led to an increase in centriole number. Surprisingly, the novel DN mutant does not interact with STIL protein but localizes on the centrioles indicating only a limited role of STIL in CPAP recruitment. The work further identified the involvement of an alternative CEP152-dependent route for CPAP centriole recruitment. *In silico* analysis of zebrafish equivalent CPAP G-box crystal structure from PDB predicts destabilizing effects of EV and stabilizing effects of DN mutation on the protein structure. Subsequently, we demonstrated that the CPAP C-terminal region (895-1338 amino acid) is sufficient for CPAP centriole localization and is capable of independent regulation of various CPAP functions. Further work revealed that CPAP induced centriole defects in length and number lead to multipolar spindle organization and decreases cell viability. Hence, the study provides mechanistic insights into CPAP MCPH mutations causing primary microcephaly.

3.3 Materials and Methods

3.3.1 Cell culture

HeLa Kyoto cells were cultured in DMEM medium (Himedia, #AL111) supplemented with 10% fetal bovine serum (Merck, #12103C), 2 mM antibiotic penicillin and streptomycin solution (Merck, #P4333), and 2 mM L-glutamine (Merck, #G7513) in a humidified incubator with 5% CO₂ at 37 °C.

3.3.2 cDNAs and Cloning

Full-length and C-terminal coding sequences (residues 895-1338, CP3) were PCR amplified from the siRNA-resistant CPAP, which were fused in the same reading frame to EGFP in pCDNA5/FRT/TO or to 6xHis-MBP tag in pET Duet vector. Full-length and

mutants (E1235V and D1196N) were generated by site-directed mutagenesis using the Phusion High-Fidelity DNA Polymerase (New England Biolabs, #M0530S), followed by DpnI digestion (Takara, #1235A). The cDNA of STIL (Addgene, #80266) and the CR2 fragment (residues 375-510) sequences were PCR amplified and cloned in-frame to the GST tag in pGEX-6P1 vector. The siRNA-resistant CPAP cDNA containing plasmid was obtained from Addgene (#46390).

3.3.3 Protein expression and purification

E. coli (BL21) cells were transformed with the recombinant plasmid pETDuet_His-MBP_CPAP-CP3 (wt and mutants) and pGEX6P1_GST_STIL-CR2. The bacterial cultures were grown at 37 °C until an OD₆₀₀ of 0.4-0.8. Protein expression was induced by adding either 0.4 mM IPTG to the cell culture with pET_Duet_His MBP_CPAP-CP3 (wt and mutants) or 1 mM IPTG, 0.1 mM ZnCl₂, 2% glucose solution and 1.5% ethanol solution for pGEX6P1_GST_STIL-CR2. The cells were grown at 18 °C for 18 hours. For affinity purification of the His-tagged proteins, the bacterial cells were lysed using the lysis buffer (50 mM Tris-Cl pH 8.0, 300 mM NaCl, 5 mM Imidazole). The cell lysate was sonicated at 35% amp on ice, thrice for 10 seconds with a 10 seconds break. The cell debris was pelleted down at 13,000 rpm, for 10 minutes 4 °C and the supernatant was incubated with pre-equilibrated Ni-NTA agarose beads (Qiagen, #1018244) with shaking on ice for 3 hours. The beads were spun down at 2000 rpm for 2 minutes to remove unbound fraction. This was followed by three washing steps with the washing buffer (50 mM Tris-Cl pH 8.0, 300 mM NaCl, and 20 mM Imidazole). Finally, the protein was eluted from the beads using the elution buffer (50 mM Tris-Cl pH 8.0, 300 mM NaCl, 300 mM Imidazole). GST-tagged proteins were purified by the same protocol using the GST lysis buffer (50 mM Tris-Cl pH 8.0, 100 mM NaCl, 1 Triton-X, 1mM DTT, 1 mM PMSF), the glutathione superflow resin (Takara, #635607), and the GST wash buffer (1x PBST, 1% NP-40, 2 mM DTT).

3.3.4 Cell Transfection and synchronization

For plasmid transfections, HeLa Kyoto cells at 80-90% confluency were transfected using Lipofectamine 3000 as per the manufacturer's protocol and analyzed after 48 hours. To visualize metaphase cells, they were synchronized for the G₂/M phase by using 5 uM RO-3306 (Merck, #SML0569) for 18 hours followed by 45 minutes of release. Cells were synchronized at the S-phase using the double thymidine block (2 mM) to detect endogenous

STIL. Cells were incubated with thymidine for 18 hours followed by 9 hours of release and the second block was introduced for additional 15 hours followed by cell fixation for immunofluorescence. Double-stranded CPAP siRNA oligonucleotides were synthesized with 3'-UU overhangs with the sequence 5'-AGAAUUAGCUCGAAUAGAA-3' (Dharmacon). STIL and CEP152 siRNA duplexes were synthesized with the sequence 5'-GUUUAAGGGAAAAGUUAUU-rU-rU-3'; and 5'-GGAGGACCAUGUCAUUAGACU-rU-rU-3', respectively (Merck).

3.3.5 Immunofluorescence

HeLa cells grown on coverslips were washed with 1X PBS to remove media. They were permeabilized using 0.1% TritonX-100 prepared in 1XPBS (1XPBST) and fixed with 4% (v/v) paraformaldehyde solution prepared in 1X PBS for 10 minutes at room temperature. The cells were washed thrice with 1XPBST and blocked using 1% BSA in 1XPBST for 1 hour. The cells were washed thrice with 1XPBST and left in the primary antibody overnight at 4 °C. After primary antibody incubation, cells were washed thrice and incubated in the secondary antibody for an hour at room temperature. Cells were stained with DAPI (0.5 µg/ml) prepared in 1XPBS for 15 minutes and mounted on glass slides with the help of mounting solution (0.62 g DABCO, 2 g mowiol, 6.25 ml glycerol, 25 ml 0.2 M Tris-Cl pH 8.5). The fluorescent images were acquired using the 100X oil, 1.35 NA objective lens of the epifluorescence microscope (Olympus IX83).

Primary antibodies used for immunofluorescence were mouse anti-centrin1 (Merck, #04-1624, 1:500), rabbit anti-STIL (Abcam, #Ab89314, 1:100), mouse anti- α -tubulin (Merck, #T6074, 1:500), mouse anti-tubulin acetylated (Merck, #T6793, 1:2000), rabbit anti-CENPJ (Invitrogen, #PA597577, 1:300), SAS6 (Abcam, # Ab85993, 1:200), CEP135 (Proteintech, #24428-1-AP, 1:100). Secondary antibodies used were anti-rabbit IgG (Fc) TRITC (Merck, #SAB3700846, 1:500), anti-mouse IgG (Fc) TRITC (Merck, #SAB3701020, 1:500) and anti-rabbit Alexa Fluor 405 (Invitrogen, #A-31556, 1:250). Images are represented as maximum projections of Z-stacks and analyzed using ImageJ. The bar graphs were plotted using Microsoft excel software and the figures were arranged on Microsoft PowerPoint.

3.3.6 GST-pull down assay and immunoblotting

GST-pull down assay was performed to test the interaction between 6xHis-MBP tagged CPAP-CP3 (wt or mutants) with the GST_STIL-CR2 region. Bacterial cell lysates

expressing His-MBP_CPAP-CP3 constructs were incubated with immobilized GST_STIL-CR2 fusion protein at 4 °C for 3 hours in binding buffer (50 mM Tris-Cl pH 8.0, 100 mM NaCl, 1% Triton-X, 1mM DTT, 1 mM PMSF). After incubation, the beads were washed thrice with GST wash buffer. The proteins were eluted from the beads by incubating with the SDS laemmli buffer and followed by boiling at 95 °C for 10 minutes. The proteins were resolved on a 10% SDS-PAGE. The protein bands were transferred to the nitrocellulose membrane and analyzed by western blotting [Green et al. 2012]. Primary antibodies used for immunoblotting were rabbit anti-CENPJ (Invitrogen, #PA597577, 1:1000), mouse anti-His (Invitrogen, #MA121315, 1:1000), rabbit anti-GST (Merck, #ZRB1223, 1:1000), mouse anti- β -actin (Santa Cruz Biotechnology, #SC-47778, 1:1000). Secondary antibodies used were anti-mouse IgG HRP linked (Cell Signalling Technology, #CST 7076S, 1:10,000) and anti-rabbit IgG HRP linked (Cell Signaling Technology, #CST 7074S, 1:10,000).

3.3.7 Molecular Dynamics Simulations

Normal mode-based geometric simulation method was used for multi-scale modeling of protein conformational changes. The X-ray diffraction structures of the wild type (PDB ID: 4BXP) and E1021 (PDB ID: 4BXQ) (equivalent to human CPAP E1235V mutant residue position) of CPAP TCP10/G-box domain of *Danio rerio* were used for the simulation. The D982N structure (equivalent to human CPAP D1196N mutant residue position) was generated by mutating the 982 residues of wild-type TCP10 in UCSF Chimera. The protein structures were submitted to the NMSim webserver for rigid cluster normal-mode analysis and simulated for 5000 cycles with 0.5 step size. Other parameters were kept at default during submission. The RMSD and RMSF graphs were plotted using Origin software. Additionally, the simulation results were validated using mCSM which uses the concept of graph-based structural signatures to study and predict the impact of single-point mutations on protein stability.

3.3.8 MTT assay

Approximately 5000 cells were seeded per well in a 96-well tissue culture plate (Tarsons, #980040) and allowed to settle overnight. The next day, cells were transfected using Lipofectamine 3000 as per the manufacturer's protocol. Each transfection was performed in duplicates. The cells were treated for 48 hours at 37 °C in a 5 % CO₂ atmosphere. Then, 10 μ l of MTT [3-(4,5- dimethylthiazol-2-yl)-2,5-diphenyl tetrazolium

bromide; Invitrogen, #M6494] stock solution prepared in 1XPBS at 5 mg/ml was added to each well in 100 μ l serum-free medium. The plate was incubated at 37 °C in humidified incubator chamber for 4 hours. Acidic Isopropanol (100 μ l) was added to each well and mixed thoroughly in a shaker to dissolve the purple formazan crystals. The absorbance was measured using a microplate reader (Biotek Synergy H1) at a wavelength of 570 nm. The results were analyzed and plotted in Microsoft Excel.

3.3.9 Quantification and Statistical Analyses

All the bar graphs were generated in Microsoft Excel which plots the mean and standard deviation for each condition. Origin software was used to plot RMSF and RMSD graphs. All experiments were repeated at least twice and the number of cells (n) analyzed was >50 for each experiment. p-values were generated using a two-tailed student's t-test. The p-values for centriole and spindle pole counting experiments were generated using the chi-square test.

3.4 Results and Discussions

3.4.1 CPAP is essential for centriole organization

CPAP localizes on the centrioles and helps in linking the proteinaceous cartwheel to the outer centriole microtubule triplet [Hatzopoulos et al. 2013]. To test the role of CPAP on centriole organization, we used small interfering RNA (siRNA) to knock-down CPAP (Fig. 3.2 a-b). We observed statistically significant ($p < 0.001$) reduction in CPAP levels at 80 nM siRNA concentration as compared to control and used the same siRNA concentration for subsequent experiments. The number of centrioles in the cells was counted using centrin-1 as a centriole marker [Paoletti et al. 1996]. In EGFP expressing control HeLa cells, the majority of interphase cells (84%) contain two centrioles while only a minor population has more than two (6%) or less than two (10%) centrioles (Fig. 3.2 c-d). However, the depletion of endogenous CPAP showed a significant increase in the interphase cell population (49%) with less than two centrioles. More importantly, we were able to rescue the CPAP depletion effect on centriole number by expressing ectopic siRNA resistant EGFP tagged CPAP (EGFP_CPAP) where the majority of interphase cells (74%) show two centrioles again, thus highlighting that the observed phenotype was due to CPAP knock-down.. However, over expression of EGFP_CPAP, did not affect the centriole number. The data is indicative that

CPAP is essential for centriole cartwheel assembly as its depletion reduces centriole number in interphase cells.

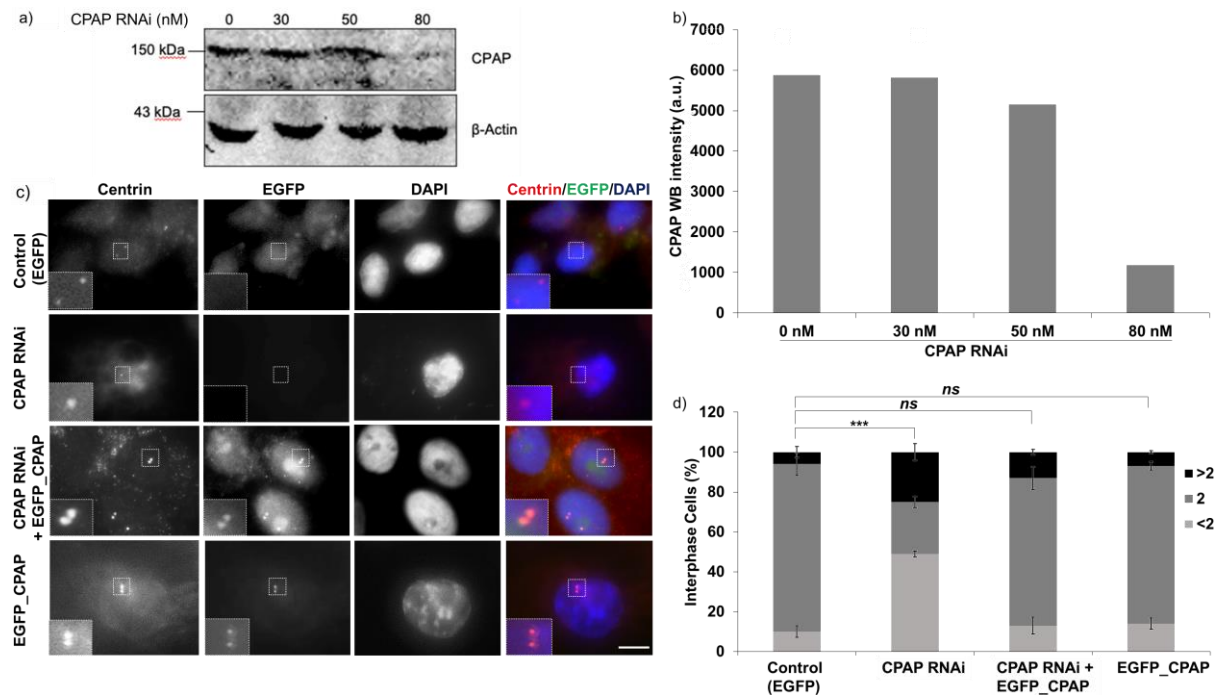


Figure 3.2: Optimization of endogenous CPAP knockdown in HeLa cells and the role of CPAP for centriole organization– (a) Western blot analysis of total cell lysate showing knockdown of endogenous CPAP using siRNA. Significant knockdown was observed at 80 nM, which was used for subsequent experiments. (b) Quantification of CPAP western blot band intensity in experiment (a). (c) Immunofluorescence image of representative transfected HeLa cells stained for centrin-1 (centriole marker) and DAPI (nuclear marker). EGFP channel represents respective constructs in each condition. The small insert shows a magnified view of the centriole in respective channels. The scale bar is 5 μ m. (d) The bar graph represents the percentage of interphase cells with <2 (light grey), 2 (medium grey) and >2 (black) centrioles in respective conditions in experiment (c). Values are average percentages $\pm$ SD from two independent experiment (n=50–100). The statistical significance represents p-value of a Chi-square test. [p>0.05: not significant (ns); p<0.001: 3 stars]

The C-terminal region of CPAP, G-box, exhibits MCPH-associated mutations. The MCPH6 patients have a single missense mutation which results in a single amino acid residue change (E1235V) in the G-box. This mutation causes an increase in the length of centrioles [Bond et al. 2005; Kitagawa et al. 2011] by an unknown mechanism. Recently, another MCPH mutation (D1196N) has been identified in the CPAP G-box [Makhdoom et al. 2021]. However, the effect of DN mutation on CPAP function remains unclear. We compared the CPAP amino acid sequences harboring these amino acid residues from organisms such as humans, mice, zebrafish and flies (Fig. 3.3) which shows the evolutionarily conserved nature of the residues.

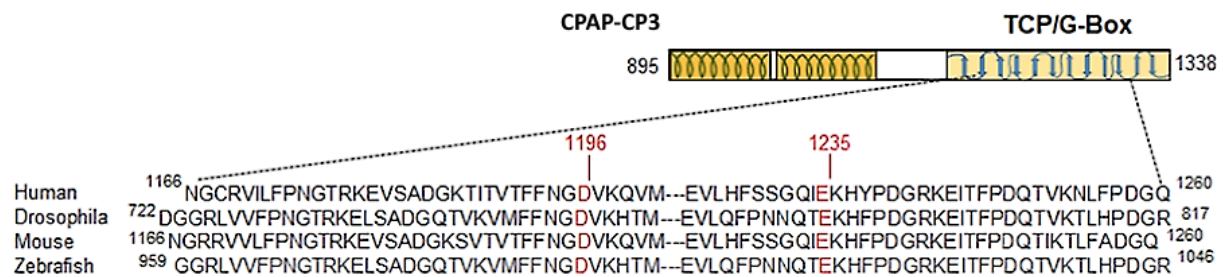


Figure 3.3: Schematic representation of CPAP-CP3 referring to the C-terminal (895-1338) part of CPAP comprising CC4-5 and the G-box domain. The alignment below shows the conserved nature of MCPH-linked D1196 and E1235 amino acid residues (marked in red) in humans, Drosophila, mouse and zebrafish.

3.4.2 Distinct CPAP MCPH mutations in the G-box affect different functions of CPAP at the centriole

Previous work has reported that over-expression of CPAP affects centriole length with nominal effects on centriole numbers [Kohlmaier et al. 2009]. Hence, we tested the role of CPAP MCPH mutations on centriole length and number using acetylated-tubulin and centrin-1 as respective markers in fluorescence microscopy. Endogenous CPAP depletion by siRNA in HeLa cells expressing RNAi-resistant wild-type EGFP_CPAP showed that the majority of the cell population (82.5%) has two centrioles (Fig. 3.4 a-b). Interphase cells over-expressing CPAP EV showed no increase in centriole numbers when compared to control cells. Contrastingly, interphase cells over-expressing EGFP_CPAP_DN showed an eight-fold increase in a cell population with more than two centrioles. Early reports have shown that CPAP EV mutation leads to elongated centrioles [Kitagawa et al. 2011]. Accordingly, we also show that expressing CPAP EV led to cells with elongated centrioles Fig. (3.4 c-d). However, expressing CPAP DN had nominal effects on the centriole length. When we expressed only the C-terminal CP3 region (895-1338 residues) harboring MCPH mutations, similar effects were observed as full-length CPAP (Fig. 3.4 e-h). These results indicate that two different MCPH mutations affect separate CPAP functions in regulating centriole length and number.

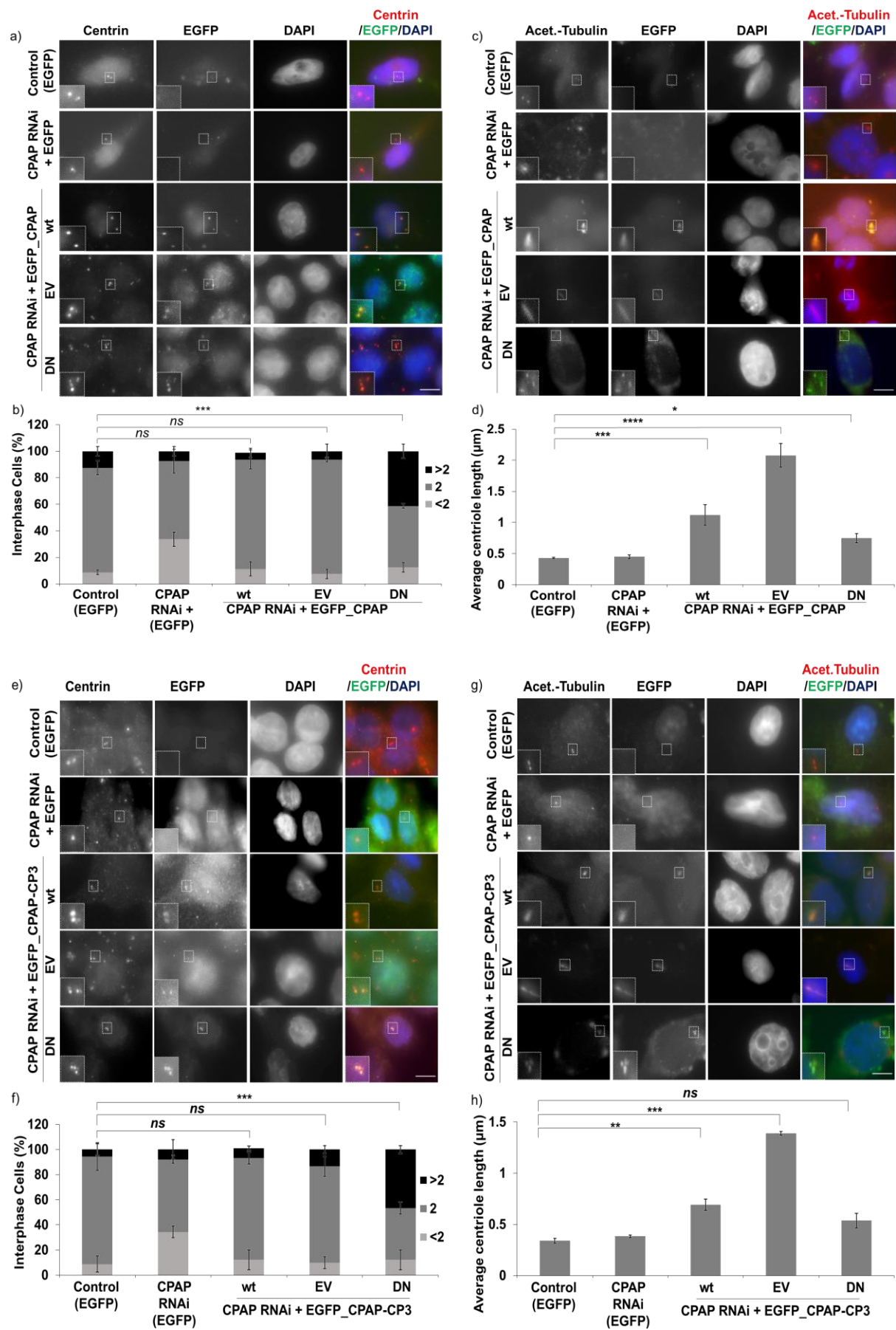


Figure 3.4: Separation of function phenotypes by CPAP and CP3 MCPH, E1235V (EV) and D1196N (DN) mutations– (a) Immunofluorescence image of representative transfected HeLa cells stained for centrin–1 (centriole marker) and DAPI (nuclear marker). EGFP channel represents respective constructs in each condition. The small insert shows a magnified view of the centriole in respective channels. The scale bar is 5 μ m. (b) Bar graphs representing the percentage of interphase cells with <2 (light grey), 2 (medium grey) and >2 (black) centrin–1 dots in respective conditions in experiment (a). Values are average percentages $\pm$ SD from two independent experiments (n=50–100). The statistical significance represents the p–value of a Chi–square test [p>0.05: not significant (ns); p<0.001: 3 stars]. (c) Immunofluorescence image of representative transfected HeLa cells stained for acetylated–tubulin (centriole length marker) and DAPI (nuclear marker). EGFP channel represents respective constructs in each condition. The small insert shows a magnified view of the centriole in respective channels. The scale bar is 5 μ m. (d) Bar graphs representing the average centriole length observed for the HeLa transfected with respective constructs in experiment (c). Values are average centriole length $\pm$ SD from two independent experiments (n=50–100). (e–h) MCPH mutations in the C–terminus of CPAP (CP3) affect distinct functions in regulating centriole number and length in respective conditions as demonstrated in experiments (a–d). The statistical significance represents the p–value of a student t–test [p<0.001: 3 stars; p<0.0001: 4 stars; p<0.05: 1 star].

3.4.3 CPAP MCPH mutations reduce direct interaction with STIL but differ in their centriole localization pattern

The EV mutation is known to affect CPAP direct interaction with centriole cartwheel protein STIL [Kitagawa et al. 2011; Tang et al. 2011]. However, the effect of DN mutation on CPAP–STIL interaction has not been tested earlier. STIL–CR2 region (375–510) residues directly interact with the CPAP–CP3 region (895–1338 residues) [Tang et al. 2011]. We faced challenges in expressing full-length recombinant CPAP or STIL, hence we used GST pull-down assay to test the direct interaction between recombinant His–MBP tagged CPAP–CP3 and GST-tagged STIL–CR2 expressed in *Escherichia coli*. We observed that DN also reduces CPAP–STIL interaction in the *in vitro* pull-down assay (Fig. 3.5 a). STIL protein is required for centriole localization of CPAP, we used fluorescence microscopy to diligently analyze the localization of CPAP MCPH mutants which do not interact with STIL. We found comparable expression levels of N-terminal EGFP tagged CPAP wild-type, EV and DN mutants in the transfected HeLa cells (Fig. 3.5 b). Moreover, the fusion proteins (EGFP–CPAP wild-type, EV and DN) localized to the centrioles where EV protein levels were significantly reduced, but wild-type and DN centriole protein levels were comparable (Fig. 3.5. c–d). Similar effects were observed for transiently transfected HeLa cells with EGFP_CPAP–CP3 (wild-type and mutants) (Fig. 3.5 e–f). This motivated us to further investigate the molecular mechanism involved in the recruitment of CPAP at the centrioles.

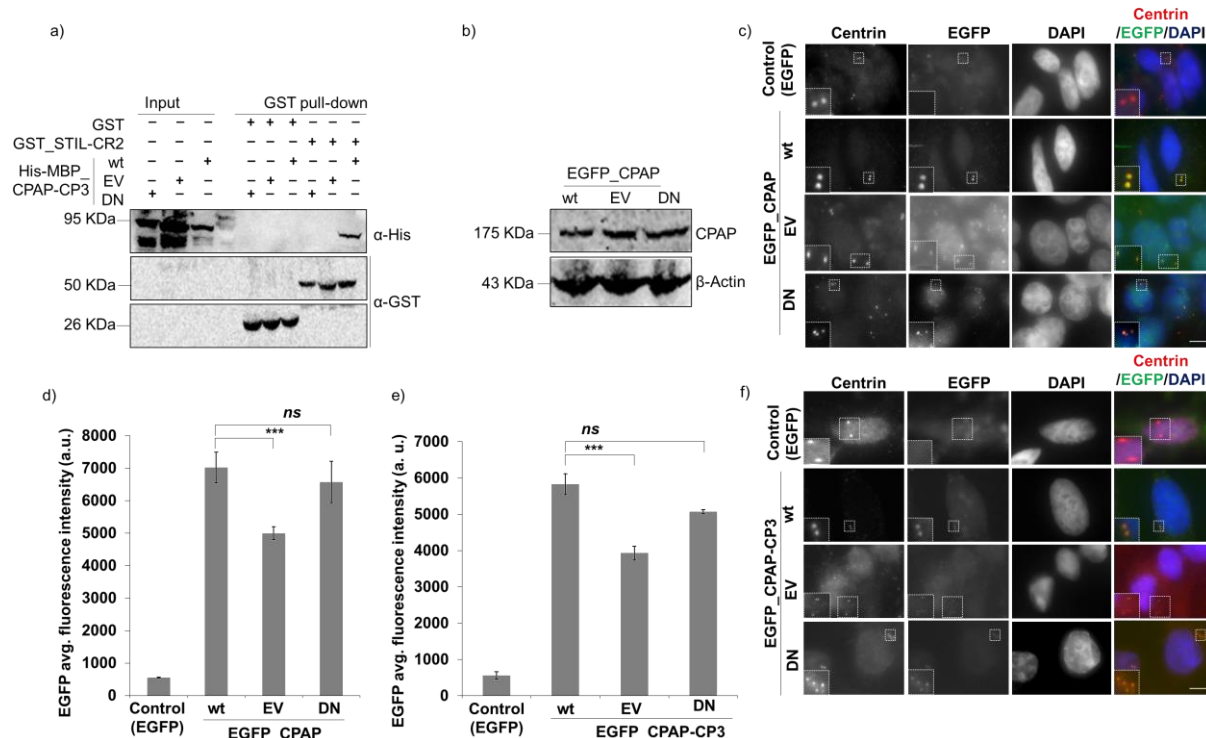


Figure 3.5: CPAP MCPH mutations, E1235V (EV) and D1196N (DN) affect direct interaction with the cartwheel STIL protein and have distinct centriole localization– (a) The western blot showing loss of interaction between GST_STIL–CR2 region and His–MBP_CPAP–CP3 EV and His–MBP_CPAP–CP3 DN as compared to the His–MBP_CPAP–CP3 wild-type (wt) in the GST–pull down lanes. (b) The western blot analysis shows equal expression levels in the total cell lysates of transfected HeLa cells with EGFP_CPAP wt, EGFP_CPAP EV and EGFP_CPAP DN. (c and f) Immunofluorescence image of representative transfected HeLa cells stained for centrin–1 (centriole marker) and DAPI (nuclear marker). EGFP channel represents respective constructs in each condition. The small insert shows a magnified view of the centriole in respective channels. The scale bar is 5 μm. (d and e) Bar graphs representing the average EGFP fluorescence intensity at centrioles in respective conditions of the experiment (c, and (f, respectively). Values are average fluorescence intensity ± SD from two independent experiments (n=50–100). The statistical significance represents the p–value of a student t–test. [p>0.05: not significant (ns); p<0.001: 3 stars].

3.4.4 CPAP D1196N mutant reveals CEP152-dependent alternate route involved in centriole recruitment of CPAP

The STIL protein is required in CPAP localization at the procentriole [Tang et al. 2011]. So, mutations affecting this interaction should affect the localization of CPAP, which is observed for the EV mutation (Fig. 3.5 c, panel 3). However, the DN microcephaly mutation in CPAP, which is not interacting with STIL (Fig. 3.5 a) can localize normally to the centrioles (Fig. 3.5 c panel 4), which challenges the previous notion. To test this carefully, we depleted endogenous STIL using 90 nM siRNA (Fig. 3.6 a-b) and observed centriole localization of transfected CPAP (wt and mutants) by fluorescence microscopy. Depleting STIL had partially reduced levels of EGFP_CPAP at the centrioles (Fig. 3.6 c-d). Similar

results were also obtained in cells transfected with the C-terminal EGFP_CPAP-CP3 region (Fig. 3.6 e-f). This data suggest an involvement of some alternate route to STIL responsible for CPAP localization at the centrioles.

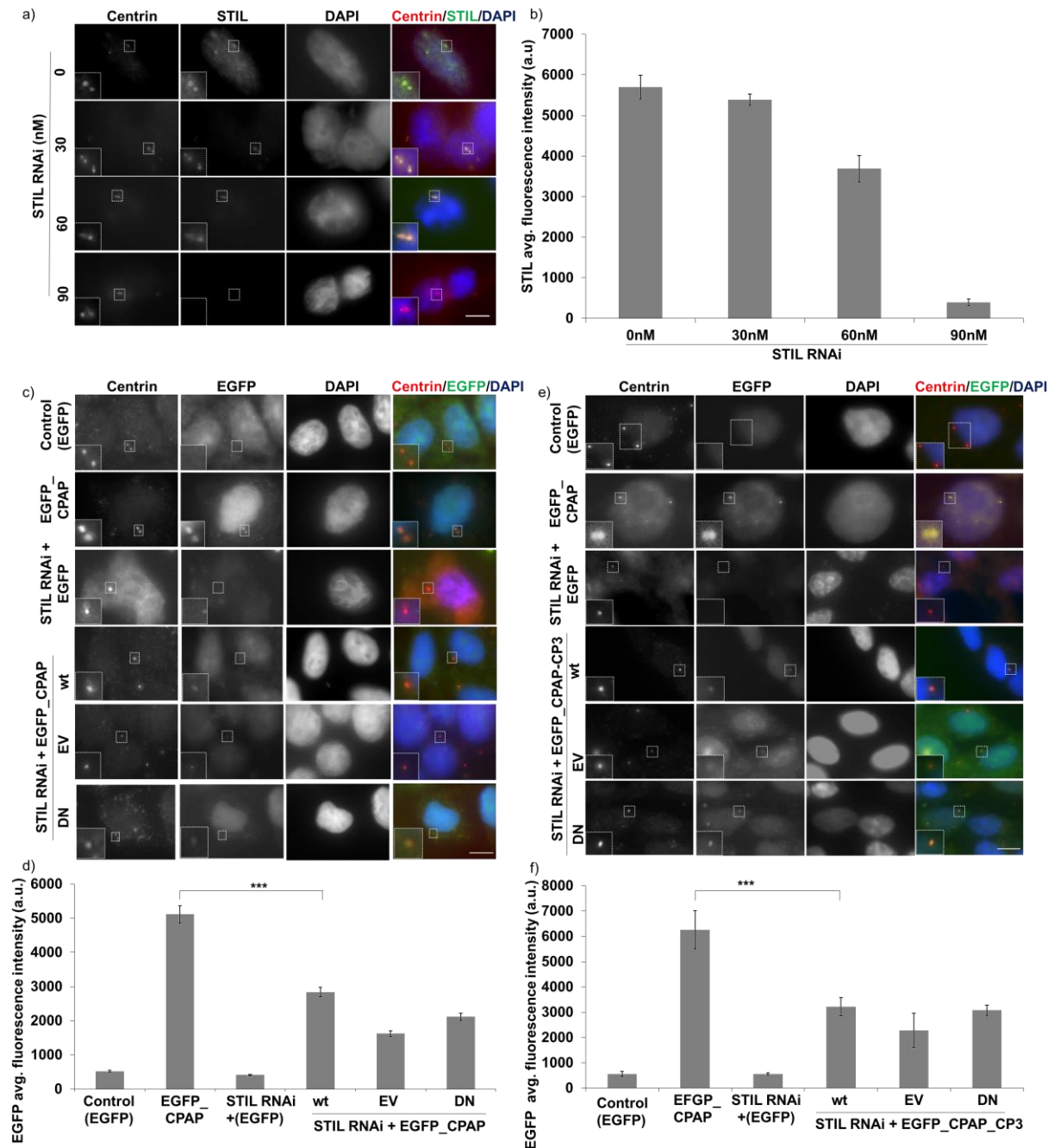


Figure 3.6: Optimization of endogenous STIL knockdown in HeLa cells and role of STIL in CPAP recruitment to centrioles– (a) Immunofluorescence image showing a dose-dependent decrease in the STIL protein levels at the centrioles in the respective siRNA treated HeLa cells. Small inserts provide a magnified view of the centrioles. A significant reduction in STIL level at the centrioles was observed at 90 nM, which is used for the subsequent experiments. The scale bar is 5 μ m. (b) Quantification of the STIL protein absolute fluorescence intensity at centrioles. (c and e) Immunofluorescence image of representative transfected HeLa cells stained for centrin-1 (centriole marker) and DAPI (nuclear marker). EGFP channel represents respective constructs in each condition. The small insert shows a magnified view of the centriole in

respective channels. The scale bar is 5 μm . (d and f) Bar graphs representing the average EGFP fluorescence intensity at centrioles in respective conditions of the experiment (c, and (e, respectively. Values are average fluorescence intensity $\pm$ SD from two independent experiments (n=50–100). The statistical significance represents the p-value of a student t-test. [p<0.001: 3 stars; p<0.0001: 4 stars].

The parent centriole protein, CEP152, is an upstream regulator, which also forms a complex with the N-terminus of CPAP [Cizmecioglu et al. 2010]. Next, we depleted CEP152 using siRNA (Fig. 3.7 a-b) and observed a drastic reduction in the levels of transfected CPAP at the centrioles (Fig. 3.7 c-d). Similar effects were observed when CEP152 was depleted in CPAP-CP3 (wt and mutants) expressing HeLa cells (Fig. 3.7 e-f). The data suggest that STIL is not working alone to recruit CPAP at centrioles, and there must be an alternate route that is CEP152 dependent.

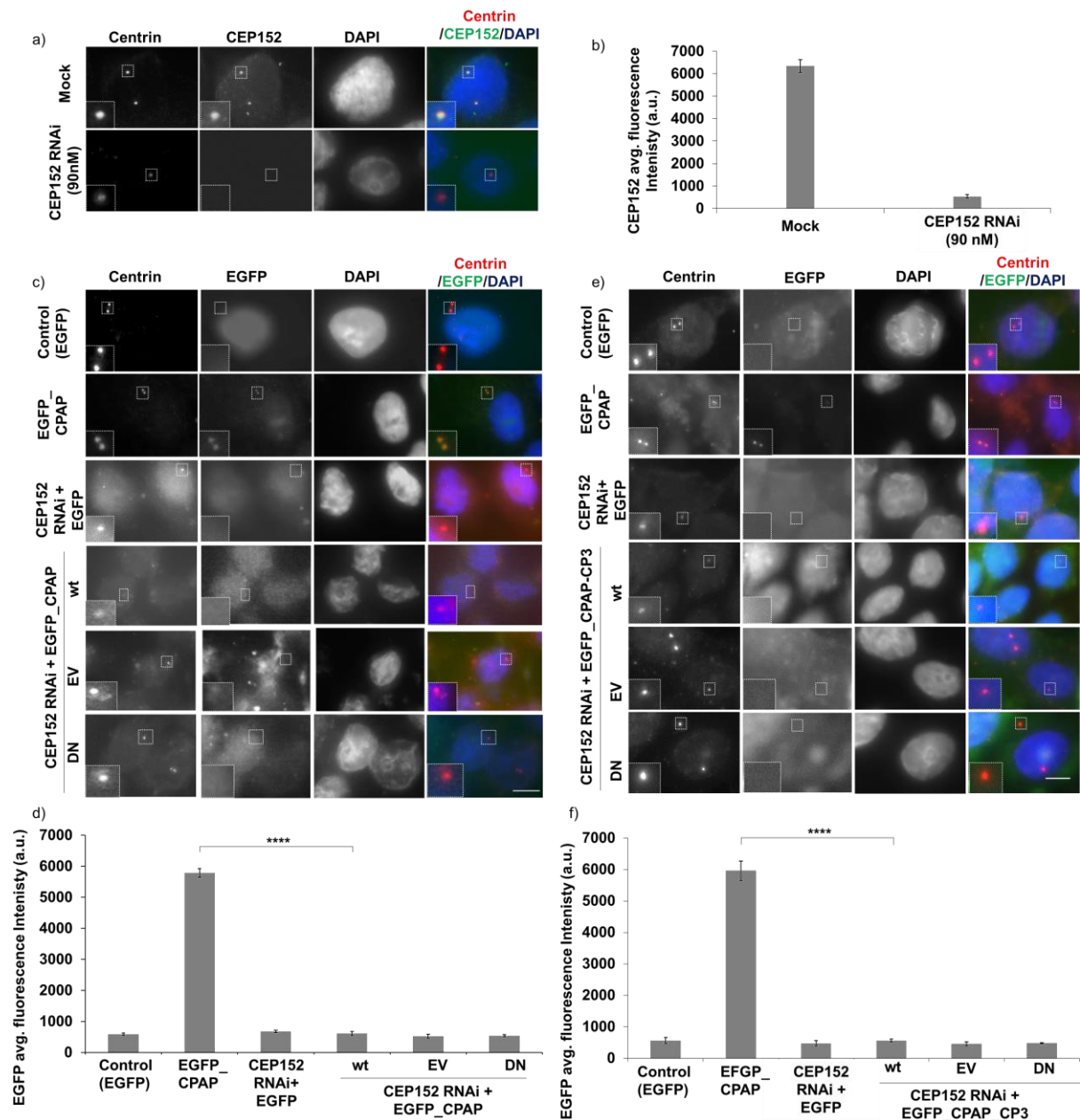


Figure 3.7: Optimization of endogenous CEP152 knockdown in HeLa cells and CEP152 dependent alternate route for CPAP localization– (a) Immunofluorescence image showing a reduction in CEP152 protein at centrioles on knockdown with 90 nM siRNA. Small inserts provide a magnified view of the centrioles. The scale bar is 5 μ m. (b) Quantification of the CEP152 protein absolute fluorescence intensity at the centrioles. (c and e) Immunofluorescence image of representative transfected HeLa cells stained for centrin–1 (centriole marker) and DAPI (nuclear marker). EGFP channel represents respective constructs in each condition. The small insert shows a magnified view of the centriole in respective channels. The scale bar is 5 μ m. (d and f) Bar graphs representing the average EGFP fluorescence intensity at centrioles in respective conditions of the experiment (c, and (e, respectively). Values are average fluorescence intensity $\pm$ SD from two independent experiments (n=50–100). The statistical significance represents the p–value of a student t–test. [$p < 0.001$: 3 stars; $p < 0.0001$: 4 stars].

Simultaneously, we observed that SAS6 localization remains unaffected in cells rescued with either the wild-type or mutant CPAP proteins (Fig. 3.8 a-b). The localization of CEP135 was not affected by CPAP siRNA, however, expression of EV mutant in the

complementation experiment causes a significant reduction in CEP135 levels at the centrosome as compared to the control cells. The DN mutant had a very low effect on CEP135 levels (Fig. 3.8 c-d). This suggests a possible dominant negative effect of the EV mutation.

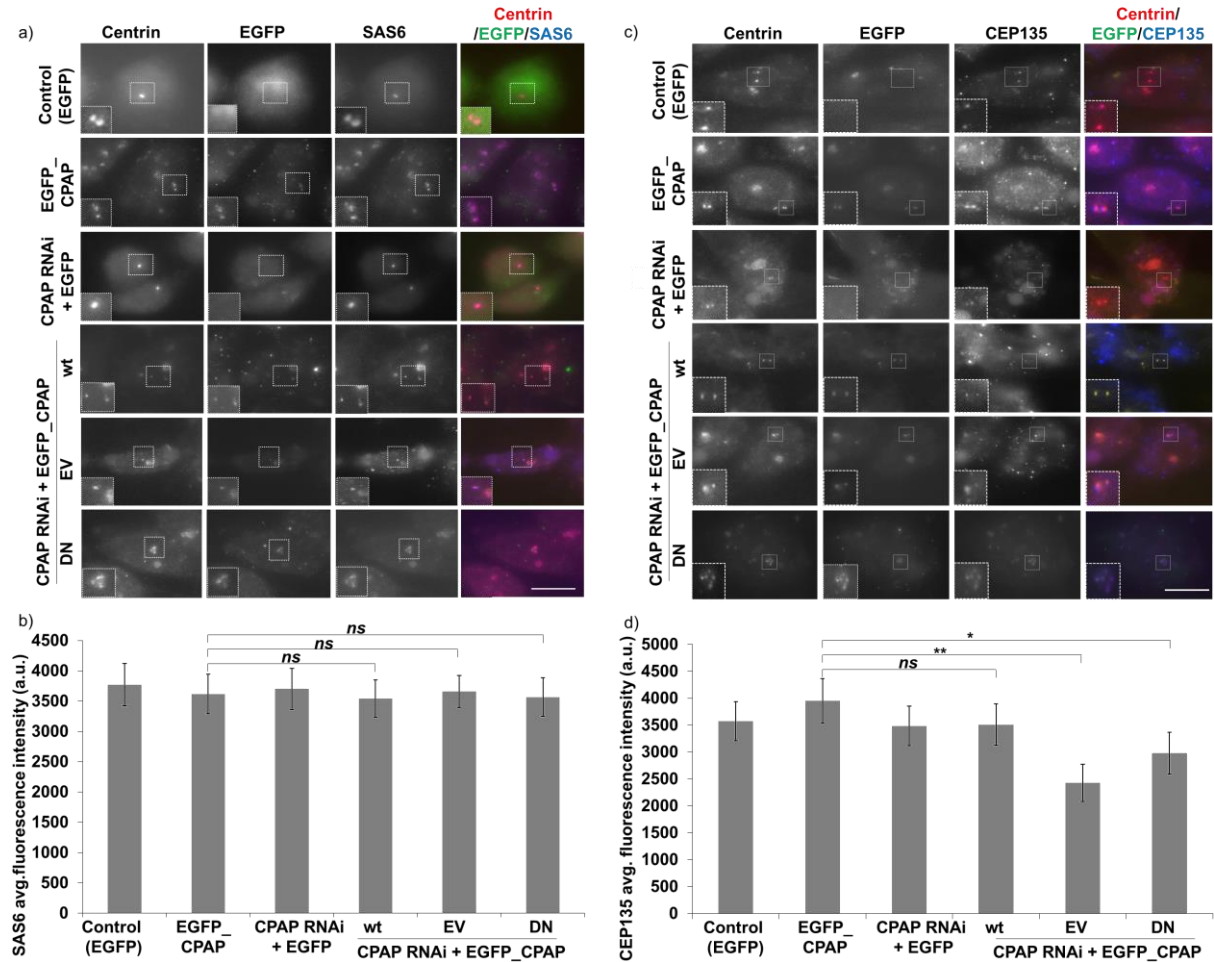


Figure 3.8: CPAP E1235V has a dominant negative effect-(a and c) Immunofluorescence image of representative transfected HeLa cells stained for centrin-1 (centriole marker) and DAPI (nuclear marker). EGFP channel represents respective constructs in each condition. The small insert shows a magnified view of the centrosome in respective channels. The scale bar is 5 μ m. (b and d) Bar graphs representing the average EGFP fluorescence intensity at centrosomes in respective conditions of the experiment (a) and (c), respectively. Values are average fluorescence intensity $\pm$ SD from two independent experiments (n=50-100). The statistical significance represents the p-value of a student t-test. [p<0.001: 3 stars; p<0.0001: 4 stars].

To understand the effect of MCPH mutation on the protein dynamics, we performed molecular dynamics simulation using the available crystal structure of wild-type (PDB ID: 4BXP) and E1021 (PDB ID: 4BXQ) equivalent to human EV residue mutation, (Fig. 3.9 a) of TCP-10/G-box domain of *Danio rerio*. The D982N structure is equivalent to human DN residue mutation (Fig. 3.8 a) and was generated by mutating the respective residue in the

wild-type TCP-10 structure in UCSF Chimera. The normal mode-based geometric simulation was performed using the NMsim webserver for 5000 cycles with 0.5 step size. The root means square deviation (RMSD) determines the deviation of the protein relative to its native conformation during its simulation. The RMSD profile indicates that the wild-type structure reaches equilibration after approximately 3500 conformations similar to D982N, however, E1021V did not attain equilibration even at 5000 conformations (Fig. 3.9 b). To analyze the effect of mutation on the residue-wise mobility of the proteins, we performed root mean square fluctuations (RMSF) analysis. The RMSF graph indicated higher fluctuations for E1021V as compared to the wild-type protein structure, whereas the D982N mutant protein had the lowest fluctuations. In the wild-type TCP-10, major fluctuations were observed near the site (residues 60-120) which harbors the D982 (residue no. 65 in the graph) and E1021 (residue no. 104 in the graph) indicating conformational flexibility of the region (Fig.3.9 b and c).

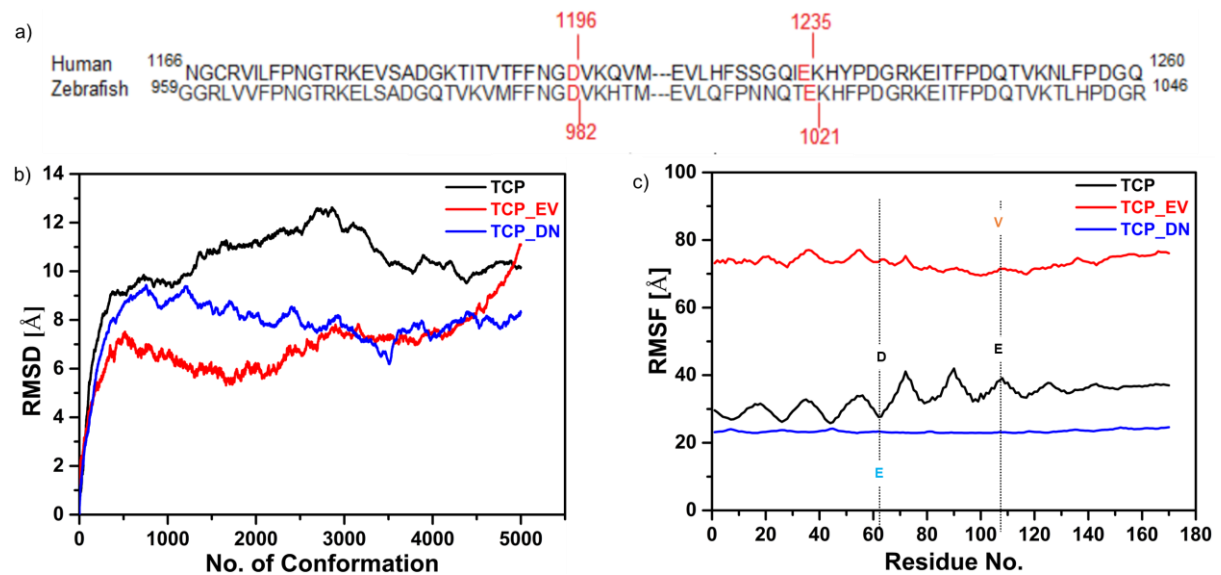


Figure 3.9: Molecular dynamic simulations TCP10/G-box domain (wt, D982, E1021) of Danio rerio– (a) Sequence alignment of human and zebrafish CPAP regions showing MCPH associated residues marked in red and their respective position number. (b) RMSD plot of TCP10/G-box from zebrafish versus the number of conformations in molecular dynamics simulation. (c) RMSF plot of zebrafish TCP10/G-box of wild type (black), E1201V (red) and D982N (blue) versus a number of residues.

Simultaneously, we also analyzed the above structures using the mutation Cutoff Scanning Matrix (mCSM) which utilizes graph-based atom distance pattern to predict change in Gibbs free energy ($\Delta\Delta G$), an indicator of protein stability. The E1021V has $\Delta\Delta G < 0$ denoting the unstability of the protein, whereas D982N has $\Delta\Delta G > 0$, which indicates the stabilizing effect of this mutation (Table 3.2).

Table 3.2: Table summarizing the change in Gibbs free energy ($\Delta\Delta G$) as predicted by mCSM webserver .

Wild-Type residue	Mutated Residue	Position	$\Delta\Delta G$ (kcal/mol)	Predicted Stability Change
Glu	Val	1021	-0.386	Destabilizing
Asp	Asn	982	0.135	Stabilizing

3.4.5 CPAP MCPH mutations alter bipolar spindle organization in dividing cells and reduce cell viability

Proper centriole organization is crucial for the formation of a bipolar spindle in dividing cells. Hence, we tested the effect of CPAP MCPH mutations on spindle organization. Our results show that depleting endogenous CPAP by siRNA leads to a seven-fold increase in the population of mitotic cells with monopolar spindles (stained with α -tubulin) (Fig. 3.10 a-b). This is because of the decrease in centriole number as observed in figure 3.1. We could rescue spindle phenotype on the expression of exogenous siRNA-resistant EGFP_CPAP. However, when we complemented with EGFP_CPAP EV and DN mutants we could observe a three-fold increase in cell population with multipolar spindle formation. These results indicate that both CPAP microcephaly mutations ultimately lead to aberrant spindle organization even though they have different effects on centriole.

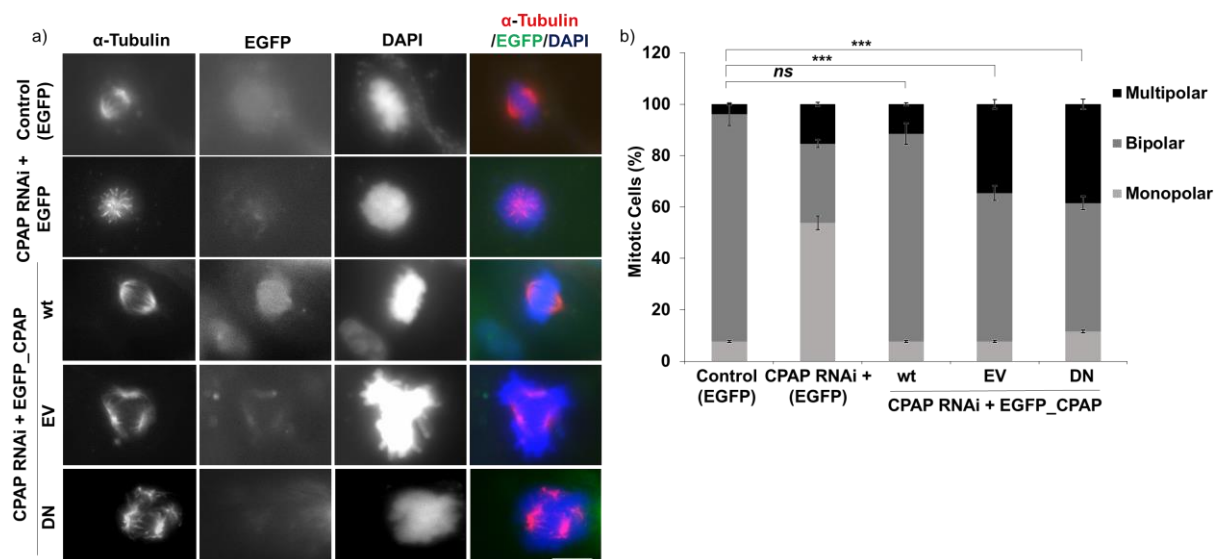


Figure 3.10: CPAP MCPH mutation causes an increase in multipolar spindles in mitotic cells– (a) Immunofluorescence image of HeLa cells transfected with respective constructs and synchronized at M-phase to visualize mitotic spindle using α -tubulin as a marker and DAPI (nuclear marker) . EGFP channel represents respective constructs in each condition . The scale bar is 5 μ m. (b) Bar graph representing the average percentage of mitotic cells with monopolar (light grey), bipolar (medium grey) and multipolar (black) spindles in respective conditions quantified from experiment (a) . Error bars are $\pm$ SD from two independent experiments (n=50). The statistical significance represents the p-value of a student's t-test [$p > 0.05$: not significant (ns) ; $p < 0.001$: 3 stars].

The MCPH mutations of CPAP should affect the pool of neural stem cells, which causes a reduction in brain size and head circumference in patients. We performed an MTT assay which measures the cellular metabolic activity to further understand the effects of CPAP MCPH mutations on cell viability and proliferation (Fig. 3.11). We show that CPAP MCPH mutations cause a decrease in cell viability which could explain their contribution to the etiology of primary microcephaly.

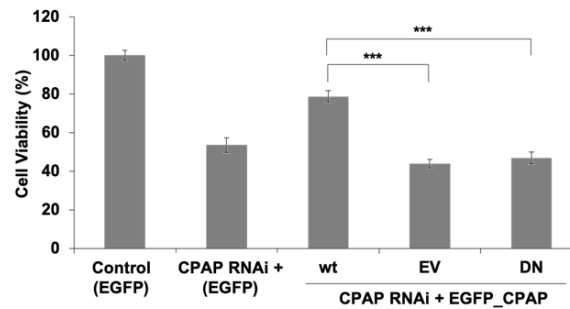


Figure 3.11: CPAP MCPH mutation decreases cell viability in mitotic cells–Bar graph representing a percentage of cell viability using the MTT assay for indicated conditions . The error bars indicate $\pm$ SD .

3.5 Conclusion

The core structure of the centrosome is made of MT-based structures called centrioles. The centriole MTs are more stable and slow growing than the cytoplasmic MTs. However, it remains puzzling as to how these centrioles achieve definite size and stability. CPAP has emerged as an essential protein linking central cartwheel proteins to the growing outer MTs and possibly also stacking these cartwheels along the length of the centriole lumen. The CPAP MT-destabilizing, PN2-3, and stabilizing A5N domains are known to interact directly with centriole MTs. The C-terminal G-box domain of CPAP is engaged in direct interaction with the core centriole cartwheel protein STIL, which is involved in CPAP localization and procentriole generation. Here, we simultaneously characterized the two reported MCPH mutations in the G-box of CPAP EV, and DN, which revealed that despite affecting the same domain of the proteins, they have different effects on CPAP functions. According to previous literature, the CPAP EV (MCPH6) caused overly long centrioles with no effect on the centriole number. Interestingly, CPAP DN does not affect centriole length, however, causes a significant increase in their number. We show using *in vitro* pull-down assay that both CPAP MCPH mutations affect direct interaction with the STIL protein. Moreover, we observed using fluorescence microscopy that EV reduces CPAP centriole localization, whereas DN does not affect its localization. Further work revealed that STIL is

only partially required for the centriole localization of CPAP, and we show an alternative CEP152-dependent pathway that is simultaneously involved in this process (Fig. 3.12).

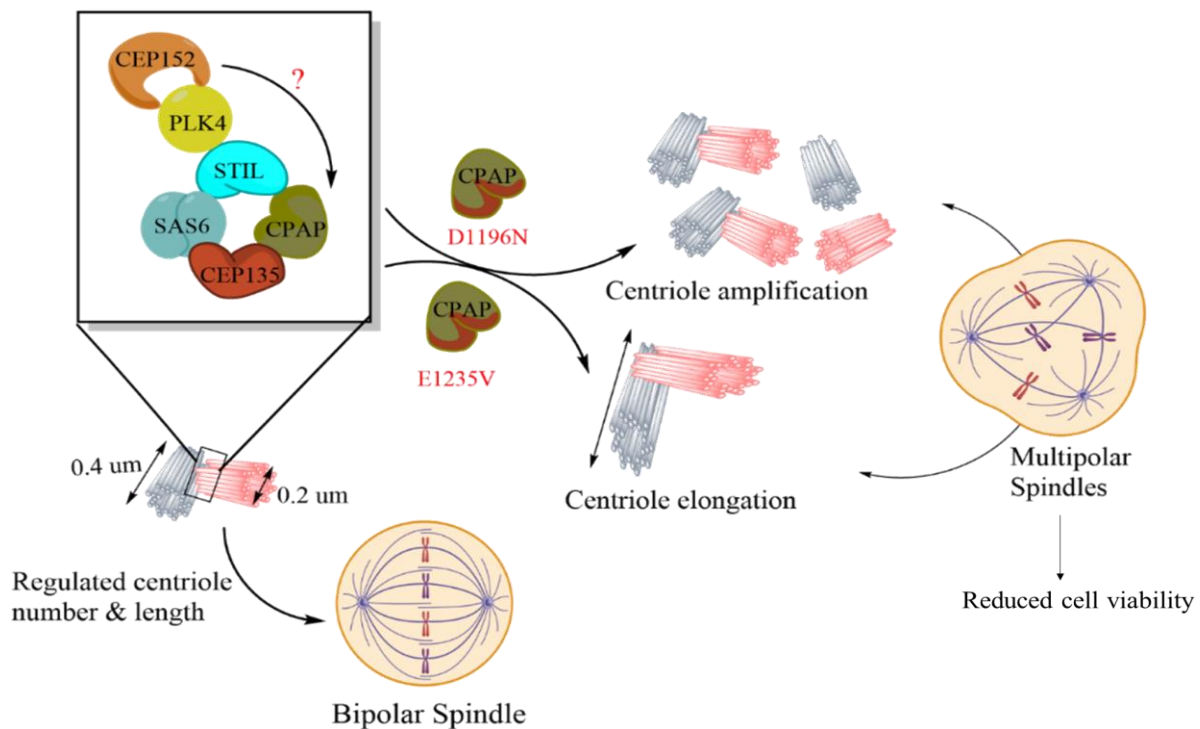


Figure 3.12: The proposed model, showing alternate CEP152 dependent route utilized by CPAP for centriole recruitment. The two MCPH mutations in CPAP cause distinct effects on CPAP functioning resulting in abnormal spindle organization and reduced viability of dividing cells.

The molecular dynamics simulations of zebrafish G-box crystal structure, wild-type, E1201V and predicted D982N, suggest that these MCPH mutations affect protein stability differently. The EV destabilizes the protein, whereas DN has a stabilizing effect on the CPAP protein. This work highlights the importance of the CPAP G-box domain in regulating two distinct functions of CPAP, i.e., centriole number and centriole size length. Since important functional mutations map to the G-box domain of CPAP, further work is required to understand the involvement of this region in the CPAP activity. It would be interesting to look at the structural effects of EV and DN on the G-box organization. In addition, both investigated MCPH mutations substitute negatively charged amino acids for uncharged amino acids, which points towards the involvement of protein interactions. Several centriolar proteins found in complexes with CPAP, like, CEP135 [Lin et al. 2013], Centrobins [Gudi et al. 2011], and CEP120 [Comartin et al. 2013] are also involved in regulating centriole length. These would be attractive candidates for gaining mechanistic insights and their possible involvement in the CPAP centriole localization and functioning.

3.6 Scientific doctoral publications related to the chapter

- **Jaiswal S.**, Sanghi S., Singh P. Separation of function MCPH mutations in CPAP affect centriole number and length. *Journal of Cell Science*, 2023, doi: <https://doi.org/10.1242/jcs.261297>
- **Jaiswal S.** and Singh P. Centrosome dysfunction in human diseases. *Seminars in Cell & Developmental Biology*, 2020, 110, 113-122.
- **Jaiswal S.**, Kasera H., Jain S., Khandelwal S. and Singh P. Centrosome: A Microtubule Nucleating Cellular Machinery. *Journal of Indian Institute of Science*, 2021, 101, 5-18.

