HHS Public Access

Author manuscript

Structure. Author manuscript; available in PMC 2018 February 07.

Published in final edited form as:

Structure. 2017 February 07; 25(2): 287–294. doi:10.1016/j.str.2016.12.008.

STK40 is a pseudokinase that binds the E3 ubiquitin ligase COP1

Izabela Durzynska[1,2], **Xiang Xu**[2], **Guillaume Adelmant**[1,2], **Scott B. Ficarro**[1,2], **Jarrod A. Marto**[1,2], **Piotrek Sliz**[2], **Sacha Uljon**[1,2,3], and **Stephen C. Blacklow**[1,2,3,*]

[1]Department of Cancer Biology, Dana Farber Cancer Institute, Boston, MA 02215

[2]Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston MA, 02115

[3]Department of Pathology, Brigham and Women's Hospital and Harvard Medical School, Boston MA 02115

SUMMARY

Serine/threonine kinase 40 (STK40) was originally identified as a distant homolog of Tribbles-family proteins. Despite accumulating data attesting to the importance of STK40 in a variety of different physiologic processes, little is known about its biological activity or mechanism of action. Here, we show that STK40 interacts with Constitutive Photomorphogenic Protein 1 (COP1), relying primarily on a C-terminal sequence analogous to the motif found in Tribbles proteins. In order to further elucidate structure-function relationships in STK40, we determined the crystal structure of the STK40 kinase homology domain at 2.5 Å resolution. The structure, together with ATP-binding assay results, show that STK40 is a pseudokinase, in which substitutions of conserved residues within the kinase domain prevent ATP binding. Though the structure of the kinase homology domain diverges from the analogous region of Trib1, the results reported here suggest functional parallels between STK40 and tribbles-family proteins as COP1 adaptors.

Graphical abstract

Correspondence to SCB or SU: stephen_blacklow@hms.harvard.edu or suljon@partners.org.
*Lead contact

ACCESSION NUMBERS

The accession number for the coordinates for the X-ray structure of the STK40 kinase homology domain reported in this paper is PDB: 5L2Q.

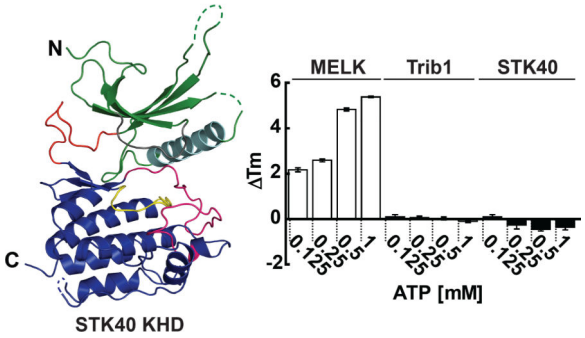
SUPPLEMENTAL INFORMATION

Supplemental Information includes 3 figures and 1 table and can be found with this article online at xxxx.

AUTHOR CONTRIBUTIONS

Conceptualization, S.U., S.C.B., and I.D.; Methodology, S.U., S.C.B., G.A., P.S. and I.D.; Investigation, all authors; Writing and Editing, S.U., S.C.B., and I.D.; Supervision, S.C.B., and S.U.; Funding, S.C.B. and S.U.

Tagged "Bait"	Recovered COP1 peptides Cytoplasmic (Nuclear)
Flag-HA-STK40 FL (1)	12 (14)
Flag-HA-STK40 FL (2)	15 (14)
Flag-HA-STK40 FL (3)	13 (16)
Control HeLa	0 (0)



STK40 is a pseudokinase that binds COP1

INTRODUCTION

Protein kinases regulate complex biological processes, including proliferation, differentiation, and cell death. Approximately 10% of proteins containing kinase homology domains are predicted to be enzymatically inactive (pseudokinases), owing to the absence of at least one of the conserved motifs required for efficient ATP binding or catalysis. Nearly half of these pseudokinases are reported to bind ATP, but the functional role of nucleotide binding in such cases remains unknown (Hammaren et al., 2016; Manning et al., 2002).

STK40 (serine/threonine kinase 40), also known as SgK495 (Sugen kinase 495) or SHIK (SINK-homologous inhibitory kinase), was originally identified as a homolog of Trib3 (SINK) (Huang et al., 2003). Though Tribbles-family (Trib) pseudokinases are the closest relatives of STK40, they share only 20% (Trib1), 19% (Trib2), and 21% (Trib3) sequence identity.

The precise function of STK40 is unknown, but current evidence points toward a role in promoting cell differentiation. For example, studies in knockout mice show that loss of STK40 results in respiratory failure and perinatal lethality due to the failure of maturation in lung epithelial cells, a phenotype similar to that of mice lacking the differentiation factor C/EBPα in respiratory epithelial cells (Martis et al., 2006; Yu et al., 2013). Forced expression of STK40 in embryonic stem cells induces extraembryonic endoderm differentiation, a process that may be mediated through the Erk/MAPK pathway (Li et al., 2010; Zhang et al., 2014). Conversely, knockdown of STK40 in cultured mouse embryonic fibroblasts (MEFs) promotes differentiation into adipocytes, attributed in part to translational control of the levels of C/EBPβ and C/EBPδ (Yu et al., 2015). A growing number of studies also suggest roles for STK40 in other cellular contexts. STK40 may be a regulatory target of miR-31, and suppression of STK40 expression correlates with the development of psoriasis and esophageal squamous cell carcinoma (ESCC) in a Zn-deficient rat model (Taccioli et al., 2015; Xu et al., 2013).

Despite this growing evidence attesting to the importance of STK40 in a variety of biological processes, there is no structural and little biochemical evidence to provide a molecular mechanistic understanding of STK40 function. Here, we show that STK40 binds to the COP1 E3-ubiquitin ligase using a C-terminal COP1-binding sequence analogous to motifs present in Tribbles proteins and Jun and ETS transcription factors, and report the crystal structure of the STK40 kinase homology domain to 2.5 Å resolution, which, together with biochemical studies, reveals that STK40 is a pseudokinase. Despite structural divergence of the kinase homology domain from the analogous region of Trib1, the similar modular configuration of a pseudokinase preceding a COP1-binding motif hints at the potential for functional parallels between STK40 and Trib-family proteins.

RESULTS

STK40 binds to COP1 in human cells

In order to begin deciphering the function of STK40, we performed a proteome-wide search for STK40 binding proteins in human cells. We generated HeLa cells stably expressing a full-length STK40 containing N-terminal FLAG, HA tags (Figure 1A,B), and then purified STK40 complexes using sequential anti-Flag and anti-HA immunoprecipitations (IPs). Parental HeLa cells were similarly processed in parallel as negative controls. Peptides eluted from the tandem IPs, analyzed according to an established protocol (Adelmant et al., 2012), detected a number of proteins that associate with STK40, including COP1 (also known as RFWD2 in humans) (Table S1 and Figure 1B). COP1 peptides were well represented in both the cytoplasmic and nuclear fractions in three independent samples, whereas no COP1 peptides were detected in the control IPs (Figure 1B, Figure S1, Table S1).

STK40 interacts with the WD40 domain of COP1 *via* a C-terminal binding motif

In order to assess further the interaction between STK40 and COP1, we performed Flag-immunoprecipitations on HEK293T cells co-transfected with Flag-STK40 and COP1 (Figure 1C). Full-length STK40 indeed associates with full-length COP1 (Figure 1C, Lane 5). Domain mapping studies show that the WD40 domain of COP1 is sufficient for binding full-length STK40 (Figure 1C, lane 6). Reassuringly, full length STK40 also immunoprecipitates with endogenous COP1 (Figure 1C, black arrow).

Alignment of the STK40 and Tribbles sequences reveals a potential COP1 binding motif in STK40 immediately C-terminal to the kinase homology domain (amino acids 339–351, SSLSGPLQVVPDI) (Figure 1A). This motif aligns well to the COP1-binding sequence of Trib1 and contains the signature VP sequence structurally critical for peptide binding (Uljon et al., 2016), but differs from the “canonical” COP1 binding motif because it lacks the acidic residue(s) typically found 2–3 amino acids N-terminal to the VP(D/E) motif. We first tested whether this region of STK40 binds to COP1 in a pull-down assay using a 12 residue peptide fused to GFP, and observed that the isolated peptide recovers the COP1 WD40 domain, but less efficiently than full-length STK40 (Figure 1C, lane 8). Conversely, a construct of STK40 containing only the central kinase homology domain (KHD) is unable to recover detectable quantities of the COP1 WD40 domain (Figure 1C, Lane 7). We then determined the affinity of the isolated STK40 peptide for the purified COP1 WD40 domain

using a previously described fluorescence polarization assay (Uljon et al., 2016). The STK40 peptide bound to the WD40 domain of COP1 with a K_d of 3.1 ± 0.6 μM. This affinity is not as strong as that of an analogous peptide from Trib1 (Uljon et al., 2016), which binds with a K_d of 0.5 ± 0.1 μM under these conditions. In competition assays using purified STK40 proteins, the kinase homology domain (KHD) of STK40 does not compete with peptide binding (and appears to increase the polarization change). In contrast, the KHD containing a C-terminal extension of 20 amino acids encompassing the COP1-binding peptide successfully displaces the fluorescently labeled peptide with an apparent K_i of 1.2 ± 0.1 μM, with a lower apparent K_i than the isolated peptide in the same competition assay (Figure 1E). Lastly, peptides in which the core VP dipeptide sequence is mutated fail to compete in this binding assay, indicating that peptide binding relies on the integrity of those key COP1 interacting amino acids (Figure 1E).

Structure of STK40

To gain insight into the mechanism of action of STK40, we solved the crystal structure of a portion of the STK40 protein that includes the complete kinase homology domain (Figure 2A) to 2.5 Å resolution (Table 1). STK40 adopts the overall architecture of a protein kinase with canonical N- and C-terminal lobes (Figure 2B). The asymmetric unit (ASU) contains four molecules, which superimpose with a backbone root-mean-square deviation (RMSD) of <1.1 Å for the core regions (aa 37–116 and 136–335, using the SSM [secondary structure matching] algorithm)(Krissinel and Henrick, 2004). [Note that the extent of disorder varies among the four copies in some regions due to crystal packing interactions (Figure S2); the coordinates from chain A, which exhibits the lowest overall B factors, are used for all figures except one (Figure 3A).] The N-lobe (Figure 2B, green) consists of a five-stranded anti-parallel beta sheet and a single conserved helix, the canonical kinase “C-helix” (Figure 2B, cyan). The helical C-lobe (figure 2B, blue) includes a well-structured activation loop (Figure 2B, pink).

The C-lobe of STK40 superimposes well with the C-lobes of prototypical kinase domains, represented here by the enzymatically active Protein Kinase A (pdb code 1ATP) (Zheng et al., 1993) and the pseudokinase Trib1 (pdb code 5CEM) (Murphy et al., 2015), with backbone RMSD values of 1.4 Å and 1.2 Å, respectively over 170 residues (SSM, STK40 aa 159–333; PKA aa 128–298; Trib1 aa 166–339). Surprisingly, the N-lobe of STK40 more closely resembles that of the active kinase PKA (N-lobe LSQ [least squares superposition] RMSD, 3.6 Å over 37 Cα atoms; Figure 2C) than that of the related pseudokinase Trib1 (N-lobe LSQ RMSD, 11.4 Å over 35 Cα atoms; Figure 2D). The αC helix of STK40 is rotated inwards, an orientation typically associated with the active state of a true kinase like PKA. This conformation of the C-helix is stabilized by a salt bridge between Lys66 on β3 and Glu93 on αC, analogous to the Lys72-Glu91 interaction in PKA (Figure 2C, wheat). In contrast, the STK40 N-lobe superimposes poorly onto that of Trib1, primarily because Trib1 contains a truncated, bent αC helix, incapable of forming a stabilizing salt bridge required for efficient nucleotide binding (Figure 2D, light pink).

STK40 possesses two atypical sequence insertions when compared with the canonical active kinase PKA. First, there is an extended loop between β4 and β5 (aa 116–135, Figure 3A).

This loop, which contains a predicted nuclear localization signal (NLS) (black bracket, Figure 3A), is likely disordered in the protein, but is visible in two of the four copies (chains C and D) within the asymmetric unit because of crystal packing interactions as part of an extended beta-hairpin. A DALI search using the STK40 structure reveals one other distantly related pseudokinase, ROP2 (pdb code 2W1Z), that contains a similar β4 and β5 strand extension (Figure 3B) (Holm and Rosenstrom, 2010; Labesse et al., 2009). In addition, STK40 contains a unique 10 amino acid insert in the hinge region (aa 145–155, Figure 3B). When the structures of PKA and STK40 are superimposed, the unique extended hinge region of STK40 does not align with the hinge region of PKA (Figure 3C, red). Rather, the extended hinge region of STK40 more closely resembles a portion of the C-terminal extension of PKA (residues 316–327, cyan, Figure 3C). This “C-tail” is present in the AGC-family kinases and is important for correct positioning of the C-helix in the active conformation (Kannan et al., 2007; Taylor et al., 2013). In STK40, however, it appears to be a structurally analogous element buttressing the N-lobe, rather than a modulator of kinase activity (see below).

STK40 contains sequence substitutions of kinase residues critical for catalytic activity

The STK40 kinase homology domain maintains the overall fold of an active kinase and the catalytic loop contains the HRD sequence found in active kinases (HRDxKxxN) (Figure 4A). Hydrophobic residues corresponding to the regulatory spine positions of active kinases make packing interactions that stabilize the conformational relationship between the N- and C-lobes of STK40. However, STK40 bears mutations that degrade key motifs important for kinase activity, consistent with its designation as a pseudokinase. First, the glycine-rich loop (or “P-loop”) of STK40 is highly divergent from canonical kinases, with two of the canonical glycine residues replaced by a serine and a proline (LGNSPVP) (Figure 4A). Second, whereas active kinases envelop the adenine ring of ATP using hydrophobic residues (Figure 4B), STK40 contains substitutions in the β2 (V) and β3 (AxK) loops, as the alanine residue in β3 is replaced by glutamine (Q64). Substitution of this crucial position is a sensitive indicator of loss of ATP binding. When ATP is modeled into our STK40 structure, it is evident that Q64 projects into the site normally occupied by ATP in an active kinase (Figure 4B and Figure S3). In ROP2, the analogous E280 also partially occludes the nucleotide binding site (Labesse et al., 2009). In the case of VRK3, the alanine is mutated to a small polar residue, serine (S201), that forms a hydrogen bond with the backbone amide of L262, also resulting in partial occlusion of the ATP binding site (Scheeff et al., 2009). By contrast, Trib1 has no steric hindrance to ATP binding (Figure 4B). Rather, the C-helix is so degraded that the ATP pocket can scarcely be said to exist at all.

Another striking change in the putative ATP binding pocket in STK40 is the modification of the highly conserved magnesium binding motif (DFG) important for the catalytic function of active kinases. In place of the “DFG” motif, STK40 has an NFC sequence, eliminating the charged aspartate participating in magnesium coordination. In addition the glycine is replaced by a cysteine (C218), contributing to further degradation of the site for ATP binding and phosphate transfer (Figure 4A, Figure S3). The related Trib1 kinase-homology domain harbors an unusual SLE sequence in this position, which also disrupts the magnesium ion-coordination site.

To determine whether the purified STK40 pseudokinase domain binds ATP or the promiscuous ATP-competitor staurosporine at various concentrations, we performed a fluorescent thermal shift assay using the crystallized construct (residues 22–339). The serine/threonine maternal embryonic leucine zipper kinase (MELK), used as a positive control, binds ATP in a magnesium dependent manner with a thermal shift (ΔTm) of 2–5.5°C, and exhibits a thermal shift of 15– 16°C upon binding staurosporine (Figure 4C, D white bars). In contrast, neither Trib1 pseudokinase (grey bars, used here as a negative control), nor the STK40 (black bars) exhibits a significant thermal shift in the presence of magnesium-ATP or staurosporine. These results are consistent with a prior report showing no ATP binding by a partially purified full-length STK40 at a single ATP concentration (0.2mM) (Murphy et al., 2014).

DISCUSSION

Here we demonstrate an interaction in human cells between the pseudokinase STK40 and the E3 ubiquitin ligase COP1. We have determined that the STK40 C-terminal peptide (339–351) recapitulates binding in a purified system and contributes to binding in cells, and that the COP1 WD40 domain is necessary and sufficient for binding. These results echo our recent findings for the binding of the pseudokinase Trib1 to the WD40 domain of COP1 (Uljon et al., 2016), though the functional significance of the STK40-COP1 interaction is still unknown. Intriguingly, the observation that both STK40 and Tribbles pseudokinases bind to the same site on the COP1 WD40 domain, together with the lack of evidence that either is a target of COP1-mediated degradation, raises the possibility that STK40 and Tribbles may act as adaptors for COP1 substrate selection or as regulators of COP1 ligase activity.

The structure of the STK40 kinase homology domain reveals the position of key kinase motifs otherwise not readily identifiable by multiple sequence alignment due to its sequence divergence. Even though substitutions of key residues conserved in active kinases degrade the ATP binding site and render the protein incapable of ATP binding, it is remarkable that the fold of STK40 nevertheless closely resembles that of an active kinase and differs notably from its closest structural pseudokinase relative Trib1, especially in the conformation of the αC helix in the N-lobe. (Figure 2C, D). The substitutions responsible for inactivating the ATP binding site are evolutionarily conserved in the furthest identifiable STK40 orthologue in frogs, indicating loss of kinase activity occurred early in evolution.

STK40 contains two unusual sequence insertions, one in the N-lobe between β4 and β5 and another in the hinge loop. Within the β4 and β5 insertion, we identify a potential nuclear localization signal. The loop is highly solvent exposed and two serine residues present upstream of the NLS could function as a switch for cellular localization of STK40 by undergoing phosphorylation to neutralize the positive change of the putative NLS, though it remains to be determined whether this predicted NLS motif is active in cells and whether the phosphorylation of the upstream serines plays a role in nuclear localization. The hinge insertion of STK40 is highly unusual and no other structures containing a hinge insertion were found in a DALI search. We speculate that it serves a purely structural role in stabilizing the conformation of the N-lobe.

Overall, the resemblance of the STK40 global fold to that of the active kinase PKA in the "closed state" is remarkable, but belies its inability to bind ATP. Because our structure contains only the kinase homology domain, and our ATP binding experiments employ the crystallized construct, we cannot rule out the possibility that full length STK40, with or without cellular partners, may in some cases function as an active kinase. However, the most parsimonious interpretation is that STK40 belongs to the growing list of inactive pseudokinases that act as scaffolds to alter the activity of their binding partners.

METHODS

Protein Production

STK40 constructs with N-terminal Histidine tags (22–359, and 22–339) were expressed and purified from Hi5 cells. After lysis in a buffer containing 50mM Tris-HCl (pH 8.0), 500mM NaCl, 5% glycerol, 10mM β-mercaptoethanol, and protease inhibitors, STK40 was purified using Ni-NTA affinity, followed by cation-exchange chromatography, and then by gel filtration in buffer containing 30 mM Tris-HCl pH 7.5, 250 mM NaCl, 5% glycerol, and 2 mM TCEP. The histidine tag was cleaved with TEV protease after the Ni-NTA affinity step. Purified proteins were concentrated to 7.5 mg/ml for crystallization trials. N-terminally tagged selenomethionine (SeMet)-labeled STK40 was produced in BL21 cells and purified using the same method. Tribbles and COP1 proteins used in biochemical studies were produced as previously described (Uljon et al., 2016).

Crystallization

Native crystals of STK40 22–339 were grown in 48-well format hanging drops at 20°C in 21% PEG 4000, 5% tert-butanol, 200 mM NaCl, 0.1 M Tris-HCl (pH 7.5), 6% glycerol, 1 mM TCEP and cryoprotected using the well solution with the concentration of PEG 4000 supplemented to 30%. SeMet crystals of STK40 were grown in 48-well format hanging drops at 20°C in 18% PEG 4000, 9% tert-butanol, 0.175 M NaCl, 0.1 M Tris-HCl (pH 7.5), and 1 mM TCEP, and picked straight from the drop. Crystals were flash frozen in liquid nitrogen.

Data collection and processing

An initial native STK40 data set (2.7 Å) was collected at APS beamline 19-ID. Diffraction images were processed and scaled with iMosflm (Battye et al., 2011) and the STK40 structure was solved by molecular replacement using Phaser (in Phenix) using the protein kinase CAMKII (PDB: 2VZ6) as a search model (McCoy et al., 2007). Automatic model building was carried out by AutoBuild (Phenix) and resulted in R_{work}/R_{free} values of 0.32/0.37. This model was subsequently used as a search model for molecular replacement after reprocessing the same data set with Xia2 (Winter et al., 2013), which improved data processing and scaling over iMosflm processing. During refinement of the molecular replacement model, a data set from a crystal grown using SeMet lableled STK40 was collected at APS beamline 24-ID-E, which was used to validate the molecular replacement solution we initially obtained for the native data set.. Selenium atoms were located by generating a difference map using the CAD and FFT modules in ccp4, where all structure factors and model phases were obtained from the native set-based solution and only

anomalous difference factors were taken from the SeMet data (Potterton et al., 2003). This step identified the positions of 15 of the 24 SeMet sites (the remaining sites lie in unstructured regions of the four molecules in the asymmetric unit). Manual building and refinement of the model from the native data set was carried out in Coot using $2F_o-F_c$, F_o-F_c, and composite-omit maps. Refinement (xyz coordinates, individual B-factors, NCS restraints, TLS parameters) was carried out using Phenix refine (Afonine et al., 2012) to a final R_{work}/R_{free} of 0.27/0.30. A full data set from a third crystal, which diffracted to higher resolution (2.53 Å), was collected at APS beamline 24-ID-C, and the data were processed with Xia2. The previous working model was then used for molecular replacement, and refined as above using Phenix refine to a final R_{work}/R_{free} of 0.27/0.29 (Table 1).

Fluorescence polarization assay

Fluorescence polarization assays were performed as previously described (Uljon et al., 2016). Briefly, for the initial binding experiment, 25 nM of FITC-STK40 or Trib1 peptide was mixed with increasing concentrations (0–10.5 μM) of purified COP1 WD40 domain (386–731) in HBS-P buffer (GE Healthcare Life Sciences; catalog # BR100368) supplemented with 1 mM TCEP and 0.02% Tween20. FITC-labeled and unlabeled, competing peptides were obtained from LifeTein at a purity of >95%, as assessed by reverse-phase HPLC, and used without further purification. KHD and KHD+20 STK40 constructs used in the competition assay were purified to homogeneity. In the competition assay, concentrations of 25 nM FITC-peptide and 1500 nM COP1 WD40 (386–731) were used. Experiments were conducted in a 30 μl well volume in 384-well plates. Polarization values were read at 538 nM on a Spectramax M5 plate reader (Molecular Devices). Plots present data from three independent experiments. Error bars represent SEM.

Tandem-IP experiments

HeLa cell lines were transduced with retroviruses expressing the STK40 full length protein, and stable cell line was generated by sorting for co-expressed Interleukin-2 receptor (IL2R). For each proteomic experiment, the STK40 stable cell line, together with HeLa controls, was grown in 15 cm plates to confluency and harvested. After dounce homogenization and separation of the nuclear and cytoplasmic fractions by centrifugation, Flag immunoprecipitations were performed using Flag-conjugated agarose beads (Sigma cat #A2220) in a buffer containing 50 mM Tris pH 7.5, 150 mM NaCl, 1 mM EDTA, 0.5% NP40, and 10% glycerol supplemented with EDTA-free protease inhibitor tablets (Roche) and phosphatase inhibitor cocktail (Cell Signaling). The beads were washed three times and the immunoprecipitated protein was eluted with Flag peptide (Sigma #F3290). The eluate was then subjected to a second HA pull-down using similar conditions and HA-bound protein and its interactors were eluted with HA peptide.

Tandem Affinity Purified (TAP) samples were reduced with 10 mM DTT for 30 min at 56°C in the presence of 0.1% RapiGest SF (Waters). Cysteines were alkylated with 22.5 mM iodoacetamide for 20 min at room temperature in the dark. Samples were digested overnight at 37°C with 4 μg trypsin (Promega). Tryptic peptides were acidified and purified by batch mode reversed phase and strong cation exchange chromatography (PMID: 22535209). Purified peptides were loaded onto a precolumn (4 cm POROS 10R2, Applied Biosystems)

and eluted with an HPLC gradient (NanoAcquity UPLC system, Waters; 2%–35% B in 45 min; A = 0.2 M acetic acid in water, B = 0.2 M acetic acid in acetonitrile) (Ficarro et al., 2009). Peptides were resolved on a self-packed analytical column (50 cm Monitor C18, Column Engineering) and introduced in the mass spectrometer (Q-Exactive HF, Thermo Scientific) equipped with a Digital PicoView electrospray source platform (New Objective, ESI spray voltage = 2.8 kV). The mass spectrometer was programmed to perform data-dependent MS/MS on the ten most abundant precursors in each MS1 scan using higher energy dissociation (HCD with 30% normalized collision energy). MS spectra were converted into a Mascot generic file format (.mgf) using multiplierz scripts (Parikh et al., 2009). Search parameters included trypsin specificity with up to two missed cleavages, fixed carbamidomethylation (C, +57 Da) and variable oxidation (M, +16 Da). Precursor and product ion mass tolerance mass tolerances were set to 15 ppm and 0.02 Da, respectively. The search databases consisted of human protein sequences (downloaded from RefSeq on 07/11/2011) and protein sequences for common lab contaminants both appended to their own decoy database. Sequence matches to the decoy databases were used to implement a global 1% false discovery rate (FDR) filter for the resulting peptide identifications. A fast peptide matching algorithm was used to map peptide sequences to all possible human genes (Askenazi et al., 2010). We discarded candidate proteins that were detected in a large compendium of negative TAP controls with a frequency greater than 1% (Rozenblatt-Rosen et al., 2012). Three sets of negative control and STK40 were independently purified from cytoplasmic and nuclear extracts and analyzed by LC-MS/MS. Tryptic peptides for one pair of negative control and STK40 samples purified form cytoplasmic extracts were analyzed in two replicate LC-MS/MS runs.

293T Flag-IPs

293T cells were grown in standard medium in 10 cm dishes. Cells were transfected with a pcDNA3.0 vector expressing the Flag-STK proteins (Invitrogen) and a pTriEx1.1 vector for the COP1 proteins (Novagen). An anti-Flag immunoprecipitation was performed as above and the resulting lysates and eluates were run on a 4–20% gradient gel blotted with anti-Flag (Sigma #A8592) or anti-COP1 (Abcam 56400) antibodies.

Differential Scanning Fluorimetry (DSF)

Purified proteins [STK40 KHD (22–339), Trib1 84–342 and MELK 1–342] were added to a buffer containing 50 mM Hepes, pH 7.4, 250 mM NaCl, 5% glycerol, 0.02% P20 surfactant, 5 mM $MgCl_2$ and 2 mM TCEP to a final protein concentration of 5 μM. The thermal melt was performed using a 7500 Fast Real-time PCR System (Applied Biosystems) in 96 well plate format in 20 μl volumes with sypro orange dye (ThermoFisher) and increasing concentrations of ATP or staurosporine as indicated. Thermal Shift Software was used for data analysis. Column bars present data from at least three independent experiments. Error bars represent standard deviation.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank beamline staff at the Advanced Photon Source for their technical assistance and support. We thank SBGrid for computing support and assistance with data processing. We thank Richard Gildea and Greame Winter for their assistance in modifying the xia2 software to enable prcessing of the initial STK40 native data set. We thank Michael Eck and Andrew Kruse for assistance with structure refinement and the members of the Blacklow lab, particularly Brandon Zimmerman, for helpful discussions and review of the manuscript. This work was supported in part by NIH awards R01 CA092433 (to SCB), P50 GM107618 (to SCB), and K08 CA166227 (to SU). This work is based upon research conducted at the Northeastern Collaborative Access Team beamlines, which are funded by the National Institute of General Medical Sciences from the National Institutes of Health (P41 GM103403). The Pilatus 6M detector on 24-ID-C beam line is funded by a NIH-ORIP HEI grant (S10 RR029205). This research used resources of the Advanced Photon Source, a U.S. Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under Contract No. DE-AC02-06CH11357.

REFERENCES

Adelmant G, Calkins AS, Garg BK, Card JD, Askenazi M, Miron A, Sobhian B, Zhang Y, Nakatani Y, Silver PA, et al. DNA ends alter the molecular composition and localization of Ku multicomponent complexes. Molecular & cellular proteomics : MCP. 2012; 11:411–421. [PubMed: 22535209]

Afonine PV, Grosse-Kunstleve RW, Echols N, Headd JJ, Moriarty NW, Mustyakimov M, Terwilliger TC, Urzhumtsev A, Zwart PH, Adams PD. Towards automated crystallographic structure refinement with phenix.refine. Acta Crystallogr D Biol Crystallogr. 2012; 68:352–367. [PubMed: 22505256]

Askenazi M, Marto JA, Linial M. The complete peptide dictionary--a meta-proteomics resource. Proteomics. 2010; 10:4306–4310. [PubMed: 21082763]

Battye TG, Kontogiannis L, Johnson O, Powell HR, Leslie AG. iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. Acta Crystallogr D Biol Crystallogr. 2011; 67:271–281. [PubMed: 21460445]

Ficarro SB, Zhang Y, Lu Y, Moghimi AR, Askenazi M, Hyatt E, Smith ED, Boyer L, Schlaeger TM, Luckey CJ, et al. Improved electrospray ionization efficiency compensates for diminished chromatographic resolution and enables proteomics analysis of tyrosine signaling in embryonic stem cells. Analytical chemistry. 2009; 81:3440–3447. [PubMed: 19331382]

Hammaren HM, Virtanen AT, Silvennoinen O. Nucleotide-binding mechanisms in pseudokinases. Bioscience reports. 2016; 36:e00282.

Holm L, Rosenstrom P. Dali server: conservation mapping in 3D. Nucleic acids research. 2010; 38:W545–W549. [PubMed: 20457744]

Huang J, Teng L, Liu T, Li L, Chen D, Li F, Xu LG, Zhai Z, Shu HB. Identification of a novel serine/threonine kinase that inhibits TNF-induced NF-kappaB activation and p53-induced transcription. Biochemical and biophysical research communications. 2003; 309:774–778. [PubMed: 13679039]

Kannan N, Haste N, Taylor SS, Neuwald AF. The hallmark of AGC kinase functional divergence is its C-terminal tail, a cis-acting regulatory module. Proceedings of the National Academy of Sciences of the United States of America. 2007; 104:1272–1277. [PubMed: 17227859]

Krissinel E, Henrick K. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. Acta Crystallogr D Biol Crystallogr. 2004; 60:2256–2268. [PubMed: 15572779]

Labesse G, Gelin M, Bessin Y, Lebrun M, Papoin J, Cerdan R, Arold ST, Dubremetz JF. ROP2 from Toxoplasma gondii: a virulence factor with a protein-kinase fold and no enzymatic activity. Structure. 2009; 17:139–146. [PubMed: 19141290]

Li L, Sun L, Gao F, Jiang J, Yang Y, Li C, Gu J, Wei Z, Yang A, Lu R, et al. Stk40 links the pluripotency factor Oct4 to the Erk/MAPK pathway and controls extraembryonic endoderm differentiation. Proceedings of the National Academy of Sciences of the United States of America. 2010; 107:1402–1407. [PubMed: 20080709]

Manning G, Whyte DB, Martinez R, Hunter T, Sudarsanam S. The protein kinase complement of the human genome. Science. 2002; 298:1912–1934. [PubMed: 12471243]

Martis PC, Whitsett JA, Xu Y, Perl AK, Wan H, Ikegami M. C/EBPalpha is required for lung maturation at birth. Development. 2006; 133:1155–1164. [PubMed: 16467360]

McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. Phaser crystallographic software. J Appl Crystallogr. 2007; 40:658–674. [PubMed: 19461840]

Murphy JM, Nakatani Y, Jamieson SA, Dai W, Lucet IS, Mace PD. Molecular Mechanism of CCAAT-Enhancer Binding Protein Recruitment by the TRIB1 Pseudokinase. Structure. 2015; 23:2111–2121. [PubMed: 26455797]

Murphy JM, Zhang Q, Young SN, Reese ML, Bailey FP, Eyers PA, Ungureanu D, Hammaren H, Silvennoinen O, Varghese LN, et al. A robust methodology to subclassify pseudokinases based on their nucleotide-binding properties. The Biochemical journal. 2014; 457:323–334. [PubMed: 24107129]

Parikh JR, Askenazi M, Ficarro SB, Cashorali T, Webber JT, Blank NC, Zhang Y, Marto JA. multiplierz: an extensible API based desktop environment for proteomics data analysis. BMC bioinformatics. 2009; 10:364. [PubMed: 19874609]

Potterton E, Briggs P, Turkenburg M, Dodson E. A graphical user interface to the CCP4 program suite. Acta Crystallogr D Biol Crystallogr. 2003; 59:1131–1137. [PubMed: 12832755]

Rozenblatt-Rosen O, Deo RC, Padi M, Adelmant G, Calderwood MA, Rolland T, Grace M, Dricot A, Askenazi M, Tavares M, et al. Interpreting cancer genomes using systematic host network perturbations by tumour virus proteins. Nature. 2012; 487:491–495. [PubMed: 22810586]

Scheeff ED, Eswaran J, Bunkoczi G, Knapp S, Manning G. Structure of the pseudokinase VRK3 reveals a degraded catalytic site, a highly conserved kinase fold, and a putative regulatory binding site. Structure. 2009; 17:128–138. [PubMed: 19141289]

Taccioli C, Garofalo M, Chen H, Jiang Y, Tagliazucchi GM, Di Leva G, Alder H, Fadda P, Middleton J, Smalley KJ, et al. Repression of Esophageal Neoplasia and Inflammatory Signaling by Anti-miR-31 Delivery In Vivo. Journal of the National Cancer Institute. 2015; 107

Taylor SS, Zhang P, Steichen JM, Keshwani MM, Kornev AP. PKA: lessons learned after twenty years. Biochimica et biophysica acta. 2013; 1834:1271–1278. [PubMed: 23535202]

Uljon S, Xu X, Durzynska I, Stein S, Adelmant G, Marto JA, Pear WS, Blacklow SC. Structural Basis for Substrate Selectivity of the E3 Ligase COP1. Structure. 2016; 24:687–696. [PubMed: 27041596]

Winter G, Lobley CM, Prince SM. Decision making in xia2. Acta Crystallogr D Biol Crystallogr. 2013; 69:1260–1273. [PubMed: 23793152]

Xu N, Meisgen F, Butler LM, Han G, Wang XJ, Soderberg-Naucler C, Stahle M, Pivarcsi A, Sonkoly E. MicroRNA-31 is overexpressed in psoriasis and modulates inflammatory cytokine and chemokine production in keratinocytes via targeting serine/threonine kinase 40. Journal of immunology. 2013; 190:678–688.

Yu H, He K, Li L, Sun L, Tang F, Li R, Ning W, Jin Y. Deletion of STK40 protein in mice causes respiratory failure and death at birth. The Journal of biological chemistry. 2013; 288:5342–5352. [PubMed: 23293024]

Yu H, He K, Wang L, Hu J, Gu J, Zhou C, Lu R, Jin Y. Stk40 represses adipogenesis through translational control of CCAAT/enhancer-binding proteins. Journal of cell science. 2015; 128:2881–2890. [PubMed: 26065429]

Zhang J, Zhang J, Zhao C, Shen R, Guo X, Li C, Ling X, Liu C. Analysis of transcription factor Stk40 expression and function during mouse pre-implantation embryonic development. Molecular medicine reports. 2014; 9:535–540. [PubMed: 24276375]

Zheng J, Knighton DR, ten Eyck LF, Karlsson R, Xuong N, Taylor SS, Sowadski JM. Crystal structure of the catalytic subunit of cAMP-dependent protein kinase complexed with MgATP and peptide inhibitor. Biochemistry. 1993; 32:2154–2161. [PubMed: 8443157]

Highlights

- STK40 binds the COP1 WD40 domain using a VPD/E motif in its C-terminal tail
- X-ray structure reported for the kinase homology domain of STK40
- STK40 has a partially occluded active site and lacks a DFG motif
- The STK40 kinase homology domain does not bind ATP

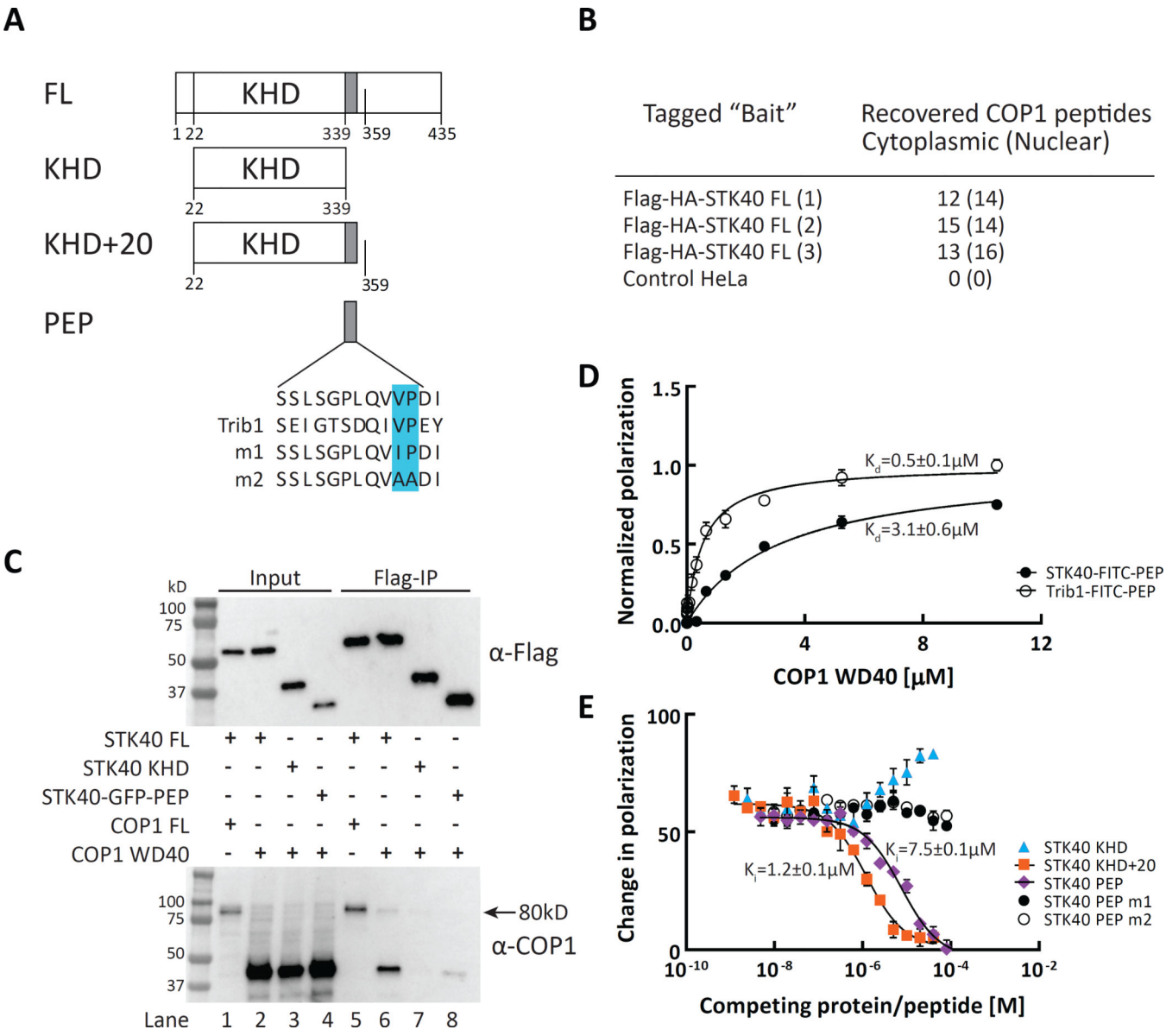


Figure 1. STK40 binds to COP1 in human cells

(A) Schematic diagrams of human STK40 constructs used in the accompanying binding studies. The putative COP1 binding motif in full-length STK40 (SSLSGPLQVVPDI) is aligned with that of Trib1 with the critical VP colored in cyan. (B) Results of tandem immunoprecipitation experiments in HeLa cells using STK40 as "bait." The number of unique endogenous COP1 peptides recovered from the cytoplasmic fraction is listed for each of three independent experiments. Endogenous COP1 peptides recovered from the nuclear fraction are in parentheses. (C) Flag-STK40 immunoprecipitations carried out in 293T cells. N-terminal Flag-tagged FL, KHD and STK40 proteins were recovered using anti-Flag resin. A Flag-GFP-fused peptide derived from STK40 ("PEP," amino acids 339–351) was also recovered using anti-FLAG resin. The cleared lysates (Input) and the immunoprecipitates (Flag IP) were western blotted and probed with anti-Flag (top) and anti-COP1 antibodies (bottom). The position of endogenous COP1 is indicated by a black arrow. (D) Fluorescence

polarization assay measuring the binding affinity of the human COP1 WD40 β propeller (386–731) for the STK40 peptide FITC-SSLSGPLQVVPD. The change in fluorescence polarization is plotted as a function of COP1 protein concentration. Trib1 peptide (FITC-SEIGTSGQIVPEY) is included for comparison. (E) Competition assay comparing binding affinities of STK40 proteins, including the KHD, KHD+20 and STK40 peptides (wild-type PEP and the two STK40 VP mutants labeled m1, m2) to the COP1 WD40 domain. Displacement of the consensus STK40 peptide FITC-SSLSGPLQVVPD was monitored using fluorescence polarization. The change in polarization is plotted for the indicated competitor as a function of unlabeled competitor concentration.

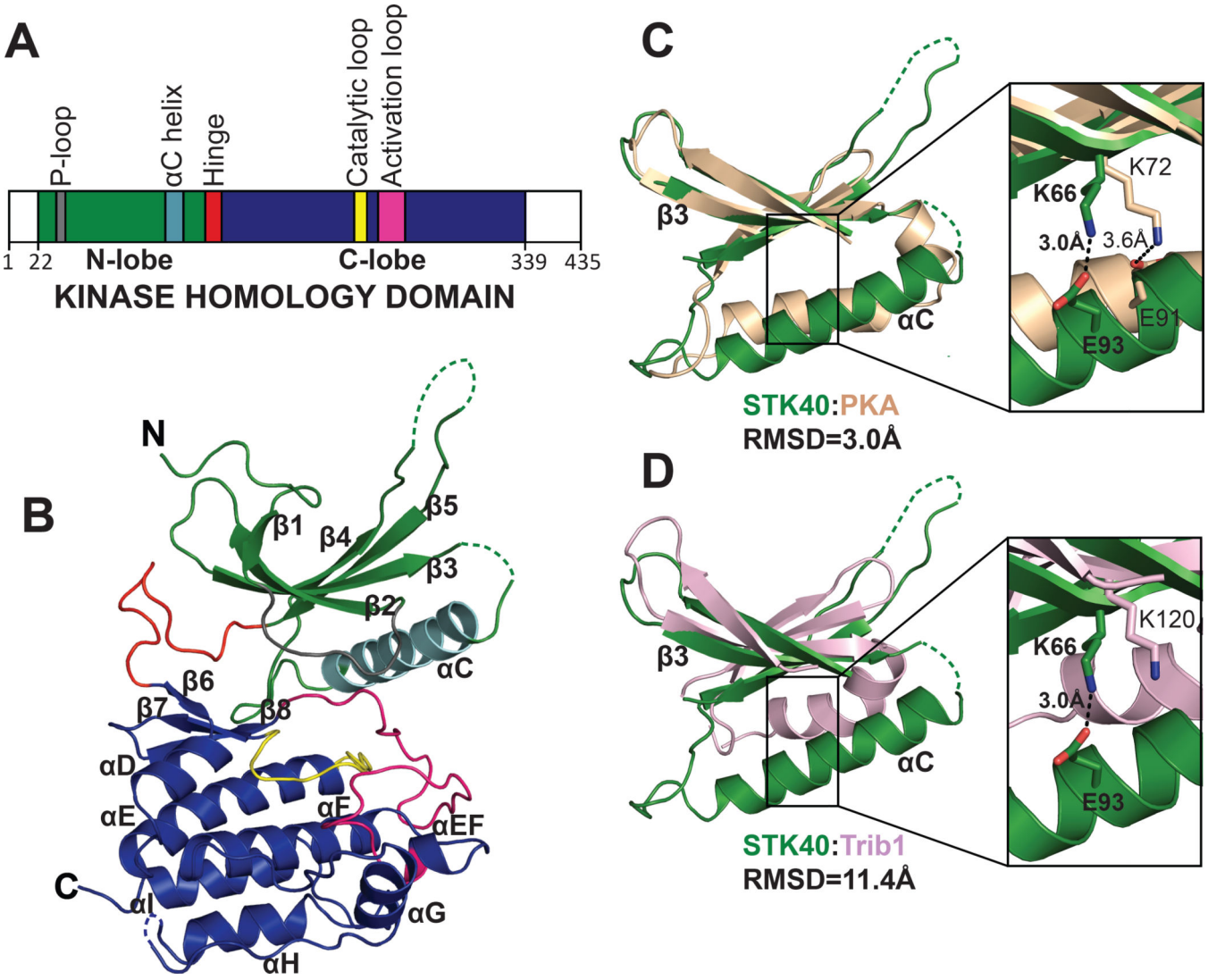


Figure 2. Crystal structure of STK40

(A) Domain organization of STK40 full-length protein with a short N-terminus (1–21), serine/threonine kinase homology domain (KHD, 22–339), and C-terminus (340–435). The crystal structure includes the entire KHD. The N-lobe is colored in green and the C-lobe is in blue (aa 156–336). Within the N-lobe, the glycine-rich P-loop is colored in grey (40–47), the αC helix is in cyan (84–98), and the hinge region is in red (142–155). Within the C-lobe, the catalytic loop is colored in yellow (191–202) and the activation loop is in magenta (215–243). (B) Ribbon diagram of the crystal structure of STK40 KHD using the same color scheme. (C,D) Least squares superposition (LSQ) of STK40 N-lobe (green, residues 35–71) on the N-lobe of PKA (wheat, residues 42–75) and on the N-lobe of Trib1 (light pink, residues 94–123). Coot least squares alignment (LSQ) was used for superpositions. RMSD values are listed.

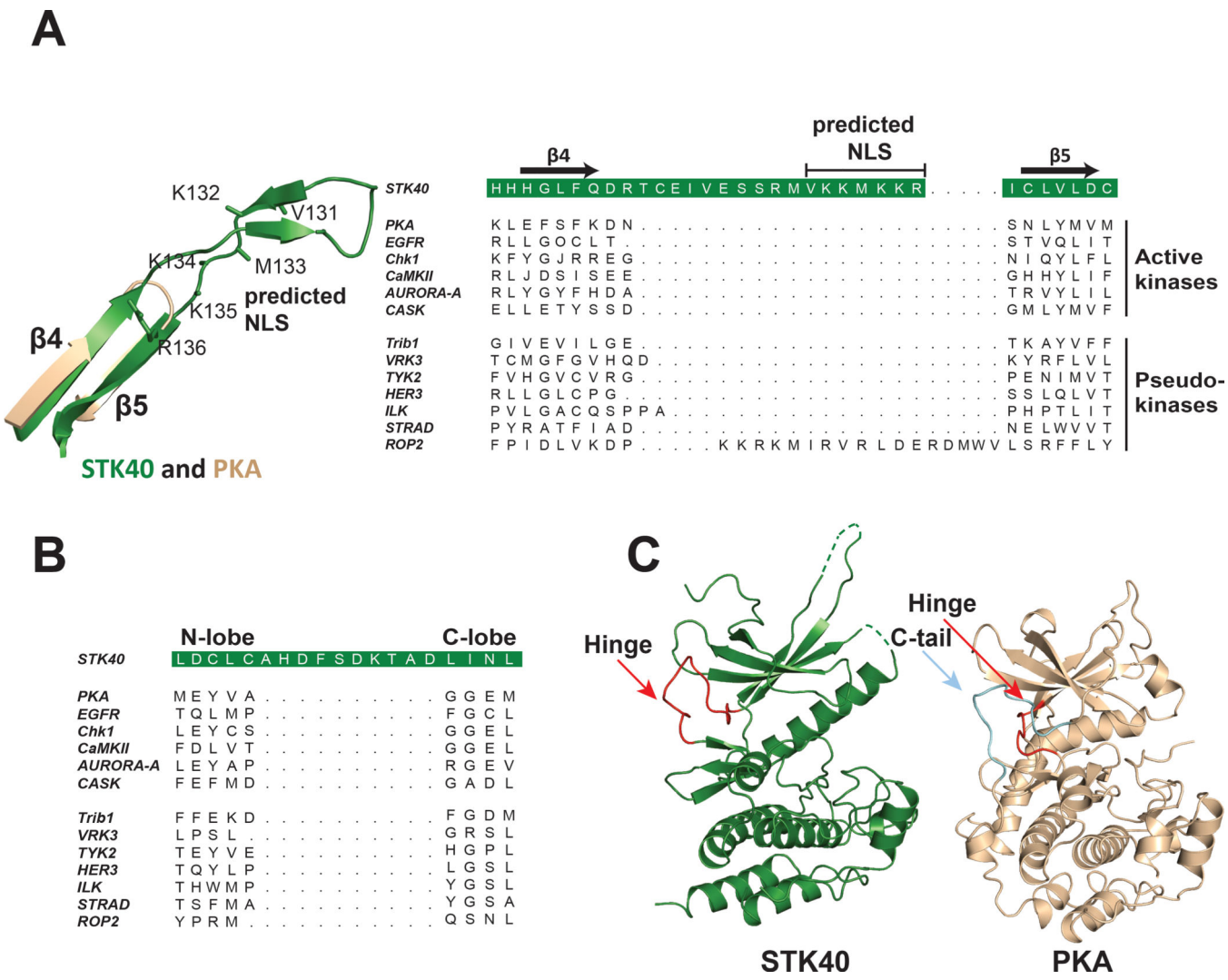


Figure 3. Atypical sequence insertions in STK40 structure

(A) Structural and sequence alignment of the β4–β5 loop extension. Left: STK40 N-lobe (chain D) β4–β5 loop (green) is aligned to the PKA short connecting loop (wheat), with residues of the predicted NLS represented as sticks. Right: Sequence insertion of STK40 aligned to selected active kinases: PKA (pdb code:1ATP), EGFR (1M14), Chk1 (1IA8), CaMKII (2VZ6), Aurora-A (2WTV), CASK (3MFS) and pseudokinases: Trib1 (5CEK), VRK3 (2JII), TYK2 (3ZON), HER3 (4RIW), ILK (3REP), STRADα (3GNI), ROP2 (2W1Z). Brackets indicate predicted nuclear localization signal (NLS). (B) Sequence alignment of hinge loop insertion. Residues flanking the unique hinge insertion in STK40 are aligned to analogous residues in known structures as in (A). (C) Structure of the hinge region of STK40 (red arrow) compared to the hinge region (red arrow on PKA) and the C-terminal tail of PKA (cyan). Coot secondary structure alignment (SSM) was used for superpositions.

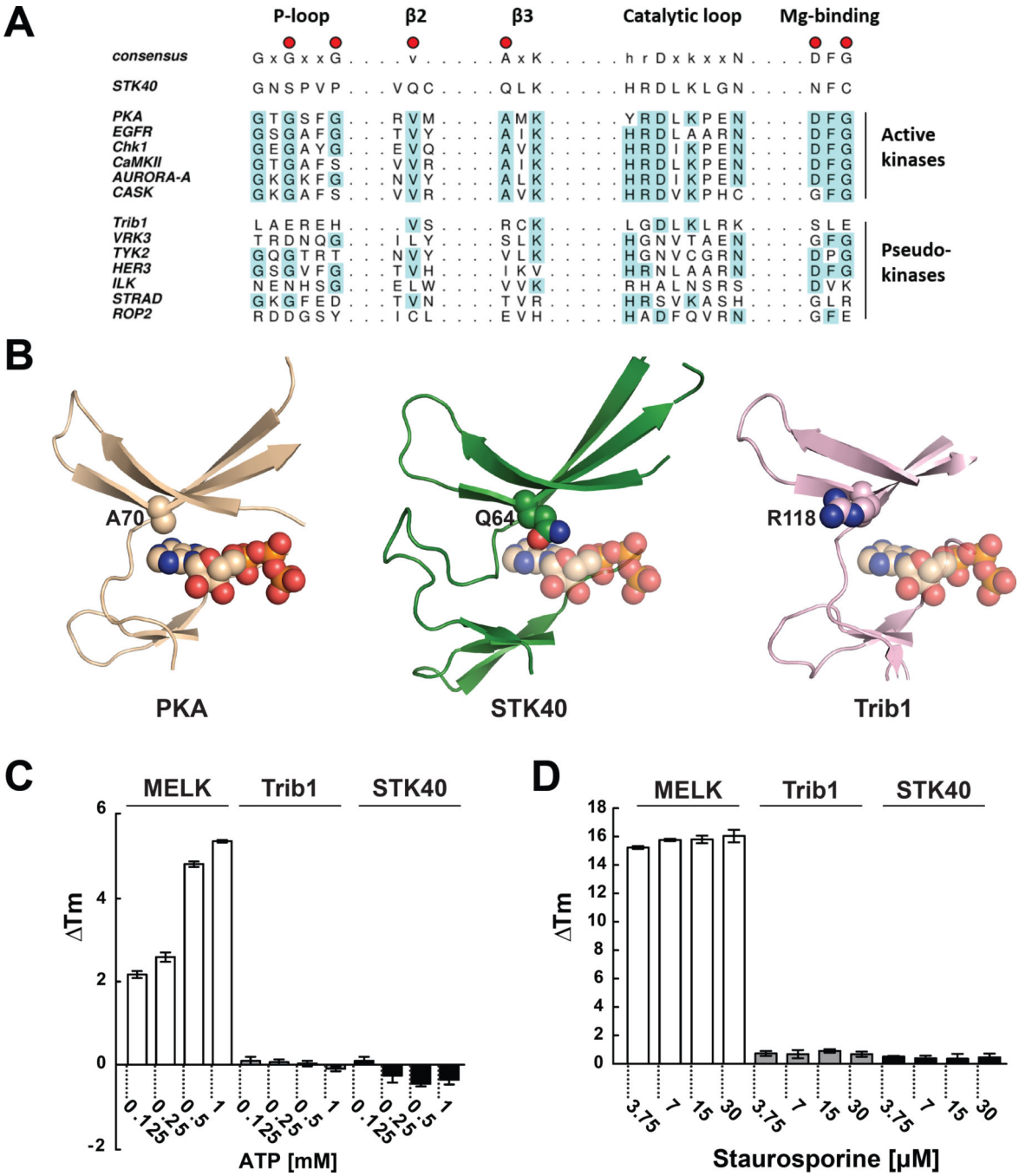


Figure 4. Degradation of the STK40 ATP binding pocket

(A) Multiple sequence alignment of canonical kinase motifs. The STK40 sequence (bottom) is aligned to the active kinases PKA, EGFR, Chk1, CamKII, Aurora-A, and CASK, and the pseudokinases: Trib1, VRK3, TYK2 pseudokinase domain, HER3, ILK, STRADα, and ROP2. Identical residues are highlighted in cyan. Divergent residues in STK40 are indicated by red dots. (B) Comparison of ATP binding pockets among PKA (wheat), STK40 (green), and Trib1 (light pink). Superposition of STK40 and Trib1 on PKA was performed using SSM in Coot. The position of ATP in STK40 and Trib1 is modeled based on its position in

the aligned, ATP-bound structure of PKA (PDB code 1ATP). ATP, A70 (PKA), Q64 (STK40) and R118 (Trib1) are represented as spheres. (C–D) Thermal shift assay. (C) Shift in melting temperature (ΔTm) of the STK40 kinase-homology domain (aa 22–339, black bars), MELK kinase (aa 1–342, white bars), and the Trib1 kinase-homology domain (aa 84–342, grey bars) upon addition of ATP as ligand. Ligand concentration varies from 0–1 mM as indicated on the x-axis. (D) Thermal shift using the ATP analogue staurosporine, titrated from 0–30 μM. Graphs were produced in GraphPad Prism, plotting the mean and standard deviation of T_m from three independent experiments.

Table 1

Data collection and refinement statistics

	Native STK40	SeMet STK40
Data collection		
Space group	P 1 2_1 1	P 1 2_1 1
Cell dimensions		
a, b, c (Å)	55.55, 237.65, 55.61	56.10, 241.91, 56.20
α, β, γ (°)	90.00, 106.32, 90.00	90.00, 104.78, 90.00
Resolution (Å)[a]	59.41−2.53 (2.62−2.53)	52.93−3.15 (3.26−3.15)
Wavelength	0.98	0.98
R_{merge}[b]	0.06 (0.78)	0.16 (0.58)
$CC_{1/2}$	0.99 (0.60)	0.98 (0.77)
I/σ(I)	12.52 (1.49)	7.58 (2.19)
Completeness (%)	98.59 (99.21)	98.39 (99.16)
Redundancy	3.40 (3.50)	3.80 (3.80)
Refinement		
Resolution (Å)	59.41−2.53	
No. of unique reflections	45250 (4527)	
R_{work} / R_{free}	0.27/0.29	
No. atoms	8350	
Protein	8284	
Water	66	
B-factors		
Protein		
Chain A	70.36	
Chain B	74.69	
Chain C	88.17	
Chain D	102.67	
Water	56.48	
R.m.s. deviations		
Bond lengths (Å)	0.009	
Bond angles (°)	0.94	
Ramachandran favored (%)	93.68	
Ramachandran allowed (%)	6.32	
Ramachandran outliers (%)	0.00	
Rotamer outliers (%)	1.67	
Clashscore	13.60	

[a] Values in parentheses are for highest resolution shell.

[b] A single crystal was used for each structure determination.