



OPEN ACCESS

EDITED BY

Maureen Wirschell,
McKendree University, United States

REVIEWED BY

Bénédicte Durand,
Université Claude Bernard Lyon 1, France
Dorota Włoga,
Polish Academy of Sciences, Poland

*CORRESPONDENCE

Gang Dong,
✉ gang.dong@meduniwien.ac.at

RECEIVED 07 October 2024

ACCEPTED 09 January 2025

PUBLISHED 13 February 2025

CITATION

Stadler A, Gabriel HB, De Liz LV, Alonso-Gil S, Deng X, Crickley R, Korbula K, Mikolaskova B, Vaughan S, Huang K, Žagrović B, Sunter JD and Dong G (2025) CFAP410 has a bimodular architecture with a conserved surface patch on its N-terminal leucine-rich repeat motif for binding interaction partners.

Front. Cell Dev. Biol. 13:1507470.
doi: 10.3389/fcell.2025.1507470

COPYRIGHT

© 2025 Stadler, Gabriel, De Liz, Alonso-Gil, Deng, Crickley, Korbula, Mikolaskova, Vaughan, Huang, Žagrović, Sunter and Dong. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

CFAP410 has a bimodular architecture with a conserved surface patch on its N-terminal leucine-rich repeat motif for binding interaction partners

Alexander Stadler^{1,2}, Heloisa B. Gabriel³, Laryssa V. De Liz³, Santiago Alonso-Gil^{1,4}, Xuan Deng⁵, Robbie Crickley³, Katharina Korbula^{1,2}, Barbora Mikolaskova^{1,2}, Sue Vaughan³, Kaiyao Huang⁵, Bojan Žagrović^{1,4}, Jack D. Sunter³ and Gang Dong^{1,2*}

¹Max Perutz Labs, Vienna Biocenter Campus (VBC), Vienna, Austria, ²Center for Medical Biochemistry, Medical University of Vienna, Vienna, Austria, ³Department of Biological and Medical Sciences, Oxford Brookes University, Oxford, United Kingdom, ⁴Department of Structural and Computational Biology, University of Vienna, Vienna, Austria, ⁵Institute of Hydrobiology, Chinese Academy of Sciences, Wuhan, Hubei, China

Cilia and flagella associated protein 410 (CFAP410) is a protein localized at the basal body of cilia/flagella and plays essential roles in ciliogenesis. Multiple single amino acid mutations in CFAP410 have been identified in patients. However, the molecular mechanism for how the mutations cause these disorders remains poorly understood due to a lack of high-resolution structures of the protein. Our studies demonstrate that CFAP410 adopts a bimodular architecture. We have previously reported our structural studies on the C-terminal domain (CTD) of CFAP410 from various organisms. Here we report a 1.0-Å resolution crystal structure of the N-terminal domain (NTD) of *Trypanosoma brucei* CFAP410. We further examined how the disease-causing mutations in this domain may affect the folding and structural stability of CFAP410. Our results suggest that the single-residue mutations in the CFAP410-NTD cause human diseases by destabilizing the structure that subsequently disrupts its interaction with other partners.

KEYWORDS

CFAP410, cilium, ciliopathies, flagellum, protein, structure

1 Introduction

Cilia- and flagella-associated protein 410 (CFAP410) was initially called C21orf2 as it is encoded by a gene which was found in open reading frame 2 of chromosome 21 (Krohn et al., 1997). CFAP410 was shown to localize to the basal body and interacts with NEK1 and SPATA7 during ciliogenesis and DNA damage repair (Wheway et al., 2015; Wang et al., 2017b; Fang et al., 2015). CFAP410 is well conserved across eukaryotes and is found in the genomes of all ciliated organisms. *Homo sapiens* CFAP410 (HsCFAP410, accession code: NP_004919) is 256 amino acids long and a number of single residue mutations, including I35F, C61Y, R73P, Y107C, Y107H, V111M and L224P, have been associated with skeletal

and/or retinal disorders (Whewey et al., 2015; Wang et al., 2016; McInerney-Leo et al., 2017; Khan et al., 2015; Suga et al., 2016). These disorders are genetically autosomal recessive and have a range of symptoms that include rib dysplasia, vertebral bodies, ilia, proximal femora, and retinitis pigmentosa (Ehara et al., 1997; Warman et al., 2011). However, it remains unknown how these point mutations in HsCFAP410 contribute to these disorders.

We have carried out extensive structural and biochemical characterizations on three homologs of CFAP410 from *H. sapiens*, *Trypanosoma brucei* and *Chlamydomonas reinhardtii*. Our work revealed that CFAP410 is a bimodular protein comprising two highly conserved domains, including a leucine rich repeat (LRR) motif at its N-terminus and an oligomeric helical bundle at its C-terminus. We have recently reported crystal structures for the C-terminal domain (CTD) of the three homologs of CFAP410 and how the disease-causing L224P mutation abolishes its localization to the basal body (Stadler et al., 2024). Here we report our further structural studies on the N-terminal domain (NTD) of *T. brucei* CFAP410 (TbCFAP410, accession code XP_011776929/Tb927.9.12170). The structure allows us to examine how the mutations might affect the folding of the protein. We further checked the localization of the wild-type and disease-causing mutations in *T. brucei* and analyzed its potential interaction *in silico* with NEK1 via a conserved surface patch. Taken together, our work provides a potential explanation how single-residue mutations in CFAP410 might cause ciliopathies.

2 Materials and methods

2.1 Cloning and site-directed mutagenesis of TbCFAP410

Sequence encoding full-length TbCFAP410 was amplified from *T. brucei* genomic DNA using PCR and ligated into the pET15b vector (Novagen). All truncated constructs were subsequently generated from the full-length construct using either pET15b or a custom vector that provides a maltose-binding protein (MBP) and His₁₀ tag to the N-terminus of the target proteins. The MBP-His₁₀ tag can be cleaved off by the tobacco etch virus (TEV) protease. All single- and multi-residue mutations of TbCFAP410-NTD were generated using the megaprimer-based mutagenesis strategy as reported (Tyagi et al., 2004). Incorporation of mutations was confirmed by DNA sequencing.

2.2 Protein expression and purification

All recombinant TbCFAP410 proteins were expressed in *Escherichia coli* (strain BL21-DE3). The transformed cells were grown in Luria-Bertani (LB) medium at 37°C to an OD₆₀₀ of 0.6–0.8, and then subjected to cold shock on ice for 30 min. Protein expression was induced by addition of 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG), and cell cultures were further incubated at 18°C overnight. Cells were harvested by centrifugation in a Sorvall GS3 rotor (6,000×g, 12 min, 4°C). Cell pellets were resuspended in 10 mL of lysis buffer [20 mM HEPES pH 7.5, 100 mM NaCl, 20 mM imidazole, 1 mM β-

TABLE 1 Data collection and refinement statistics.

Beamline	ESRF ID23-1
Wavelength (Å)	0.972
Resolution range (Å)	19.90 - 1.0 (1.04 - 1.0)
Space group	P 2 ₁ 2 ₁ 2 ₁
Unit cell <i>a</i> , <i>b</i> , <i>c</i> (Å) <i>α</i> , <i>β</i> , <i>γ</i> (°)	39.80, 52.96, 61.43 90, 90, 90
Total reflections	789,670 (30,477)
Unique reflections	69,533 (5,835)
Multiplicity	11.4 (5.2)
Completeness (%)	98.17 (83.75)
Mean I/σ(I)	10.84 (0.61)
R-merge	0.099 (1.695)
R-meas	0.104 (1.882)
R-pim	0.030 (0.792)
CC _{1/2}	0.998 (0.386)
CC*	1.00 (0.746)
Reflections used in refinement	69,519 (5,834)
Reflections used for R-free	1,989 (168)
R-work	0.152
R-free	0.162
Number of heavy atoms in proteins	1,302
Number of solvent	157
RMS (bonds)	0.010
RMS (angles)	1.22
Ramachandran favored (%)	96.23
allowed (%)	3.77
outliers (%)	0.00

Statistics for the highest-resolution shell are shown in parentheses.

Mercaptoethanol, 1 mM phenylmethylsulfonyl fluoride (PMSF)] per L of cell culture.

Resuspended cells were lysed in an EmulsiFlex-C3 homogenizer (Avestin) at 40 bars and cell debris was pelleted by centrifugation (40,000 × g, 40 min, 4°C). The supernatant was filtered (0.45-μm pore size, Amicon) and loaded onto a Ni-HiTrap[®] HP column pre-equilibrated with the same lysis buffer in an ÄKTA[™] Pure chromatography system (Cytiva). The column with bound proteins was washed with 2× column volume (cv) of the same lysis buffer. Bound protein was subsequently eluted using a linear gradient concentration of imidazole (20–500 mM, 20 × cv) in the lysis buffer.

The N-terminal MBP-His₁₀ tag of all TbCFAP410 truncated proteins were cleaved off by incubating the pooled fractions of interest (20–40 mL) with approximately 1% (w/w) of the TEV protease in a dialysis bag against 2–4 L of dialysis buffer containing 20 mM HEPES (pH 7.5) and 100 mM NaCl (4°C, overnight). The dialyzed samples were loaded on a Ni-HiTrap[®] HP column to get rid of the cleaved tag and any uncut proteins.

Target proteins in the flow-through were further purified on a Superdex-200 16/60 column (Cytiva) with a running buffer containing 20 mM HEPES (pH 7.5), 100 mM NaCl, and 1 mM dithiothreitol (DTT). The eluted proteins were then pooled and concentrated, and final concentration was determined using Ultraviolet (UV) absorbance at 280 nm. The samples were either used freshly or stored at -80°C for subsequent experiments.

2.3 Crystallization and structural determination

Purified TbCFAP410-NTD proteins were subjected to extensive crystallization trials using several commercial crystallization kits from Hampton Research. Conditions of the initial hits were further optimized to obtain single crystals. The final condition for crystallizing TbCFAP410-NTD contained 0.1 M sodium citrate (pH 5.5) and 1 M $\text{NH}_4\text{H}_2\text{PO}_4$ at 22°C . X-ray diffraction data were collected at the ESRF synchrotron site and processed using XDS (Kabsch, 2010). Crystal structure was determined by the molecular replacement method using a homology model of TbCFAP410 (aa16-138) generated by SWISS-MODEL (Waterhouse et al., 2018) based on the crystal structure of the *C. reinhardtii* axonemal dynein light chain-1 (PDB code: 6L4P) (Toda et al., 2020). The resulting model was optimized by multiple rounds of manual rebuilding in COOT (Emsley and Cowtan, 2004) and refinement in Phenix (Adams et al., 2010), and finally validated in MolProbity (Chen et al., 2010). Details of data collection and refinement are summarized in Table 1.

2.4 Generation of constructs for *in vivo* assays in *Trypanosoma brucei*

DNA sequence of TbCFAP410 was recoded in wild-type and all mutations for *in vivo* assays. The recoded DNA sequences were chemically synthesized (Twist Bioscience) and cloned into the plasmid pDEX577-mNG. All constructs were subsequently linearized with Not I and then electroporated using a BTX Gemini Twin Wave with 3 pulses of 1.7 kV for 100 μs as previously described (Dean et al., 2015). Cells were recovered in SDM-79 for 8 h before selection with phleomycin (5 $\mu\text{g}/\text{mL}$).

2.5 Cell culture of *Trypanosoma brucei*

SmOxP9 procyclic *T. brucei* expressing TbSAS6 endogenously tagged with mScarlet were used for all experiments (Atkins et al., 2021). These were derived from TREU 927 strain, expressing T7 RNA polymerase and tetracycline repressor (Poon et al., 2012) and were grown at 28°C in SDM-79 medium supplemented with 10% FCS (Brun and Schonberger, 1979). Cell concentration was determined in a Z2 Coulter Counter particle counter.

2.6 Fluorescence microscopy

Cells were induced overnight with doxycycline (0.5 $\mu\text{g}/\text{mL}$) and harvested by centrifugation at 800 g for 5 min. For live-cell

microscopy, cells were washed two times in PBS supplemented with 10 mM glucose and 46 mM sucrose (vPBS). In the last wash, the DNA was stained using 10 $\mu\text{g}/\text{mL}$ Hoechst 33342 (Sigma-Aldrich), re-suspended in vPBS and then mounted with a coverslip and immediately imaged. For cytoskeletons, cells were washed in PBS, settled onto glass slides for 5 min and treated with 1% (v/v) NP-40 in PEME for 1 min. Cytoskeletons were then fixed in -20°C methanol for 20 min and rehydrated in PBS for 10 min. After fixation, the DNA was stained with 20 $\mu\text{g}/\text{mL}$ Hoechst 33342 in PBS, washed in PBS for 10 min or overnight and mounted before imaging. Images were taken using a Zeiss Axio Imager.Z1 microscope equipped with an ORCA-Flash 4.0 CMOS camera using a Plan-Apochromat 63 \times /1.4 NA oil objective. Images were acquired and analysed with ZEN 2 PRO software. For each induction 2 sets of 100 cytoskeletons were counted for each induced cell line using ImageJ (NIH, MD). To measure the CFAP410 basal body fluorescence signal intensity, two sets of 30 cytoskeletons were analyzed using ImageJ. For each cell, a circle was drawn around the basal body region and the signal intensity measured. The same area was used to measure the signal within the anterior region of cell (background). The background signal was subtracted from its corresponding basal body signal. The mean values for each mutant were compared to the wild type using a Student t-test.

2.7 Western blotting

After 24 h induction with doxycycline (0.5 $\mu\text{g}/\text{mL}$), parasites were harvested by centrifugation at $800 \times g$ for 3 min and washed twice with vPBS. Parasites were resuspended in an equal volume of vPBS and $2 \times \text{Laemmli}$ buffer to a final concentration of 2×10^8 cells/mL, heated at 95°C for 5 min and stored at -20°C . An amount equivalent to 1×10^6 parasites was loaded onto a 10% (w/v) SDS-PAGE gel and electrophoresed before being transferred to a PVDF membrane. The membrane was blocked with 5% (w/v) milk in Tris-buffered saline with Tween-20 (TBS-tween) overnight at 4°C , followed by incubation with BB2 (anti-Ty) (1:100 in 5% milk) or L8C4 (anti-PFR2) (1:100 in 5% milk) for 1 h at room temperature. After incubation, the membrane was washed five times for 5 min with TBS-tween, followed by incubation with the secondary antibody anti-mouse HRP (1:5,000 in 5% milk) for 1 h at room temperature. Finally, the membrane was washed five times, incubated with the Chemiluminescent Horseradish Peroxidase substrate for 2 min, and visualized using a G:BOX Chemi XRQ machine (Syngene).

2.8 *Chlamydomonas reinhardtii* strain and culture growth

Chlamydomonas reinhardtii strain CC-5325 was grown photoheterotrophically in tris-acetate-phosphate (TAP) medium at 25°C under constant illumination at 50 $\mu\text{mol photons}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$.

2.9 SpCas9 protein purification and *in vitro* transcription of sgRNA

SpCas9 protein purification were performed as described Xue et al. (2019). The sgRNA target site (GTGAAGAACCCTCAACCTC

TGGGG) for the CrCFAP410 gene (Cre14.g625550) in *C. reinhardtii* was selected using CRISPR RGEN Tools (<http://www.rgenome.net/cas-designer/>). Transcription of sgRNA was accomplished using the HiScribe™ T7 Quick High Yield RNA Synthesis Kit (NEB) and Monarch® RNA Cleanup Kit (NEB).

2.10 *In vitro* digestion of target DNA with RNP

Cells were picked into 50 μ L 5% (w/v) Chelex100, heated to 95°C for 10 min and then on the ice for 1 min to generate crude genomic DNA. The template DNA contained the Cas9 cleavage site was amplified using primers (PAM-F: 5'-TGTCACCTTCAC CTGCTGCGA-3'; PAM-R: 5'-CCAAGTTCGGCAGCTACAGT-3') with the crude genomic DNA as a template. *In vitro* cleavage assays were according to the previous protocol with the following modifications (Shin et al., 2016). 500 ng gRNA and 700 ng Cas9 were preincubated in Cas9 cleavage buffer (20 mM HEPES pH 7.5, 100 mM NaCl, 5 mM MgCl₂, 0.5 mM DTT, 0.1 mM EDTA) at 37°C for 15 min, then 500 ng template DNA was added, and the mixtures were incubated at 37°C for 1 h, then 10 \times STOP solution (0.5 M EDTA, 80% (v/v) glycerol, and 10% (w/v) SDS) was added. The final products were separated using 2% (w/v) agarose gel electrophoresis.

2.11 Preparation of dsDNA donor containing hygromycin resistance cassette with microhomologies arms for targeting CrCFAP410

The hygromycin resistance cassette was amplified from the plasmid pHK266 (Cheng et al., 2017). The 20 bp microhomologies based on the flank sequence of the Cas9 cleavage site were added to both sides of the hygromycin resistance cassette by PCR using primers (donor-F: 5'-GAGGTGAAGAACCTCAACCTTCTTCTTGCGCT ATGACACT-3'; donor-R: 5'-CGTCCAGGTCCTGGCCCCAGT GGAGAAAGAGGCCAAAATC-3'). The PCR product was purified using Gel Extraction Kit (QIAGEN).

2.12 Transformation of *Chlamydomonas reinhardtii* cells

Cells were cultured for 3 days to a cell density of 2×10^6 cells/mL in TAP medium. The cells were then concentrated to 2×10^8 cells/mL in TAP medium with 60 mM sorbitol. For each shot, Cas9 protein (50 μ g, 0.53 nmol) and sgRNA (1.6 nmol) as 1:3 M ratio at 37°C for 15 min to assemble RNP complexes. 20 μ L RNP and 2 μ g dsDNA donor were added to 80 μ L cell cultures in the 4 mm cuvette. Electroporation was conducted using BTX ECM® 630 Electroporation System (600 V, 50 μ F, 350 Ω). Immediately after electroporation, 600 μ L of TAP medium with 60 mM sorbitol was added. Cells were recovered overnight (24 h) in 10 mL TAP with 60 mM sorbitol and plated onto TAP media supplemented with 20 μ g/mL hygromycin. Colonies appeared after 5–7 days.

2.13 Genotyping and sequencing of potential mutants

To determine whether the dsDNA donor inserted into the cleavage site, the crude genomic DNA from colonies was used as templates, and PAM-F and PAM-R were used as PCR primers. The PCR products were then analyzed using 1% (w/v) agarose gel. For sequencing of CrCFAP410 mutations, about 2,188 bp PCR bands were gel purified and sequenced using the primers PAM-F and PAM-R.

2.14 Flagellar percentage and length measurements

The methods of flagellar length measurements and percentage of the flagellated cells were described previously (Wang et al., 2019). For each sample, at least 100 cells and 200 flagella were measured.

2.15 Molecular dynamics (MD) simulations

Atomistic models for wild-type TbCFAP410-NTD and all mutants (V37F, C63Y, R75P, Y107C, Y107H and V111M) were constructed starting from the crystal structure by *in silico* mutagenesis using Pymol (<http://www.pymol.org>). All further calculations were performed using GROMACS 2019 package (Van Der Spoel et al., 2005) and Amber99SB-ILDN force-field parameters (Lindorff-Larsen et al., 2010). In all cases, the simulated system was placed into a $7 \times 7 \times 7$ -nm³ octahedral box, energy minimized and solvated with TIP3P water molecules. Additionally, Na⁺ and Cl⁻ ions were added to achieve electroneutrality at the final salt concentration of 150 mM. The complete system was then energy-minimized using position restraints placed on protein Ca-atoms, and subjected to 5,000 steps of NVT molecular dynamics (MD) with a 0.5-fs time step and subsequent 250,000 steps of NPT MD with a 1-fs time step. After this initial equilibration, a 1 μ s NPT production MD run was simulated (5×10^8 steps with a 2-fs time step) for each system. All analysis was performed on the final 750 ns of each simulated trajectory. A twin-range (10 Å, 10.5 Å) spherical cut-off function was used to truncate the van der Waals interactions. Electrostatic interactions were treated using the particle-mesh Ewald summation (real space cut-off of 10-Å and 1.2-Å grid with fourth-order spline interpolation). MD simulations were carried out using three-dimensional periodic boundary conditions, in the isothermal-isobaric (NPT) ensemble with an isotropic pressure of 1.013 bar and a constant temperature of 295 K. The pressure and temperature were controlled using a Nose-Hoover thermostat (Nose, 1984; Hoover, 1985) and a Parrinello-Rahman barostat with 0.5-ps and 10-ps relaxation parameters, respectively, and a compressibility of 4.5×10^{-5} bar⁻¹ for the barostat.

3 Results

3.1 Crystal structure of TbCFAP410-NTD reveals a canonical LRR motif with a highly conserved surface patch

Our bioinformatic analysis suggests that all CFAP410 orthologs have a bimodular organization consisting of

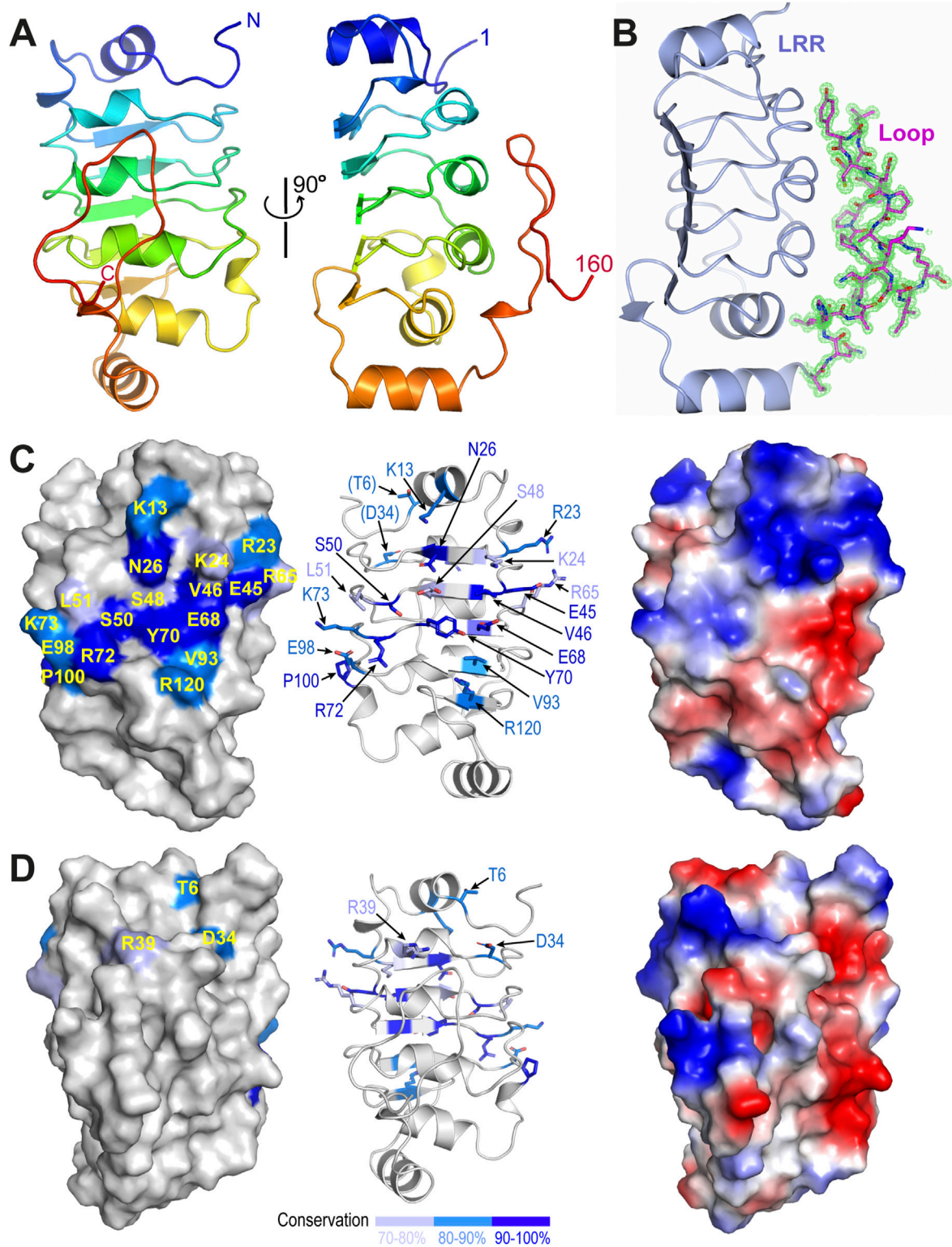


FIGURE 1

Crystal structure of the N-terminal LRR motif of TbCFAP410. **(A)** Ribbon diagram of TbCFAP410-NTD (aa1-160) color-ramped from blue (N-terminus) to red (C-terminus). **(B)** Cartoon of TbCFAP410-NTD with its C-terminal loop (aa141-160) shown as sticks together with $2F_o - F_c$ electron density map contoured at 1.5σ level. **(C)** A conserved surface patch on the flat β sheet of the LRR motif. Conserved surface residues are colored in blue, marine and light blue based on their sequence identities. **(D)** The opposite side of the LRR motif covered by the loop is not conserved. An electrostatic surface plot in the same orientation is shown at the right side in both **(C, D)**.

two folded domains that are connected by a disordered loop. We have previously determined high-resolution crystal structures for the CTD of three CFAP410 proteins from *H. sapiens*, *T. brucei* and *C. reinhardtii* (Stadler et al., 2024). Here, we report our further investigation on the NTD of CFAP410. We made multiple truncation constructs based on our bioinformatic analysis and tried different fusion tags, including His₆, SUMO, and maltose-binding protein (MBP). No soluble proteins could be obtained for the constructs of HsCFAP410 and CrCFAP410. For TbCFAP410 we managed to produce soluble proteins for four of the six truncations we generated, which cover residues 1–177, 1–170, 1–164 and 1–160, whereas two other shorter constructs spanning residues 1–149 and 1–140 were insoluble upon overexpression in bacteria.

Purified proteins of all these four soluble TbCFAP410-NTD truncations were subjected to extensive crystallization screenings. In the end we obtained crystals only for the construct spanning residues 1–160, which is what NTD refers to in all subsequent structural analyses of TbCFAP410. Single crystals of TbCFAP410-NTD diffracted X-rays to approximately 1.0 Å resolution at the ESRF synchrotron site (Table 1). The crystal structure was determined by the molecular replacement method using a homology model of residues 16–138 of TbCFAP410 generated by SWISS-MODEL (Waterhouse et al., 2018) based on the crystal structure of the *C. reinhardtii* axonemal dynein light chain-1 (PDB code: 6L4P) (Toda et al., 2020).

The refined model shows that all residues in the construct could be traced and confidently built (Figure 1A). Residues 1–138 fold into a canonical leucine-rich repeat (LRR), with five tandem repeats forming a calzone-like structure. The LRR motif contains five parallel arranged β strands forming a flat sheet on one side and an array of four short α helices on the other. The space between the α helices and the β strands constitutes the hydrophobic core of the LRR motif and is packed with many highly conserved leucine residues. Additionally, there is a short α helix on either end of the LRR motif that acts as a cap and is oriented roughly perpendicular to the β sheet. The overall structure is slightly bent toward the β sheet side as seen in other LRR structures (Kobe and Deisenhofer, 1995). Additionally, we found that residues 139–160 of TbCFAP410 forms a long hairpin-like loop, which is tightly packed on top of the helical surface of the LRR motif (Figure 1B). The loop makes extensive contacts with the LRR motif via six hydrogen bonds and numerous hydrophobic interactions (Supplementary Figure S1A). Those residues in the loop involved in its interaction with the LRR motif are highly conserved in all homologs from various kinetoplastids, so as the interacting residues on the LRR side (Supplementary Figures S1B–D).

Sequence alignments of CFAP410 homologs from mammals, trypanosomes and algae reveal an extensive surface patch on the flat β sheet of the LRR motif. The patch includes eighteen residues that are highly conserved in all homologs across these phyla (Figure 1C; Supplementary Figures S1A, B). The opposite side of the LRR motif is covered by a long hairpin-like loop in TbCFAP410 as shown above. Although this loop is conserved in kinetoplastids, it is essentially devoid of any conserved residues in all other homologs from animals and algae (Figure 1D; Supplementary Figure S2C).

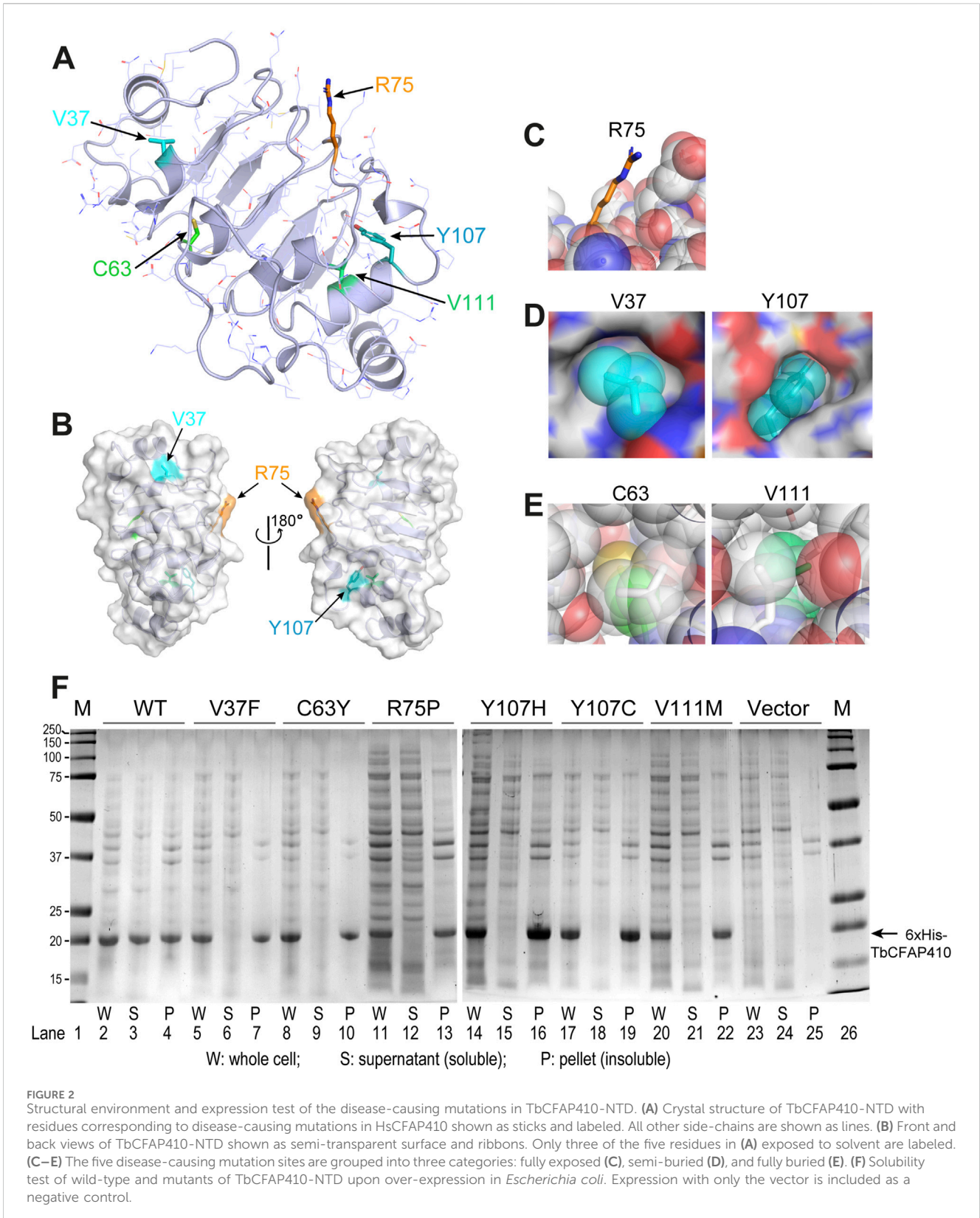
3.2 Examination of disease-causing mutations on CFAP410-NTD

Multiple single-residue mutations in HsCFAP410 have been identified in patients. Except for L224P located at the CTD, all other mutations are in the N-terminal LRR motif of HsCFAP410, including I35F, C61Y, R73P, Y107C, Y107H and V111M (Khan et al., 2015; Suga et al., 2016; Wheway et al., 2015; Wang et al., 2016; McInerney-Leo et al., 2017). Sequence alignment shows that the LRR motif is highly conserved in animals, trypanosomes and algae (Supplementary Figure S2A). Except for R73, all sites associated with disease-causing mutations are highly conserved hydrophobic residues. The LRR motifs of *H. sapiens* and *T. brucei* CFAP410 share 41% identity and 60% similarity (Supplementary Figure S3A). Comparison of the crystal structure of TbCFAP410-NTD with structural models of HsCFAP410 (aa1-130) generated by SWISS-MODEL (template-based) (Waterhouse et al., 2018) and AlphaFold2 (*ab initio*) (Jumper et al., 2021) shows that they all share a highly similar three-dimensional structure (Supplementary Figure S3B). Particularly, all the five disease-causing mutation sites mentioned above adopt essentially the same conformation in these structures. Statistics reports show that both predicted models of HsCFAP410 were built with high confidence (Supplementary Figures S3C–E). These allow us to examine how the mutations of single residues in HsCFAP410 might affect its structure and function based on the crystal structure we have determined for TbCFAP410-NTD.

Location of individual residues corresponding to the five disease-causing mutations in HsCFAP410 can be clearly visualized in TbCFAP410 (Figures 2A, B). These residues can be grouped into three categories, including one surface residue that is fully exposed to solvent (R75), two semi-exposed residues (V37 and Y107), and two hidden residues that are completely buried in the hydrophobic core of the structure (C63 and V111) (Figures 2C–E).

Given that most of these disease-causing mutants are hydrophobic and partially or fully buried in the structure, we therefore examined whether these mutations affect the folding or structural stability of the protein. Wild-type and mutant TbCFAP410-NTD proteins with an N-terminal His₆ tag were expressed in *E. coli*. After breaking the cells and spinning down the cell lysate, we checked the soluble fractions in the supernatant and the insoluble parts in the pellet. Our data show that wild-type TbCFAP410-NTD were present in both fractions, suggesting that it is at least partially soluble (Figure 2F, lanes 2–4). In contrast, however, all the six mutants were present only in the pellets but not in the supernatants (Figure 2F, lanes 5–22), demonstrating that the proteins are insoluble and thus likely mis-folded.

We further examined the crystal structure trying to understand how these mutations may cause changes in the folding of the protein. Firstly, mutation R73P of HsCFAP410, equivalent to residue R75 in TbCFAP410, is located in a short linker tightly packed between an α helix and a β strand in the structure (Figure 3A). We found that its fully solvent exposed side-chain forms a stacking interaction with that of residue E53 via two salt bridges. Additionally, the amide proton of residue R75 forms a hydrogen bond with the carbonyl group of residue L54. Mutating R75 to proline would disrupt all these interactions. Furthermore, the circular side-chain of proline in the mutant would additionally cause clashes with the neighboring



residues in the tightly packed region of the structure, and thus destabilizes the structure of the protein.

All the other five ciliopathy mutation sites within HsCFAP410-NTD occur at highly conserved hydrophobic residues

(Supplementary Figure S2A). Our structural analyses on the equivalent residues in TbCFAP410-NTD show that residue V37 (corresponding to I35 in HsCFAP410) is packed tightly against a small hydrophobic cleft. Its replacement by phenylalanine (i.e., I35F

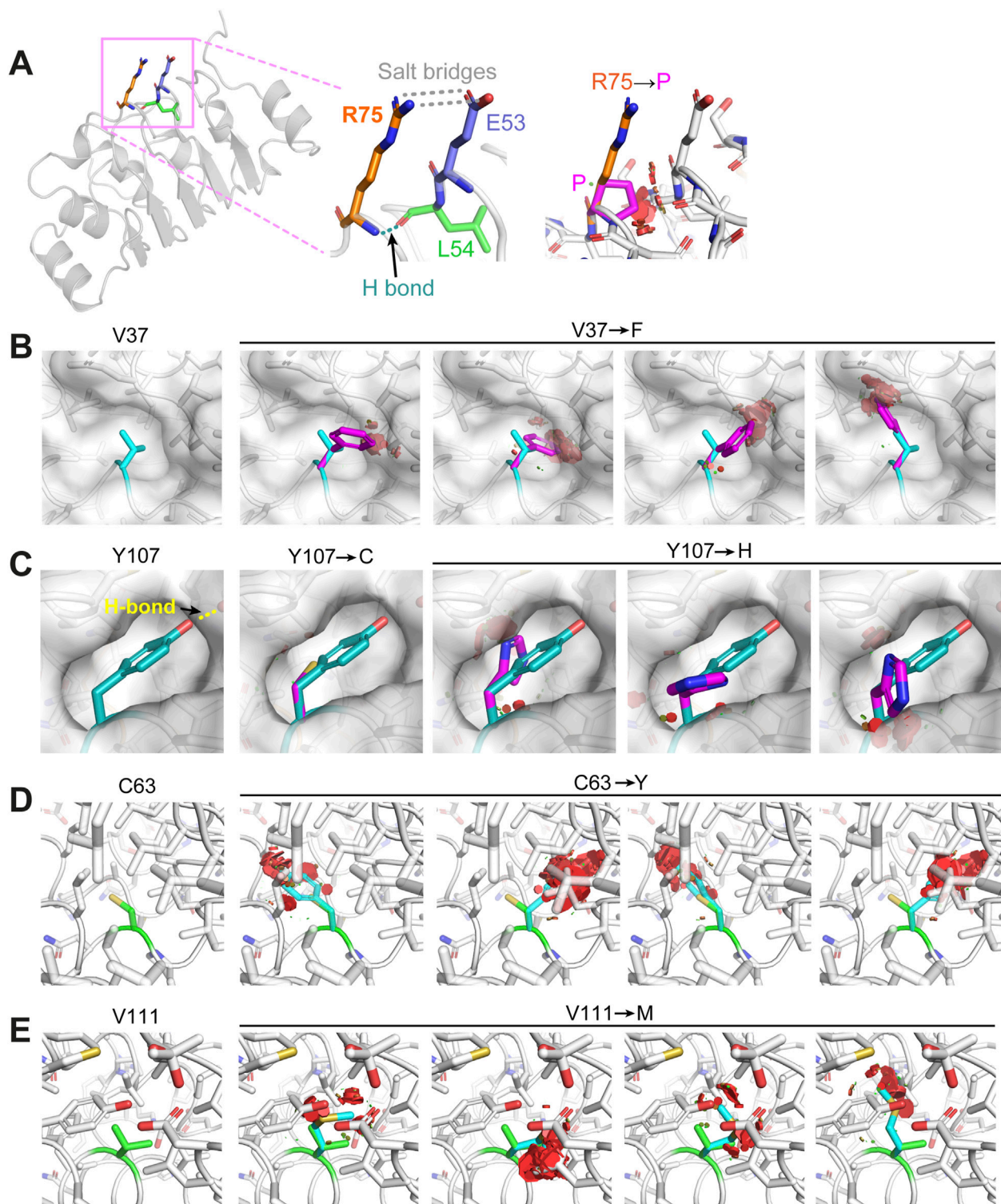


FIGURE 3
Structural examination of how the disease-causing mutants on TbCFAP410-NTD may affect protein folding or structural stability. **(A)** Residue R75 is in a rigid linker and forms hydrogen bonds with residues from the neighboring linker. Its mutation to proline would disrupt the salt bridges and the hydrogen bond, as well as cause severe clashes with neighboring residues. **(B)** Residue V37 is located in a small cleft. Its replacement by phenylalanine would clash with the surrounding residues. **(C)** Residue Y107 locates in a semi-open eggshell-like pocket and the hydroxyl group on its side-chain forms a hydrogen bond with a neighboring residue. Both Y107C and Y107H mutants would disrupt this hydrogen bond and generate a cavity or clash with other residues surrounding the pocket. **(D, E)** Residues C63 and V111 are both buried deeply in the hydrophobic core of the structure. Their replacements by residues with bulkier side-chains (i.e., tyrosine and methionine) would cause severe clashes with the neighboring residues. All images were generated using Pymol (<http://www.pymol.org>).

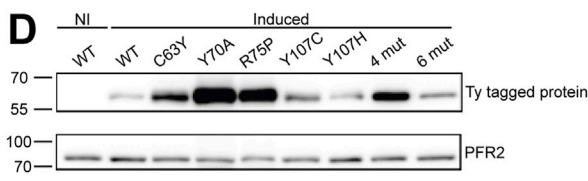
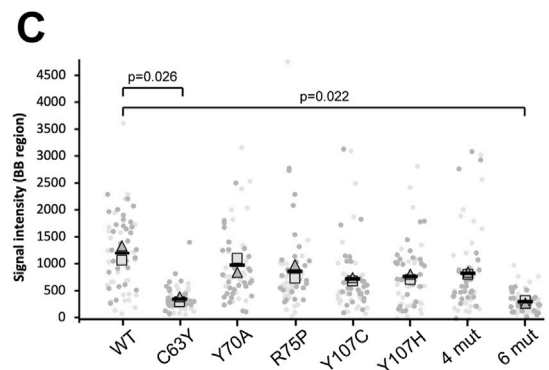
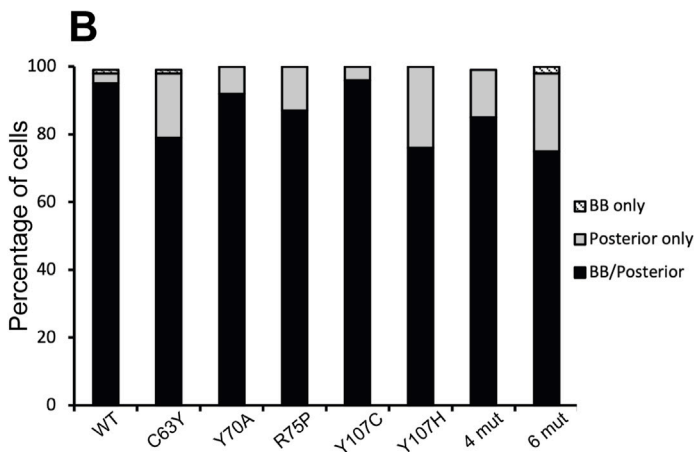
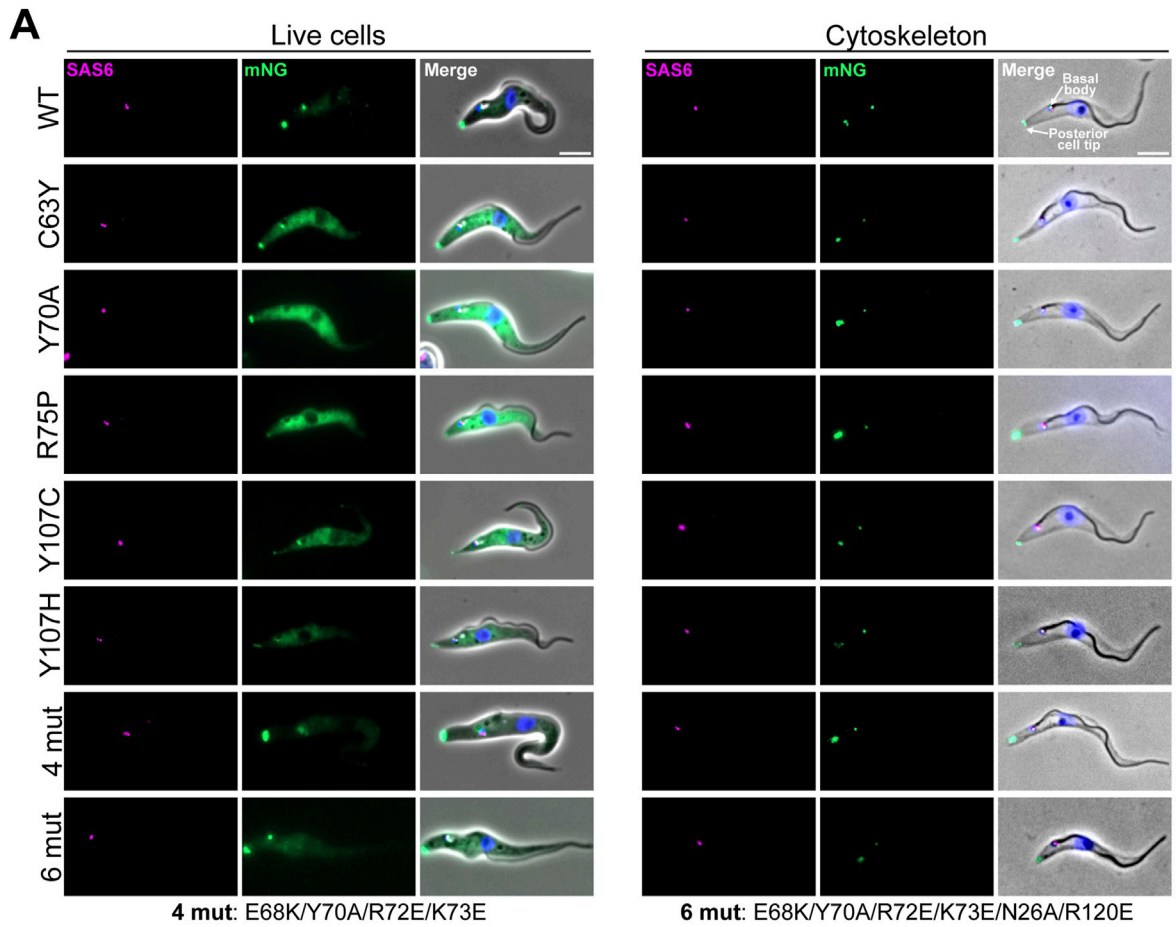


FIGURE 4 Fluorescence microscopy images from wild-type and mutants of TbCFAP410. **(A)** Images acquired from live cells or methanol fixed cytoskeletons from doxycycline induced cells expressing TbCFAP410 wild-type or TbCFAP410 mutants (green) with mScTbSAS6 - basal body marker (magenta). Scale bar: 5 μ m. **(B)** Graph of TbCFAP410 signal counts. The percentage of cytoskeletons with a basal body and posterior signal, a basal body only signal and posterior signal only was calculated for each cell line. Induction was repeated twice, with a similar result both times and data from one is shown. BB: basal body. N = 100 cells. **(C)** Signal intensity measurements from the basal body region for each cell line. Light and dark gray represents different replicates. The square is the mean from replicate one, the triangle the mean from replicate two, and the black bars the mean of both replicates. *(Continued)*

FIGURE 4 (Continued)

Measurements were compared between mutants and the wild type, and the p value of significant changes are indicated. N = 30 cells per replicate.

(D) Western blot showing the induction of expression of wild-type and mutant TbCFAP410 in *Trypanosoma brucei*. BB2 (anti-Ty, top blot) antibody was used for detecting Ty-tagged proteins, and L8C4 antibody (anti-PFR2, bottom blot) was used as a loading control. 1×10^6 parasites per lane were loaded. NI: non-induced RNAi.

in HsCFAP410) would cause multiple clashes between the bulky aromatic side-chain and the neighboring residues around the cleft (Figure 3B). Regarding the two mutations Y107C and Y107H, residue Y107 is located in a semi-open eggshell-like pocket. The hydroxyl group of Y107 forms a hydrogen bond with the side-chain of a neighboring residue. Mutation Y107C would apparently disrupt this hydrogen bond as well as generate a void cavity (Figure 3C). It is similar for mutation Y107H, but this mutation would additionally cause clashes between the side-chain of His and other hydrophobic residues around the pocket. For the other two mutants C63Y and V111M, both residues are fully buried inside the hydrophobic core of the structure (Figures 3D, E). Their replacements by residues with a bulkier side-chain would inevitably cause severe clashes and thus disrupt the folding of CFAP410-NTD. Our molecular dynamics simulation analyses indeed demonstrated significant changes in all these disease-causing mutants, particularly C63Y, in comparison with the crystal structure of wild-type TbCFAP410-NTD (Supplementary Figures S4A, B). Structural modeling by AlphaFold2 suggests that mutant C63Y would cause the flip of an α helix in the LRR motif to accommodate the large aromatic sidechain of tyrosine (Jumper et al., 2021). Such flip would subsequently destabilize the interaction between the helix and the hairpin loop to cause the exposure of the hydrophobic residues covered by the loop (Supplementary Figure S4C).

3.3 *In vivo* localization of the disease-causing mutants of TbCFAP410

To examine the impact of these disease-causing mutations on the function of TbCFAP410 *in vivo*, we generated individual mutant constructs with an N-terminal mNeonGreen fluorescent protein tag and expressed them in *T. brucei* using a tetracycline inducible expression system. All expressions were induced for 24 h. The cells, which also expressed SAS6 endogenously tagged with mScarlet as a marker of the basal body and pro-basal body, were then examined by wide-field fluorescence microscopy (Atkins et al., 2021). The wild-type protein localized to the basal and pro-basal bodies as well as the posterior cell tip, together with a weak cytoplasmic signal (Figure 4A, left), matching the localization observed in the TrypTag project (Billington et al., 2023). In trypanosomes the posterior cell tip is a site of microtubule dynamics with many microtubule-associated proteins located there, including CFAP410 (Billington et al., 2023). We found that the localization of the mutants in the NTD was on the whole similar to that of the wild-type CFAP410, with a posterior cell tip signal, a basal body signal and a cytoplasmic signal; however, there was variability in the signal strength for the mutants at these different cellular positions. Additionally, we examined a single central amino acid mutation (Y70A) and two clustered mutations within the surface patch that is highly conserved across all phyla, either mutating only the central ridge of the patch (4 mut: E68K/Y70A/R72E/K73E) or a more extensive

one with two extra flanking residues mutated as well (6 mut: E68K/Y70A/R72E/K73E/N26A/R120E). The localization patterns in the cells expressing these mutants were also like that of the wild-type protein, though again with variability in signal strength from the cytoplasm, basal body and posterior cell tip (Figure 4A, left).

To further investigate how stably the mutated proteins were associated with the cytoskeleton and to examine the basal body and posterior cell tip localizations without the cytoplasmic signal, we next generated detergent extracted cytoskeletons of the induced cells and quantified the localization of these mutants. Similar to what was seen in live cells, the wild-type and mutant proteins generally remained stably associated with both the basal body and the posterior cell tip, with a variability in signal strength seen, with for example, a weak posterior cell tip signal seen for the cells expressing the mutant proteins C63Y, Y107C, Y107H and 6mut (Figure 4A, right). The distribution of the localization of wild-type and all mutant TbCFAP410 was quantified (Figure 4B). For all mutants except Y107C, there was a slight reduction in the localization of the protein to the basal body. We also noted there was a level of variability in the basal body signal strength, with a weaker signal seen in the mutants C63Y and 6mut. We examined the basal body localization in more detail by measuring the signal strength for each mutant in 1K1N cytoskeletons (Figure 4C). This confirmed that there was a reduction in basal body signal strength in cells expressing the mutant proteins C63Y and 6mut. Expression of the wild-type and all mutant proteins after induction was confirmed by Western blot, which showed they all had the expected molecular weight (Figure 4D). Overall, these data suggest that the NTD of TbCFAP410 may contribute to the stabilization of the protein at the basal body.

3.4 Depletion of CFAP410 causes no significant changes in ciliogenesis in *Chlamydomonas reinhardtii*

To check if CFAP410 plays an essential role in ciliary assembly in another organism, we carried out depletion tests in *C. reinhardtii*. We successfully generated two targeted knockout mutants of CrCFAP410 using CRISPR/Cas9 (Figures 5A–C). However, although there was notable reduction in both flagellum length and percentage of flagellated cells in the two knockout cell lines in comparison to the wild-type cells, the differences were not as dramatic as what would be expected for an essential protein (Figures 5D, E).

4 Discussion

CFAP410 is a protein present in all sequenced ciliates and plays an essential role in ciliogenesis (Wheway et al., 2015). Located mainly at the

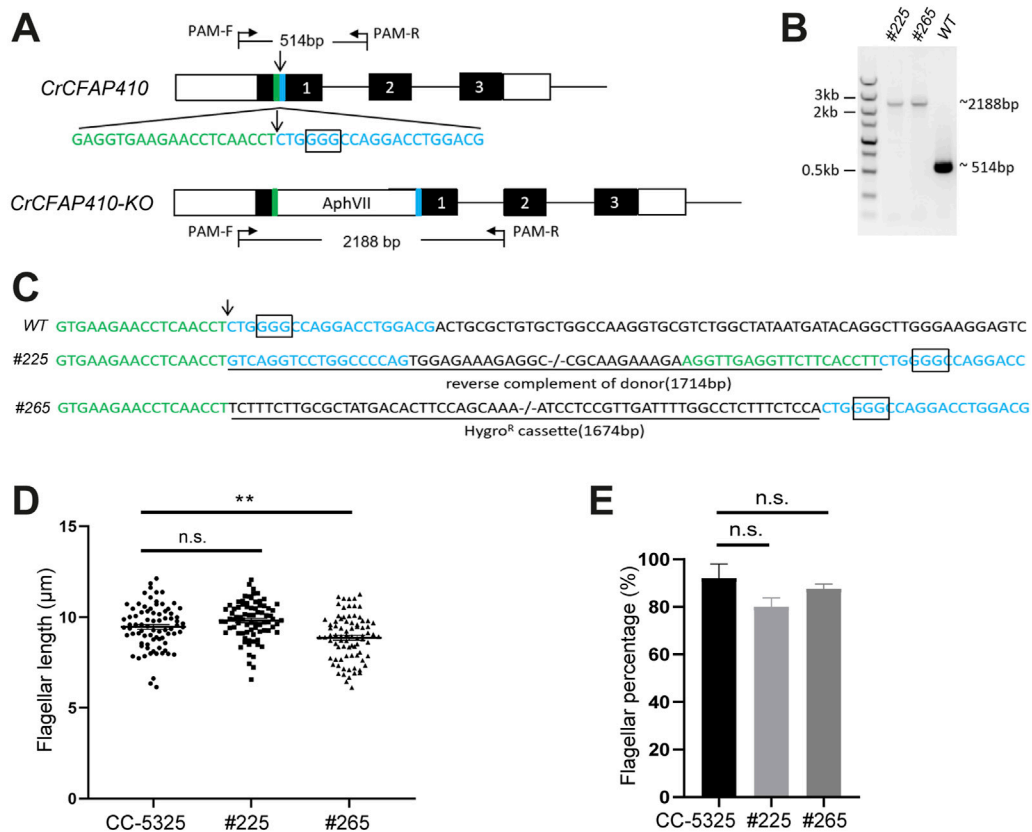


FIGURE 5
 Targeted knockout of CrCFAP410 by co-delivery of RNP complex and microhomologous dsDNA donor. **(A)** Schematic of the CrCFAP410 loci with dsDNA donor carrying a hygromycin resistance cassette inserted into the Cas9 cleavage site. Exons are indicated as black boxes, together with 5' and 3'UTR (white boxes), the cleavage site (arrow), the PAM site (the rectangular box) and 20 bp homologous arms (green and blue). The predicted sizes of the PCR fragments were 514 bp and 2,188 bp using PCR primers PAM-F and PAM-R in wild-type CC-5325 and in the knock-out mutant, respectively. **(B, C)** PCR results **(B)** and DNA sequences of PCR products **(C)** for wild-type and two mutants (#225 and #265). The 1,714 bp reverse complement of donor containing the two homologous arms was imprecisely integrated into the Cas9 cleavage site in mutant #225. The 1,674 bp hygromycin resistance cassette was precisely inserted in mutant #265. The insertion at the cleavage site was underlined. **(D, E)** The percentage of flagellated cells **(D)** and length **(E)** in CC-5325, CrCFAP410 knock-out mutants #225 and #265. Data are represented as mean ± standard error (S.E.). Statistical analysis was performed using a two-tailed unpaired Student's *t*-test. ***p* < 0.005, n.s., not significant.

basal body, the protein is highly conserved across phyla including algae, protists and animals. Our analysis shows that CFAP410 is a bimodular protein comprising two folded domains, i.e., NTD and CTD, which are connected via a variable unstructured loop. All single-residue mutations causing ciliopathies are located in the NTD or CTD. We have previously reported crystal structures of the CTD of three CFAP410 homologs from *T. brucei*, *C. reinhardtii*, and *H. sapiens*, which all adopt a similar tetrameric conformation (Stadler et al., 2024). Multiple single-residue mutations in HsCFAP410 have been identified in patients with spondylometaphyseal dysplasia, axial (SMDAX) or retinal dystrophy with or without macular staphyloma (RDMS) diseases (Khan et al., 2015; McInerney-Leo et al., 2017; Suga et al., 2016). Except for L224P, all other six ciliopathy mutations locate in the NTD of HsCFAP410. Here we report a high-resolution crystal structure of TbCFAP410-NTD, which shares 41% identities with the HsCFAP410. Comparison of our crystal structure with the predicted models of HsCFAP410 demonstrates a similar conformation. Therefore, we could extrapolate from the structure of TbCFAP410-NTD how the disease-causing mutations might affect the structure and function of HsCFAP410.

To check how the disease-causing mutations affect the function of CFAP410 in the native environment, we further carried out *in vivo* assays on TbCFAP410. Our results showed that unlike the mutant L272P that completely lost its localization at the basal body and the posterior cell tip, none of the mutants in the NTD showed such a dramatic change in their localization pattern. Our solubility tests for recombinant protein in *E. coli* demonstrate that all the mutants in the NTD of TbCFAP410 cause structural instability as mutant proteins were completely insoluble upon overexpression in *E. coli*. However, the defect in the folding of the NTD did not have a strong effect on the specific localization of TbCFAP410 in the cell. Moreover, we did not observe stable interaction between the individually purified NTD and CTD of TbCFAP410 (data not shown). All these suggest that the two modules in CFAP410 likely function independently, with the CTD controlling its localization and the NTD regulating its other functions.

In the crystal structure of TbCFAP410-NTD there is an extensive surface patch across all five strands of the β sheet. This patch consists of eighteen highly conserved residues. We

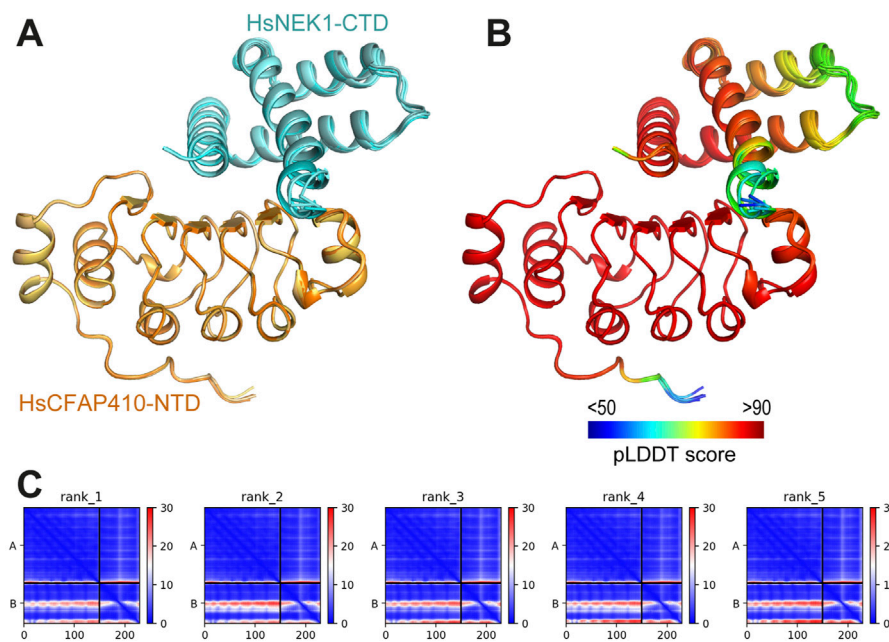


FIGURE 6

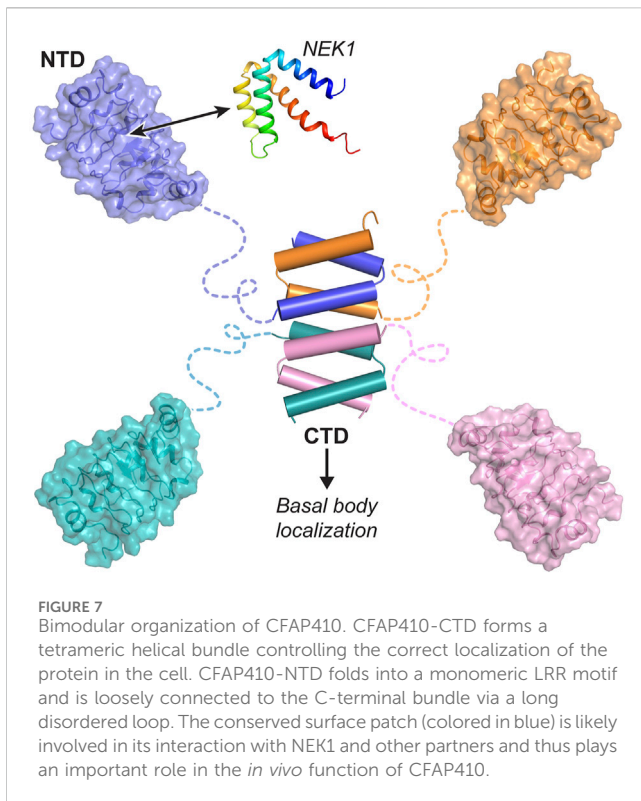
Predicted interaction between HsCFAP410-NTD and HsNEK1-CTD. **(A)** Superimposed cartoons of the top-ranked five models of the HsCFAP410-NTD (aa1-150) and HsNEK1-CTD (aa1208-1286) complex generated by AlphaFold2-multimer. **(B)** The same view as in **(A)** but colored by the predicted local distance difference test (pLDDT) score at each position. **(C)** Prediction aligned error (PAE) score for the five models, which indicates the calculated error of the predicted distance for each pair of residues. The uncertainty in the predicted distance of two amino acids is color coded from blue (0 Å) to red (30 Å), as shown in the right bar. The upper left quadrant corresponds to errors in the distances of residues within HsCFAP410-NTD, the lower right quadrant to errors within HsNEK1-CTD, and both the upper right and lower left quadrants to errors between CFAP410 and NEK1. The differences among the five models are rather marginal.

speculate that this surface patch might play an important role for the function of CFAP410. None of the mutations within the patch we tested caused strong defects in cellular localization of the protein in *T. brucei*; therefore, we think that the function of the conserved patch in CFAP410 is for its other functions. Previous studies have shown that CFAP410 is a component of a retinal ciliopathy-associated protein complex containing both NEK1 and SPATA7 (Eblimit et al., 2015). Direct interaction between CFAP410 and the C-terminal end of NEK1 has been observed in human cells (Gregorczyk et al., 2023). We therefore speculate that the conserved surface patch on the NTD of CFAP410 is possibly involved in its interaction with NEK1. We carried out structural modelling using AlphaFold-multimer for the complex of HsCFAP410-NTD (aa1-150) and HsNEK1-CTD (aa1208-1286). In the prediction, the multiple sequence alignment (MSA) identified approximately 3,000 and 1,000 sequences similar to HsCFAP410-NTD and HsNEK1-CTD, respectively. The five generated models of the complex could be perfectly superimposed (Figure 6A). The predicted local distance difference test (pLDDT) score was above 90 for most parts of both molecules, indicating very high accuracy that is equivalent to structures determined by experiments (Figure 6B). The prediction aligned error (PAE) score, which displays the calculated errors for the predicted distance of each pair of residues, was also substantially low for the residues on the interface in all the five models (Figure 6C). All these suggest that the conserved surface on the NTD of CFAP410 is likely the binding site of NEK1. Consistently, recent studies demonstrate

that mutation of multiple residues in the conserved surface patch on human CFAP410-NTD significantly reduced its interaction with NEK1-CTD (Gregorczyk et al., 2023).

Although the helical side of the LRR motif lacks conservation across phyla, a cross-species conservation pattern exists in all kinetoplastids. The hairpin-like loop in the crystal structure of TbCFAP410-NTD spans over 20 residues (aa139-160) and makes extensive contacts with the LRR motif via multiple hydrogen bonds as well as numerous hydrophobic interactions. Those residues in the loop involved in its interaction with the LRR motif are highly conserved in all kinetoplastid homologs, so as the corresponding interaction residues on the LRR motif. All these suggest that this extensive loop is a unique feature present only in kinetoplastid species. As mentioned above, all TbCFAP410 truncations with the intact loop were soluble, whereas those lacking parts or the whole loop, i.e., constructs covering residues 1–149 and 1–140, were all insoluble upon expression in *E. coli*. This further suggests that the loop is essential to facilitate the folding and/or to stabilize the structure of the LRR motif in TbCFAP410. The stabilizing loop might be an ancient element retained during evolution in the early branched flagellated Kinetoplastida.

CFAP410 is a ciliary protein localized to the base of cilia in human cells and involved in ciliopathies characterized with retinal and skeletal impairment (Khan et al., 2015; Wheway et al., 2015). Previous comparative proteomics studies revealed a 1.59-fold increase in CFAP410 levels during flagellar disassembly in *C. reinhardtii* (Wang et al., 2017a), suggesting a potential role in ciliogenesis. To investigate this, we generated two targeted knockout mutants of CFAP410 in



C. reinhardtii. Surprisingly, no significant changes in flagellar length or the percentage of flagellated cells were observed compared to wild-type cells. This may reflect a cell cycle-dependent role for CFAP410, as its effects on flagellar length could become apparent only at specific stages of the cell cycle. Alternatively, CFAP410 may regulate flagellar movement by influencing flagellar beating or participate in signaling pathways, as many flagella also function as sensory antennae via receptor localization in the flagellar membrane (Mitchell, 2007). Notably, CFAP410 interacts with NEK1 and SPATA7 (Eblimit et al., 2015; Fang et al., 2015; Wang et al., 2017b), further supporting its potential roles in ciliary signaling and motility.

Similar to *C. reinhardtii*, there is also a paralog of CFAP410 in *T. brucei* (accession code: XP_011777799/Tb927.10.5130). This paralog also localizes to the basal body, pro-basal body and posterior cell tip in same way as TbCFAP410 (Dean et al., 2017). Sequence alignment shows that the paralog has a similar size to TbCFAP410 and their N-terminal regions adopt a similar conformation (Supplementary Figures S5A, B). Furthermore, we found that most residues within the surface patch on TbCFAP410 are also conserved in the paralog (Supplementary Figures S5C, D). It is possible that the two proteins might function redundantly in *T. brucei*.

While *T. brucei* contains only a single paralog of CFAP410, human cells express eleven isoforms of the protein. Unlike the *T. brucei* paralog, which features a distinct CTD, all HsCFAP410 isoforms share an identical CTD (Supplementary Figure S6A). However, these isoforms exhibit diversification in their N-terminal regions or the linker connecting the two folded domains. A tetrameric model of full-length HsCFAP410, generated using AlphaFold-multimer, reveals that the CTD forms the same tightly packed helical bundle observed in the crystal structures. The N-terminal LRR motifs of the four subunits are loosely associated with the bundle, connected via a long, unstructured loop (Supplementary

Figure S6B). The large number of CFAP410 isoforms in humans might be formed from alternative splicings, variable promoter usage, or other post-transcriptional modifications of the same gene, with each unique mRNA sequence subsequently produces a specific form of the protein (Brett et al., 2002; Kpebe and Rabinow, 2008; Uhlen et al., 2015). Whether these isoforms exhibit tissue-specific expression, as observed with many other human proteins, remains an open question for future investigation. The fact that they all contain the same CTD that forms an oligomeric structure suggests that this domain likely plays an important role for the function of HsCFAP410 in humans. Despite variability in the connecting loops, most of these isoforms also contain an intact NTD with nearly identical sequences in their structural cores. This conservation further underscores the likely importance of the NTD in the functional roles of HsCFAP410 isoforms.

In summary, our extensive structural, biochemical, and *in vivo* studies reveal that CFAP410 forms a tetramer via its CTD, which assembles into a tightly packed eight-helix bundle. This structure clusters four globular NTDs into a compact protein complex (Figure 7). The NTD of CFAP410 adopts a canonical LRR motif with an extended surface patch that is highly conserved across diverse phyla, including algae, protist parasites, and mammals. This conservation suggests a critical functional role, such as mediating interactions with NEK1 or other binding partners. Our findings further elucidate the molecular mechanisms underlying disease-associated mutations in human CFAP410. Mutations like L224P disrupt the oligomeric assembly of the CTD, leading to mislocalization within the cell (Stadler et al., 2024). Meanwhile, mutations in the NTD (e.g., I35F, C61Y, R73P, Y107C, Y107H, and V111M) destabilize its structure, impairing interactions with essential partners. These disruptions likely drive the pathogenesis of CFAP410-related ciliopathies.

Data availability statement

The datasets presented in this study can be found in online repositories (PDB code for TbCFAP410-NTD: 8AXJ). The names of the repository/repositories and accession number(s) can be found in the article/Supplementary Material.

Ethics statement

The manuscript presents research on animals that do not require ethical approval for their study.

Author contributions

AS: Investigation, Methodology, Writing—original draft. HG: Investigation, Methodology, Writing—original draft. LD: Investigation, Methodology, Writing—original draft. SA-G: Investigation, Methodology, Writing—original draft, Writing—review and editing. XD: Investigation, Methodology, Writing—original draft. RC: Investigation, Writing—original draft. KK: Investigation, Methodology, Writing—original draft. BM: Investigation, Writing—original draft. SV: Writing—original draft, Conceptualization. KH: Formal Analysis, Supervision, Writing—original draft, Project administration. BŽ:

Formal Analysis, Supervision, Writing—original draft, Writing—review and editing. JS: Conceptualization, Formal Analysis, Project administration, Resources, Supervision, Writing—original draft, Writing—review and editing. GD: Conceptualization, Data curation, Formal Analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing—original draft, Writing—review and editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This work was supported by the grant I5960-B2 from the Austrian Science Fund (FWF) to GD and a Newton International Fellowship (NIF\R1\191618) to HG. SA acknowledges the VIP2 fellowship 2020 funded by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 847548. JDS acknowledges the Wellcome Trust and Oxford Brookes University for their support. RC was supported by a Nigel Groome PhD studentship.

Acknowledgments

The authors are grateful to the staff at the beamlines of ID23-1 at the European Synchrotron Radiation Facility (ESRF) for their help with X-ray diffraction data collection.

References

- Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., et al. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D. Biol. Crystallogr.* 66, 213–221. doi:10.1107/S0907444909052925
- Atkins, M., Tyc, J., Shafiq, S., Ahmed, M., Bertiaux, E., De Castro Neto, A. L., et al. (2021). CEP164C regulates flagellum length in stable flagella. *J. Cell Biol.* 220, e202001160. doi:10.1083/jcb.202001160
- Billington, K., Halliday, C., Madden, R., Dyer, P., Barker, A. R., Moreira-Leite, F. F., et al. (2023). Genome-wide subcellular protein map for the flagellate parasite *Trypanosoma brucei*. *Nat. Microbiol.* 8, 533–547. doi:10.1038/s41564-022-01295-6
- Brett, D., Pospisil, H., Valcarcel, J., Reich, J., and Bork, P. (2002). Alternative splicing and genome complexity. *Nat. Genet.* 30, 29–30. doi:10.1038/ng803
- Brun, R., and Schonberger, (1979). Cultivation and *in vitro* cloning or procyclic culture forms of *Trypanosoma brucei* in a semi-defined medium. Short communication. *Acta Trop.* 36, 289–292. doi:10.5169/seals-312533
- Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., et al. (2010). MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D. Biol. Crystallogr.* 66, 12–21. doi:10.1107/S0907444909042073
- Cheng, X., Liu, G., Ke, W., Zhao, L., Lv, B., Ma, X., et al. (2017). Building a multipurpose insertional mutant library for forward and reverse genetics in *Chlamydomonas*. *Plant Methods* 13, 36. doi:10.1186/s13007-017-0183-5
- Dean, S., Sunter, J. D., and Wheeler, R. J. (2017). TrypTag.org: a trypanosome genome-wide protein localisation resource. *Trends Parasitol.* 33, 80–82. doi:10.1016/j.pt.2016.10.009
- Dean, S., Sunter, J., Wheeler, R. J., Hodkinson, I., Gluenz, E., and Gull, K. (2015). A toolkit enabling efficient, scalable and reproducible gene tagging in trypanosomatids. *Open Biol.* 5, 140197. doi:10.1098/rsob.140197
- Eblimit, A., Nguyen, T. M., Chen, Y., Esteve-Rudd, J., Zhong, H., Letteboer, S., et al. (2015). Spata7 is a retinal ciliopathy gene critical for correct RPGRIPI localization and protein trafficking in the retina. *Hum. Mol. Genet.* 24, 1584–1601. doi:10.1093/hmg/ddu573
- Ehara, S., Kim, O. H., Maisawa, S., Takasago, Y., and Nishimura, G. (1997). Axial spondylometaphyseal dysplasia. *Eur. J. Pediatr.* 156, 627–630. doi:10.1007/s004310050679
- Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr. D. Biol. Crystallogr.* 60, 2126–2132. doi:10.1107/S0907444904019158
- Fang, X., Lin, H., Wang, X., Zuo, Q., Qin, J., and Zhang, P. (2015). The NEK1 interactor, C21ORF2, is required for efficient DNA damage repair. *Acta Biochim. Biophys. Sin. (Shanghai)* 47, 834–841. doi:10.1093/abbs/gmv076
- Gregorczyk, M., Pastore, G., Muñoz, I., Carroll, T., Streubel, J., Munro, M., et al. (2023). Functional characterization of C21ORF2 association with the NEK1 kinase mutated in human in diseases. *Life Sci. Alliance* 6, e202201740. doi:10.26508/lsa.202201740
- Hoover, W. G. (1985). Canonical dynamics - equilibrium phase-space distributions. *Phys. Rev. A* 31, 1695–1697. doi:10.1103/physreva.31.1695
- Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., et al. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature* 596, 583–589. doi:10.1038/s41586-021-03819-2
- Kabsch, W. (2010). Xds. *Acta Crystallogr. D. Biol. Crystallogr.* 66, 125–132. doi:10.1107/S0907444909047337
- Khan, A. O., Eisenberger, T., Nagel-Wolfrum, K., Wolfrum, U., and Bolz, H. J. (2015). C21orf2 is mutated in recessive early-onset retinal dystrophy with macular staphyloma and encodes a protein that localises to the photoreceptor primary cilium. *Br. J. Ophthalmol.* 99, 1725–1731. doi:10.1136/bjophthalmol-2015-307277
- Kobe, B., and Deisenhofer, J. (1995). Proteins with leucine-rich repeats. *Curr. Opin. Struct. Biol.* 5, 409–416. doi:10.1016/0959-440x(95)80105-7
- Kpebe, A., and Rabinow, L. (2008). Alternative promoter usage generates multiple evolutionarily conserved isoforms of *Drosophila* DOA kinase. *Genesis* 46, 132–143. doi:10.1002/dvg.20374
- Krohn, K., Ovod, V., Vilja, P., Heino, M., Scott, H., Kyriakou, D. S., et al. (1997). Immunochemical characterization of a novel mitochondrially located protein encoded by a nuclear gene within the DFNB8/10 critical region on 21q22.3. *Biochem. Biophys. Res. Commun.* 238, 806–810. doi:10.1006/bbrc.1997.7352
- Lindorff-Larsen, K., Piana, S., Palmo, K., Maragakis, P., Klepeis, J. L., Dror, R. O., et al. (2010). Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins* 78, 1950–1958. doi:10.1002/prot.22711
- McInerney-Leo, A. M., Wheeler, L., Marshall, M. S., Anderson, L. K., Zankl, A., Brown, M. A., et al. (2017). Homozygous variant in C21orf2 in a case of Jeune syndrome

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fcell.2025.1507470/full#supplementary-material>

- with severe thoracic involvement: extending the phenotypic spectrum. *Am. J. Med. Genet. A* 173, 1698–1704. doi:10.1002/ajmg.a.38215
- Mitchell, D. R. (2007). The evolution of eukaryotic cilia and flagella as motile and sensory organelles. *Adv. Exp. Med. Biol.* 607, 130–140. doi:10.1007/978-0-387-74021-8_11
- Nose, S. (1984). A unified formulation of the constant temperature molecular-dynamics methods. *J. Chem. Phys.* 81, 511–519. doi:10.1063/1.447334
- Poon, S. K., Peacock, L., Gibson, W., Gull, K., and Kelly, S. (2012). A modular and optimized single marker system for generating *Trypanosoma brucei* cell lines expressing T7 RNA polymerase and the tetracycline repressor. *Open Biol.* 2, 110037. doi:10.1098/rsob.110037
- Shin, S. E., Lim, J. M., Koh, H. G., Kim, E. K., Kang, N. K., Jeon, S., et al. (2016). CRISPR/Cas9-induced knockout and knock-in mutations in *Chlamydomonas reinhardtii*. *Sci. Rep.* 6, 27810. doi:10.1038/srep27810
- Stadler, A., De Liz, L. V., Gabriel, H. B., Alonso-Gil, S., Crickley, R., Korbula, K., et al. (2024). The C-terminus of CFAP410 forms a tetrameric helical bundle that is essential for its localization to the basal body. *Open Biol.* 14, 240128. doi:10.1098/rsob.240128
- Suga, A., Mizota, A., Kato, M., Kuniyoshi, K., Yoshitake, K., Sultan, W., et al. (2016). Identification of novel mutations in the LRR-cap domain of C21orf2 in Japanese patients with retinitis pigmentosa and cone-rod dystrophy. *Invest Ophthalmol. Vis. Sci.* 57, 4255–4263. doi:10.1167/iovs.16-19450
- Toda, A., Nishikawa, Y., Tanaka, H., Yagi, T., and Kurisu, G. (2020). The complex of outer-arm dynein light chain-1 and the microtubule-binding domain of the gamma heavy chain shows how axonemal dynein tunes ciliary beating. *J. Biol. Chem.* 295, 3982–3989. doi:10.1074/jbc.RA119.011541
- Tyagi, R., Lai, R., and Duggleby, R. G. (2004). A new approach to 'megaprimer' polymerase chain reaction mutagenesis without an intermediate gel purification step. *BMC Biotechnol.* 4, 2. doi:10.1186/1472-6750-4-2
- Uhlen, M., Fagerberg, L., Hallstrom, B. M., Lindskog, C., Oksvold, P., Mardinoglu, A., et al. (2015). Proteomics. Tissue-based map of the human proteome. *Science* 347, 1260419. doi:10.1126/science.1260419
- Van Der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A. E., and Berendsen, H. J. (2005). GROMACS: fast, flexible, and free. *J. Comput. Chem.* 26, 1701–1718. doi:10.1002/jcc.20291
- Wang, L., Gu, L., Meng, D., Wu, Q., Deng, H., and Pan, J. (2017a). Comparative proteomics reveals timely transport into cilia of regulators or effectors as a mechanism underlying ciliary disassembly. *J. Proteome Res.* 16, 2410–2418. doi:10.1021/acs.jproteome.6b01048
- Wang, Q., Peng, Z., Long, H., Deng, X., and Huang, K. (2019). Polyubiquitylation of α -tubulin at K304 is required for flagellar disassembly in *Chlamydomonas*. *J. Cell Sci.* 132, jcs229047. doi:10.1242/jcs.229047
- Wang, Z., Horemuzova, E., Iida, A., Guo, L., Liu, Y., Matsumoto, N., et al. (2017b). Axial spondylometaphyseal dysplasia is also caused by NEK1 mutations. *J. Hum. Genet.* 62, 503–506. doi:10.1038/jhg.2016.157
- Wang, Z., Iida, A., Miyake, N., Nishiguchi, K. M., Fujita, K., Nakazawa, T., et al. (2016). Axial spondylometaphyseal dysplasia is caused by C21orf2 mutations. *PLoS One* 11, e0150555. doi:10.1371/journal.pone.0150555
- Warman, M. L., Cormier-Daire, V., Hall, C., Krakow, D., Lachman, R., Lemerrer, M., et al. (2011). Nosology and classification of genetic skeletal disorders: 2010 revision. *Am. J. Med. Genet. A* 155A, 943–968. doi:10.1002/ajmg.a.33909
- Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., et al. (2018). SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res.* 46, W296–W303. doi:10.1093/nar/gky427
- Wheway, G., Schmidts, M., Mans, D. A., Szymanska, K., Nguyen, T. T., Racher, H., et al. (2015). An siRNA-based functional genomics screen for the identification of regulators of ciliogenesis and ciliopathy genes. *Nat. Cell Biol.* 17, 1074–1087. doi:10.1038/ncb3201
- Xue, J. H., Chen, G. D., Hao, F., Chen, H., Fang, Z., Chen, F. F., et al. (2019). A vitamin-C-derived DNA modification catalysed by an algal TET homologue. *Nature* 569, 581–585. doi:10.1038/s41586-019-1160-0