HHS Public Access

Author manuscript

Cytoskeleton (Hoboken). Author manuscript; available in PMC 2021 January 01.

Published in final edited form as:

Cytoskeleton (Hoboken). 2020 January ; 77(1-2): 25–35. doi:10.1002/cm.21591.

The Outer Dynein Arm Assembly Factor CCDC103 Forms Molecular Scaffolds Through Multiple Self-Interaction Sites

Stephen M. King[*], **Ramila S. Patel-King**

Department of Molecular Biology and Biophysics, University of Connecticut Health Center, 263 Farmington Avenue, Farmington, CT 06030-3305, USA.

Abstract

CCDC103 is a small protein with unusual biophysical properties that is required for outer dynein arm assembly on ciliary axonemes. Mutations in both human and zebrafish CCDC103 proteins lead to primary ciliary dyskinesia. Previous studies revealed that this protein can oligomerize and appears to be arrayed along the entire length of the ciliary axoneme. CCDC103 also binds purified microtubules directly and indeed stabilizes them. Here we use biochemical approaches to identify two regions of CCDC103 that mediate self-interaction. In both cases, these associations are stable to heating in the presence of detergent and are not disrupted by strong reducing agents. One interaction region consists of a 27-residue inherently disorder segment that can mediate heat/detergent-resistant dimerization when attached to unrelated monomeric proteins. The second interface includes the C-terminal RPAP3_C alpha helical domain. Our data suggest that CCDC103 can form an unconventional polymer and we propose models for how the monomers might be organized. We also use molecular modeling of the RPAP3_C domain to determine the structural consequences of the pathogenic H154P mutation found in human PCD patients.

Keywords

Axoneme; *Chlamydomonas*; Cilia; Dynein; Flagella; Microtubule

INTRODUCTION

Building ciliary axonemes which contain many hundreds of different proteins presents a daunting problem in macromolecular assembly (Pazour *et al.*, 2005; Nicastro *et al.*, 2006; Ma *et al.*, 2019; Pazour and King, 2019). This issue is readily exemplified by the complexity of axonemal inner and outer dynein arm incorporation onto the nine doublet microtubules. Within every 96 nm axonemal repeat, there are four outer arms which each contain 2–3 motor units depending on species, a single inner arm I1/f which has two motor units, and six

[*]To whom correspondence should be addressed (king@uchc.edu).

AUTHOR CONTRIBUTIONS

Conceptualization, SMK; Methodology, SMK and RSP-K; Investigation, SMK and RSP-K; Writing – original draft, SMK; Writing – review and editing, SMK and RSP-K; Resources, SMK; Funding Acquisition, SMK.

DECLARATION OF INTERESTS

The authors declare no competing interests.

DATA AVAILABILITY

Data are available on request from the authors.

different monomeric inner arm dynein motors (Nicastro *et al.*, 2006; Bui *et al.*, 2008). Furthermore, there is radial asymmetry and in *Chlamydomonas*, for example, outer arms are missing from microtubule doublet #1 and instead a bridge structure is incorporated (Hoops and Witman, 1983). Monomeric inner arm patterning also changes depending on the doublet microtubule number (Bui *et al.*, 2009) and even exhibits longitudinal variation between the proximal and medial/distal portions of the axoneme (Piperno and Ramanis, 1991; Fliegauf *et al.*, 2005; Yagi *et al.*, 2009; Ishikawa, 2018). Thus, the assembly process is clearly exquisitely controlled; indeed, although many mutant strains fail to incorporate one or more dynein types, to our knowledge the ectopic localization of dynein arms at inappropriate axonemal locations has not been reported.

For the inner arm system, a molecular ruler composed of two elongated coiled coil proteins (CCDC39/CCDC40[footnote 1] *aka* FAP59 and FAP172) is involved in patterning the inner arm assembly process and defines the axonemal repeat length (Oda *et al.*, 2014). In *Chlamydomonas*, the outer dynein arm consists of three different ~540 kDa heavy chain motor units (HCs), two WD-repeat intermediate chains (ICs) and eleven distinct light chains (LCs) (King, 2018). However, this ~2 MDa structure is not sufficient for axonemal attachment and requires other components for its precise localization. For example, the trimeric outer arm docking complex (ODA-DC), which is built from two coiled coil proteins and a calmodulin homologue that can bind both Ca^{2+} and Mg^{2+}, is necessary for assembly (Takada and Kamiya, 1994; Casey *et al.*, 2003; Ide *et al.*, 2013; Owa *et al.*, 2014). Indeed, outer arms in extracts that would also have contained the ODA-DC can bind all around the circumference of cytoplasmic microtubules *in vitro* (Haimo and Fenton, 1988). Thus, although the ODA-DC can form linear arrays along microtubules in a cooperative manner and direct assembly of outer arms onto microtubules (Owa *et al.*, 2014), it is clearly not sufficient to define a precise incorporation site on axonemal microtubules. Currently, the mechanism of ODA-DC localization to specific protofilaments of the outer doublet A-tubule remains unknown.

Analysis of the *schmalhans* mutation in zebrafish lead to the identification of a small coiled coil-containing protein CCDC103 as important for axonemal assembly of outer dynein arms in vertebrates (Panizzi *et al.*, 2012). This zebrafish mutation results in a premature stop codon at residue Q27 and causes motile cilia phenotypes including kidney cysts, hydrocephalus and curved body axis. Similarly, mutations within CCDC103 have been identified in several human families who also exhibit classic phenotypes associated with primary ciliary dyskinesia (PCD) including respiratory problems and *situs inversus* (Panizzi *et al.*, 2012; Shoemark *et al.*, 2018). In both zebrafish and humans, CCDC103 mutations were found to result in the loss of outer dynein arms from ciliary axonemes (Panizzi *et al.*, 2012).

Human CCDC103 (HsCCDC103) consists of 242 residues, with predicted N- and C-terminal coiled coil regions and an internal α-helical RPAP3_C domain; this domain is also

[1]Abbreviations used: CCDC, coiled coil domain containing; FAP, flagella-associated protein; HC, heavy chain; IC, intermediate chain; LC, light chain; MBP, maltose-binding protein; ODA-DC, outer dynein arm docking complex; RPAP3, RNA polymerase II-associated protein 3; TCEP, Tris(2-carboxyethyl)phosphine.

found in the eponymous RNA polymerase II associated protein 3 and in SPAG1 – a known dynein assembly factor. CCDC103 exhibits unusual biophysical properties including the formation of dimers and higher order structures that are stable to, or reform after, heating in the presence of detergent (Panizzi *et al.*, 2012; King and Patel-King, 2015). Further analysis using circular dichroism spectroscopy revealed that this protein has a melting temperature above 85°C and can refold to its native state even after heating to 100°C. Biochemical and immunofluorescence studies in *Chlamydomonas* identified CCDC103 in both the cytoplasm and the cilium where it appeared to be arrayed along the entire axonemal length (King and Patel-King, 2015). CCDC103 has also been identified by mass spectrometry in human airway cilia (Blackburn *et al.*, 2017). Axonemal CCDC103 was found to be stable to high salt treatment but could be readily extracted with 0.3% Sarkosyl as were the outer dynein arms. *In vitro* studies demonstrated that CCDC103 binds brain microtubules directly in linear arrays, and indeed stabilizes them against cold depolymerization (King and Patel-King, 2015). Together, these observations lead to a model in which CCDC103 forms a self-assembling microtubule-binding polymer which might guide incorporation of the ODA-DC and/or outer arms.

Here we further examine this model and identify regions of human CCDC103 (HsCCDC103) involved in dimer/oligomer formation. We find that dimerization is mediated through an inherently disordered region located between the N-terminal coiled coil and the RPAP3_C domain. Interestingly, the more C-terminal region involving the RPAP3_C domain is also capable of independently forming oligomers. In the *Chlamydomonas* protein (CrCCDC103), we identify a C-terminal segment capable of independent dimerization which is consistent with the ability of CrCCDC103, but not HsCCDC103, to form stable tetramers *in vitro* (King and Patel-King, 2015). Thus, with multiple self-interaction sites CCDC103 can form repetitive molecular scaffolds and we propose two general models for how these domains might interact to generate large-scale polymeric structures. We also use molecular modeling to examine the C-terminal RPAP3_C domain of HsCCDC103 and to identify subtle structural changes caused by the pathogenic H154P mutation that leads to PCD in humans (Panizzi *et al.*, 2012) and which disrupts oligomerization (Shoemark *et al.*, 2018).

RESULTS AND DISCUSSION

Domain Organization of CCDC103

Bioinformatics analyses suggest that human and zebrafish CCDC103 contain both N- and C-terminal coiled coils as well as an RPAP3_C domain (residues 96–189 of Hs_CCDC103) adjacent to the C-terminal coiled coil (Fig. 1a). Intriguingly, although CrCCDC103 appears very similar to the vertebrate proteins at the secondary structure level (King and Patel-King, 2015), neither coiled coil is predicted with high probability by the COILS algorithm although these regions are likely α-helical. However, like the human protein, CrCCDC103 does have an RPAP3_C domain located towards the C-terminal region (Fig. 1b). The segment between the N-terminal coiled coil and the RPAP3_C domain lacks a strong secondary structure prediction. Indeed, in both human and *Chlamydomonas* proteins, this

region is likely to be inherently disordered (Fig. 1a, b). There is also a second region of predicted disorder near the C-terminal coiled coil.

CrCCDC103 is 25 residues longer than the human protein. Much of this (16 residues) forms an extended Gly/Ala/Glu-rich loop within the RPAP3_C domain; intriguingly, this extension occurs immediately adjacent to the site of the human PCD-causing mutation H154P (King and Patel-King, 2015). Domain analysis using cDART also identified a DUF2587 domain C-terminal of the RPAP3_C region of CrCCDC103 (residues 208–249) which appears to be lacking in vertebrate CCDC103; this domain is annotated as bacterial in origin and its tentative identification in the *Chlamydomonas* protein may merely reflect the presence of multiple Ala/Pro residues.

The Central Inherently Disordered Region of HsCCDC103 Mediates Dimer Formation

The unusual biophysical behavior and self-interactions of intact HsCCDC103 prompted us to attempt to define which part(s) of the molecule were responsible for these properties. Previous studies of the N- and C-terminal coiled coil regions of HsCCDC103 revealed that neither participated in homodimer formation nor did they interact with each other when attached to heterologous proteins (King and Patel-King, 2015). Extending that analysis, we have now generated a soluble fusion protein consisting of HsCCDC103 residues 1–93 attached to the N-terminal region of outer arm dynein light chain LC1. Although LC1 on its own (Fig. 2 *upper panel inset*) or attached to the N-terminal coiled coil of HsCCDC103 (King and Patel-King, 2015), behaves solely as a monomer, the HsCCDC103(1–93)+LC1 protein migrated as two clear peaks when fractionated by gel filtration chromatography on a Superdex 200 Increase column. Electrophoretic analysis revealed that the larger peak consisted of both monomeric and dimeric proteins indicating that the HsCCDC103(1–93) region allows for dimerization that is resistant to, or can reform after, heating in the presence of detergent and strong reducing agents (Fig. 2 *upper panel*). The smaller peak consisted of only monomeric protein. An additional fusion protein HsCCDC103(68–94)+LC1 behaved similarly (Fig. 2 *lower panel*), demonstrating that this short inherently disordered region is sufficient for heat/detergent-resistant dimer formation.

Self-Interactions of the C-terminal Region

To assess the properties of the C-terminal portion of HsCCDC103, the region encompassing residues 93–242 (plus His-Met as preceding residues derived from the pET16b expression vector) was prepared. This recombinant protein was folded as judged by NMR spectroscopy and yielded a well-dispersed ^{1}H–^{15}N HSQC spectrum (Fig. 3a). Although numerous through-bond datasets (including HNCA, HNCACB, HNCO, and HNCACO) were obtained from ^{15}N, ^{13}C-labeled samples, only approximately half of the backbone atom resonances could be connected and assigned (Fig. 3b), possibly due to the presence of multiple conformational states. Thus, it was not feasible to analyze through-space connections and calculate the solution structure.

Gel filtration chromatography of HsCCDC103(93–242) revealed two very distinct peaks (Fig. 3c). One migrated as a monomer and indeed only monomeric protein was detected following electrophoresis. The second peak eluted with the void volume indicating that it

contained very high molecular weight material; electrophoretic analysis revealed a mixture of monomeric and oligomeric species (Fig. 3c). Thus, this region of HsCCDC103 can form heat/detergent-resistant oligomers. These do not represent disulfide-bonded assemblies as they occur even in the presence of the strong reducing agent Tris(2-carboxyethyl)phosphine (TCEP). Furthermore, when the monomer peak was re-chromatographed, only a single monomeric protein peak was obtained (Fig. 3c inset). Together, these observations suggest that the recombinant HsCCDC103(93–242) region can exist in at least two distinct conformational states, only one of which forms oligomeric structures.

Soluble CrCCDC103 was somewhat more difficult to prepare than the human protein. However, our previous gel filtration data revealed a clear tetrameric form of CrCCDC103 in addition to both dimers and higher order oligomers; an equivalent tetramer peak was not seen in similar experiments with HsCCDC103 (King and Patel-King, 2015). In an attempt to investigate this dichotomy, we tested two CrCCDC103 fusions with maltose-binding protein (MBP) to assess self-interactions of the N- and C-termini of the *Chlamydomonas* protein (Fig. 4). Electrophoretic analysis revealed that both fusion proteins were sensitive to proteolysis apparently losing the attached region of CrCCDC103 (these bands are marked by * in Fig. 4). The N-terminal fusion protein MBP+CrCCDC103(1–56) showed no evidence of self-interaction; it migrated solely as a monomer on gel filtration and only monomeric protein (and the smaller proteolytic product) were detected by gel electrophoresis; this is completely consistent with data obtained previously for the equivalent region of HsCCDC103 (King and Patel-King, 2015). In contrast, a C-terminal construct MBP +CrCCDC103(216–267) which contains the region homologous to the HsCCDC103 C-terminal coiled coil as well as the Gly/Ala/Glu-rich region unique to CrCCDC103, exhibited two peaks plus a third clearly resolved proteolytic fragment peak of smaller mass. Both peaks containing the full-length fusion protein showed evidence for SDS/heat-resistant dimer formation, and indeed for the larger peak only dimeric protein was observed following SDS-PAGE. Intriguingly, there is a discrepancy in the estimated mass from gel filtration of this oligomer (~150 kDa) and the apparent M_r following SDS-PAGE of ~95,000. This may reflect a higher order oligomer formed from more than one MBP+CrCCDC103(216–267) dimer that is not stable to detergent and heating, or that the fusion protein dimer has an elongated conformation and thus migrates on gel filtration with a larger hydrodynamic diameter than would be predicted; a similar property was found for the LC1 dynein light chain which is highly elongated and has a long axis twice the length of the two shorter axes (Wu *et al.*, 2000; Wu *et al.*, 2003). Thus, although variations in oligomerization properties are observed *in vitro*, self-association of the C-terminal region of CCDC103 is a general feature that has been conserved between vertebrates and the chlorophyte algae.

Features of the HsCCDC103 RPAP3_C Domain and Consequences of the H154P Mutation

Although the structure of the RPAP3_C domain of HsCCDC103 could not be solved (see above), an NMR solution structure of the closely related domain from RPAP3 is now available (pdb # 6EZ4; (Maurizy *et al.*, 2018)) and confirms that it is mainly α-helical as predicted previously (King and Patel-King, 2015). Based on this structure, homology models for this 92-residue region from wildtype and H154P mutant HsCCDC103 were calculated (Fig. 5). The domain consists of a six-helix bundle. In a peptide backbone, amino acid

residues rotate around the Cα-amide N bond (the Φ angle) and the Cα-carbonyl C bond (the Ψ angle), but only certain angle combinations are sterically feasible for L-amino acids; these occupy specific regions of Ramachandran plots where all possible Φ and Ψ angles are plotted against each other. Thus, a key parameter defining the quality of the models is whether the Φ/Ψ angles of the peptide backbone are within these allowable regions. For the wildtype protein model, all Φ/Ψ angles are within favored/allowable regions of the Ramachandran space plots. For the H154P mutant form, Φ/Ψ angles for the residue (D153) immediately preceding the mutant-derived Pro are only just within the allowable region, suggesting that this mutation induces a significant conformational change; in the model this leads to premature termination of the fourth helix within this domain. This local alteration in conformation is further indicated by analysis of the molecular surface and Poisson/Boltzmann electrostatics (Fig. 6). Exchange of His154 for Pro results in formation of a cavity in the mutant protein and leads to exposure of a basic patch missing from the wildtype protein surface. As this H154P mutant HsCCDC103 can dimerize but exhibits defective oligomerization (Shoemark *et al.*, 2018), one possible interpretation is that this structural change in the RPAP3_C domain is directly responsible for the alterations in bulk biochemical properties.

The H154P missense mutation causes PCD in humans and is strongly associated with *situs inversus*. It is quite prevalent among certain population groups of South Asian origin and results in a partial dynein assembly phenotype that is difficult to diagnose clinically by electron microscopy (Shoemark *et al.*, 2018). The H154P alteration is hypomorphic as the human mRNA partially rescues the *schmalhans* mutation when injected into mutant zebrafish (Panizzi *et al.*, 2012). Our structural models combined with previous biochemical data (Panizzi *et al.*, 2012; King and Patel-King, 2015) suggest that H154P alters the conformation of the RPAP3_C domain and consequently the ability of the mutant CCDC103 protein to form high order oligomers. Nasal biopsies from patients homozygous for H154P exhibit a mixture of both static (*i.e.* immotile) and motile cilia which might reflect stochastic failure of CCDC103 polymer assembly and consequently axonemal dynein arm incorporation during formation of individual cilia.

Potential Modes of CCDC103 Polymerization and Predictions for Axonemal Function

CCDC103 oligomers form spontaneously from recombinant protein *in vitro* and are also detectable in cell extracts. For example, analysis of *Chlamydomonas* cytoplasm identified an ~0.5 MDa CrCCDC103-containing complex that migrated as a mix of monomer and dimer upon gel electrophoresis (Panizzi *et al.*, 2012). The presence of at least two regions capable of self-association suggests several ways in which CCDC103 oligomers might potentially occur. These models, referred to as the “molecular ladder” and ‘beads on a string”, are illustrated in Fig. 7. In both cases, dimers form through the inherently disordered region. In the case of “beads on a string”, the C-terminal region also dimerizes which would be consistent with the formation of stable CrCCDC103 tetramers (King and Patel-King, 2015). This would result in a linear array of interconnecting dimers comprised of monomers aligned in an antiparallel manner – overall, this arrangement does not result in a polar linear structure. An alternative model envisages the inherently disordered region dimerizing two arrays of oligomers that associate through their RPAP3_C domains and/or more C-terminal

regions. Potentially, the ladder arrangement might yield a polar polymer. In both cases though, the N- and C-terminal helical/coiled coil regions would remain available for interaction with other axonemal components and thus might guide dynein arm incorporation. Our previous negative stain electron microscopy studies (King and Patel-King, 2015) of *in vitro* assembled CCDC103/microtubules complexes revealed CCDC103 particles arrayed along microtubules with a 12-nm spacing which would be consistent with a beads-on-a-string type arrangement. Potentially, further higher resolution electron microscopic approaches might allow the two model arrangements to be distinguished. Given their very unusual biochemical and biophysical properties, CCDC103 assemblies appear to be unconventional polymers.

Our current data thus support the previous prediction (King and Patel-King, 2015) that CCDC103 self-assembling polymers underlie the placement of key dynein assembly factors. *In vitro*, CCDC103 binds microtubules directly and indeed stabilizes them against cold depolymerization (King and Patel-King, 2015). This is consistent with the observation that immunofluorescence detection of CrCCDC103 in isolated axonemes was weak in regions where outer arms had assembled, presumably due to limited antibody accessibility, but much stronger at both the axonemal base and tip where these motors were absent. To date, no cryo-EM densities assigned to CCDC103 structures have been reported (*e.g.* Ma *et al.*, 2019), and so whether these CCDC103 oligomers are arrayed on the microtubule surface or embedded within the doublets as are many other axonemal microtubule-associated proteins remains to be determined.

Our data also reveal that attachment of the short HsCCDC103(68–94) region ($_{68}$KRTVPWNCHTIQGRTFQDVATEISPEK$_{94}$) is sufficient to direct the self-interaction of unrelated normally monomeric proteins through an interface resistant to denaturation by heat and detergent. Potentially this unusual property might be useful to help assess whether two tagged proteins expressed in the same cell actually are physically close *in vivo* and capable of interaction even though such interactions might not normally survive biochemical purification.

In conclusion, we demonstrate that the enigmatic outer dynein arm assembly factor CCDC103 is capable of self-assembly through two distinct interaction interfaces. One derives from a short inherently disordered region and the second from the C-terminal section that includes the RPAP3_C domain. We propose models for how these self-interactions might lead to formation of a polymer that could guide the assembly of axonemal dynein arms and their associated components.

METHODS

Protein Expression and Refolding

His-tagged N-terminal LC1 fusions and HsCCDC103 (93–242) were expressed from pET16b constructs in *Escherichia coli* BL21 (DE3); C-terminal maltose-binding protein (MBP) fusions were expressed from the pMAL-c2 vector in the same strain. *E. coli* were grown in LB medium containing ampicillin to an optical density of 0.6, and protein expression induced by the addition of 1–2 mM β-D-1-thiogalactopyranoside. Following

centrifugation, bacterial pellets were frozen at −20°C until use. Once defrosted, pellets were resuspended in 20 mM Tris.Cl pH8.0, 150 mM NaCl, 1 mM TCEP and sonicated. Following centrifugation to pellet cell debris, soluble LC1 fusions were then purified by affinity chromatography on Ni^{2+} resin. Samples were loaded onto the column in 20 mM Tris.Cl pH 8.0, 500 mM NaCl, 1 mM TCEP and 20 mM imidazole, washed extensively and protein eluted with the same buffer containing 250 mM imidazole. The His tag was removed by digestion with Factor Xa. Soluble MBP fusion proteins were purified by affinity chromatography using amylose resin; protein was eluted with 20 mM Tris.Cl pH8.0 buffer containing 150 mM NaCl, 1 mM TCEP and 10 mM maltose.

The His-tagged HsCCDC103(93–242) protein was initially obtained from inclusion bodies. Pellets remaining after sonication were resuspended overnight in 20 mM Tris.Cl pH8.0, 150 mM NaCl, 1 mM TCEP containing 8 M urea. Protein refolding was achieved by adding the urea-solubilized sample dropwise to 1l of 20 mM Tris.Cl pH8.0, 150 mM NaCl, 1 mM TCEP that was being vigorously stirred to allow for very rapid dilution. Following centrifugation to remove aggregated material, the refolded protein was purified by Ni^{2+} affinity chromatography as described above.

Gel Filtration Chromatography and Electrophoresis

All samples were fractionated in a Superdex 200 Increase 10/300 gel filtration column attached to an Äkta Purifier 100 chromatography workstation. 0.5 ml samples in 20 mM Tris.Cl pH8.0, 150 mM NaCl, 1 mM TCEP were injected and 0.5 ml fractions collected and analyzed. The Superdex 200 Increase column was calibrated using the following standard proteins: thyroglobulin, 669 kDa; apo-ferritin, 443 kDa, β-amylase, 200 kDa, alcohol dehydrogenase, 150 kDa; bovine serum albumin, 66 kDa; carbonic anhydrase, 29 kDa. Samples were routinely electrophoresed in 10 or 12.5% polyacrylamide SDS gels and stained with Coomassie blue.

^{13}C, ^{15}N-labeling and NMR Data Collection

To prepare samples for NMR spectroscopy, the C-terminal region (93–242) of HsCCDC103 subcloned in pET16b was expressed in *E. coli* BL21(DE3) cells grown in minimal M9 medium containing either $[^{15}N]NH_4Cl$ or both $[^{15}N]NH_4Cl$ and [U-^{13}C]-glucose. Proteins were solubilized in 8 M urea and refolded and purified as described above. After removal of the His tag by Factor Xa digestion, protein was concentrated to ~4.5 mg/ml (0.27 mM) in 40 mM Tris.Cl pH6.5, 100 mM NaCl, 0.1 mM TCEP using an Amicon Ultra-4 ultrafiltration concentrator with a nominal molecular weight cutoff of 10 kDa; protein samples (500 μl) were made to a final volume of 600 μl by addition of 67 μl buffer and 33 μl D_2O. NMR data collection was performed on an 800 MHz Varian spectrometer equipped with a cryogenic probe. Datasets obtained included a 2-dimensional ^{1}H–^{15}N HSQC, and 3-dimensional HNCA, HNCACB, HNCO and HNCOCA spectra.

Bioinformatics

Coiled coil sequences within HsCCDC103 (NP_001245324) and CrCCDC103(Pr46b) (XP_001696391) were predicted using the Coils algorithm (https://embnet.vital-it.ch/software/COILS_form.html) (Lupas *et al.*, 1991). Intrinsically disordered regions within

both proteins were identified with IUPRED2A using the short disorder and ANCHOR2 context-dependent prediction parameters (https://iupred2a.elte.hu) (Mészáros *et al.*, 2018). Domain analysis was performed using the conserved domain architecture retrieval tool (https://www.ncbi.nlm.nih.gov/Structure/lexington/lexington.cgi) (Geer *et al.*, 2002). Homology models of the Hs_CCDC103 RPAP3_C domain and the H154P pathogenic variant were built using Swiss-Model (https://swissmodel.expasy.org/) using pdb 6EZ4 (Maurizy *et al.*, 2018) as the modeling template. Ramachandran plots of modeled Φ/Ψ angles were generated by Swiss-Model. Full Poisson-Boltzmann electrostatics were calculated with the APBS tools plugin within PyMOL using an AMBER force-field. Molecular structures and surfaces were displayed and painted using the PyMOL molecular graphics system (Schrödinger, LLC).

ACKNOWLEDGEMENTS

We thank Mark Maciejewski (NMR Facility, University of Connecticut Health Center) for collecting and processing the NMR spectra. This study was supported by grant GM051293 (to SMK) from the National Institutes of Health.

REFERENCES

Blackburn K, Bustamante-Marin X, Yin W, Goshe MB, and Ostrowski LE (2017). Quantitative proteomic analysis of human airway cilia identifies previously uncharacterized proteins of high abundance. J Proteome Res 16, 1579–1592. [PubMed: 28282151]

Bui K, Sakakibara H, Movassagh T, Oiwa K, and Ishikawa T (2008). Molecular architecture of inner dynein arms *in situ* in *Chlamydomonas reinhardtii* flagella. J Cell Biol 183, 923–932. [PubMed: 19029338]

Bui K, Sakakibara H, Movassagh T, Oiwa K, and Ishikawa T (2009). Asymmetry of inner dynein arms and inter-doublet links in *Chlamydomonas* flagella. J Cell Biol 186, 437–446. [PubMed: 19667131]

Casey D, Yagi T, Kamiya R, and Witman G (2003). DC3, the smallest subunit of the Chlamydomonas flagellar outer dynein arm-docking complex, is a redox-sensitive calcium-binding protein. J Biol Chem 278, 42652–42659. [PubMed: 12920131]

Fliegauf M, Olbrich H, Horvath J, Wildhaber JH, Zariwala MA, Kennedy M, Knowles MR, and Omran H (2005). Mislocalization of DNAH5 and DNAH9 in respiratory cells from patients with primary ciliary dyskinesia. Am. J. Respir. Crit. Care Med 171, 1343–1349. [PubMed: 15750039]

Geer LY, Domrachev M, Lipman DJ, and Bryant SH (2002). CDART: protein homology by domain architecture. Genome Res 12, 1619–1623. [PubMed: 12368255]

Haimo LT, and Fenton RD (1988). Interaction of *Chlamydomonas* dynein with tubulin. Cell Motil Cytoskeleton 9, 129–139. [PubMed: 2965993]

Hoops HJ, and Witman GB (1983). Outer doublet heterogeneity reveals structural polarity related to beat direction in *Chlamydomonas* flagella. J Cell Biol 97, 902–908. [PubMed: 6224802]

Ide T, Owa M, King SM, Kamiya R, and Wakabayashi K (2013). Protein–protein interactions between intermediate chains and the docking complex of *Chlamydomonas* flagellar outer arm dynein. FEBS Letters 587, 2143–2149. [PubMed: 23747306]

Ishikawa T (2018). Organization of dyneins in the axoneme In: Dyneins: Structure, Biology and Disease. Volume 1 - The Biology of Dynein Motors, ed. King SM, Oxford, UK: Elsevier, Academic Press, pp.203–217.

King SM (2018). Composition and assembly of axonemal dyneins In: Dyneins: Structure, Biology and Disease. Volume 1 - The Biology of Dynein Motors, ed. King SM, Oxford, UK: Elsevier, Academic Press, pp.163–201.

King SM, and Patel-King RS (2015). The oliogmeric outer arm dynein assembly factor CCDC103 is tightly integrated within the ciliary axoneme and exhibits periodic binding to microtubules. J Biol Chem 290, 7388–7401. [PubMed: 25572396]

Lupas A, M VD, and J S (1991). Predicting coiled coils from protein sequences. Science 252, 1162–1164. [PubMed: 2031185]

Ma M, Stoyanova M, Rademacher G, Dutcher SK, Brown A, and Zhang R (2019). Structure of the decorated ciliary doublet microtubule. Cell 179, 909–922. [PubMed: 31668805]

Maurizy C, Quinternet M, Abel Y, Verheggen C, Santo PE, Bourguet M, C.F. Paiva A., Bragantini B, Chagot M-E, Robert M-C, Abeza C, Fabre P, Fort P, Vandermoere F, M.F. Sousa P., Rain J-C, Charpentier B, Cianférani S, Bandeiras TM, Pradet-Balade B, Manival X, and Bertrand E (2018). The RPAP3-C terminal domain identifies R2TP-like quaternary chaperones. Nature Communications 9, 2093.

Mészáros B, Erdős G, and Dosztányi Z (2018). IUPred2A: context-dependent prediction of protein disorder as a function of redox state and protein binding Nucleic Acids Research 46, W329–W337. [PubMed: 29860432]

Nicastro D, Schwartz C, Pierson J, Gaudette R, Porter ME, and McIntosh JR (2006). The molecular architecture of axonemes revealed by cryoelectron tomography. Science 313, 944–948. [PubMed: 16917055]

Oda T, Yanagisawa H, Kamiya R, and Kikkawa M (2014). A molecular ruler determines the repeat length in eukaryotic cilia and flagella. Science 346, 857–860. [PubMed: 25395538]

Owa M, Furuta A, Usukura J, Arisaka F, King SM, Witman GB, Kamiya R, and Wakabayashi K. -i. (2014). Cooperative binding of the outer arm-docking complex underlies the regular arrangement of outer arm dynein in the axoneme. Proc Natl Acad Sci USA 111 9461–9466 [PubMed: 24979786]

Panizzi J, Becker-Heck A, Castleman V, Al-Mutairi D, Liu Y, Loges NT, Pathak N, Austin-Tse C, Sheridan E, Schmidts M, Olbrich H, Werner C, Haffner K, Hellman NE, Chodhari R, Gupta A, Kramer-Zucker A, Olale F, Burdine R, Schier A, O'Callaghan C, Chung EMK, Reinhardt R, Mitchison H, King SM, Omran H, and Drummond I (2012). CCDC103 mutations cause primary ciliary dyskinesia by disrupting assembly of ciliary dynein arms. Nature Genetics 44, 714–719. [PubMed: 22581229]

Pazour G, Agrin N, Leszyk J, and Witman G (2005). Proteomic analysis of a eukaryotic flagellum. J Cell Biol 170, 103–113. [PubMed: 15998802]

Pazour GJ, and King SM (2019). Ciliary doublet microtubules at near atomic resolution. Cell 179, 805–807. [PubMed: 31675493]

Piperno G, and Ramanis Z (1991). The proximal portion of *Chlamydomonas* flagella contains a distinct set of inner dynein arms. J Cell Biol 112, 701–709. [PubMed: 1825211]

Shoemark A, Moya E, Hirst RA, Patel MP, Robson EA, Hayward J, Scully J, Fassad MR, Lamb W, Schmidts M, Dixon M, Patel-King RS, Rogers AV, Rutman A, Jackson CL, Goggin P, Rubbo B, Ollosson S, Carr S, Walker W, Adler B, Loebinger MR, Wilson R, Bush A, Williams H, Boustred C, Jenkins L, Sheridan E, Chung EMK, Watson CM, Cullup T, Lucas JS, Kenia P, O'Callaghan C, King SM, Hogg C, and Mitchison HM (2018). High prevalence of *CCDC103* p.His154Pro mutation causing primary ciliary dyskinesia disrupts protein oligomerisation and is associated with normal diagnostic investigations. Thorax 73, 157–166. [PubMed: 28790179]

Takada S, and Kamiya R (1994). Functional reconstitution of *Chlamydomonas* outer dynein arms from alpha- beta and gamma subunits: requirement of a third factor. J Cell Biol 126, 737–745. [PubMed: 8045937]

Wu H, Blackledge M, Maciejewski MW, Mullen GP, and King SM (2003). Relaxation-based structure refinement and backbone molecular dynamics of the dynein motor domain-associated light chain. Biochemistry 42, 57–71. [PubMed: 12515539]

Wu H, Maciejewski MW, Marintchev A, Benashski SE, Mullen GP, and King SM (2000). Solution structure of a dynein motor domain-associated light chain. Nature Struct Biol 7, 575–579. [PubMed: 10876244]

Yagi T, Uematsu K, Liu Z, and Kamiya R (2009). Identification of dyneins that localize exclusively to the proximal portion of *Chlamydomonas* flagella. J Cell Sci 122, 1306–1314. [PubMed: 19351714]

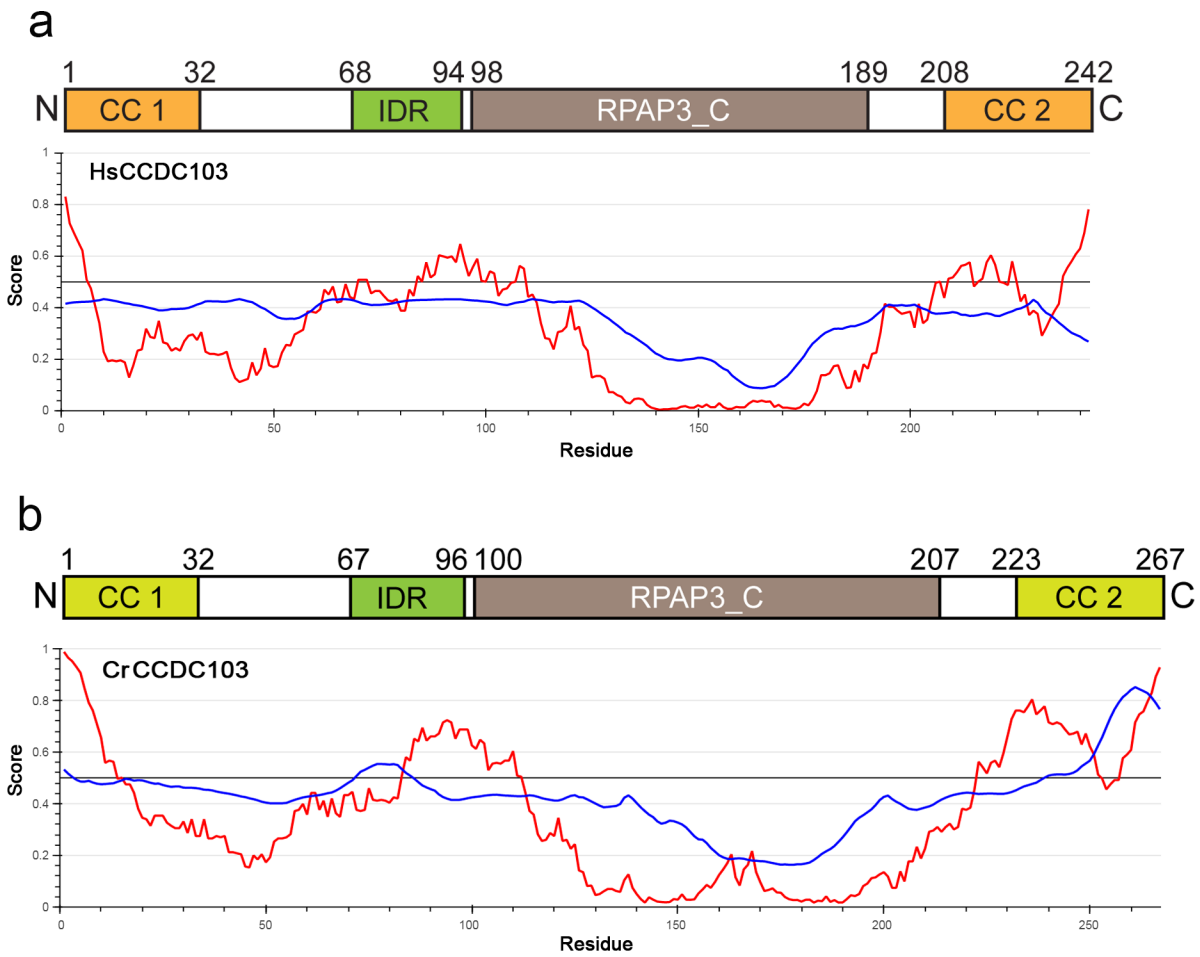

Fig. 1. Domain Organization of HsCCDC103 and CrCCDC103

a) Diagram of HsCCDC103 domain organization. The two confidently predicted coiled coils are shown in orange, and the RPAP3_C domain in brown. An inherently disordered region (shown in green) located immediately N-terminal of the RPAP3_C domain is predicted by the IUPred2A short disorder parameter algorithm (red plots); blue traces represent the ANCHOR2 context-dependent parameter. There is also prediction of a disordered region near the C-terminus which overlaps with the predicted coiled coil. b) A similar map and disorder plots for CrCCDC103. The N- and C-terminal helical domains are indicated in light green as, although apparently homologous to the coiled coils of HsCCDC103, neither is confidently predicted as a coiled coil. The RPAP3_C domain is enlarged compared to HsCCDC103 due to a 16-residue insertion.

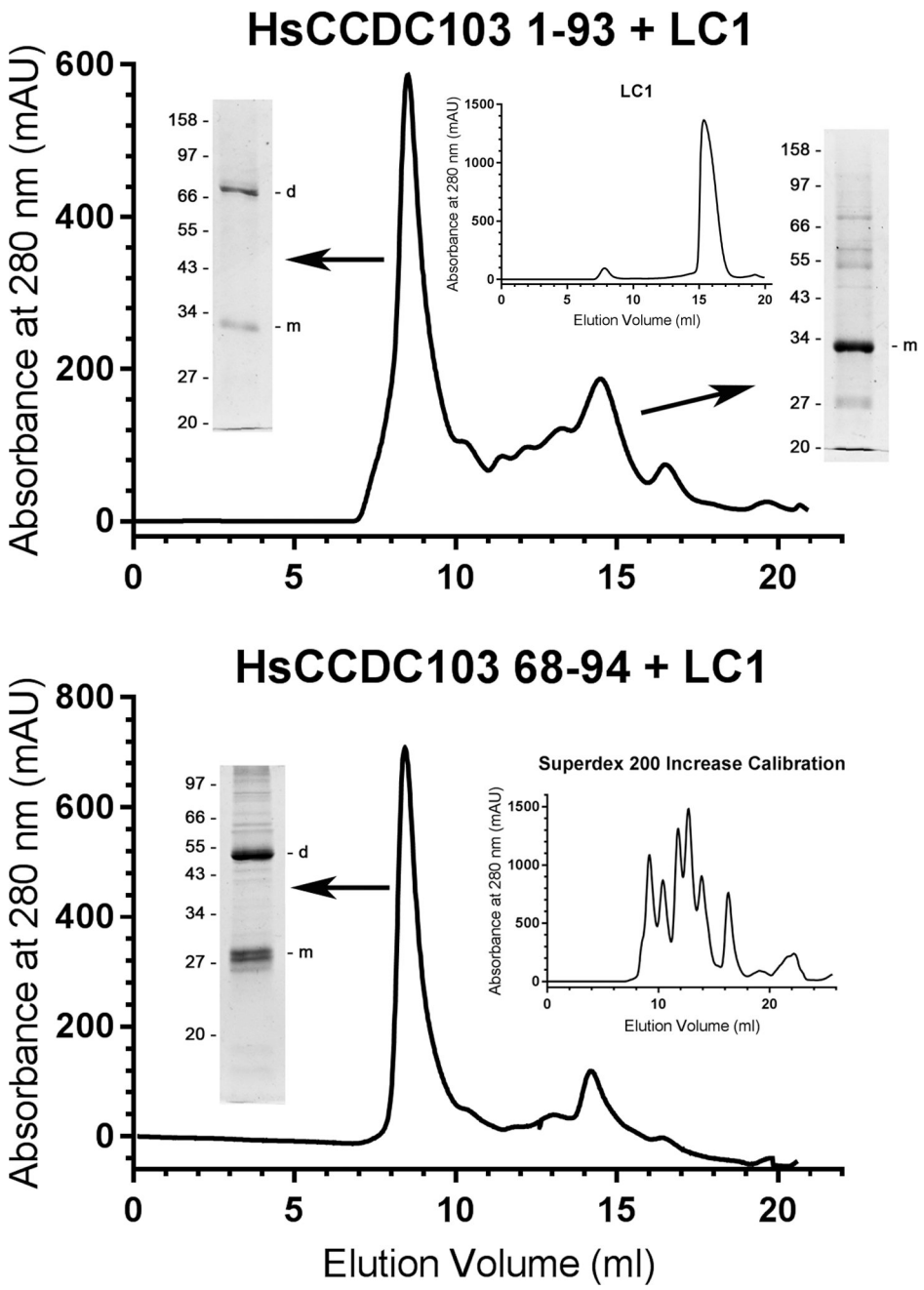

Fig. 2. The Inherently Disordered Region of HsCCDC103 Mediates Dimer Formation

HsCCDC103 regions 1–93 and 68–94 were fused to the N-terminus of the *Chlamydomonas* LC1 dynein light chain which on its own behaves solely as a monomer (*upper panel inset*) (Wu *et al.*, 2000; King and Patel-King, 2015). The fusion proteins were fractionated on a Superdex 200 Increase gel filtration column and samples from the peaks subject to SDS-PAGE. Both proteins gave high molecular mass peaks that were capable of forming detergent/heat resistant dimers. Previously, the N-terminal coiled coil was found to behave as a monomer (King and Patel-King, 2015). Thus, the 68–94 disordered region is sufficient to explain the unusual dimerization properties of HsCCDC103.

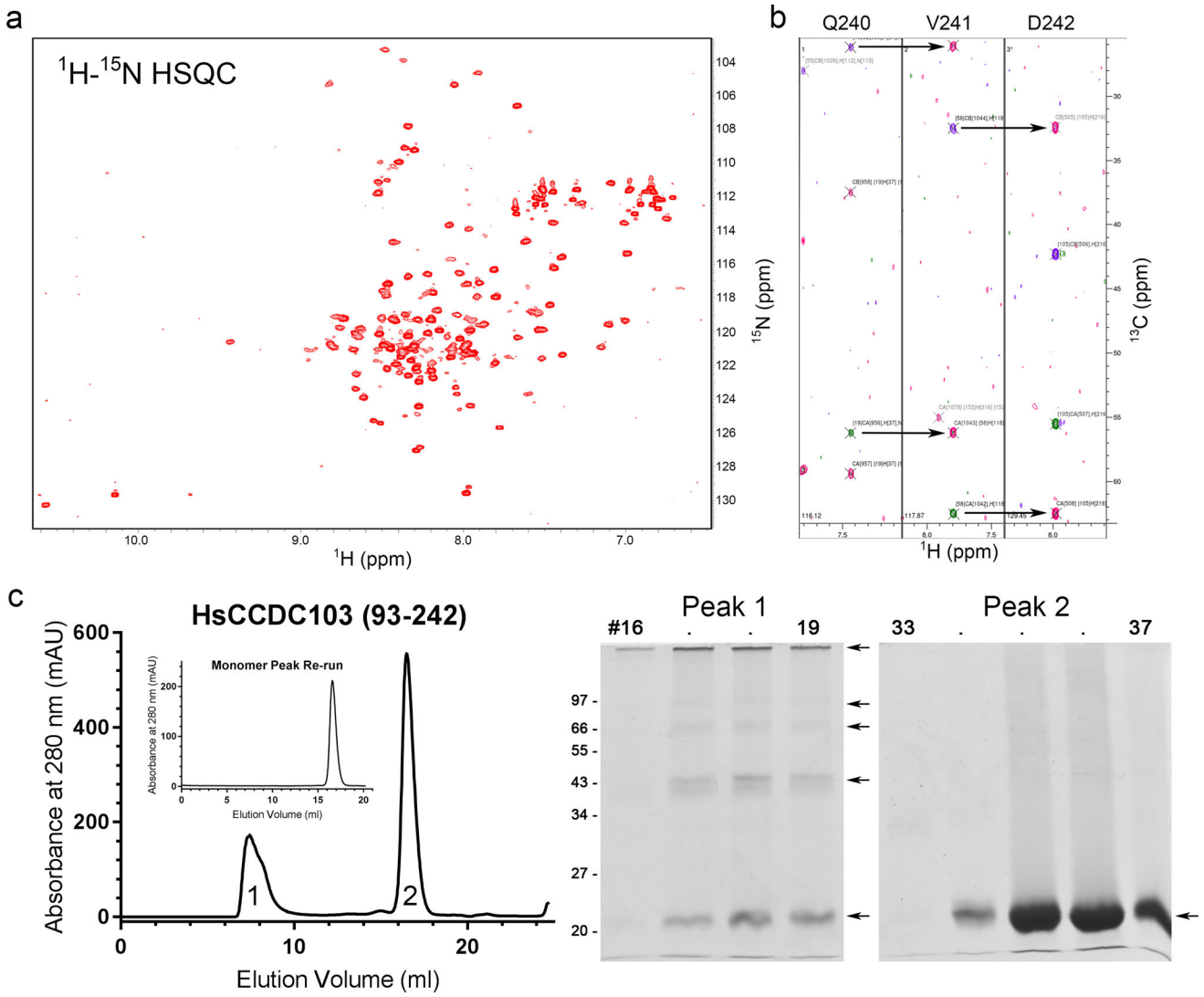

Fig. 3. The RPAP3_C Domain Plus C-terminal Region of HsCCDC103 Forms Oligomers

a) ^{1}H–^{15}N HSQC spectrum for HsCCDC103(93–242). The spectrum is resolved and dispersed indicative of a well-folded protein. b) Strips from the HNCACB spectrum illustrating the peak connectivities of the C_{α} and C_{β} atoms for the C-terminal residues Q240-D242; each strip shows the C_{α} and C_{β} peaks for the indicated residue and the preceding residue. c) The HsCCDC103(93–242) preparation was fractionated by gel filtration. Two peaks were observed that represent monomeric and oligomeric forms. When the monomer peak was re-chromatographed no oligomeric forms were observed. SDS-PAGE analysis revealed that while the monomer peak migrated solely as a monomer, the larger molecular mass peak consisted of some monomer and numerous oligomeric forms; dimers, trimers and tetramers are clearly resolved. These data suggest the 93–242 region of HsCCDC103 can adopt at least two distinct conformations only one of which is capable of oligomerization.

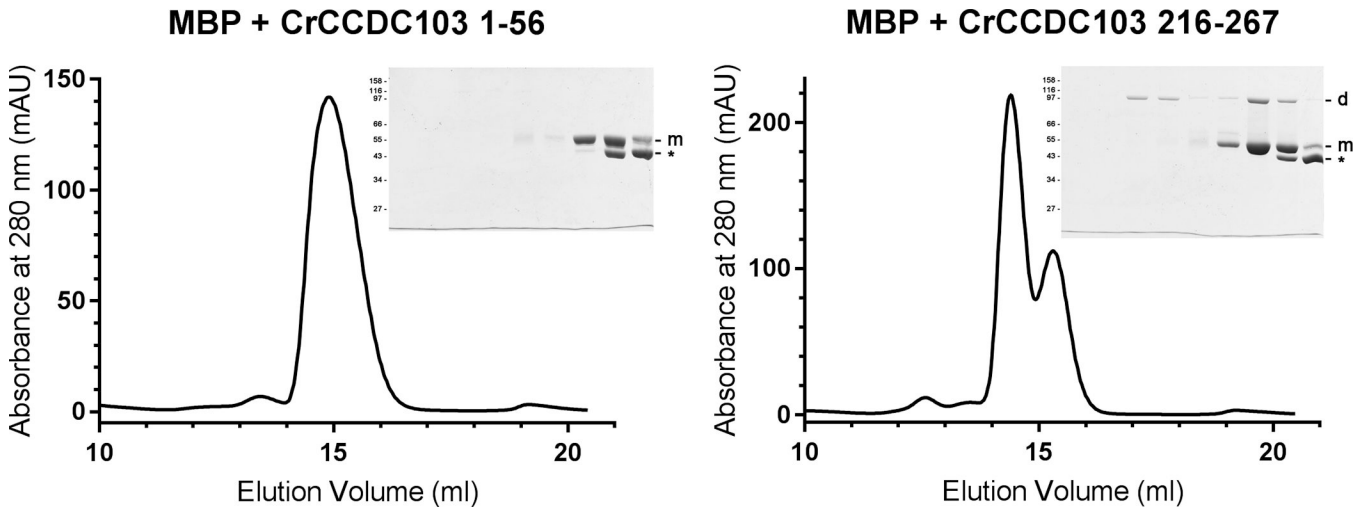

Fig. 4. CrCCDC103 Contains a C-terminal Dimerization Domain

Two segments of CrCCDC103 encoding the terminal domain residues 1–56 and 216–267 were fused to MBP and fractionated by gel filtration. The 1–56 fusion protein migrated as a single monomer peak. SDS-PAGE revealed the MBP-CrCCDC103 monomer band and also a proteolytic fragment (indicated by *). In contrast, the C-terminal CrCCDC103 fusion yielded three peaks. SDS-PAGE analysis revealed that the largest consisted only of a fusion protein dimer; the second contained both dimer and monomer and was clearly resolved from the third peak that was almost completely formed from the proteolytic fragment alone. Thus, this C-terminal region of CrCCDC103 is capable of dimer formation.

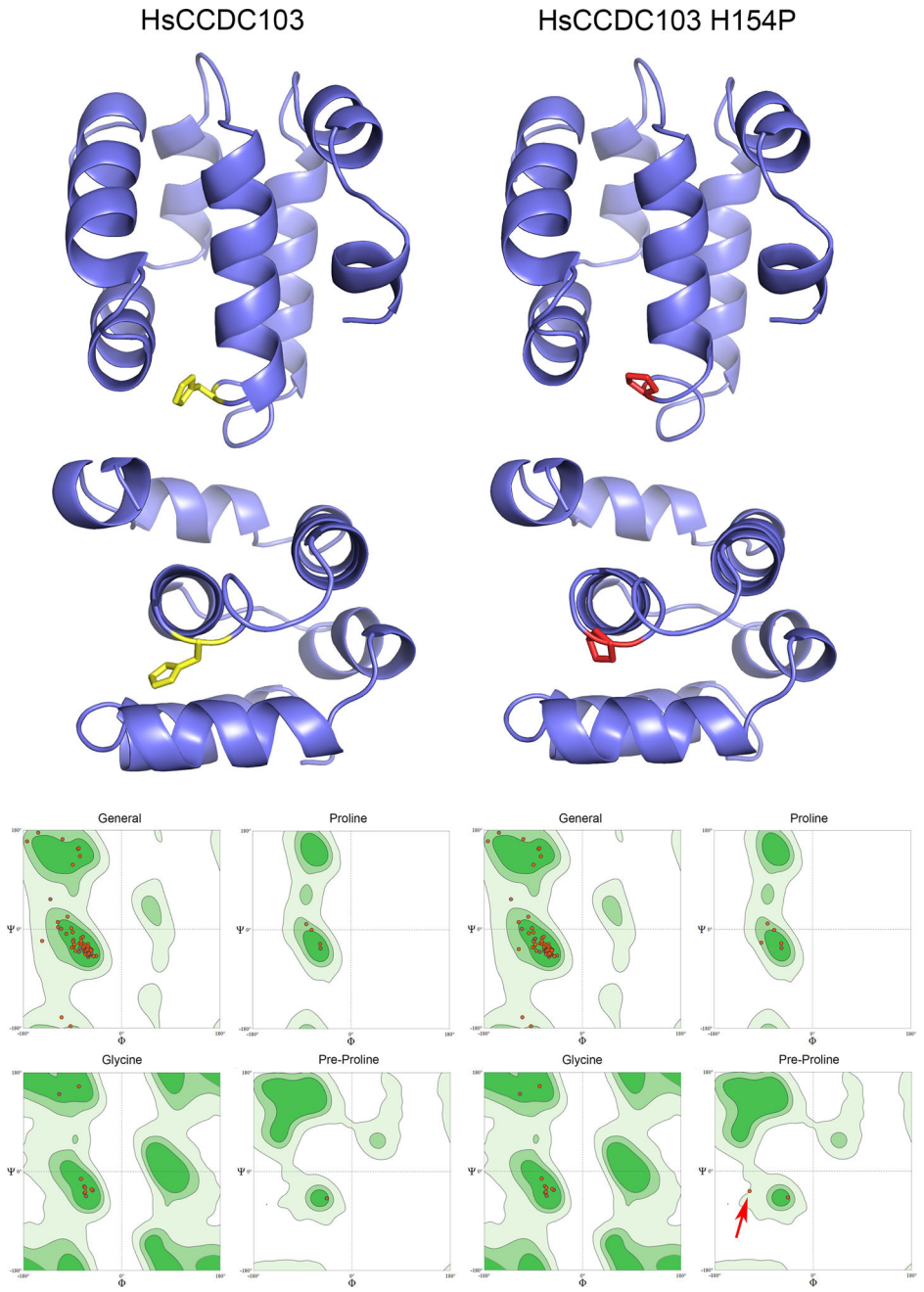

Fig. 5. Molecular Models of the RPAP3_C Domains from Wildtype and H154P Mutant HsCCDC103

Molecular models for the RPAP3_C domains of wildtype and H154P mutant HsCCDC103 were calculated using SWISSMODEL based on the NMR solution structure of this domain from RPAP3 (PDB 6EZ4). The domain consists of six tightly packed helices. Residue H154 is shown in yellow and when mutated to Pro in red. Ramachandran space plots correlating the Φ and Ψ angles in the peptide backbone for general residues, prolines, glycines and residues preceding proline (when not Gly) are shown for both models; light green indicates sterically allowable regions of the plots; dark green areas are favored regions. In the H154P model, with the exception of the connection between the Pro introduced by the H154P

mutation and its preceding residue (red arrow), all Φ/Ψ angles are in allowed and/or favorable regions.

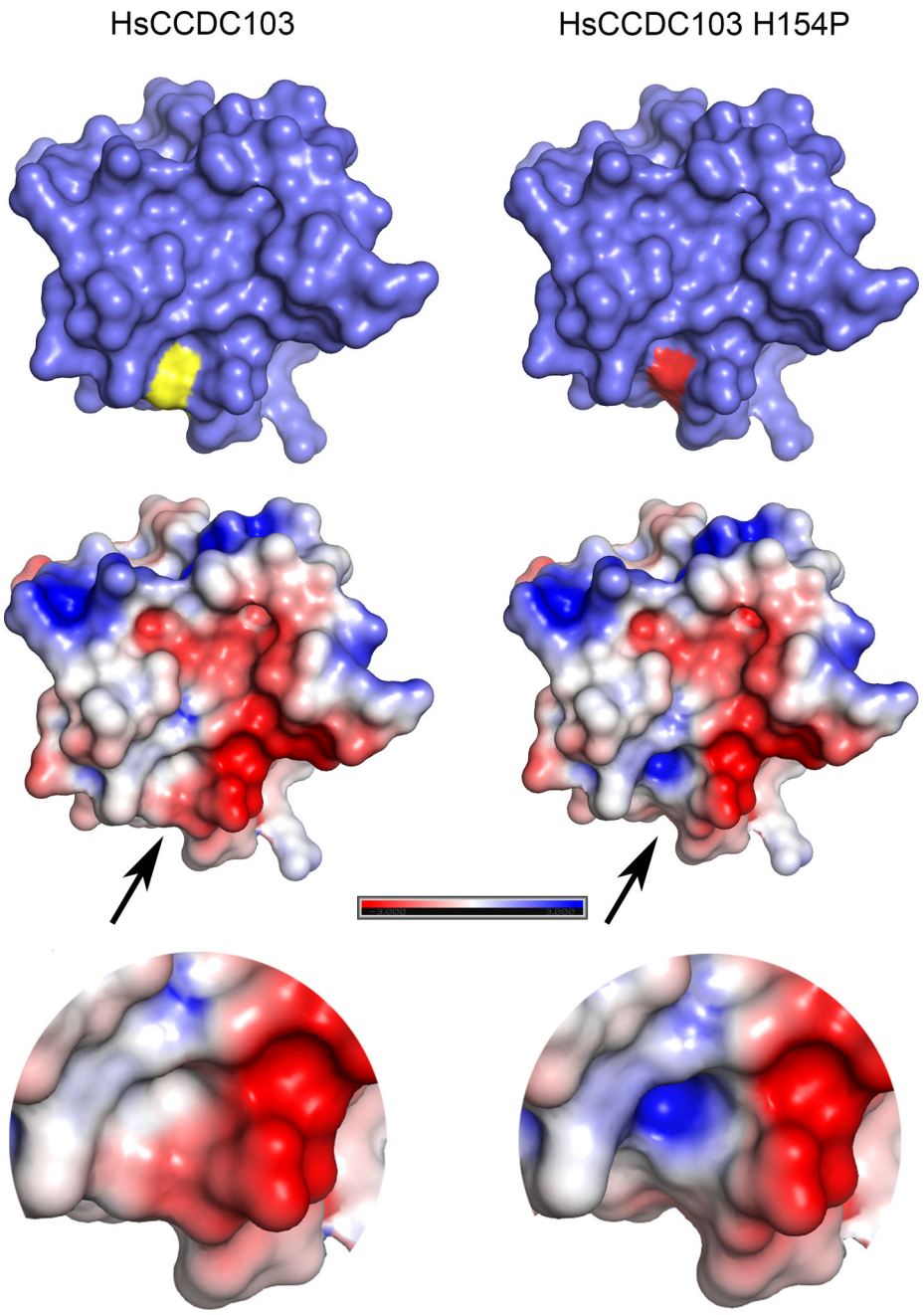

Fig. 6. Structural Consequences of the Pathogenic H154P HsCCDC103 Mutation

The van der Waal molecular surface for both homology models is shown. In the representations at top this reveals that H154 (yellow) is exposed and when converted to Pro (red) results in an indentation. The lower representations show the molecular surface painted with full Poisson-Boltzmann electrostatics calculations painted onto the surface; the color map indicates basic regions in blue, acidic regions in red and neutral/uncharged regions in white. This analysis identified a large acidic patch that extends to the location of H154. In the H154P mutant, the alteration in conformation leads to exposure of a basic patch within the indentation which might affect protein-protein interactions.

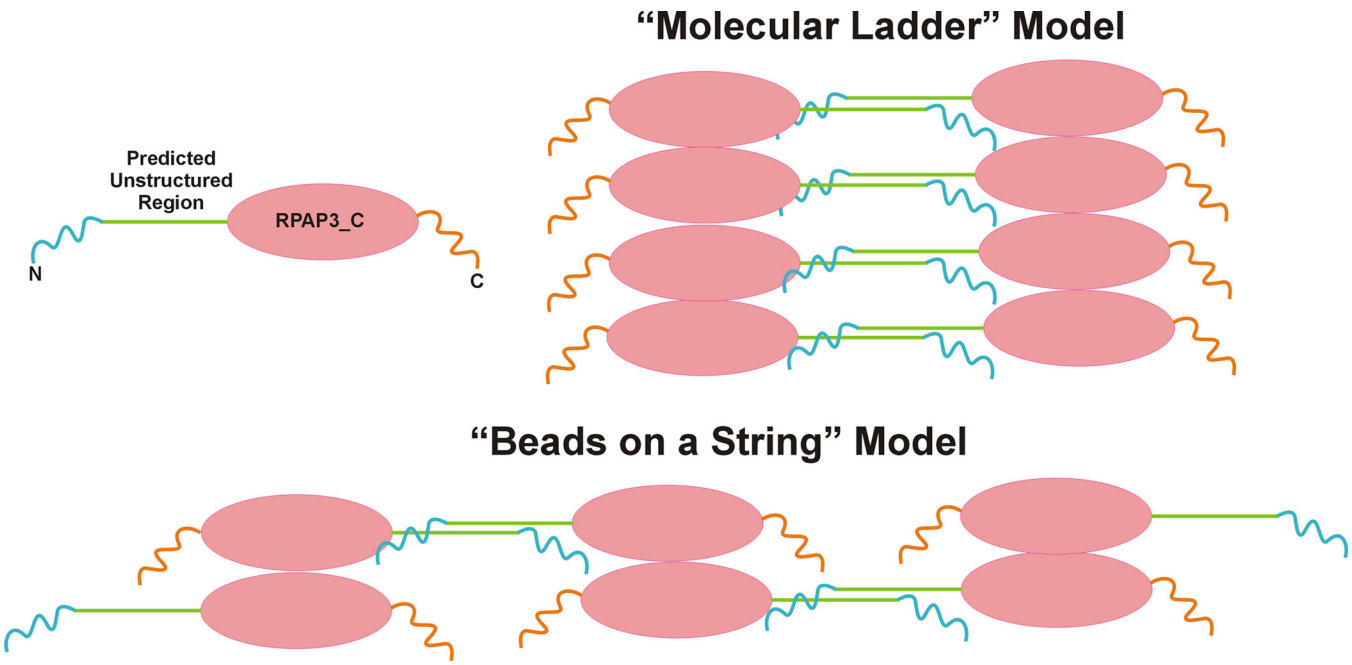

Fig. 7. Models for CCDC103 Molecular Scaffolds

Two possible mechanisms by which CCDC103 might oligomerize to form scaffolds. The “molecular ladder” model envisages oligomerization through the RPAP3_C domain/C-terminal regions combined with dimerization of two scaffolds through the inherently disordered regions. In contrast, the “beads on a string” model would be compatible with the presence of two dimerization interfaces and the ability of CrCCDC103 to form stable tetramers. However, in both cases CCDC103 oligomers would retain unassociated N- and C-terminal coiled coil domains which remain available for interactions with other axonemal components.