Supplementary Material

Structural basis for autoinhibition and phosphorylation-dependent activation of c-Cbl

Hao Dou, Lori Buetow, Andreas Hock, Gary J. Sibbet, Karen H. Vousden and Danny T. Huang

Table of Contents

Supplementary Note

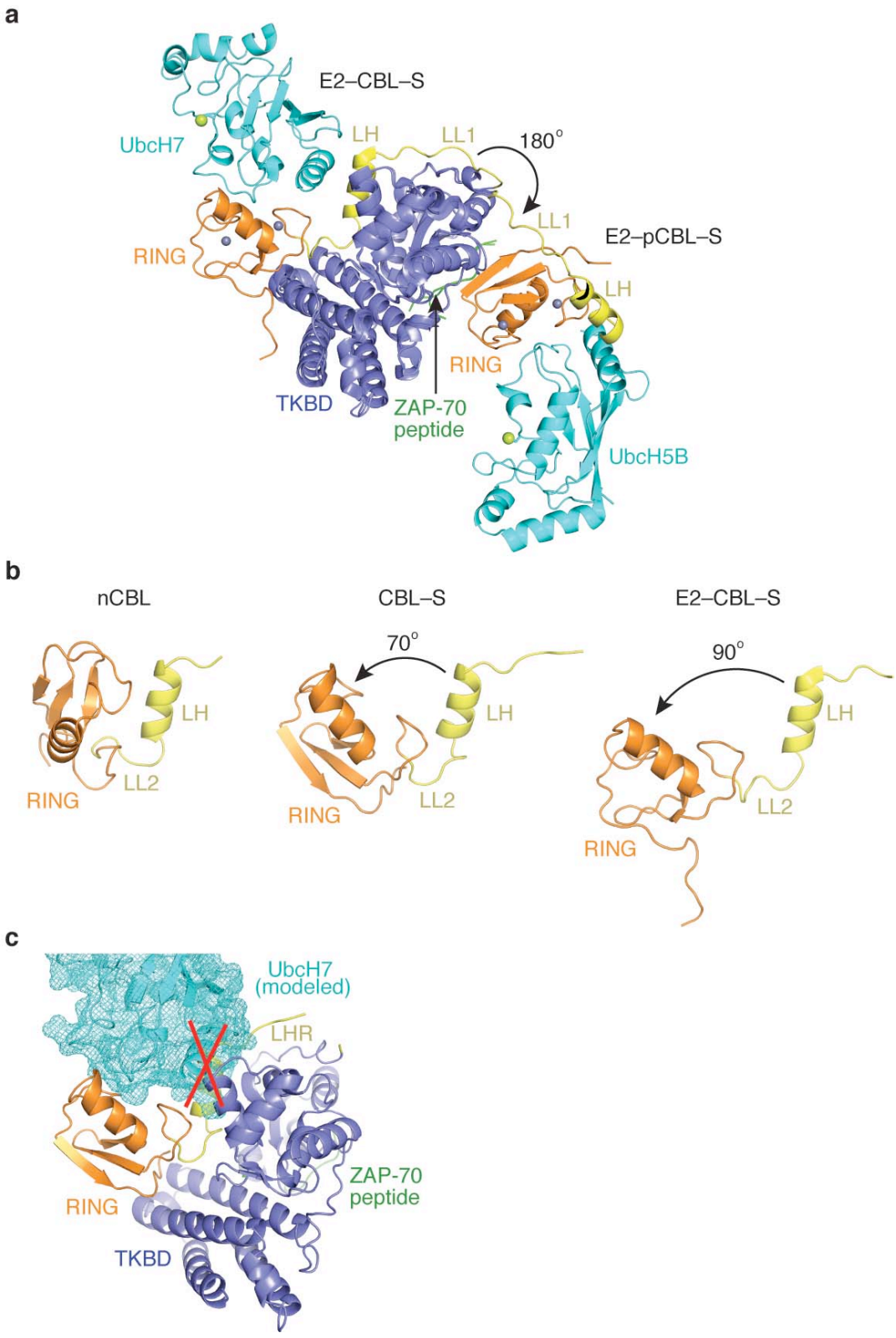
CBL's RING domain adopts different orientations in the structures of nCBL, CBL–S, E2–CBL–S and E2–pCBL–S. To determine whether CBL's RING domain rotates in solution in the presence and absence of E2, ZAP-70 peptide or phosphorylation, we developed a disulfide assay using cysteine mutants of CBL. The experiments described here use $CBL_{47–435}$.

We selected a pair of residues where one residue is located on the TKBD and the other on the C-terminal region of the RING domain such that their Cβ atoms are in close proximity in the structure of nCBL, but widely separated in other CBL structures (**Fig. 3a**). We then mutated these residues individually and in pairs to cysteine (Cys). In solution, we expected the double Cys mutant to form an intramolecular disulfide bond upon adopting the closed conformation. In the presence of E2, ZAP-70 peptide or Tyr371 phosphorylation, we predicted that disulfide bond formation would be slowed or prevented because the distance between the two Cys residues increases as the RING domain opens and the LHR conformation changes. Criteria required that the residues not participate in E2 recruitment, substrate binding or stabilization of the closed conformation. Several Cys pair mutants were engineered. Some formed the engineered intramolecular disulfide bond rapidly, even in the presence of reducing agent, while others formed non-specific disulfide bonds. CBL R139C D435C did not exhibit these problems. This construct is referred to as diCys. It is noteworthy that the side chain of Asp435 does not make contact with the RING domain in any of the CBL structures presented here. Thus mutation of this residue is unlikely to affect folding of the RING domain.

Given that CBL is a single polypeptide protein, we anticipated that visualizing the disulfide bond would be challenging. Thus, we inserted a unique thrombin

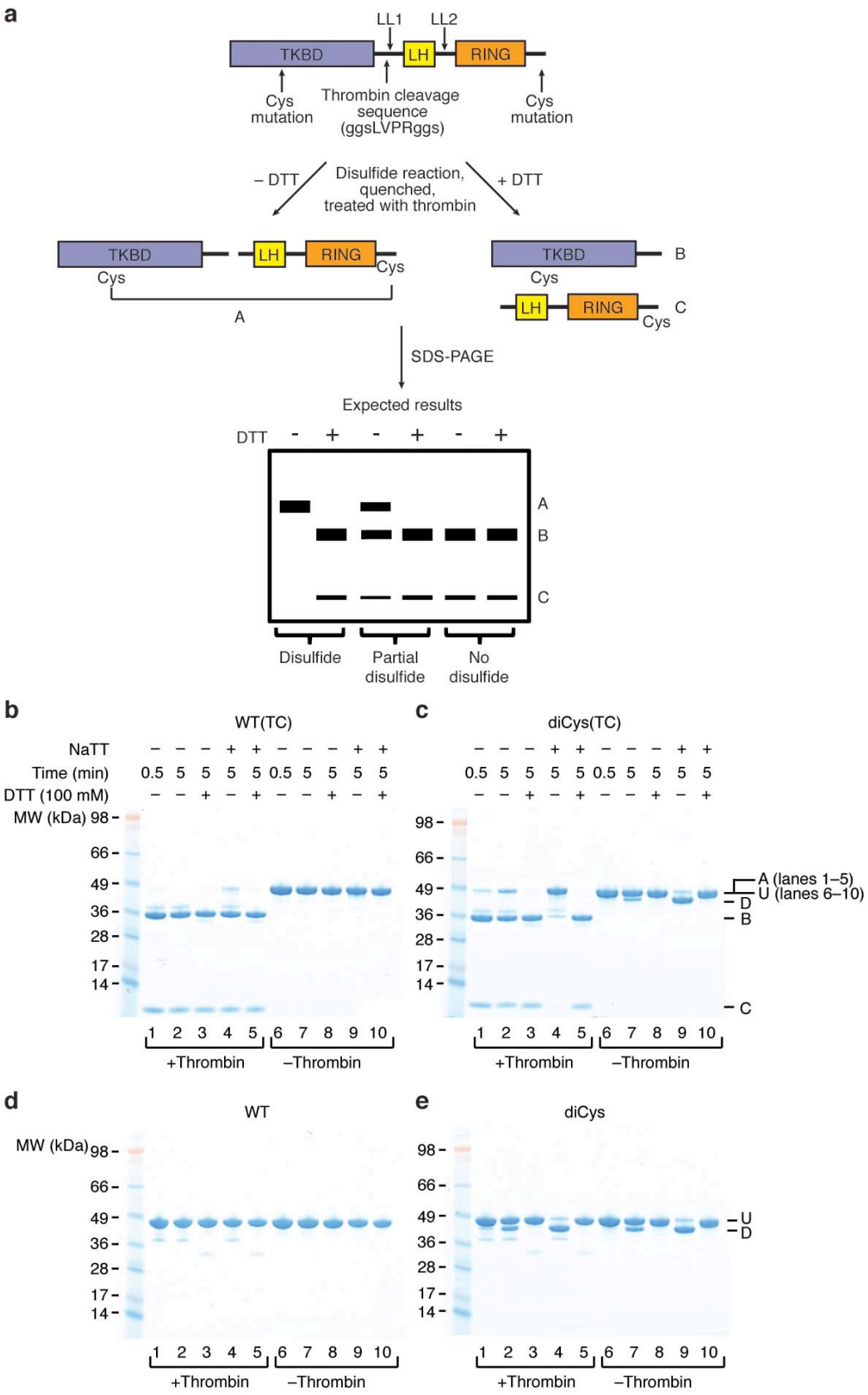
cleavage (TC) sequence (ggsLVPRggs) in LL1 between Glu354 and Pro355 (**Supplementary Fig. 2a**). These two residues do not contact any other residues in all CBL structures, and, based on the structure of E2–CBL–S, an insertion between these two residues was not expected to perturb the TKBD fold. CBL R139C D435C with a TC sequence is referred to as diCys(TC). Upon thrombin treatment, we expected CBL to yield two smaller fragments, the TKBD (36 kDa) and LHR-RING domain (9 kDa), which are separable by SDS-PAGE. If R139C and D435C formed an intramolecular disulfide bond, then the cleavage fragments will be tethered together and run as a large molecular weight band (45 kDa) that only dissociates into two smaller fragments in the presence of reducing agent (**Supplementary Fig. 2**).

Following thrombin treatment, diCys(TC) migrated as a single large molecular weight species and dissociated into two smaller fragments only in the presence of reducing agent, indicative of intramolecular disulfide bond formation (**Supplementary Fig. 2c**). In contrast, wild-type (WT) CBL containing a TC sequence migrated as two smaller fragments in both the presence and absence of reducing agent (**Supplementary Fig. 2b**). In the absence of thrombin treatment, the disulfide-bonded diCys(TC) migrated faster than the unmodified diCys(TC) (**Supplementary Fig. 2c** lanes 7 and 9). This pattern is also observed in diCys, which lacks the TC insertion sequence (**Supplementary Fig. 2e**). This fortuitous finding simplified the visualization of disulfide bind formation. In all subsequent disulfide assays, the fast migrating band was used to monitor intramolecular disulfide bond formation.



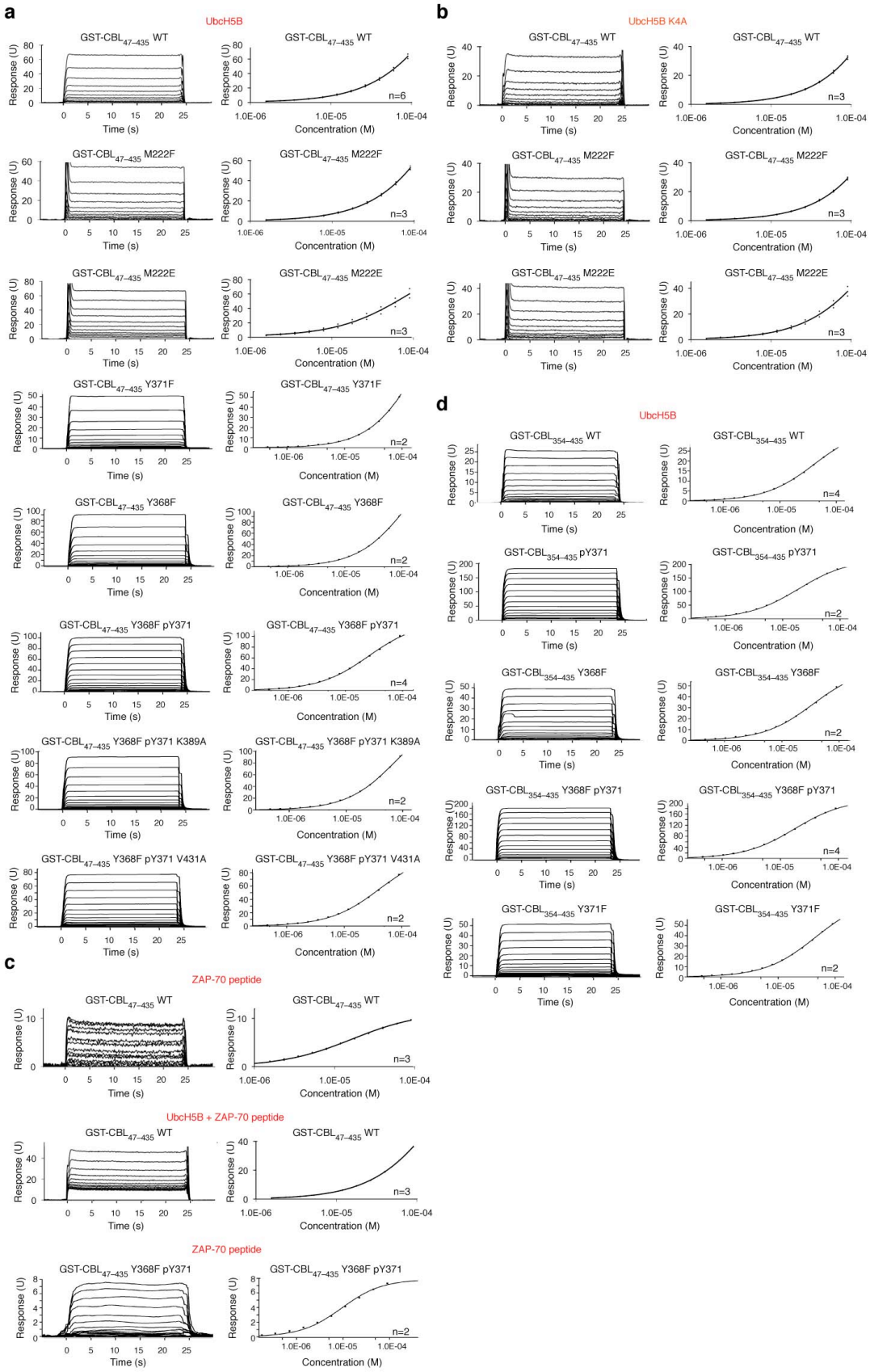
Supplementary Figure 1 LHR and RING domain rotation in CBL structures**.** (**a**) Superposition of the TKBD of E2–CBL–S and E2–pCBL–S reveals 180° rotation about LL1. (**b**) LH, LL2, and RING domains from the structures of nCBL (left), CBL–S (middle), and E2–CBL–S (right) displayed in the same orientation as in **Fig.**

1b–d. (**c**) Model of UbcH7 with CBL–S. The RING domains of CBL–S and E2–CBL–S were superimposed to generate this model. UbcH7 from E2–CBL–S is displayed as mesh in cyan. A red cross indicates the clash between UbcH7 and LH, suggesting that further conformational shifts are required to accommodate E2. All coloring is as described in **Fig. 1**.



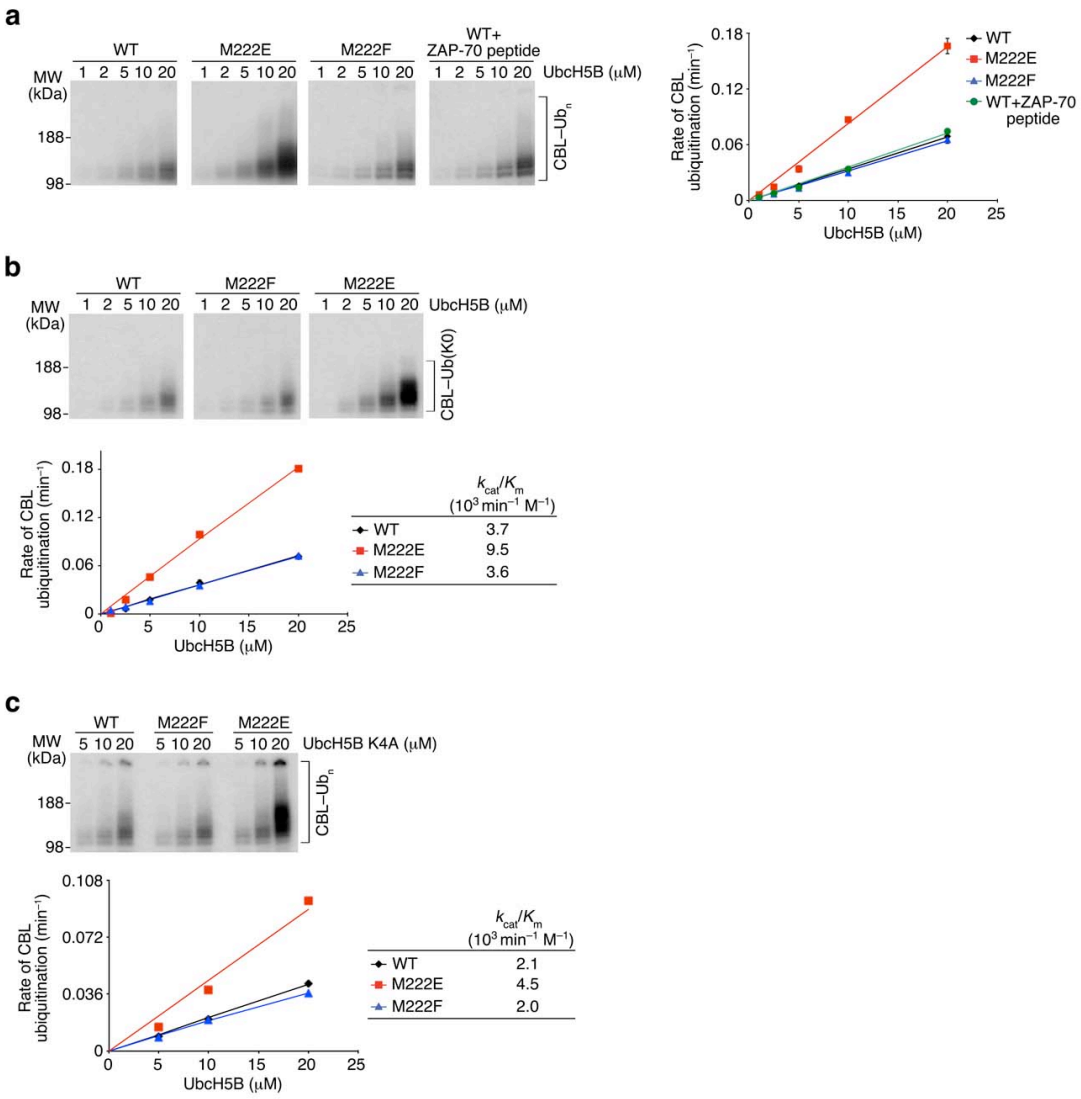
Supplementary Figure 2 CBL variants disulfide bond formation. (**a**) Strategy for visualization of intramolecular disulfide bond formation. Cys mutations were

introduced in the TKBD and C-terminus of the RING domain. A unique thrombin cleavage sequence (ggsLVPRggs) was inserted between Glu354 and Pro355 in LL1. Thrombin treatment is expected to separate CBL into bands B and C. If the engineered disulfide bond forms then CBL remains intact as A, but dissociates into B and C in the presence of a reducing agent such as DTT. Complete disulfide bond formation will produce band A and partial disulfide bond formation will yield bands A, B and C in the absence of DTT. In the case of no disulfide bond formation, bands B and C are expected in the presence or absence of DTT. Disulfide bond formation for WT(TC) (**b**), diCys(TC) (**c**), WT (**d**), and diCys (**e**) in the presence and absence of cross-linking reagent, NaTT. Lanes 1–5 and 6–10 were treated with and without thrombin, respectively. Bands A, B and C only refer to lanes 1–5 in **b** and **c** and represent species of CBL as in **a**. For lanes 6–10 in **b** and **c**, the slow migrating (upper) band corresponds to uncleaved diCys(TC) that did not form a disulfide bond and has been relabelled as U (Unmodified); the fast migrating (lower) band corresponds to uncleaved diCys(TC) that formed a disulfide bond and is relabelled as D (Disulfide). WT and diCys lacking an inserted TC sequence yield slow and fast migrating bands, which correspond to unmodified and disulfide-bonded species, respectively. The disulfide-bonded species is readily reduced to the unmodified species in the presence of DTT. Molecular weight standards are indicated. 0.5 min sample was taken immediately after buffer exchange.



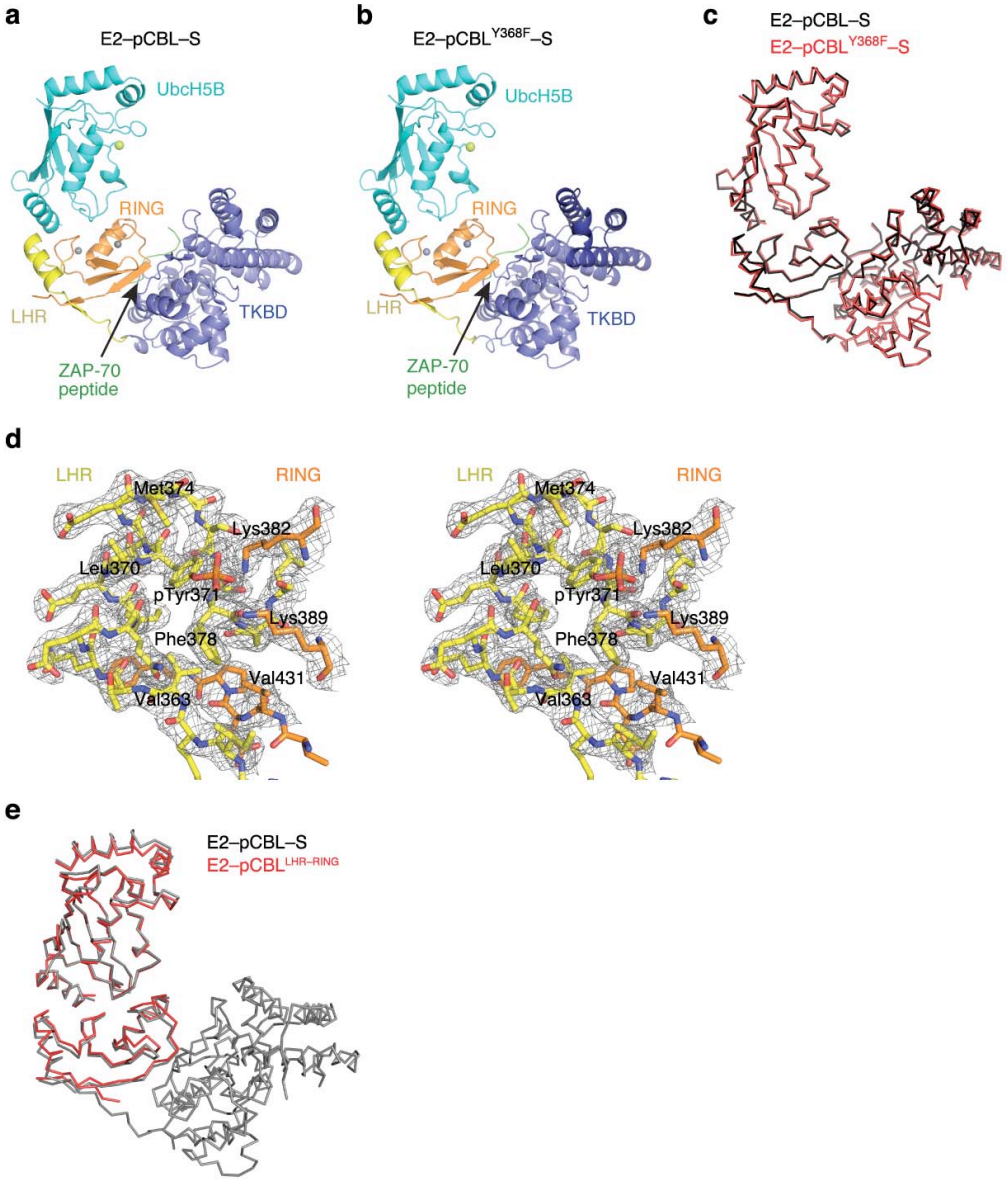
Supplementary Figure 3 Surface plasmon resonance analyses of CBL-E2 and CBL-ZAP-70 peptide binding. Representative sensorgrams (left) and binding curves (right)

for GST-$CBL_{47–435}$ variants with UbcH5B (**a**), UbcH5B K4A (**b**), ZAP-70 peptide in the presence or absence of UbcH5B (**c**), GST-$CBL_{354–435}$ variants with UbcH5B (**d**), as indicated. GST-ligands are indicated above the sensorgrams and binding curves. Analytes are in red. The number of independent replicates is indicated in the binding curve.

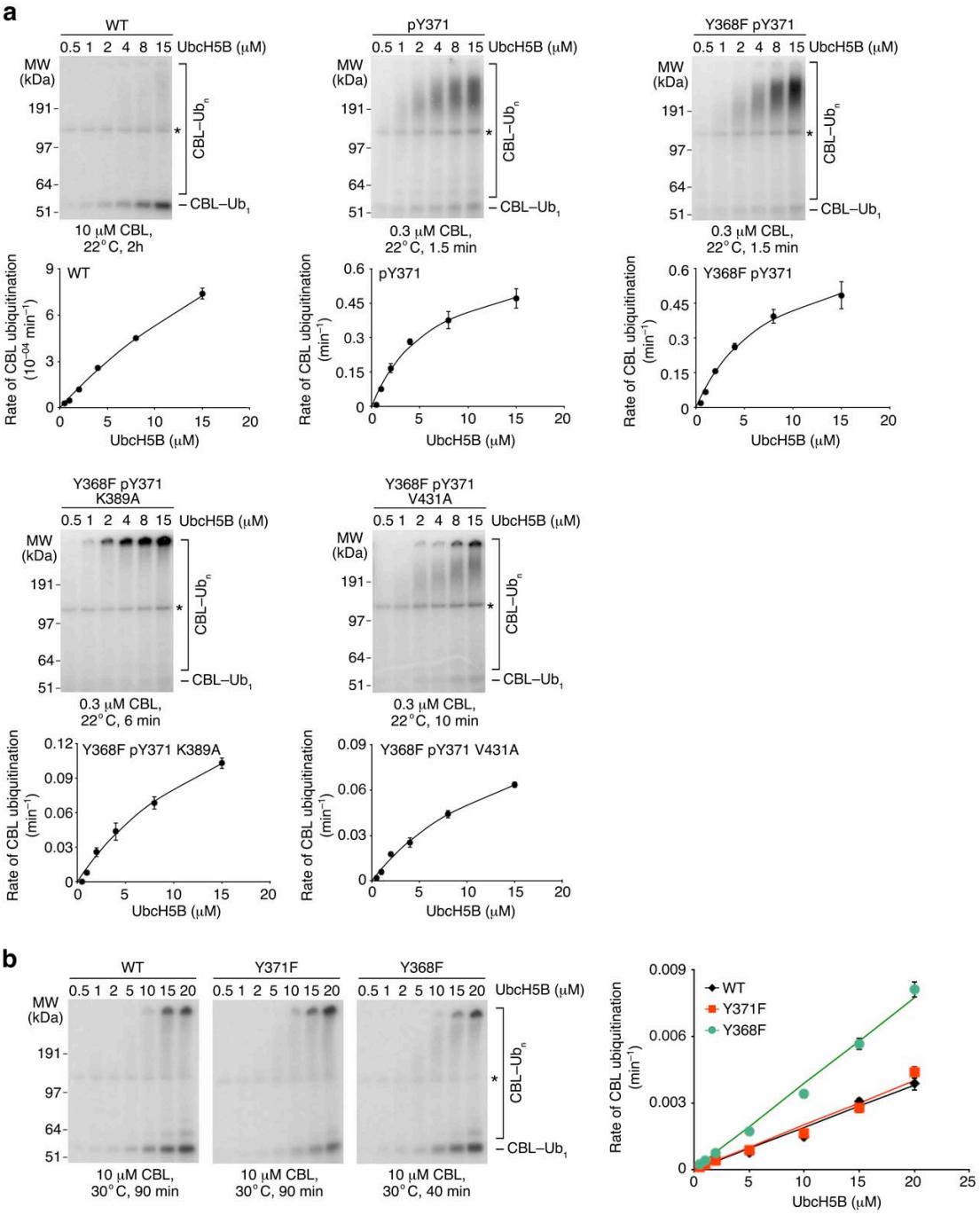


Supplementary Figure 4 Full-length CBL variants autoubiquitination. (**a**) Reduced autoradiograms showing the formation of ^{32}P-Ub products with full-length CBL variants (3 μM) and varying UbcH5B concentrations at 30°C for 15 min. The rate of CBL ubiquitination was plotted against UbcH5B concentration. (**b**) Reduced autoradiograms showing the formation of ^{32}P-K0-Ub products with full-length CBL variants (3 μM) and varying UbcH5B concentrations at 30°C for 15 min. The rate of CBL ubiquitination was plotted against UbcH5B concentration. k_{cat}/K_m values were estimated by fitting data points onto the Michaelis-Menten equation. Using chain-terminating K0-Ub, we found that full-length CBL is ubiquitinated at several lysine residues. (**c**) Reduced autoradiograms showing the formation of ^{32}P-Ub products with full-length CBL variants (3 μM) and indicated UbcH5B K4A concentrations at 30°C for 60 min. The rate of CBL ubiquitination was plotted against UbcH5B

concentration. k_{cat}/K_m values were estimated by fitting data points onto the Michaelis-Menten equation. Error bars show s.e.m. determined from two independent sets of reactions.

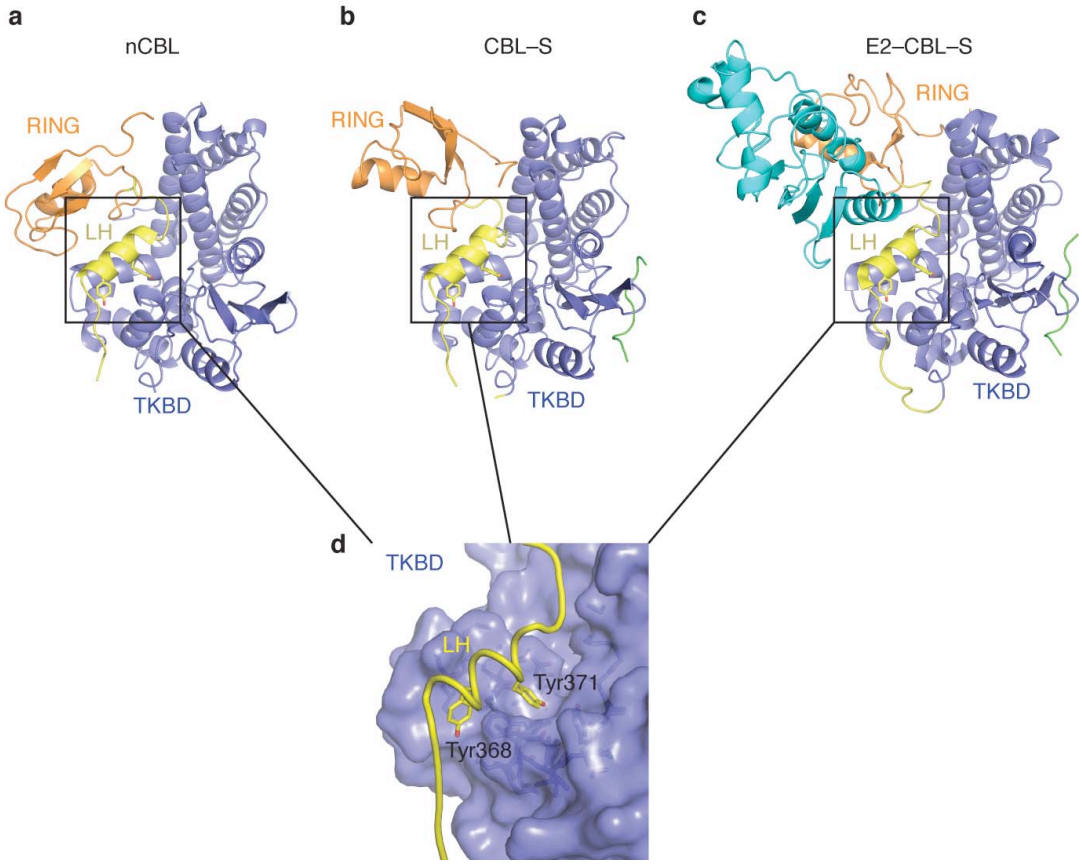


Supplementary Figure 5 The structures of E2–pCBL^{Y368F}–S and E2–pCBL$^{LHR–RING}$. (**a**) Structure of E2–pCBL–S. (**b**) Structure of E2–pCBL^{Y368F}–S. (**c**) Superposition of E2–pCBL–S (red) and E2–pCBL^{Y368F}–S (black) (r.m.s.d. of 0.64 Å for all Cα atoms). (**d**) Stereo view of the pTyr371-binding site in E2–pCBL$^{LHR–RING}$ complex. 2Fo-Fc electron density (grey) contoured at 1σ is shown. Key residues that stabilize pTyr371 are indicated. (**e**) Superposition of E2–pCBL$^{LHR–RING}$ complex (red) onto the E2–pCBL$^{LHR–RING}$ portion of E2–pCBL–S (grey) (r.m.s.d. of 1.1 Å for all Cα atoms). Coloring in **a**,**b**,**d** is as described in **Fig. 1**.

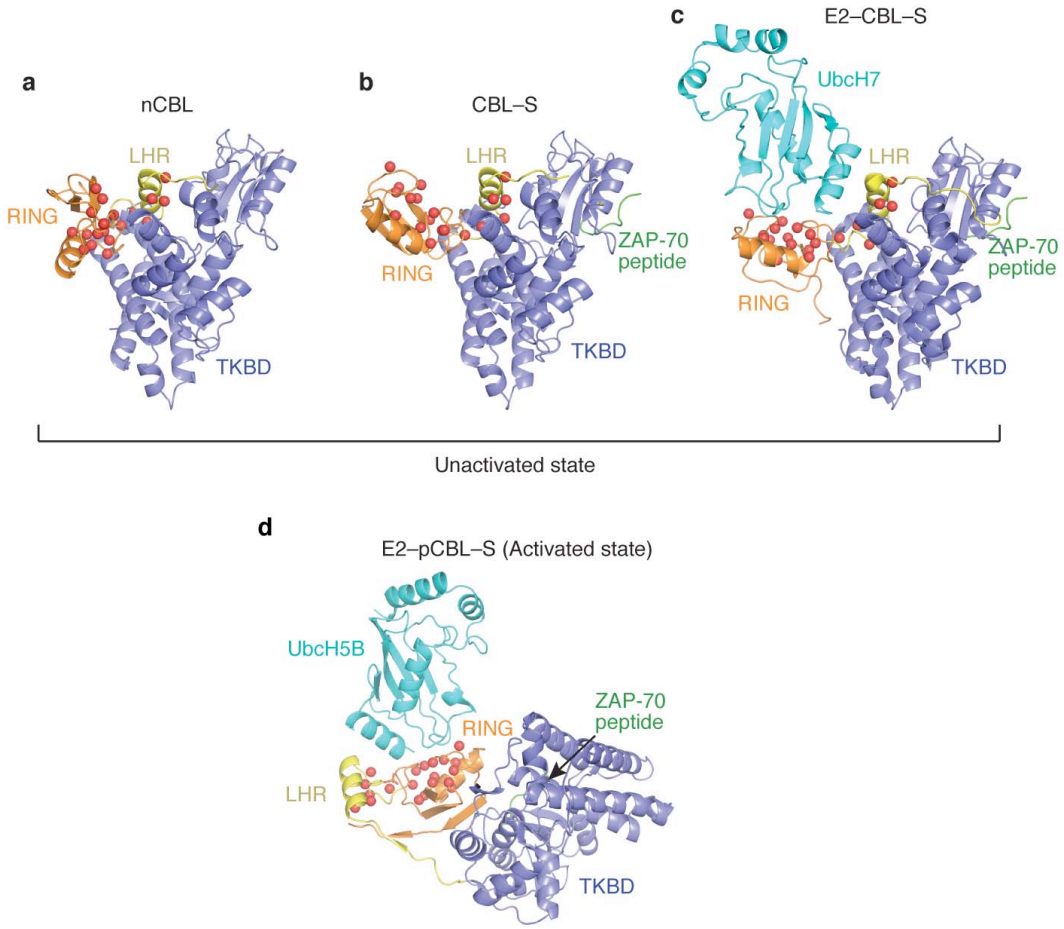


Supplementary Figure 6 $CBL_{47–435}$ variants autoubiquitination. (**a**) Reduced autoradiograms showing the formation of ^{32}P-Ub products with $CBL_{47–435}$ variants with or without Tyr371 phosphorylation and varying UbcH5B concentrations at 22°C. The rate of CBL ubiquitination was plotted against UbcH5B concentration. (**b**) Reduced autoradiograms showing the formation of ^{32}P-Ub products with $CBL_{47–435}$ variants and varying UbcH5B concentrations at 30°C. The rate of CBL ubiquitination was plotted against UbcH5B concentration. Asterisk indicates non-reducible E1-Ub band that formed in the absence of CBL and was omitted from the quantification.

Error bars shown in the curves are s.e.m. determined from two independent sets of reactions.



Supplementary Figure 7 LH-TKBD interactions in unphosphorylated CBL structures. Structures of nCBL (**a**), CBL–S (**b**) and E2–CBL–S (**c**) displayed in the same orientation and colored as described in **Fig. 1**. (**d**) Close up view of the inset in **a**–**c** showing LH's Tyr368 and Tyr371 are buried within the TKBD, thereby "fastening" LH to the TKBD.



Supplementary Figure 8 Locations of CBL mutational hot spots. Locations of CBL mutations in human patients with myeloproliferative diseases[1] are displayed as red spheres on the structures of nCBL (**a**), CBL–S (**b**), E2–CBL–S (**c**) and E2–pCBL–S (**d**). All coloring is as described in **Fig. 1**. Unactivated and activated CBL states are indicated.

Supplementary Methods

Protein preparation

Constructs were generated by standard PCR-ligation techniques and sequences verified by automated sequencing. All CBL variants comprising residues 47–435 were cloned into pGEX4T1 (GE Healthcare) which contains an N-terminal glutathione S-transferase (GST) tag followed by a TEV protease or thrombin cleavage site. Proteins were expressed in *E. coli* BL21 (DE3) Gold or BL21 (DE3) RIL (Stratgene). For crystallization, CBL was expressed from pGEX4T1, purified by glutathione-affinity chromatography, treated with TEV protease[2] to cleave the GST-tag, and further purified by desalting and glutathione-affinity pass-back and size exclusion chromatography. CBL pY371 variants were generated by coexpressing CBL variants in pGEX4T1 vector with a thrombin cleavage site and mouse Src 84–526 in pRSFDuet vector containing an N-terminal His-MBP tag. CBL pY371 variants were then purified by glutathione-affinity chromatography, treated with thrombin to release the GST-tag, and then separated from unphosphorylated CBL using anion exchange chromatography followed by size exclusion chromatography. Untagged UbcH5B was expressed from pRSF1b, purified by cation exchange and size exclusion chromatography. For disulfide assays, CBL variants were expressed and purified as described for crystallization except cleavage was performed on glutathione sepharose beads and the pass-back step omitted. For ubiquitination and disulfide assays, CBL pY371 variants were purified as described except size exclusion chromatography was omitted. For Biacore analyses, GST-CBL variants were expressed from pGEX4T1 and purified by glutathione-affinity and size exclusion chromatography; GST-CBL pY371 variants were coexpressed as described above then purified by glutathione-affinity and MonoQ chromatography. For Biacore analyses and ubiquitination assays,

UbcH5B variants were expressed from pRSFDUET, purified by Ni^{2+}-NTA affinity chromatography, treated with TEV protease, and further purified by cation exchange and size exclusion chromatography. For ubiquitination assays, human UBA1 and ^{32}P-Ub were prepared as described previously[3]. Chain-terminating K0-Ub was expressed in pGEX2TK vector and purified by glutathione-affinity chromatography, treated with TEV protease to cleave the GST-tag, and further purified by desalting and glutathione-affinity pass-back and size exclusion chromatography. His-tag Ub was expressed in pRSF1b vector and purified by Ni^{2+}-NTA affinity and size exclusion chromatography. Full-length CBL variants were expressed from pRSFDUET containing a TEV-cleavable N-terminal His-tag and purified by Ni^{2+}-NTA affinity chromatography followed by anion exchange and size exclusion chromatography. Protein concentrations were determined by Bradford assay using BSA as a standard[4] and Ub concentration was determined as described previously[5]. Proteins were stored in 25 mM Tris-HCl (pH 7.6), 0.15 M NaCl and 1 mM DTT or 25 mM HEPES (pH 7.0), 0.15 M NaCl and 1 mM DTT at –80°C. ZAP-70 phosphotyrosine peptide (sequence TLNSDG(p)YTPEPA) was purchased from Alta Bioscience.

In vitro EGFR ubiquitination assay

EGFR was purified from clarified lysate of A-431 cells by immunoprecipitation with agarose immobilized with anti-EGFR (Affibody, AB) as described previously[6]. Reactions (20 μL) were performed in 50 mM Tris-HCl (pH 7.6), 50 mM NaCl, 5 mM $MgCl_2$, 5 mM ATP, 0.3 U ml^{-1} creatine kinase, 0.3 U ml^{-1} inorganic pyrophosphatase, 5 mM creatine phosphate, 1 mg ml^{-1} BSA, AG1478 (100 μM; Merck), human UBA1 (100 nM), His-Ub (100 μM), $CBL_{47–435}$ variants (0.5 μM), UbcH5B (0.5 μM) and 3.5 μL EGFR at RT (22°C) for 30 min or 1h. For **Fig. 5b** right

panel, reactions were performed with UbcH5B (10 μM) at 30°C for 2h. The reactions were quenched with 6.6 M Guanidine HCl, incubated with Ni^{2+}-NTA beads and washed in the Micro Bio-Spin chromatography column (BioRAD) as follows: three times with wash buffer (25 mM Tris-HCl, 0.15M NaCl, 15 mM imidazole and 5 mM β-mercaptoethanol) containing 6 M Guanidine HCl, two times with wash buffer containing 3 M Guanidine HCl and then five times with wash buffer. The Ni^{2+}-NTA beads were then incubated with gel loading buffer containing 300 mM imidazole for 20 min and centrifuged to recover the sample. All samples were resolved by 6% SDS-PAGE and transferred to nitrocellulose membrane. EGFR was detected by Western blotting with EGFR antibody (MZ6847, Millipore).

***In vivo* EGFR ubiquitination assay**

U2OS cells were transfected with CMV-driven expression plasmids expressing His-Ub, EGFR-GFP and Myc-tagged full-length CBL as indicated. 24h after transfection, FCS was withdrawn for 18h. Cells were then induced with EGF (30 ng ml^{-1}) for 10 min, collected in PBS and 5% was kept as input. The rest was centrifuged for 5 min and the cell pellet lysed in 700 μl of UBA buffer (6 M Guanidine HCl, 0.3 M NaCl, 50 mM phosphate pH 8.0, 100 μg ml^{-1} NEM) and sonicated for 5 min at 20% amplitude. Lysates were incubated over night with Dynabeads His-tag matrix (Invitrogen), washed once each with UBA, UBB, UBC and PBS (UBB: UBA and UBC 1:1; UBC: 0.3 M NaCl, 50 mM phosphate pH 8.0, 100 μg ml^{-1} NEM) and resolved by 8% SDS–PAGE, followed by immunoblotting with GFP antibody (assay; 11814460001, Roche) or GFP, CBL (sc-170, Santa Cruz) and Actin (MAB1501, Millipore) antibodies (input).

Mass spectrometry

Protein was initially reduced with 50 mM DTT and then carbamyidomethylated with 50 mM iodoacetamide followed by a solution digest with trypsin. Resulting peptides were analyzed with LC-MS/MS with an Orbitrap Velos. Data were processed with Raw2MSM[7] and searched with Mascot 2.3 (ref. 8) against an *E. coli* database containing the sequence of the fragment of CBL used for crystallization and biochemical assays. For phosphorylated $CBL_{47–435}$, the standard score was 2659 and 77% of the sequence was covered; the absence of phosphorylation on Tyr191 was verified by examining electron density maps as the peptide containing this residue was not covered. For phosphorylated $CBL_{354–435}$, the standard score was 6953 and 85% of the sequence was covered, which included all possible tyrosine phosphorylation sites.

Structural determination

The data were integrated and scaled with HKL2000 (ref. 9) or integrated with Mosflm[10] or automated XDS[11] and scaled using the CCP4 program suite[12]. nCBL crystals belong to space group $C222_1$ with six molecules in the asymmetric unit. CBL–S crystals belong to space group $P6_5$ with two complexes per asymmetric unit. E2–pCBL–S and E2–$pCBL^{Y368F}$–S crystals belong to space group $P6_522$ with one complex in the asymmetric unit. E2–$pCBL^{LHR–RING}$ crystals belongs space group $P3_221$ with one complex in the asymmetric unit. Initial phases of nCBL and CBL–S were obtained by molecular replacement with PHASER[13] using PDB 1B47 (ref. 14; CBL 47–350) and PDB 2CBL[14] (CBL 47–350 and ZAP-70 peptide 4–12) as the search models, respectively. The structure of CBL's RING domain (381–428; PDB 1FBV[15])

was manually moved into the electron densities of nCBL and CBL–S. Initial phases of E2–pCBL–S and E2–pCBL^{Y368F}–S were obtained by molecular replacement with PHASER using PDB 2CBL[14] (CBL 47–350 and ZAP-70 peptide 4–12), PDB 2ESK[16] (UbcH5B) and PDB 1FBV[15] (CBL 381–428). Initial phase of E2–pCBL$^{LHR–RING}$ was determined by molecular replacement with PHASER using PDB 2ESK[16] (UbcH5B) and PDB 1FBV[15] (CBL 381–428). All models were built in COOT[17] and refined using CCP4, CNS[18] and PHENIX[19].

The structure of nCBL (Chains A–F) was refined at a resolution of 2.67 Å and the model contained Chain A residues 47–353 and 359–435, Chain B residues 47–355 and 360–435, Chain C residues 47–354 and 359–435, Chain D residues 47–354 and 359–435, Chain E residues 47–353 and 360–435, and Chain F residues 47–353 and 359–435. PHENIX Xtriage analyses indicated that CBL–S crystals exhibited merohedral twinning with twin operator (h,–h–k,–l) and an estimated twin fraction of 0.395. The structure of CBL–S was refined to 2.0 A resolution with PHENIX using a merohedral twinning protocol. The final model contained CBL (Chain A residues 52–354 and 362–435 and Chain C residues 52–354 and 359–435) and ZAP-70 peptide (Chains B and D residues 4–12). The final model of E2–pCBL–S contained CBL (Chain A residues 48–435), ZAP-70 peptide (Chain B residues 4–12), UbcH5B (Chain C residues 2–128 and 130–147). The final model of E2–pCBL^{Y368F}–S contained CBL (Chain A residues 48–435), ZAP-70 peptide (Chain B residues 4–12), UbcH5B (Chain C residues 2–147). The final model of E2–pCBL$^{LHR–RING}$ contained CBL (Chain A residues 359–435) and UbcH5B (Chain B residues 2–147). Details of the refinement statistics are shown in **Table 1**. All figure models were generated using PYMOL (Schrödinger). Figures were assembled with Adobe Illustrator CS5.1 and Adobe Photoshop CS5.1 (Adobe).

Supplemental References

1. Kales, S.C., Ryan, P.E., Nau, M.M. & Lipkowitz, S. Cbl and human myeloid neoplasms: the Cbl oncogene comes of age. *Cancer Res.* **70**, 4789–94 (2010).
2. Nallamsetty, S. *et al.* Efficient site-specific processing of fusion proteins by tobacco vein mottling virus protease in vivo and in vitro. *Protein Expr. Purif.* **38**, 108–15 (2004).
3. Huang, D.T., Zhuang, M., Ayrault, O. & Schulman, B.A. Identification of conjugation specificity determinants unmasks vestigial preference for ubiquitin within the NEDD8 E2. *Nat. Struct. Mol. Biol.* **15**, 280–7 (2008).
4. Bradford, M.M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72**, 248–54 (1976).
5. Bohnsack, R.N. & Haas, A.L. Conservation in the mechanism of Nedd8 activation by the human AppBp1-Uba3 heterodimer. *J. Biol. Chem.* **278**, 26823–30 (2003).
6. Schmidt, M.H. & Dikic, I. Assays to monitor degradation of the EGF receptor. *Methods Mol. Biol.* **327**, 131–8 (2006).
7. Olsen, J.V. *et al.* Parts per million mass accuracy on an Orbitrap mass spectrometer via lock mass injection into a C-trap. *Mol. Cell. Proteomics* **4**, 2010–21 (2005).
8. Perkins, D.N., Pappin, D.J., Creasy, D.M. & Cottrell, J.S. Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* **20**, 3551–67 (1999).
9. Otwinowski, Z. & Minor, W. Processing of X-ray Diffraction Data Collected in Oscillation Mode. *Methods in Enzymology* **276**, 307–326 (1997).
10. Leslie, A.G.W. Recent changes to the MOSFLM package for processing film and image plate data *Joint CCP4 + ESF-EAMCB Newsletter on Protein Crystallography, No. 26.* (1992).
11. Kabsch, W. Xds. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 125–32 (2010).
12. Collaborative Computational Project, N. "The CCP4 Suite: Programs for Protein Crystallography". *Acta Cryst.* **D50**, 760–763 (1994).
13. Storoni, L.C., McCoy, A.J. & Read, R.J. Likelihood-enhanced fast rotation functions. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 432–8 (2004).
14. Meng, W., Sawasdikosol, S., Burakoff, S.J. & Eck, M.J. Structure of the amino-terminal domain of Cbl complexed to its binding site on ZAP-70 kinase. *Nature* **398**, 84–90 (1999).
15. Zheng, N., Wang, P., Jeffrey, P.D. & Pavletich, N.P. Structure of a c-Cbl-UbcH7 complex: RING domain function in ubiquitin-protein ligases. *Cell* **102**, 533–9 (2000).
16. Ozkan, E., Yu, H. & Deisenhofer, J. Mechanistic insight into the allosteric activation of a ubiquitin-conjugating enzyme by RING-type ubiquitin ligases. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 18890–5 (2005).
17. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–32 (2004).

18. Brunger, A.T. *et al.* Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D Biol. Crystallogr.* **54**, 905–21 (1998).
19. Adams, P.D. *et al.* PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr. D Biol. Crystallogr.* **58**, 1948–54 (2002).