Appendix Data for

Narcis A. Petriman1\*,Marta Loureiro-López2, Michael Taschner3, Nevin K. Zacharia1, Magdalena M. Georgieva4, Niels Boegholm1, Jiaolong Wang1, André Mourão5, Robert B. Russell4, Jens S. Andersen2 and Esben Lorentzen1\*

1 Department of Molecular Biology and Genetics, Aarhus University, Universitetsbyen 81, 8000 Aarhus C, Denmark

2 Department for Biochemistry and Molecular Biology, University of Southern Denmark, Campusvej 55, 5230 Odense M, Denmark

3 Department of Fundamental Microbiology, University of Lausanne, CH-1015 Lausanne, Switzerland

4 BioQuant, Heidelberg University, Im Neuenheimer Feld 261, 69120, Heidelberg, Germany

5 Institute of Structural Biology, Helmholtz Zentrum München, Ingolstädter Landstrasse 1, 85764 Neuherberg, Germany

\* Corresponding authors: E-mail: narcispetriman@mbg.au.dk and el@mbg.au.dk

Contents

[Appendix Figure S1: Purification and crosslinking of CrIFT88sumo-1-437/70/52281-430/46188-319/81/74128-C/27/251-136/22 and CrIFT88/521-335-GST/57/38/54/20 complexes 2](#_Toc116548903)

[Appendix Figure S2: AF model of IFT81/74/27/25/22 and validation of the IFT22 binding site 5](#_Toc116548904)

[Appendix Figure S3: AF model of IFT81/74/27/25/46/52 complex 6](#_Toc116548905)

[Appendix Figure S4: CrIFT81/74 interacts with the CrIFT88/70/52/46 complex through either their N- or C- termini 8](#_Toc116548906)

[Appendix Figure S5: AlphaFold predicted model of the CrIFT57/38/54135-C/20 tetramer colored according to the pLDDT scores 9](#_Toc116548907)

[Appendix Figure S6: AlphaFold predictions of CrIFT172/80, CrIFT172/57 and CrIFT80/38 complexes 11](#_Toc116548908)

[Appendix Figure S7: IFT88 bridges the interaction between IFT-B1 and IFT-B2 complexes 12](#_Toc116548909)

[Appendix Figure S8: Global mapping of DSBU crosslinking onto the structural model of the 15 subunits IFT-B complex 14](#_Toc116548910)

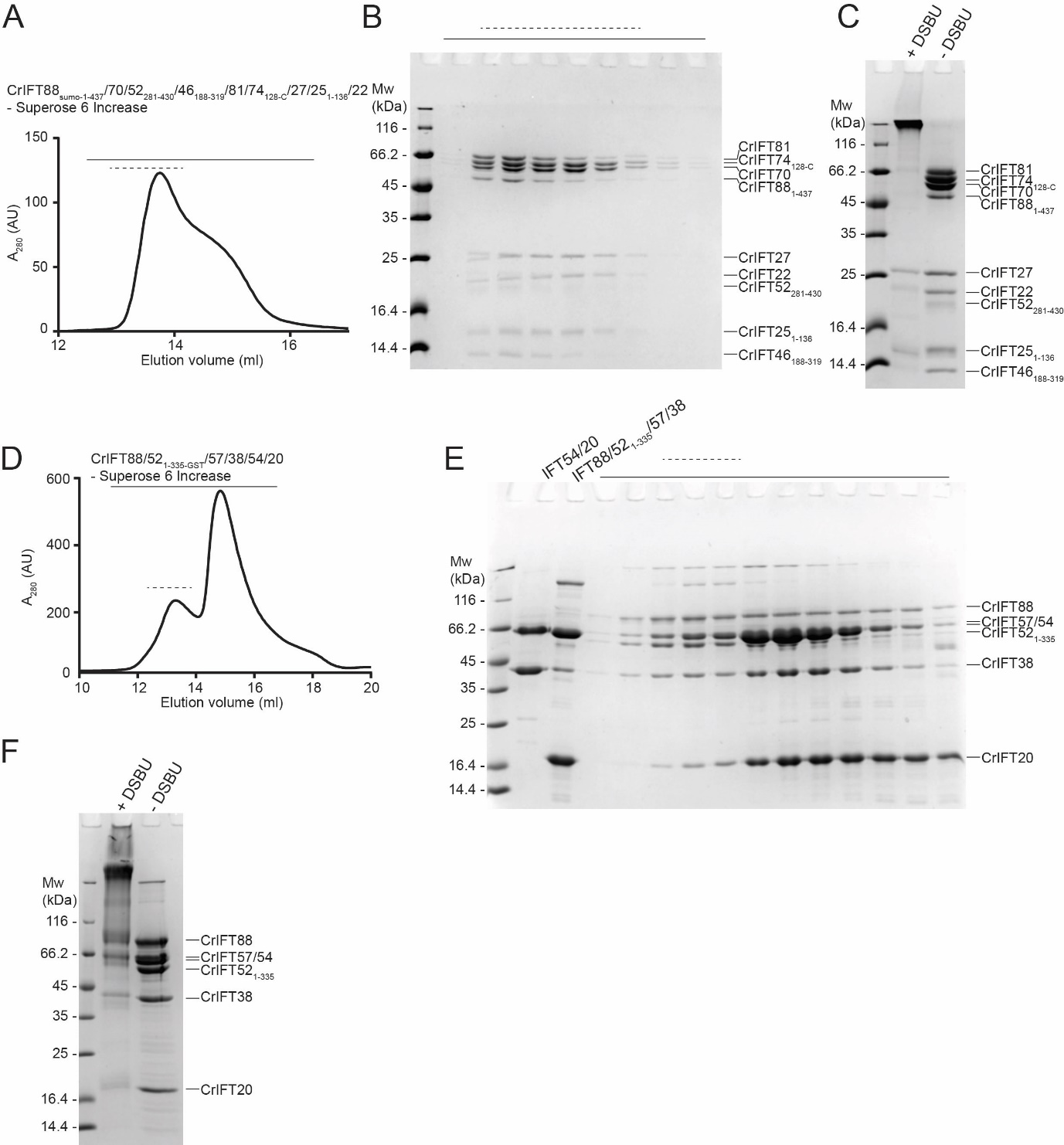
[Appendix Figure S9: Conservation and electrostatic plots of the 15-subunit IFT-B complex structure 16](#_Toc116548911)

[Appendix Figure S10: Benign (GnomAD) variants for human IFT-B proteins mapped onto *Chlamydomonas* 15-subunit IFT-B structure 16](#_Toc116548912)

[Appendix Table S1: X-ray data collection and refinement statistics for the CrIFT70/52/88 and CrIFT70/52/46 crystals. 17](#_Toc116548913)

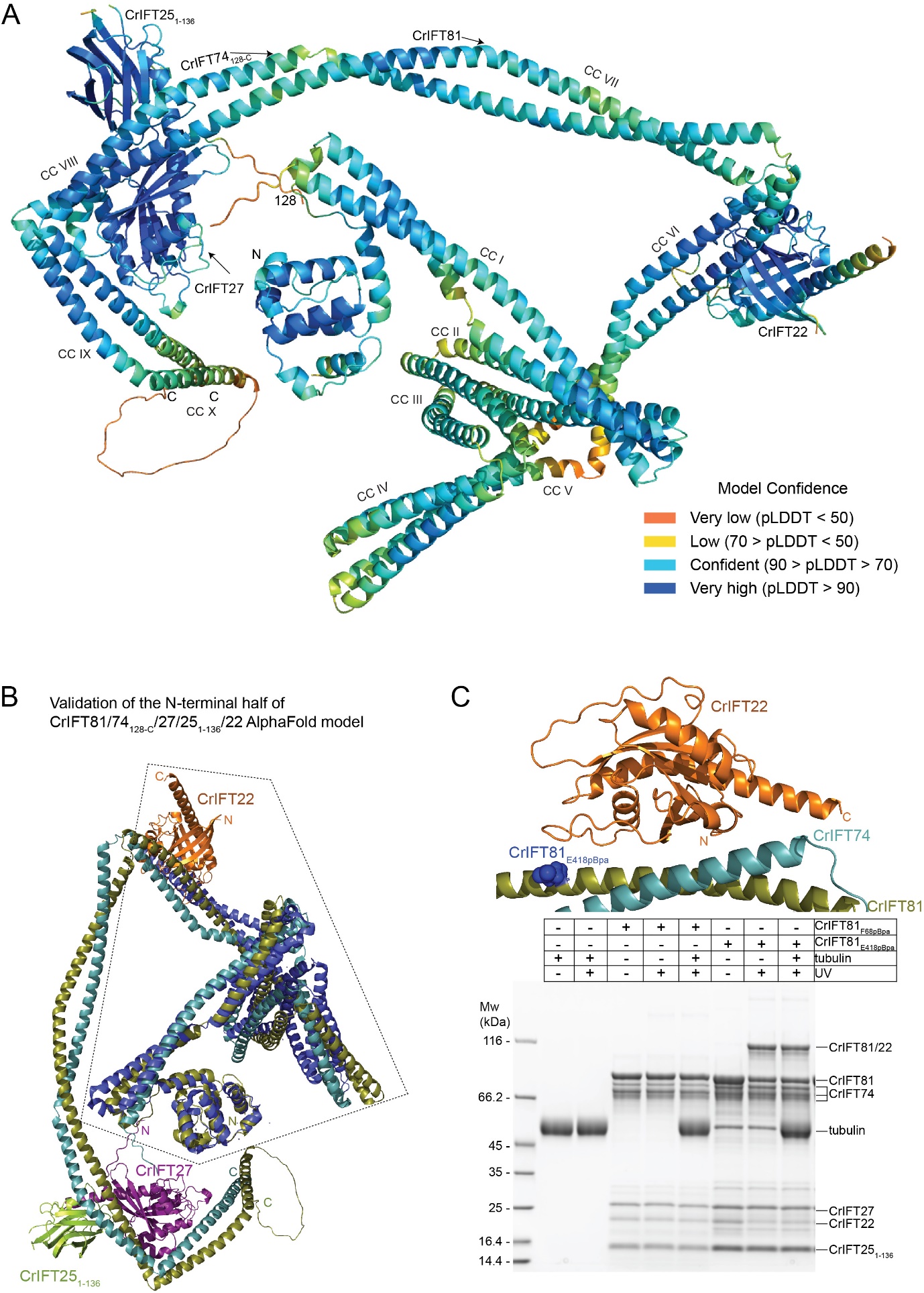
[Appendix Table S2: SAXS Data collection parameters for IFT381-133 and IFT801-582/381-133 complex 19](#_Toc116548914)

[Supplementary References 19](#_Toc116548915)



# Appendix Figure S1: Purification and crosslinking of CrIFT88sumo-1-437/70/52281-430/46188-319/81/74128-C/27/251-136/22 and CrIFT88/521-335-GST/57/38/54/20 complexes

**(A)** SEC profile of purified CrIFT88sumo-1-437/70/52281-430/46188-319/81/74128-C/27/251-136/22 complex. **(B)** All fractions illustrated above the SEC profile in panel (A) with a solid horizontal line are verified for purity on SDS PAGE and stained with Coomassie. The dashed lines in (A) and (B) represent the SEC fractions that were pooled for crosslinking. **(C)** The pooled SEC fractions were crosslinked with 0.25mM DSBU and subjected to MS analysis. **(D)** SEC with incubated CrIFT54/20 and CrIFT88/521-335-GST/57/38 complexes to yield a CrIFT88/521-335-GST/57/38/54/20hexameric complex. The SEC fractions highlighted with a solid line were monitored on SDS PAGE for purity (E) and fractions containing all 6 subunits indicated by a dashed line were pooled, crosslinked as shown in (F) and subjected to MS analysis.



# Appendix Figure S2: AF model of IFT81/74/27/25/22 and validation of the IFT22 binding site

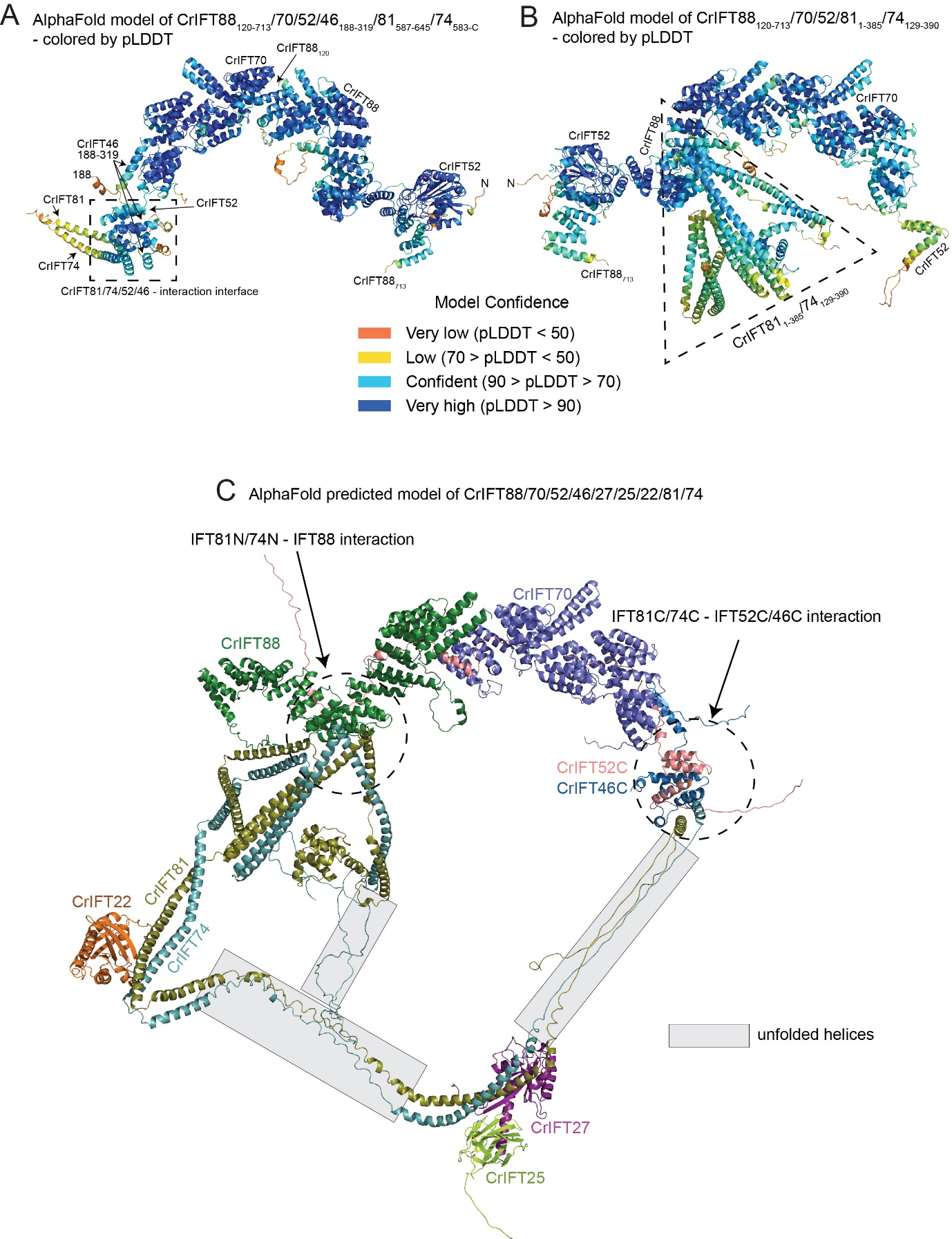
**(A)** AlphaFold predicted structure of CrIFT81/74128-C/27/251-136/22 colored according to pLDDT scores. **(B)** The N-terminal half of the predicted CrIFT81/74128-C/27/251-136/22 model superimposed onto the crystal structure of TbIFT81N/74N/22 (PDB accession number 6ian; colored blue). **(C)** The position of CrIFT22 in the predicted AlphaFold model is validated by site directed photo-crosslinking. The recombinant CrIFT81/74/27/251-136/22 complex containing acid p-benzoyl-L-phenyl-alanine (pBpa) instead of its natural E418 amino acid in CrIFT81 was activated by UV light (365nm wavelength) and showed a strong CrIFT81-CrIFT22 crosslink as illustrated on SDS PAGE and Coomassie staining (last two lanes). The CrIFT81-CrIFT22 crosslink formed at the interaction interface predicted by AlphaFold and was independent of tubulin cargo, which binds to the N-termini of CrIFT81/74. As a negative control, we used a recombinant CrIFT81/74/27/251-136/22 complex that had the F68 residue located at the N-terminus of CrIFT81, far from the CrIFT22 binding site, replaced by pBpa (lanes 3 - 5).

**Diagram

Description automatically generated**

# Appendix Figure S3: AF model of IFT81/74/27/25/46/52 complex

**(A**) The AF structure prediction of CrIFT81460-C/74460-C/27/251-136/46148-328/52382-C complex colored according to chains. **(B**) The CrIFT81460-C/74460-C/27/251-136/46148-328/52382-C structure is predicted with very high confidence (pLDDT scores >70 except for less ordered termini) and illustrates the docking of CrIFT27/251-136 heterodimer and CrIFT52382-C/46148-328 on the C-terminal half of CrIFT81/74. **(C)** The PAE plot of CrIFT81460-C/74460-C/27/251-136/46148-328/52382-C complex.



# Appendix Figure S4: CrIFT81/74 interacts with the CrIFT88/70/52/46 complex through either their N- or C- termini

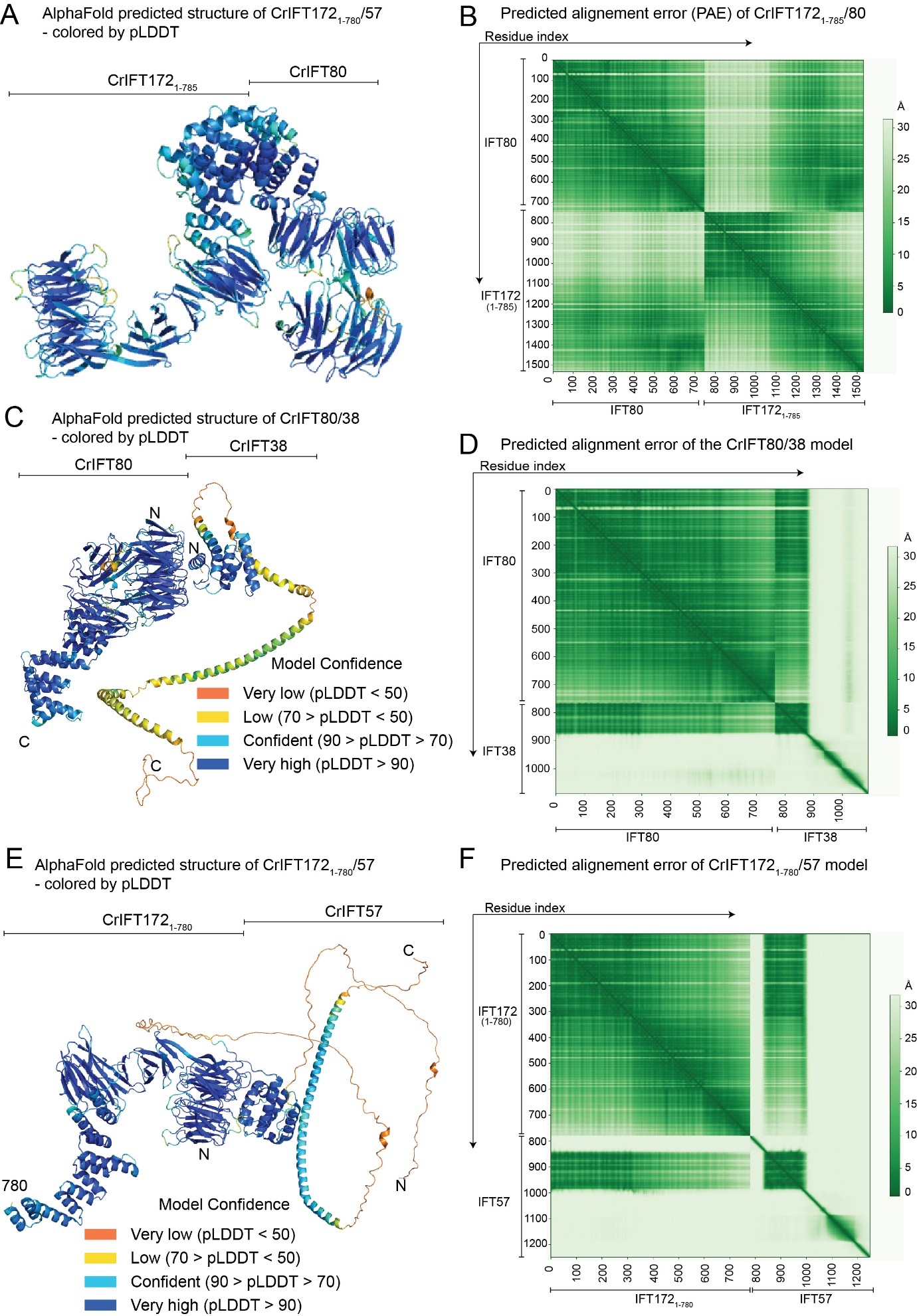
**(A)** The predicted structure of the CrIFT88120-713/70/52/46188-319/81587-645/74583-C complex colored according to the pLDDT confidence score. The lower pLDDT score for part of the IFT81/74 helices is likely a result of the short constructs of IFT81/74 and the missing IFT27/25 subunits in this prediction. **(B)** The predicted structure of the CrIFT88120-713/70/52/811-385/74129-390 complex colored according to the pLDDT scores. The dashed triangle represents the CrIFT81N/74N structure that intercalates in between TPRs 8-10 of IFT88 using the loops between CC I and CC II domains. **(C)** AlphaFold structure prediction of a CrIFT88/70/52/46/27/25/22/81/74 complex containing both binding site of panels A and B. The structural model shows simultaneous IFT81C/74C-IFT52C/46C and IFT81N/74N-IFT88 association at the expense of unfolding several CC domains of IFT81/74 as indicated by grey boxes.

**Map

Description automatically generated**

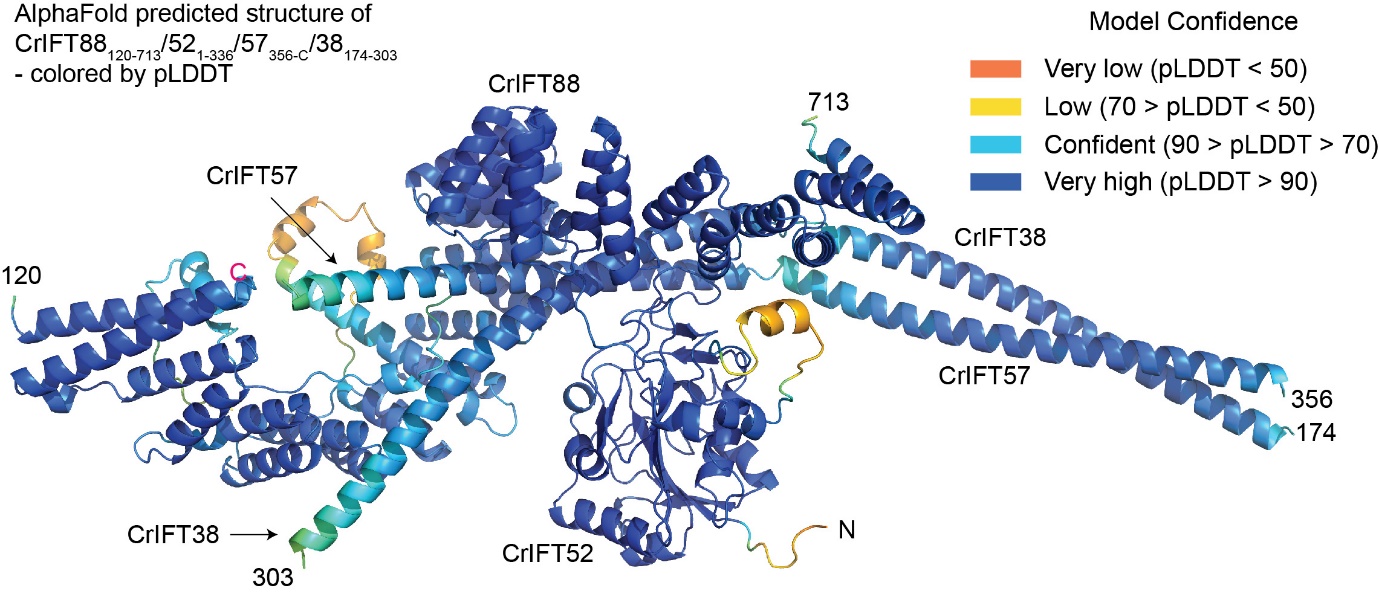
# Appendix Figure S5: AlphaFold predicted model of the CrIFT57/38/54135-C/20 tetramer colored according to the pLDDT scores

The N-terminal CH-domain of CrIFT57 as well as the coiled-coils of all four subunits are predicted with high confidence (pLDDT>70). Intrinsically disorder structural elements between CH and CC domains are colored orange and have pLDDT scores <50.



# Appendix Figure S6: AlphaFold predictions of CrIFT172/80, CrIFT172/57 and CrIFT80/38 complexes

**(A)** The predicted structure of IFT1721-785 in complex with IFT80 colored according to pLDDT scores. **(B)** The predicted alignment error plot for the IFT1721-785/80 structure. **(C)** The predicted AlphaFold model of CrIFT80 and CrIFT38 colored according to the pLDDT score. Structures of CrIFT80 and the CH-domain of CrIFT38 are predicted with high-confidence scores, whereas very low confidence scores accompany the structure of the C-terminal CC of CrIFT38 in the absence of its interacting partners. **(D)** The predicted alignment error plot for CrIFT80/30 model. The 1-300 indexed residues corresponding to the first β-propeller of CrIFT80 and the 748-875 indexed residues corresponding to the CH-domain of CrIFT38 have low PAE scores demonstrating high confidence for the multimer structure prediction. **(E)** The predicted structural model of CrIFT1721-780 in complex with CrIFT57 colored according to the pLDDT scores. Very low confidence scores are observed for the linker region between the CH-domain of CrIFT57 and the helical C-terminal domain. Otherwise, the folding of the complex is predicted with high confidence. **(F)** The predicted alignment error of the CrIFT1721-780/57. Low PAE scores are calculated between residues corresponding to CrIFT1721-780 (indexed as 1-300) and CrIFT57 (indexed as 850-900).



# Appendix Figure S7: IFT88 bridges the interaction between IFT-B1 and IFT-B2 complexes

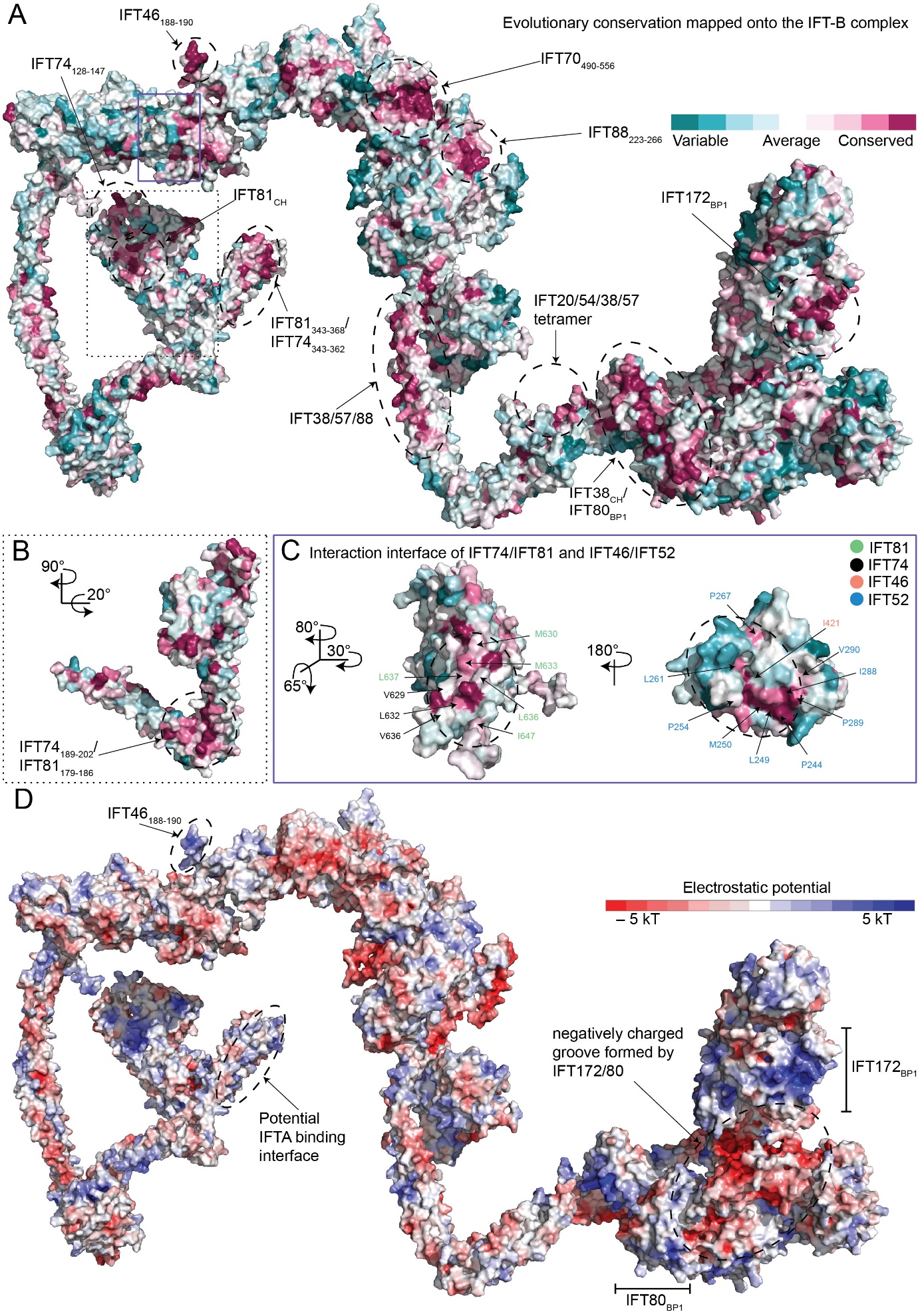
The AlphaFold predicted structure of the CrIFT88120-713/521-336/57356-C/38174-303 complex colored according to the pLDDT confidence score.

A picture containing shape

Description automatically generated

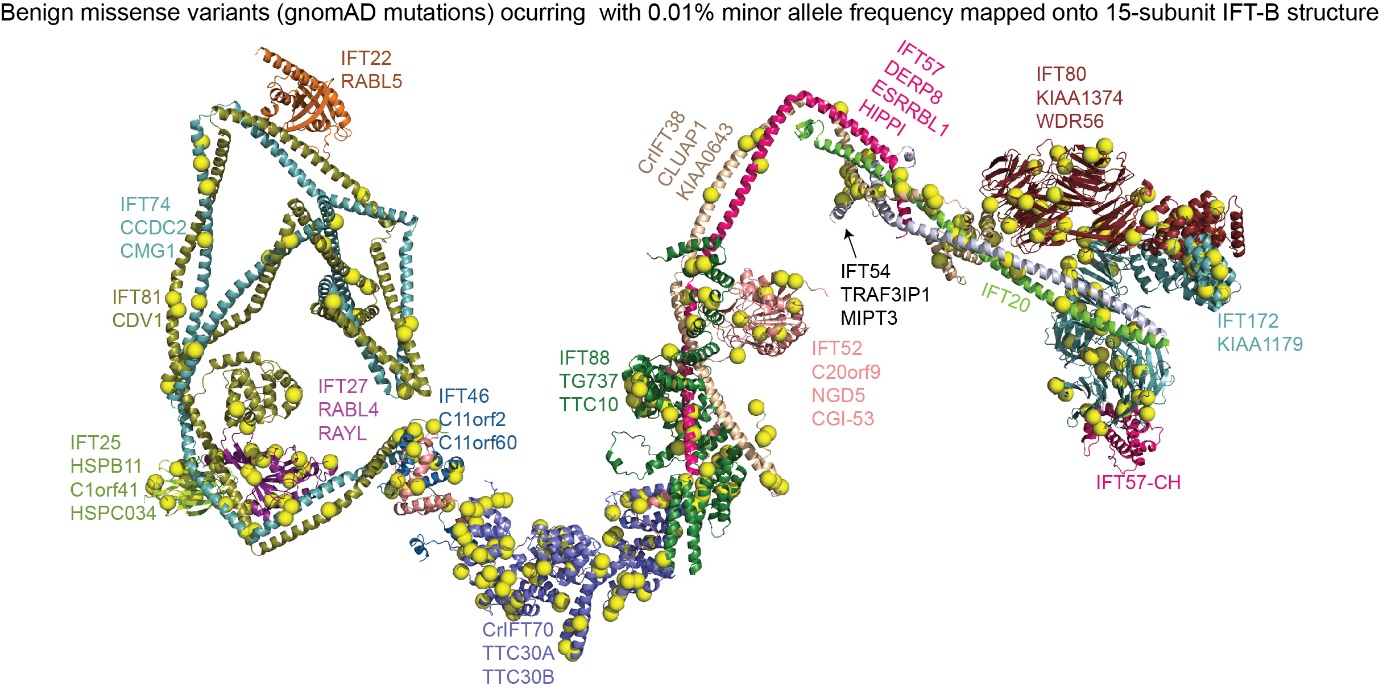
# Appendix Figure S8: Global mapping of DSBU crosslinking onto the structural model of the 15 subunits IFT-B complex

**(A)** Complete short-range (<32Å distance) intramolecular crosslinks identified by MS mapped onto the IFT-B 15mer structure. **(B)** Long-range (>32Å distance) intramolecular crosslinks identified by MS mapped onto the IFT-B 15mer structure. **(C)** Complete intermolecular crosslinks identified by MS that fall within 32Å. **(D)** Complete intermolecular crosslinks identified by MS that fall beyond 32Å.



# Appendix Figure S9: Conservation and electrostatic plots of the 15-subunit IFT-B complex structure

**(A)** Conservation analysis of the IFT-B highlights evolutionary conserved patches on the IFT-B complex. The analysis was performed with the ConSurf server (Ashkenazy et al., 2016, 2010; Celniker et al., 2013). Dashed and solid rectangles outline the regions which are discussed further in (B) and (C). **(B)** Side view of the N-terminal regions of IFT81/74 emphasizing the CH-domain of IFT81 and the conserved patch on CCs I and II where the predicted interaction with IFT88 takes place. **(C)** The conservation surface plot illustrates the interaction interface between the CC X of IFT81/74 and the C-terminal domains of IFT52/46. Conserved hydrophobic residues at the interaction interface are labeled for IFT81 (green) and IFT74 (black) in the left figure and for IFT52 (salmon) and IFT46 (blue) in the right figure. **(D)** Electrostatic plot highlighting the charged regions on IFT81/74 CC V, the platform on IFT80/172 where cargo-dynein is proposed to bind and the positively charged region on IFT46.



# Appendix Figure S10: Benign (GnomAD) variants for human IFT-B proteins mapped onto *Chlamydomonas* 15-subunit IFT-B structure

The benign single-nucleotide variants (human gnomAD database) with at least 0.01% minor allele frequency reported by genome sequencing in control patients to their aligned positions are mapped onto the *Chlamydomonas* IFT-B complex structure as yellow spheres. Unlike the case for disease variants, the benign mutations show no significant aggregation of mutations in IFT80 and IFT172.

# Appendix Table S1: X-ray data collection and refinement statistics for the CrIFT70/52/88 and CrIFT70/52/46 crystals.

|  |  |  |
| --- | --- | --- |
| **Protein complex** | **CrIFT70\_52\_88** | **CrIFT70\_52\_46** |
| **Wavelength (Å)** | 1.00 | 1.00 |
| **Resolution range (Å)** | 10 - 3.75 (4.02 - 3.75) | 90 - 4.0 (4.24 - 4.0) |
| **Space group** | C 1 2 1 | P 1 21 1 |
| **Unit cell (a,b,c,Å)** | 118,0, 269.5, 95.8, 90.11 | 98.4, 113.4, 374.5, 90.06 |
| **Total reflections** | 85471 (16160) | 476912 (47298) |
| **Unique reflections** | 24628 (4890) | 70497 (10960) |
| **Multiplicity** | 3.4 (3.3) | 6.8 (6.4) |
| **Completeness (%)** | 95.2 (94.7) | 99.4 (91.4) |
| **Mean I/sigma(I)** | 7.8 (1.5) | 7.7 (0.6) |
| **Wilson B-factor (Å2)** | 145 | 184 |
| **Twin law** | NA | h,-k,-l |
| **R-pim** | 0.054 (0.63) | 0.086 (1.6) |
| **CC1/2** | 0.998 (0.626) | 0.972 (0.325) |
| **Reflections used in refinement** | 24628 (3616) | 70497 (10960) |
| **Reflections used for R-free** | 1442 (210) | 2913 (188) |
| **R-work** | 0.358 (0.433) | 0.313 (0.346) |
| **R-free** | 0.379 (0.462) | 0.346 (0.385) |
| **Number of non-hydrogen atoms** | 12261 | 22980 |
| **macromolecules** | 12261 | 22980 |
| **Ligands** | 0 | 0 |
| **Solvent** | 0 | 0 |
| **Protein residues** | 1548 | 2976 |
| **RMS(bonds)** | 0.004 | 0.006 |
| **RMS(angles)** | 0.77 | 1.32 |
| **Ramachandran favored (%)** | 95.60 | 91.49 |
| **Ramachandran allowed (%)** | 4.20 | 7.92 |
| **Ramachandran outliers (%)** | 0.20 | 0.59 |
| **Rotamer outliers (%)** | 0.84 | 0.04 |
| **Clash score** | 14.9 | 29.9 |
| **Average B-factor (Å2)** | 185 | 198 |

# Appendix Table S2: SAXS Data collection parameters for IFT381-133 and IFT801-582/381-133 complex

|  |  |  |  |
| --- | --- | --- | --- |
| **Data-collection parameters** | |  | |
| Instrument: | | ESRF BM29 | |
| Wavelength (Å) | | 0.99 | |
| q-range (Å-1) | | 0.0032 – 0.49 | |
| Sample-to-detector distance | | 2.867 m | |
| Exposure time (sec) | | 1 per frame | |
| Temperature (K) | | 283 | |
| Detector | | Pilatus 1M (Dectris) | |
| Flux (photons/s) | | 1 × 1012 | |
| Beam size (µm2) | | 172 × 172 | |
| **Structural parametes** | | IFT381-133 | IFT801-582/381-133 |
| Type of experiment | | CS | CS |
| Concentration used (mg/mL) | | 0.4-6.4 (0.4) | 0.37-6.0 (0.75) |
|  | NaCl concentration (mM) | 150 | 150 |
| From p(r) | Rg (Å) | 21.5 | 29.2 |
| Dmax (Å) | 70.4 | 102.2 |
| Porod volume Vp x103 (Å3) | | 21.97 | 106.37 |
| Molecular mass (kDa) from Vp | | 13.7 | 66.4 |
| Molecular mass (kDa) from sequence | | 15 | 80 |
| **Modeling** | |  |  |
| Dammif | NSD | 0.557 | 0.628 |

# Supplementary References

Ashkenazy H, Abadi S, Martz E, Chay O, Mayrose I, Pupko T, Ben-Tal N. 2016. ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic Acids Research* **44**:W344–W350. doi:10.1093/nar/gkw408

Ashkenazy H, Erez E, Martz E, Pupko T, Ben-Tal N. 2010. ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. *Nucleic Acids Research* **38**:W529–W533. doi:10.1093/nar/gkq399

Celniker G, Nimrod G, Ashkenazy H, Glaser F, Martz E, Mayrose I, Pupko T, Ben-Tal N. 2013. ConSurf: Using Evolutionary Data to Raise Testable Hypotheses about Protein Function. *Israel Journal of Chemistry* **53**:199–206. doi:10.1002/ijch.201200096