CtIP tetramer assembly is required for DNA-end resection and repair

Owen R. Davies[1,4]*, Josep V. Forment[1,2,3]*, Meidai Sun[1], Rimma Belotserkovskaya[1,2], Julia Coates[1,2], Yaron Galanty[1,2], Mukerrem Demir[1,2], Christopher Morton[1], Neil Rzechorzek[1], Stephen P. Jackson[1,2,3], Luca Pellegrini[1]

*These authors contributed equally to this work

[1]Department of Biochemistry, University of Cambridge, Cambridge CB2 1GA, UK.

[2]The Gurdon Institute, University of Cambridge, Cambridge CB2 1QN, UK.

[3]The Wellcome Trust Sanger Institute, Hinxton CB10 1SA, UK.

[4]Present address: Institute for Cell and Molecular Biosciences, Faculty of Medical Sciences, Newcastle University, Framlington Place, Newcastle upon Tyne NE2 4HH, UK.

Authors for Correspondence

Luca Pellegrini: lp212@cam.ac.uk

Stephen P. Jackson: s.jackson@gurdon.cam.ac.uk

Mammalian CtIP protein plays major roles in DNA double-strand break (DSB) repair. While it is well established that CtIP promotes DNA-end resection in preparation for homology-dependent DSB repair, the molecular basis for this function remains unknown. Here we show by biophysical and X-ray crystallographic analyses that the N-terminal domain of human CtIP exists as a stable homotetramer. Tetramerisation results from interlocking interactions between the N-terminal extensions of the CtIP coiled coil region, leading to a 'dimer-of-dimers' architecture. Through interrogation of the structure, we identify a CtIP mutation that abolishes tetramerisation of the N-terminal domain while preserving its dimeric structure *in vitro*. Importantly, we establish that this mutation abrogates CtIP oligomer assembly in cells, leading to strong defects in DNA-end resection and gene conversion. These findings indicate that a CtIP tetramer architecture is essential for effective DSB repair by homologous recombination.

Keywords

CtIP/RBBP8, double-strand DNA break repair, DNA-end resection, gene conversion, homologous recombination.

The successful repair of DNA lesions is essential for cellular survival and the prevention of cancer[1]. Double-strand breaks (DSBs) are highly dangerous DNA lesions that are induced by various exogenous and endogenous agents, including ionising radiation and collapsed DNA replication forks. During late S and G2 phases of the cell cycle, high-fidelity repair of DSBs is promoted by homologous recombination (HR), in which the newly synthesised sister chromatid acts as a DNA repair template[2]. In contrast, during G1 and early S phases, the lack of a suitable homologous template means that DSBs must be repaired by the error-prone mechanism of non-homologous end joining (NHEJ)[3]. Optimal maintenance of genomic integrity thus depends on the successful execution of HR and NHEJ, and the appropriate choice of DSB repair pathway on the basis of cell cycle stage.

CtIP/RBBP8 was initially identified as a binding partner of the retinoblastoma protein (RB1)[4], CtBP (C-terminal binding protein 1)[5] and BRCA1[6], and as a transcription factor with regulatory functions in G1/S cell cycle progression and DNA damage-induced activation of the G2/M checkpoint[7-10]. Subsequent studies revealed its direct involvement in DNA DSB repair and the prevention of genomic instability[11]. It is now well established that CtIP plays key roles in DSB repair by HR and microhomology-mediated end joining (MMEJ)[12-15]. The DSB repair functions of CtIP are attributed to its role in promoting DNA-end resection, acting in functional and physical association with the MRE11-RAD50-NBS1 (MRN) complex[11]. CtIP recruitment to DSBs is required for effective formation of the RPA-coated 3'-ssDNA extensions that are a key intermediate in HR-mediated repair, and for activation of DNA-damage signalling by the protein kinase ATR[11,16]. CtIP activity is regulated in a cell-cycle dependent manner, which affects the choice of DSB repair pathway: in S and G2 phases, cyclin-dependent kinase (CDK) phosphorylation of CtIP residues S327 (which is necessary for interaction with BRCA1) and T847 promotes HR[12,13,17]. In addition, CtIP is important for the limited DNA-end resection that promotes MMEJ during G1 phase[12,14,15].

Extensive regions of amino acid conservation among vertebrate CtIP sequences are observed only at the N- and C-terminal ends of the molecule. Limited conservation of the CtIP C-terminal region extends to Sae2, the functional homologue of CtIP in *Saccharomyces cerevisiae*, which contains a CDK phosphorylation site that is functionally analogous to T847 of human CtIP[13,17]. The N- and C-terminal regions of CtIP are independently capable of interacting with the MRN complex, and deletion of either region is sufficient to abrogate the function of CtIP in DNA-end resection and G2/M checkpoint activation[11,18]. Primary sequence analysis of the conserved C-terminal region of CtIP has not provided clear insights into its mechanism of action, but it is likely to act as a recruitment module for macromolecular

interactions with protein and nucleic acids that are regulated by phosphorylation. In contrast, considerable α-helical structure is predicted for the N-terminal region, suggesting that it might mediate key aspects of CtIP structure and function.

Biophysical studies in solution have provided evidence that amino acid residues 45 to 160 are responsible for CtIP dimerisation via formation of an extended parallel coiled coil, interrupted in the middle by a putative zinc-binding motif Cys-X-X-Cys (residues 89-92, where X is any amino acid)[19,20]. However, recent evidence suggests that dimer formation is insufficient to explain the essential role of the CtIP N-terminal region in DNA repair, as deletion of residues 20-45 or 22-45, which precede the dimerisation domain, impairs DSB end-resection, HR, MMEJ and G2/M checkpoint activation[18,21]. Here, we demonstrate through biophysical and crystallographic analysis that the CtIP N-terminal region adopts a tetrameric structure, mediated by antiparallel association of helical extensions protruding from the N-termini of two parallel coiled-coil dimers. Furthermore, we show that a point mutation that disrupts tetramer formation, while retaining CtIP dimerisation, leads to failure of DNA-end resection and defective HR. We discuss potential functions for CtIP 'dimer-of-dimers' assembly and why it is crucial for effective DNA repair.

RESULTS

The CtIP N-terminal region forms a tetramer

To gain insight into the structure of the CtIP N-terminal domain (CtIP-NTD), we expressed the region spanning CtIP residues 18 to 145 in *E. coli* and purified it to homogeneity (**Fig. 1a** and **Supplementary Fig. 1**). Analysis by size-exclusion chromatography multi-angle light scattering (SEC-MALS) revealed that the CtIP-NTD forms a single tetrameric species corresponding to a molecular mass of 68.7 kDa (**Fig. 1b**). Furthermore, we found that the tetrameric structure is almost entirely α-helical and shows high thermal stability (**Supplementary Fig. 2a, b**). Notably, we identified by microbeam proton-induced X-ray emission (microPIXE) analysis the presence of bound zinc with an estimated CtIP-NTD:Zn^{2+} ratio of 2:1 (**Supplementary Fig. 3**), a conclusion supported by spectrophotometric assays based on 4-(2-pyridylazo)resorcinol (PAR)[22], where zinc was detected at 40% of the molar protein concentration (**Fig. 1c**). Furthermore, PAR analysis of a C89 and C92 double alanine mutant (CtIP-NTD-DM) detected only background quantities of zinc (**Fig. 1c**), identifying C89 and C92 as ligands for the metal. The tetrameric structure and α-helical content of CtIP-NTD-DM were essentially the same as the non-mutated protein (**Fig. 1b** and

Supplementary Fig. 2a), indicating that zinc is not important for tetramerisation. These data thus established that the CtIP N-terminal domain adopts a helical tetrameric structure with two bound zinc ions.

Defining the CtIP tetramerisation motif

To identify the region mediating CtIP tetramerisation within the CtIP-NTD, we truncated the N-terminal sequence that precedes the coiled-coil region (**Fig. 1a**). SEC-MALS analysis of the ensuing region, comprising residues 52 to 145 (CtIP-cNTD), revealed a dimeric structure of 23.6 kDa (**Fig. 1d**), indicating that residues N-terminal to the coiled-coil sequence are essential for tetramerisation. We found that, like CtIP-NTD, CtIP-cNTD is almost entirely α-helical, shows high thermal stability (**Supplementary Fig. 2c, d**) and contains bound zinc, in agreement with a stoichiometry of one zinc atom per CtIP-cNTD dimer (**Fig. 1e**). Importantly, while the double C89A C92A mutant protein no longer bound zinc (**Fig. 1e**), it retained a dimeric structure, indicating that zinc is not required for dimerisation. Nevertheless, the double C89A C92A mutation caused an increase in the SEC elution volume of both CtIP-NTD and CtIP-cNTD, accompanied by reduced thermal stability (**Fig. 1b, d** and **Supplementary Fig. 2**), suggesting that the zinc-binding motif may act as a rigid strut that holds the two coiled-coil segments in an extended conformation.

To investigate the nature of CtIP tetramerisation, we analysed the N-terminal sequence encoded by residues 18-52 (CtIP-nNTD). Remarkably, this short sequence of 35 amino acids was sufficient to support tetramerisation, as demonstrated by SEC-MALS analysis of CtIP-nNTD, both as a free polypeptide and when fused to a maltose binding protein (MBP) tag (**Fig. 1f** and **Supplementary Fig. 1c**). We found that this tetrameric structure is α-helical in nature (**Supplementary Fig. 2e, f**) and, in keeping with our other data, does not contain bound zinc (**Fig. 1e**). Collectively, these findings demonstrated that the CtIP N-terminal sequence comprising residues 18-52 is necessary and sufficient for CtIP tetramerisation, mediated by head-to-head association between CtIP dimers.

Crystal structure of the CtIP tetramerisation motif

To define the molecular basis for CtIP tetramerisation, we determined the X-ray crystal structure of human CtIP-nNTD at a resolution of 1.9 Å (**Table 1** and **Supplementary Fig. 1d**). Crystallographic analysis showed that the CtIP-nNTD adopts an entirely α-helical tetrameric structure, in which two CtIP dimers associate head-to-head through the

interdigitating interaction of their splayed N-terminal tails (**Fig. 2**). In this 'dimer-of-dimers' assembly, the antiparallel interactions between residues 18-31 in the CtIP-nNTD are responsible for tetramerisation, whereas residues 32-52 form parallel dimeric coiled-coils (**Fig. 2**). Each tetramerisation helix is involved in two distinct sets of antiparallel interactions with the two N-terminal helices of the opposing CtIP dimer. In one set of interactions, the aromatic side chains of tryptophan residues at position 24 in each monomer pack against each other in a stacking interaction that is reinforced by reciprocal hydrogen bonding between the indole ring nitrogen and the main-chain carbonyl group of the opposing tryptophan (**Fig. 3a**); in the second set, leucine residues 23 and 27 interdigitate in canonical antiparallel coiled-coil fashion (**Fig. 3b**). F20 contributes to both interfaces, by packing against the side chain of K28 (**Fig. 3a**) and by adding one coiled-coil interaction with C30 of the antiparallel chain (**Fig. 3b**). Thus, we conclude that the combination of anti-parallel interactions provides a hydrophobic core that buries the side chains of F20 and L27 and likely drives tetramer formation.

The crystal structure further reveals that the C-terminal limit of the CtIP tetrameric interface is defined by H31, which packs against F20 of the opposing anti-parallel chain to provide the outermost hydrophobic interaction of the tetramer as well as the first dimerisation contact in the coiled coil (**Fig. 3c**). A parallel coiled-coil extends from H31 to the end of CtIP-nNTD, with residues V35, L38, V42 and L45 occupying canonical heptad-repeat positions at the hydrophobic interface (**Fig. 3c**). The heptad repeat continues in the CtIP sequence until the zinc-binding site defined by C89-X-X-C92 (**Supplementary Fig. 1a**), indicating that the coiled-coil of the CtIP-nNTD is likely to be continuous with the coiled-coil structure of the CtIP-cNTD dimer. Taken together, our crystallographic and biophysical data lead to a structural model in which the CtIP-NTD undergoes tetramerisation through head-to-head 'dimer-of-dimers' association mediated by the short N-terminal CtIP-nNTD sequence.

L27E mutation blocks CtIP tetramerisation

Having defined the structural basis of CtIP tetramerisation, we sought to test its physiological importance. Thus, we designed a point mutation replacing buried L27 with glutamate (L27E), to reverse the chemical nature of the amino acid side chain. The mutation aimed to disrupt tetramerisation while preserving the dimeric structure of CtIP, by introducing a charged residue in the hydrophobic core of the CtIP tetramerisation motif (**Fig. 4a**). Importantly, SEC-MALS analysis confirmed that a CtIP-NTD protein bearing the L27E mutation can only form dimeric species, and thereby had an equivalent effect on preventing dimer-of-dimers

formation to deletion of amino acids 20-27, which removes the majority of the tetramerisation helix (**Fig. 4b**).

While we had shown that the L27E mutation does not interfere with CtIP-NTD dimerisation *in vitro*, we wanted to confirm that this is also the case for full-length CtIP *in vivo*. Thus, we stably transfected human U2OS cells with FLAG-tagged versions of full-length wild-type CtIP, the single-point mutant L27E and a truncated form of the protein lacking the entire N-terminal domain, Δ1-140 CtIP (**Fig. 4c**, top panel). Green-fluorescent protein (GFP)-tagged wild-type CtIP was then transiently expressed in these cells and immunoprecipitated. As revealed by their co-immunoprecipitation, both wild type and L27E forms of FLAG-CtIP interacted with GFP-CtIP, whereas Δ1-140 CtIP did not (**Fig. 4c**, bottom panel). Thus, the L27E mutation does not affect the ability of CtIP to dimerise in cells, in agreement with our biophysical data.

To assess the impact of the L27E mutation in cells, the FLAG-epitope tagged construct of either wild type or L27E mutant full-length CtIP protein was transfected transiently into human HEK-293 cells. FLAG-CtIP proteins were purified under stringent conditions so as to maintain only specific interactions (see Methods) and then subjected to size-exclusion chromatography. The later chromatographic elution of L27E CtIP indicated a substantial reduction in size to approximately half the apparent molecular weight of the native molecule (**Fig. 4d**). These findings demonstrated that full-length FLAG-tagged CtIP exists in cells in an oligomeric state, and established that oligomerisation is disrupted by the L27E mutation, consistent with head-to-head tetramerisation mediated by its N-terminal domain occurring *in vivo*.

CtIP tetramerisation is required for HR and promotes MMEJ

To test whether the DNA repair function of CtIP was affected by the L27E mutation, we took advantage of the fact that we had stably transfected wild-type and mutated CtIP constructs in a U2OS cell line containing the Traffic Light Reporter (TLR) system[23] integrated into its genome. This system consists of an expression cassette containing a truncated *eGFP* gene fused to an *mCherry* gene that is out-of-frame translationally, and hence not expressed (**Fig. 5a**). Transfection of cells with the endonuclease I-SceI generates a DSB in the *eGFP* gene that, when repaired by gene conversion (a form of HR) of an exogenous *eGFP* template, results in the production of a functional *eGFP* gene and cells acquiring green fluorescence. Alternatively, if the DSB is repaired by accurate NHEJ, the cassette is restored and no

fluorescence is detected. If, however, mutagenic DSB repair takes place that results in a +2 translational frameshift, the *mCherry* reading frame is restored, causing cells to fluoresce red (**Fig. 5a** and **Supplementary Fig. 4**). The TLR system therefore allows the simultaneous measurement of DSB repair in cells by either mutagenic end-joining (mutEJ) or HR via gene conversion[23].

To assess the effect of CtIP depletion in the TLR system, we performed the assay in control cells not expressing any recombinant form of CtIP and which had been treated with a control small-interfering RNA (siRNA) or an siRNA targeting the CtIP coding sequence. While CtIP depletion dramatically reduced the ability of cells to perform gene conversion (**Fig. 5b, left panel**), as described previously in other HR reporter systems[11], it slightly increased the occurrence of mutEJ events (**Fig. 5b, right panel**), an effect that can be explained by CtIP depletion causing an imbalance in the HR/NHEJ/mutEJ ratio[23,24]. Next, we carried out TLR assays with the U2OS-TLR cells that we had engineered to stably express siRNA-resistant wild type, L27E or Δ1-140 CtIP proteins (see 'FLAG' inputs in **Fig. 4c** for expression levels of each recombinant protein), and which had then been treated with an siRNA to deplete endogenous CtIP. As shown in **Fig. 5c**, whereas wild-type CtIP restored HR proficiency to cells depleted of endogenous CtIP protein, both L27E and Δ1-140 CtIP failed to do so, essentially behaving as null mutants. (**Supplementary Fig. 4i** shows that a similar trend was observed in regards to decreasing the rate of mutEJ). Taken together, these results indicate that dimerisation, which can still take place in the L27E CtIP mutant, is not sufficient for CtIP function in gene conversion and support a model in which CtIP 'dimer-of-dimers' tetramerisation is required for this HR pathway.

CtIP's function during DNA repair is not exclusively linked to HR, as cells depleted of CtIP also show defects in MMEJ[12-15]. MMEJ is an alternative end-joining pathway that requires limited resection of the DSB, and differs from the classical NHEJ pathway in that it is independent of the Ku heterodimer and DNA-PK. In MMEJ, CtIP-mediated limited resection is believed to unveil regions of micro-homology that are used to anneal and ligate the broken DNA ends[25]. We used a U2OS-based system (**Fig. 5d, left panel**; ref. 21) to assess the impact of CtIP depletion on MMEJ (**Fig. 5d, right panel**). We then generated U2OS-MMEJ cell lines stably expressing siRNA-resistant FLAG-tagged versions of CtIP and performed rescue experiments on cells depleted of endogenous CtIP by siRNA treatment. As shown in **Figure 5e**, and in agreement with previous reports[18], wild type CtIP largely complemented CtIP deficiency, while Δ1-140 CtIP behaved as a null mutant. Interestingly, two independent L27E CtIP clones (**Supplementary Fig. 5a**) partially rescued the MMEJ deficiency of CtIP-

depleted cells (**Fig. 5e**), raising the possibility that CtIP dimers could be functional to some degree in this DNA repair pathway.

CtIP tetramerisation promotes DNA-end resection

CtIP localisation to sites of DNA damage is crucial for its function during DNA repair[18,21]. To test whether CtIP tetramerisation is important for the accumulation of CtIP at damaged sites, laser micro-irradiation was performed on U2OS cells stably expressing wild type or L27E GFP-CtIP (**Supplementary Fig. 5b**). Strikingly, the L27E GFP-CtIP mutant protein showed a pronounced defect in localising to sites of DNA damage when compared to wild type GFP-CtIP (**Fig. 6a**). Similar effects were observed when cells were treated with other DNA-damaging agents such as hydroxyurea or mitomycin C (**Fig. 6b** and **Supplementary Fig. 5c;** note in **Fig. 6b** that the L27E mutation did not affect the amount of DNA damage induced by such agents, as assessed by induction of the S139-phosphorylated form of histone H2AX (γH2AX), a well-established marker of DNA damage[26]). It has been reported that the N-terminal part of CtIP, where the tetramerisation motif resides, is important for the constitutive interaction of CtIP with the MRN complex[18] and that the MRN complex is required for localisation of CtIP to DNA-damage sites[16]. However, the L27E GFP-CtIP mutant protein interacts with the components of the MRN complex similarly to wild-type GFP-CtIP (**Fig. 6c**).

The inability of CtIP-depleted cells to perform HR is at least in part explained by a defect in DNA-end resection[11]. Resection of a DSB is believed to be essential for all forms of HR, and involves the 5'-to-3' nucleolytic degradation of DNA ends to generate single-stranded DNA (ssDNA) that is then rapidly bound by the ssDNA-binding protein replication protein A (RPA). Subsequently, RPA is replaced by RAD51 to form the nucleoprotein filaments that are necessary for strand invasion and completion of gene conversion[27]. To test whether the HR defect we observed in cells expressing the L27E CtIP mutant reflected impaired resection, we performed a flow-cytometry assay to detect RPA retention on chromatin[28]. RPA is present at all cell-cycle stages but only becomes chromatin-bound during S phase and when ssDNA is exposed under specific circumstances, such as when a DSB is resected. When CtIP is depleted, chromatin-bound RPA can still be detected in S-phase cells, but there is a dramatic reduction in DNA-damage dependent RPA chromatinization[28]. The DNA topoisomerase I inhibitor camptothecin generates DSBs at DNA replication forks[29] that can only be repaired by HR and are channelled into the end-resection pathway[30]. We found that, after camptothecin treatment, CtIP-depleted cells exhibited a dramatic reduction of RPA chromatinisation that was rescued by expression of wild-type CtIP (**Fig. 6d**, compare bottom

panels 1 and 2 showing an increase from 0.5 to 25% RPA-positive cells). Notably, however, neither L27E CtIP nor Δ1-140 CtIP rescued the resection defect, despite all cell populations containing similar levels of DNA damage as scored by γH2AX immunostaining (**Fig. 6d** and **Supplementary Fig. 6**). Thus, the HR defect sustained by cells expressing L27E CtIP or Δ1-140 CtIP correlated with their inability to engage DSBs in DNA-end resection. Collectively, these results implied that CtIP dimers are not functional during DNA-end resection and that tetramerisation is crucial for CtIP localisation at DNA-damage sites, resection and HR.

DISCUSSION

We have demonstrated that mutation of CtIP to prevent its tetrameric assembly abrogates CtIP function in DNA-end resection and HR. Our crystallographic analysis shows that CtIP tetramer formation depends on interlocking interactions between N-terminal ends of opposing coiled-coil dimers. Although the crystal structure includes only residues up to position 52, the parallel coiled-coil structure is likely to extend to the whole CtIP N-terminal region up to residue 145, with the only interruption caused by the zinc-binding motif centred on C89 and C92. While the role of zinc coordination by CtIP is unclear (a C89A C92A CtIP mutant is proficient in DNA-end resection, our unpublished results), our data suggest that it contributes to the conformational rigidity of the coiled-coil region.

Importantly, our analyses indicate that full-length CtIP undergoes multimeric assembly in cells, in a manner consistent with tetramerisation through head-to-head assembly of CtIP dimers. Indeed, we find that the apparent molecular weight of full-length CtIP expressed in cells is reduced by approximately half upon introduction of the single-point mutation L27E, which we established biochemically to block tetrameric assembly while preserving dimerisation of the CtIP N-terminal domain. As the L27E mutation behaves as a null CtIP mutant in DNA-end resection and HR, this leads us to conclude that the 'dimer-of-dimers' tetramer represents the functional oligomeric state of CtIP in these processes.

In agreement with our findings, previous biochemical studies identified a dimeric structure for CtIP residues 45-160[19,20]. It was further reported that sedimentation equilibrium data of a CtIP protein spanning residues 22 to 160 could not be interpreted as a single species due to non-specific aggregation of the sample[19]. We note that this construct, while including the majority of residues involved in tetramerisation, omits F20. Importantly, we have found that mutating F20 to glutamate leads to a mixed population of CtIP-NTD dimers and tetramers (**Supplementary Fig. 7a**), demonstrating that F20 is required for stable tetrameric

assembly. In previous studies, deletions of CtIP residues 20-45, 22-45 or the combined mutations L27P, K28A, E29A, E34A were also found to block CtIP function in DNA-end resection and HR[18,21]. On the basis of our *in vitro* and *in vivo* data, it is now apparent that the effects of such mutations can be explained by disruption of CtIP tetramerisation.

Why is the L27E mutant protein unable to fulfill CtIP's roles in DNA-end resection and repair? It has been recently reported that CtIP possesses endonuclease activity[31,32]. However, CtIP's function as an endonuclease seems to be dispensable for DNA-end resection and, importantly, is not affected by the deletion of the entire N-terminal region of CtIP[32]. Thus, the functional effects we have observed when abrogating CtIP tetramerisation cannot reflect a possible negative impact on CtIP's putative nuclease activity. Perhaps surprisingly, we found that although L27E CtIP remains capable of interacting with the MRN complex, its accumulation at damage sites is impaired. This observation highlights the importance of CtIP's tetramerisation state for its effective accrual at DNA-damage sites, although the precise nature of the underlying molecular mechanism remains unclear. In this regard, we note that CtIP function in DNA repair is intimately associated with that of the MRN complex. A well-established feature of the MRN complex is its ability to undergo long-range self-association mediated by a 'zinc hook' in the centre of a 600Å coiled-coil insertion in its RAD50 subunit[33]. These long-range interactions are thought to enable the complex to simultaneously engage multiple DNA molecules, a function that appears essential for effective HR repair. The molecular model of CtIP emerging from our data presents intriguing similarities to the functional architecture of the MRN complex. Given that the CtIP tetrameric structure spans 80 Å and each coiled-coil dimer contains up to eleven heptad repeats, a linear organisation of the CtIP N-terminal domain may reach up to ~300 Å in length (**Fig. 7**). Thus, a CtIP tetramer might be able to to position its C-terminal domains simultaneously onto distinct DNA molecules, where they would cooperate with the MRN complex and other factors to promote repair. In agreement with this model, we have found that a region of CtIP spanning residues 769 to 897 (C-terminal end of the protein) possesses DNA-binding ability (**Supplementary Fig. 7b**).

Finally, we note that Sae2, the functional orthologue of CtIP in *S. cerevisiae*[34,35], has also been shown to assemble into multimers[36], and it was recently reported that post-translational modifications of Sae2 regulate its oligomeric state and function[37]. Although Sae2 is considerably smaller than CtIP, its N-terminal region is predicted to be helical, suggesting that Sae2 might also adopt a tetrameric structure. This therefore raises the prospect that similar regulatory mechanisms could act to switch CtIP between dimeric and tetrameric states and thereby control CtIP function in human cells. In this respect, we note that, while

we have demonstrated that the 'dimer-of-dimers' tetrameric assembly is essential for end resection and HR, the results of the MMEJ assay suggest that CtIP dimers could function in other cellular processes, such as alternative modes of DNA repair or in gene transcription.

Accession codes. Coordinates and structure factors have been deposited in the Protein Data Bank under accession code 4D2H.

ACKNOWLEDGEMENTS

We thank Mairi Kilkenny for help with the collection of X-ray diffraction data, Andrew Sharff and Peter Keller for help with X-ray data processing and Joseph Maman for assistance with SEC-MALS. This work was supported by a Wellcome Trust Senior Fellowship award in basic biomedical sciences to L.P., an Isaac Newton Trust research grant to L.P. and O.R.D. and a Cambridge Overseas Trust PhD studentship to M.D.S. Research in the S.P.J. laboratory is funded by Cancer Research UK programme grant C6/A11224, the European Research Council, and the European Community Seventh Framework Programme grant agreement no. HEALTH-F2-2010-259893 (DDResponse). Core funding is provided by CRUK (C6946/A14492) and the Wellcome Trust (WT092096). S.P.J. receives his salary from the University of Cambridge, UK, supplemented by CRUK. J.V.F. is funded by Cancer Research UK programme grant C6/A11224 and the Ataxia Telangiectasia Society. R.B. and J.C. are funded by Cancer Research UK programme grant C6/A11224. Y.G. and M.D. are funded by the ERC grant DDREAM.

AUTHOR CONTRIBUTIONS

O.R.D. and M.S. expressed and purified the CtIP proteins and performed the biochemical and biophysical experiments; C.R.M. and N.J.R. performed the EMSA assays; O.R.D. determined the X-ray crystal structure of CtIP-nNTD; J.V.F. performed the TLR and MMEJ assays, the DNA-end resection assays, the immunofluorescence (with help from Y.G.), the GFP-CtIP quantification on FACS and the co-immunoprecipitation experiments; R.B. performed the gel filtration analysis of FLAG-CtIP from HEK-293T cell extracts and created the FLAG-tagged CtIP constructs; J.C. established the U2OS-TLR and the U2OS-MMEJ stable cell lines; M.D. established the U2OS GFP-CtIP cell lines; J.V.F and Y.G. established the TLR system in U2OS cells; O.R.D., J.V.F., S.P.J. and L.P. designed experiments and wrote the paper.

References

1. Jackson, S. P. & Bartek, J. The DNA-damage response in human biology and disease. *Nature* **461,** 1071–1078 (2009).
2. Krejci, L., Altmannova, V., Spirek, M. & Zhao, X. Homologous recombination and its regulation. *Nucleic Acids Res.* **40,** 5795–5818 (2012).
3. Lieber, M. R. The Mechanism of Double-Strand DNA Break Repair by the Nonhomologous DNA End-Joining Pathway. *Annu. Rev. Biochem.* **79,** 181–211 (2010).
4. Fusco, C., Reymond, A. & Zervos, A. S. Molecular cloning and characterization of a novel retinoblastoma-binding protein. *Genomics* **51,** 351–358 (1998).
5. Schaeper, U., Subramanian, T., Lim, L., Boyd, J. M. & Chinnadurai, G. Interaction between a cellular protein that binds to the C-terminal region of adenovirus E1A (CtBP) and a novel cellular protein is disrupted by E1A through a conserved PLDLS motif. *J. Biol. Chem.* **273,** 8549–8552 (1998).
6. Wong, A. K. *et al.* Characterization of a carboxy-terminal BRCA1 interacting protein. *Oncogene* **17,** 2279–2285 (1998).
7. Li, S. *et al.* Functional link of BRCA1 and ataxia telangiectasia gene product in DNA damage response. *Nature* **406,** 210–215 (2000).
8. Chen, P. L. *et al.* Inactivation of CtIP leads to early embryonic lethality mediated by G1 restraint and to tumorigenesis by haploid insufficiency. *Mol. Cell. Biol.* **25,** 3535–3542 (2005).
9. Liu, F. & Lee, W.-H. CtIP activates its own and cyclin D1 promoters via the E2F/RB pathway during G1/S progression. *Mol. Cell. Biol.* **26,** 3124–3134 (2006).
10. Yu, X. & Chen, J. DNA damage-induced cell cycle checkpoint control requires CtIP, a phosphorylation-dependent binding partner of BRCA1 C-terminal domains. *Mol. Cell. Biol.* **24,** 9478–9486 (2004).
11. Sartori, A. A. *et al.* Human CtIP promotes DNA end resection. *Nature* **450,** 509–514 (2007).
12. Yun, M. H. & Hiom, K. CtIP-BRCA1 modulates the choice of DNA double-strand-break repair pathway throughout the cell cycle. *Nature* **459,** 460–463 (2009).
13. Huertas, P. *et al.* CDK targets Sae2 to control DNA-end resection and homologous recombination. *Nature* **455,** 689 (2008).
14. Zhang, Y. & Jasin, M. An essential role for CtIP in chromosomal translocation formation through an alternative end-joining pathway. *Nat. Struct. Mol. Biol.* **18,** 80–84 (2011).
15. Lee-Theilen, M., Matthews, A. J., Kelly, D., Zheng, S. & Chaudhuri, J. CtIP promotes microhomology-mediated alternative end joining during class-switch recombination. *Nat. Struct. Mol. Biol.* **18,** 75–79 (2011).
16. You, Z. *et al.* CtIP Links DNA Double-Strand Break Sensing to Resection. *Mol. Cell* **36,** 954–969 (2009).
17. Huertas, P. & Jackson, S. P. Human CtIP mediates cell cycle control of DNA end resection and double strand break repair. *J. Biol. Chem.* **284,** 9558–9565 (2009).
18. Yuan, J. & Chen, J. N terminus of CtIP is critical for homologous recombination-mediated double-strand break repair. *J. Biol. Chem.* **284,** 31746–31752 (2009).
19. Dubin, M. J. *et al.* Dimerization of CtIP, a BRCA1- and CtBP-interacting protein, is mediated by an N-terminal coiled-coil motif. *J. Biol. Chem.* **279,** 26932–26938 (2004).
20. Stokes, P. H., Thompson, L. S., Marianayagam, N. J. & Matthews, J. M. Dimerization of CtIP may stabilize in vivo interactions with the Retinoblastoma-pocket domain. *Biochem. Biophys. Res. Commun.* **354,** 197–202 (2007).

21. Wang, H. *et al.* CtIP Protein Dimerization Is Critical for Its Recruitment to Chromosomal DNA Double-stranded Breaks. *J. Biol. Chem.* **287,** 21471–21480 (2012).
22. Säbel, C. E., Shepherd, J. L. & Siemann, S. A direct spectrophotometric method for the simultaneous determination of zinc and cobalt in metalloproteins using 4-(2-pyridylazo)resorcinol. *Anal. Biochem.* **391,** 74–76 (2009).
23. Certo, M. T. *et al.* Tracking genome engineering outcome at individual DNA breakpoints. **8,** 671–676 (2011).
24. Gomez-Cabello, D., Jimeno, S., Fernández-Ávila, M. J. & Huertas, P. New tools to study DNA double-strand break repair pathway choice. *PLoS ONE* **8,** e77206 (2013).
25. McVey, M. & Lee, S. E. MMEJ repair of double-strand breaks (director's cut): deleted sequences and alternative endings. *Trends Genet.* **24,** 529–538 (2008).
26. Sharma, A., Singh, K. & Almasan, A. Histone H2AX phosphorylation: a marker for DNA damage. *Methods Mol. Biol.* **920,** 613–626 (2012).
27. Huertas, P. DNA resection in eukaryotes: deciding how to fix the break. *Nature structural & molecular biology* **17,** 11–16 (2010).
28. Forment, J. V., Walker, R. V. & Jackson, S. P. A high-throughput, flow cytometry-based method to quantify DNA-end resection in mammalian cells. *Cytometry A* **81,** 922–928 (2012).
29. Pommier, Y. Topoisomerase I inhibitors: camptothecins and beyond. *Nat Rev Cancer* **6,** 789–802 (2006).
30. Saleh-Gohari, N. *et al.* Spontaneous homologous recombination is induced by collapsed replication forks that are caused by endogenous DNA single-strand breaks. *Mol. Cell. Biol.* **25,** 7158–7169 (2005).
31. Makharashvili, N. *et al.* Catalytic and Noncatalytic Roles of the CtIP Endonuclease in Double-Strand Break End Resection. *Mol. Cell* **54,** 1022–1033 (2014).
32. Wang, H. *et al.* CtIP Maintains Stability at Common Fragile Sites and Inverted Repeats by End Resection-Independent Endonuclease Activity. *Mol. Cell* **54,** 1012–1021 (2014).
33. Hopfner, K.-P. *et al.* The Rad50 zinc-hook is a structure joining Mre11 complexes in DNA recombination and repair. *Nature* **418,** 562 (2002).
34. Lobachev, K. S., Gordenin, D. A. & Resnick, M. A. The Mre11 Complex Is Required for Repair of Hairpin-Capped Double-Strand Breaks and Prevention of Chromosome Rearrangements. *Cell* **108,** 183–193 (2002).
35. Clerici, M., Mantiero, D., Lucchini, G. & Longhese, M. P. The Saccharomyces cerevisiae Sae2 Protein Promotes Resection and Bridging of Double Strand Break Ends. *J. Biol. Chem.* **280,** 38631–38638 (2005).
36. Lengsfeld, B. M., Rattray, A. J., Bhaskara, V., Ghirlando, R. & Paull, T. T. Sae2 Is an Endonuclease that Processes Hairpin DNA Cooperatively with the Mre11/Rad50/Xrs2 Complex. *Mol. Cell* **28,** 638–651 (2007).
37. Fu, Q. *et al.* Phosphorylation-regulated transitions in an oligomeric state control the activity of the Sae2 DNA repair enzyme. *Mol. Cell. Biol.* **34,** 778–793 (2014).
38. Peranen, J., Rikkonen, M., Hyvonen, M. & Kaariainen, L. T7 vectors with modified T7lac promoter for expression of proteins in Escherichia coli. *Anal. Biochem.* **236,** 371–373 (1996).
39. Sreerama, N. & Woody, R. W. Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. *Anal. Biochem.* **287,** 252–260 (2000).
40. Whitmore, L. & Wallace, B. A. Protein secondary structure analyses from circular dichroism spectroscopy: methods and reference databases. *Biopolymers* **89,** 392–400 (2008).
41. Kabsch, W. Xds. *Acta Crystallogr D Biol Crystallogr* **66,** 125–132 (2010).

42. Collaborative Computational Project, N. 4. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr* **D50,** 760–763 (1994).
43. McCoy, A. J. *et al.* Phaser crystallographic software. *Journal of applied crystallography* **40,** 658–674 (2007).
44. Kelley, L. A. & Sternberg, M. J. E. Protein structure prediction on the Web: a case study using the Phyre server. *Nat Protoc* **4,** 363–371 (2009).
45. Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66,** 213–221 (2010).
46. Blanc, E. *et al.* Refinement of severely incomplete structures with maximum likelihood in BUSTER-TNT. *Acta Crystallogr D Biol Crystallogr* **D60,** 2210–2221 (2004).
47. Emsley, P. & Cowtan, K. Coot: Model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60,** 2126–2132 (2004).
48. Dignam, J. D., Lebovitz, R. M. & Roeder, R. G. Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. *Nucleic Acids Res.* **11,** 1475–1489 (1983).
49. Richardson, C., Moynahan, M. E. & Jasin, M. Double-strand break repair by interchromosomal recombination: suppression of chromosomal translocations. *Genes Dev.* **12,** 3831–3842 (1998).
50. Limoli, C. L. & Ward, J. F. A new method for introducing double-strand breaks into cellular DNA. *Radiation research* (1993).
51. Bekker-Jensen, S. *et al.* Spatial organization of the mammalian genome surveillance machinery in response to DNA strand breaks. *J. Cell Biol.* **173,** 195–206 (2006).

FIGURE LEGENDS

Figure 1 CtIP-NTD (amino acids 18-145) exists as a tetramer in solution. (**a**) Schematic diagram of the human CtIP sequence. Based on bioinformatics analysis, the CtIP N-terminal domain (NTD) can be divided in a helical (α) N-terminal region (CtIP-nNTD) and a C-terminal region (CtIP-cNTD) in which a central C89-X-X-C92 motif is flanked by predicted coiled coils (CC). The conserved C-terminal domain (CTD) and the two CDK phosphorylation sites are also depicted. (**b**) Size-exclusion chromatography multi-angle light scattering (SEC-MALS) analysis of wild type and C89A/C92A double mutant (DM) recombinant CtIP-NTD. CtIP-NTD and CtIP-NTD-DM are tetrameric species of 68.7 kDa and 62.1 kDa, respectively. The light scattering (LS) as relative Raleigh ratio and the differential Reflective Index (dRI) are drawn as solid and dashed lines, respectively. The values for the fitted molecular masses (M_r) of CtIP-NTD and CtIP-NTD-DM are shown as diamond shapes across the respective elution peaks. The predicted molecular masses of the monomeric species are shown in brackets above each elution peak. (**c**) Spectrophotometric determination of zinc content using metal chelator 4-(2-pyridylazo) resorcinol (PAR), with zinc standards shown in grey. Zinc content was estimated as a percentage of the molar protein concentration at 40% and 0.5% for CtIP-NTD (solid line) and CtIP-NTD-DM (dashed line) respectively. (**d**) SEC-MALS analysis of

CtIP-cNTD and CtIP-cNTD-DM, revealing dimeric species of 23.6 kDa and 21.5 kDa, respectively. (**e**) PAR analysis in which zinc content was estimated at 33%, 0.6% and 0.4% for CtIP-cNTD (solid line), CtIP-cNTD-DM (dashed line) and CtIP-nNTD (dotted line) respectively. (**f**) SEC-MALS analysis of CtIP-nNTD reveals a tetrameric species of 17.3 kDa.

Figure 2 Crystal structure of the tetramerisation domain of human CtIP (CtIP-nNTD; amino acids 18-52). The CtIP-nNTD tetramer is formed through the head-to-head association between two CtIP dimers, drawn as ribbons and coloured orange and purple. The N- and C-termini of each CtIP molecule are labelled. The central tetrameric assembly consists of interlocking, antiparallel interactions between the two pairs of N-terminal helices comprising amino acids 18-31. The C-terminal amino acids 32-52 are part of the region responsible for CtIP dimerisation via parallel coiled-coil interactions. Two views of the CtIP-nNTD structure are shown, related by 90° rotation around the long axis of the tetramer.

Figure 3 Details of the homotypic interactions responsible for tetrameric assembly (**a, b**) and dimerisation (**c**) of CtIP. Each panel shows a ribbon drawing of the CtIP-nNTD (top) and side-chain details of CtIP interactions responsible for self-association (bottom). For clarity, only the two CtIP chains for which details of side-chain interactions are shown are coloured (as in Figure 2), whereas the other two chains are in grey. (**a**) Antiparallel interactions in the tetramerisation motif (residues 18-31) involve side-chain stacking of tryptophan residues at position 24 and reciprocal hydrogen bonding (depicted as yellow springs) between their indole ring nitrogens and main-chain carbonyl groups. The side chains of interacting residues F20 and K28 are also shown. (**b**) Coiled-coil contacts in the tetramerisation motif involve residues F20, L23, L27 and C30 from adjacent pairs of antiparallel helices. (**c**) The coiled-coil dimer is formed between parallel C-terminal sequences (residues 31-52), with residues V35 and V42 at the 'a' position and residues H31, L38 and L45 at the 'd' position of the heptad repeat. Histidine residue 31 marks the transition point between the tetramerisation motif and the coiled-coil dimer structure.

Figure 4 CtIP mutation L27E abrogates tetramer assembly while preserving coiled-coil dimerisation. (**a**) L27E introduces a charged residue in the hydrophobic core of the tetramer. CtIP is drawn as ribbon; tetramerisation region is coloured red, the rest of the chain in yellow. The side chain of L27 is shown in green as stick drawing. (**b**) SEC-MALS analysis revealing that CtIP-NTD L27E and Δ20-27 proteins are dimeric species. The LS as relative Raleigh ratio, the dRI and the M_r are drawn as in Figure 1. The predicted molecular masses of the monomeric species are shown in brackets above each elution peak. (**c**) *Top panel:* diagram of the FLAG-tagged CtIP derivatives used in the immunoprecipitation (IP)

experiments. The CtIP-NTD region is coloured as in panel **a**. *Bottom panel:* western blot showing co-immunoprecipitation of wild-type (WT) GFP-CtIP with the WT and L27E versions of FLAG-CtIP, but not with CtIP Δ1-140. Control IP was performed with cells that express WT FLAG-CtIP transfected with a GFP-expressing plasmid. (**d**) Western blot of the fractions from a gel filtration experiment shows that FLAG-tagged WT CtIP (upper panel, peak at fraction 3) elutes earlier than the corresponding L27E mutant (lower panel, peak at fraction 7). Inputs are shown in the first lane. The elution volume of each fraction, the elution position for size markers ferritin (440 kDa) and aldolase (58 kDa) and the estimated sizes of WT FLAG-CtIP and L27E FLAG-CtIP are also shown. Uncropped images for all western blot data can be seen in Supplementary Fig. 8.

Figure 5 CtIP mutation L27E impairs HR and, to a lesser extent, MMEJ. (**a**) Schematic of the Traffic Light Reporter (TLR) system (adapted from ref. 23; see also Supplementary Fig. 4). (**b**) CtIP depletion reduces HR efficiency (left panel) and slightly increases the rate of mutagenic end-joining (right panel) in the TLR system. (**c**) Only expression of wild-type (WT) FLAG-tagged CtIP rescues the HR deficiency caused by endogenous CtIP depletion. All values are normalised to cells transfected with control siRNA. (**d**) *Left panel*: schematic of the MMEJ system (adapted from ref. 21). *Right panel*: CtIP depletion impairs MMEJ. (**e**) L27E FLAG-CtIP is not completely defective in MMEJ. EV: empty vector. All quantifications are averages of n = 3 biological replicates (independent siRNA transfections). Error bars are ± SEM.

Figure 6 CtIP mutation L27E impairs CtIP accumulation at DNA-damage sites and DNA-end resection. (**a**) L27E GFP-CtIP shows defective accumulation at laser-induced DNA-damage sites. Samples were fixed 30 min after irradiation. Cells were stained with γH2AX (to detect damaged cells) and cyclin A (to identify cells in S and G2 phases of the cell cycle) antibodies. Note that CtIP is not recruited to DNA-damage sites in G1 cells. (**b**) L27E GFP-CtIP also shows defective accumulation in other DNA-damaging conditions. Cells were left untreated or treated for 24 h with 1 mM hydroxyurea (HU) or 120 ng/ml mytomycin C (MMC). Quantification of γH2AX-positive cells is shown as control for DNA damage (see also Supplementary Fig. 5c) (**c**) L27E GFP-CtIP interacts with the MRN complex. (**d**) Only expression of wild-type (WT) FLAG-tagged CtIP rescues the resection defect caused by endogenous CtIP depletion. *Left panel:* dot plots representing the intensity of the signal of RPA staining (*y* axis) against the DNA content (*x* axis). Quantification gates were established in untreated samples (Supplementary Fig. 6). *Right panel:* quantification of RPA-positive cells. All values are normalised to the percentage of γH2AX-positive cells and to

cells transfected with control siRNA (Supplementary Fig. 6). EV: empty vector. All quantifications are averages of n = 3 biological replicates (independent siRNA transfections). Error bars are ± SEM. Uncropped images for all western blot data can be seen in Supplementary Fig. 8.

Figure 7 Model for the functional architecture of human CtIP. The central tetrameric assembly of the NTD holds CtIP dimers in a head-to-head orientation. The two ends of this central tetramer extend into coiled-coil structures that define the CtIP dimers. Each dimer comprises two coiled-coil regions flanking a central zinc-binding site, in which one zinc ion is co-ordinated by C89 and C92 of the two CtIP chains. The CtIP-NTD may span a total distance of up to 300 Å.

ONLINE METHODS

Recombinant protein expression and purification

Sequences corresponding to human CtIP residues 1-145, 18-145, 18-52, 52-145 and 769-897, including point mutants L27E, F20E, C89A/C92A and Δ20-27, were cloned into pMAT11 vectors[38] for expression as N-terminal His_6-MBP fusion proteins. Constructs were expressed in BL21(DE3)Rosetta2 cells (Novagen®), in 2xYT media, induced with 0.5 mM IPTG for 16 hours at 20°C. Fusion proteins were purified from clarified lysate through consecutive Ni-NTA (Qiagen) and amylose (NEB) affinity chromatography. Affinity tags were removed by incubation with TEV protease (Invitrogen) and reverse binding to Ni-NTA resin. Further purification was achieved through anion exchange chromatography (Resource Q: GE Healthcare) and size-exclusion chromatography (HiLoad™ 16/600 Superdex 200: GE Healthcare) in 20 mM Tris pH 8.0, 150 mM NaCl, 2 mM DTT. Protein samples were concentrated to 10 mg/ml using Amicon Ultra® 10,000 MWCO centrifugal filter units (Millipore), and were stored at -80˚C following flash-freezing in liquid nitrogen.

Circular dichroism (CD) spectroscopy

Far-UV CD spectroscopy data were collected on an Aviv 410 spectropolarimeter (Biophysics facility, Department of Biochemistry, University of Cambridge). CtIP samples were analysed at 0.15 mg/ml, in 10mM Na_2HPO_4/ NaH_2PO_4 pH 7.5, using a 1 mm path-length quartz cuvette (Hellma). CD spectra were recorded at 0.5 nm intervals between 260 and 185 nm (with 1 nm slit width and 1 s averaging time), at 5°C. Mean residue ellipticity ([θ]) was calculated from smoothed corrected raw data averaged across three recordings, and

deconvolution was performed using the CDSSTR algorithm of the Dichroweb server (http://dichroweb.cryst.bbk.ac.uk)[39,40]. CD thermal denaturation data were recorded at 222 nm, at 1°C intervals between 5 and 95°C (1°C per minute ramping rate, with 30 s incubation time). Raw data were converted to mean residue ellipticity ($[\theta_{222}]$).

Size-exclusion chromatography multi-angle light scattering (SEC-MALS)

The absolute molecular masses of recombinant CtIP protein samples were determined by size-exclusion chromatography multi-angle light scattering (SEC-MALS). 100 *µ*l protein samples (at approximately 2 mg/ml) were loaded onto a Superdex™ 200 10/300 GL size exclusion chromatography column (GE Healthcare) in 20 mM Tris pH 8.0, 150 mM NaCl, 2 mM DTT, at 0.5 ml/min using an ÄKTA™ Purifier (GE Healthcare). The column output was fed into a DAWN® HELEOS™ II MALS detector (Wyatt Technology), followed by an Optilab® T-rEX™ differential refractometer (Wyatt Technology). Light scattering and differential refractive index data were collected and analysed using ASTRA® 6 software (Wyatt Technology). Molecular masses and estimated errors were calculated across individual eluted peaks by extrapolation from Zimm plots using a dn/dc value of 0.1850 ml/g. SEC-MALS data are presented with light scattering (LS) and differential refractive index (dRI) plotted alongside fitted molecular weights (M_W).

Spectrophotometric analysis of zinc content

A method for the spectrophotometric determination of zinc content was adapted from the published protocol[22]. CtIP recombinant protein samples at 0.5-0.1 mg/ml were digested with 0.6 *µ*g/*µ*l proteinase K (NEB) at 60°C for 1 hour. For each digested protein sample, in addition to standard solutions containing 0-100 *µ*M zinc acetate, 10 *µ*l of supernatant was added to 80 *µ*l of 50 *µ*M 4-(2-pyridylazo)-resorcinol (PAR) in 20 mM Tris pH 8.0, 300 mM NaCl, incubated for 5 minutes at room temperature, and UV absorbance spectra recorded between 600 and 300 nm (Varian Cary 50 spectrophotometer). Zinc concentrations were estimated by comparison of $A_{492nm}/(A_{414nm}+A_{492nm})$ with lines of best fit obtained from analysis of 0-100 *µ*M zinc acetate solutions. Protein concentrations were determined by UV spectrophotometry, with extinction coefficients and molecular weights calculated by ProtParam (http://web.expasy.org/protparam/).

Protein crystallisation and X-ray structure solution

Protein crystals were obtained through hanging drop vapour diffusion in which CtIP-nNTD at 4.7 mg/ml was mixed with equal volumes of reservoir solution (200 mM lithium sulphate, 100 mM sodium acetate pH 3.6, 32% (v/v) PEG 400) and equilibrated for 7-10 days. Suitable crystals were incubated in cryoprotectant (20 mM Tris pH 8.0, 150 mM sodium chloride, 200 mM lithium sulphate, 100 mM sodium acetate pH 3.6, 32% (v/v) PEG 400, 20% (v/v) glycerol) and frozen in liquid nitrogen. X-ray diffraction data were collected at 1.03970 Å wavelength as 1000 contiguous frames with 0.2 second exposure and 0.2 degrees of oscillation on a Pilatus 6M detector at beamline PROXIMA1 of the Soleil synchrotron in Paris, France. The data were indexed, integrated and scaled in XDS[41], and merged using Aimless (CCP4 software package)[42]. The data could be indexed equally well in a monoclinic or orthorhombic lattice, and the crystal was initially assigned to orthorhombic space group $P2_12_12_1$, with cell dimensions: a=36 Å, b=38 Å, c= 98 Å and four predicted CtIP molecules in the asymmetric unit. The structure was solved with the molecular replacement method in Phaser[43], using as search model a homology-model helical template spanning amino acids 18-52 of human CtIP, built by fold-recognition server PHYRE[44]. Phaser placed four copies of the helical template, arranged as an antiparallel dimer of two dimeric parallel coiled-coils. The tetrameric assembly was refined to 2.1 Å resolution with Phenix[45] and Buster[46], using Coot[47] for manual model building. Reanalysis of the data processing during refinement highlighted the fact that the merging statistics were better in the case of monoclinic symmetry ($P2_12_12_1$: R_{merge}=0.180, R_{meas}=0.195, R_{pim}=0.073; $P2_1$: R_{merge}=0.056, R_{meas} =0.065; R_{pim}=0.034), and therefore refinement was continued in space group $P2_1$, with two copies of the tetramer in the asymmetric unit. The final crystallographic model was refined to R and R_{free} values of 0.237 and 0.251 respectively, at a resolution of 1.9 Å, with all residues of the asymmetric unit in the favoured region of the Ramachandran plot.

Cell lines, culture conditions, drugs and transfections

Human U2OS and HEK-293T cells were grown in DMEM supplemented with 10% FBS, penicillin, streptomycin and glutamine. U2OS-TLR cells were selected in the presence of 2 µg/ml puromycin. U2OS-TLR cells stably transfected with FLAG-tagged versions of CtIP were grown in the presence of 0.5 µg/ml puromycin and 0.5 mg/ml geneticin. U2OS-MMEJ cells were described previously[21]. U2OS-MMEJ cells stably transfected with FLAG-tagged versions of CtIP were grown in the presence of 150 µg/ml hygromycin and 0.25 mg/ml geneticin. U2OS cells stably transfected with GFP-tagged versions of CtIP were grown in the presence of 0.5 mg/ml geneticin. Camptothecin, hydroxyurea and mytomycin C were from

Sigma. Plasmid transfections were performed with Fugene HD (Promega) following manufacturer's conditions. siRNA transfections were performed with Lipofectamine RNAiMax (Life Technologies) using manufacturer's conditions.

Plasmids, siRNAs and antibodies

3xFLAG-CtIP vector was generated by replacing *eGFP* of GFP-CtIP[11] with the 3xFLAG linker at the NheI and SalI sites. L27E was introduced in 3xFLAG-CtIP or GFP-CtIP by site-directed mutagenesis with Pfu Polymerase AD (Agilent Technologies). Δ1-140 3xFLAG-CtIP was generated by PCR with the 3xFLAG-CtIP plasmid as template. Antibodies for western blotting were: anti-FLAG M2 (Sigma F3165, 1:1000 dilution), anti-GFP 7.1/13.1 (Roche 11814460001, 1:1000), anti-RAD50 13B3 (Abcam ab89, 1:1000; see manufacturer's web page for validation details), anti-MRE11 (Abcam ab33125, 1:10000; see manufacturer's web page for validation details), anti-NBS1 1D7 (Abcam ab7860, 1:1000; see manufacturer's web page for validation details). siRNA sequences: Control siRNA (Luciferase 5'-CGUACGCGGAAUACUUCGA-3'), CtIP siRNA (5'-GCUAAAACAGGAACGAAUC-3'). All CtIP constructs are siRNA-resistant[11].

Gel filtration

Nuclear extracts from HEK-293T cells transiently transfected with plasmids expressing wild type or L27E versions of FLAG-CtIP were prepared as described previously[48]. NaCl concentration in extracts was adjusted to 300 mM and FLAG-tagged proteins were isolated by EZview Red anti-FLAG M2 Affinity gel (Sigma). Anti-FLAG beads were washed three times with lysis buffer containing 300 mM NaCl. Following elution with 3xFLAG peptide (Sigma) proteins were fractionated on a Superdex 75 HR 10/30 column (GE Healthcare). Relative molecular weights of the FLAG-CtIP protein complexes were calculated based on the elution profiles of the calibration standards (GE Healthcare), which were as follows: thyroglobulin (669 kDa), ferritin (440 kDa), aldolase (158 kDa), conalbumin (75 kDa), ovalbumin (44 kDa).

Immunoprecipitations (IPs)

For CtIP dimer analyses, U2OS-TLR cells expressing versions of 3xFLAG-CtIP were transfected with a wild-type form of GFP-CtIP. 48 h after transfection, cells were scraped off

15 cm plates, pelleted and resuspended in 0.5 ml of lysis buffer (20 mM Tris-HCl pH 7.5, 40 mM NaCl, 2 mM $MgCl_2$, 10% glycerol, 0.5% NP-40, EDTA-free protease inhibitors (Roche), 10 mM N-ethylmaleimide (Sigma-Aldrich), phosphatase inhibitor cocktail (42 µg/ml NaF, 0.22 mg/ml beta-glycerophosphate, 37 µg/ml Na_3VO_4, 2.2 mg/ml sodium pyrophosphate)) containing 100 U of benzonase (Merck). After a 5 min incubation at room temperature, 50 µl of 5 M NaCl were added to the mix and tubes were left on ice for 15 min. Extracts were pelleted at 15 000 rpm at 4 °C for 45 min. Supernatants were recovered and diluted with 0.5 ml of lysis buffer with no benzonase. 1 mg of protein extract and 10 µl of GFP-Trap (ChromoTek) bead slurry were used per IP. IPs were incubated with rotation for 1.5 h at 4 °C. Beads were washed 5x with lysis buffer, changing tubes after the first wash. Immunoprecipitated proteins were eluted from the beads using 1x Lämmli buffer and boiling for 5 min. Samples were run in a 6% SDS-PAGE gel. For co-immunoprecipitations with known interactors, HEK-293T cells were transiently transfected with wild type or L27E GFP-CtIP, or the plasmid backbone expressing GFP only as a control. The rest was performed as described above. Samples were analysed in a 6-18% SDS-PAGE gel.

Generation of U2OS-TLR cells

Viral particles containing pCVL Traffic Light Reporter 1.1 (Sce target) Ef1a Puro[23] were generated by transfection of HEK-293T cells together with pMD2.G and psPAX2 (Addgene). Media containing the viral particles were collected 48 h after transfection, filtered and used to replace the medium of U2OS cells seeded the previous day. 48 h after media exchange, 2 µg/ml puromycin was added to select for integration events. Several clones were picked 10-15 days after starting selection and tested by transfection of pRRL sEF1a HA.NLS.Sce(opt).T2A.IFP (I-SceI nuclease plasmid) and pRRL SFFV d20GFP.T2A.mTagBFP Donor[23]. One of the clones displaying no background mCherry fluorescence in non-transfected cells and non-coincidental eGFP and mCherry fluorescence in cells transfected with both IFP-nuclease and BFP-Donor (indicative of a single integration of the TLR cassette) was selected for further experimentation.

TLR assay

U2OS-TLR cells were stably transfected with empty vector or forms of 3xFLAG-CtIP. Clones with similar expression levels of 3xFLAG-CtIP were selected (**Fig. 4c**). 350 000 cells were seeded in 6-cm dishes and transfected with control siRNA (siLuciferase) or siRNA targeting

CtIP. 6 h after transfection, plates were washed and cells were transfected with 2 µg of BFP-Donor and 3 µg of IFP-nuclease. 72 h after siRNA transfection, cells were collected by trypsinization, resuspended in 300 µl of 1x PBS + 1 mg/ml BSA and immediately analyzed in an LSR-Fortessa Cell Analyzer (BD Biosciences) (**Supplementary Fig. 4b**). Alternatively, cells were fixed for 20 min at room temperature in 2% paraformaldehyde in 1x PBS, washed, resuspended in 300 µl of 1x PBS + 1 mg/ml BSA and kept at 4 °C for posterior analysis.

Flow cytometry analysis was performed by exciting BFP with a 405 nm laser and emissions collected via a 450/50 filter. IFP was excited with a 635 nm laser and emissions collected via a 670/30 filter. eGFP was excited with a 488 nm laser and emissions collected via a 530/30 filter. mCherry was excited with a 561 nm laser and emissions collected via a 610/20 filter. For a representative example of the gating strategy, see **Supplementary Fig. 4**. Background signal was subtracted for each sample and the intensity of the BFP signal was used to normalise the percentages of eGFP-positive cells, as the efficiency of homologous recombination increases with increasing amounts of transfected BFP-Donor[23]. Data analysis was performed using FlowJo X (Tree Star).

MMEJ assay

U2OS-MMEJ cells[21] were stably transfected with empty vector or forms of 3xFLAG-CtIP. Clones with similar expression levels of 3xFLAG-CtIP were selected (**Supplementary Fig. 5a**). 350 000 cells were seeded in 6-cm dishes and transfected with control siRNA (siLuciferase) or siRNA targeting CtIP. 6 h after transfection, plates were washed and cells were transfected with 3 µg of pCBASce, a vector that expresses I-SceI nuclease[49]. 72 h after siRNA transfection, cells were collected by trypsinization, resuspended in 300 µl of 1x PBS + 1 mg/ml BSA and immediately analyzed in an LSR-Fortessa Cell Analyzer (BD Biosciences). Flow cytometry analysis was performed by exciting eGFP with a 488 nm laser and emissions collected via a 530/30 filter. Data analysis was performed using FlowJo X (Tree Star).

Laser micro-irradiation and immunofluorescence

For generation of localized damage in cellular DNA by exposure to a UV-A laser beam[50,51], cells were plated on glass-bottomed dishes (Willco-Wells) and presensitized with 10 μM BrdU (Sigma-Aldrich) in phenol red-free medium (Invitrogen) for 24 h at 37°C. Laser micro-irradiation was performed by using a FluoView 1000 confocal microscope (Olympus) equipped with a 37°C heating stage (Ibidi) and a 405-nm laser diode (6 mW) focused

through a 603 UPlanSApo/1.35 oil objective to yield a spot size of 0.5–1 mm. The time of cell exposure to the laser beam was ~250 msec (fast scanning mode). Laser settings (0.4-mW output, 50 scans) were chosen to generate a detectable damage response restricted to the laser path in a presensitization-dependent manner without noticeable cytotoxicity. Imaging was done on the same microscope by using the objective and software as described above.

30 min after irradiation, cells were washed three times with PBS followed by fixation with 2% paraformaldehyde (w/v) in PBS for 10 min. Following three washes with PBS, Cells were permeabilised in 0.5% Triton X-100 in PBS for 5 min. Following three washes with PBS, cells were co-stained with primary antibodies against γH2AX (Cell Signaling 2577, 1:500 dilution; see manufacturer's web page for validation details) and cyclin A H432 (Santa Cruz sc-751, 1:500; see manufacturer's web page for validation details) in 5% FBS in PBS for 1 h at RT. Following three washes with PBS, cells were stained with goat anti-rabbit Alexa Fluor 594 secondary antibody (Molecular Probes, 1:1000) in 5% FBS in PBS for 30 min at RT.

Quantification of GFP-CtIP accumulation in damaged cells

U2OS cells stably expressing wild type or L27E mutant GFP-CtIP were left untreated or treated for 24 h with 1 mM hydroxyurea or 120 ng/ml mytomycin C. After treatment cells were trypsinised, collected and resuspended in extraction buffer (25 mM HEPES pH 7.4, 50 mM NaCl, 1 mM EDTA, 3 mM $MgCl_2$, 300 mM sucrose, 0.5% Triton X-100, protease inhibitors) for 5 min on ice. After washing with 1xPBS + 1 mg/ml BSA (PBS-B), cells were fixed for 20 min at room temperature in 2% paraformaldehyde in 1xPBS. After washing with PBS-B, cells were incubated with antibodies against γH2AX (Cell Signalling 2577, 1:200). Antibody detection was by incubating cells with goat anti-rabbit Alexa Fluor 594 (Molecular probes, 1:1000). DNA was stained with 4',6-diamidino-2-phenylindole (DAPI). Samples were analyzed in an LSR-Fortessa Cell Analyzer (BD Biosciences). Data analysis was performed using FlowJo X (Tree Star).

DNA-end resection assay

Sample preparation was as described previously[28]. Briefly, U2OS-TLR cells expressing forms of 3xFLAG-CtIP were transfected with control siRNA (siLuciferase) or siRNA targeting CtIP. 48 h after transfection, cells were left untreated or treated with 1 μM camptothecin for 1 h. After permeabilization and fixation, cells were incubated with antibodies against an RPA subunit (RPA2 9H8; Abcam ab2175 1:500 dilution; see manufacturer's web page for

validation details) and γH2AX (Cell Signalling 2577, 1:200). Antibody detection was by incubating cells with goat anti-mouse Alexa Fluor 488 (Molecular probes, 1:1000) and goat anti-rabbit Alexa Fluor 594 (Molecular probes, 1:1000). DNA was stained with 4',6-diamidino-2-phenylindole (DAPI). Samples were analyzed in an LSR-Fortessa Cell Analyzer (BD Biosciences). Data analysis was performed using FlowJo X (Tree Star).

Electrophoretic mobility shift assay

CtIP constructs for the NTD and CTD were incubated with 1μM fluorescein-labelled 200bp double-stranded DNA at concentrations: 0, 0.5, 1.0, 2.0 , 5.0 and 10μM in 20 mM Tris pH 8.0, 150mM NaCl for 60 minutes in darkness at room temperature. To the 10μl reaction volume, 2μl of gel loading buffer comprising of 250mM Tris pH 8.0, 40% glycerol and bromophenol blue was added and samples were analysed by electrophoresis on a 1.5% (wt/vol) agarose gel (Ultra Pure Invitrogen) in 0.5xTB buffer at 25V for 2.5h at 4°C. DNA was visualised in a Gel Doc-It visualisation system using UV light.

SUPPLEMENTARY FIGURE LEGENDS

Supplementary Figure 1. Sequence alignment of the CtIP N-terminal domain, summary of recombinant protein constructs, SEC-MALS analysis of MBP-CtIP-nNTD and details of the $2F_o$ - F_c electron density map of CtIP-NTD. (**a**) Multiple sequence alignment, including the predicted secondary structure (red rods indicate regions of predicted α-helix) and coiled-coil regions; the probability of coiled-coil formation is indicated by '-' (<50%), 'c' (50-90%) or 'C' (>90%). Protein sequence analysis was performed with the Jalview suite of bioinformatics tools (www.jalview.org). *Hs: Homo sapiens, Mm: Mus musculus, Ec: Equus caballus, Oc: Oryctolagus cuniculus, Gg: Gallus gallus, Xl: Xenopus laevis, Dr: Danio rerio*. (**b**) SDS-PAGE analysis of purified recombinant CtIP protein samples: NTD, NTD-DM (C89A C92A), nNTD, cNTD, cNTD-DM (C89A C92A) and CTD visualised by Coomassie staining. (**c**) MBP-CtIP-nNTD eluted in a single peak corresponding to a molecular weight of 187 kDa; its theoretical tetramer size is 197 kDa. For comparison, free MBP eluted in a single peak of 44.5 kDa, corresponding to its theoretical monomer size of 44.7 kDa. The light scattering (LS) as relative Raleigh ratio and the differential Reflective Index (dRI) are drawn as solid and dashed lines, respectively. The value for the fitted molecular mass (M_r) of MBP-CtIP-nNTD is shown as diamond shapes across its elution peak. The predicted molecular masses of

monomeric MBP-CtIP-nNTD and free MBP are shown in brackets above the respective elution peak. (**d**) Details of the $2F_o$ - F_c electron density map, contoured at 1.2 rmsd, superimposed on the refined crystallographic model of CtIP-NTD. The figure shows the map in the vicinity of amino acid W24. The map is shown as thin mesh in light purple, the protein atoms are shown in ball-and-stick representation and coloured according to chemical identity. The figure was prepared in Coot[47].

Supplementary Figure 2. Circular dichroism (CD) analysis of CtIP N-terminal domain constructs. **(a, c, e)** Far UV CD spectra recorded between 260 and 185 nm, in mean residue ellipticity, MRE ([θ]) (1000x $deg.cm^2.dmol^{-1}.residue^{-1}$), with α-helical content calculated through deconvolution using the CDSSTR algorithm with normalised root-mean-square deviation values (nrmsd) as shown. (**b, d, f**) CD thermal denaturation data recorded between 5 and 95°C, in mean residue ellipticity at 222 nm ($[\theta]_{222}$). (**a**) CD spectra and (**b**) thermal denaturation of CtIP-NTD (solid line) and CtIP-NTD-DM (dashed line). (**c**) CD spectra and (**d**) thermal denaturation of CtIP-cNTD (solid line) and CtIP-cNTD-DM (dashed line). (**e**) CD spectra and (**f**) thermal denaturation of CtIP-nNTD.

Supplementary Figure 3. Microbeam proton induced X-ray emission (microPIXE) analysis of the CtIP N-terminal domain in buffer containing KBr. **(a)** Rutherford backscattering (RBS) spectrum and **(b)** PIXE spectrum with raw data shown as black dots and the modelling fitting as red lines. **(c)** Zinc was detected at a mean atomic ratio to sulphur of 0.095, which given the presence of three methionine and three cysteine residues per chain, corresponds to a CtIP-NTD:Zn^{2+} ratio of 1.8:1.

Supplementary Figure 4. Experimental procedure and gating scheme for the Traffic Light Reporter (TLR) system. (**a**) TLR cells are not fluorescent due to the fact that the TLR cassette contains an in-frame truncated version of the *eGFP* gene and full-length *mCherry* is out of frame. When TLR cells are transfected with an I-SceI nuclease encoded in a plasmid also expressing an infrared fluorescent protein (IFP) and an exogenous donor template containing the missing part of *eGFP* that also expresses a blue fluorescent protein (BFP), the different repair outcomes will give different fluorescent cells. If the double-strand break (DSB) is repaired by gene conversion (homologous recombination), *eGFP* is restored and the cells fluoresce in green. If the DSB is repaired by mutagenic end joining causing a +2 frameshift, *mCherry* gets in frame and its expression yields red fluorescence. The T2A "dis-linker" between *eGFP* and *mCherry* allows the downstream-encoded *mCherry* to escape degradation of the misfolded protein encoded in the +3 reading frame of *eGFP*[23]. (**b**) Experimental outline of the TLR assay. (**c-h**) Gating scheme for a TLR assay. (**c**)

Discrimination of cells over debris using a forward scatter vs. side scatter plot (areas). (**d**) Discrimination of cell singlets over doublets using a forward scatter (area) vs. side scatter (width) plot. (**e**) Selection of events for quantification. Dot plot showing the intensity of BFP (donor; *x* axis) vs. intensity of IFP (nuclease; *y* axis). Cells with negative staining for both BFP and IFP are depicted in (**f**) to show the absence of eGFP or mCherry fluorescence in non-transfected cells, and to establish quantification gates. Cells counted positive for both BFP and IFP (that is, transfected with both the donor and the nuclease) are represented in (**g**), where quantification gates are the same as the ones established in (**f**). At least 10,000 events were analysed in the double-positive population. (**h**) Due to the fact that the amount of donor transfected in the cells affects the outcome of the assay[23], the intensity of the BFP signal in the double-positive cells (geometric mean of BFP signal in 'BFP+ IFP+' cells) was calculated for normalisation purposes (see Methods). (**i**) Mutagenic end-joining (mutEJ) rates in the TLR-CtIP complementation system. Only expression of the wild-type version of FLAG-CtIP reduced the rate of mutEJ to values similar to the control (cells transfected with control siRNA). EV: empty vector. All quantifications are shown as the average of three independent experiments. Error bars are ± SEM.

Supplementary Figure 5. CtIP expression levels in the different stable cell lines generated and FACS quantification of GFP-CtIP accumulation on damaged cells. (**a**) CtIP expression levels in the different U2OS-MMEJ FLAG-CtIP clones. Samples were collected 48 h after siRNA transfection. CtIP antibody was a gift from R. Baer. PARP-1 antibody (Cell Signalling) was used as loading control. (**b**) GFP-CtIP expression levels in the U2OS stable cell lines. RPA34-20 (Merck) antibody was used as loading control. (**c**) A representative complete dataset for quantification of GFP-CtIP accumulation on damaged cells. Dot plots representing, for each dataset, gating and quantification (in the *y* axis) of the GFP-positive cells (top panels) and the γH2AX-positive cells (bottom panels) taking into account their DNA content (*x* axis). Quantification gates were always established using the untreated samples of each dataset. At least 10,000 events were analysed per sample.

Supplementary Figure 6. A representative complete dataset for the DNA-end resection assay. Dot plots representing, for each dataset, gating and quantification (in the *y* axis) of the RPA-positive cells (top panels) and the γH2AX-positive cells (bottom panels) taking into account their DNA content (*x* axis). Quantification gates were always established using the untreated samples of each dataset and were defined as to avoid quantification of RPA chromatinisation due to DNA replication. Note that as camptothecin (CPT) only causes DNA

double-strand breaks in S-phase cells, both RPA- and γH2AX-positive cells only appear in that cell cycle stage. At least 10,000 events were analysed per sample.

Supplementary Figure 7. (**a**) SEC-MALS analysis of F20E CtIP-NTD. F20E CtIP-NTD forms two distinct species of 62.3 kDa and 33.2 kDa, corresponding to its theoretical tetramer and dimer sizes of 63.2 kDa and 31.6 kDa, respectively. The light scattering (LS) as relative Raleigh ratio and the differential Reflective Index (dRI) are drawn as solid and dashed lines, respectively. The values for the fitted molecular masses (M_r) of dimeric and tetrameric F20E CtIP-NTD are shown as diamond shapes across the respective elution peaks. The predicted molecular mass of monomeric F20E CtIP-nNTD is shown in brackets above the elution peaks. (**b**) Electrophoretic mobility shift assay measuring the ability of CtIP NTD and CTD regions to interact with linear, double-stranded DNA. In the experiment, increasing amounts (0, 0.5, 1.0, 2.0, 5.0 and 10μM) of the CtIP protein were incubated with a 1μM sample of 5' fluorescein-labelled 200bp double-stranded DNA. After incubation, each sample was resolved by electrophoresis on agarose gel and visualised under UV light.

Supplementary Figure 8. Uncropped images for all western blot data. (**a**) Co-immunoprecipitations shown in Figure 4c. (**b**) Gel filtration experiment shown in Figure 4d. (**c**) Co-immunoprecipitations shown in Figure 6c.

Table 1 Data collection and refinement statistics

	CtIP-NTD (amino acids 18-52)
Data collection	
Space group	$P2_1$
Cell dimensions	
a, *b*, *c* (Å)	35.97, 37.91, 97.79
a, b, g (°)	90.00, 90.02, 90.00
Resolution (Å)	28.98 - 1.90 (1.94 - 1.90)
R_{merge}	0.063 (0.919)
R_{meas}	0.074 (1.074)
R_{pim}	0.038 (0.547)
I / sI	10.2 (1.6)
CC(1/2)	0.998 (0.770)
Completeness (%)	99.5 (99.5)
Multiplicity	3.6 (3.6)
Refinement	
Resolution (Å)	28.97 - 1.90
No. reflections	20969
R_{xpct}[a] / R_{free}[b]	0.212/0.251
No. atoms	2474
Protein	2267
Ligand/ion	70
Water	137
B-factors	49.2
Protein	47.4
Ligand/ion	103.4
Water	52.5
R.m.s. deviations	
Bond lengths (Å)	0.012
Bond angles (°)	1.34

Values in parentheses are for highest-resolution shell.

[a] $R_{xpct} = \frac{\sum_{hkl} || F_{obs} | - | F_{xpct} ||}{\sum_{hkl} | F_{obs} |}$, where F_{obs} and F_{xpct} are the observed structure-factor amplitude and the expectation of the model structure factor amplitude, respectively[46].

[b] R_{free} equals R_{xpct} of the test set (5% of the data that were excluded from refinement).

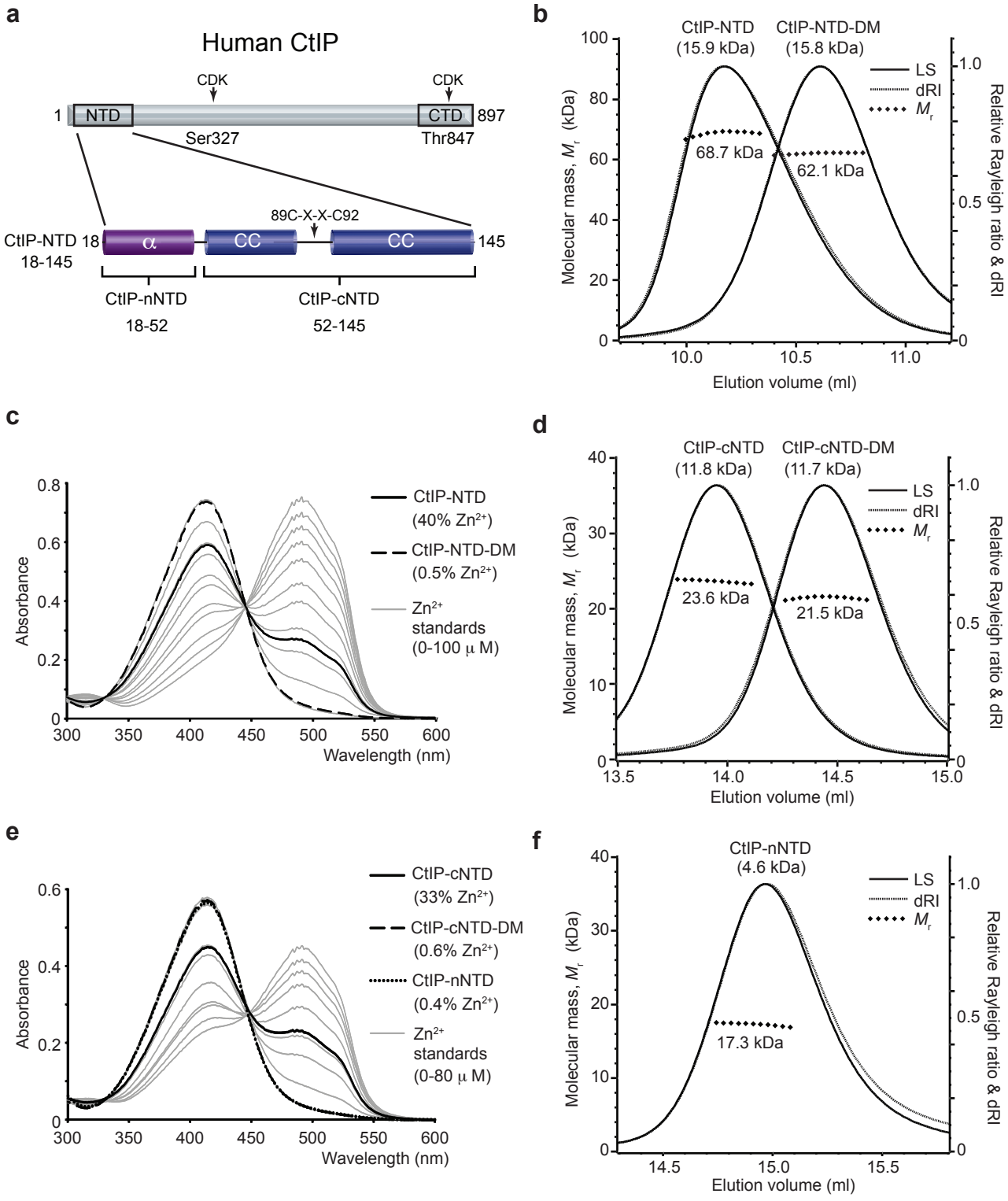

a
Human CtIP
CDK
CDK
1
NTD
CTD
897
Ser327
Thr847
89C-X-X-C92
CtIP-NTD 18
18-145
α
CC
CC
145
CtIP-nNTD
18-52
CtIP-cNTD
52-145
b
CtIP-NTD
(15.9 kDa)
CtIP-NTD-DM
(15.8 kDa)
LS
dRI
M_r
68.7 kDa
62.1 kDa
Molecular mass, M_r (kDa)
Relative Rayleigh ratio & dRI
Elution volume (ml)
c
CtIP-NTD
(40% Zn^{2+})
CtIP-NTD-DM
(0.5% Zn^{2+})
Zn^{2+}
standards
(0-100 μ M)
Absorbance
Wavelength (nm)
d
CtIP-cNTD
(11.8 kDa)
CtIP-cNTD-DM
(11.7 kDa)
LS
dRI
M_r
23.6 kDa
21.5 kDa
Molecular mass, M_r (kDa)
Relative Rayleigh ratio & dRI
Elution volume (ml)
e
CtIP-cNTD
(33% Zn^{2+})
CtIP-cNTD-DM
(0.6% Zn^{2+})
CtIP-nNTD
(0.4% Zn^{2+})
Zn^{2+}
standards
(0-80 μ M)
Absorbance
Wavelength (nm)
f
CtIP-nNTD
(4.6 kDa)
LS
dRI
M_r
17.3 kDa
Molecular mass, M_r (kDa)
Relative Rayleigh ratio & dRI
Elution volume (ml)

Figure 1

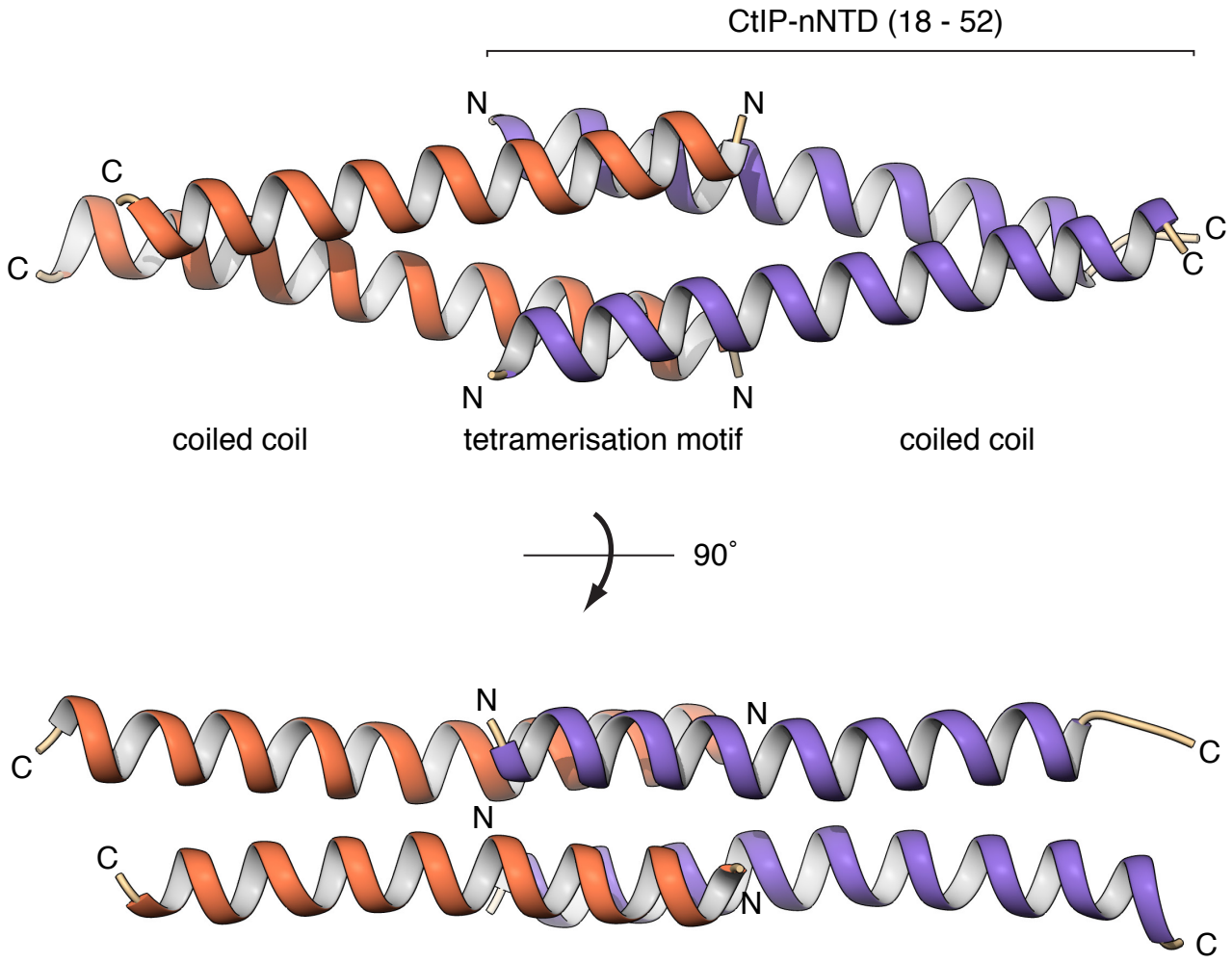

CtIP-nNTD (18 - 52)
N
N
C
C
C
N
N
coiled coil
tetramerisation motif
coiled coil
90˚
N
N
C
C
N
N
C
C

Figure 2

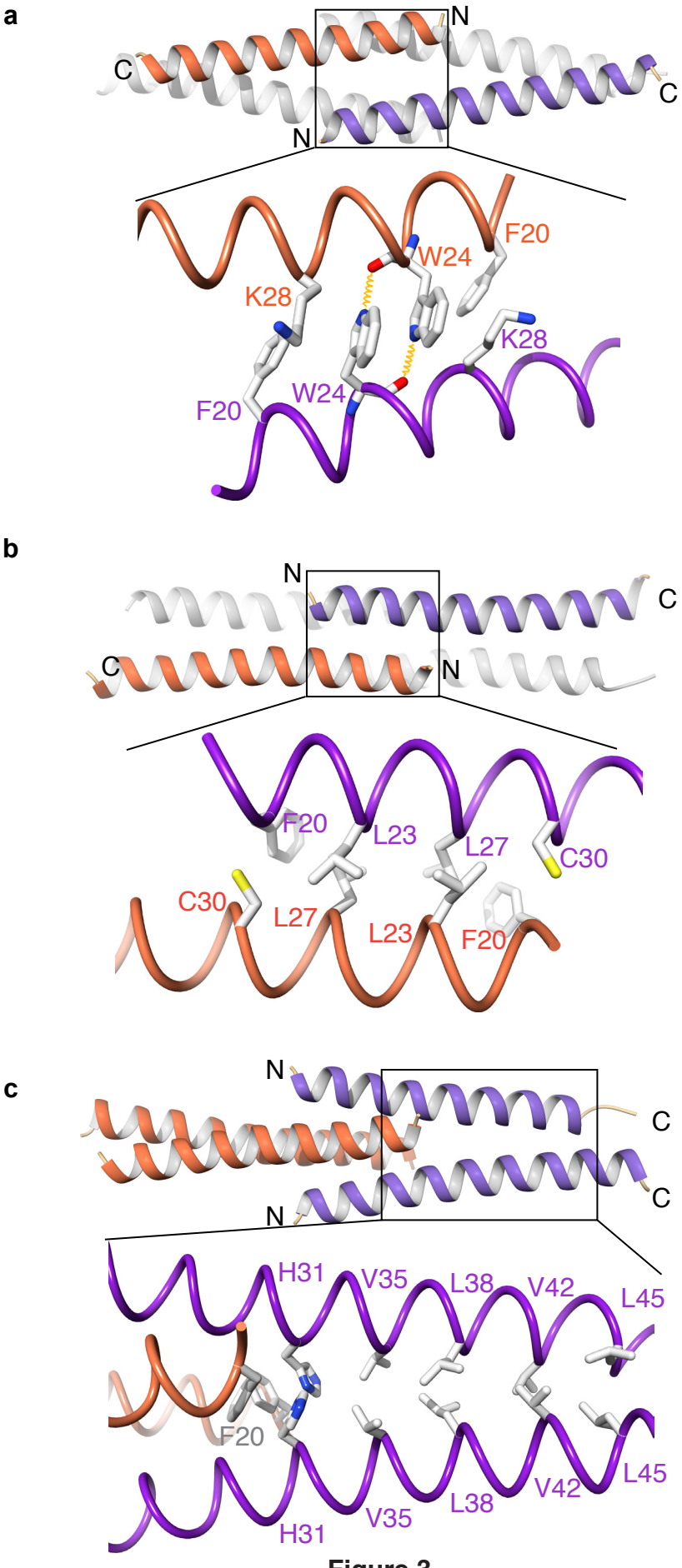

a
N
C
C
N
F20
W24
K28
K28
W24
F20
b
N
C
C
N
F20
L23
L27
C30
C30
L27
L23
F20
c
N
C
C
N
H31
V35
L38
V42
L45
F20
L45
V42
L38
V35
H31

Figure 3

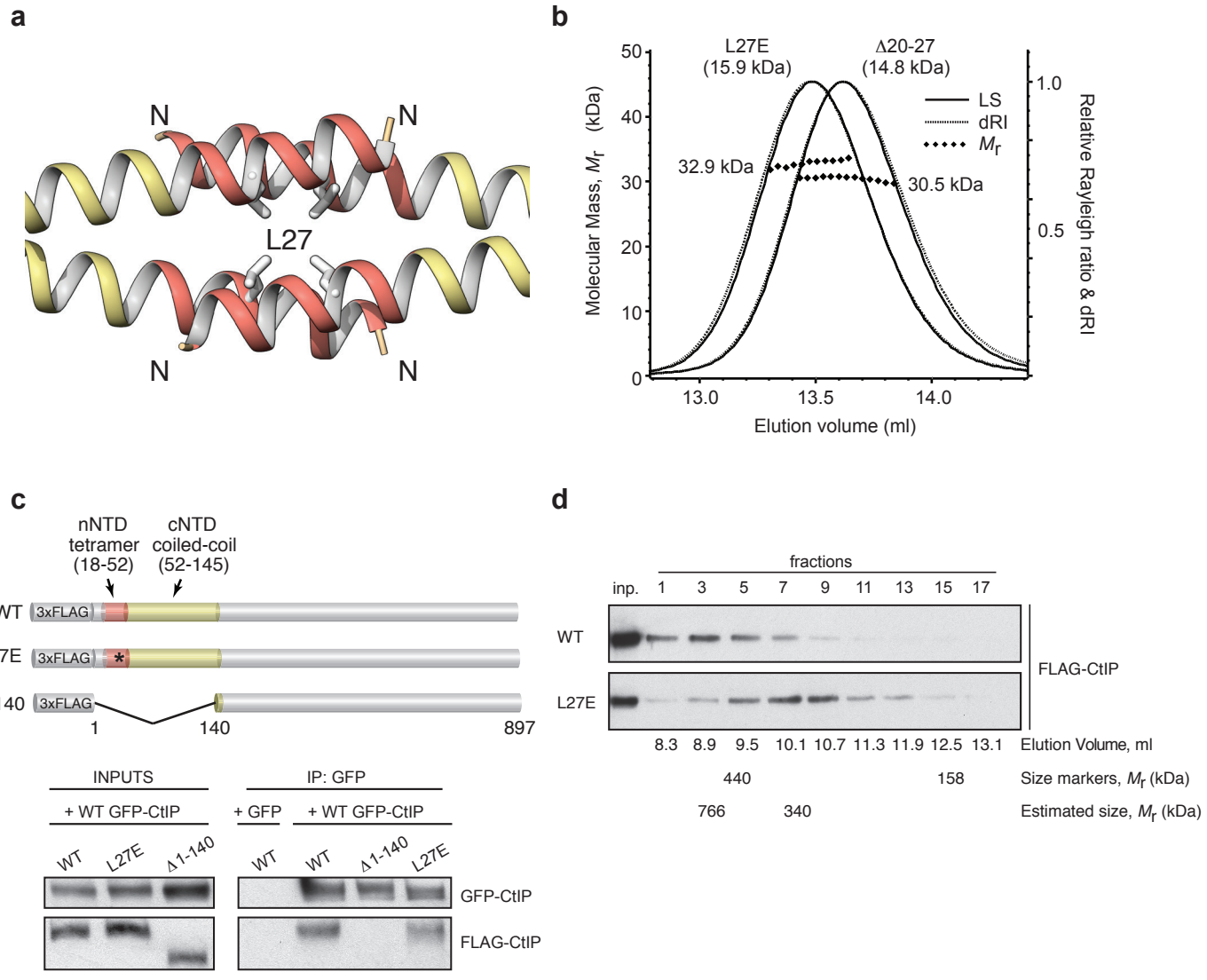
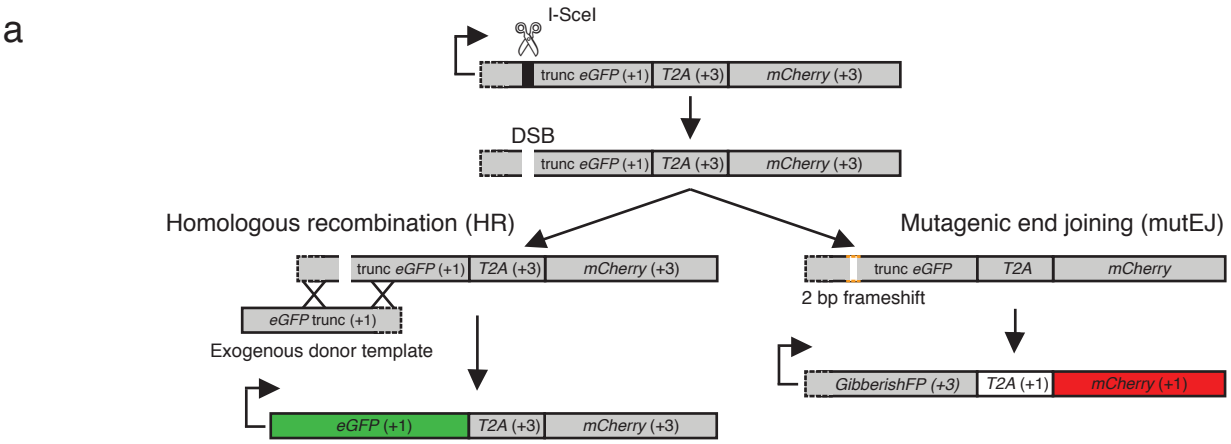
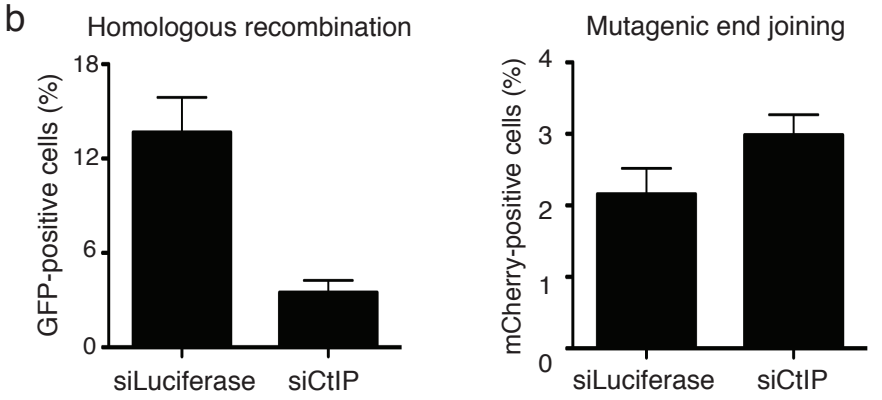
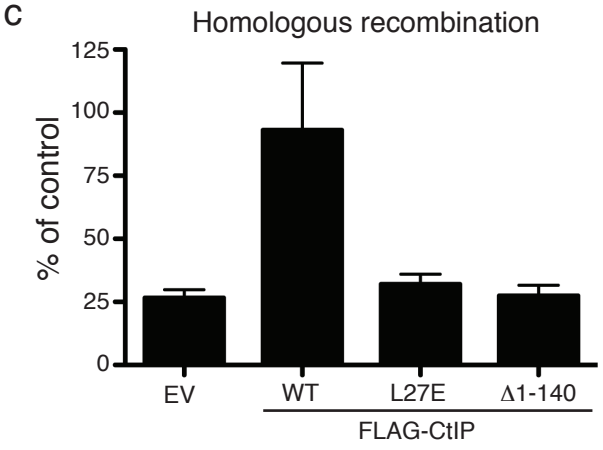
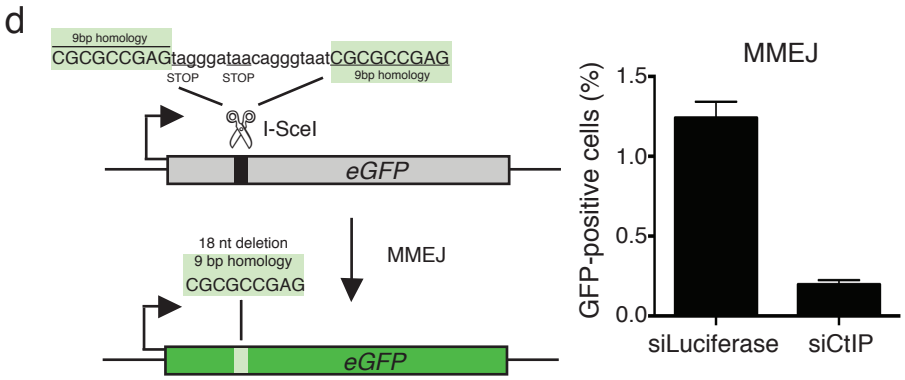

a
N
N
L27
N
N
b
L27E
(15.9 kDa)
Δ20-27
(14.8 kDa)
LS
dRI
M_r
32.9 kDa
30.5 kDa
Molecular Mass, M_r (kDa)
Relative Rayleigh ratio & dRI
Elution volume (ml)
c
nNTD
tetramer
(18-52)
cNTD
coiled-coil
(52-145)
WT
L27E
Δ1-140
3xFLAG
1
140
897
INPUTS
IP: GFP
+ WT GFP-CtIP
+ GFP
+ WT GFP-CtIP
WT
L27E
Δ1-140
GFP-CtIP
FLAG-CtIP
d
fractions
inp.
WT
L27E
FLAG-CtIP
Elution Volume, ml
Size markers, M_r (kDa)
Estimated size, M_r (kDa)

Figure 4

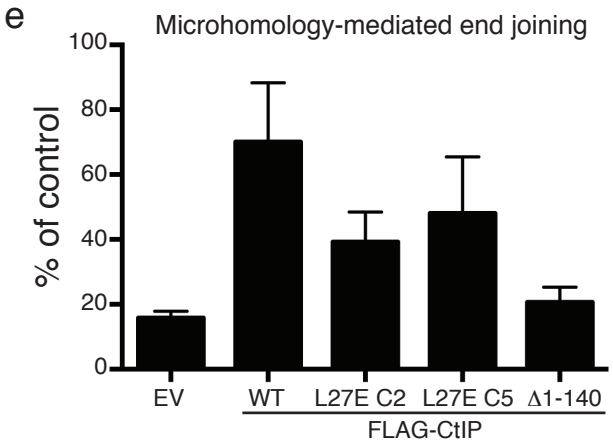

a
I-SceI
trunc eGFP (+1)
T2A (+3)
mCherry (+3)
DSB
Homologous recombination (HR)
Mutagenic end joining (mutEJ)
eGFP trunc (+1)
Exogenous donor template
eGFP (+1)
trunc eGFP
T2A
mCherry
2 bp frameshift
GibberishFP (+3)
T2A (+1)
mCherry (+1)

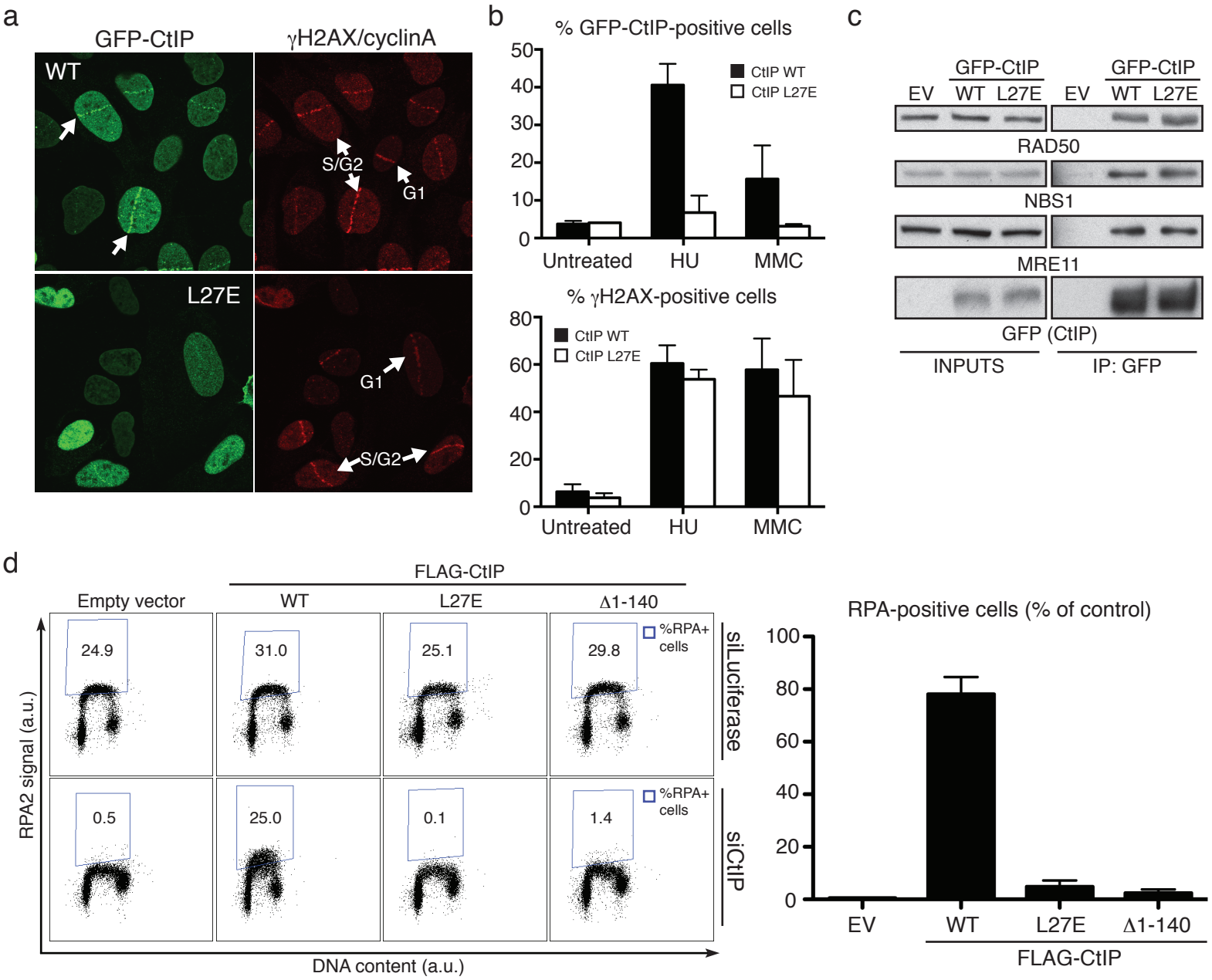

b
Homologous recombination
GFP-positive cells (%)
18
12
6
0
siLuciferase
siCtIP
Mutagenic end joining
mCherry-positive cells (%)
4
3
2
1
0
siLuciferase
siCtIP


c
Homologous recombination
% of control
125
100
75
50
25
0
EV
WT
L27E
Δ1-140
FLAG-CtIP


d
9bp homology
CGCGCCGAGtagggataacagggtaatCGCGCCGAG
9bp homology
STOP
STOP
I-SceI
eGFP
18 nt deletion
9 bp homology
CGCGCCGAG
MMEJ
eGFP

MMEJ
GFP-positive cells (%)
1.5
1.0
0.5
0.0
siLuciferase
siCtIP


e
Microhomology-mediated end joining
% of control
100
80
60
40
20
0
EV
WT
L27E C2
L27E C5
Δ1-140
FLAG-CtIP


a
GFP-CtIP
γH2AX/cyclinA
WT
L27E
S/G2
G1
G1
S/G2
b
% GFP-CtIP-positive cells
CtIP WT
CtIP L27E
Untreated
HU
MMC
% γH2AX-positive cells
CtIP WT
CtIP L27E
Untreated
HU
MMC
c
GFP-CtIP
GFP-CtIP
EV
WT
L27E
EV
WT
L27E
RAD50
NBS1
MRE11
GFP (CtIP)
INPUTS
IP: GFP
d
FLAG-CtIP
Empty vector
WT
L27E
Δ1-140
24.9
31.0
25.1
29.8
%RPA+ cells
siLuciferase
0.5
25.0
0.1
1.4
%RPA+ cells
siCtIP
RPA2 signal (a.u.)
DNA content (a.u.)
RPA-positive cells (% of control)
EV
WT
L27E
Δ1-140
FLAG-CtIP

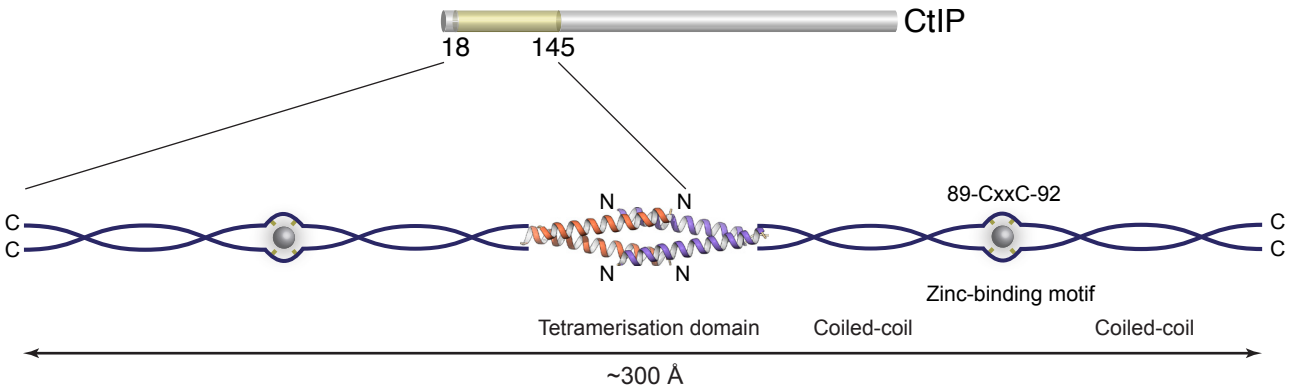
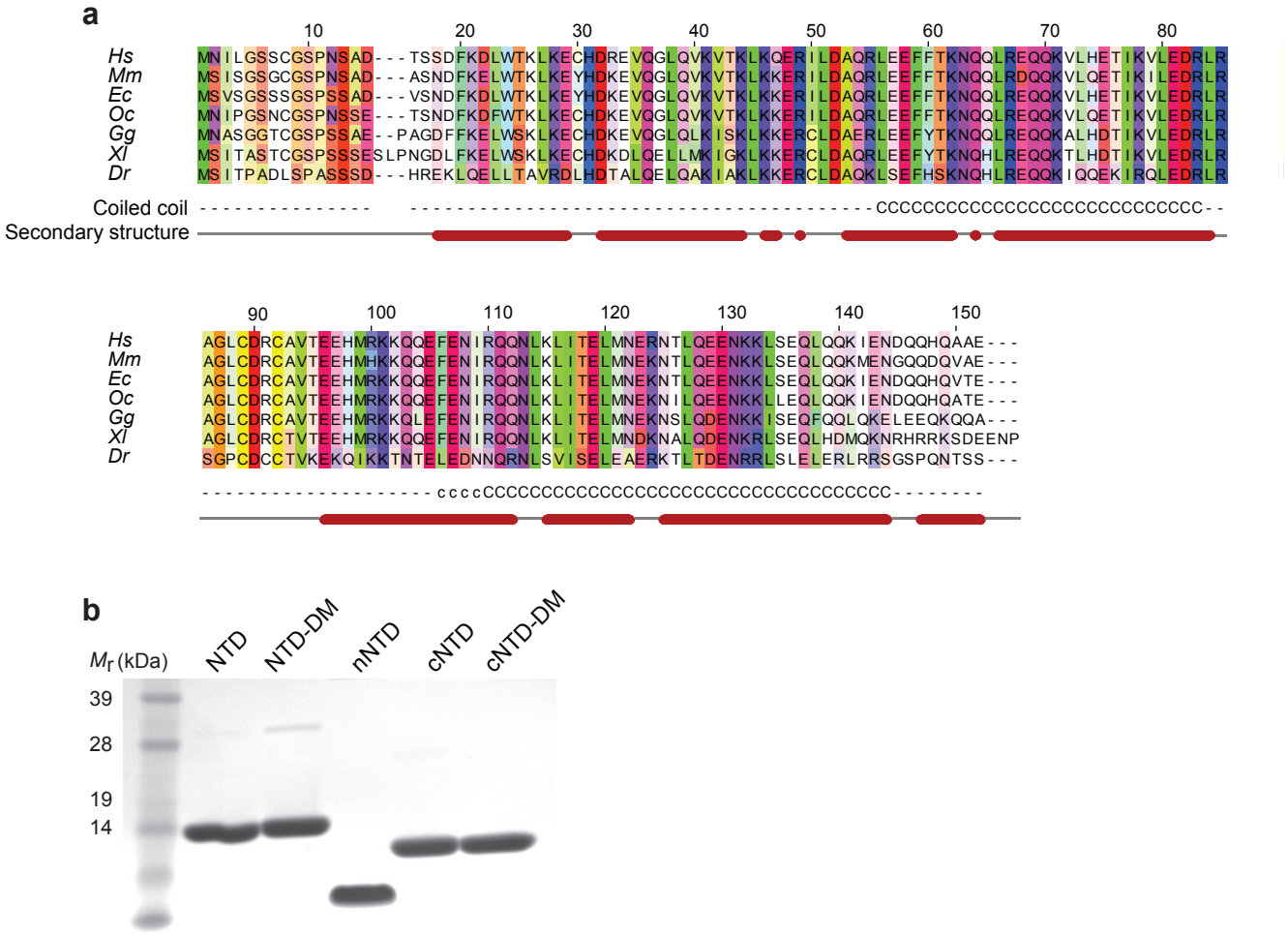
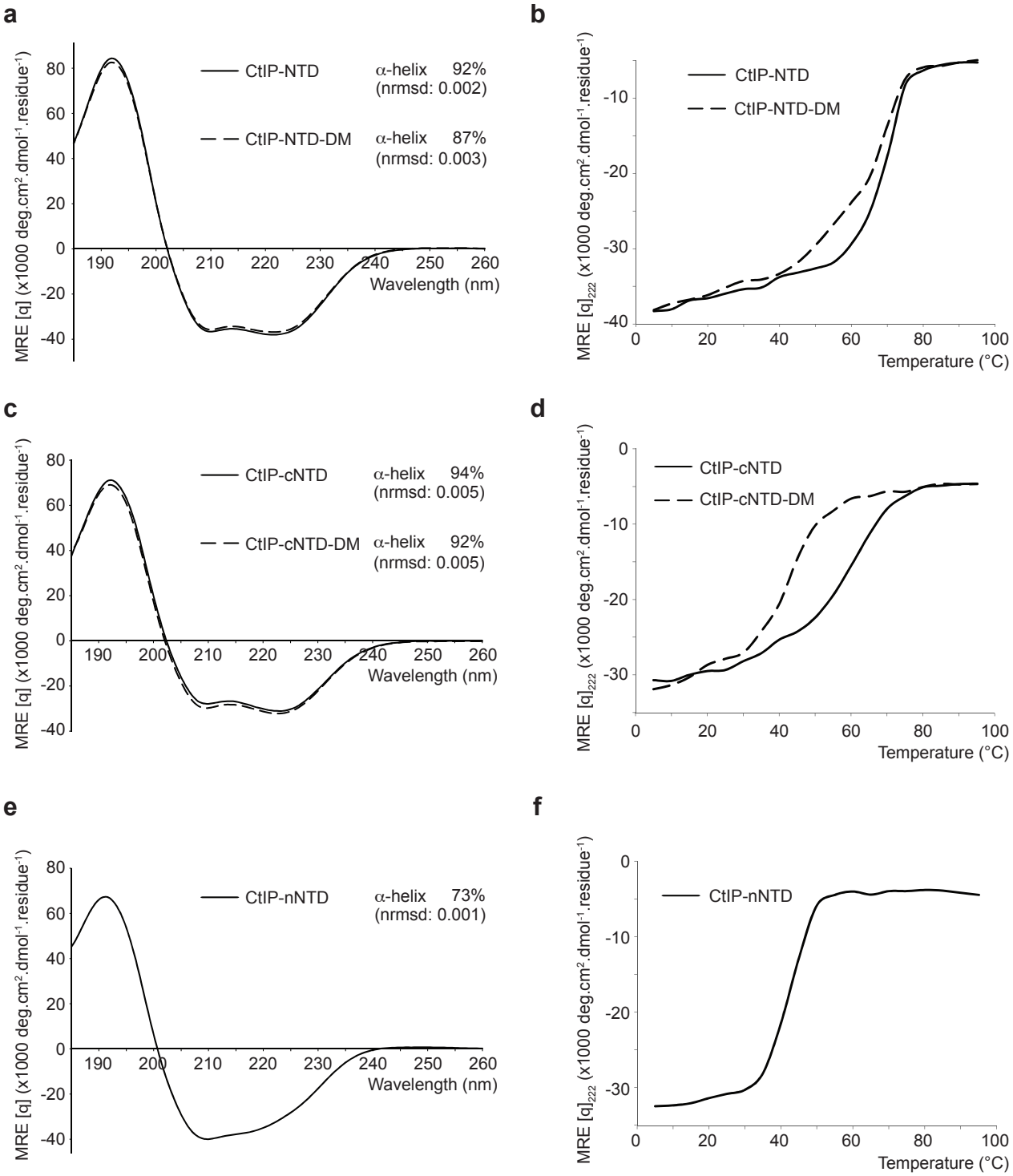

CtIP
18
145
C
C
N
N
N
N
89-CxxC-92
C
C
Zinc-binding motif
Tetramerisation domain
Coiled-coil
Coiled-coil
~300 Å

Figure 7

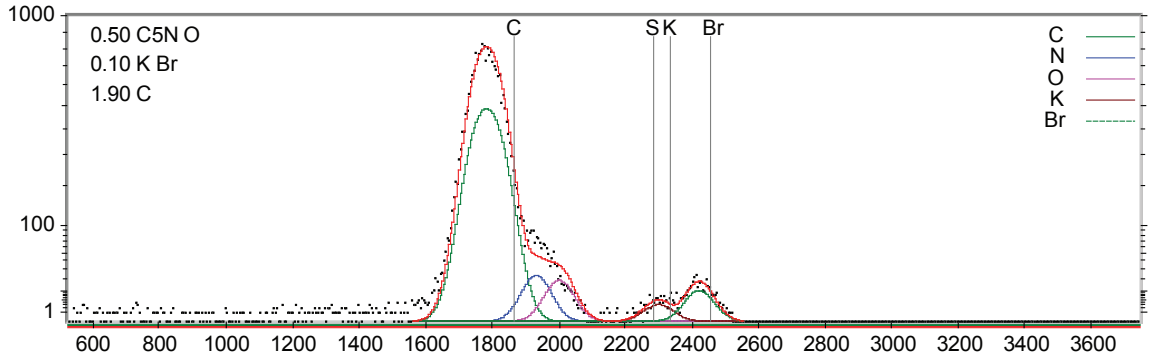


Supplementary Figure 1

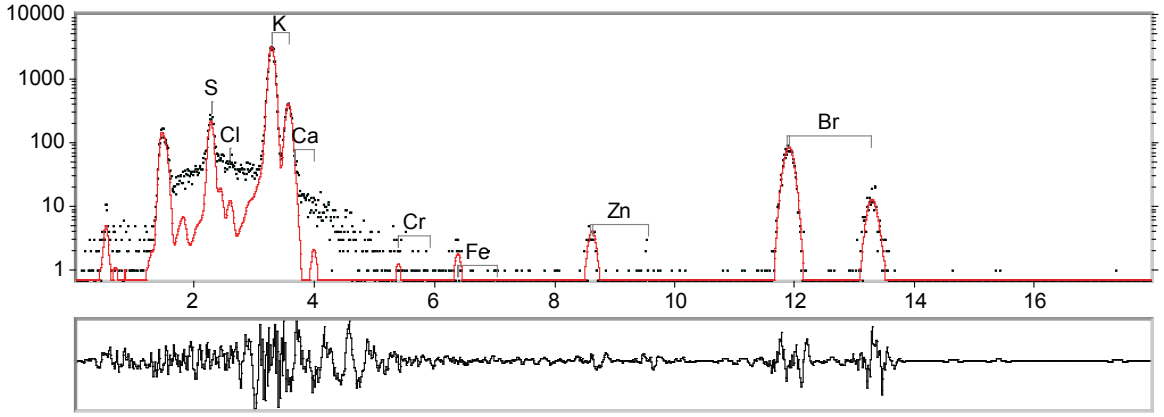


Supplementary Figure 2

a



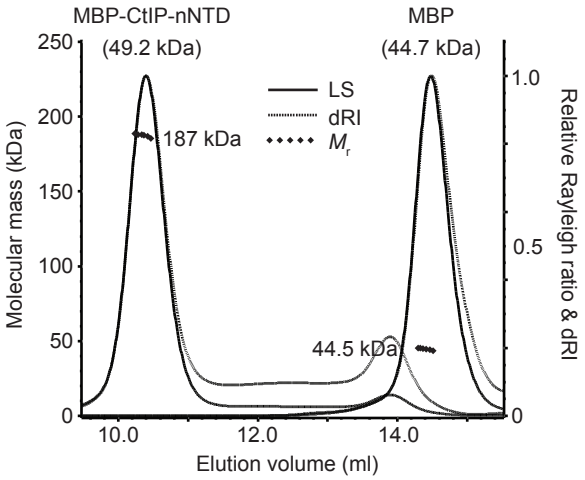
b



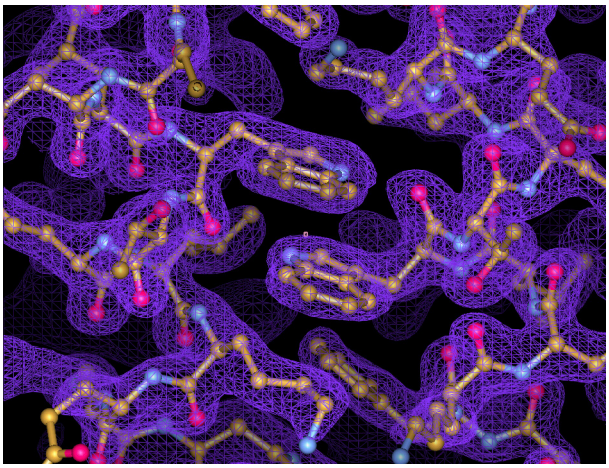
c

Element	Mass ratio to S (mean)	Atomic ratio to Sulphur	Mean MDL (at/S)	Std Dev (rel %)	Mean fitting error%
Al	0.243	0.288	2.266	173.2%	108.610
Si	0.013	0.015	0.134		346.470
P	0.005	0.006	0.086	173.2%	22.027
S	**1.000**	**1.000**	**0.028**	**0.0%**	**3.513**
Cl	0.042	0.038	0.025	58.1%	35.947
K	9.601	7.857	0.028	57.5%	1.063
Ca	0.058	0.047	0.038	37.1%	28.843
Ti	0.006	0.004	0.010	109.9%	94.137
Cr	0.061	0.038	0.027	134.4%	59.287
Mn	0.000	0.000	0.010	173.2%	62.323
Fe	0.033	0.019	0.010	46.1%	34.460
Co	0.028	0.015	0.045	173.2%	56.577
Ni	0.000	0.000	0.039		0.000
Cu	0.017	0.009	0.014	87.1%	40.747
Zn	**0.195**	**0.095**	**0.024**	**25.2%**	**21.780**
Br	17.787	7.124	0.133	47.8%	4.267

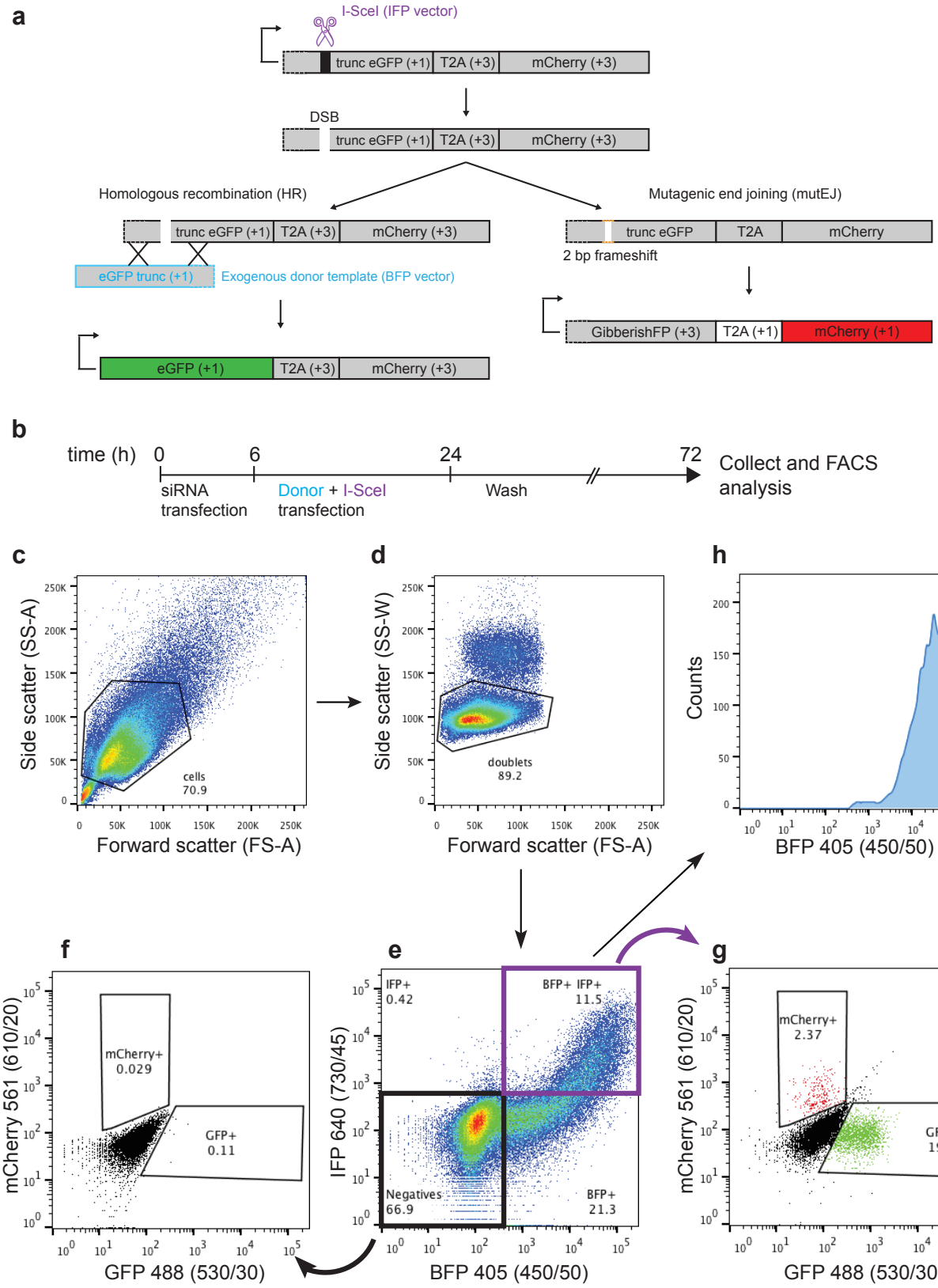
Supplementary Figure 3



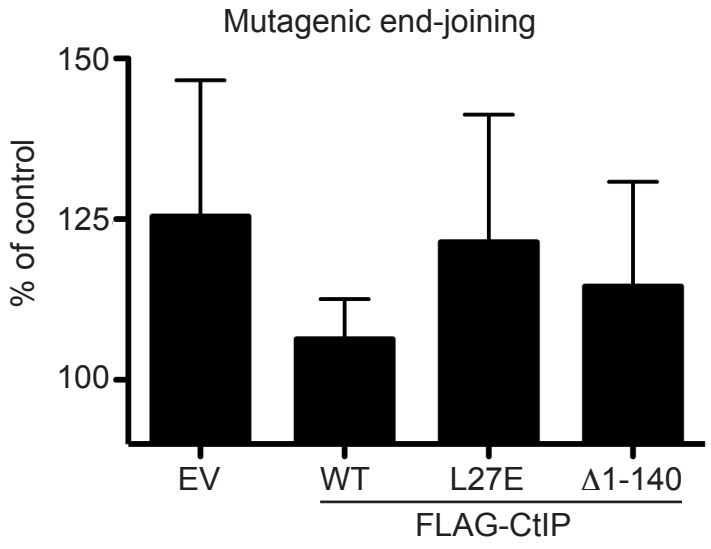
Supplementary Figure 4



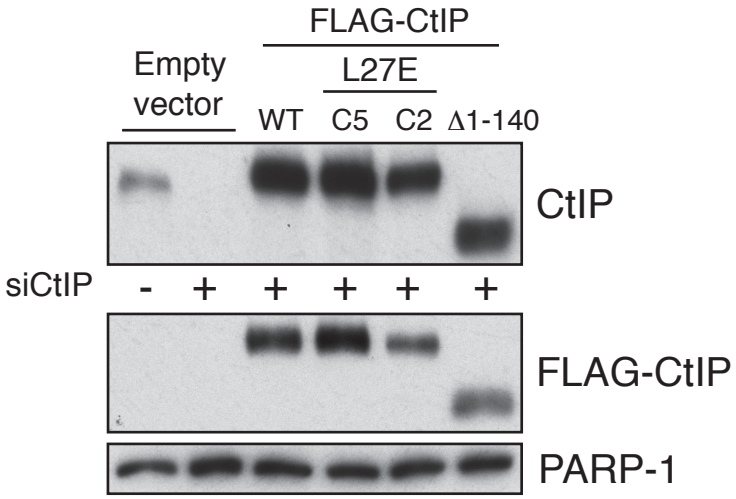
Supplementary Figure 5



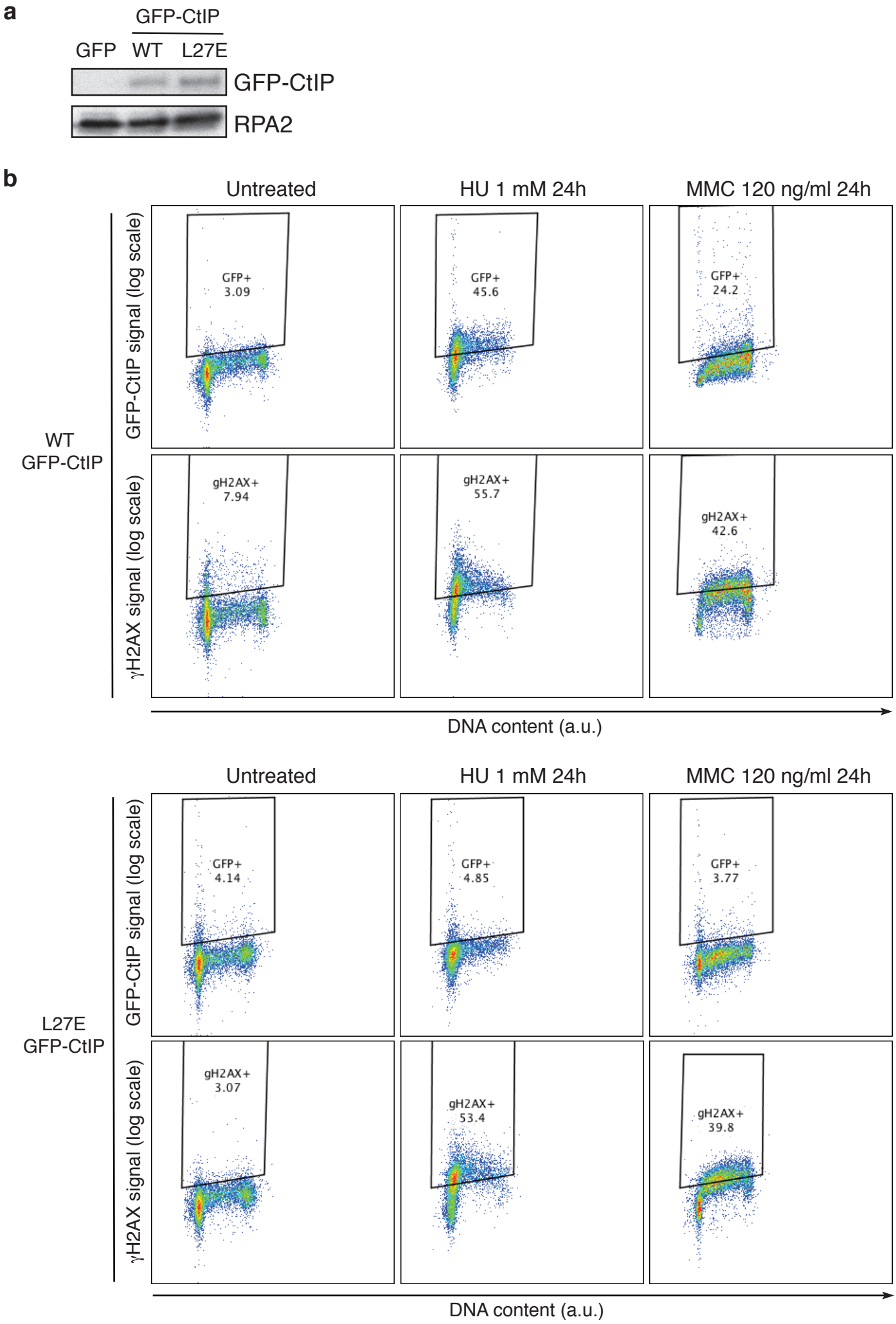
Supplementary Figure 6



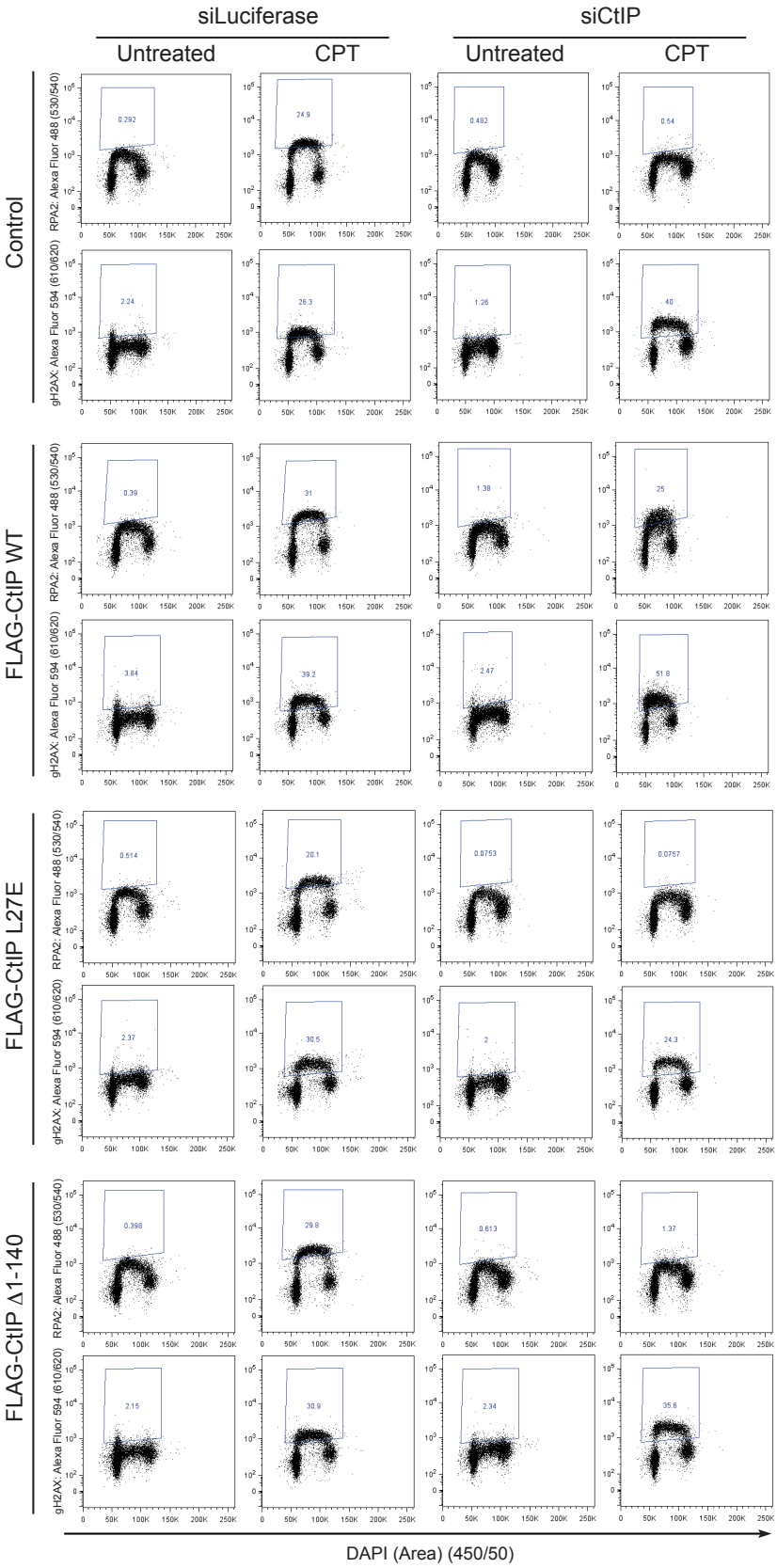
Supplementary Figure 7



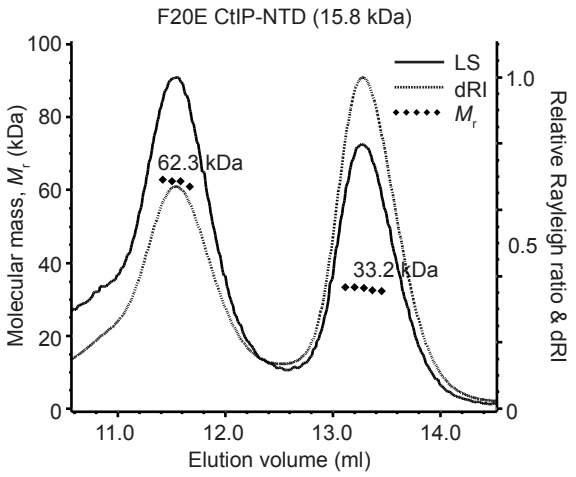
Supplementary Figure 8



Supplementary Figure 9

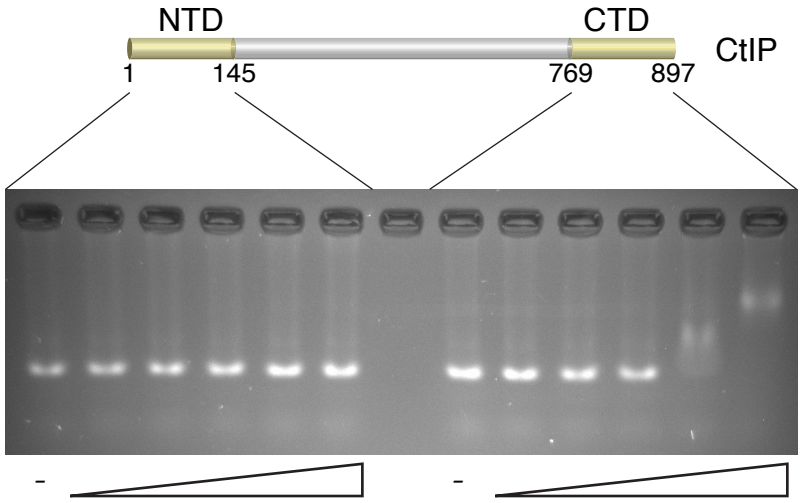


Supplementary Figure 10

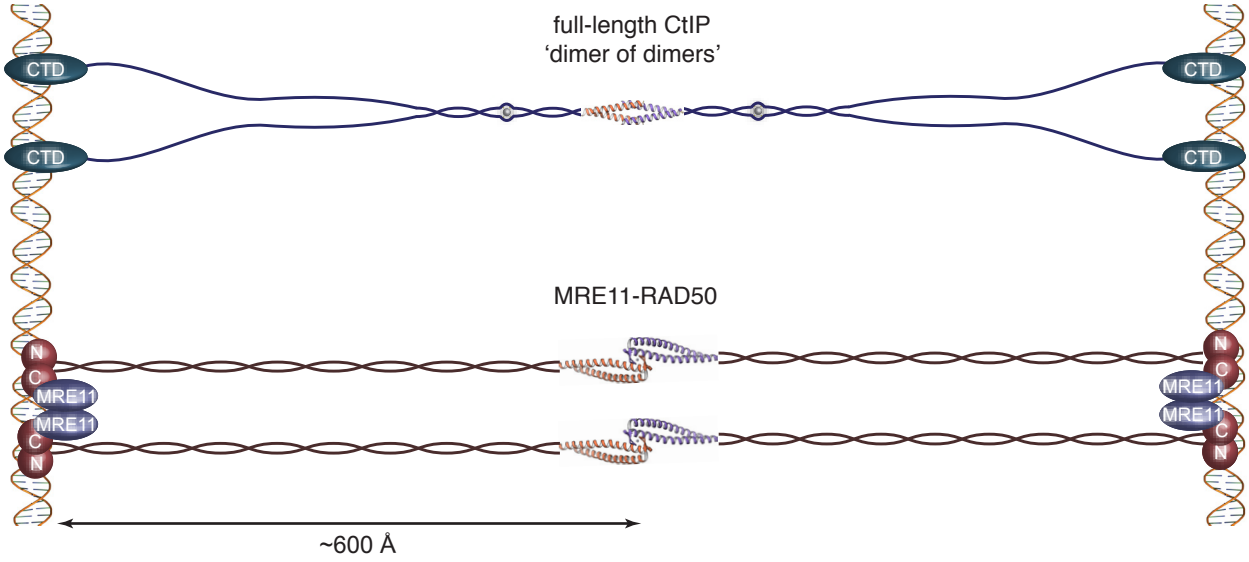


Supplementary Figure 11

a



b



Supplementary Figure 12