Cryo-EM structure of the human CST–Polα/primase complex in a recruitment state

Sarah W. Cai, John C. Zinder, Vladimir Svetlov, Martin W. Bush, Evgeny Nudler, Thomas Walz and Titia de Lange

The CST–Pol/primase complex is essential for telomere maintenance and functions to counteract resection at double-strand breaks. We report a 4.6-Å resolution cryo-EM structure of human CST–Pol/primase, captured prior to catalysis in a recruitment state stabilized by chemical cross-linking. Our structure reveals an evolutionarily conserved interaction between the C-terminal domain of the catalytic POLA1 subunit and an N-terminal expansion in metazoan CTC1. Cross-linking mass spectrometry and negative-stain EM analysis provide insight into CST binding by the flexible POLA1 N-terminus. Finally, Coats plus syndrome disease mutations previously characterized to disrupt formation of the CST–Pol/primase complex map to protein–protein interfaces observed in the recruitment state. Together, our results shed light on the architecture and stoichiometry of the metazoan fill-in machinery.

Human telomeric DNA terminates in a 3′ overhang of the G-rich strand, which is required for t-loop formation and telomere protection. The mature 3′ overhang must be generated during each cell cycle in a controlled manner. Following replication, nucleolytic resection of the 5′ strand can result in excessively long overhangs. The loss of sequences from the 5′ strand is counteracted through fill-in DNA synthesis by the CTC1–STN1–TEN1 (CST) complex and DNA polymerase α–primase (Polα/primase) (Fig. 1a), which are recruited to telomeres by the shelterin complex. Dysfunctional fill-in, primarily driven by mutations in CST, causes telomere protection defects in BRCA1-deficient cells. Although a structure of decameric CST bound to a short oligonucleotide has been determined, the molecular details of how CST interacts with Polα/primase are largely unknown. Here, we present a combination of structural, biochemical, and biophysical data describing the molecular basis of their interaction. Our structure of the complex in a recruitment state reveals a novel interface between Polα/primase and CST that evolved in metazoans and informs on CP mutations found in the N-terminal region (POLA1 N, 1–335 aa) that is dispensable for catalysis and omitted in most structures of the enzyme (Fig. 1a). We purified Polα/primase lacking POLA1 N (referred to hereafter as PPΔN), reconstituted a CST–PPΔN (ssDNA) complex (Fig. 1a) as for full-length Polα/primase (referred to hereafter as PPΔ), and collected cryo-EM data (Extended Data Fig. 3b). The omission of POLA1 N resulted in lower CST occupancy, so we introduced additional classification steps to select for particles containing intact CST (Extended Data Fig. 3c). We used 131,850 particles to generate the final map with a global resolution of 4.6 Å (Fig. 1b and Extended Data Fig. 3d). Local resolution estimates revealed lower resolution for the peripheral regions of CST and PPΔN (Extended Data Fig. 3e), likely due to flexibility, as suggested by the blurred-out regions in the 2D class averages (Extended Data Fig. 3d).

The crystal structure of apo PPΔN (PDB ID: 5EXR) and the cryo-EM structure of an ssDNA-bound CST monomer extracted from the decamer structure (PDB ID: 6WGW) could readily be docked into our density map. We then substituted the CTC1 structure with a model from the AlphaFold 2 database, which provides information about the CTC1 N-terminus that was poorly resolved in the published CST cryo-EM map. The ssDNA was included in the complex, as evidenced by the native gel of the GraFix fractions (Extended Data Fig. 3a), but no reliable density could be found for it owing to the low resolution of CST in the cryo-EM map. However, we observe low occupancy of the STN1 C-terminal half (STN1 C-

1Laboratory of Cell Biology and Genetics, The Rockefeller University, New York, NY, USA. 2Laboratory of Molecular Electron Microscopy, The Rockefeller University, New York, NY, USA. 3Department of Biochemistry and Molecular Pharmacology, New York University Grossman School of Medicine, New York, NY, USA. 4Howard Hughes Medical Institute, Chevy Chase, MD, USA. 5E-mail: twalz@rockefeller.edu; delange@rockefeller.edu
184–398 aa) in that region (Fig. 1c) and, because ssDNA competes with STN1\(^\circ\) for that binding site\(^1\), we retained the 4 nt of ssDNA from the previously determined CST structure\(^1\) in the model we docked into our map (Fig. 1c).

After initial rigid-body docking followed by flexible fitting and refinement, the overall conformations of the two subcomplexes did not show major changes from their structures in isolation (Fig. 1c and Extended Data Fig. 3f). In the complex, Pol\(\alpha\)/primase remains in the occluded state\(^1\), in which the POLA2 subunit is blocking entry of DNA into the active site of POLA1 (Fig. 1c). This finding is consistent with reported results showing that cross-linking preferentially stabilizes the more compact, occluded state compared with the flexible, extended state of the enzyme\(^2\). Thus, we conclude that our structure likely captures a recruitment state of the complex that forms prior to active RNA and DNA synthesis by Pol\(\alpha\)/primase.

**Fig. 1** | Architecture of the human CST–Pol\(\alpha\)/primase complex. **a**, Domain schematics for proteins used in this study. Gray bars indicate regions modeled in **c**. NTD, N-terminal domain; CTD, C-terminal domain; OB, oligonucleotide/oligosaccharide binding fold; PDE, phosphodiesterase domain; wH, winged helix-turn-helix motif. **b**, Orthogonal views of the cryo-EM map of CST–PP\(^\Delta\)N, segmented and colored by subunit as in **a**. **c**, Model of CST–PP\(^\Delta\)N shown in cartoon representation.
Structural and evolutionary analysis of the CTC1–POLA1 interface. The primary interaction interface observed in our structure occurs between the C-terminal domain of POLA1 (POLA1^{CTD}, 1,265–1,462 aa) and the N-terminal OB folds of CTC1 (Figs. 1c and 2a). The resolution is limiting for rigorous analysis of amino acid interactions, and analysis of surface electrostatic potential suggests that this interface is not driven by a dominant hydrophobic or charged interaction, but rather by shape complementarity of the two proteins, burying 1,250 Å² of solvent-accessible surface area (Fig. 2a and Extended Data Fig. 4a,b).

We identified a CTC1-recognition loop (CRL, 1,400–1,424 aa) in the POLA1^{CTD} that is shifted relative to its position in the apo structure^{17} to contact CTC1 (Fig. 2a). Sequence conservation analysis of the interface revealed that both CTC1 and the interacting region on the POLA1^{CTD} have low conservation at the primary sequence level (Extended Data Figs. 4c and 5a). However, the CRL is identifiable by an insertion of uniform length in metazoans, and we find that, when modeled using predictions from AlphaFold 2 (refs. 19,20), the CRL is structurally conserved in metazoans. Furthermore, we find that the presence of a CRL feature correlates with an expansion of metazoan CTC1 to contain the N-terminal OB folds that interact with the CRL. In unicellular eukaryotes, this loop diverges greatly between species and can be either shorter (for example, in Tetrahymena thermophila and Schizosaccharomyces pombe) or longer (for example, in Saccharomyces cerevisiae) and adopts a different predicted structure compared with that in the metazoan CRL (Fig. 2b and Extended Data Fig. 5b).

To assess the involvement of the CRL in CST binding, we generated human POLA1^{CTD} constructs with the wild-type CRL (CRL^{WT}, 1,265–1,462 aa), with the CRL swapped for a GGSGGS-linker (CRL^{GGSGGS}, 1,265–1,402–GGSGGS–1,423–1,462 aa), or with the CRL swapped for the S. pombe short loop (CRL^{S. pombe}, 1,265–1,399–QTGTGAT–1,425–1,462 aa) (Extended Data Fig. 6a–d). Although the POLA1^{CTD} constructs ran as heterogeneously sized smears in SDS–PAGE, they compressed to single bands in native PAGE (Extended Data Fig. 6a,b). Sharp symmetric peaks in SEC elution profiles (Extended Data Fig. 6c) and spectrophotometric (Nanodrop) quantification indicated that the protein was pure and free from nucleic acid contamination, respectively. We measured the affinity of the interaction between the POLA1^{CTD} constructs and fluorescently labeled CST using microscale thermophoresis (MST) (Fig. 2e). Two distinct binding events between the CRL^{WT} protein and CST were observed: the first binding event (1; Fig. 2d) is higher affinity and characterized by a positive change in the normalized fluorescence (Fig. 2d), and the second binding event (2) is lower affinity and characterized by a negative change in the normalized fluorescence (Fig. 2e). We separated the data on the basis of the two inflection points^{23} and calculated dissociation constant (K_D) values of ~3.3 μM and ~122 μM for binding events 1 and 2, respectively. In contrast, the CRL^{GGSGGS} and CRL^{S. pombe} mutant proteins displayed binding in only the second event (Fig. 2e,d). For binding event 2, the determined K_D values for the CRL^{GGSGGS} and CRL^{S. pombe} proteins were ~96 μM and ~88 μM, respectively.

The CRL and one Zn^{2+}-binding domain in POLA1^{CTD} form the module that fits into a complementary cleft in CTC1. We speculate that the high-affinity binding mode (event 1; Fig. 2d) observed in the MST experiments involves both the CRL and Zn^{2+}-binding domain, whereas the lower affinity binding mode (event 2; Fig. 2e) reflects the interaction of the Zn^{2+}-binding domain with CTC1, which is CRL-independent. In this second binding mode, CST could rock about the Zn^{2+}-binding domain, echoing the hinge-like flexible motion of CST about POLA1 suggested by the cryo-EM data (Extended Data Fig. 3d,e). CTC1^{OB-D}, an elongated OB fold that shares no homology with known OB folds^{18}, forms the major interaction with POLA1^{CTD} (Fig. 2a and Extended Data Fig. 4). This finding is particularly interesting given the proposal that CST and Pol6 co-evolved in eukaryotes^{24}. Our structural model would then capture a metazoan-specific development in this trajectory.

Our structural model also places regions of CTC1 near POLA2 and PRIM2 (Fig. 1c). These two potential interaction sites are not

| Table 1 | Cryo-EM data collection, refinement and validation statistics |
|-----------------------------------|-------------------|-------------------|-------------------|
| **CST–PP^{PL}** (dataset 1)       | **CST–PP^{PL}** (dataset 2) | **CST–PP^{LN}** (dataset 3) |
| Article   | EMD-26347         | EMD-26347         | EMD-2646 |
| Magnification (x)                | x28,000           | x53,000           | x64,000 |
| Voltage (kV)                     | 200               | 300               | 300    |
| Electron exposure (e−/Å²)        | 53                | 57                | 52     |
| Defocus range (μm)               | -1.5 to -3        | -1 to -2.2        | -1 to -2.2 |
| Pixel size (Å)                   | 1.5               | 1.32              | 1.08   |
| Symmetry imposed                 | C₁                | C₁                | C₁     |
| Initial particle images (no.)     | 307,358           | 441,335           | 2,515,853 |
| Final particle images (no.)       | 109,224 (with no. 2) | 109,224 (with no. 1) | 131,850 |
| Map resolution (Å)               | 16                | 16                | 4.6    |
| Map sharpening B factor (Å⁻²)     | -311              | -311              | -193   |
| Model composition                | N/A               | N/A               | N/A    |
| Non-hydrogen atoms               | 30,077            | 3740              | 5      |
| Protein residues                 | 103.2             | 110.7             |        |
| Ligands                          | 0.006             | 0.849             |        |
| Validation                       | N/A               | N/A               | N/A    |
| MolProbity score                 | 2.60              | 27.01             | 0.12   |
| Clashscore                       | 83.69             | 16.31             | 0.00   |

Refinement

| Initial model used (PDB code)     | PDB 5EXR          | PDB 5EXR          | PDB 6W6W   |
|-----------------------------------|-------------------|-------------------|------------|
| Model resolution (Å)               | N/A               | N/A               | 4.7        |
| Final particle images (no.)        | 109,224 (with no. 2) | 109,224 (with no. 1) | 131,850 |
| Map resolution range (Å)           | Not calculated    | Not calculated    | 4.0–8.0   |
| Model composition                | N/A               | N/A               | N/A       |
| Non-hydrogen atoms               | 30,077            | 3740              | 5        |
| Protein residues                 | 103.2             | 110.7             |          |
| Ligand                           | 0.006             | 0.849             |          |
| Validation                       | N/A               | N/A               | N/A      |
| MolProbity score                 | 2.60              | 27.01             | 0.12     |
| Clashscore                       | 83.69             | 16.31             | 0.00     |
Fig. 2 | Structural analysis of the CTC1–PoLa1CTD interface. 

**a,** Zoomed-in view of the interaction of PoLa1CTD (cartoon) with CTC1 (cartoon and surface) superposed with PoLa1CTD from the apo Polα/primase structure (PDB ID: 5EXR). The CTC1-recognition loop (CRL, 1400–1424 aa) is highlighted in cyan in the cartoon representation. The corresponding cleft in CTC1 is indicated with a cyan arc. **b,** Comparison of PoLa1CTD models generated with AlphaFold 2 (refs. 19,20) from several species alongside domain comparisons of corresponding CTC1 orthologs. The metazoan CTC1 N-terminal expansion is indicated, and the binding pocket is represented by a cyan arc as in **a** or a dashed fuchsia arc in the non-metazoan species. **c,** Microscale thermophoresis measuring binding of PoLa1CTD proteins to RED-tris-NTA-labeled His6-MBP-CTC1–STN1–TEN1. Error bars represent s.e.m. for each data point, calculated from three independent thermophoresis measurements (Extended Data Fig. 6e). **d,** Zoom-in of binding events 1 and 2 from **c,** respectively. $K_D$ values were calculated (N.D., not determined) with the MO Affinity Analysis (Nanotemper) software (split into two events for the CRLWT and generated with all data for the CRLGGSGGS and CRL$^{\Delta20}$ mutants). $F_{norm}$ normalized fluorescence.
well resolved in our cryo-EM map, but it is possible that these interactions are weak and/or more transient as CST flexes about the hinge generated by the POLA1CTD-CTC1 interaction.

**CST–Pol/primase maintains a 1:1 stoichiometry.** The CST–Pol/primase complex is sterically incompatible with the previously reported ssDNA-bound CST decamer\(^1\), as it would bind in the center of the ring and sterically clash with neighboring CST subunits (Fig. 3a). CST has also been shown to dimerize upon ssDNA binding\(^2\), but two additional lines of evidence suggest that active CST–Pol/primase has a 1:1 stoichiometry. First, we do not observe CST dimers in our 2D class averages (Extended Data Figs. 1 and 2). Second, we characterized the CST–POLA1N interaction to understand why CST–PPFL did not yield a high-resolution map although PPFL forms a more stable interaction with CST. We reconstituted a native complex of CST and MBP-tagged POLA1N and analyzed it by negative-stain EM and cross-linking mass spectrometry (CX-MS) (Extended Data Fig. 7). With MBP as a mass label, we localized the N-terminus of POLA1N to the primary CST dimerization interface (Fig. 3b). CX-MS analysis suggests that POLA1N binds in multiple modes to CST, which could partially explain the heterogeneity observed with CST–PPFL (Fig. 3c,d). It is possible that POLA1N binding is restrained in the presence of the full complex, but we observe similar cross-links between CTC1 and POLA1N in CX-MS analysis of CST–PPFL (Extended Data Fig. 7c). Thus, we conclude that POLA1N binds heterogeneously to CST in the region of the dimer interface, and the CST–Pol/primase fill-in machinery functions as a 1:1 complex.

**Coats plus mutations map to recruitment interfaces.** Three CP point mutations (p.A227V, p.V259M, and p.V665G) in CTC1 were previously described to disrupt Pol/primase association with CST\(^1\) and to the start of RNA/DNA synthesis by the enzyme. Notably, this occurs during recruitment of Pol/primase to the telomere, prior to the start of RNA/DNA synthesis by the enzyme. Notably, this interface is formed by the N-terminal four OB folds of CTC1, a beta-sheet hinge generated by the POLA1CTD-CTC1 interaction. We propose that this interaction is specific to metazoans and provides a mechanism to recruit Pol/primase to the telomere. Substitution of the CRL with a short loop, either a GGSGGS linker or the orthologous S. pombe sequence, abrogates binding in the high-affinity (\(K_d = 1-10 \mu M\))

**Discussion.** In this study, we report a novel structure of the human CST–Pol/primase fill-in machinery. By chemically cross-linking the complex, we captured a conformational state that reveals how CST can recognize and bind the occluded state of Pol/primase through a newly uncovered interaction between CTC1 and POLA1CTD. We propose that this interaction is specific to metazoans and occurs during recruitment of Pol/primase to the telomere, prior to the start of RNA/DNA synthesis by the enzyme. Notably, this interface is formed by the N-terminal four OB folds of CTC1, a metazoan expansion of the subunit that further differentiates it from the paralogous RPA large subunit and from CTC1 homologs found in unicellular eukaryotes\(^13,24\). Our evolutionary conservation analysis identified a complementary species-specific loop in the POLA1CTD CRL, termed the CRL, that appears to have co-evolved with the expansion in CTC1. Substitution of the CRL with a short loop, either a GGSGGS linker or the orthologous S. pombe sequence, abrogates binding in the high-affinity (\(K_d = 1-10 \mu M\))
binding mode between POLA1CTD and CST. MST experiments also captured an order-of-magnitude-lower affinity ($K_D \approx 100 \mu\text{M}$) binding mode that is CRL-independent. The two binding modes are consistent with our cryo-EM structure, in which POLA1 CTD utilizes both the CRL and a Zn$^{2+}$-binding motif to interact with CTC1.

**Fig. 4 | Disease mutation mapping and fill-in model.** **a**, Summary of CP point mutations in CTC1. **b**, Mapping of mutations affecting Polα/primase binding on the CST–PPαΔN structure. Colors as in a. Box, zoomed-in view of the primary CTC1OB–POLA1CTD interaction. **c**, Comparison of the CST–PPαΔN and CST–PPαΔ (Extended Data Fig. 2a) cryo-EM density maps, segmented and colored according to docked CST and PPαΔN models. **d**, Proposed model for the handoff of Polα/primase to shelterin-bound CST following telomere replication.
Further work is needed to delineate the functional role of POLA1N in regulating the fill-in machinery. We uncovered an interaction between POLA1N and CST by using a combination of negative-stain EM, CX-MS, and cryo-EM with CST bound to POLA1N in isolation and additionally bound in the context of PPFL. The CST–PPFL complex associated more robustly than CST–PP, as inferred from higher CST occupancy in cryo-EM 2D averages and stability during native SEC. However, we observed greater structural heterogeneity in the cryo-EM data for CST–PPFL. The addition of POLA1N may allow the complex to sample a larger conformational space, interfering with accurate alignment of the particles in the CST–PPFL complex. POLA1N is responsible for the flexible tethering of Polα/primase to the replisome via an interaction with AND-1 (refs. 27,28), and its extensive yet flexible interaction with CST suggests a potential spatiotemporal regulation of fill-in after telomere replication, where the replisome hands Polα/primase off to a shelterin-bound CST for C-strand synthesis (Fig. 4d).

Finally, our model informs on CP mutations (p.A227V, p.V259M, p.V665G) previously characterized to disrupt CST–Polα/primase association23,25. It is unlikely that these single point mutations result in complete loss of function, as such mutations would likely be lethal. Consistent with this framework of mild dysfunction, we observe that the three CP mutations map close to interaction interfaces but are not necessarily responsible for direct interaction with Polα/primase. For example, V665 resides on a β-strand in Ctc107,111 that is not directly on the interface, but the glycine substitution could destabilize the OB fold and weaken the interaction. Similarly, the Ctc107,111 mutations reside near the bridge observed in the CST–PPFL cryo-EM map. Since those mutants (p.A227V, p.V259M, p.V665G) were characterized14,23, three more CP mutations have been reported in CTC107,111 (refs. 23,25). Presumably, these would be deleterious in a manner resembling that of p.A227V and p.V259M. Future genetic and functional studies are required to elucidate the precise mechanism by which these mutations cause CP, but our structural results provide a framework for understanding the molecular basis of human disease linked to CST–Polα/primase, its role in telomere maintenance, and its contribution to DSB repair.

Online content
Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41594-022-00766-y.

Received: 17 December 2021; Accepted: 30 March 2022; Published online: 16 May 2022

References
1. de Lange, T. Shelterin-mediated telomere protection. Annu Rev Genet 52, 223–247 (2018).
2. Chen, L. Y., Redon, S. & Lingner, J. The human CST complex is a terminator of telomerase activity. Nature 488, 540–544 (2012).
3. Wan, M., Qin, J., Songyang, Z. & Liu, D. OB fold-containing protein 1 (OBFC1), a human homolog of yeast Snt1, associates with TP1 and is implicated in telomere length regulation. J. Biol. Chem. 284, 26725–26731 (2009).
4. Wu, P., Takai, H. & de Lange, T. Telomeric 3' overhangs derive from resection by ExoI and Apollo and fill-in by POT1β-associated CST. Cell 150, 39–52 (2012).
5. Takai, H. et al. A POT1 mutation impairs defective telomere end fill-in and telomere truncations in Coats plus. Genes Dev. 30, 812–826 (2016).
6. Stewart, J. P. et al. Human CST promotes telomere duplex replication and general replication restart after fork stalling. EMBO J. 31, 3537–3549 (2012).
7. Peng, X., Hsu, S. J., Kasbek, C., Chadiken, M. & Price, C. M. CTC1-mediated C-strand fill-in is an essential step in telomere length maintenance. Nucleic Acids Res. 45, 4281–4293 (2017).
8. Anderson, B. H. et al. Mutations in CTC1, encoding conserved telomere maintenance component 1, cause Coats plus. Nat. Genet. 44, 338–342 (2012).
9. Savage, S. A. Beginning at the ends: telomeres and human disease. F1000Res 3, https://doi.org/10.12688/f1000research.14068.1 (2018).
10. Mirman, Z. et al. 53BP1-RIFI-shieldin counteracts DSBR resection through CST- and Polα-dependent fill-in. Nature 560, 112–116 (2018).
11. Mirman, Z. & de Lange, T. 53BP1: a DSB escort. Genes Dev. 34, 7–23 (2020).
12. Mirman, Z., Sasi, N. K., King, A., Chapman, J. R. & de Lange, T. 53BP1-shieldin-dependent DSB processing in BRCA1-deficient cells requires CST–Polα-primase fill-in synthesis. Nat. Cell Biol. 24, 51–61 (2022).
13. Lim, C. J. et al. The structure of human CST reveals a decameric assembly bound to telomeric DNA. Science 368, 1081–1085 (2020).
14. Chen, L. Y., Majerski, J. & Lingner, J. Molecular basis of telomere syndrome caused by CTC1 mutations. Genes Dev. 27, 2099–2108 (2013).
15. Ganduri, S. & Lue, N. F. STN1-POLA2 interaction provides a basis for primase-pol α stimulation by human STN1. Nucleic Acids Res. 45, 9455–9466 (2017).
16. Kastner, B. et al. GraFix: sample preparation for single-particle electron cryomicroscopy. Nat. Methods 5, 53–55 (2008).
17. Mechnovski, V. G. et al. Mechanisms of concerted RNA–DNA primer synthesis by the human primosome. J. Biol. Chem. 291, 10006–10020 (2016).
18. Coloma, J., Johnson, R. E., Prakash, L., Prakash, S. & Aggarwal, A. K. Human DNA polymerase α in binary complex with a DNA:DNA template-primer. Sci. Rep. 6, 23784 (2016).
19. Jumper, J. et al. Applying and improving AlphaFold at CASP14. Proteins 89, 1712–1721 (2021).
20. Varadi, M. et al. AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. Nucleic Acids Res. gkb1061 (2021).
21. Kilkenny, M. L. et al. Structural basis for the interaction of SARS-CoV-2 virulence factor nsp1 with DNA polymerase α-primase. Protein Sci. 31, 333–344 (2022).
22. Núñez-Ramírez, R. et al. Flexible tethering of primase and DNA Pol α in the eukaryotic primosome. Nucleic Acids Res. 39, 8187–8199 (2011).
23. Seidel, S. A. et al. Microscale thermophoresis quantifies biomolecular interactions under previously challenging conditions. Methods 59, 301–315 (2013).
24. Lue, N. F. Evolving linear chromosomes and telomerases: a C-strand-centric view. Trends Biochem. Sci. 43, 314–326 (2018).
25. Gu, P. et al. CTC1-STN1 coordinates G- and C-strand synthesis to regulate telomere length. Aging Cell 17, e12783 (2018).
26. Huang, H. et al. Structure of a DNA polymerase alpha-primase domain that docks on the SV40 helicase and activates the viral primosome. J. Biol. Chem. 285, 17112–17122 (2010).
27. Villa, F. et al. Ctf4 is a hub in the eukaryotic replisome that links multiple CIP-box proteins to the CMG helicase. Mol. Cell. 63, 385–396 (2016).
28. Kilkenny, M. L. et al. The human CTF4-orthologue AND-1 interacts with DNA polymerase α/primase via its unique C-terminal HMG box. Open Biol. 7, 170217 (2017).
29. Walne, A. J. et al. Mutations in the telomere capping complex in bone marrow failure and related syndromes. Haematologica 98, 334–358 (2013).
30. Hidalgo-Sanz, J. et al. Infant with intracranial calcifications and retinopathy. Rev. Neurol. 69, 289–292 (2019).

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.
Methods

DNA construct and baculovirus generation. DNA fragments encoding human CTC1, STN1, TEN1, POLA1 (POLA1N, 1–335–1492 aa), POLA2, PRIM1, and PRIM2 were cloned using Gibson assembly into the pBac vectors with affinity tags. The plasmids used in this study are pBGlAa (His6–MBP–PreSc–CtC1/STN1/TEN1), pBGl2ab (His6–MBP–PreSc–POLA1/1/2A) and pBGl3ab (His6–MBP–PreSc–POLA1/1/2A) and pBGl3ab (His6–MBP–PreSc–POLA1/1/2A). Recombinant bacmids were generated from these plasmids using the InvoHiBac competent cells (Gibco; cat. no. 10361012) and transfected into S9 insect cells (Gibco; cat. no. 11496015) with CellecHt II Reagent (Gibco) to generate a P1 baculovirus stock. P1 baculovirus was amplified in deficient S9 insect cells to generate P2 and P3 stocks, and the P3 virus was used to infect Tni suspension insect cell culture (Expression Systems; cat. no. 94-002S) for protein expression. 

Constructs for expression ofGST-1CTD inE. coli (CRLT: 1265–1462 aa; CRLGGSGGS: 1265–1402–GGSGGS–1423–1462 aa; CRLGGSGGS: 1265–1423–1462 aa; CRLGGSGGS: 1265–1423–1462 aa; CRLGGSGGS: 1265–1423–1462 aa) were cloned with a twofold molar excess of His6–MBP–PreSc–POLA1 for GST-1CTD and incubated for 1 h at 4 °C prior to being loaded onto a HiLoad Superdex 200 16/600 GL column equilibrated with a buffer containing 20 mM HEPES (pH 7.5), 150 mM NaCl, 0.5 mM TCEP, and 5% glycerol. The fractionated fractions (Extended Data Fig. 5) were pooled and concentrated. Purified GST-1CTD was mixed with a twofold molar excess of His6–MBP–PreSc–POLA1 and incubated for 1 h at 4 °C prior to being loaded onto a HiLoad Superdex 200 16/600 GL column equilibrated with a buffer containing 20 mM HEPES (pH 7.5), 150 mM NaCl, 0.5 mM TCEP, and 5% glycerol. The fractionated fractions (Extended Data Fig. 5) were pooled and concentrated.

Negative-stain EM sample preparation, data collection, and image processing. Protein samples for negative-stain EM (3.5–μL drops) in a concentration range of 0.01–0.05 mg/mL were adsorbed to glow-discharged carbon-coated copper grids with a collodion film, washed with three drops of deionized water, and stained with two drops of freshly prepared 1% (w/v) uranyl acetate. Specimens were imaged at room temperature using a Phillips CM10 electron microscope equipped with a tungsten filament and operated at an acceleration voltage of 80 kV. The magnification used corresponds to a calibrated pixel size of 2.8 Å. Particle coordinates were auto-picked using the Swarm picker in EMAN2 (ref. 3). Particle extraction and 2D classification were performed in RELION 3.0 or RELION 3.1, as indicated.

Grafix stabilization of CST–Polε/primase complexes. Purified CST and Polε/primase (PP or PPε) were first mixed in equimolar amounts (for a 1.5 μL final protein concentration of each component) and incubated on ice for 1 h before the addition of 0.5 μL DNase I (GTTTAG), 3′ ssDNA at a 2.25-fold molar excess. The protein–ssDNA mixture was incubated on ice for 2 h and then loaded on top of a 11-ml 20 mM HEPES, pH 7.5, 150 mM NaCl, 0.5 mM tris(2-carboxyethyl)phosphine (TCEP), 250 mM imidazole, and 5% glycerol. 

After elution from the Ni-NTA resin, His6–MBP–PreSc–CtC1/STN1/TEN1 was incubated with rhinovirus 3C protease overnight at 4 °C to remove the His6–MBP tag. The cleaved protein was loaded onto a HiTrap Heparin HP column (GE Healthcare) equilibrated with a buffer containing 20 mM HEPES (pH 7.5), 150 mM NaCl, 0.5 mM TCEP, and 5% glycerol and subsequently washed with 20-50 column volumes (CV) of the same buffer. Bound protein was eluted in a buffer containing 20 mM HEPES (pH 7.5), 150 mM NaCl, 0.5 mM tris(2-carboxyethyl)phosphine (TCEP), 250 mM imidazole, and 5% glycerol.

Cryo-EM sample preparation and data collection. Four microliters of the Grafix-stabilized samples were applied to Quantifoil R 1.2/1.3 mesh Au400 holey carbon grids covered with a graphene oxide support layer (EMS), blotted for 1–1.5 sec, and plunge frozen in liquid ethane using a Vitrobot Mark IV (Thermo Fisher Scientific) operated at 4 °C and 100% humidity. Cryo-EM imaging was performed in the Cryo-EM Resource Center at the Rockefeller University using SerialEM4. Data-collection parameters are summarized in Table 1.

For CST–PP, one dataset (dataset 1) was collected on a 200-kV Talos Arctica electron microscope (Thermo Fisher Scientific) at a nominal magnification of ×28000 (TEM nanoprobe), corresponding to a calibrated pixel size of 1.5 Å on the specimen level. Images were collected using a defocus range from −1.5 to −3 μm with a K2 Summit direct electron detector (Gatan) in super-resolution counting mode. Exposures of 10 sec were dose-fractionated into 50 frames (200 ms per frame) with a dose rate of 1.4 electrons/pixel/sec (approximately 0.7 electrons per Å2 per frame), resulting in a total dose of 53 electrons per Å2. The second dataset (dataset 2) was collected on a 300-kV Titan Krios electron microscope at a nominal magnification of ×53000 (EFTEM nanoprobe), corresponding to a calibrated pixel size of 1.32 Å on the specimen level. Images were collected using a defocus range from −1 to −2.2 μm with a K3 direct electron detector (Gatan) in super-resolution counting mode. Exposures of 10 sec were dose-fractionated into 50 frames (60 ms per frame) with a dose rate of 33 electrons/pixel/sec (approximately 1.14 electrons per Å2 per frame), resulting in a total dose of 57 electrons per Å2.
Cryo-EM data processing. For all datasets, movie stacks were motion-corrected with the RELION-3 (ref. 3) implementation of MotionCor2 and motion-corrected micrographs were manually inspected and curated (graphene oxide coverage of grids was important to C2T correction with STEMHIDE4 (ref. 4) implemented in RELION-3. Particles were automatically picked with Gautomatch and extracted in RELION-3 for all further 2D and 3D processing steps. Auto-picked particles were examined by 2D classification, and particles in ‘bad’ classes corresponding to ice contamination or graphene oxide oxide fold lines were discarded. The first reference model was generated using RELION-3’s 3D initial model job and subsequently improved as continued 3D classification produced better maps.

For the CST–PPN complex, multiple processing strategies were pursued to generate higher-resolution maps but were unsuccessful owing to substantial heterogeneity among the particles. The reported standard image-processing pipeline was performed in two steps, first using C2T classification with the previously determined cryo-EM map and not modeled, so it was substituted in our initial model with a crystal structure as poor and it was substituted in our initial model with a crystal structure (refs. 19,20). All subunits were docked into the previously determined cryo-EM map and not modeled, so it was substituted in our initial model with a crystal structure. The resulting density map was sharpened by post-processing, and Fourier shell correlation (FSC) curves and an improved model job and subsequently improved as continued 3D classification produced better maps.

Starting models used in this study can be found in the Protein Data Bank under accession code 4Q2MR and 4Q2M3-F1. The cryo-EM map generated in this study has been deposited at the Electron Microscopy Data Exchange (EMD) under accession code EMD-26347 (CST–PPF) and EMD-26347 (CST–PPN), and the CST–PPM coordinates have been deposited in the Protein Data Bank under accession code PDB 7USC. Source data are provided with this paper.

References
31. Weissmann, F. et al. bigBac enables rapid gene assembly for the expression of large multisubunit protein complexes. Proc. Natl Acad. Sci. USA 113, E264–E269 (2016).
32. Tang, G. et al. EMA2: an extensible image processing suite for electron microscopy. J. Struct. Biol. 157, 38–46 (2007).
33. Mastronarde, D. N. Automated electron microscope tomography using robust prediction of specimen movements. J. Struct. Biol. 152, 36–51 (2005).
34. Zivanov, J. et al. New tools for automated high-resolution cryo-EM structure determination in RELION-3. J. Struct. Biol. 160, 1605–1612 (2004).
35. Holzer, S. et al. Structural basis for inhibition of human primase by arabinofuranosyl ribavirine. ACS Chem. Biol. 14, 1904–1912 (2019).
36. Croll, T. I. ISOLDE: a physically realistic environment for model building into low-resolution electron-density maps. Acta Crystallogr D. Struct. Biol. 74, 519–530 (2018).
39. Pettersen, E. F. et al. UCSF ChimeraX: structure visualization for researchers, educators, and developers. *Protein Sci.* 30, 70–82 (2021).
40. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr D. Biol. Crystallogr* 66, 486–501 (2010).
41. Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D. Biol. Crystallogr* 66, 213–221 (2010).
42. Waterhouse, A. M., Procter, J. B., Martin, D. M., Clamp, M. & Barton, G. J. Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics* 25, 1189–1191 (2009).
43. Larkin, M. A. et al. Clustal W and Clustal X version 2.0. *Bioinformatics* 23, 2947–2948 (2007).
44. Meyer, A. et al. Systematic analysis of protein-detergent complexes applying dynamic light scattering to optimize solutions for crystallization trials. *Acta Crystallogr F. Struct. Biol. Commun.* 71, 75–81 (2015).
45. Chi, H. et al. Comprehensive identification of peptides in tandem mass spectra using an efficient open search engine. *Nat. Biotechnol.* https://doi.org/10.1038/nbt.4236 (2018).
46. Chen, Z. L. et al. A high-speed search engine pLink 2 with systematic evaluation for proteome-scale identification of cross-linked peptides. *Nat. Commun.* 10, 3404 (2019).
47. Combe, C. W., Fischer, L. & Rappsilber, J. xiNET: cross-link network maps with residue resolution. *Mol. Cell Proteom.* 14, 1137–1147 (2015).

**Acknowledgements**

We are grateful to H. Suzuki and Y. Zhang for guidance with data processing. We thank M. Ebrahim, J. Sotiris, and H. Ng at the Evelyn Gruss Lipper Cryo-EM Resource Center of The Rockefeller University for assistance with cryo-EM data collection. We thank I. Nudelman of the Fisher Drug Discovery Resource Center of The Rockefeller University for providing training and guidance on the microscale thermophoresis experiments. S.W.C. is supported by the David Rockefeller Graduate Program and a National Science Foundation Graduate Research Fellowship under grant no. 1946429. J.C.Z. was previously supported by an NIH cancer biology training grant (T32 CA009673) and is currently the Lorraine Egan Fellow of the Damon Runyon Cancer Research foundation (DRG 2337-19). This work is supported by grants from the NIH (5 R35 CA210036) and the Breast Cancer Research Foundation (BCRF-19-036) to T.d.L. and by funding from the Howard Hughes Medical Institute, NIH (5 R01 GM126891) and Blavatnik Family Foundation to E.N. and V.S.

**Author contributions**

This study was conceived by T.d.L., T.W., J.C.Z., and S.W.C. S.W.C. purified and reconstituted protein complexes, performed EM studies, and carried out computational and biophysical analyses. J.C.Z. and S.W.C. designed and cloned expression constructs. CX-MS data were obtained by E.N. and V.S.M.W.B. assisted with negative-stain EM experiments. S.W.C. wrote the manuscript with input from all authors.

**Competing interests**

T.d.L. is a member of the SAB of Calico Life Sciences LLC, San Francisco. The other authors have no conflicts to declare.

**Additional information**

Extended data is available for this paper at https://doi.org/10.1038/s41594-022-00766-y.

**Supplementary information**

The online version contains supplementary material available at https://doi.org/10.1038/s41594-022-00766-y.

**Correspondence and requests for materials**

should be addressed to Thomas Walz or Titia de Lange.

**Peer review information**

*Nature Structural and Molecular Biology* thanks the anonymous reviewers for their contribution to the peer review of this work. Editor recognition statement Beth Moorefield was the primary editor on this article and managed its editorial process and peer review in collaboration with the rest of the editorial team.

**Reprints and permissions information** is available at www.nature.com/reprints.
Extended Data Fig. 1 | Reconstitutions of CST–PP.

(a) SEC-elution profile of CST–PP (top). The grey line indicates the fractions visualized on the SYPRO Ruby-stained SDS-PAGE gel (bottom). This reconstitution was independently reproduced at least three times with similar results.

(b) Negative-stain EM image-processing pipeline and averages of the 35 (of 50) most populated 2D classes of the reconstituted complex characterized in (a). The size of the 128-px box corresponds to 358 Å. The reconstitution was independently reproduced at least three times with similar results.

(c) Cryo-EM image-processing pipeline and 2D-class averages obtained after vitrification and imaging of the GraFixed complex shown in (c). See Table 1 for full data collection and processing details.
Extended Data Fig. 2 | Cryo-EM image-processing pipeline of CST–PP\textsuperscript{L}. \textbf{a}, Cryo-EM image-processing pipeline used for the CST–PP\textsuperscript{L} complex. Two datasets were combined in RELION 3.1 as separate optics groups, and particles were refined together with a reference generated from dataset no. 1. See Table 1 for full data collection and processing details. \textbf{b}, 2D-class averages from a cleanup step prior to 3D classification. Classes clearly representing ice contamination or graphene oxide edges were removed during cleanup.
Extended Data Fig. 3 | See next page for caption.
Extended Data Fig. 3 | GraFix preparation of CST--PPAN complex and analysis by cryo-EM. a, GraFix preparation of CST--PPAN--ssDNA. Top: Coomassie blue-stained SDS-PAGE gel (4–12% Tris-Glycine; Invitrogen) of GraFix fractions, showing formation of cross-linked CST--PPAN species. Fractions marked with an asterisk (*) were pooled for analysis by cryo-EM. Bottom: SYBR Gold-stained native PAGE gel (4–20% TBE; Invitrogen) of GraFix fractions, showing free and bound ssDNA. This reconstitution was independently reproduced at least three times with similar results. b, Selected high defocus motion-corrected cryo-EM micrograph of CST--PPAN (representative of 9,282 curated micrographs). c, Cryo-EM image-processing pipeline used for the CST--PPAN complex, including supervised 3D classification and focused 3D classification steps used to select particles with intact CST. Gold-standard FSC curves for the final map of CST--PPAN indicate a global resolution of 4.6 Å (FSC = 0.143). See Table 1 for full data collection and processing details. d, 2D-class averages obtained with the particle stack (131,850 particles) used to generate the final CST--PPAN map show blurring at the peripheral regions, indicating flexibility. e, Local resolution estimates of the CST--PPAN map. f, Orthogonal views of the CST--PPAN--ssDNA model fit into the cryo-EM map.
Extended Data Fig. 4 | Analysis of the CTC1-POLA1<sup>CTD</sup> interface surfaces. Surface representations of the CTC1-POLA1<sup>CTD</sup> interface shown in two orientations related by a 180° rotation and views CTC1 and POLA1<sup>CTD</sup> after separation to reveal the interface surfaces (outlined in black).<br>a, Surface overlaid with cartoon representation of CTC1–POLA1<sup>CTD</sup>. Identification of CTC1 and POLA1<sup>CTD</sup> binding surfaces, colored in dark blue.<br>b, Surface representation colored according to electrostatic potential.<br>c, Surface representation colored according to sequence conservation for metazoans using an alignment containing all species listed in Supplementary Table 1.
Extended Data Fig. 5 | See next page for caption.
Extended Data Fig. 5 | Conservation analysis of the CTC1-POLA1<sup>CTD</sup> interface. **a**, Multiple sequence alignment showing the CRL region in POLA1. Residues are colored according to the color scheme used by Clustal X<sup>43</sup>. Accession codes for sequences are listed in Supplementary Table 1. **b**, Sequences for POLA1<sup>CTD</sup> and CTC1 (when present) were modeled in AlphaFold<sup>2</sup><sup>19,20</sup> and aligned in PyMOL (Schrödinger) (Supplementary Table 1). The CRL in metazoans is colored cyan, and the corresponding insertion region in unicellular eukaryotes is colored in fuchsia. For CTC1, the elongated OB-D is colored yellow and the rest of the N-terminus (OB-A, OB-B, OB-C, three-helix bundle) is colored cyan.
Extended Data Fig. 6 | See next page for caption.
Extended Data Fig. 6 | Purification and microscale thermophoresis analysis of POLA1CTD CRL mutants. a, Coomassie blue-stained SDS-PAGE gels (4–12% (w/v) Bis-Tris, run in MES-SDS; Invitrogen) showing proteins used for MST analyses. POLA1CTD constructs run as heterogeneously sized smears on SDS-PAGE. Gels are representative of three independent protein preparations. b, Native PAGE (4–16% Bis-Tris; Invitrogen) gel showing POLA1CTD constructs and a 1 μg BSA loading control next to a NativeMark unstained protein standard. Gels are representative of three independent protein preparations. c, SEC-elution profiles of the POLA1CTD constructs (estimated MW ~20 kDa) superposed with gel-filtration standards (BioRad). d, Models generated with AlphaFold 2⁰⁵ of the human POLA1CTD constructs encoding CRLWT (cyan), CRLGGSGGS (red), or CRLS. pombe (fuchsia), demonstrating the accommodation of shortened linkers. e, Independent microscale thermophoresis experiments with POLA1CTD constructs and RED-tris-NTA-labeled His₆-MBP-CTC1-STNI-TEN1. Dissociation constant (K_D) values were calculated with the MO Affinity Analysis (Nanotemper) software. Error bars represent the s.e.m. for each data point calculated from three capillary scans. The combined data are shown in Fig. 2c.
Extended Data Fig. 7 | Reconstitution and analysis of the CST-POLA1$^\text{N}$ interaction. 

**a**, SEC-elution profile of reconstituted His$_6$-MBP-POLA1$^\text{N}$-CST (top). The grey line indicates the fractions visualized on the SYPRO Ruby-stained SDS-PAGE gel (4–12% Bis-Tris, run in MOPS-SDS; Invitrogen) (bottom). The black bar indicates the fractions pooled for negative-stain EM and CX-MS analysis (see Fig. 3). This reconstitution was independently reproduced at least three times with similar results. 

**b**, Negative-stain EM image-processing pipeline and averages of the 35 (of 50) most populated classes of the reconstituted complex characterized in **a**. The size of the 128-px box corresponds to 358 Å$^2$. 

**c**, Cross-links between POLA1$^\text{N}$ and CST subunits identified by CX-MS analysis of the CST–PPF$^\text{L}$ complex (Extended Data Fig. 1a). Blue solid lines: cross-links between CST and POLA1$^\text{N}$ shared with the His$_6$-MBP-POLA1$^\text{N}$-CST dataset (Fig. 3c). Blue dashed lines: cross-links between CST and POLA1$^\text{N}$ not shared with the His$_6$-MBP-POLA1$^\text{N}$-CST dataset (Fig. 2c). Black dashed lines: all other inter-subunit cross-links.
Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

| n/a | Confirmed |
|-----|-----------|
| ❌   | The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement |
| ❌   | A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly |
| ❌   | The statistical test(s) used AND whether they are one- or two-sided |
| ❌   | Only common tests should be described solely by name; describe more complex techniques in the Methods section. |
| ❌   | A description of all covariates tested |
| ❌   | A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons |
| ❌   | A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| ❌   | For null hypothesis testing, the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted |
| ❌   | Give P values as exact values whenever suitable. |
| ❌   | For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings |
| ❌   | For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes |
| ❌   | Estimates of effect sizes (e.g. Cohen’s d, Pearson’s r), indicating how they were calculated |

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection

SerialEM v3.6, Nanotemper MO Affinity Control v2.3

Data analysis

Relion 3.0, Relion 3.1, Gaitmatch 0.56, Cliffrad 1.06, UCSF Chimera 1.13, Phenix 1.17.1, UCSF ChimeraX 0.93, EMAN 2.1, Nanotemper MO Affinity Analysis v2.3, pFind3, xieNET (http://crosslinviewer.org/), PyMOL v2.1.1, GraphPad Prism v9.3.1, AlphaFold v2.0, Jaliwet 2.111.0, MUSCLE (https://www.ebi.ac.uk/Tools/msa/muscle/)

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub), See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

Starting models used in this study can be found in the Protein Data Bank under the accession codes PD3-GW6W, PD3-S6XR, and PD6 6R84 and in the Alphafold Protein Structure Database under accession code AF-O2WJ3-F1. The cryo-EM maps generated in this study have been deposited at the Electron Microscopy Data Bank under accession codes EMD-26346 (CST+PPE) and EMD-26347 (CST+FPI), and the CST+PPE coordinates have been deposited in the Protein Data Bank under accession code PDB-7USC. Source data for figures 2, 3, and Extended Data Figures 1, 3, 6, and 7 are provided with this paper.
Field-specific reporting

Please select the one that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☑ Life sciences  ☐ Behavioural & social sciences  ☐ Ecological, evolutionary & environmental sciences

*For a reference copy of the document with all sections, see [nature.com/documents/mr-reporting-summary-flat.pdf](http://nature.com/documents/mr-reporting-summary-flat.pdf)*

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

| Sample size | Sample sizes for the Cryo-EM datasets were determined by the need to obtain meaningful structures and the availability of cryo-EM time. For the CST-PP-Fl dataset, 6340 movie stacks were collected and 109,224 particles were used for the final reconstruction. For the CST-PP-deltaN dataset, 17732 movie stacks were collected and 131,850 particles were used for the final reconstruction, which was sufficient to yield a 4.6-Å resolution structure. For other experiments, no sample-size determination was performed, as the biochemical experiments followed standard practices. |
| Data exclusions | Micrographs clearly suffering from astigmatism, image drift, poor graphene oxide coverage, ice contamination, and/or cubic ice formation were excluded during the micrograph curation step in all cryo-EM datasets analyzed. Particles in 2D classes showing no secondary structural features and in 3D classes showing unsatisfactory structural features were excluded from the final reconstructions in all datasets analyzed. For the MST measurements, single point measurements were excluded if they were found to have aggregation auto-detected by the instrument. |
| Replication | For the high-resolution structure obtained in this study, a small dataset was first collected on a 200-kV Talos Arctica electron microscope. The structures were then confirmed and improved by data collection on a 300-kV Titan Krios electron microscope. For the MST measurements, three independent experiments were performed for each construct, and each experiment consisted of three technical replicate scans. Each biochemical reconstitution was performed at least three times and was reproducible. XLMS experiments were performed 1-5 times independently with the same protein preparation for depth coverage and source data are included. |
| Randomization | Not applicable to this study in general since it did not allocate for experimental groups. In the 3D refinement, particle images were split into two half groups. |
| Blinding | Blinding is generally not applicable to biochemical studies and structure determination by single particle cryo-EM and is not used in standard protocols for the experiments described here. Similarly, the MST analysis was not blinded. |

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

| Materials & experimental systems | Methods |
|----------------------------------|---------|
| n/a | n/a |
| ☒ Antibodies | ☐ ChiP-seq |
| ☒ Eukaryotic cell lines | ☒ Flow cytometry |
| ☒ Palaeontology and archaeology | ☐ MRI-based neuroimaging |
| ☐ Animals and other organisms | |
| ☐ Human research participants | |
| ☐ Clinical data | |
| ☐ Dual use research of concern | |

Eukaryotic cell lines

*Policy information about [cell lines](http://example.com/cell-lines)*

| Cell line source(s) | Commercial insect cell lines were used for baculovirus production (Gibco Sf9) and protein expression (Expression Systems TnI). |
| Authentication | All cell lines are commercially available and authenticated by the manufacturers by isozyme and karyotype analysis. |
| Mycoplasma contamination | Cell lines were tested by the manufacturer for mycoplasma contamination. |
| Commonly misidentified lines (see [ICLC register](http://example.com/iclc-register)) | None. |