

Nucleotide binding, evolutionary insights and interaction partners of the pseudokinase Unc-51-like kinase 4

Franziska Preuss^{1,2&}, Deep Chatterjee^{1,2&}, Sebastian Mathea^{1,2}, Safal Shrestha³, Jonathan St-Germain⁵, Manipa Saha⁶, Natarajan Kannan³, Brian Raught⁵, Robert Rottapel^{6,7,8}, Stefan Knapp^{1,2,4*}

¹*Institute for Pharmaceutical Chemistry, Johann Wolfgang Goethe-University, Max-von-Laue-Str. 9, D-60438 Frankfurt am Main, Germany.*

²*Buchmann Institute for Molecular Life Sciences, Structural Genomics Consortium, Johann Wolfgang Goethe-University, Max-von-Laue-Str. 15, D-60438 Frankfurt am Main, Germany.*

³*Institute of Bioinformatics & Department of Biochemistry and Molecular Biology, 120 Green Street, University of Georgia, Athens, Georgia, USA 30602-7229*

⁴*German Cancer Consortium (DKTK) and Frankfurt Cancer Institute (FCI), Frankfurt am Main, Germany*

⁵*Princess Margaret Cancer Centre 200 Elizabeth Street, Toronto M5G 2C4, Canada*

⁶*Princess Margaret Cancer Centre, University Health Network, Toronto, Ontario, Canada.*

⁷*Departments of Medicine, Immunology and Medical Biophysics, University of Toronto, Toronto, Ontario, Canada*

⁸*Division of Rheumatology, St. Michael's Hospital, Toronto, Ontario, Canada.*

[&]*shared first authors*

*Correspondence: Stefan Knapp

e-mail: knapp@pharmchem.uni-frankfurt.de

phone: 0049 69 798-29871

Abstract

Unc-51-like kinase 4 (ULK4) is a pseudokinase that has been linked to the development of several diseases. Even though sequence motifs required for ATP binding in kinases are lacking, ULK4 still tightly binds ATP and the presence of the cofactor is required for structural stability of ULK4. Here we present a high-resolution structure of a ULK4-ATP γ S complex revealing a highly unusual ATP binding mode in which the lack of the canonical VAIK motif lysine is compensated by K39, located N-terminal to α C. Evolutionary analysis suggests that degradation of active site motifs in metazoan ULK4 has co-occurred with an ULK4 specific activation loop, which stabilizes the C-helix. In addition, cellular interaction studies using BioID and biochemical validation data revealed high confidence interactors of the pseudokinase and armadillo repeat domains. Many of the identified ULK4 interaction partners were centrosomal and tubulin associated proteins and several active kinases suggesting new roles for ULK4.

Keywords: Unc-51-like kinase, ULK4, pseudokinase, evolution, Bio-ID, ATP binding, interaction partners

Highlights:

Structure of the ULK4 ATP complex reveals a unique ATP binding mode.

Disease associated mutations modulate ATP binding and ULK4 stability

Degradation of active site motifs co-occurred in evolution with an ULK4 specific activation loop

BioID suggests a role of ULK4 regulating centrosomal and cytoskeletal functions

Introduction

The human genome encodes about 520 protein kinases. Structural studies revealed that protein kinases share a common topology, the canonical 'kinase fold' and large scale sequence comparison have revealed highly conserved amino acid motifs important for kinase catalytic function including (I) the glycine-rich loop, (II) the VAIK motif containing a lysine to bridge the β 3 strand to the α C helix in active kinases coordinating the Mg²⁺/triphosphate moiety of the ATP cofactor, (III) the HRD motif with the catalytically indispensable catalytic base, (IV) a conserved

asparagine prior to the $\beta 7$ strand that orients the ATP phosphates and finally (V) the DFG motif at the N-terminus of the activation loop that is important for binding of the phosphate moieties of ATP and Mg^{2+} (Manning et al., 2002). However, about 10% of kinases lack one or more of these conserved motifs, which often renders them catalytically inactive or with significantly reduced catalytic activity (Boudeau et al., 2006; Kwon et al., 2019). These kinases are therefore referred to as pseudokinases. Instead of catalyzing the phosphoryl transfer reaction, their physiological role mediating cellular signaling is scaffolding and to serve as allosteric regulators of active enzymes (recently reviewed by Jacobsen *et al.* (Jacobsen and Murphy, 2017)).

Pseudokinases are present in all major groups of the human kinome and across diverse species (Kwon et al., 2019). The structural features of pseudokinases are accordingly diverse, sharing many regulatory mechanisms observed in canonical kinases (Ha and Boggon, 2018; Jura et al., 2009; Patel et al., 2017; Scheeff et al., 2009; Shrestha et al., 2020; Zeqiraj et al., 2009b). Interestingly, many human pseudokinases have completely lost the ability to interact with nucleotides, while others still bind to and are stabilized by ATP interaction (Murphy et al., 2014). The subgroup of ATP-binding pseudokinases are thought to fulfil their physiological tasks in the nucleotide-bound form. There are several variations of how pseudokinases exert their signaling roles (Murphy et al., 2017). For instance, the transmembrane growth factor receptor HER3 regulates the MAPK and PI3K pathways through interactions with the surface of its pseudokinase C-lobe which constitutes a docking site for client proteins (Citri et al., 2003). When binding, the clients are forced into a specific conformation, resulting in client activation (Jura et al., 2009).

An intriguing example of a complex and dynamic regulatory mechanism mediated by a pseudokinase is Mixed Lineage Kinase domain-Like (MLKL), a key regulator of the non-apoptotic, kinase dependent programmed cell death called necroptosis (Murphy et al., 2013). MLKL comprises a N-terminal 4-helix bundle (4HB) domain and a C-terminal pseudokinase which lacks catalytic activity due to loss of the DFG and HRD motifs. Induction of necroptosis relies on two active kinases, the receptor interacting protein kinase RIPK1 and RIPK3 which interact with MLKL forming the necrosome. Formation of this complex results in phosphorylation of MLKL, triggering a conformational change in MLKL that unleashes the 4HB domain resulting in

membrane recruitment and formation of membrane-disrupting pores (Sun et al., 2012; Wang et al., 2014).

Unfortunately, for many pseudokinases there is no data available on the pathways they regulate and their specific interaction partners. Since pseudokinases are often deregulated in diseases, a better understanding of pseudokinase interactions and their regulatory mechanisms would be desirable also for the evaluation of their potential as targets for the development of new therapeutics (Jacobsen and Murphy, 2017; Ribeiro et al., 2019).

Unc-51-like kinase 4 (ULK4) harbors sequence variations and lacks essential residues in all 5 conserved kinase motifs that are important for catalytic activity (**Figure 1A**). In particular, the substitutions of the VAIK lysine with leucine and the HRD histidine with phenylalanine are rare events in the human kinome. In addition, the lack of the essential ATP binding motifs such as VAIK (VAIL in ULK4) and DFG (NFC in ULK4) suggests that ULK4 has compromised ATP binding. However, strong ATP binding has recently been reported by a comparative binding study using temperature shift assays (Lucet and Murphy, 2017). A recent structure of the catalytic domain in complex with weakly binding inhibitors revealed a canonical kinase fold and may provide the foundation for the development of ULK4 inhibitors that modulate ULK4 scaffolding function (Khamrui et al., 2020).

The ULK family of kinases comprises the catalytically active members ULK1, ULK2, ULK3 and STK36. The kinase domains in ULKs are located at the N-termini of all family members. Usually, the regions C-terminal to the kinase domain contain protein interaction motifs important for substrate recruitment (**Figure 1B**). The role of the C-terminal substrate recruitment domains has been demonstrated for ULK3 (Caballe et al., 2015). Its second microtubule interacting and trafficking (MIT) domain binds to the protein IST1, thus enabling the ULK3 kinase domain to phosphorylate IST1 (Caballe et al., 2015). This mechanism enables ULK3 to regulate the timing of membrane remodeling by the ESCRTIII machinery. Analogously, the ULK1 C-terminal domain-like (CTD-like) domain interacts with autophagy-related protein 13 (ATG13). When activated, the ULK1 kinase domain phosphorylates the bound ATG13, which is part of the cascade ultimately leading to autophagosome formation (Zachari and Ganley, 2017). Despite its lack of kinase activity, a similar recruitment mechanism for substrates or interacting proteins might be also relevant for ULK4 which

contains armadillo repeat protein interaction domains. However, no interaction partners for the ULK4 C-terminal armadillo repeats which are also present in C-terminus of STK36 have been thus far reported.

ULK4 has been linked to several human disorders, highlighting the developmental function of this catalytically inactive kinase scaffold protein. Single nucleotide polymorphisms (SNPs) in the ULK4 gene have been linked to blood pressure regulation and hypertension (Levy et al., 2009). As a consequence, certain ULK4 alleles increase the risk of acute aortic dissections (Guo et al., 2016). Genome wide association studies have identified a ULK4 polymorphism associated with increased risk for the development of multiple myeloma (Broderick et al., 2011). Deletion of ULK4 in mice results in hydrocephalus, a condition linked to impaired cilia development (Vogel et al., 2012). Multiple roles for ULK4 in brain development and neuronal function have been reported, including neuronal motility (Lang et al., 2014), cortex development (Lang et al., 2016), hypomyelination (Liu et al., 2018b) and GABAergic signaling (Liu et al., 2018a). However, a molecular mechanism of how ULK4 mediates these cellular signaling functions is lacking. ULK4 falls within a class of “dark” kinases for which there is little or no functional data (Oprea, 2019).

To elucidate the mechanistic basis for ULK4 function, we solved the crystal structure of the ULK4 pseudokinase domain in complex with ATP γ S to gain insights into its unique nucleotide binding properties and identified interaction partners using mass spectrometry. The structure of the ULK4 pseudokinase domain revealed a unique nucleotide-binding mode that results in tight metal ion independent binding of the co-factor, also providing a rationale for the disease-relevant polymorphisms affecting ATP binding residues. Through deep evolutionary analysis, we demonstrate that loss of catalytic activity of mammalian ULK4 occurred progressively in evolution with ancestral ULK4 orthologues in plants and protists retaining the canonical active site residues. Remarkably, the selective loss of ULK4 active site residues in mammals and metazoans correlated with the emergence of an extended activation segment that structurally compensates for the loss of the canonical salt bridge between the VAIK lysine and the α C glutamate by stabilizing the regulatory C-helix in an active conformation. Identification of ULK4 binding partners by BioID mass spectrometry suggested roles in microtubule and centrosome function. We propose specific

docking/protein-protein interaction sites by delineating the selective constraints imposed on the ULK4 catalytic domain.

Results and Discussion

Based on secondary structure predictions and sequence alignments with ULK kinases we designed several constructs for expression in *E. coli*. Good yields of soluble protein were observed for the construct containing the pseudokinase domain residues 1-288 (ULK4_{PD}). This sequence was expressed in frame with a cleavable (TEV: tobacco etch virus protease) N-terminal purification tag (His₆-TEV) allowing efficient purification of the recombinant protein. However, removal of the N-terminal tag by TEV protease was difficult, indicating that the cleavage site was not accessible. The tag was therefore not removed for crystallization and functional studies. In size-exclusion chromatography (SEC), ULK4_{PD} eluted as a monomeric protein (**Supplemental Figure S1**). The identity of ULK4_{PD} was confirmed by mass spectrometry, confirming the correct mass of the un-cleaved recombinant protein which however underwent N-terminal methionine excision. No post-translational modifications were detected. The recombinant protein that was more than 95% pure as judged by SDS PAGE readily crystallized in the presence of non-hydrolysable ATP γ S yielding crystals that diffracted to 1.9 Å resolution. Data collection statistics and refinement is summarized in **Supplemental Table S1**.

Structural features of the ULK4 pseudokinase domain

The structure of ULK4_{PD} revealed the canonical bilobal domain architecture of protein kinases (PDB-ID 6TSZ, **Figure 1C**) that was similar to a recent ULK4 structure in complex with an inhibitor (pdb-code: 6U5L) (Khamrui et al., 2020) but a larger portion of the activation segment was visible in the ATP complex. The structures of the active ULK family members ULK1 and ULK2 superimposed well with ULK4_{PD} despite the low sequence identity (29%) shared between these two kinases (Chaikuad et al., 2019; Lazarus et al., 2015) (**Figure 1D, Supplemental Figure S2**). However, major structural differences were observed in the length of the α C helix, the β 7- β 8 loop length, the length and topology of the activation loop as well as the orientation of the α G helix is also notably different (**Figure 1**). This helix does not form extensive crystal contacts in ULK

structures and the ULK4 α G helix remained stable in a molecular dynamics study (see below), suggesting that the captured conformations in the compared crystal structures are also present in solution. The glycine rich loop contains a threonine residue in ULK4 (T16) at a position where in active kinases usually a glycine is found. The introduction of this bulky residue resulted in a rotation of the β 1 strand. A similarly rotated β 1 strand in combination with a glycine rich loop threonine has been observed in structures of the TYK2, JAK1 and JAK2 pseudokinase domains (Lupardus et al., 2014). The canonical VAIK motif is replaced by VAIL in ULK4 replacing the conserved lysine residue by a leucine (L33). This substitution has not been reported for any other human kinase. In active kinase conformations, the VAIK lysine stabilizes the “in” conformation of helix α C by forming a highly conserved salt bridge with an invariant glutamate present in α C. In addition, this lysine residue is essential for coordinating the nucleotide phosphates in the ATP bound state. Structural superimposition with ULK1 showed that the α C glutamate residue is also missing in ULK4 and it has been replaced by a tryptophan (W46) (**Figure 2A**). The VAIL leucine is however too far away (4 Å) to form efficient hydrophobic interactions with W46. Instead, W46 extends the regulatory spine through interaction with F140, the central residue in the degenerated DFG motif (NFC in ULK4), most likely contributing to the stability of the observed active-like state of the pseudokinase domain. It is interesting to note that the pseudokinase domain of STK40 is the only other human kinase that also bears an NFC motif (Durzynska et al., 2017). Nevertheless, α C is in an active “in” conformation similar to the one observed in phosphorylated ULK1 (pdb code: 4WNO). Helix α C is one turn shorter than in ULK1. However, the length in α C is highly variable in kinases and may depend on the activation state (Eswaran et al., 2007). In ULK2, α C is even shorter than in ULK4 and contains a helical insert in the linker connecting α C with β 3 (**Supplemental Figure S2**). The C-terminus of α C is stabilized by interactions of the degenerated NFC motif residue F140 which assumes an NFC “in” conformation.

The catalytic loop that comprises the conserved HRD motif in canonical kinases is replaced by an FCD motif in ULK4. However, the phenylalanine (F119) within the FCD motif still anchors the catalytic loop to the kinase core through hydrophobic interactions with residues L109, L112, I122, F137 and F140. The HRD arginine residue that often links the catalytic loop to phosphorylated residues in the activation segment is replaced by a cysteine, but this position is quite variable also

in active kinases. The kinase catalytic aspartate, however, is conserved in ULK4. There is no indication that this aspartate can function as a catalytic base, but it still forms a commonly seen salt bridge to the +5 residue (K126) in the catalytic loop. This residue in the very C terminus of the catalytic loop is typically an asparagine in canonical kinases, but it is a lysine (K126) in ULK4. As a result, there is a network of salt bridges linking K126 with the co-factor ATP and D121 in the catalytic loop. To date, the only other pseudokinase for which an aspartate to lysine salt bridge in the catalytic loop has been reported is TRIB1 which maintains a similar catalytic loop conformation as observed in ULK4 (pdb code: 5CEM). The salt bridge is however tilted towards the core of TRIB1 due to a side chain flip of TRIB1 D205 (Murphy et al., 2015).

In comparison with the canonical kinase fold, the ULK4 activation segment contains a large insert with a helical structure (L150 to E161). The activation segment N-terminus contains the unusual degenerated DFG motif (NFC) followed by a typical short sheet region that anchors that activation segment to the lower kinase lobe by hydrogen bonds formed by main chain interactions between two antiparallel short β -sheets. At the N-terminus of the short sheet structure there is a large helix inserted that forms an intricate network of salt bridges with α C possibly stabilizing its inward oriented conformation. The interaction with α C is further stabilized by hydrophobic interactions that form a small hydrophobic core involving residues F153, F154 and L150 at the center of this activation segment helix and P41, L49 located in α C. A salt bridge links E148 located in the loop N-terminal to the activation segment helix with α C R48. The hydrophobic environment surrounding this salt bridge suggest a strong polar interaction of the R48/E148 salt bridge (**Figure 2B**). The helix is connected to a coil structure bypassing the catalytic loop. Intriguingly, S182 located in the activation segment 5 residues N-terminal to the APE helix is positioned opposite the catalytic base of the degenerated HRD motif (FCD) mimicking a substrate bound state. Thus, the substrate binding site in ULK4 is blocked by the activation segment. The activation segment terminates with a canonical APE motif.

Degradation of the active site has co-evolved with an extended activation segment

We performed phylogenetic analysis of ULK4 orthologs from diverse organisms to investigate the origin and evolution of ULK4 as a pseudokinase. We identified ULK4 orthologs from diverse taxonomic groups including metazoans, plants and protists. Phylogenetic analysis of these

sequences reveals clear separation of metazoan ULK4 from plants and protists (**Supplemental Figure S3**). Further annotation of the sequences based on the presence or absence of the canonical active residues indicates species specific variations in the canonical active site residues namely K72, E91, D166 and D184 (Protein Kinase A numbering). While metazoan ULK4s lack three of these canonical residues except for the catalytic aspartate (D166) (cyan circular stripes in **Supplemental Figure S3**), orthologues in plant and protists conserve these canonical residues with few exceptions, suggesting that loss of catalytic function of ULK4 emerged later in evolution, presumably for developing specialized metazoan ULK4 pseudokinase functions. In addition to divergence in the canonical active site residues, metazoan ULK4's displayed variations in the activation loop as well. Comparison of the activation loop lengths across ULK4 orthologues indicates an extended activation segment (mean length of 49.6 amino acids) in pseudo-metazoans compared to ULK4 orthologues from other species (**Figures 2C, S3 and S4**).

The extended segment (L150-E161) forms a helical insert in human ULK4 and packs against the C-helix in the crystal structure. In order to test the stability of this conformation, we performed molecular dynamics (MD) simulation of human ULK4 in the presence of ATP (PDB code: 6TSZ). The simulation indicates that while the inserted helical segment (L150-V157) stably packs against the C-helix, the rest of the loop (A158-E161) is flexible as shown by the lack of secondary structure propensity during the course of simulation (**Figure 2D**). A helical activation loop has also been reported for mouse MLKL (PDB code: 4BTF) (Murphy et al., 2013). This helical segment assumes however a different orientation compared to the ULK4 activation helix. In the mouse MLKL structure, the α C helix is moved outwards making room for the MLKL activation segment helix which resides in a similar position as the ULK4 α C helix. It is interesting to note that in human MLKL, no helical insert has been reported. These data suggest that α C and the activation loop helix may serve pseudokinase specific functions that may even differ between species as observed for MLKL (Petrie et al., 2018). Our evolutionary analysis suggests that the loss of lysine-glutamate salt bridge in metazoan ULK4 required the stabilization of α C by the activation segment helix which in metazoan ULK4 has co-occurred with an extended activation loop, presumably compensating for the loss of the canonical K-E salt bridge. However, since the motif 'LEEFFALVAAEE' forming

the helical insert in human ULK4 is present in metazoan ULK4 pseudokinases, it is likely that the helical activation loop structure is a common feature of ULK4 sequences that lack the β 3-Lysine.

Evolutionary constraints distinguishing ULK4 from other paralogs.

We next performed a Bayesian statistical analysis to identify the evolutionary constraints distinguishing ULK4 from the closely related ULK1-3 and STK36 sequences. This revealed strong ULK4-specific constraints imposed on residues in the N-terminal ATP binding lobe of the pseudo kinase domain. Specifically, these constraints map to residues in the β 2- β 3 loop as well as in β 4 and β 5 strands. Some of these constrained residues include R22, R23, and K24 that form a solvent exposed surface patch with hydrophobic residues I27 and F29, which are also ULK4-specific (**Figure 3A**). In addition, an extended network of ULK4 specific residues (F61, W64, E66, and L71) structurally couple the surface patch to the C-helix. These residues are conserved across diverse ULK4 orthologs (including plants and protists) and divergent from the closely related ULK1-3 and STK36 (**Figure 3B**), implying that they are constrained for important ULK4-specific functions. A possible role of this ULK4-specific surface patch could be that it serves as a docking site similar to the one reported for Aurora A with the microtubule associated protein TACC3 (Burgess et al., 2018).

ULK4 has an unusual ATP binding mode

Because conserved motifs known to be essential for nucleotide binding are absent in ULK4, we were particularly interested in how ULK4 interacts with the co-factor ATP. After purification of the recombinant protein the absorbance spectra of ULK4_{PD} showed an unusual ratio at 260 and 280 nm (A_{260}/A_{280}) suggesting co-purification of nucleotides. Most proteins exhibit maxima at 280 nm due to the presence of tryptophan residues, with A_{260}/A_{280} ratio usually ranging from 0.45 to 0.55. In ULK4_{PD}, this ratio was 0.76 (**Figure 4A**). An explanation for an aberrant A_{260}/A_{280} ratio might be the amino acid composition of the protein. This is not the case in ULK4_{PD} which contains 2.0% tryptophan residues. Another explanation might be the covalent attachment of a co-factor such as ADP-ribose but this scenario can be excluded because any covalent modification would have been detected in our mass spectrometry analysis. We therefore assume that ULK4_{PD} tightly, but non-covalently, bound to a nucleotide metabolite from *E. coli*, and that this interaction was still present at least partially after dialysis and several chromatographic purification steps.

In order to assess the durability of ATP binding we performed a Molecular Dynamics (MD) simulation of the crystal structure of ULK4 with ATP bound to it. The co-factor remained stably bound during this 350 ns simulation supporting the high affinity of the co-factor (**Supplemental Figure S5**). A movie illustrating structural changes is also available in the supplemental material.

In order to identify nucleotides that are most likely bound to ULK4 in cells, we performed a co-factor screen using the most common nucleotide-based co-factors and their counter ions. For this analysis we performed thermal shift assays, a screening technology that is well established in our laboratory and that detects molecular interactions based on their effects on protein stability (Fedorov et al., 2012). Recombinant ULK4_{PD} exhibited a melting temperature (T_M) of 30°C. In comparison to most kinases (40°C < T_M < 60°C) ULK4_{PD} was highly unstable after purification. Both adenine and guanine nucleotides strongly stabilized ULK4_{PD}. By far the highest T_M shift was observed using ATP (20°C), while the ADP stabilization was 9°C. This highlighted the importance of the ATP γ -phosphate for binding. ADP derivatives such as ADP-ribose and acetyl-CoA stabilized to a similar extent as ADP alone, indicating that the substituents did not interfere strongly with binding (**Figure 4B**). Interestingly, Mg²⁺ ions did not promote nucleotide binding in accordance with previously published results (Murphy et al., 2014). In contrast, the presence of 5 mM Mg²⁺ ions significantly destabilized the ULK4_{PD}-ATP complex by 8°C. The pseudokinases calcium/calmodulin-dependent serine protein kinase (CASK), STE20 Related Adaptor Beta (STRAD β), Mixed Lineage Kinase Domain Like Pseudokinase (MLKL) and Ephrin Receptor B6 (EphB6) all lack a functional DFG motif and show metal independent co-factor binding (Bailey et al., 2015; Mukherjee et al., 2008; Murphy et al., 2014). Since ATP is highly abundant in cells, it is likely that the physiological ligand of ULK4_{PD} is indeed ATP. We therefore used the highly stabilizing non-hydrolysable ATP analogue ATP γ S for crystallization studies. Screening of a kinase targeted library of 1500 compounds using DSF resulted only in small temperature shifts for some inhibitors (data not shown) suggesting weak interaction with typical kinase inhibitors.

An intriguing aspect of the ULK4_{PD}-ATP γ S structure was the unprecedented binding mode of the nucleotide. In an active kinase, both the ATP orientation and the triphosphate moiety need to be adjusted perfectly to allow for catalysis. For instance, in the first structure of a kinase ATP complex, protein kinase A (PKA, PDB code: 1ATP), the co-factor binds two Mg²⁺ ions coordinated by the

ATP phosphates, the DFG aspartate (D184), the catalytic base in the HRD motif (D166) and N171 and K168 located in the catalytic loop. The glycine rich loop interacts closely with the ATP phosphates forming several main chain interactions. The adenine ring is anchored to the hinge region by two conserved hydrogen bonds with main chain atoms of E121 and V123 (**Figure 4C**). While some interactions between ULK4 and ATP γ S are inspired by the kinase ATP binding mode, other interactions have not been described in active kinases and pseudokinase co-factor complexes before (**Figure 4D**). Similar to active kinases, the ATP adenine is anchored to the ULK4 hinge main chain by hydrogen bonding to E76 and C78. Hydrophobic interactions stabilize the adenine ring system which is sandwiched between hydrophobic residues in the β 2 strand (V18), the VAIL motif (A31 and L33) as well as the β 7 strand (L128). Intriguingly, the lack of the VAIK motif lysine is compensated by K39, capping the α C helix. The ATP γ S triphosphate is embedded into an intricate hydrogen-bonding network involving side chains from the tip of α C helix (K39), the catalytic loop (R125 and K126) and the DFG motif (N139). As a result, the ATP γ S molecule decorates the back of a groove in ULK4_{PD} with the ribose hydroxyl groups and several phosphate oxygen atoms exposed. A similar groove is observed in the pseudokinase STRAD α . Binding of both ATP and the allosteric activator protein MO25 allegedly forces STRAD α into an active-like conformation (**Figure 4E**). The ternary complex then binds to and activates the client kinase STK11/LKB1 (Serine/Threonine Kinase 11/Liver Kinase B1) (Zeqiraj et al., 2009b). In this process, the role of ATP is solely to maintain STRAD α in its binding-competent form. The ATP molecule interacted directly neither with the client kinase nor with the activator.

Recently, genome-wide association studies (GWAS) linked several single-nucleotide polymorphisms (SNPs) in ULK4 to diseases such as high blood pressure (Ehret et al., 2016; Levy et al., 2009) and sporadic thoracic aortic dissection (STAD) (Guo et al., 2016). Intriguingly, SNP rs2272007 is localized in the pseudokinase domain of ULK4 affecting residue 39. As described above, in the most common ULK4 variant, K39 takes over the role of the missing VAIK lysine. According to “The 1000 Genomes Project”, 68% of sequenced individuals carry the nucleotide T in rs2272007 and thus express ULK4 protein with a lysine in position 39, and 32% carry the nucleotide C and express an arginine in this position. We were therefore interested in how the R39 polymorphism could affect ATP binding, and generated a site-directed mutant carrying this

mutation. We also produced another reported variant: N139L. This variant protein was unstable during purification and precipitated quickly, precluding analysis. However, since N139 replaces the aspartate in the degenerated ULK4 DFG motif and the N139 side chain form intricate interactions with the ATP alpha and beta phosphates, the mutation into a hydrophobic lysine is likely to severely compromise ATP binding affecting structural stability of the ULK4 pseudokinase domain. Temperature shift experiments using the R39 variant showed that the un-ligated protein has similar stability as the K39 protein, and the presence of 1 mM ATP resulted in similar temperature shifts suggesting that both lysine and arginine can form efficient salt bridges to the co-factor and that there was enough space in the ATP binding pocket to accommodate the slightly bulkier arginine side chain. Interestingly, a significant difference was observed in ADP binding. While ULK4_{PD} K39 was stabilized by only 9°C, ULK4_{PD} R39 was stabilized by 13°C, suggesting tighter binding of ADP. It is likely that the side chain of K39 is too short to form efficient salt bridges with the ADP β-phosphate, but R39 has sufficient length to form more efficient contacts (**Figure 4F**). Thus, ULK4 K39 and R39 differed in their affinities for ADP whereas the variant N139L was highly unstable due to compromised ATP binding highlighting the importance of ATP binding for ULK4 function and making a compelling case that altered nucleotide affinity is possibly responsible for disease development as suggested by GWAS studies.

BioID identifies interaction partners of ULK4

To date, no validated ULK4 interacting partners have been reported. Knockdown studies revealed a role in neural cell migration pointing to a cytoskeletal function as well as stimulating role on MAPK signaling (Lang et al., 2014). To better understand the ULK4 interaction landscape, we performed BioID (proximity-dependent biotin identification) (Roux, 2013), a proximity-dependent labelling technique that identifies proteins in the vicinity of a target polypeptide in living cells. The technique makes use of a mutant form of the *E. coli* biotin ligase enzyme BirA* (R118G), which is fused to a protein of interest. The R118G BirA mutant is unable to maintain an interaction with the activated biotinoyl-AMP molecule. This highly reactive intermediate is thus released into the vicinity of the fused bait protein and reacts with amine groups of nearby polypeptides. Biotinylated proteins can then be purified using streptavidin, and identified by mass spectrometry (Gingras et al., 2019). N- and C-terminally FlagBirA* tagged full length ULK4 was expressed in Flp-In T-Rex 293 cells.

Immunofluorescence (anti-Flag) indicated that the tagged ULK4 protein is localized in the cytoplasm (**Supplemental Figure S6**), in agreement with data available at the human protein atlas (<https://www.proteinatlas.org>).

ULK4 BioID was performed using both N- and C-terminally tagged FlagBirA* fusion proteins, and each analysis was performed twice. Data were compared to BioID conducted on 293 Flp-In T-Rex 293 cells expressing the FlagBirA* tag alone to control for endogenously biotinylated proteins and polypeptides that interact non-specifically with the streptavidin-sepharose matrix material. Data were analyzed using the SAINT (statistical analysis of interactomes (Choi et al., 2011; Teo et al., 2014)) algorithm to identify high confidence proximity interactors (displaying a Bayesian false discovery rate ≤ 0.01 **Figure 5; Supplemental Table S2**). The proximity interactor detected with the highest number of peptides was calmodulin regulated spectrin associated protein 1 (CAMSAP1), a large cytoplasmic scaffolding polypeptide that binds microtubules via its CKK domain, and controls non-centrosomal MT minus-end dynamics (Atherton et al., 2019; Baines et al., 2009). CAMSAP family members are involved in the regulation of a number of key cellular functions, including cell polarity, regulation of neuronal differentiation and axonal regeneration (Akhmanova and Hoogenraad, 2015; Marcette et al., 2014; Pongrakhananon et al., 2018). The related family member CAMSAP3 was also identified as a ULK4 proximity interactor. Intriguingly, many tubulin binding and centrosomal proteins were also detected, including: HAUS2 and HAUS8 (HEC1/NDC80-Interacting Centrosome-Associated Protein), components of the microtubule-binding augmin/HAUS complex, which is essential for mitotic spindle assembly (Goshima et al., 2008; Uehara et al., 2009); centrosomal proteins such as CCP110 (Centriolar Coiled-Coil Protein 110), CEP97 (Centrosomal Protein 97), CSPP1 (Centrosome and Spindle Pole Associated Protein 1) and OFD1 (OFD1 centriole and centriolar satellite protein); and motor proteins of the kinesin family (KIF1B, KIF3B and KIFAP3).

Other interacting partners included several protein kinases implicated in microtubular function, such as the Microtubule Associated Serine/Threonine Kinase 2 MAST2, Rho associated coiled-coil containing protein kinase 2 (ROCK2) and Rho GTPase activating protein 29 (ARHGAP29). In addition, PTPN14 (protein tyrosine phosphatase, non-receptor type 14) and the ULK family

member STK36, an important regulator of the sonic hedgehog (Shh) pathway (Han et al., 2019; Murone et al., 2000), were also detected as high confidence proximity interactors.

Consistent with the BioID data, CAMSAP1, PTPN14, ROCK1 and ROCK2 co-migrated with full length C-terminal Flag-tagged ULK4 in analytical ultra-centrifugation density gradients, as detected by western blotting (**Figure 6A**). Co-immunoprecipitation using two different methods (C-terminal FLAG-tag IP and streptavidin pulldown) validated the ULK4 interactions with CAMSAP1, ROCK1, ROCK2 and PTPN14 (**Figure 6B and 6C**). The same interactions were validated using N-terminal FLAG-tag constructs with similar results (**Supplemental Figure S7**)

To map the regions of ULK4 that mediate these interactions, the pseudokinase domain and the armadillo repeat region were separately subjected to BioID (**Figure 6D**). This analysis revealed that the armadillo repeat region interacted uniquely with CAMSAP1, OFD1, poly(A) specific ribonuclease subunit 2 (PAN2) and several other proteins. ULK4 pseudokinase domain-specific interactions were detected for STK36, PTPN14, CAMSAP3. The interaction of ULK4 pseudokinase domain with STK36 is particularly intriguing as similar to STRAD/LKB1 this interaction might indicate that ULK4 directly regulates an active kinase (Zeqiraj et al., 2009a). All domain specific detected interactions are summarized in **Supplemental Table S3**. In order to determine whether the polymorphism at position 39 affects the ULK4 protein interaction network that might explain the pathogenicity of this amino acid alteration, the R39 variant was also subjected to BioID. No significant changes in the interactome were detected (**Supplemental Table S4**). Based on our DSF binding study, the R39 variant also bound ATP with similar affinity to the K39 protein. Stability of the pseudokinase domain fold is thus unlikely to be affected by this amino acid change.

The presented structural, bioinformatics, biochemical data together with our BioID interaction study provided a diversity of high confidence domain specific interactions and suggested a microtubular and centrosomal function of ULK4. The high affinity interaction with ATP is required for the stability of the pseudokinase domain, maintaining these interactions. The high-resolution structure of the ULK4 ATP γ S complex revealed a highly unusual ATP binding mode despite the lack of canonical ATP interaction motifs demonstrating how this unusual pseudokinase redesigned interaction for co-factor binding. Our evolutionary studies indicated that ULK4 orthologs in plants and protists conserve the canonical active site residues, suggesting that they function as 'active' kinases. In

contrast, the metazoan-specific variations in the active site suggest that loss of kinase activity and emergence of ULK4 as a pseudokinase occurred later in evolution, presumably to accommodate the specialized functions of differentiated cell types such as the metazoan specific nervous and immune systems. We hope that the presented data will stimulate future research on the function of this poorly studied pseudokinase.

Significance

ULK4 is an understudied protein associated with diseases such as hypertension and psychiatric disorders. Even though the overall structure of ULK4 has been reported before (Khamrui et al., 2020), we provided in this study important new data comprising structural details of the ULK4 ATP interaction in the absence of conserved canonical ATP binding motifs, evolutionary aspects as well as interaction partners of ULK4. Despite alterations in ATP binding motifs that are essential in canonical kinases, many pseudokinases bind ATP or mimic an ATP bound state which may not require or is even opposed by Mg^{2+} binding (Murphy et al., 2014) (Scheeff et al., 2009). In ULK, polymorphism of key ATP binding residues such as K39R and N139L have been linked to disease development, suggesting that ATP binding is required for ULK4 function. Interaction of the ATP cofactor in canonical kinases leads to a closed structure that is considerably more stable. It is therefore likely that ATP binding and the significant stability increase is required for ULK4 scaffolding function and its role in cellular signaling. Additional structural elements co-emerged with the loss of canonical motifs in evolution that compensate for the loss of stabilizing effects. For instance, our interspecies analysis of ULK4 sequences showed that alteration of the active site motifs in metazoan ULK4 has evolved with an extended activation loop, which stabilizes the flexible C-helix and compensates for the loss of VAIK lysine and the α C glutamate. The diversity of structural mechanisms that show how pseudokinases maintain strong binding activity for ATP in the absence of canonical binding motifs is fascinating and suggest an essential role of stable domain structures in pseudokinases which we confirmed by MD simulations. In order to shed light on signaling pathways regulated by ULK4 we performed interaction studies using BioID mass spectroscopy. These data together with biochemical validation revealed high confidence

interaction partners that were mapped to the pseudokinase and to the armadillo repeat domains, respectively. Many of the identified ULK4 interaction partners have centrosomal function suggesting a centrosomal role of ULK4. In addition, a number of active kinases have been identified as ULK4 interactors including the ULK family member STK36. Our studies provide a structural framework for targeting ULK4 in diseases.

Acknowledgments

SK, FP, DC, SM are grateful for support by the SGC, a registered charity (number 1097737) that receives funds from AbbVie, Bayer Pharma AG, Boehringer Ingelheim, Canada Foundation for Innovation, Eshelman Institute for Innovation, Genome Canada, Innovative Medicines Initiative (EU/EFPIA EUBOPEN), Janssen, Merck KGaA Darmstadt Germany, MSD, Novartis Pharma AG, Ontario Ministry of Economic Development and Innovation, Pfizer, São Paulo Research Foundation-FAPESP, Takeda, and Wellcome. Funding for NK from the NIH (1U01CA239106-01: A data analytics framework for mining the dark kinome) is acknowledged. SK is grateful for support by the Collaborative Sonderforschungsbereich 1177 Autophagy (SFB1177) at Frankfurt University, as well as the German Cancer Consortium (DKTK). The data collection at SLS has been supported by the funding from the European Union's Horizon 2020 research and innovation program under grant agreement number 730872, project CALIPSOplus.

Author Contributions

SK, SM, BR, RR designed research, FP, DC performed protein biochemistry and crystallization, SM solved and refined the ULK4 crystal structure, SS and NK performed the bioinformatics analysis. BR, J-St-G and MS performed the BioID study and validation data on interaction partners. SK, SM and NK wrote manuscript, which was approved by all authors.

Declaration of Interests

The authors declare no competing interests.

References

- Akhmanova, A., and Hoogenraad, C.C. (2015). Microtubule minus-end-targeting proteins. *Curr Biol* 25, R162-171.
- Atherton, J., Luo, Y., Xiang, S., Yang, C., Rai, A., Jiang, K., Stangier, M., Vemu, A., Cook, A.D., Wang, S., *et al.* (2019). Structural determinants of microtubule minus end preference in CAMSAP CKK domains. *Nat Commun* 10, 5236.
- Bailey, F.P., Byrne, D.P., Oruganty, K., Eyers, C.E., Novotny, C.J., Shokat, K.M., Kannan, N., and Eyers, P.A. (2015). The Tribbles 2 (TRB2) pseudokinase binds to ATP and autophosphorylates in a metal-independent manner. *Biochem J* 467, 47-62.
- Baines, A.J., Bignone, P.A., King, M.D., Maggs, A.M., Bennett, P.M., Pinder, J.C., and Phillips, G.W. (2009). The CKK domain (DUF1781) binds microtubules and defines the CAMSAP/ssp4 family of animal proteins. *Mol Biol Evol* 26, 2005-2014.
- Boudeau, J., Miranda-Saavedra, D., Barton, G.J., and Alessi, D.R. (2006). Emerging roles of pseudokinases. *Trends Cell Biol* 16, 443-452.
- Broderick, P., Chubb, D., Johnson, D.C., Weinhold, N., Forsti, A., Lloyd, A., Olver, B., Ma, Y., Dobbins, S.E., Walker, B.A., *et al.* (2011). Common variation at 3p22.1 and 7p15.3 influences multiple myeloma risk. *Nat Genet* 44, 58-61.
- Burgess, S.G., Mukherjee, M., Sabir, S., Joseph, N., Gutierrez-Caballero, C., Richards, M.W., Huguenin-Dezot, N., Chin, J.W., Kennedy, E.J., Pfuhl, M., *et al.* (2018). Mitotic spindle association of TACC3 requires Aurora-A-dependent stabilization of a cryptic alpha-helix. *EMBO J* 37.
- Caballe, A., Wenzel, D.M., Agromayor, M., Alam, S.L., Skalicky, J.J., Kloc, M., Carlton, J.G., Labrador, L., Sundquist, W.I., and Martin-Serrano, J. (2015). ULK3 regulates cytokinetic abscission by phosphorylating ESCRT-III proteins. *Elife* 4, e06547.
- Chaikuad, A., Koschade, S.E., Stolz, A., Zivkovic, K., Pohl, C., Shaid, S., Ren, H., Lambert, L.J., Cosford, N.D.P., Brandts, C.H., *et al.* (2019). Conservation of structure, function and inhibitor binding in UNC-51-like kinase 1 and 2 (ULK1/2). *Biochem J* 476, 875-887.
- Choi, H., Larsen, B., Lin, Z.Y., Breikreutz, A., Mellacheruvu, D., Fermin, D., Qin, Z.S., Tyers, M., Gingras, A.C., and Nesvizhskii, A.I. (2011). SAINT: probabilistic scoring of affinity purification-mass spectrometry data. *Nat Methods* 8, 70-73.
- Citri, A., Skaria, K.B., and Yarden, Y. (2003). The deaf and the dumb: the biology of ErbB-2 and ErbB-3. *Exp Cell Res* 284, 54-65.
- Durzynska, I., Xu, X., Adelmant, G., Ficarro, S.B., Marto, J.A., Sliz, P., Uljon, S., and Blacklow, S.C. (2017). STK40 Is a Pseudokinase that Binds the E3 Ubiquitin Ligase COP1. *Structure* 25, 287-294.
- Ehret, G.B., Ferreira, T., Chasman, D.I., Jackson, A.U., Schmidt, E.M., Johnson, T., Thorleifsson, G., Luan, J., Donnelly, L.A., Kanoni, S., *et al.* (2016). The genetics of blood pressure regulation and its target organs from association studies in 342,415 individuals. *Nat Genet* 48, 1171-1184.
- Eswaran, J., Lee, W.H., Debreczeni, J.E., Filippakopoulos, P., Turnbull, A., Fedorov, O., Deacon, S.W., Peterson, J.R., and Knapp, S. (2007). Crystal Structures of the p21-activated kinases PAK4, PAK5, and PAK6 reveal catalytic domain plasticity of active group II PAKs. *Structure* 15, 201-213.
- Fedorov, O., Niesen, F.H., and Knapp, S. (2012). Kinase inhibitor selectivity profiling using differential scanning fluorimetry. *Methods Mol Biol* 795, 109-118.
- Gingras, A.C., Abe, K.T., and Raught, B. (2019). Getting to know the neighborhood: using proximity-dependent biotinylation to characterize protein complexes and map organelles. *Curr Opin Chem Biol* 48, 44-54.
- Goshima, G., Mayer, M., Zhang, N., Stuurman, N., and Vale, R.D. (2008). Augmin: a protein complex required for centrosome-independent microtubule generation within the spindle. *J Cell Biol* 181, 421-429.

- Guo, D.C., Grove, M.L., Prakash, S.K., Eriksson, P., Hostetler, E.M., LeMaire, S.A., Body, S.C., Shalhub, S., Estrera, A.L., Safi, H.J., *et al.* (2016). Genetic Variants in LRP1 and ULK4 Are Associated with Acute Aortic Dissections. *Am J Hum Genet* 99, 762-769.
- Ha, B.H., and Boggon, T.J. (2018). The crystal structure of pseudokinase PEAK1 (Sugen kinase 269) reveals an unusual catalytic cleft and a novel mode of kinase fold dimerization. *J Biol Chem* 293, 1642-1650.
- Han, Y., Wang, B., Cho, Y.S., Zhu, J., Wu, J., Chen, Y., and Jiang, J. (2019). Phosphorylation of Ci/Gli by Fused Family Kinases Promotes Hedgehog Signaling. *Dev Cell* 50, 610-626 e614.
- Jacobsen, A.V., and Murphy, J.M. (2017). The secret life of kinases: insights into non-catalytic signalling functions from pseudokinases. *Biochem Soc Trans* 45, 665-681.
- Jura, N., Shan, Y., Cao, X., Shaw, D.E., and Kuriyan, J. (2009). Structural analysis of the catalytically inactive kinase domain of the human EGF receptor 3. *Proc Natl Acad Sci U S A* 106, 21608-21613.
- Khamrui, S., Ung, P.M.U., Secor, C., Schlessinger, A., and Lazarus, M.B. (2020). High-Resolution Structure and Inhibition of the Schizophrenia-Linked Pseudokinase ULK4. *J Am Chem Soc* 142, 33-37.
- Kwon, A., Scott, S., Taujale, R., Yeung, W., Kochut, K.J., Eysers, P.A., and Kannan, N. (2019). Tracing the origin and evolution of pseudokinases across the tree of life. *Sci Signal* 12.
- Lang, B., Pu, J., Hunter, I., Liu, M., Martin-Granados, C., Reilly, T.J., Gao, G.D., Guan, Z.L., Li, W.D., Shi, Y.Y., *et al.* (2014). Recurrent deletions of ULK4 in schizophrenia: a gene crucial for neurogenesis and neuronal motility. *J Cell Sci* 127, 630-640.
- Lang, B., Zhang, L., Jiang, G., Hu, L., Lan, W., Zhao, L., Hunter, I., Pruski, M., Song, N.N., Huang, Y., *et al.* (2016). Control of cortex development by ULK4, a rare risk gene for mental disorders including schizophrenia. *Sci Rep* 6, 31126.
- Lazarus, M.B., Novotny, C.J., and Shokat, K.M. (2015). Structure of the human autophagy initiating kinase ULK1 in complex with potent inhibitors. *ACS Chem Biol* 10, 257-261.
- Levy, D., Ehret, G.B., Rice, K., Verwoert, G.C., Launer, L.J., Dehghan, A., Glazer, N.L., Morrison, A.C., Johnson, A.D., Aspelund, T., *et al.* (2009). Genome-wide association study of blood pressure and hypertension. *Nat Genet* 41, 677-687.
- Liu, M., Fitzgibbon, M., Wang, Y., Reilly, J., Qian, X., O'Brien, T., Clapcote, S., Shen, S., and Roche, M. (2018a). Ulk4 regulates GABAergic signaling and anxiety-related behavior. *Transl Psychiatry* 8, 43.
- Liu, M., Xu, P., Guan, Z., Qian, X., Dockery, P., Fitzgerald, U., O'Brien, T., and Shen, S. (2018b). Ulk4 deficiency leads to hypomyelination in mice. *Glia* 66, 175-190.
- Lucet, I.S., and Murphy, J.M. (2017). Characterization of Ligand Binding to Pseudokinases Using a Thermal Shift Assay. *Methods Mol Biol* 1636, 91-104.
- Lupardus, P.J., Ultsch, M., Wallweber, H., Bir Kohli, P., Johnson, A.R., and Eigenbrot, C. (2014). Structure of the pseudokinase-kinase domains from protein kinase TYK2 reveals a mechanism for Janus kinase (JAK) autoinhibition. *Proc Natl Acad Sci U S A* 111, 8025-8030.
- Manning, G., Whyte, D.B., Martinez, R., Hunter, T., and Sudarsanam, S. (2002). The protein kinase complement of the human genome. *Science* 298, 1912-1934.
- Marcette, J.D., Chen, J.J., and Nonet, M.L. (2014). The *Caenorhabditis elegans* microtubule minus-end binding homolog PTRN-1 stabilizes synapses and neurites. *Elife* 3, e01637.
- Mukherjee, K., Sharma, M., Urlaub, H., Bourenkov, G.P., Jahn, R., Sudhof, T.C., and Wahl, M.C. (2008). CASK Functions as a Mg²⁺-independent neurexin kinase. *Cell* 133, 328-339.
- Murone, M., Luoh, S.M., Stone, D., Li, W., Gurney, A., Armanini, M., Grey, C., Rosenthal, A., and de Sauvage, F.J. (2000). Gli regulation by the opposing activities of fused and suppressor of fused. *Nat Cell Biol* 2, 310-312.

Murphy, J.M., Czabotar, P.E., Hildebrand, J.M., Lucet, I.S., Zhang, J.G., Alvarez-Diaz, S., Lewis, R., Lalaoui, N., Metcalf, D., Webb, A.I., *et al.* (2013). The pseudokinase MLKL mediates necroptosis via a molecular switch mechanism. *Immunity* 39, 443-453.

Murphy, J.M., Farhan, H., and Eyers, P.A. (2017). Bio-Zombie: the rise of pseudoenzymes in biology. *Biochem Soc Trans* 45, 537-544.

Murphy, J.M., Nakatani, Y., Jamieson, S.A., Dai, W., Lucet, I.S., and Mace, P.D. (2015). Molecular Mechanism of CCAAT-Enhancer Binding Protein Recruitment by the TRIB1 Pseudokinase. *Structure* 23, 2111-2121.

Murphy, J.M., Zhang, Q., Young, S.N., Reese, M.L., Bailey, F.P., Eyers, P.A., Ungureanu, D., Hammaren, H., Silvennoinen, O., Varghese, L.N., *et al.* (2014). A robust methodology to subclassify pseudokinases based on their nucleotide-binding properties. *Biochem J* 457, 323-334.

Neuwald, A.F. (2014). A Bayesian sampler for optimization of protein domain hierarchies. *J Comput Biol* 21, 269-286.

Oprea, T.I. (2019). Exploring the dark genome: implications for precision medicine. *Mamm Genome* 30, 192-200.

Patel, O., Griffin, M.D.W., Panjekar, S., Dai, W., Ma, X., Chan, H., Zheng, C., Kropp, A., Murphy, J.M., Daly, R.J., *et al.* (2017). Structure of SgK223 pseudokinase reveals novel mechanisms of homotypic and heterotypic association. *Nat Commun* 8, 1157.

Petrie, E.J., Sandow, J.J., Jacobsen, A.V., Smith, B.J., Griffin, M.D.W., Lucet, I.S., Dai, W., Young, S.N., Tanzer, M.C., Wardak, A., *et al.* (2018). Conformational switching of the pseudokinase domain promotes human MLKL tetramerization and cell death by necroptosis. *Nat Commun* 9, 2422.

Pongrakhananon, V., Saito, H., Hiver, S., Abe, T., Shioi, G., Meng, W., and Takeichi, M. (2018). CAMSAP3 maintains neuronal polarity through regulation of microtubule stability. *Proc Natl Acad Sci U S A* 115, 9750-9755.

Ribeiro, A.J.M., Das, S., Dawson, N., Zaru, R., Orchard, S., Thornton, J.M., Orengo, C., Zeqiraj, E., Murphy, J.M., and Eyers, P.A. (2019). Emerging concepts in pseudoenzyme classification, evolution, and signaling. *Sci Signal* 12.

Roux, K.J. (2013). Marked by association: techniques for proximity-dependent labeling of proteins in eukaryotic cells. *Cell Mol Life Sci* 70, 3657-3664.

Scheeff, E.D., Eswaran, J., Bunkoczi, G., Knapp, S., and Manning, G. (2009). Structure of the pseudokinase VRK3 reveals a degraded catalytic site, a highly conserved kinase fold, and a putative regulatory binding site. *Structure* 17, 128-138.

Shrestha, S., Byrne, D.P., Harris, J.A., Kannan, N., and Eyers, P.A. (2020). Cataloguing the dead: breathing new life into pseudokinase research. *FEBS J*.

Sun, L., Wang, H., Wang, Z., He, S., Chen, S., Liao, D., Wang, L., Yan, J., Liu, W., Lei, X., *et al.* (2012). Mixed lineage kinase domain-like protein mediates necrosis signaling downstream of RIP3 kinase. *Cell* 148, 213-227.

Teo, G., Liu, G., Zhang, J., Nesvizhskii, A.I., Gingras, A.C., and Choi, H. (2014). SAINTexpress: improvements and additional features in Significance Analysis of INTeractome software. *J Proteomics* 100, 37-43.

Uehara, R., Nozawa, R.S., Tomioka, A., Petry, S., Vale, R.D., Obuse, C., and Goshima, G. (2009). The augmin complex plays a critical role in spindle microtubule generation for mitotic progression and cytokinesis in human cells. *Proc Natl Acad Sci U S A* 106, 6998-7003.

Vogel, P., Read, R.W., Hansen, G.M., Payne, B.J., Small, D., Sands, A.T., and Zambrowicz, B.P. (2012). Congenital hydrocephalus in genetically engineered mice. *Vet Pathol* 49, 166-181.

Wang, H., Sun, L., Su, L., Rizo, J., Liu, L., Wang, L.F., Wang, F.S., and Wang, X. (2014). Mixed lineage kinase domain-like protein MLKL causes necrotic membrane disruption upon phosphorylation by RIP3. *Mol Cell* 54, 133-146.

Zachari, M., and Ganley, I.G. (2017). The mammalian ULK1 complex and autophagy initiation. *Essays Biochem* 61, 585-596.

Zeqiraj, E., Filippi, B.M., Deak, M., Alessi, D.R., and van Aalten, D.M. (2009a). Structure of the LKB1-STRAD-MO25 complex reveals an allosteric mechanism of kinase activation. *Science* 326, 1707-1711.

Zeqiraj, E., Filippi, B.M., Goldie, S., Navratilova, I., Boudeau, J., Deak, M., Alessi, D.R., and van Aalten, D.M. (2009b). ATP and MO25alpha regulate the conformational state of the STRADalpha pseudokinase and activation of the LKB1 tumour suppressor. *PLoS Biol* 7, e1000126.

Figure legends

Figure 1: Structural features of ULK kinases. A: Structure based alignment of the kinase domains of ULK1 and ULK4. The main secondary structural elements are highlighted and conserved catalytic domain motifs are labelled. Amino acid residue changes in conserved motifs are indicated by a green arrow. **B:** Domain architecture of the human ULK family. Protein interaction domains are labelled as C-terminal domain (CTD; ULK1 and ULK2), microtubule interacting and trafficking molecule (MIT; ULK3) as well as armadillo repeat domain (Arm, STK36 and ULK4). **C:** Structure of the ULK4 pseudokinase domain in complex with ATP γ S. Position of residues in conserved catalytic domain motifs that are altered in ULK4 are highlighted by blue spheres. **D:** Structure of ULK1 for comparison. Additional supplemental figures (**Supplemental Figure S2**) show superimpositions of ULK4 with ULK1 and ULK2, respectively.

Figure 2: Details of ULK4 structural features and evolutionary conservation. A: Comparison of interaction between the ULK1 VAIK motif and α C showing the canonical salt bridge of α C in conformations and ULK4 in which the VAIK lysine is replaced by a leucine residue (L33) and the α C glutamate (W46) by a tryptophan. **B:** Details of the interaction of the ULK4 activation segment helix with α C. Interactions in the catalytic loop are also shown. **C:** Histogram showing the lengths of the activation loop in ULK4 orthologs. Canonical Plants, Pseudo Metazoans, and Pseudo Protists/Plants are colored as light red, light purple, and green, respectively **D:** Secondary structure plot of the activation loop of ULK4. Define Secondary Structure of Proteins (DSSP) was used to

define the secondary structure. A maximum likelihood tree of available ULK4 sequences of different species are shown in **Supplemental Figure S3**.

Figure 3: ULK4 specific constraints are in the N-lobe. A: Sequence constraints that distinguish ULK4 sequences from closely related ULK1-3 sequences are shown in a contrast hierarchical alignment (CHA). The CHA alignment shows selected ULK4 sequences from diverse organisms as the display alignment. The foreground alignment consists of 277 ULK4 sequences while the background alignment contains 1732 ULK1-3 sequences. The foreground and background alignments are shown as residue frequencies below the display alignment in integer tenths (1–9). The histogram (red) above the display alignment indicates the extent to which distinguishing residues in the foreground alignment diverge from the corresponding position in the background alignment. Black dots mark the alignment positions used by BPPS (Neuwald, 2014) procedure when classifying ULK4 from other ULK sequences. Alignment number is based on the human ULK4 sequence. **B:** ULK4 specific residues mapped onto the crystal structure of human ULK4. The distinguishing residues are shown as sticks and surface electrostatic. The activation loop is colored green. A detailed view is shown in the right panel. Carbon atoms of the residues are colored in magenta. A consensus sequence of ULK4 activation segment is shown in **Supplemental Figure S4**.

Figure 4: Nucleotide binding to ULK4. A: Comparison of UV spectra of ULK1 and ULK4. The high absorbance of ULK4 at 260 nM suggests the presence of ATP after purification from bacteria. **B:** Affinity of nucleotides and effect of metals on binding. Shown are temperature shifts at 1mM nucleotide concentration as well as SEM (n=3). **C:** Canonical binding mode of ATP exemplified by the PKA ATP complex (pdb-ID: 1ATP). Main interactions are shown in stick representation and are labelled. Some side chains have been deleted for clarity. Major hydrophobic contacts interacting with the adenine ring in the low lobe are shown as spheres. **D:** Binding mode of ATP γ S in ULK4. Orientations and labels are similar as shown in C. **E:** Nucleotide binding in STRAD α (PDB-ID 3GNI). **F:** Effect of K39R polymorphism on nucleotide binding assess by temperature shift experiments. Shown are temperature shifts at 1mM nucleotide concentration as well as SEM (n=3).

Figure 5. ULK4 Proximity Interaction Network. Proximity interactors of FlagBirA*-ULK4 and/or ULK4-BirA*Flag fusion proteins expressed in Flp-In T-REx 293 cells are shown. Probability of specific interactors was determined using Bayesian statistics against a panel of FlagBirA* only BioID samples. Proteins with a Bayesian FDR of interaction of ≤ 0.01 were considered statistically significant putative proximity interactors. Border (in red) thickness of nodes is proportional to the number of PSMs for each protein. Proteins are categorized according to GO-based functions. ULK4 cellular localization data are shown in **Supplemental Figure S6**.

Figure 6. Association studies using density gradients and immunoprecipitation. A: Whole cell lysate from tetracycline-treated Flp-In T-REx 293 cells expressing C-terminally FLAG-tagged full-length ULK4 (ULK4-BirA*Flag) was used for density gradient ultra-centrifugation. Immunoblot showed that fractions containing ULK4 also contained PTPN14, ROCK1, ROCK2 and CAMSAP1. **B:** ULK4 was immunoprecipitated from tetracycline-treated Flp-In T-REx 293 cells expressing full-length ULK4-BirA*Flag. Subsequent immunoblotting showed co-immunoprecipitation of PTPN14, ROCK1, ROCK2 and CAMSAP1. **C:** Flp-In T-REx 293 cells expressing tetracycline-inducible Flag- and BirA- tagged full length ULK4 (ULK4-BirA*Flag) were treated with biotin for 24 hrs. Proteins thus biotinylated were pulled down using streptavidin-conjugated beads. Immunoblot showed the presence of PTPN14, ROCK1, ROCK2 and CAMSAP1 in the eluate. **D:** Domain Mapping of ULK4 Proximity Interacting Proteins. Proximity interactors of FlagBirA*-ULK4 and/or ULK4-BirA*Flag pseudokinase domain or Armadillo repeats (Arm) domains fusion proteins expressed in Flp-In T-REx 293 cells are shown. Probability of specific interactors was determined using Bayesian statistics against a panel of FlagBirA* only BioID samples. Proteins with a Bayesian FDR of interaction of ≤ 0.01 were considered statistically significant putative proximity interactors.

Figure 1

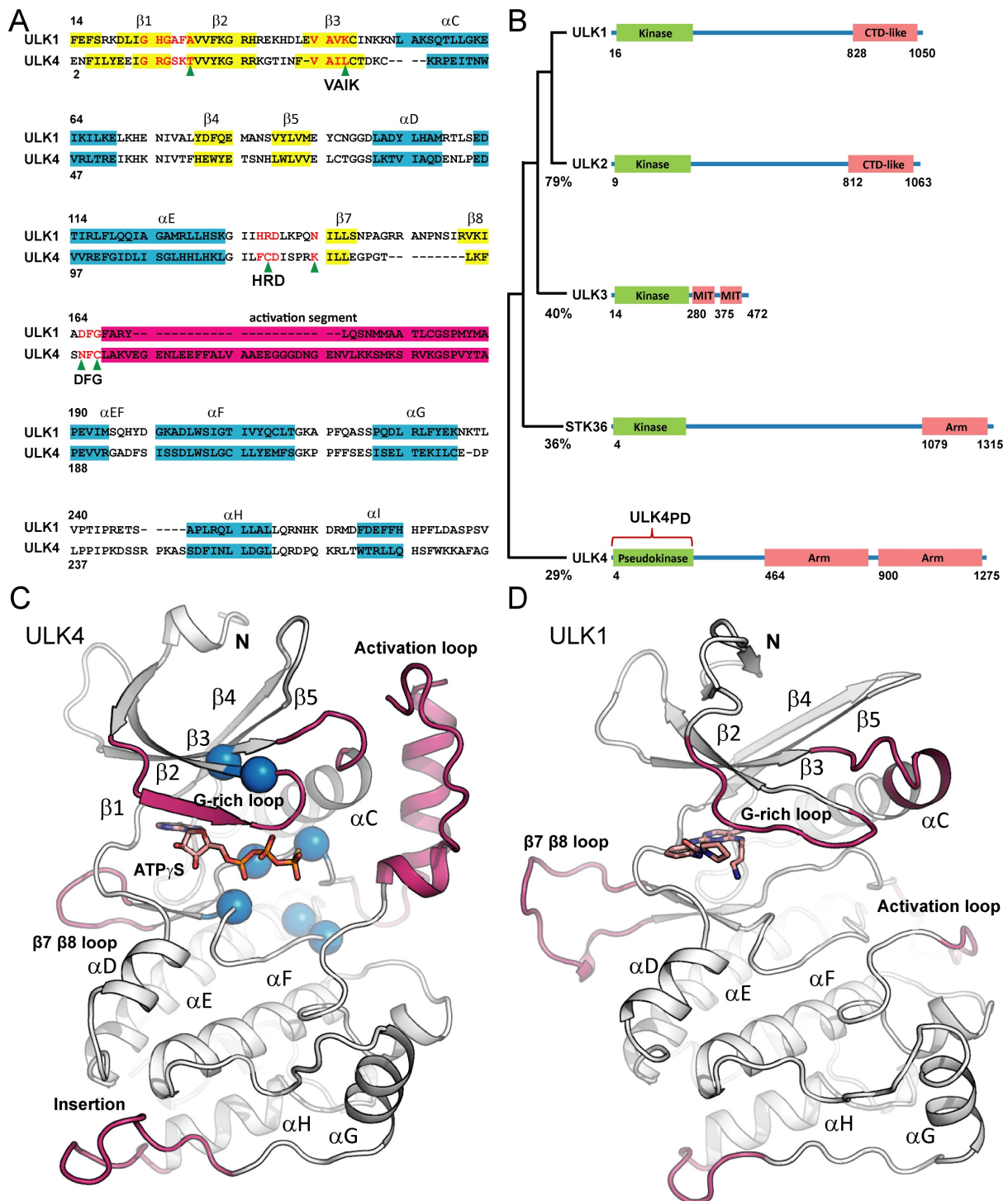


Figure 2

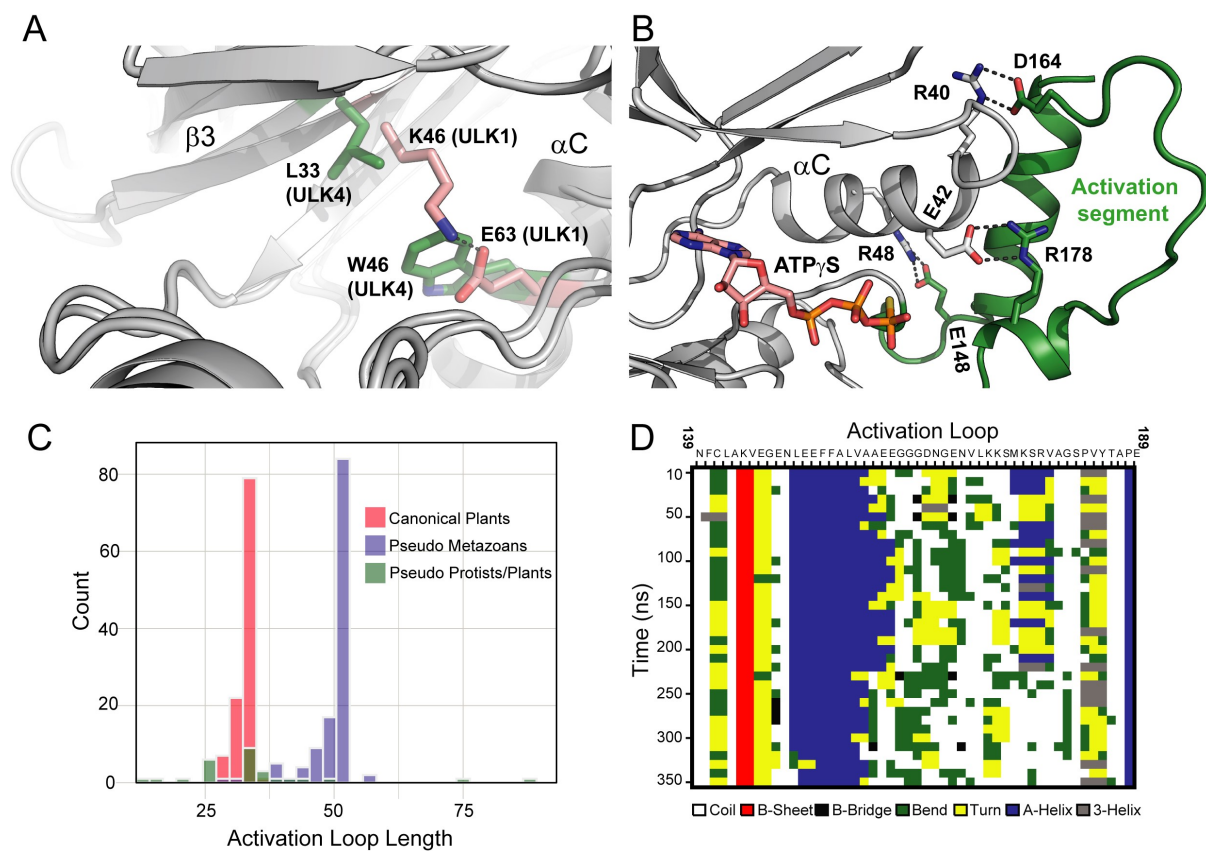


Figure 3

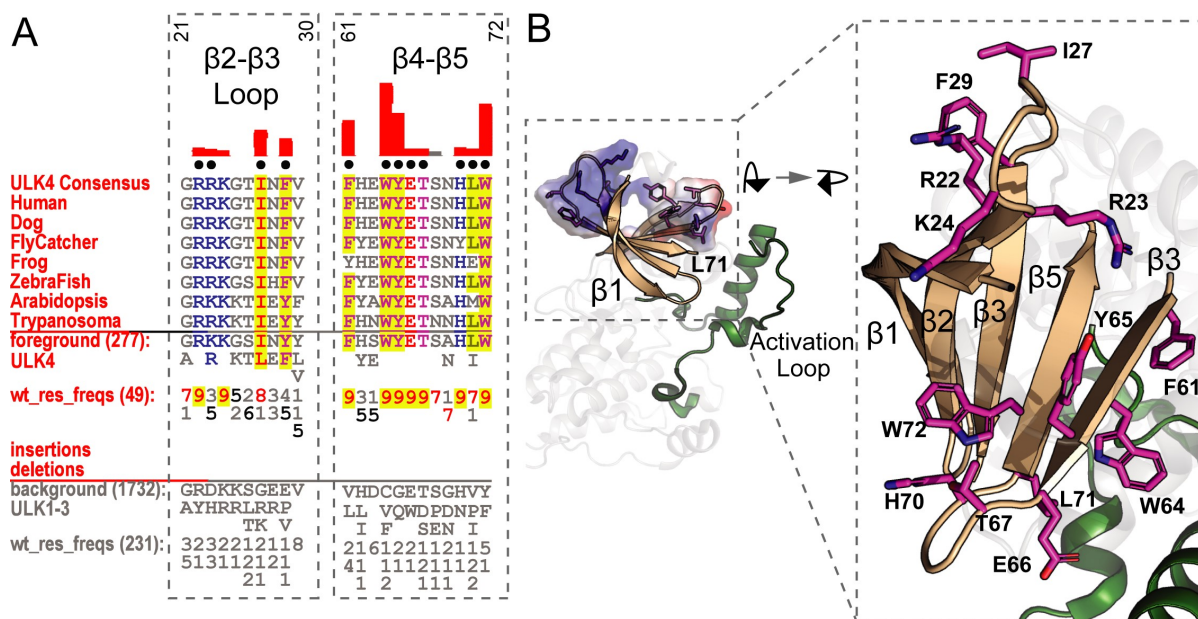


Figure 4

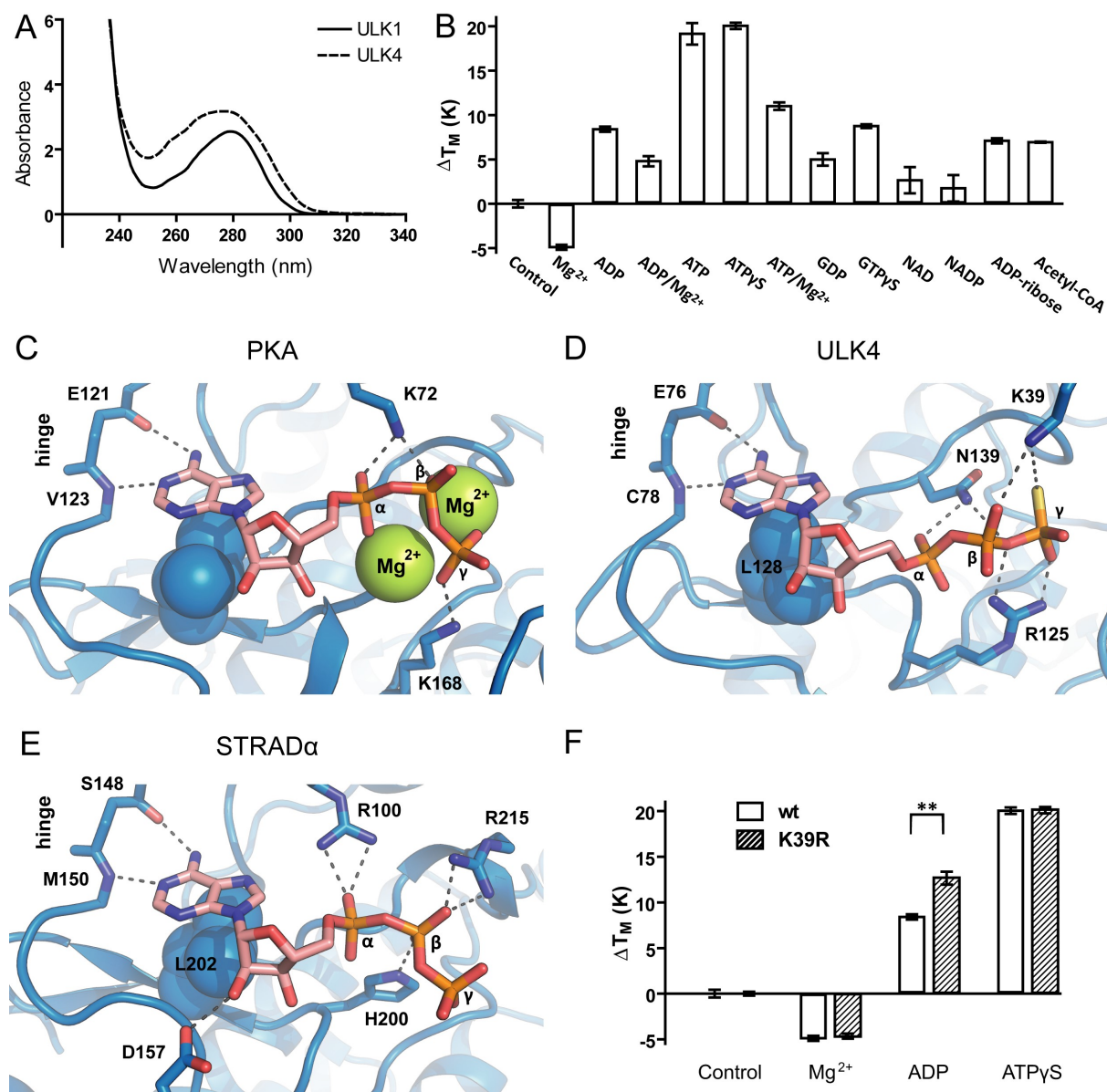


Figure 5

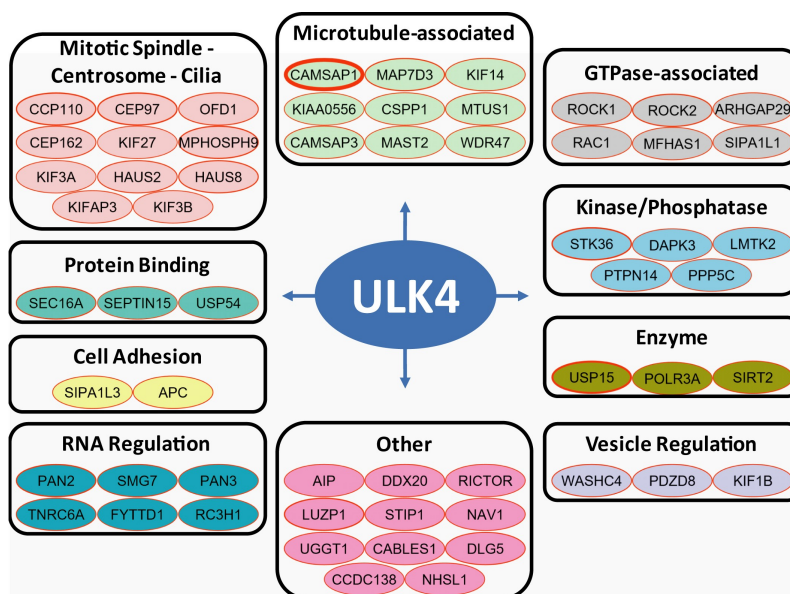


Figure 6

