Domain-level interaction of FAP174 (MYCBP-1) and FAP147 (MYCBPAP) proteins of the C2a projection of *Chlamydomonas* cilia

[1]Sneha Desai, [1]Abhay Pal[$], [2]Hridhya Nair[$], [3,4]Sarath Chandra Dantu and [1]Jacinta S. D'Souza*

[$]Equal contribution

[1]School of Biological Sciences, University of Mumbai-Department of Atomic Energy Centre for Excellence in Basic Sciences, Kalina campus, Santacruz (E), Mumbai 400098, India

[2]School of Chemical Sciences, University of Mumbai-Department of Atomic Energy Centre for Excellence in Basic Sciences, Kalina campus, Santacruz (E), Mumbai 400098, India

[3]Department of Computer Science, Brunel University London, Uxbridge UB8 3PH, United Kingdom

[4]The Thomas Young Centre for Theory and Simulation of Materials, London SW7 2AZ, United Kingdom

*Corresponding author: jacinta@cbs.ac.in

Phone (M): +91-9820770314

Running title: MycBP-1 and MycBPAP interactions from *Chlamydomonas* flagella

Abstract

The C2a projection of the central pair of flagella in *Chlamydomonas reinhardtii* harbours the A-Kinase anchoring protein FAP65, FAP174, FAP147, and FAP70. FAP174, an RII-like protein with its N-terminal dimerization and docking domain, binds to the amphipathic helices of FAP65. Cryo-EM data do not reveal the entire sequences for FAP174 and FAP147. Hence, the interacting domains within this scaffold remain elusive. This study has identified the interacting domains of FAP174 with FAP147. The FAP147 protein and its MYCBPAP domain (129-639 a.a.) bind to the C-terminus of FAP174 (47-92 a.a.). *In silico* docking analyses using CABS-Dock to delineate the interaction identified several MYCBPAP-derived peptides, such as p3 (310–339), p4 (319–348), p9 (547–576), p13 (528–557), and p15 (350–379), to form stable interacting complexes with RMSD <3 Å 2-3 times, and are potentially amphipathic. To gain atomistic details of the interaction, molecular dynamics (MD) simulations of the FAP147 MYCBPAP domain in complex with the FAP174 C-terminus were performed. It revealed stable interfacial contacts, a subset of which overlap with residues within the p15 peptide region of the MYCBPAP domain, while identifying G48, S49, P52, Y55, L79, Q80, and V83 as key interacting residues within the C-terminus of FAP174. Individual alanine substitution of the FAP174 residues, followed by overlay assay with FAP147, retained the interaction, indicating that the interaction does not depend solely on discrete amino acids but on broader interface interactions.

Keywords: *Chlamydomonas*, FAP147, FAP174, Flagella, MYCBP-1, MYCBPAP

Introduction

Proteins, primarily consisting of amino acid (a.a.) residues, are the macromolecules that play a crucial role in most biological activities within cells and rarely perform their function as a single entity. Instead, they interact with a range of small molecules and biomolecules, such as ligands, other proteins, peptides, nucleic acids, nucleotides, hemes, and ions, to form stable complexes that enable their roles in cellular processes (Alberts, 1998). Biological processes widely recognized as outcomes of the continuous interaction between functional nucleic acid stretches and specific proteins involve DNA replication, transcription, gene regulation, chromosome packaging, repair, and rearrangement, along with RNA activities like splicing, transport, translation, and silencing (Marko, 2015). Such crosstalk between interaction networks ensures the stability and functionality of organelle development and function. For example, in organelles such as centrosomes, the pericentriolar material constitutes a protein complex that facilitates microtubule nucleation through protein-protein interactions (Lee and Young, 2013; Nigg and Stearns, 2011). The current study focuses on protein-protein interaction in the 9+2 motile cilium.

Cilia/flagella are small, hair-like, microtubule-based, evolutionarily conserved structures found in diverse cell types, ranging from single-celled protozoans to multicellular tissues in mammals (Satir and Christensen, 2007). The axoneme, originating from the basal body, constitutes the basic structure of cilia. Based on the microtubule arrangement in the axoneme, cilia can either be immotile (primary cilia) or motile. Non-motile cilia are composed of 9+0 axonemes that are devoid of the central pair of microtubules, radial spokes (RS), and dynein arms and are present as a solitary structure on most epithelial cells, fibroblasts, and neurons, where they primarily have a sensory function (Mitchell, 2007). The 9+2 motile cilia are composed of nine outer doublet microtubules (DMT), inner and outer dynein arms (IDA and ODA), the nexin-dynein regulatory

complex (N-DRC), radial spokes (RS), and the two central microtubules. These are present in cells of the respiratory tract, fallopian tube, sperm, and brain ventricles, where they function to transport fluid across cellular surfaces and motility (Satir and Christensen, 2007). Together, the substructures that make up the 9+2 motile cilia generate and regulate motility. The precise mechanism of this regulation remains uncertain. Dysfunction of both non-motile and motile cilia leads to phenotypically and genotypically heterogeneous disorders collectively known as ciliopathies. Primary ciliary dyskinesia (PCD), hydrocephalus, and various forms of male infertility (asthenozoospermia, teratozoospermia) are ciliopathies that are caused by the dysfunction of motile cilia (Lechtreck and Witman, 2007; Lee *et al*., 2008; Finn *et al*., 2014; Sha *et al*., 2020). PCD is a rare heterogeneous multisystem disorder with a clinical phenotype such as middle ear disease, rhinosinusitis, bronchiectasis, and subfertility (Bhatt and Hogg, 2020).

It is hypothesized that the central pair apparatus (CPA) acts as a mechanotransducer that interacts with the RS to transmit mechanochemical signals to the N-DRC and IDA, regulating ODA activity (Smith, 2002; Smith, 2002; Smith and Yang, 2004). The central microtubules consist of 11 architecturally distinct projections (C1a-f, C2a-e) and large protein complexes with 16- or 32-nm longitudinal periodicity. This structure is collectively known as the CPA. Mutations in CPA genes cause PCD and affect organs such as the ear, eyes, brain, reproductive, and respiratory systems (Bhatt and Hogg, 2020). Research on *Chlamydomonas reinhardtii* shows that cilia comprise 450 proteins (Pazour *et al*., 2005). CPA includes over 66 proteins, with 48 mapped by cryogenic electron microscopy (cryo-EM), each contributing uniquely to signal transduction and motility (Zhao *et al*., 2019; Han *et al*., 2022; Gui *et al*., 2022). Studies on CPA gene mutations have increased; for example, Hydin, an interactor of KLP1, localizes to the C2b projection and spans the bridge linking C1 and C2 microtubules, serving as a structural tether. In *C. reinhardtii*

hydin mutants, the C2b projection is impaired, and the C2c projection is compromised, indicating Hydin's role in stabilizing both projections or as part of C2c (Lechtreck and Witman, 2007). Hydin mutation causes hydrocephalus in mice and chronic respiratory infections in humans (Lechtreck *et al*., 2008; Shapiro *et al*., 2023). PF6, another CPA protein, exists in C1a and C1e projections as part of a complex with FAP101, FAP119, FAP114, FAP227, and calmodulin. Its absence in *C. reinhardtii* reduces motility and loss of the C1a projection. The C-terminal domain of PF6 is crucial for motility and assembly of the C1a projection, while the N-terminal half contributes to the stability of the PF6 sub-complex and beat frequency (Goduti and Smith, 2012). SPAG17, the mammalian ortholog of PF6, is associated with PCD and asthenozoospermia in mice (Kazarian *et al*., 2018; Abdelhamed *et al*., 2020). FAP74 forms a part of the multi-protein complex at the base of the C1d projection that consists of calmodulin-binding proteins such as FAP54, FAP46, and FAP221. Its mutant lacks the C1d projection, the C1f sheath (C1d to C1b), with a reduction in the swimming velocity and an asymmetrical beat frequency (DiPetrillo and Smith 2010; Gui *et al*., 2022). In humans, CFAP74 results in multiple morphological abnormalities of the sperm flagella ((MMAF); Sha *et al*., 2020; Biebach *et al*., 2022). Mutations in FAP81 cause structural and phenotypic defects, including the absence of C1c and C1e projections and flagella with reduced beat frequency, altered waveforms, and asynchrony (Fu *et al*., 2019). DLEC1/CFAP81 is linked to cancer (Daigo *et al*., 1999; Kwong *et al*., 2006; Zhang *et al*., 2010; Ying *et al*., 2009; Zhang *et al*., 2015). FAP70, a C2a protein, forms a complex with FAP65 and FAP147, with mutants lacking the C2a projection and displaying impaired motility (Hou *et al*., 2021). The *cpc1* mutant demonstrates reduced beat frequency and swim velocity and lacks the C1b projection (Zhang and Mitchell, 2004). FAP65, also known as AKAP240, localizes to the C2a projection of CPA, where its amphipathic helices (AH) interact with the RII-domain of FAP174 (Rao *et al*., 2024). FAP174

is localized to three different positions within the CPA (Han *et al*., 2022; Gui *et al*., 2022). The *fap174* mutant exhibits poor or no motility and shorter flagella, with no structural changes seen under transmission electron microscopy (TEM), suggesting involvement in flagellar assembly (Rao *et al*., 2024). The *fap42* mutant exhibits reduced swimming speed and impaired ciliary beating, accompanied by structural defects in the CPA, such as decreased densities in the C1b arm, as observed using cryo-electron tomography (cryo-ET; Cai *et al*., 2021). In addition to FAP42, mutations in several other proteins located outside the CPA can also cause CPA defects. These include the *pf15*, *pf18*, *pf19*, and *pf20* mutants in *C. reinhardtii*, all of which display flagellar paralysis resulting from the absence of a functional CPA. *PF15* encodes the p80 subunit of katanin, a microtubule-severing protein (Dymek *et al*., 2004), while *PF19* encodes the p60 catalytic subunit of katanin. Both proteins localize to the flagella, where they play essential roles in CPA assembly and stability by promoting microtubule remodelling and the proper incorporation of CPA microtubules into the axoneme. In contrast, the gene corresponding to *PF18* remains unidentified (Dymek and Smith, 2012). PF20 is located along the length of the C2 microtubule, at inter-microtubule bridges that connect the core microtubules (Smith and Lefebvre, 1997).

The CPA projections containing protein complexes, therefore, contribute to the structural integrity and motility of the cilia. Loss of any protein and/or their interacting domains from such complexes may affect its function and, therefore, motility. Such domains within the proteins in the complex also serve to stabilize the projection (Goduti and Smith, 2012). Domain-domain mapping will aid in elucidating the protein's function in the complex, specifically, structural stability or motility. To completely understand the mechano-transduction role of the CPA, mapping the proteins in its projections is therefore essential. While most proteins have been mapped, some gaps do exist (Han *et al*., 2022; Gui *et al*., 2022). For example, in the C2a projection, certain a.a.

stretches in FAP65 and FAP147 remain 'unmodelled'. The a.a. stretches of FAP65 that are 'unmodelled' include 1-1107, 1426-1442, 1438-1560, 1778-2257 and are therefore not evident in the cryo-EM images (Han *et al*., 2022; Gui *et al*., 2022). Also, the a.a. residues 48-89 of the sequence in the cryo-EM study that the authors consider as FAP174 do not match exactly with the sequence of FAP174 annotated in NCBI, Phytozome and our earlier study (Rao *et al*., 2024; Rao *et al*., 2016). Additionally, the stretch corresponding to 1-47 a.a. is not revealed in the cryo-EM (Han *et al*., 2022). The C2a projection consists of FAP65 (A-kinase anchoring protein), FAP147 (ASH domain-containing and mammalian MYCBPAP homologue), FAP174 [mammalian MYCBP-1 homologue and RII-like, (Yogesha *et al*., 2017; Rao *et al*., 2024), FAP70 (tetratricopeptide-containing), PF20 (WD repeats-containing), FAP239 and FAP20 (CFAP20 domain). Of these, FAP174, FAP147, FAP70 and FAP65 form a multiprotein complex (Hou *et al*., 2021; Lyu *et al*., 2026), suggesting their possible interaction. However, the exact domain interaction of FAP174 and FAP147 in *C. reinhardtii* and mammals remains elusive. Previous studies have shown that the N-terminus of FAP174 harbours the dimerization and docking domain and binds to amphipathic helices of FAP65 (Rao *et al*., 2024; Rao *et al*., 2016; Yogesha *et al*., 2017). Since the cryo-EM analysis of this projection has not been completely mapped, an approach toward domain-domain mapping is pivotal for proving the interactions between these proteins. The current study elucidates the interaction between two proteins in this scaffold, FAP147 and FAP174 and bridges the gap of domain mapping not revealed for some proteins in cryo-EM.

Results

In-silico analysis of FAP147

The C2a projection of the central pair apparatus of *C. reinhardtii* flagella consists of a multiprotein complex consisting of FAP65, FAP70, FAP147, and FAP174 (Hou *et al*., 2021). FAP147 is a 936 a.a. protein encoded by a 2,811 bp ORF region of the *C. reinhardtii* genome, with a gene GC content of approximately 72%. FAP147 has an a.a. composition characterized by 49.39% hydrophobic, 13.7% basic, and 12.1% acidic residues (Table 1). On the other hand, domain analysis highlights the presence of the MYCBP-associated protein (MYCBPAP) domain, ASH (Aspm-SPD-2 Hydin) domain, Siah-interacting protein, N-terminal, and domain of unknown function on the FAP147 protein (Figure 1A), and further sequence alignment using MultAlin software displayed significant conservation of the MYCBPAP and the ASH domain (Figure 1B) across various taxa, including plants, animals, fungi, and protozoans. Comparative domain alignment revealed that the MYCBPAP domain of *C. reinhardtii* shares 60-93% sequence similarity, with the highest similarity with closely related Chlorophyta, followed by invertebrate deuterostomes and non-mammalian chordates within the phylum Chordata. A BLASTP search further indicated that FAP147 is homologous to MYCBPAP-like proteins from both ciliary and non-ciliary organisms, confirming its identity as a MYCBPAP homologue (Table 2). Phylogenetic analysis using the Maximum Likelihood method and Jones-Taylor-Thornton (JTT) matrix-based model (Jones *et al*., 1992) showed two distinct clusters, reflecting evolutionary divergence between higher eukaryotes and unicellular or primitive multicellular organisms. The first cluster comprises higher eukaryotes, including vertebrates and early-diverging metazoans, while the second cluster includes unicellular and primitive multicellular eukaryotes, such as green algae, ciliates, and other organisms that belong to Amoebozoa, Stramenopiles, and Apicomplexa. *C. reinhardtii* FAP147 falls within this protist cluster, specifically grouped in the algal lineage (Figure 1C).

Generating gene and protein resources, followed by quality checks

It is worth noting that when this study commenced, a single Fap147 sequence, annotated as the human homologue of MYCBPAP, was available on NCBI and published in 2020 by the Witman laboratory (Zhao *et al*., 2020). This full-length ORF was gene-synthesized for this study. For the overexpression and/or solubilization and purification of FAP147 and its variants, the following approach was undertaken. The full-length ORF was synthesized and cloned into pET-28a(+). After the plasmid underwent appropriate clinical restriction enzyme digestion, it was induced with IPTG in *E. coli* BL21(DE3) (Figure S1A). The overexpression posed a challenge; hence, the 6xHis-tagged FAP147 full-length protein was overexpressed in LB media supplemented with 3% ethanol (Figure S1B, C; Chhetri *et al*., 2015). Further, the nucleotide sequence corresponding to the conserved MYCPBAP domain (385-1917 bp) was successfully cloned into pGEX-4T-1 using cDNA as a template (Figure S2A), and the overexpressed recombinant protein showed a polypeptide at ~81.5 kDa on SDS-PAGE (Figure S2B). The GST-tagged recombinant polypeptide was evident upon western blot analysis, and some degradation was also observed (Figure S2B). The recombinant polypeptide was insoluble, but solubilization with 2% sarkosyl, followed by affinity purification, yielded the anticipated molecular weight, along with additional low-molecular-weight bands (Figures S2C, D). Similarly, the DNA fragment corresponding to the ASH domain (1582-1686 bp) was cloned into the pGEX-4T-1 vector after performing a PCR using cDNA as a template (Figure S3A). The resulting overexpressed recombinant polypeptide, anticipated to have a molecular weight of ~31.3 kDa, was soluble and successfully purified using affinity chromatography (Figure S3B-D). The DNA sequence (1-897 bp) corresponding to the N-terminus (1-299 a.a.) was independently cloned into pET-28a(+), yielding comparable solubilization results (Figures S4A-C). The FAP147 C-terminus, characterized by a larger

amplicon size (~1900 bp) relative to other FAP147 fragments, was successfully cloned into the pET-28a(+) vector (Figure S5A), and the recombinant polypeptide was overexpressed in M9 minimal media, which was not evident on the Coomassie-stained gel (Figure S5B). Western blot analysis using an anti-6xHis antibody revealed the recombinant polypeptide of the expected Mr (~66 kDa; Figure S5B). The protein was found to be soluble (Figure S5C).

FAP174 full-length and its corresponding fragments used in this study were already cloned into pET-28a(+), (Rao *et al*., 2024) and purified to homogeneity. The purity and molecular mass of the FAP174 full-length and its corresponding fragments used for overlay studies were confirmed using silver staining. The *Mr* for these fusion proteins was confirmed using silver staining of the gels onto which the purified proteins were electrophoresed (Figure S6). Similarly, the purity of the FAP147 fragments used for pull-down studies was ascertained by silver staining (Figure S6).

Binding of FAP174 and its variants with FAP147 and its variants

The purified recombinant full-length and variants (N-terminus, Middle portion, C-terminus, C46A) of FAP174 were used as individual baits in an overlay assay wherein the overexpressed FAP147 full-length and its variants (N-terminus, MYCBPAP domain, ASH domain, C-terminus) were presented on a nitrocellulose membrane and transferred after electrophoresis on an SDS-PAGE gel. FAP174, C46A and its C-terminus showed binding to FAP147, whereas the N-terminal and middle regions exhibited no binding to FAP147. This suggests that the primary interaction interface resides within the C-terminus of FAP174 (Figure 2). On the other hand, the N-terminus and C-terminus fragments of the FAP147 protein showed no interaction with FAP174 and its variants; however, the MYCBPAP domain of FAP147 exhibited specific binding to FAP174 and its C-terminus, with no notable interactions detected with other variants of FAP174 (Figure

3A&B). This interaction between the FAP147 MYCBPAP domain and FAP174 and its C-terminus was further ascertained using a pull-down assay.

The pull-down assay using GST beads with MYCBPAP as a bait and FAP174 and its C-terminus demonstrated that both FAP174 and its C-terminus exhibited specific binding to the MYCBPAP domain of FAP147, corroborating the findings from the overlay assay (Figure 4A&B). These findings show that the C-terminal region of FAP174 mediates its interaction with the MYCBPAP domain of FAP147.

Binding of the specific amino acid stretch of the MYCBPAP domain with the C-terminus of FAP174, an *in silico* insight

CABS-Dock web server generated 10,000 model structures of the query protein sequence. The top-ranked models of FAP174 and FAP147 peptides generated by the AlphaFold 3 structures were used in CABS-Dock studies. In this study, the cluster with the highest density was selected, and the average RMSD was used as the measure of the binding quality. Structures with RMSD <3 Å were considered high-quality predictions, representing the most probable models for the protein-peptide complexes (Kurcinski *et al*., 2015). Docking of the conserved regions of the MYCBPAP domain with the FAP174 monomer identified peptides p4, p3, p7, p13, and p9 with increasing RMSD values, all <3 Å (Figure S7A; Table 3). The docking interface primarily involved residues 1 and 32 of the FAP174 chain, exhibiting a contact frequency > 0.5 (Figure S7B). Among the predicted models, the p4 peptide formed a helical structure and showed the lowest RMSD (1.43 Å), suggesting the strongest binding with the N-terminus of the FAP174 monomer. In contrast, when the conserved regions of the MYCBPAP domain were docked against the C-terminal region of the FAP174 monomer (FAP174 CT), none of the peptides exhibited RMSD < 3 Å (Table 3), indicating no significant binding occurs at this interface.

On the other hand, when the FAP174 dimer was used as a receptor and the conserved regions of the MYCBPAP domain were used as peptides, the peptides p9, p4, p14, p18, p5, p1, p13 p3, and p15 showed RMSD values <3 Å (Figure S7C; Table 3), with p9 showing the lowest and p15 showing the highest RMSD value. The docking site is between residues 1 and 32 of the FAP174 chains with a contact frequency >0.5 (Figure S7D). From the predicted structures, a model with p9 as the peptide showing the lowest RMSD value suggests that the conserved p9 binds at the N-terminus of FAP174. Earlier studies have shown that the N-terminus of FAP174 dimer interacts with the amphipathic helix (AH) of FAP65 (Rao *et al*., 2024) and probably occupies it *in vivo*. Hence, it was decided to use the FAP174 dimer bound to AH1 used as the receptor, with the conserved regions of the MYCBPAP domain as peptides; the peptides p9, p6, p12, p13 and p15 show RMSD <3Å (Figure S7E; Table 3). The docking occurs between the residues 38 and 92 of the FAP174 chains with a contact frequency > 0.5 (Figure S7F). From the predicted structures, a model with p9 as the peptide exhibits a helical structure and the lowest RMSD value, suggesting that the conserved peptide p9 spans the N-terminus and the C-terminus of FAP174. Together, five peptides (p9, p4, p13, p3, p15) demonstrated significant interaction with the receptor protein in CABS docking studies. Subsequently, all five peptides were explored for common features. Since AH1 was known to bind to FAP174, the peptides were assessed for their potential to form amphipathic helices using HeliQuest software (Figure S7G-K). The peptides exhibited a strong to moderate propensity for amphipathic helix (AH) formation, with p13 demonstrating the least propensity due to a less distinct segregation of hydrophobic and hydrophilic a.a.(Figure S7I). Peptides p9, p4, p3 and p15 exhibited appropriate segregation of hydrophobic residues, resulting in the formation of a hydrophobic face (Figure S7G, H, J, K), which supports their binding to the

N-terminus of the FAP174 monomer and dimer, especially in the absence of AH1 complexed to FAP174.

Molecular docking simulation studies to predict amino acid level interactions of MYCBPAP domain and the C-terminus of FAP174

To further study the contact residues, MD simulations were conducted on the experimentally mapped interface, utilizing the MYCBPAP domain of FAP147 and the C-terminus of FAP174 whose structural models were generated by Protenix software. The complex exhibited a binding free energy, i.e. ΔG of ~ -14.5 kcal/mol (Figure 5A) with MD simulations revealing reproducible interaction patterns across three independent trajectories. Backbone RMSD analysis showed stabilization within ~0.4-1.0 A^{o} after equilibration, indicating overall structural integrity of the truncated complex while allowing for conformational flexibility over the 100-nanosecond time scale (Figure 5B). Contact frequency analysis using get_contacts identified recurring residue-level interactions between the MYCBPAP domain and the FAP174 C-terminus. Quantification of hydrogen bonds, hydrophobic interactions, salt bridges, and vdw interactions across three distinct MD simulations revealed that hydrophobic, salt bridges, and vdw interactions are mainly involved in the complex. Several residue pairs G(48), S(49), P(50), P(52), Y(55), L(79), Q(80), and V(83) engaged with the MYCBPAP domain across trajectories and form highly stable interactions with significant occupancy with Y(55) and V(83) showing high frequency of vdw contacts (Figure 5C, S8A). The following FAP174 C-terminus – MYCBPAP domain complex residues formed contacts with high frequency vdw interactions: G(48)-D(213) and A(214), L(79)-V(165), P(52)-I(362), Y(55)-A(358) and V(83)-A(663) (Figure S8A). These FAP174 C-terminus residue contacts were subsequently used to experimentally assess their contribution to the interaction.

Validation of the interacting residues of the FAP174 protein by site-directed mutagenesis

Site-directed mutagenesis was performed on residues in the C-terminal region of FAP174 that were anticipated by molecular dynamics simulations to interact with FAP147's MYCBPAP domain. All variant constructs were sequence validated and successfully overexpressed in *E. coli* BL21 (DE3) (Figure S9A-E). The mutant FAP174 proteins were soluble, comparable to the wild-type recombinant FAP174 polypeptide, and were purified to homogeneity using Ni–NTA affinity chromatography. SDS-PAGE gel revealed that the FAP174 variants exhibited both monomeric and dimeric forms like those observed for wild-type FAP174, indicating that the mutations did not affect the dimeric state (Figure S9F-J). Overlay assays using FAP147 as prey and the FAP174 variant polypeptides as bait showed interactions with all variants (Figure 3C).

Discussion

Ciliary motility in *C. reinhardtii* is orchestrated by the precise management of signalling pathways through protein–protein interactions within the CPA projections and radial spokes. Among these, A-kinase anchoring proteins (AKAPs) act as critical scaffolds, organizing signalling complexes essential for motility regulation. MYCBPAP, a protein implicated in mammalian spermatogenesis, is part of a multiprotein complex including KIF9, SPAG16, CFAP65, CFAP70, and MYCBP-1. Disruption of *Mycbpap* in mice leads to male infertility, reduced sperm motility, and aberrant sperm morphology, revealing its functional significance. Homologues of MYCBPAP and MYCBP-1, namely FAP147 and FAP174, from the *C. reinhardtii* flagellar proteome (Pazour *et al*., 2005) have been localized specifically to the C2a projection (Han *et al*., 2022; Gui *et al*., 2022). In human HeLa cells, MYCBPAP and MYCBP-1 co-localize within the cytoplasm. In mammals, both MYCBPAP and MYCBP-1 exhibit testis-specific expression, predominantly during the post-meiotic stages of spermatogenesis (Yukitake *et al*., 2002). MYCBP-1, which contains the

conserved dimerization and docking domain, forms protein complexes with various AKAPs, indicating its organelle-specific signalling roles. Notably, MYCBP-1 interacts with the transcription factor c-MYC, modulating E-box–mediated transcriptional activation (Taira *et al*., 1998; Furusawa *et al*., 2001). In the nucleus, MYCBP-1, together with AKAP95 and PKA, forms ternary complexes that regulate PKA kinase activity (Furusawa *et al*., 2002). Beyond the nucleus, MYCBP-1 is found in sperm mitochondria—interacting with AKAP149 and its splice variant S-AKAP84—and in the trans-Golgi network, where it associates with BIG2 to potentially coordinate ARF-mediated membrane trafficking (Furusawa *et al*., 2001; Ishizaki *et al*., 2006).

FAP174, an RII-like protein with an N-terminal dimerization and docking domain, interacts with the amphipathic helices of AKAPs such as FAP65 in C2a, FAP297 in C1d/f and CPC1 in C1b (Rao *et al*., 2024). Cryo-EM analyses have indicated that FAP147 and FAP174 are proximal, suggesting a potential interaction; however, the precise domains mediating this interaction remain elusive. Domain mapping of these scaffolding proteins is crucial for understanding the structural basis of CPA complexes, as these projections function as multiprotein assemblies that regulate signalling and flagellar motility.

In *C. reinhardtii*, FAP147 functions within a FAP65 scaffold complex together with FAP70 and FAP174 in the C2a projection of the central pair apparatus. The present study provides the first biochemical evidence defining the interaction domains of FAP147 and FAP174. Sequence and domain analyses identified a conserved MYCBPAP domain in FAP147, consistent with orthologues across higher eukaryotes (Figure 1). Phylogenetic analysis further supports this conservation, revealing two major clusters: one comprising higher eukaryotes, and another containing unicellular and primitive multicellular organisms such as green algae and ciliates. The high GC content (~72%) and hydrophobic nature (~58.8% hydrophobic a.a.; Table 1) of *C.*

reinhardtii FAP147 posed challenges for recombinant expression in *E. coli*, resulting predominantly in inclusion body formation (Figure S1). Employing denaturing solubilization (2 M urea) and surfactant-based extraction (2% Sarkosyl) enabled solubilization of the full-length protein (Figure S1C) and partial recovery and purification of its ASH and MYCBPAP domains (Figure S2C-D; S3C-D; 6B). Domain-specific constructs revealed that the ASH domain expressed in soluble form under native conditions, whereas the MYCBPAP domain required harsher solubilization to yield usable quantities. These results emphasize the need for tailored expression strategies for high-GC, hydrophobic proteins such as FAP147. For comparison, human MYCBPAP exhibits a lower GC content (54.09%) and reduced hydrophobicity, yet exhibits *in vitro* instability (instability index 51.93; data not shown). The optimized expression and purification framework described here establishes a foundation for *in vitro* interaction studies of MYCBPAP homologues that were previously impractical. The FAP174, its variants, and FAP147, its variants, thus purified, were ~90% pure (Figure S6).

Cryo-EM reconstructions revealed that FAP147 and FAP174 are in close spatial arrangement within the C2a projection. This study sheds light on the unexplored interactions of the MYCBPAP domain of FAP147 of FAP174, as revealed by overlay (Figures 2-3) and pull-down assays (Figure 4). This biochemical evidence explains the evolutionary conservation of the MYCBPAP domain across species and reinforces its functional relevance. The mapped amino acid residues of FAP147 known to interact with FAP65 (416–420, 449–454, 481–486, 488, 489, 491, 493, 494, 517–519, and 521) and FAP70 (434, 437, 536, 538, 540–543) (Gui *et al*., 2022) frame its broader role within the CPA. The residues interacting with FAP174 (21, 22, 27, 29–40, 43, 45, 47, 52, and 161) clarify its position within the complex. Sequence-level discrepancies within the FAP174 region (1647–1660) require further structural resolution to reconcile these details. Thus,

based on the findings of the present study and the cryo-EM studies, it is evident that the conserved MYCBPAP domain of FAP147 protein harbours residues critical for interaction with other C2a projection proteins, such as FAP65, FAP70, and FAP174.

As identified through sequence analysis, the MYCBPAP domain spans from 129-639 a.a. within the FAP147 primary sequence. To further predict the precise interacting regions of the MYCBPAP domain and FAP174 C-terminus, *in silico* analyses were performed. Docking simulations using CABS-dock, which models flexible peptide-protein interactions without prior knowledge of the binding site (Blaszczyk *et al*., 2016), revealed a preference for binding of MYCBPAP-derived peptides to the N-terminal region of FAP174, especially for shorter peptides (≤30 a.a.). This apparent difference can be attributed to several factors, including the limitation of peptide sequences (30 a.a.) in docking with CABS-dock which commonly adopt α-helical conformations that is enriched in residues that form amphipathic helices *in silico*, thus mimicking amphipathic helices that bind to the R2 D/D fold at the N-terminus of FAP174. Additionally, the overlay test involves the entire domain contacts. In contrast, CABS-dock assesses local peptide flexibility and may overestimate interactions with accessible or highly charged surfaces, such as the N-terminus. Notably, the C-terminal region of FAP174 was identified as the predominant contact interface when the FAP174 dimer and MYCBPAP peptides were docked with a modelled amphipathic helix (representing a more native environment), which was consistent with the overlay results (Figure S7E, F). The N-terminal interaction observed in docking (Figure S7A-D) may therefore reflect a methodological bias of peptide-based docking, which favours amphipathic helical configurations (Figure S7G-K) that may not be prevalent in the full-length protein environment. Taken together, these findings indicate that FAP174 can interact with various amino acid stretches; however, under experimental conditions, the arrangement of the domains favours

binding *via* the C-terminus (Figures 3-4). CABS-dock simulations with MYCBPAP-derived peptides (p9, p4, p13, p3 and p15) revealed stable interaction modes, characterized by low RMSD values across various FAP174 conformations (Table 3). The peptides exhibit amphipathic and electrostatically favourable motifs, aligning with flexible yet structured binding interfaces (Figure S7 A-K).

Given that CABS-dock does not account for the structural constraints in the MYCBPAP domain, an MD simulation of MYCBPAP and the C-terminus of FAP174 was conducted to further define the experimentally relevant interface (Karaca *et al*., 2022). MD simulations of FAP147 MYCBPAP domain and FAP174 C-terminus showed certain residues making contact and forming a stable protein-protein complex (Figure 5). The following FAP174 C-terminus - MYCBPAP domain complex residues formed contacts with high frequency vdw interactions: G(48)-D(213) and A(214), L(79)-V(165), P(52)-I(362), Y(55)-A(358) and V(83)-A(663) (Figure S8A). These residues overlap with conserved peptide segments highlighted by CABS-dock as high-confidence interaction motifs (RMSD < 3 Å). Mutations in FAP174 C-terminus a.a. residues, individually or in combination using the full-length protein to assess the contribution of these residues to binding under native folding conditions, followed by overexpression, solubilization and purification showed that these FAP174 variants exhibited monomeric and dimeric forms, characteristics like those of wild-type FAP174, indicating that the mutations did not significantly perturb dimer formation (Figure S9). An overlay assay using these variants as individual baits revealed that these maintained contact with FAP147, suggesting that no individual residue is exclusively responsible for binding (Figure 3C). This is consistent with a binding interface distributed across multiple residues in the native proteins. Transient interactions are essential for signalling, regulation, and dynamic assembly of protein complexes. These interactions are stabilized by interfaces distributed

across the proteins, so mutation of a single amino acid on an interface does not necessarily abolish the interaction (Ozbabacan *et al*., 2011). Such interactions are essential when a protein interacts with multiple proteins in the multiprotein complex that is involved in dynamic processes such as signalling. In this scenario, FAP147 and FAP174 are CPA-harboured proteins and thus hypothesized to be involved in signalling and regulation of motility. Cryo-EM data reveal that FAP147 is an elongated protein that serves as a central structural component in the C2a projection, thereby acting as a framework for organizing C2a arm proteins, such as FAP65 and FAP70. It also interacts with FAP213, a microtubule inner protein (MIP), by penetrating through the cleft between C2 microtubule protofilaments (Han *et al*., 2022). FAP147 thus acts as a structural scaffold, possessing the ability to span various structural layers of the C2 microtubule and thereby forming the central axis of the C2a projection. It thus contributes to the structural framework of the C2a projection, which in turn helps regulate ciliary motility, involving ciliary waveform, beat frequency, and velocity. Our data support the structural role of the FAP147 protein. The *in vitro* assays suggest that the MYCBPAP domain of FAP147 binds the C-terminus of FAP174, which may be part of the C2a arm complex. The rachis-like structure of FAP147 likely positions MYCBPAP to interact with FAP174's C-terminus, possibly stabilizing or orienting the binding site. According to cryo-EM, FAP147 has charged patches that facilitate interactions with other proteins (Han *et al*., 2022). These domain mapping findings support the concept of FAP147 as a scaffold that facilitates regulated binding events, aligning with its suggested rachis-like function within the central pair.

In conclusion, this study closes a critical gap left by incomplete cryo-EM models by demonstrating that the conserved MYCBPAP domain of FAP147 interacts specifically with the C-terminal region of FAP174 (Figure 6). Together, these findings advance our understanding of how

the FAP65–FAP147–FAP174 complex assembles and functions within the central pair apparatus of *C. reinhardtii*. This work lays a solid foundation for future structural and functional analyses, extending insights to their mammalian counterparts and enhancing our comprehension of the conserved signalling mechanisms that regulate ciliary motility, particularly for rachis-like proteins such as FAP147.

Materials and Methods

In-silico analyses of FAP147

National Centre for Biotechnology Information (NCBI) BLAST (https://blast.ncbi.nlm.nih.gov/Blast.cgi), ProtParam (https://web.expasy.org/protparam/), Clustal Omega (https://www.ebi.ac.uk/jdispatcher/msa/clustalo), and MultAlin (http://multalin.toulouse.inra.fr/multalin/) software were used to perform bioinformatic analysis. The NCBI database (https://www.ncbi.nlm.nih.gov/) and *C. reinhardtii* flagellar proteome were used to retrieve gene, ORF, and protein sequences. Gene and protein sequences of other organisms were obtained from the NCBI (https://www.ncbi.nlm.nih.gov/). For *in silico* analysis of FAP147, the following parameters were studied: (i) Sequence retrieval: the primary sequence of the *C. reinhardtii* FAP147 protein was retrieved in FASTA format from the NCBI database and Phytozome v13 database (https://phytozome-next.jgi.doe.gov/) and used for further *in silico* *a*nalysis. (ii) Analysis of physicochemical properties: Physical and chemical properties, including molecular weight, isoelectric point (pI), extinction coefficients, aliphatic index (AI), and GRAVY (grand average of hydropathy) of *C. reinhardtii* FAP147 protein were analyzed using the ProtParam tool of ExPASy. (iii) The family, domains, and motifs of FAP147 were determined using the Conserved Domain Database (CDD) and Pfam databases. (iv) Multiple sequence

alignment: A BLASTp search from NCBI (http://www.ncbi.nlm.nih.gov/) against the non-redundant database with default parameters was performed to find the protein's homologs. Multiple-sequence alignments and a phylogenetic tree were constructed using the MultAlin and MEGAX software, respectively.

AlphaFold 3 structure prediction, molecular docking, and identification of the amphipathic helix

AlphaFold3 webserver (https://alphafoldserver.com/; Abramson *et al*., 2024) was used to predict the protein/peptide structures. Each structure was generated by specifying the seed parameter as one of the controls for reproducibility. The protein and ligand structures submitted for docking were those generated by AlphaFold 3. The CABS-Dock webserver (https://biocomp.chem.uw.edu.pl/CABSdock/; Kurcinski *et al*., 2015; Blaszczyk *et al*., 2016) was used to conduct flexible protein-peptide docking. The input receptor was predicted using AlphaFold 3, submitted directly to the server without further modification, and the ligand peptide was input as a sequence of up to 30 amino acids. For each protein-peptide pair, CABS-Dock was run without prior knowledge of the binding site. Docked models were evaluated based on the CABS-Dock average root mean square deviation (RMSD) of the cluster. Default parameters were applied to all docking runs.

Molecular Docking Simulation of FAP147 MYCBPAP- FAP174 C-terminus complex

Protein-protein complex structures were generated using the Protenix (https://www.biorxiv.org/content/10.1101/2025.01.08.631967), an AlphaFold-based multimer modelling framework. The binding free energy (ΔG) of the predicted complexes was estimated using the PRODIGY webserver (Vangone and Bonvin, 2015; Xue *et al*., 2016), which predicts protein-protein binding affinity based on interfacial contacts. The selected protein-protein complex

was prepared for MD simulations using the CHARMM-GUI v3.7 solution builder (Park *et al*., 2023). The complex was solvated in an octahedral simulation box with a minimum edge distance of 10 Å using explicit TIP3P (Jorgensen *et al*., 1983; MacKerell *et al*., 1998) water molecules, and appropriate counter ions, NaCl, were added to neutralize the system. The CHARMM36 force field (Bussi *et al*., 2007) was employed for proteins, and MD simulations were performed using GROMACS 2022 (Lee *et al*., 2016). Energy minimization was carried out using the steepest descent algorithm, followed by equilibration under the NVT ensemble at 298.15 K (25 °C) using the V-rescale thermostat (Essmann *et al*., 1995), and subsequent pressure equilibration under periodic boundary conditions. Long-range electrostatic interactions were treated using the particle mesh Ewald (PME) (Hess *et al*., 1997) method with a real-space cutoff of 1.2 A^{o}, while vdw interactions were handled using a force-switch scheme with a cutoff of 1.2 A^{o}. All hydrogen bonds were constrained using the LINCS algorithm (Hess *et al*., 1997). Three independent production MD simulations were carried out for 100 nanoseconds each. The analysis of the MD simulation results was performed using GROMACS Tools, in-house Python scripts, and PyMol 3.1.0 for visualization (Schrödinger, LLC, Version 1.8, 2023). Atomistic interaction analysis across simulation trajectories was performed using the get_contacts framework, enabling time-resolved identification of inter-residue interactions at the binding interface.

Cloning the full-length Fap147 gene and its DNA fragments

The ORF for the full-length Fap147 was identified from NCBI (www.ncbi.nlm.nih.gov) as a homologue of the human MYCBPAP gene and was gene-synthesized. RNA extraction was performed from the *C. reinhardtii* wild-type strain CC4533 using the Qiagen RNeasy Plant mini-Kit (Hilden, Germany, catalogue no. 74904) according to the manufacturer's protocol. RNA was

further converted to cDNA using Thermo Scientific™ RevertAid first-strand cDNA synthesis kit (Vilnius, Lithuania, catalogue no. K1621) and used as a template for amplifying Fap147 DNA fragments (Table S1). For genes to be cloned into the pET-28a(+) vector, specific primers were designed for Fap147 DNA fragments flanking *Nco*I and *Xho*I restriction sites. Similarly, for the Fap147 gene fragments to be cloned into the pGEX-4T-1 vector, specific primers flanking *Bam*HI and *Xho*I restriction sites were designed. The cDNA of wild-type *C. reinhardtii* was used as a template to PCR amplify the Fap147 gene fragments with the NEB Q5® High-Fidelity DNA polymerase (Ipswich, Massachusetts, USA, catalogue no. M0491S). The PCR-amplified products were visualized using agarose gel electrophoresis, and the DNA band corresponding to the expected amplicon was gel-extracted using the Cytiva GFX PCR DNA and gel band purification kit (UK, product code 28903470), followed by quantification in the Tecan Infinite M Plex plate reader. Similarly, the vectors pET-28a(+) and pGEX-4T-1 were subjected to restriction digestion using *Nco*I, *Xho*I and *Bam*HI, *Xho*I enzymes, respectively, followed by treatment with Thermo Scientific™ FastAP Thermosensitive alkaline phosphatase enzyme (Lithuania, catalogue no. EF0651) and heat inactivation of the enzyme at 80°C for 10 minutes. The restriction enzyme (RE) digested vectors were checked for RE digestion on agarose gel electrophoresis, and the RE-digested vector was gel extracted and quantified using a protocol like the one used for PCR products. The fap147 gene fragments were then cloned into the respective RE-digested vectors using NEB Gibson Assembly® Master Mix (Ipswich, Massachusetts, USA, catalogue no. E2611L) using a 1:10 and 1:5 insert:vector ratio according to the length of the DNA fragment by incubation of a 1:1 ratio of insert, vector, and Gibson assembly master mix at 50°C for 1 hour. This reaction mixture was transformed into *E. coli* DH5α strain using the chemical transformation protocol, and the transformed cells were plated on Luria-Bertani (LB) agar supplemented with 30 mg/mL

kanamycin [for pET-28a(+)] and 100 mg/mL ampicillin (for pGEX-4T-1). The transformed clones were then checked for the presence of the gene by colony PCR, in which the colony was mixed with 90 µL of nuclease-free water (NFW), heated at 95 ℃ for 5 minutes to lyse the cells, and then used as a template for PCR. Positive clones, which showed the expected amplicon of corresponding DNA fragments, were screened for the presence of the gene in the plasmid by performing plasmid PCR and RE digestion of the plasmid. The clones that showed positive results in colony PCR, plasmid PCR, and RE digestion were sequenced and checked for the correct sequence. *E. coli* BL21(DE3) expression host was transformed with the plasmid.

Overexpression of the full-length recombinant FAP147 protein, and the fragments of Fap147

The recombinant FAP147 protein and its fragment were overexpressed using 0.5 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG) at 37℃ or 18℃ in Luria–Bertani (LB) or Minimal media for 6 hours or overnight. The induced pellet was resuspended in 1X SDS-PAGE sample loading buffer and checked for the induction of SDS-PAGE. Immunoblotting with anti-6xHis and anti-GST antibodies [for pET-28a(+) and pGEX-4T-1, respectively] was done to confirm the induction. The overexpressed FAP147 protein and its fragment with a 6xHis tag were solubilized using TNG (50 mM Tris, 300 mM NaCl, 10% Glycerol) buffer containing 20 mM imidazole, 10 mM β-mercaptoethanol, and 1 mM phenylmethylsulfonyl fluoride (PMSF). For 6xHis-tag proteins that were not soluble in this buffer, the overexpressed pellet was solubilized using 2 M Urea. The GST-tagged recombinant FAP147 fragments were solubilized using TEN (10 mM Tris, 5 mM EDTA, 250 mM NaCl) buffer with 10 mM β-mercaptoethanol and 1 mM PMSF. Similarly, for GST-tagged proteins that were not soluble in this buffer, the overexpressed pellet was solubilized using 2% Sarkosyl. Western blotting and probing with anti-6xHis (Cambridge, UK, catalogue no.

ab200537) and anti-GST antibody (Rockford, USA, catalogue no. MA4-004) were performed to check for solubilization.

Quality checks for all the recombinant proteins

Silver staining was performed to check the purity of FAP174 recombinant proteins and their fragments. For silver staining, the purified recombinant proteins were run on SDS-PAGE, following which the gel was washed with MQ and then fixed in a fixing solution (5:4:1 of millQ:methanol:glacial acetic acid) for 1 hour. The fixing solution was decanted, followed by a 30 sec MilliQ wash. To the staining solution (150 mg silver nitrate in 100 mL milliQ, i.e. 8.85 mM), 57 µL of freshly prepared formaldehyde was added to stain the gel for 30 min. at RT on the rocker. The staining solution was decanted, followed by a 30 sec milliQ wash. To pre-develop (285 mM sodium carbonate, 2.55 µM solution), 57 µL of formaldehyde was added and kept until colour developed. Once colour development began, this solution was decanted, and 57 µL of formaldehyde was added to the freshly prepared developing solution. Once the gel was stained, the solution was decanted, and the terminating solution (10% Glacial acetic acid) was added.

Overlay assays for protein interaction

An overlay assay was performed to check for protein-protein interaction. A Western blot was performed for solubilized FAP147 protein and its fragment (prey protein), followed by Ponceau S staining of the nitrocellulose membrane to check for the transfer. The membrane was rinsed thrice with 1X TBST for 10 minutes with agitation on the rocker. The membrane was further blocked with a blocking solution (3% BSA in 1X TBST) at room temperature for 1 hour. The prey proteins (FAP174 and its fragments) were applied to the membranes at a concentration of 10 µg/mL and incubated overnight at 4°C with gentle agitation. After incubation, membranes were subjected to three washes with 1X TBST for 10 minutes each, followed by incubation with an anti-FAP174

antibody for 1 hour at ambient temperature with moderate agitation. The surplus antibody was eliminated by washing the membranes thrice with 1X TBST for 10 minutes each time. The blot was then incubated with HRP-labelled secondary antibody for 1 h, following which washing was done to remove excess antibody. Protein bands were identified using the ClarityTM Western ECL substrate as per the manufacturer's instructions. Negative controls comprised membranes treated without bait protein or just with a secondary antibody.

Pull-down assay

A pull-down assay was conducted to investigate the protein-interacting characteristics of FAP174 and FAP147 fragments following affinity purification. To achieve this, 150µg of purified FAP147 fragment was incubated for two hours at 4 °C with approximately 50 µl of Glutathione-Sepharose beads. After incubation, the flow-through was allowed to drain, and subsequent washes (wash buffer: 10 mM Tris, 300 mM NaCl, and 1 mM PMSF) were conducted until protein was undetectable by the Bradford reagent. The beads were then incubated overnight at 4°C with solubilized recombinant FAP174 FL and fragment, after which the flow-through and washes were collected. GST served as a negative control. The beads were resuspended in a 2X concentration of SDS-loaded dye and visualized using SDS-PAGE electrophoresis.

Site-directed mutagenesis

Mutations in the Fap174 gene were performed using the protocol mentioned in the Q5 site-directed mutagenesis kit (NEB, Ipswich, Massachusetts, USA, catalogue no. E0554S). The primers for the mutations were generated using NEBaseChanger™. PCR amplification was performed with the PCR master mix from the kit using the plasmid as the template, followed by a KLD reaction and transformation of the mixture into chemically competent *E. coli* cells, as per the protocol provided in the kit. The cells were plated on selective media containing kanamycin (30 µg/mL) and

incubated at 37 °C for 16-18 hours. The plasmids were extracted from the clones, and the clones were sequenced. The plasmid with the mutation was then transformed into *E. coli* BL21(DE3) and checked for overexpression and solubility, followed by affinity purification using Ni-NTA beads.

Acknowledgement

The authors acknowledge the funding provided by the Department of Atomic Energy (DAE), Govt. of India, for the research resources and student stipends. Part of this work was funded by the Indo-Swiss grant (Ref. no. IC-12044(11)/8/2021-ICD-DBT). This project made use of time on ARCHER2 granted via the UK High-End Computing Consortium for Biomolecular Simulation, HECBioSim (http://hecbiosim.ac.uk), supported by EPSRC (grant no. EP/X035603/1).

References

1. Abdelhamed, Z., Lukacs, M., Cindric, S., Ali, S., Omran, H., & Stottmann, R. W. (2020). A novel hypomorphic allele of *Spag17* causes primary ciliary dyskinesia phenotypes in mice. *Disease models & mechanisms*, *13*(10), dmm045344. https://doi.org/10.1242/dmm.045344
2. Abramson, J., Adler, J., Dunger, J., Evans, R., Green, T., Pritzel, A., Ronneberger, O., Willmore, L., Ballard, A. J., Bambrick, J., Bodenstein, S. W., Evans, D. A., Hung, C. C., O'Neill, M., Reiman, D., Tunyasuvunakool, K., Wu, Z., Žemgulytė, A., Arvaniti, E., Beattie, C., … Jumper, J. M. (2024). Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature*, *630*(8016), 493–500. https://doi.org/10.1038/s41586-024-07487-w

3. Acuner Ozbabacan, S. E., Engin, H. B., Gursoy, A., & Keskin, O. (2011). Transient protein-protein interactions. *Protein engineering, design & selection : PEDS*, *24*(9), 635–648. https://doi.org/10.1093/protein/gzr025
4. Alberts B. (1998). The cell as a collection of protein machines: preparing the next generation of molecular biologists. *Cell*, *92*(3), 291–294. https://doi.org/10.1016/s0092-8674(00)80922-8
5. Bhatt, R., & Hogg, C. (2020). Primary ciliary dyskinesia: a major player in a bigger game. *Breathe (Sheffield, England)*, *16*(2), 200047. https://doi.org/10.1183/20734735.0047-202
6. Biebach, L., Cindrić, S., Koenig, J., Aprea, I., Dougherty, G. W., Raidt, J., Bracht, D., Ruppel, R., Schreiber, J., Hjeij, R., Olbrich, H., & Omran, H. (2022). Recessive Mutations in *CFAP74* Cause Primary Ciliary Dyskinesia with Normal Ciliary Ultrastructure. *American journal of respiratory cell and molecular biology*, *67*(3), 409–413. https://doi.org/10.1165/rcmb.2022-0032LE
7. Blaszczyk, M., Kurcinski, M., Kouza, M., Wieteska, L., Debinski, A., Kolinski, A., & Kmiecik, S. (2016). Modeling of protein-peptide interactions using the CABS-dock web server for binding site search and flexible docking. *Methods (San Diego, Calif.)*, *93*, 72–83. https://doi.org/10.1016/j.ymeth.2015.07.004
8. Bussi, G., Donadio, D., & Parrinello, M. (2007). Canonical sampling through velocity rescaling. *The Journal of chemical physics*, *126*(1), 014101. https://doi.org/10.1063/1.2408420

9. Cai, K., Zhao, Y., Zhao, L., Phan, N., Hou, Y., Cheng, X., Witman, G. B., & Nicastro, D. (2021). Structural organization of the C1b projection within the ciliary central apparatus. *Journal of cell science*, *134*(21), jcs254227. https://doi.org/10.1242/jcs.254227
10. Chhetri, G., Kalita, P., & Tripathi, T. (2015). An efficient protocol to enhance recombinant protein expression using ethanol in Escherichia coli. *MethodsX*, *2*, 385–391. https://doi.org/10.1016/j.mex.2015.09.005
11. Daigo, Y., Nishiwaki, T., Kawasoe, T., Tamari, M., Tsuchiya, E., & Nakamura, Y. (1999). Molecular cloning of a candidate tumor suppressor gene, DLC1, from chromosome 3p21.3. *Cancer research*, *59*(8), 1966–1972.
12. DiPetrillo, C. G., & Smith, E. F. (2010). Pcdp1 is a central apparatus protein that binds Ca(2+)-calmodulin and regulates ciliary motility. *The Journal of cell biology*, *189*(3), 601–612. https://doi.org/10.1083/jcb.200912009
13. Dymek, E. E., & Smith, E. F. (2012). PF19 encodes the p60 catalytic subunit of katanin and is required for assembly of the flagellar central apparatus in Chlamydomonas. *Journal of cell science*, *125*(Pt 14), 3357–3366. https://doi.org/10.1242/jcs.096941
14. Dymek, E. E., Lefebvre, P. A., & Smith, E. F. (2004). PF15p is the chlamydomonas homologue of the Katanin p80 subunit and is required for assembly of flagellar central microtubules. *Eukaryotic cell*, *3*(4), 870–879. https://doi.org/10.1128/EC.3.4.870-879.2004
15. Essmann, U., Perera, L.E., Berkowitz, M.L., Darden, T.A., Lee, H., & Pedersen, L.G. (1995). A smooth particle mesh Ewald method. *Journal of Chemical Physics, 103*, 8577-8593.

16. Finn, R., Evans, C. C., & Lee, L. (2014). Strain-dependent brain defects in mouse models of primary ciliary dyskinesia with mutations in Pcdp1 and Spef2. *Neuroscience*, *277*, 552–567. https://doi.org/10.1016/j.neuroscience.2014.07.029

17. Fu, G., Zhao, L., Dymek, E., Hou, Y., Song, K., Phan, N., Shang, Z., Smith, E. F., Witman, G. B., & Nicastro, D. (2019). Structural organization of the C1a-e-c supercomplex within the ciliary central apparatus. *The Journal of cell biology*, *218*(12), 4236–4251. https://doi.org/10.1083/jcb.201906006

18. Furusawa, M., Ohnishi, T., Taira, T., Iguchi-Ariga, S. M., & Ariga, H. (2001). AMY-1, a c-Myc-binding protein, is localized in the mitochondria of sperm by association with S-AKAP84, an anchor protein of cAMP-dependent protein kinase. *The Journal of biological chemistry*, *276*(39), 36647–36651. https://doi.org/10.1074/jbc.M103885200

19. Furusawa, M., Taira, T., Iguchi-Ariga, S. M., & Ariga, H. (2002). AMY-1 interacts with S-AKAP84 and AKAP95 in the cytoplasm and the nucleus, respectively, and inhibits cAMP-dependent protein kinase activity by preventing binding of its catalytic subunit to A-kinase-anchoring protein (AKAP) complex. *The Journal of biological chemistry*, *277*(52), 50885–50892. https://doi.org/10.1074/jbc.M206387200

20. Goduti, D. J., & Smith, E. F. (2012). Analyses of functional domains within the PF6 protein of the central apparatus reveal a role for PF6 sub-complex members in regulating flagellar beat frequency. *Cytoskeleton (Hoboken, N.J.)*, *69*(3), 179–194. https://doi.org/10.1002/cm.21010

21. Gui, M., Wang, X., Dutcher, S. K., Brown, A., & Zhang, R. (2022). Ciliary central apparatus structure reveals mechanisms of microtubule patterning. *Nature structural & molecular biology*, *29*(5), 483–492. https://doi.org/10.1038/s41594-022-00770-2

22. Han, L., Rao, Q., Yang, R., Wang, Y., Chai, P., Xiong, Y., & Zhang, K. (2022). Cryo-EM structure of an active central apparatus. *Nature structural & molecular biology*, *29*(5), 472–482. https://doi.org/10.1038/s41594-022-00769-9

23. Hess, B., Bekker, H., Berendsen, H.J., & Fraaije, J.G. (1997). LINCS: A linear constraint solver for molecular simulations. *Journal of Computational Chemistry, 18*.

24. Hou, Y., Zhao, L., Kubo, T., Cheng, X., McNeill, N., Oda, T., & Witman, G. B. (2021). Chlamydomonas FAP70 is a component of the previously uncharacterized ciliary central apparatus projection C2a. *Journal of cell science*, *134*(12), jcs258540. https://doi.org/10.1242/jcs.258540

25. Ishizaki, R., Shin, H. W., Iguchi-Ariga, S. M., Ariga, H., & Nakayama, K. (2006). AMY-1 (associate of Myc-1) localization to the trans-Golgi network through interacting with BIG2, a guanine-nucleotide exchange factor for ADP-ribosylation factors. *Genes to cells : devoted to molecular & cellular mechanisms*, *11*(8), 949–959. https://doi.org/10.1111/j.1365-2443.2006.00991.x

26. Jo, S., Kim, T., Iyer, V. G., & Im, W. (2008). CHARMM-GUI: a web-based graphical user interface for CHARMM. *Journal of computational chemistry*, *29*(11), 1859–1865. https://doi.org/10.1002/jcc.20945

27. Jones, D. T., Taylor, W. R., & Thornton, J. M. (1992). The rapid generation of mutation data matrices from protein sequences. *Computer applications in the biosciences : CABIOS*, *8*(3), 275–282. https://doi.org/10.1093/bioinformatics/8.3.275

28. Jorgensen, W.L., Chandrasekhar, J., Madura, J.D., Impey, R., & Klein, M.L. (1983). Comparison of simple potential functions for simulating liquid water. *Journal of Chemical Physics, 79*, 926-935.

29. Kazarian, E., Son, H., Sapao, P., Li, W., Zhang, Z., Strauss, J. F., & Teves, M. E. (2018). SPAG17 Is Required for Male Germ Cell Differentiation and Fertility. *International journal of molecular sciences*, *19*(4), 1252. https://doi.org/10.3390/ijms19041252

30. Kumar, S., Stecher, G., Li, M., Knyaz, C., & Tamura, K. (2018). MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Molecular biology and evolution*, *35*(6), 1547–1549. https://doi.org/10.1093/molbev/msy096

31. Kurcinski, M., Jamroz, M., Blaszczyk, M., Kolinski, A., & Kmiecik, S. (2015). CABS-dock web server for the flexible docking of peptides to proteins without prior knowledge of the binding site. *Nucleic acids research*, *43*(W1), W419–W424. https://doi.org/10.1093/nar/gkv456

32. Kwong, J., Lee, J. Y., Wong, K. K., Zhou, X., Wong, D. T., Lo, K. W., Welch, W. R., Berkowitz, R. S., & Mok, S. C. (2006). Candidate tumor-suppressor gene DLEC1 is frequently downregulated by promoter hypermethylation and histone hypoacetylation in human epithelial ovarian cancer. *Neoplasia (New York, N.Y.)*, *8*(4), 268–278. https://doi.org/10.1593/neo.05502

33. Lechtreck, K. F., & Witman, G. B. (2007). Chlamydomonas reinhardtii hydin is a central pair protein required for flagellar motility. *The Journal of cell biology*, *176*(4), 473–482. https://doi.org/10.1083/jcb.200611115

34. Lechtreck, K. F., Delmotte, P., Robinson, M. L., Sanderson, M. J., & Witman, G. B. (2008). Mutations in Hydin impair ciliary motility in mice. *The Journal of cell biology*, *180*(3), 633–643. https://doi.org/10.1083/jcb.200710162

35. Lee, J., Cheng, X., Swails, J. M., Yeom, M. S., Eastman, P. K., Lemkul, J. A., Wei, S., Buckner, J., Jeong, J. C., Qi, Y., Jo, S., Pande, V. S., Case, D. A., Brooks, C. L., 3rd,

MacKerell, A. D., Jr, Klauda, J. B., & Im, W. (2016). CHARMM-GUI Input Generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM Simulations Using the CHARMM36 Additive Force Field. *Journal of chemical theory and computation*, *12*(1), 405–413. https://doi.org/10.1021/acs.jctc.5b00935

36. Lee, J., Cheng, X., Swails, J. M., Yeom, M. S., Eastman, P. K., Lemkul, J. A., Wei, S., Buckner, J., Jeong, J. C., Qi, Y., Jo, S., Pande, V. S., Case, D. A., Brooks, C. L., 3rd, MacKerell, A. D., Jr, Klauda, J. B., & Im, W. (2016). CHARMM-GUI Input Generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM Simulations Using the CHARMM36 Additive Force Field. *Journal of chemical theory and computation*, *12*(1), 405–413. https://doi.org/10.1021/acs.jctc.5b00935
37. Lee, L., Campagna, D. R., Pinkus, J. L., Mulhern, H., Wyatt, T. A., Sisson, J. H., Pavlik, J. A., Pinkus, G. S., & Fleming, M. D. (2008). Primary ciliary dyskinesia in mice lacking the novel ciliary protein Pcdp1. *Molecular and cellular biology*, *28*(3), 949–957. https://doi.org/10.1128/MCB.00354-07
38. Lee, T. I., & Young, R. A. (2013). Transcriptional regulation and its misregulation in disease. *Cell*, *152*(6), 1237–1251. https://doi.org/10.1016/j.cell.2013.02.014
39. Lyu, Q., Li, Q., Li, J., Luo, J., Liu, C., Nai, S., Liu, H., Zhu, X., Song, T., Liu, M., & Zhao, H. (2026). Proteomic composition and mutual assembly of the C2a projection in vertebrate motile cilia. bioRxiv. https://doi.org/10.64898/2026.01.24.701544
40. MacKerell, A. D., Bashford, D., Bellott, M., Dunbrack, R. L., Evanseck, J. D., Field, M. J., Fischer, S., Gao, J., Guo, H., Ha, S., Joseph-McCarthy, D., Kuchnir, L., Kuczera, K., Lau, F. T., Mattos, C., Michnick, S., Ngo, T., Nguyen, D. T., Prodhom, B., Reiher, W. E., … Karplus, M. (1998). All-atom empirical potential for molecular modeling and dynamics

studies of proteins. *The journal of physical chemistry. B*, *102*(18), 3586–3616. https://doi.org/10.1021/jp973084f

41. Marko J. F. (2015). Biophysics of protein-DNA interactions and chromosome organization. *Physica A*, *418*, 126–153. https://doi.org/10.1016/j.physa.2014.07.045
42. Mitchell D. R. (2007). The evolution of eukaryotic cilia and flagella as motile and sensory organelles. *Advances in experimental medicine and biology*, *607*, 130–140. https://doi.org/10.1007/978-0-387-74021-8_11
43. Nigg, E. A., & Stearns, T. (2011). The centrosome cycle: Centriole biogenesis, duplication and inherent asymmetries. *Nature cell biology*, *13*(10), 1154–1160. https://doi.org/10.1038/ncb2345
44. Park, S. J., Kern, N., Brown, T., Lee, J., & Im, W. (2023). CHARMM-GUI PDB Manipulator: Various PDB Structural Modifications for Biomolecular Modeling and Simulation. *Journal of molecular biology*, *435*(14), 167995. https://doi.org/10.1016/j.jmb.2023.167995
45. Pazour, G. J., Agrin, N., Leszyk, J., & Witman, G. B. (2005). Proteomic analysis of a eukaryotic cilium. *The Journal of cell biology*, *170*(1), 103–113. https://doi.org/10.1083/jcb.200504008
46. Ponting C. P. (2006). A novel domain suggests a ciliary function for ASPM, a brain size determining gene. *Bioinformatics (Oxford, England)*, *22*(9), 1031–1035. https://doi.org/10.1093/bioinformatics/btl022
47. Rao, V. G., Sarafdar, R. B., Chowdhury, T. S., Sivadas, P., Yang, P., Dongre, P. M., & D'Souza, J. S. (2016). Myc-binding protein orthologue interacts with AKAP240 in the

central pair apparatus of the Chlamydomonas flagella. *BMC cell biology*, *17*(1), 24. https://doi.org/10.1186/s12860-016-0103-y

48. Rao, V. G., Shendge, A. A., D'Gama, P. P., Martis, E. A. F., Mehta, S., Coutinho, E. C., & D'Souza, J. S. (2024). A-kinase anchoring proteins are enriched in the central pair microtubules of motile cilia in Chlamydomonas reinhardtii. *FEBS letters*, *598*(4), 457–476. https://doi.org/10.1002/1873-3468.14791
49. Satir, P., & Christensen, S. T. (2007). Overview of structure and function of mammalian cilia. *Annual review of physiology*, *69*, 377–400. https://doi.org/10.1146/annurev.physiol.69.040705.141236
50. Schrödinger, L., & DeLano, W. (2020). *PyMOL*. Retrieved from http://www.pymol.org/pymol
51. Sha, Y., Wei, X., Ding, L., Ji, Z., Mei, L., Huang, X., Su, Z., Wang, W., Zhang, X., & Lin, S. (2020). Biallelic mutations of CFAP74 may cause human primary ciliary dyskinesia and MMAF phenotype. *Journal of human genetics*, *65*(11), 961–969. https://doi.org/10.1038/s10038-020-0790-2
52. Shapiro, A. J., Sillon, G., D'Agostino, D., Baret, L., López-Giráldez, F., Mane, S., Leigh, M. W., Davis, S. D., Knowles, M. R., & Zariwala, M. A. (2023). *HYDIN* Variants Are a Common Cause of Primary Ciliary Dyskinesia in French Canadians. *Annals of the American Thoracic Society*, *20*(1), 140–144. https://doi.org/10.1513/AnnalsATS.202203-253RL
53. Smith E. F. (2002). Regulation of flagellar dynein by calcium and a role for an axonemal calmodulin and calmodulin-dependent kinase. *Molecular biology of the cell*, *13*(9), 3303–3313. https://doi.org/10.1091/mbc.e02-04-0185

54. Smith E. F. (2002). Regulation of flagellar dynein by the axonemal central apparatus. *Cell motility and the cytoskeleton*, *52*(1), 33–42. https://doi.org/10.1002/cm.10031

55. Smith, E. F., & Lefebvre, P. A. (1997). PF20 gene product contains WD repeats and localizes to the intermicrotubule bridges in Chlamydomonas flagella. *Molecular biology of the cell*, *8*(3), 455–467. https://doi.org/10.1091/mbc.8.3.455

56. Smith, E. F., & Yang, P. (2004). The radial spokes and central apparatus: mechano-chemical transducers that regulate flagellar motility. *Cell motility and the cytoskeleton*, *57*(1), 8–17. https://doi.org/10.1002/cm.10155

57. Taira, T., Maëda, J., Onishi, T., Kitaura, H., Yoshida, S., Kato, H., Ikeda, M., Tamai, K., Iguchi-Ariga, S. M., & Ariga, H. (1998). AMY-1, a novel C-MYC binding protein that stimulates transcription activity of C-MYC. *Genes to cells : devoted to molecular & cellular mechanisms*, *3*(8), 549–565. https://doi.org/10.1046/j.1365-2443.1998.00206.x

58. Vangone, A., & Bonvin, A. M. (2015). Contacts-based prediction of binding affinity in protein-protein complexes. *eLife*, *4*, e07454. https://doi.org/10.7554/eLife.07454

59. Xue, L. C., Rodrigues, J. P., Kastritis, P. L., Bonvin, A. M., & Vangone, A. (2016). PRODIGY: a web server for predicting the binding affinity of protein-protein complexes. *Bioinformatics (Oxford, England)*, *32*(23), 3676–3678. https://doi.org/10.1093/bioinformatics/btw514

60. Ying, J., Poon, F. F., Yu, J., Geng, H., Wong, A. H., Qiu, G. H., Goh, H. K., Rha, S. Y., Tian, L., Chan, A. T., Sung, J. J., & Tao, Q. (2009). DLEC1 is a functional 3p22.3 tumour suppressor silenced by promoter CpG methylation in colon and gastric cancers. *British journal of cancer*, *100*(4), 663–669. https://doi.org/10.1038/sj.bjc.6604888

61. Yogesha, M., Rao, V.G., Martis, E.A., Coutinho, E.C., Gohlke, H., Chidangil, S., Dongre, P.M., & D'Souza, J.S., (2017). Structural features of FAP174, a MYCBP-1 orthologue from Chlamydomonas reinhardtii, revealed by computational and experimental analyses. *RSC Adv* **7**, 51391–51402.

62. Yukitake, H., Furusawa, M., Taira, T., Iguchi-Ariga, S. M., & Ariga, H. (2002). AMAP-1, a novel testis-specific AMY-1-binding protein, is differentially expressed during the course of spermatogenesis. *Biochimica et biophysica acta*, *1577*(1), 126–132. https://doi.org/10.1016/s0167-4781(02)00411-6

63. Zhang, H., & Mitchell, D. R. (2004). Cpc1, a Chlamydomonas central pair protein with an adenylate kinase domain. *Journal of cell science*, *117*(Pt 18), 4179–4188. https://doi.org/10.1242/jcs.01297

64. Zhang, L., Zhang, Q., Li, L., Wang, Z., Ying, J., Fan, Y., He, Q., Lv, T., Han, W., Li, J., Yang, Y., Xu, B., Wang, L., Liu, Q., Sun, Y., Guo, Y., Tao, Q., & Jin, J. (2015). DLEC1, a 3p tumor suppressor, represses NF-κB signaling and is methylated in prostate cancer. *Journal of molecular medicine (Berlin, Germany)*, *93*(6), 691–701. https://doi.org/10.1007/s00109-015-1255-5

65. Zhang, Y., Miao, Y., Yi, J., Wang, R., & Chen, L. (2010). Frequent epigenetic inactivation of deleted in lung and esophageal cancer 1 gene by promoter methylation in non-small-cell lung cancer. *Clinical lung cancer*, *11*(4), 264–270. https://doi.org/10.3816/CLC.2010.n.034

66. Zhao, L., Hou, Y., McNeill, N. A., & Witman, G. B. (2020). The unity and diversity of the ciliary central apparatus. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, *375*(1792), 20190164. https://doi.org/10.1098/rstb.2019.0164

67. Zhao, L., Hou, Y., Picariello, T., Craige, B., & Witman, G. B. (2019). Proteome of the central apparatus of a ciliary axoneme. *The Journal of cell biology*, *218*(6), 2051–2070. https://doi.org/10.1083/jcb.201902017

Tables and Figures

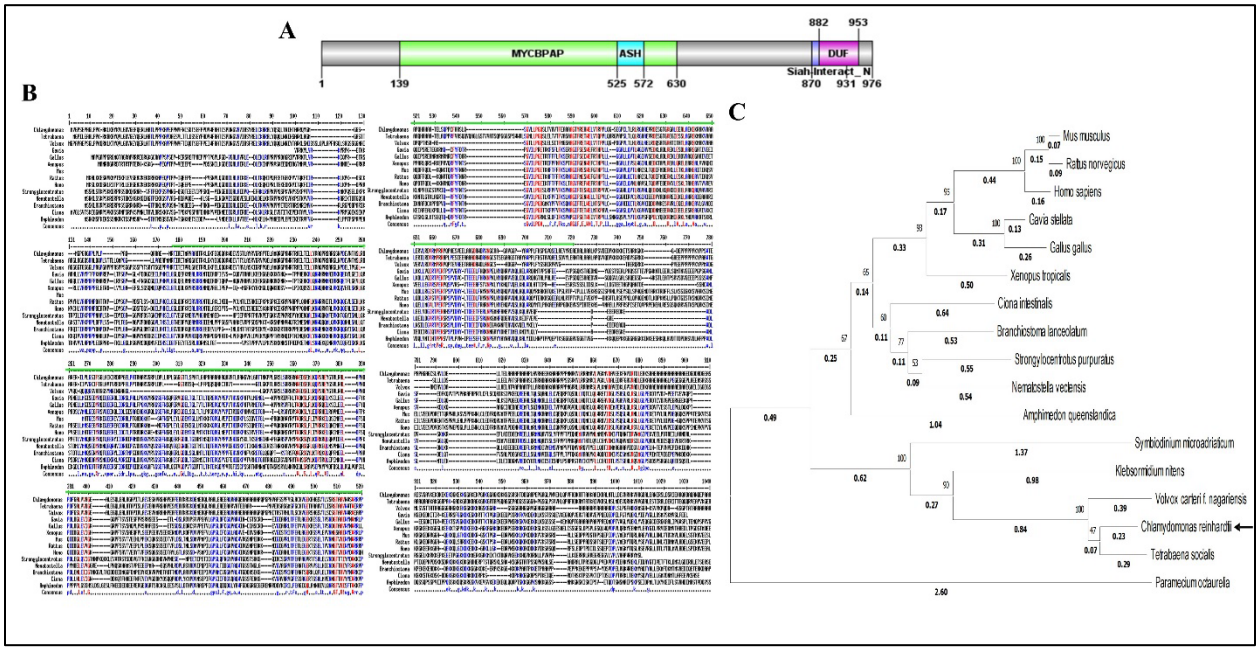

Figure 1: *In silico* analyses of FAP147. (A) Domain organization of FAP147 depicted using DOG 2.0 - Protein Domain Structure Visualization (http://dog.biocuckoo.org/). Analysis of FAP147 sequence with domain prediction databases like Pfam and CDD shows it contains a MYCBPAP domain, a domain of unknown function, i.e., DUF4407, as well as Siah-Interacting Protein, N-terminal, and ASH domain (Zhao *et al*., 2020; Ponting, 2006). **(B)** Multiple Sequence Alignment of FAP147 and Its Orthologs: Created with MultAlin software, demonstrating the alignment of the FAP147 protein with orthologs from various species. Conserved regions, such as the MYCBPAP domain, are highlighted, indicating the evolutionary conservation of this functional domain. **(C)** Phylogenetic tree of FAP147 and its orthologs. Built using the MEGA X program (Kumar *et al*., 2018) with the maximum likelihood method. The bootstrap values are shown in bold near the nodes, and the substitution per site values are near each branch. (Note: The arrow indicates *C. reinhardtii*).

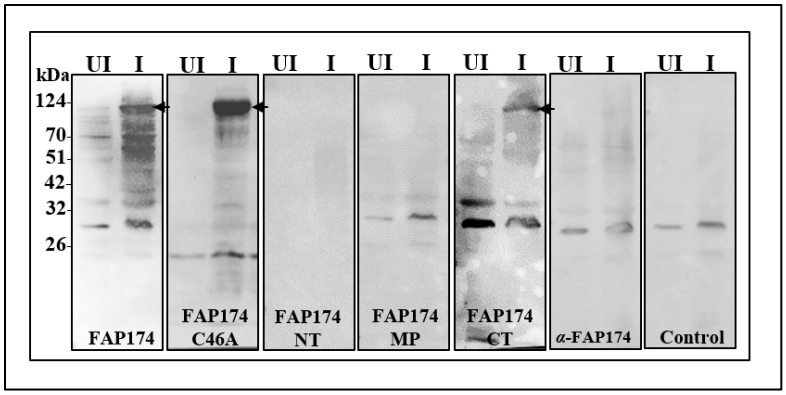

Figure 2: Interaction of FAP147 with FAP174 and its fragment polypeptides. Overlay of FAP174 and its C46A variant, N-terminus (NT), middle portion (MP), and C-terminus (CT) with uninduced (UI) and induced (I) FAP147 *E. coli* lysate. (Note: Bands represented by arrows indicate specific interactions).

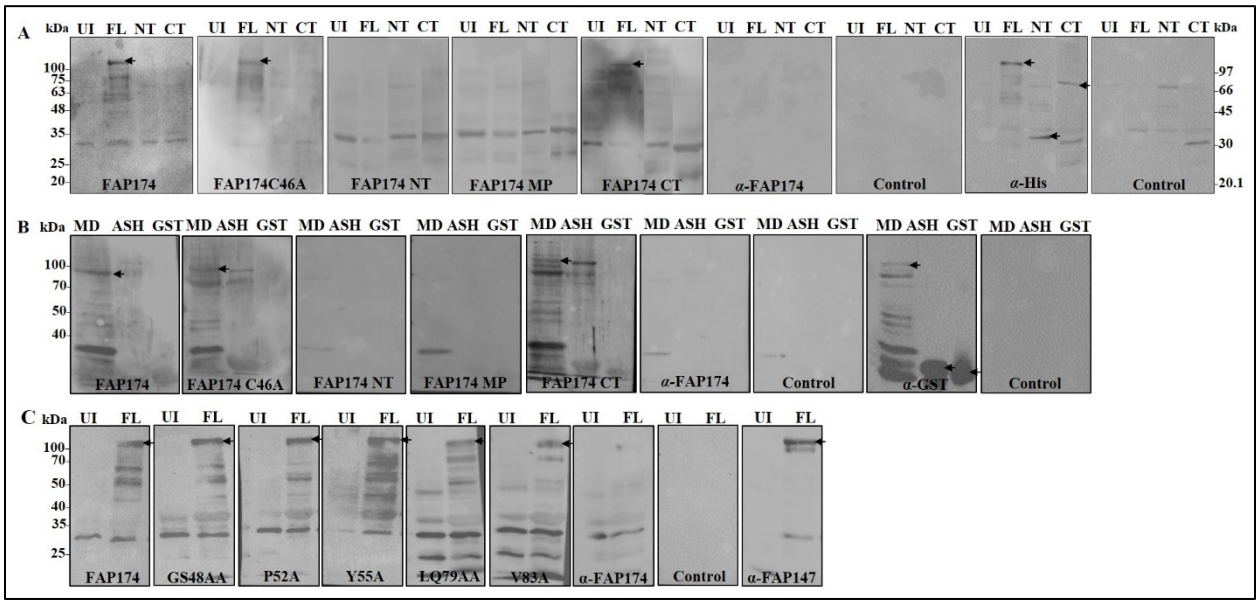

Figure 3: Interaction of FAP147 and its fragment polypeptides with FAP174 and its fragment polypeptides. (A) Overlay of 6xHis-tagged solubilized full-length (FL), N-terminus, and C-terminus fragments of FAP147 with full-length FAP174 and its fragments. **(B)** Overlay of GST-tagged purified FAP147 fragments, i.e. MYCBPAP (MD) and ASH domain, with full-length FAP174 and its fragments, using GST as a control. (**C)** Overlay of 6xHis-tagged solubilized FAP147 with purified FAP174 GS48AA, P52A, Y55A, LQ79AA and V83A variants

generated by site-directed mutagenesis (Note: Bands represented by arrow indicate specific interactions).

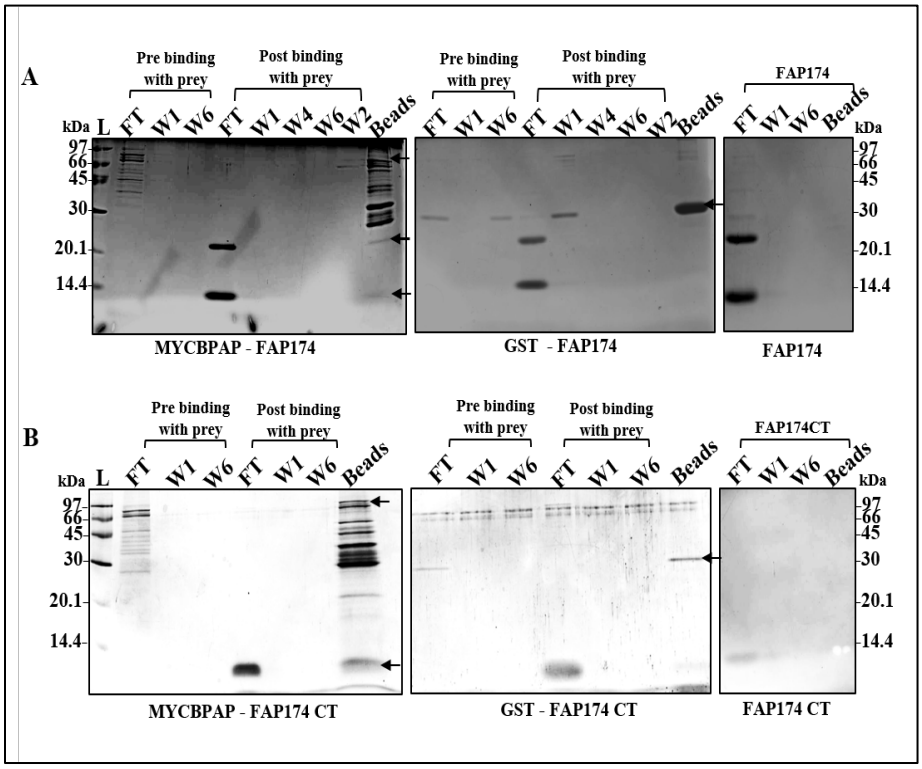

Figure 4: GST Pull-Down Assays Elucidating MYCBPAP Interaction with FAP174 and its variant. (A) Pull-down assay using GST-tagged MYCBPAP immobilized on glutathione-Sepharose (GST) beads, with FAP174 protein passed as the prey. Controls include GST alone on beads (negative control) and FAP174 directly immobilized on beads (specificity control). Bound proteins were analyzed by SDS-PAGE and visualized to assess interaction specificity **(B)** Pull-down assay using GST-tagged MYCBPAP immobilized on glutathione-Sepharose beads, with FAP174 C-terminal fragment (CT) passed as the prey. Controls include GST alone on beads (negative control) and FAP174 CT directly immobilized on beads (specificity control). Bound proteins were separated by SDS-PAGE and visualized to evaluate the interaction between MYCBPAP and the FAP174 CT construct.

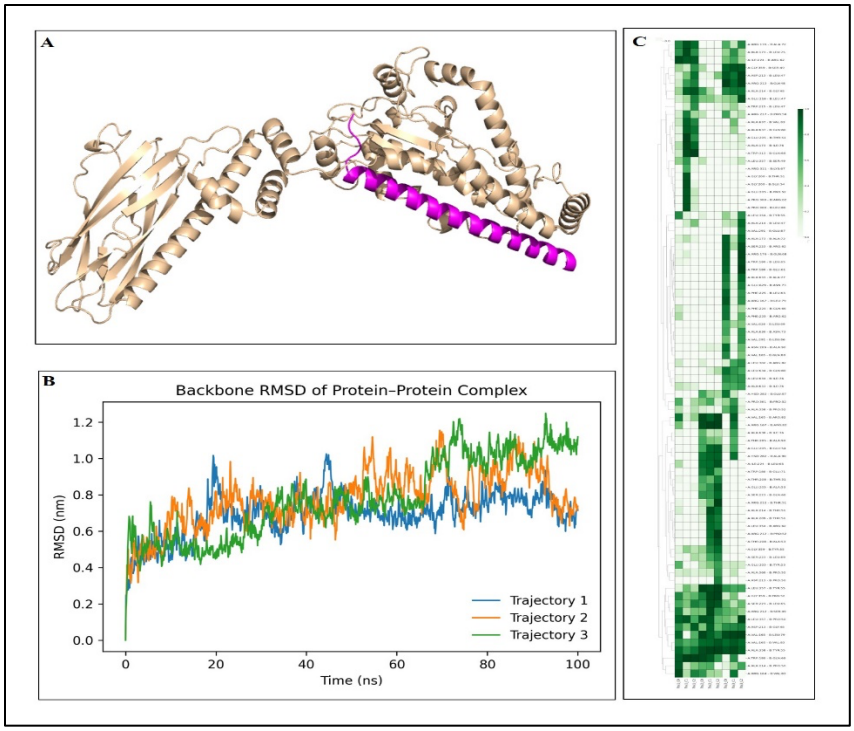

Figure 5: MD simulation studies of the FAP147 MYCBPAP domain and C-terminus of FAP174. (A) Structural model of FAP147 MYCBPAP domain and FAP174 C-terminus**.** A representative image from molecular dynamics simulations carried out for 100 ns depicting the MYCBPAP (beige) and FAP174 C-terminal region (magenta) complex. The α-helix of the FAP174 C-terminus peptide interacts with the corresponding binding interface on the FAP147 MYCBPAP domain**. (B)** Structural stability of the MD-simulated MYCBPAP domain and FAP174 C-terminus complex. The figure depicts backbone RMSD profiles from three independent 100 ns simulations, which indicate stabilization of the complex after equilibration, with fluctuations between ~0.6 and 1.0 A^{o}, suggesting a structurally stable yet flexible complex. Blue, orange, and green traces represent Trajectories 1, 2, and 3, respectively. **(C)** Residue level interaction map of FAP147 MYCBPAP domain and FAP174 C-terminus complex. The contact heat map generated using the get_contacts analysis indicates high-frequency interactions between the MYCBPAP domain and the C-terminus. (Note: Color intensity, i.e., green scale, indicates contact persistence, with darker shades representing higher contact occupancy.

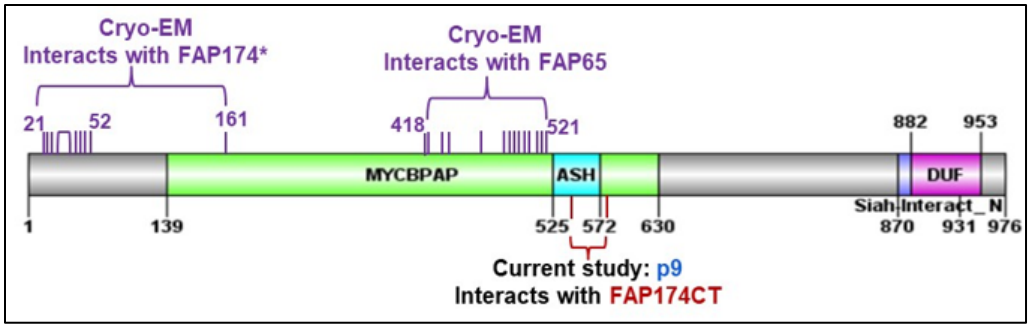

Figure 6: Schematic representation of FAP147 interacting interfaces with FAP65 and FAP174. The diagram illustrates the FAP147 protein, emphasizing critical areas associated with protein–protein interactions. Residues previously discovered by cryo-EM as interacting interface with FAP174 and FAP65 are indicated in purple. Residues found in this study that interact with the C-terminal region of FAP174, as determined by *in vitro* overlay and peptide docking investigations, are marked in blue and red, respectively. The MYCBPAP domain is highlighted in green. This model combines structural and experimental data to demonstrate suggested interaction hotspots within FAP147.

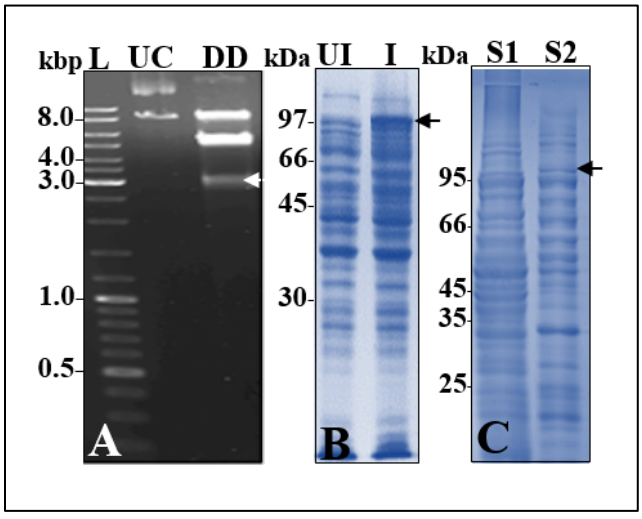

Figure S1: FAP147 Full-length (in pET-28a(+)), Amplicon –2811 bp, Expected Mw–97.04 kDa. (A) Agarose gel electrophoresis showing the restriction digestion of the target gene. The lane includes a DNA ladder, the uncut plasmid, and the restriction enzyme-digested product of ~2810bp. **(B)** A 12% SDS-PAGE gel confirming the overexpression of the recombinant protein

(~97.04kDa) in *E. coli*. Lane labelled UI and I represent the uninduced and induced lysates, respectively **(C)** SDS-PAGE (12%) gel illustrating the solubilization of the overexpressed protein (~97.04 kDa) under denaturing conditions. Lane labelled S1 and S2 denote the soluble and insoluble fractions, respectively.

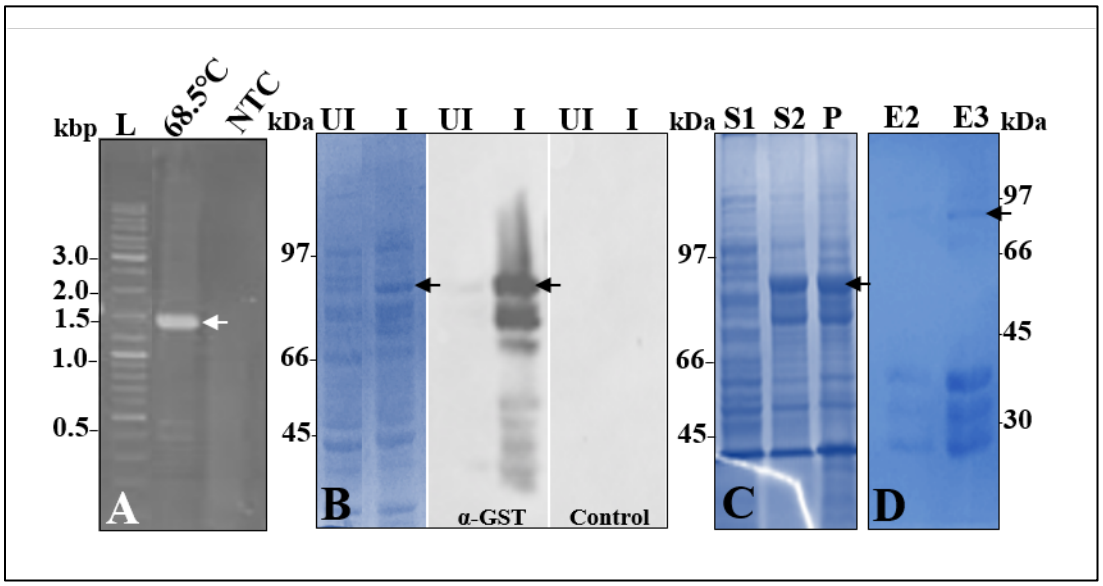

Figure S2: FAP147 MYCBPAP (in pGEX-4T-1) domain, Amplicon –1583 bp, Expected Mw –81.5 kDa (55.5 kDa +26 kDa). (A) Agarose gel electrophoresis showing the amplification of the target gene by PCR. The lane includes a DNA ladder, and the amplified product (~1583 bp) is indicated by an arrow **(B)** SDS-PAGE (10%) gel confirms the recombinant protein's overexpression (~81.5 kDa; 55.5 kDa +26 kDa) in *E. coli*. Lane UI and I represent the uninduced and induced lysates, respectively. **(C)** SDS-PAGE (10%) gel illustrating the solubilization of the overexpressed protein (~81.5 kDa) under denaturing conditions. Lane S1 and S2 denote the soluble and insoluble fractions, respectively **(D**) SDS-PAGE (12%) gel of the purified recombinant protein (~81.5 kDa) after affinity chromatography. Lanes labelled E2 and E3 show the elution consisting of the purified protein band.

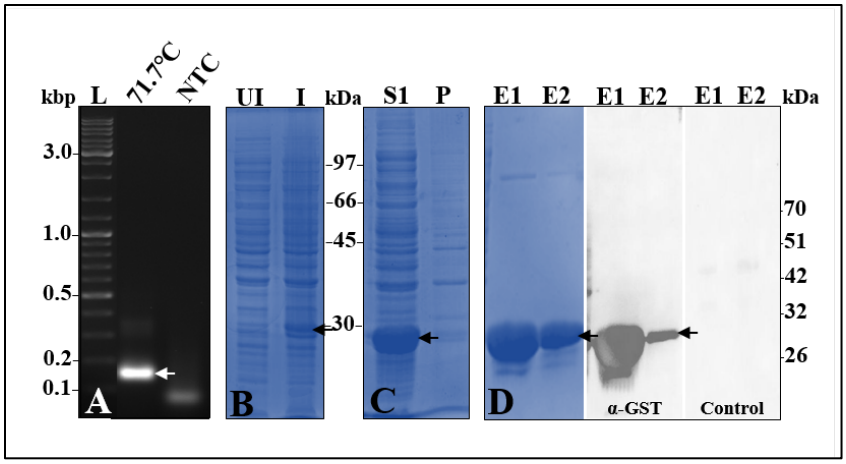

Figure S3: FAP147ASH (in pGEX-4T-1) domain, Amplicon –105 bp, Expected Mw –31.3 kDa (5.3 kDa +26 kDa). (A) Agarose gel electrophoresis showing the amplification of the target gene by PCR. The lane includes a DNA ladder, and the amplified product (~105 bp) is indicated by an arrow **(B)** SDS-PAGE (12%) gel confirms the recombinant protein's overexpression (~31.3 kDa; 5.3 kDa +26 kDa) in *E. coli*. Lane labelled UI and I represent the uninduced and induced lysates, respectively. **(C)**SDS-PAGE (12%) gel illustrating the solubilization of the overexpressed protein (~31.3 kDa) under denaturing conditions. Lane labelled S1 and P denotes the soluble and insoluble fractions, respectively **(D)** SDS-PAGE (12%) gel of the purified recombinant protein (~31.3 kDa) after affinity chromatography. Lanes labelled E1 and E2 show the elution consisting of the purified protein band.

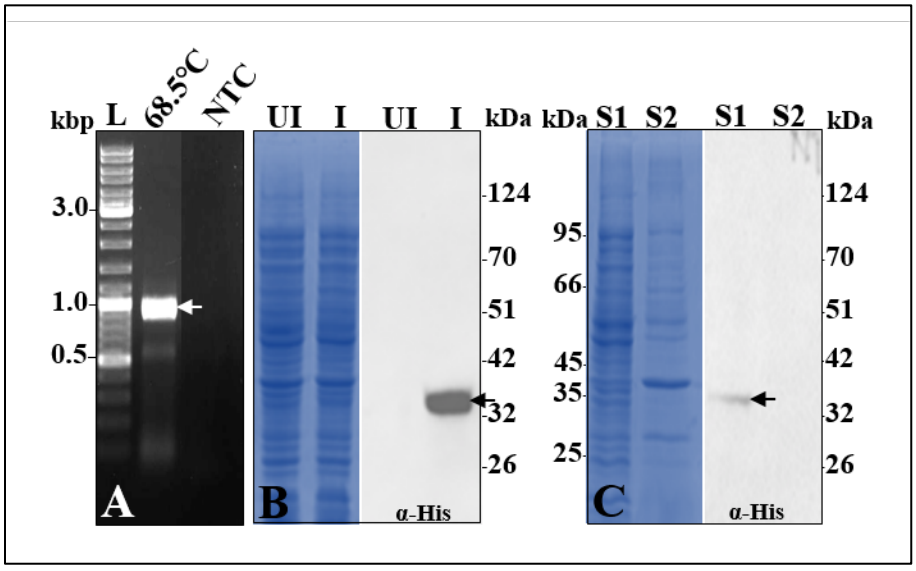

Figure S4: FAP147 N-terminus (in pET-28a(+)), Amplicon – 900 bp, Expected Mw–33.3 kDa. (A) Agarose gel electrophoresis showing the amplification of the target gene by PCR. The lane

includes a DNA ladder, and the amplified product (~900 bp) is indicated by an arrow **(B)** SDS-PAGE (12%) gel confirms the recombinant protein's overexpression (~33.3 kDa) in *E. coli*. Lane labelled UI and I represent the uninduced and induced lysates, respectively **(C)** SDS-PAGE (12%) gel illustrating the solubilization of the overexpressed protein (~33.3 kDa) under denaturing conditions. Lane labelled S1 and S2 denote the soluble and insoluble fractions, respectively.

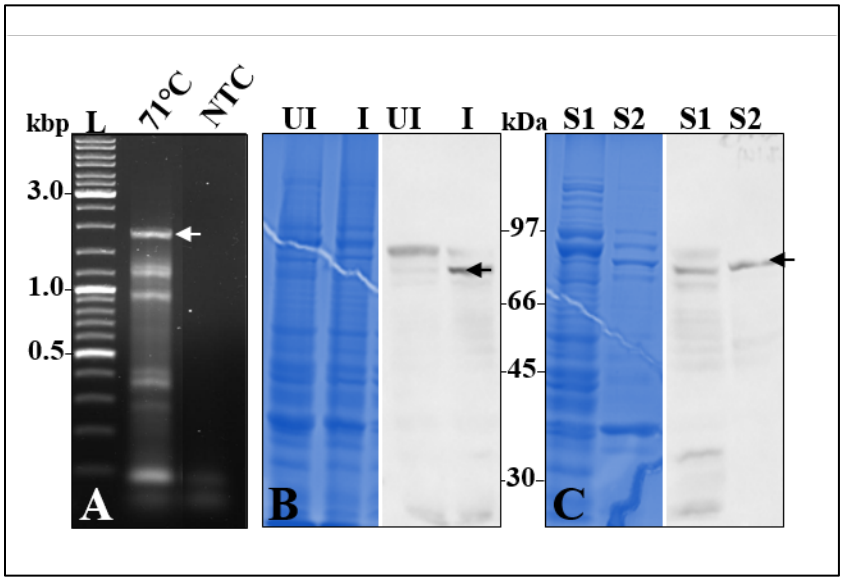

Figure S5: FAP147 C-terminus (in pET-28a(+)), Amplicon – 1902 bp, Expected Mw –66.9 kDa. (A) Agarose gel electrophoresis showing the amplification of the target gene by PCR. The lane includes a DNA ladder, and the amplified product (~1902 bp) is indicated by an arrow **(B)** SDS-PAGE (12%) gel confirms the recombinant protein's overexpression (~66.9 kDa) in *E. coli*. Lane labelled UI, and I represent the uninduced and induced lysates, respectively. **(C)**SDS-PAGE (12%) gel illustrating the solubilization of the overexpressed protein (~66.9 kDa) under denaturing conditions. Lane labelled S1 and S2 denote the soluble and insoluble fractions, respectively.

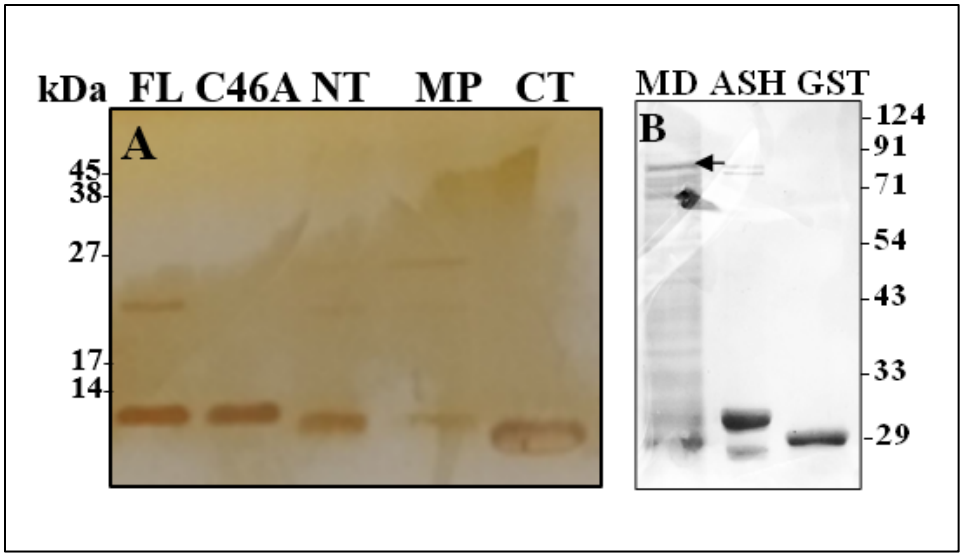

Figure S6: Silver staining of FAP174 full-length and its corresponding fragments, FAP147 fragments, i.e., MYCBPAP and ASH domain and GST. (A) An 18% SDS-PAGE gel consisting of recombinant 6xHis-tagged FAP174 full-length and its fragments. **(B)** A 12% SDS-PAGE gel consisting of GST-tagged FAP147 fragments, i.e. MYCBPAP (MD) and ASH domain. (Note – The concentration of recombinant proteins loaded onto the gel ranges from 3-5 µg, and the arrow marks the proteins with corresponding molecular mass)

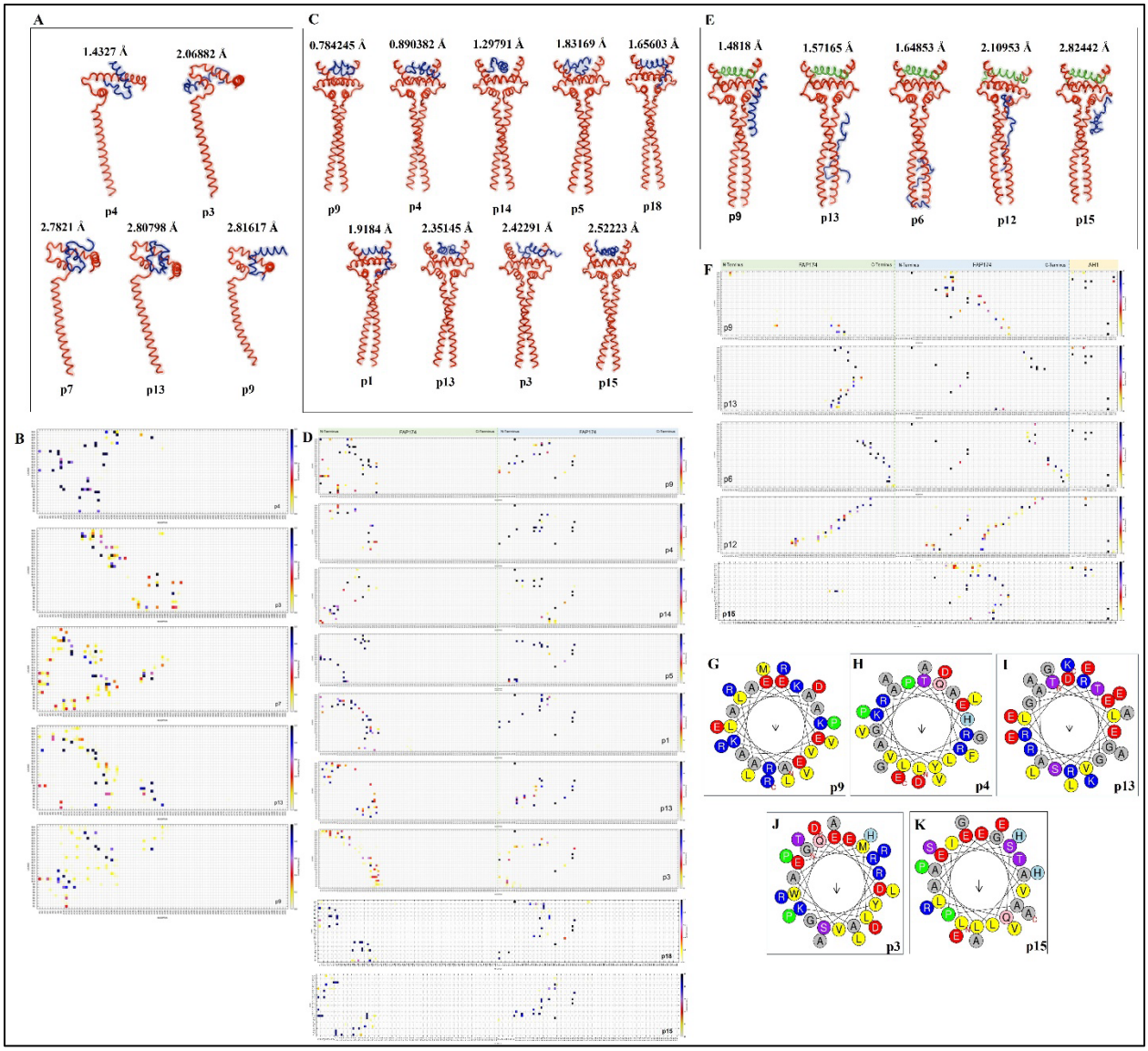

Figure S7: *In silico* docking of FAP174 with conserved amino acid stretches of FAP147 MYCBPAP domain. Predicted docking models of the FAP174 monomer **(A)**, dimer **(C)**, dimer-AH1 complex **(E)**, with conserved peptides that showed an average RMSD <3 Å. The red-coloured structure represents FAP174 or its C-terminal, the green-coloured structure represents AH1, and the docked conserved MYCBPAP peptide fragments are shown in blue colour. The FAP174 and its C-terminus (receptor) were generated using AlphaFold 3 multimer, and the conserved MYCBPAP peptides (ligand) were docked using CABS-Dock protein–peptide docking approach. The average root-mean-square deviation (RMSD) of the docked peptide conformations is indicated above each structure in Ångström (Å). **(B, D, F)** Contact map for the 1st cluster of structures arising from the docking of FAP174 monomer **(B)**, dimer **(D)**, and dimer-AH1 complex **(F)** and respective

peptide. Contact map represented for the peptides forming cluster 1 with average RMSD < 3Å. The X-axis represents the receptor residues divided into their component protein chains. The Y-axis represents the ligand peptide residues. The shade of colour represents the frequency of the contact point in the cluster. **(G-K)** Amphipathic helix prediction using the HeliQuest software for p9, p4, p13 and p3.

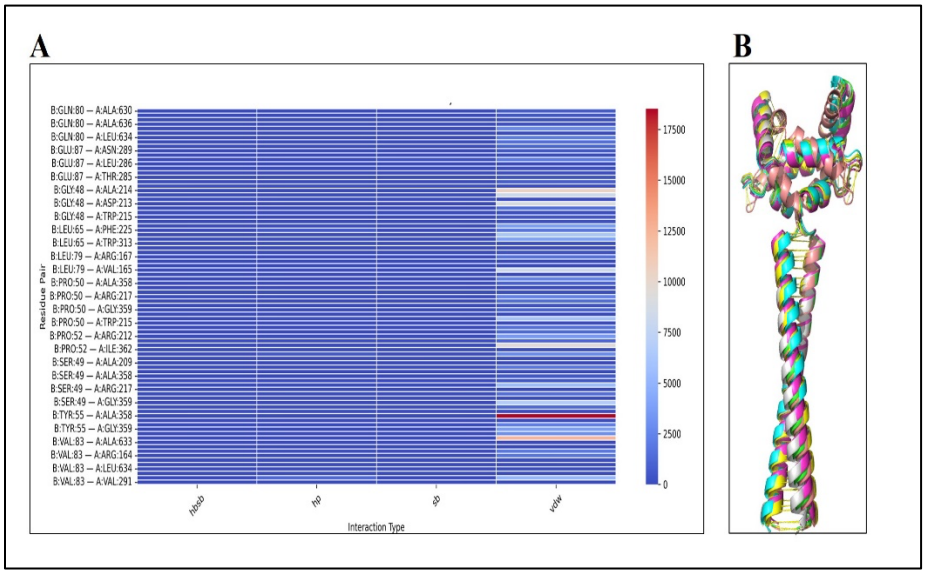

Figure S8: *In silico* representation of the FAP147 MYCBPAP domain with the C-terminus of FAP174 and the overlap of FAP174 with the variants. (A) Interaction frequency map of FAP174 C-terminal (B chain) residues with the MYCBPAP (A chain) domain. Heat map illustrating the frequency of residue–residue interactions between Chain B (FAP174 C-terminus) and Chain A (MYCBPAP domain) throughout three distinct molecular dynamics simulations. Rows denote interacting residue pairs, whereas columns specify interaction types: hydrogen bond side-chain to backbone (hbsb), hydrophobic contacts (hp), salt bridges (sb), and vdw interactions (vdw). Colour intensity represents the cumulative interaction count across trajectories, with warmer colours indicating higher interaction persistence**. (B)** Structural alignment of wild-type and variant FAP174 dimers. Ribbon representation of the superposition of wild-type FAP174 (beige) with all five variants generated in the current study. The close overlap indicates that introduced mutations do not disrupt the overall dimeric architecture of FAP174.

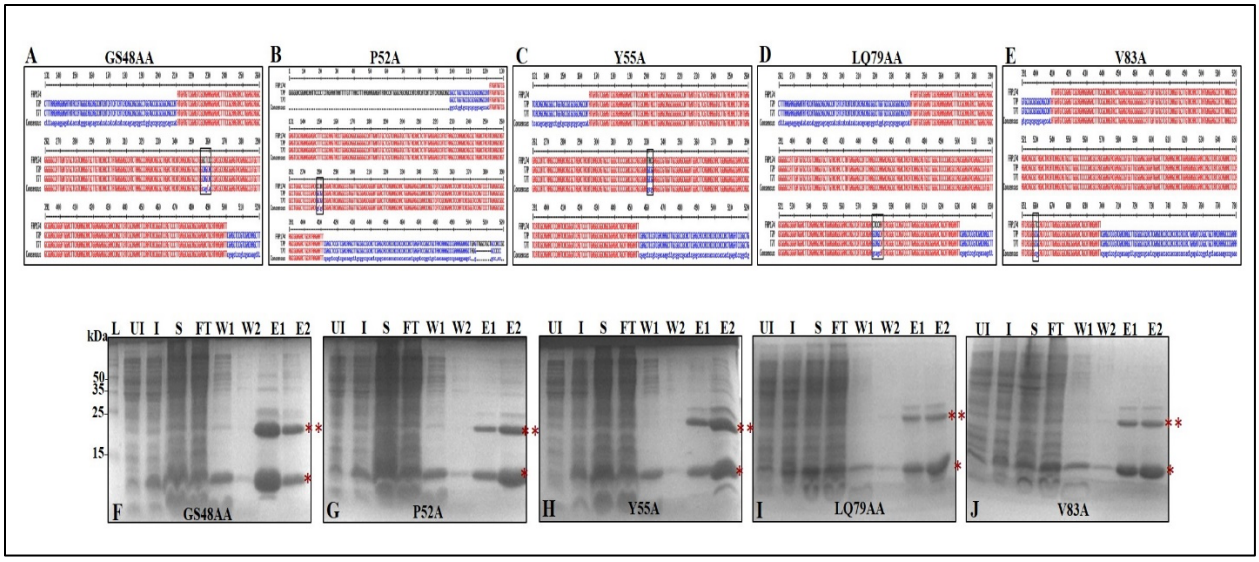

Figure S9: **Generation of FAP174 variants using recombinant DNA technology** (**A)** Sanger sequencing results of the plasmid DNA of Fap174 variants generated *via* site-directed mutagenesis. Fap174 is the wild-type gene, with the start ATG codon and the stop ATT codon. The GCG and GCC codons code for alanine. The box represents the original bases of the Fap174 gene, and the mutated bases using the T7 promoter (T7P, Forward primer) and T7 terminator (T7T, Reverse primer). (**B)** An 18% SDS-PAGE gel representing overexpression (I), solubilization (S), and purification of the FAP174 variants (FT- flowthrough, W1 and W2- washes, E1 and E2- elutions of the purified recombinant FAP174 variant proteins). The asterisk represents the recombinant variant protein of FAP174, with the monomer represented by a single asterisk and the dimer form represented by a double asterisk.

Table 1: Physicochemical properties of *C. reinhardtii* FAP147 done using the ProtParam tool (https://web.expasy.org/protparam/). The table outlines essential characteristics of the FAP147 gene and its corresponding protein, which serves as a ciliary component in *C. reinhardtii*. It includes genomic features such as the open reading frame (ORF) lengths, and GC content, which are obtained through sequence analysis. Protein properties include molecular mass derived from amino acid composition; theoretical isoelectric point (pI) calculated using pKa values of charged residues; *in vitro* and *in vivo* half-life indicating protein stability; counts of positively charged and negatively charged amino acids; instability index reflecting degradation propensity; aliphatic index

denoting thermostability; and Grand Average of Hydropathicity (GRAVY) index indicating hydrophobicity (positive) or hydrophilicity (negative).

Property of the gene/protein	**FAP147**
ORF length	2811 bp
Percentage GC content	72.04
No.of amino acid	936
Molecular weight (in dalton)	97.04
Theoretical pI	6.72
Half life	30 hours (mammalian reticulocytes, in vitro); >20 hours (yeast, in vivo); >10 hours (Escherichia coli, in vivo)
(Asp + Glu) negatively charged amino acid	113
(Arg + Lys) positively charged amino acid	110
Instability index	38.89
Aliphatic index	79.54
Grand average of hydropathicity (GRAVY)	-0.265

Table 2: Protein BLAST analysis of FAP147 using NCBI-Protein BLAST tool (https://blast.ncbi.nlm.nih.gov/Blast.cgi). Identified homologous sequences are presented along with their respective scores, query coverage percentages, identity percentages, and E-values thereby providing insights into their similarity and statistical relevance.

Accession Number	**Organism**	**Protein**	**% Query Coverage**	**% Identity**	**E-value**
XP_002951886.1	*Volvox carteri f. Nagariensis* (Green algae)	hypothetical protein VOLCADRAFT_92519	85	54.26	0

PNH10577.1	*Tetrabaena socialis* (Green algae)	MYCBP-associated protein	90	61.80	0
CAE7666388.1	*Symbiodinium microadriaticum* (Dinoflagellate, alveolata)	Mycbpap	46	26.58	3e-21
GAQ89000.1	*Klebsormidium nitens* (Green plant)	hypothetical protein KFL_004770070	39	26.52	1e-14
CAD8145242.1	*Paramecium octaurelia* (Ciliate)	unnamed protein product	13	25.19	3e-08
XP_032226527.2	*Nematostella vectensis* (Sea anemones)	MYCBP-associated protein	27	26.59	9e-13
CAH1238026.1	*Branchiostoma lanceolatum* (Lancelets)	MYCBPAP	51	31.37	1e-12
XP_026693355.1	*Ciona intestinalis* (Tunicates)	MYCBP-associated protein-like isoform X3	72	25.51	3e-11
XP_030828290.1	*Strongylocentrotus purpuratus* (Sea urchins)	MYCBP-associated protein isoform X6	32	25.70	2e-13
XP_019849715.1	*Amphimedon queenslandica* (Sponges)	PREDICTED: MYCBP-associated protein-like isoform X1	14	25.69	3e-10
KFV54418.1	*Gavia stellata* (Birds)	MYCBP-associated protein	27	31.48	5e-16
XP_040542612.1	*Gallus gallus* (Birds)	MYCBP-associated protein isoform X1	35	26.82	3e-15
XP_012808371.2	*Xenopus tropicalis* (Frogs and toads)	MYCBP-associated protein	17	32.35	5e-08

AAH27327.1	*Mus musculus* (Rodent)	MYCBP-associated protein	10	33.02	2e-07
XP_038942465.1	*Rattus norvegicus* (Rodent)	MYCBP-associated protein isoform X4	10	31.73	7e-07
NP_115509.5	*Homo sapiens* (Primate)	MYCBP-associated protein isoform 1	11	31.62	5e-07

	MYCBPAP Domain		**Avg. RMSD (Å)**				**No. of times Avg. RMSD <3.0Å**
Peptide	**Amino acids**	**Peptide sequence**	**FAP174 monomer**	**FAP174 C-Terminus**	**FAP174 dimer**	**FAP174 dimer-AH1**	
p1	307-338	SARGRAWEDSEMLKQRVAEYGTRLRALAPH	7.04	5.54	1.92	3.26	1
p2	312-341	AWEDSEMLKQRVAEYGTRLRALAPHDPDFG	3.25	6.37	4.56	4.31	0
p3	310-339	GRAWEDSEMLKQRVAEYGTRLRALAPHDPD	2.07	5.06	2.42	5.49	2
p4	319-348	LKQRVAEYGTRLRALAPHDPDFGALVVAGE	1.43	3.42	0.89	8.48	2
p5	436-461	GEKAHGSVTLVSRGTAAVNWSWRRVP	8.34	5.75	1.83	6.47	1
p6	484-513	SGVLLPGQSLTVAVTFEAAAAGTYREAWEL	5.56	8.46	5.13	1.65	1
p7	490-519	GQSLTVAVTFEAAAAGTYREAWELVTRPPL	2.78	8.51	5.15	5.78	1
p8	537-566	VRDESGTGRGALEEALAEKEKRAKVAAALE	3.48	5.67	4.08	4.53	0
p9	547-576	ALEEALAEKEKRAKVAAALERVLRDVRMPR	2.82	8.57	0.78	1.48	3
p10	575-604	PRRPQPHESVEELAAGDAWDRVNGCAAGAV	4.08	8.89	8.61	3.89	0
p11	580-609	PHESVEELAAGDAWDRVNGCAAGAVGGPYA	3.00	4.44	5.13	9.61	1
p12	598-623	GCAAGAVGGPYAWPPLFHSPGVQSEL	8.54	9.66	5.82	2.11	1

p13	528-557	TLRLRGAAEVRDESGTGRGALEEALAEKEK	2.81	4.64	2.35	1.57	3
p14	533-562	GAAEVRDESGTGRGALEEALAEKEKRAKVA	8.75	7.97	1.29	5.29	1
p15	350-379	LESQLEALAGAPITLAEVESHVRGHAPEA	4.48	4.94	2.52	2.82	2
p16	198-227	ARGSNPDEVWTASLRDAWERILPLGSIFSG	4.87	6.22	8.06	7.26	0
p17	336-366	DPDFGALVVAGEALESQLEALAGAPITLAE	3.33	7.60	3.72	3.09	0
p18	177-206	GPAWRTRAELTELLYRAQPRDARGSNPDEV	3.65	7.99	1.66	6.80	1
p19	172-201	LAMQRGPAWRTRAELTELLYRAQPRDARGS	4.78	9.24	4.17	3.25	0

Table 3: RMSD values of the interaction between FAP174 and its CT with various amino acid stretches of the FAP147 MYCBPAP domain. Docking of conserved regions (peptides) of FAP147 MYCBPAP domain docked with FAP174 monomer, FAP174 C-terminus, FAP174 dimer and a complex of FAP174 dimer-AH1 using CABS-dock. The docking was performed against the FAP174 monomer, C-terminal fragment, dimer, and a dimer complexed with AH1. The table enlists the average RMSD of each interaction. RMSD value of <3Å denotes significant interaction between the two proteins.

Protein	Sequence	Calculated MW (kDa)	Plasmid used for cloning
(A) FAP147 full-length and fragments			
FAP147 FL (936 a.a)	MVGPSEPAGLPPKAKLKAYKVLGDVGEFQERLAATLPPPKAPEPPNVFNISDTSEFPPDVWFAATISPDNGVAVIASYRELCKARALYQSQLTNIHTARRQYGMGGSKAGEVRAAASQGPAWRTRAELTELLYRAQPRDARGSNPDEVWTASLRDAWERILPLGSIFSGLAIKIRDRPGELPATRAARVGRPLDPLLAPLGGGGTTLSPATLGHPAAAAAHAQHVATLAANGVMLGATTKWQAAAGAQLISTRAGVGGLGMLKQRVAEYGTRLRALAPHDPD	97.04	pET-28a(+)

	FGALVVAGEALESQLEALAGAPITLAEVESHV RGHAPEAFEAYKAVKDEMEAQLRAALEAEE AQRAAQEAAAAAAAPQPGPHVAFSSPFVSLA CHVGEKAHGSVTLVSRGTAAVNWSWRRVPA PQHAHAATELSQPPCFAASLQSGVLLPGQSLT VAVTFEAAAAGTYREAWELVTRPPLQGSEGP CLTLRLRGAAEVRDESGTGRGALEEALAEKE KRAKVAAALERVLRDVRMPRRPQPHESVEEL AAGDAWDRVNGCAAGAVGGPYAWPPLFHSP GVQSELAEVYREAEAALAAALKPSVDEDPKK KKKGARPAASMCACALFAEGETSARKGKKG DEPPPPPPKYPPSWTEPAPAGAGCSLAVLDEL LTELAAAAPAAAAPLAVRAEEAKARARVPPN NRKVLRRAMRVLVGKMVDAVEEKFNVIRTD LEAKSAAEAAAAAAAAAAAAGEDGDDADGE ASAISVGAVKEKDKEKEKDGKEKKGGAKEKP GAKGGAGKDKKGGSASHTDDGAAPEPGAGL PNVEHLQPVFHDKATTAAKAAIKAALNTLLD EVHRHKVAAGDALEANIRSLDSRIEAAQADA QANQEPAAAEAAAAAGLALFAPAAGGVAPG PGPAARARSAELEQLYWERLCVLRTWTGLGI PTRADVR		
FAP147 NT (1-299 a.a.)	MVGPSEPAGLPPKAKLKAYKVLGDVGEFQER LAATLPPPKAPEPPNVFNISDTSEFPPDVWFAA TISPDNGVAVIASYRELCKARALYQSQLTNIH TARRQYGMGGSMGPGAGAPLPLPPAQQAAA ENRTIAETMRVAADTRKLGKFEQQQAKWDE VSSTLAYRVGRAPTELAMQRGPAWRTRAELT ELLYRAQPRDARGSNPDEVWTASLRDAWERI LPLGSIFSGLAIKIRDRPGELPATRAARVGRPL DPLLAPLGGGGTTLSPATLGHPAAAAAHAQH VATLAANGVMLGATTNK	31.77	pET-28a(+)
FAP147 MYCBPAP domain (129-639 a.a.)	RTIAETMRVAADTRKLGKFEQQQAKWDEVSS TLAYRVGRAPTELAMQRGPAWRTRAELTELL YRAQPRDARGSNPDEVWTASLRDAWERILPL GSIFSGLAIKIRDRPGELPATRAARVGRPLDPL LAPLGGGGTTLSPATLGHPAAAAAHAQHVAT LAANGVMLGATTNKPPLGRSLSARGRAWEDS EMLKQRVAEYGTRLRALAPHDPDFGALVVA GEALESQLEALAGAPITLAEVESHVRGHAPEA FEAYKAVKDEMEAQLRAALEAEEAQRAAQE AAAAAAAPQPGPHVAFSSPFVSLACHVGEKA HGSVTLVSRGTAAVNWSWRRVPAPQHAHAA TELSQPPCFAASLQSGVLLPGQSLTVAVTFEA AAAGTYREAWELVTRPPLQGSEGPCLTLRLR	54.19	pGEX-4T-1

	GAAEVRDESGTGRGALEEALAEKEKRAKVAA ALERVLRDVRMPRRPQPHESVEELAAGDAW DRVNGCAAGAVGGPYAWPPLFHSPGV QSELAEVYREAEAALAAALK		
FAP147 ASH domain (528-562 a.a.)	TLRLRGAAEVRDESGTGRGALEEALAEKEKR AKVA	3.73	pGEX-4T-1
FAP147 CT (342-976 a.a.)	ALVVAGEALESQLEALAGAPITLAEVESHVRG HAPEAFEAYKAVKDEMEAQLRAALEAEEAQ RAAQEAAAAAAAPQPGPHVAFSSPFVSLACH VGEKAHGSVTLVSRGTAAVNWSWRRVPAPQ HAHAATELSQPPCFAASLQSGVLLPGQSLTVA VTFEAAAAGTYREAWELVTRPPLQGSEGPCL TLRLRGAAEVRDESGTGRGALEEALAEKEKR AKVAAALERVLRDVRMPRRPQPHESVEELAA GDAWDRVNGCAAGAVGGPYAWPPLFHSPGV QSELAEVYREAEAALAAALKPSVDEDPKKKK KETSARKGKKGDEPPPPPPKYPPSWTEPAPAG AGCSLAVLDELLTELAAAAPAAAAPLAVRAE EAKARARVPPNNRKVLRRAMRVLVGKMVDA VEEKFNVIRTDLEAKSAAEAAAAAAAAAAAA GEDGDDADGEASAISVGAVKEKDKEKEKDG KEKKGGAKEKPGAKGGAGKDKKGGSASHTD DGAAPEPGAGLPNVEHLQPVFHDKATTAAKA AIKAALNTLLDEVHRHKVAAGDALEANIRSL DSRIEAAQADAQANQEPAAAEAAAAAGLALF APAAGGVPAPGPGPAARARSAELEQLYWERL CVLRTWTGLGIPTRADVR	65.55	pET-28a(+)
(B) FAP174 full-length and fragments			
FAP174 FL (1-93 a.a.)	MSESQKETFRKYLEQAGAIDVLVKVLVQLYE EPSKPKTALDYIKQCLGSPTPAEYEAVVAERD GLQKQLEEA NQLIAELQSRVQSLEAAAETA	10.36	pET-28a(+)
FAP174 NT (1-46 a.a.)	MSESQKETFRKYLEQAGAIDVLVKVLVQLYE EPSKPKTALDYIKQC	5.31	pET-28a(+)
FAP174 MP (24-70 a.a)	KVLVQLYEEPSKPKTALDYIKQCLGSPTPAEY EAVVAERDGLQKQLE	5.27	pET-28a(+)
FAP174 CT (47-93 a.a.)	LGSPTPAEYEAVVAERDGLQKQLEEANQLIAE LQSRVQSLEAAAETA	5.02	pET-28a(+)

FAP174 C46A (1-93 a.a.)	MSESQKETFRKYLEQAGAIDVLVKVLVQLYE EPSKPKTALDYIKQ**A**LGSPTPAEYEAVVAERD GLQKQLEEA NQLIAELQSRVQSLEAAAETA	10.29	pET-28a(+)

Table S1: List of protein sequences for FAP147 and FAP174 used in this study. The protein sequences for the full-length and corresponding DNA fragments cloned into respective plasmids, and the expected molecular mass without the tag, have been shown.