UNIVERSITY OF CALIFORNIA SAN DIEGO

Characterization of the full length p90 RSK2:ERK2 complex via network analysis, biophysical and structural methods

A dissertation submitted in partial satisfaction of the

requirements for the degree of Doctor of Philosophy

in

Chemistry

by

Evan Kobori

Committee in charge:

Professor Susan Taylor, Chair
Professor Gourisankar Ghosh
Professor Elizabeth Komives
Professor Andres Leschziner
Professor Valerie Schmidt

2023

The dissertation of Evan Kobori is approved,

and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2023

DEDICATION

I dedicate this work to my family who have helped me get this far.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

Å	Angstrom
CryoEM	Cryogenic electron microscopy
CTK	C-terminal kinase of p90 ribosomal s6 kinase
Da	Dalton
EDTA	Ethylenediaminetetraacetic Acid
ERK	Extracellular regulated kinase
FEP	Free energy pertrubation
FPLC	Fast Protein Liquid Chromatography
HDXMS	Hydrogen deuterium exchange mass spectrometry
K_d	Dissociation equilibrium constant
LSP	Local spatial pattern
MALDI-TOF	Matrix-Assisted Laser Desorption Ionization with Time of Flight Detection
MD	Molecular dynamics
MMGBSA	Molecular mechanics with generalized Born and surface area solvation
NTK	N-terminal kinase of p90 ribosomal s6 kinase
ORF45	Open reading frame protein 45 of Kaposi sarcoma herpesvirus
PDB	Protein Data Bank
PDK1	Phosphoinositide-dependent kinase 1
SAXS	Small angle x-ray scattering
SPR	Surface plasmon resonance

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

Many people have influenced me and helped me get where I am today. I first want to thank my family. My parents have been an unwavering source of support throughout my entire life, who have always done whatever they could to help, and guide me. My older brother and sister have had my back and looked out for me from the very beginning as well.

My advisor, Susan Taylor, has had a major role in my professional development and has allowed me to grow into the scientist that I am today. I initially joined her lab rather inexperienced; Susan and the lab provided me the structure and freedom to develop the necessary skills and grow into becoming a better more independent scientist. Her passion and genuine excitement for science is inspiring, and I thank her for profound role in my professional development.

Finally, I want to thank my fellow lab mates, former and current. First, Ronit Ilouz, Mira Sastri, Tsan-Wen Lu, Grace Liu, and James Hall helped me establish my footing when I first joined the lab and helped teach me some of the foundational skills in the lab. Gergo Gogl taught me how to prepare RSK and ERK samples. Phillip Aoto and Alexandr Kornev taught me a lot about how to think about my proteins from a computational perspective. Alexandr in particular was incredibly helpful in guiding me through manuscript writing. I'd like to thank Andres Leschziner and Colin Deniston for their help in the initial cryoEM analysis of this project and in particular Andres, for his continued support in terms of data processing resources. Lastly, I would like to thank all of my current lab mates who make the lab a pleasant place

every day. Katherine Chen, Rodeon Malinovski, and Hoang Nguyen comprise an excellent RSK/ERK time and are an absolute pleasure to work with. It's also been a pleasure to have Jian Wu, Jody Weng, Pallavi Kaila Sharma, Jessica Bruystens, Nathan Vu, Gerald Tan, and Frank Ma as lab mates. I'd like to thank them all for helping to create a positive and healthy work environment.

VITA

2013 Bachelors of Science, Chemical Biology, University of California, Berkeley

2013 Bachelors of Arts, Statistics, University of California, Berkeley

2017 Masters of Science, Chemistry, University of California San Diego

2023 Doctorate of Philosophy, Chemistry, University of California San Diego

PUBLICATIONS

Taylor SS, Soberg K, **Kobori E**, Wu J, Pautz S, Herberg FW, et al. The tails of PKA. Molecular Pharmacology. 2021:MOLPHARM-MR-2021-000315.

Alexa, A., Sok, P., Gross, F., Albert, K., **Kobori, E.**, et al., *A non-catalytic herpesviral protein reconfigures ERK-RSK signaling by targeting kinase docking systems in the host.* Nature Communications, 2022. **13**(1): p. 472.

Kobori, E., Aoto, P., Zhang, F., Gheyi, T., Chen, K., Antonysamy, S., Vieth, M., Kornev, A., Taylor, S. *Network analysis of the heterodimeric RSK1:ERK2 signaling complex*. In preparation

Kobori, E., Aoto, P., Wu, J., Chen, K., Malinovski, R., Taylor, S. *Hydrogen Deuterium Exchange Mass Spectrometry Analysis of the FL RSK2:ERK2 signaling complex and characterization of the conformational malleability of the RSK2 N-terminal kinase*. In preparation

FIELDS OF STUDY

Major Field: Biochemistry

Studies in Biochemistry and Biophysics

Professor Susan Taylor

HONORS AND AWARDS

2019-2020 Pharmacology Training Grant, trainee

2016-2018 Molecular Biophysics Training Grant, trainee

ABSTRACT

Characterization of the full length p90 RSK2:ERK2 complex via network analysis, biophysical and structural methods

by

Evan Kobori

Doctor of Philosophy in Chemistry

University of California San Diego, 2023

Professor Susan Taylor, Chair

Since the crystal structure of PKA was solved over thirty years ago, the overall kinase core has been well characterized; various structures have elucidated the overall catalytic cycle and internal interactions required for an active kinase. However, the function and regulation of full-length (FL) kinases and their complexes that include conserved, critically important intrinsically disordered regions (IDRs), are relatively

poorly understood.

Here we analyze p90 ribosomal S6 kinase (RSK) which is a downstream effector kinase of the Ras/MAPK signaling pathway. RSK contains two distinct, functional kinase domains, an N-terminal kinase (NTK) and a C-terminal kinase (CTK), and flexible IDRs with conserved motifs that mediate interactions necessary for complex assembly with its activating kinase ERK1/2 and to propagate the multistep RSK activation pathway. A flexible Linker connects the NTK and CTK, and a C-terminal tail contains motif that mediate high affinity ERK2 binding. RSK has an important role in cardiac function and is mis-regulated in neurodegenerative disorders and various forms of cancer. We utilize a combination of computational, biophysical, and structural techniques to study FL RSK and the RSK-ERK complex. Molecular dynamics simulations and network analysis are utilized to better understand the interfaces that comprise the CTK-ERK2 heterodimeric complex, and to identify residues that are critical for the stability of the complex but would be difficult to appreciate using other classical techniques. We utilized hydrogen deuterium exchange mass spectrometry (HDXMS) to measure solvent accessibility of the FL RSK2-ERK2 complex to determine flexible and unstructured regions in the complex. And we explored how the flexible IDRs of RSK mediate additional steps in RSK activation by peptide arrays. A non-catalytic viral protein ORF45, was found to stabilize the FL RSK2-ERK2 complex. With HDXMS and peptide arrays, we identified that ORF45 docks onto the NTK, and indirectly, enhances the RSK2-ERK2 interface. Finally, we used cryoEM on the RSK2-ERK2-ORF45 complex and solved low resolution

reconstructions of free ERK2, RSK2, and the RSK2-ERK2-ORF45 complex to evaluate their overall size and shape. A major continuing goal is to improve the resolution of our reconstructions to determine the orientation of all three kinase domains relative to one another, and identify how the IDRs of RSK are positioned to mediate additional steps in RSK activation.

Chapter I

Introduction

A. Conserved features of protein kinases

Protein kinases are among the most important enzymes throughout all of biology, and are responsible for roughly 2% of the human encoded genome (Manning, Whyte et al. 2002). Kinases transfer the γ phosphate of ATP to various substrates resulting in a conformational or functional changes (Taylor and Kornev 2011, Taylor, Keshwani et al. 2012, Kornev and Taylor 2015), and phosphorylate about 1/3 of all proteins in a human cell (Cohen 2002). Kinase misregulation is the driver behind a variety of diseases including cancer, diabetes, neurological disorders, etc. Thus, kinases are among the most popular class of enzymes targeted by drug companies mostly with small molecule inhibitors. The global market for kinase small molecule therapies consists of ~ $20 billion annually, there are 62 FDA approved kinase therapeutics, and many more in late stage clinical trials (Wu, Nielsen et al. 2015, Roskoski 2021). Although they are not efficient enzymes, kinases have evolved to be highly sensitive dynamic switches enabling them to quickly respond to a variety of cellular stimuli (Kornev and Taylor 2015). The ability of kinases to toggle between open and closed or inactive and active conformations provides regulation and the switchlike property of kinases. Typically, kinases are "switched on" and assume active conformations following phosphorylation by upstream kinases, thus multilevel kinase cascades are common (Gógl, Kornev et al. 2019).

The first protein kinase, phosphorylase kinase was identified in the 1950s (Krebs 1998) and by 1982, it was known that protein kinases were one of the largest gene families but also that the protein kinases had evolved from a common ancestor

(Walsh, Perkins et al. 1968, Barker and Dayhoff 1982). Although cAMP-dependent protein kinase (PKA) was the second protein kinase to be discovered, it the first protein kinase structure to be solved (Knighton, Zheng et al. 1991) (Knighton, Zheng et al. 1991) and it continues to serve as a prototype for the entire kinase family. The overall kinase architecture adopts a bilobal fold (Taylor, Yang et al. 2004), and conserved motifs regulate function and catalytic activity (Hanks and Hunter 1995). The smaller N-lobe is mobile and primarily responsible for ATP binding and consists of five β strands and a conserved mobile α helix (Akamine, Madhusudan et al. 2003, Taylor, Keshwani et al. 2012)(Fig 1.1A). A conserved glycine rich loop (G-loop) between β1 and β2 contributes to ATP binding and positions the γ phosphate for transfer. Moreover, $K72^{PKA}$ in the β3 strand stabilizes the α, β phosphates of ATP and forms a salt bridge with the conserved $E91^{PKA}$, bringing the αC helix in towards the kinase core (Fig 1.1B). $K72^{PKA}$ is essential for all active kinases and is typically mutated to generate a kinase dead phenotype (Gibbs and Zoller 1991). The αC helix (Fig 1.1B) is dynamic and alternates between 'in' and 'out' positions in active or inactive conformations, respectively (Taylor, Shaw et al. 2015). In rather extreme inactive cases, the αC helix can be completely unfolded (Ikuta, Kornienko et al. 2007, Utepbergenov, Derewenda et al. 2012). The linker between the N and C lobes, known as the hinge, consists of residues important for ATP binding (Liao 2007). The larger, primarily helical C-lobe is more stable and contributes more to substrate binding and other protein:protein interactions (Taylor and Kornev 2011, Taylor, Keshwani et al. 2012, Kornev and Taylor 2015). The αE, αF, and αH helices comprise the

hydrophobic core of the kinase and are shielded from solvent, while the αG is more solvent exposed (Fig 1.1A) (Yang, Garrod et al. 2005, Steichen, Iyer et al. 2010). Following the αE is the catalytic loop, which is responsible for the kinases enzymatic activity. The conserved H/YRDXKXXN motif contains the catalytic Asp, and transfers the ATP γ phosphate to the substrate (Fig 1.1B). The kinase active site lies in the cleft between the N and C lobes, and contains the activation segment. The activation segment consists of the magnesium positioning DFG motif, the Activation Loop that in most cases is phosphorylated, and the APE motif which anchors to the αF Helix the and αH-αI loop. Phosphorylation of the Activation Loop usually organizes the HRD in the Catalytic Loop and the αC helix, and positions the DFG motif which stabilizes the active conformation of the kinase (Hari, Merritt et al. 2013, Möbitz 2015) (Nolen, Taylor et al. 2004) (Kornev and Taylor 2015) (Gógl, Kornev et al. 2019).

Spanning both the N and C lobes are non-contiguous hydrophobic residues termed the regulatory (R) and catalytic (C) spines (Fig 1.1C). The R spine consists of four residues: hydrophobic residues from the β4 strand and αC helix, the His/Tyr from HRD motif, Phe from the DFG motif, and is anchored to the αF helix by Asp220. The R spine is dynamically assembled and is formed only in active kinases. Assembly of the R spine is a hallmark feature of fully active kinases that are capable of *trans*-phosphorylating a heterologous substrate. The C spine is more stable and consists of hydrophobic residues from the αF helix to the N-lobe, and is completed by the ATP ribose ring (Kornev, Haste et al. 2006, Kornev, Taylor et al. 2008).

Since the first structure of PKA was solved, structures of various conformations along with different dynamics studies (Masterson, Cheng et al. 2010, Masterson, Shi et al. 2011, Masterson, Cembran et al. 2012, Srivastava, McDonald et al. 2014) have broadly defined the overall kinase fold and catalytic cycle (Knighton, Zheng et al. 1991, Zheng, Knighton et al. 1993, Narayana, Cox et al. 1997, Narayana, Cox et al. 1997, Akamine, Madhusudan et al. 2003, Steichen, Iyer et al. 2010, Steichen, Kuchinskas et al. 2012, Bastidas, Deal et al. 2013). The next challenge and frontier in kinase signaling and biology is to characterize full length (FL) kinases and their substrate and/or inhibitor complexes to better understand the roles of flanking unstructured intrinsically disordered regions (IDRs), other domains that flank the kinase core (Gógl, Kornev et al. 2019, Taylor, Søberg et al. 2021). Embedded within IDRs are Small linear motifs (SLiMs) that include phospho-switches and regulatory elements that control the activity of the kinase domain. Kinases are often phosphorylated by other kinases, and typically assemble substrate:activator complexes (Kholodenko, Hancock et al. 2010). These flanking IDRs play crucial roles in the integration of kinases to form complex signaling networks (Wong, Jennings et al. 2004) (Newton 2018) (Taylor, Ilouz et al. 2012).

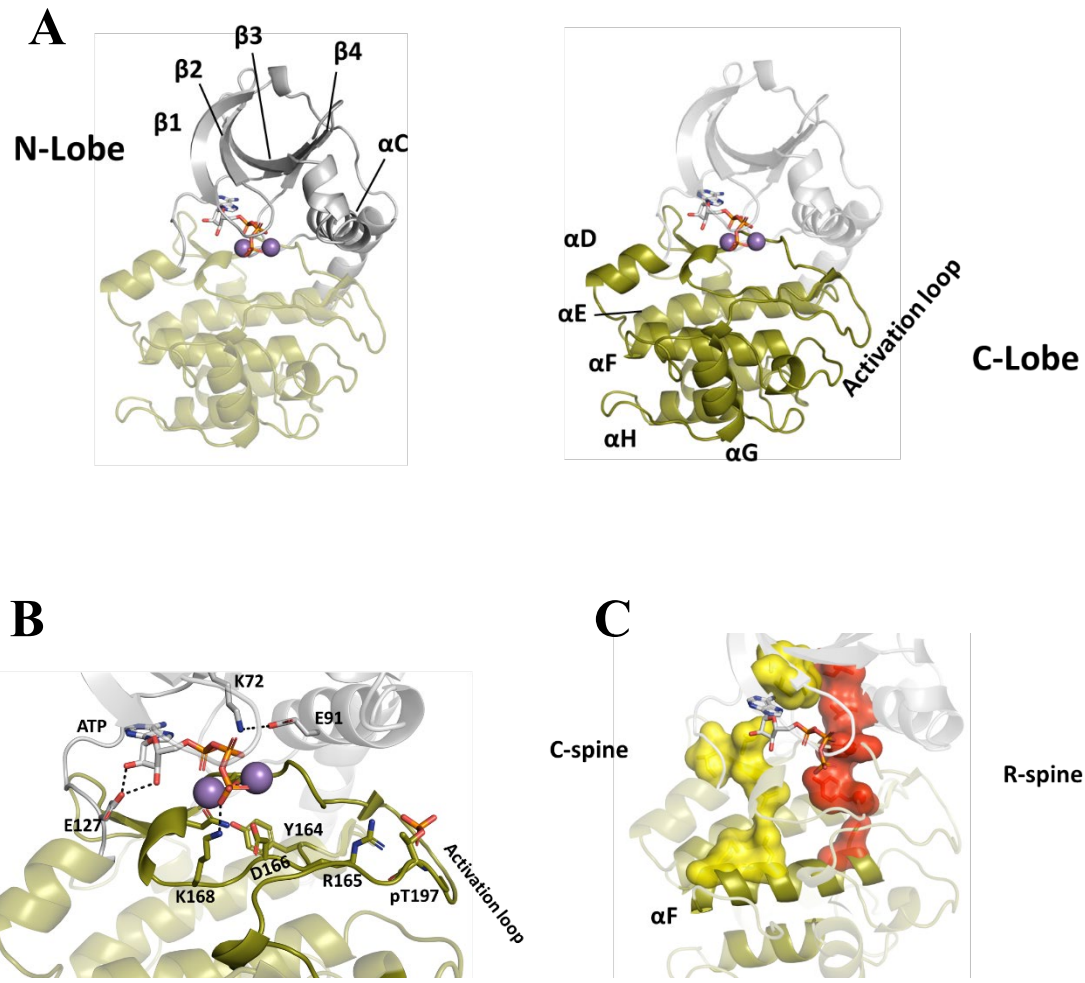


Figure 1.1. Protein kinase fold and internal architecture.
(A) Crystal structure of PKA serves as a prototype for protein kinases. Kinases have a conserved bilobal fold, and have a smaller N-lobe (Left) in white consisting of mostly β strands. The right shows the larger, mostly helical C-lobe in green.
(B) View of PKA active site displays conserved residues binding ATP, and organizing the conserved catalytic, and activation loops.
(C) Assembled hydrophobic spines in PKA. Regulatory spine residues are represented as a red surface, and the catalytic spine residues are shown as a yellow surface. Both spines anchor onto the hydrophobic αF helix.

B. RSK family biology, signaling, and tissue distribution

Characterization of p90 Ribosomal S6 Kinase (RSK) presents an opportunity to better understand how FL kinases assemble as larger complexes and integrate with unstructured/IDRs. RSK is a MAPK activated protein kinase (MAPKAPK), that lies downstream of the Ras/MAPK signaling pathway (Anjum and Blenis 2008)(Fig 1.2A). The Ras/MAPK pathway is one of the most well characterized pathways, as mutations to KRAS (Huang, Guo et al. 2021) and BRAF (Ascierto, Kirkwood et al. 2012) are among the most commonly mutated proteins in melanoma and other forms of cancer. Following growth factor like EGF binding to the EGF receptor, Ras is activated following nucleotide exchange from GDP to GTP, promoting activation of RAF. Active RAF phosphorylates and activates MEK, which phosphorylates ERK2, which then phosphorylates and activates RSK (Anjum and Blenis 2008, Romeo, Zhang et al. 2012, Gimple and Wang 2019, Zaman, Wu et al. 2019). RSK is primarily in the cytoplasm, but throughout its activation it translocates to the membrane, and then back to the cytoplasm or nucleus to phosphorylate various downstream substrates that are involved in translation, cell growth, proliferation, cell cycle control, etc.
The RSK family consists of four isoforms (1-4) that are thought to be functionally non-redundant based on differences in tissue expression levels and on different disease phenotypes associated with the various isoforms (Bjørbaek, Zhao et al. 1995, Fisher and Blenis 1996). RSKs 1-3 are ubiquitously expressed, but RSK1 is the predominant isoform in the kidney, lung, and pancreas. RSKs 2-3 are predominant in skeletal muscle, cardiac tissue, and pancreas as well (Zeniou, Ding et al. 2002). Overall

expression levels of RSK4 are lower relative to the other isoforms, but RSK4 was detected in brain, heart, cerebellum, kidney, and muscle tissue (Dümmler, Hauge et al. 2005). Each RSK isoform contains two distinct kinase domains that are each flanked by long IDRs (Fig 1.2B). RSKs contain an N-terminal kinase (NTK) and a C-terminal kinase. The NTK is flanked by an N-terminal tail and a long linker that connects the NTK and CTK. The CTK is followed by a C-terminal tail (Bjørbaek, Zhao et al. 1995, Fisher and Blenis 1996). Our interest in the RSK family stems from its unique domain organization, ability to form a stable complex with its activating kinase, ERK2, and its coordination of kinase domains and IDRs to regulate its catalytic function.

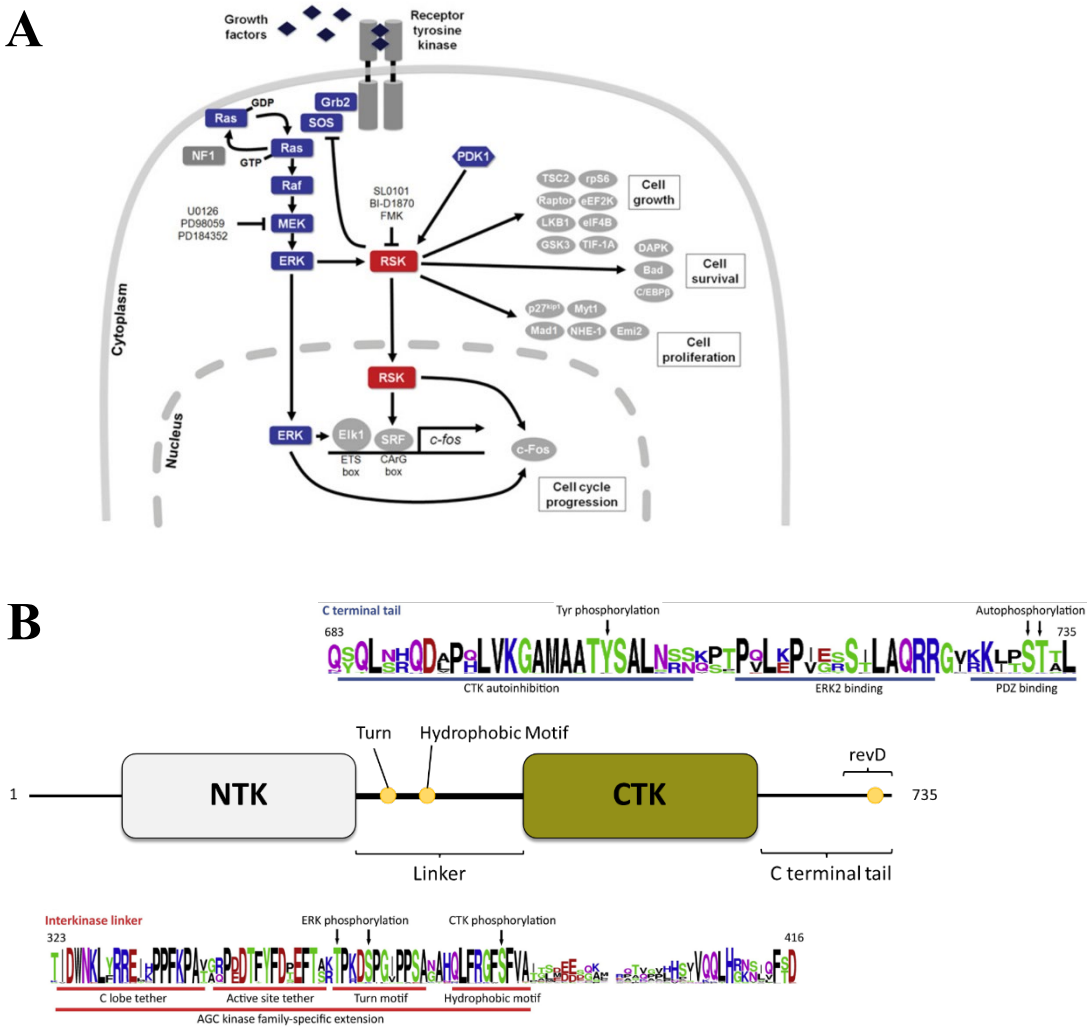


Figure 1.2. p90 RSK is an ERK effector protein and contains two distinct kinase domains.

(A) Schematic indicating that RSK lies downstream of the Ras/MAPK signaling pathway and that following activation RSK phosphorylates substrates in the cytoplasm and nucleus. Figure is from Romeo, Y., et al. *Biochem J.*, (2012).

(B) Domain organization of RSK1-4 isoforms. Each isoform has two kinase domains, an interkinase linker, and C-terminal tail. The unstructured linker and C-terminal tail contain small linear motifs (SLiMS) critical for RSK regulation.

C. RSK family domain organization

The NTK is a member of the AGC kinase family, which includes other kinases like PKA, PKC, PKG, and AKT (Pearce, Komander et al. 2010). The hallmark feature of AGC kinases is the presence of a highly conserved C-terminal *cis* regulatory tail (Kannan, Haste et al. 2007)(Fig 1.3). The tail dynamically wraps around the kinase domain, stabilizing a closed and active conformation (Batkin, Schvartz et al. 2000) (Newton 2003), and is released from the kinase core when p90Rsk is in a more open inactive conformation. The AGC tail has three important regions that interact with the kinase core, the C-lobe tether (CLT), the active site tether (AST), and N-lobe tether (NLT) (Taylor and Kornev 2011, Taylor, Keshwani et al. 2012) (Figs 1.3A,C). The CLT is the N-terminal region of the C-tail, and it is bound in a stable conformation to the C-lobe of the kinase core. Within the CLT is a conserved PXXP motif which is implicated in mediating the protein:protein interactions that anchor the CLT to the C-lobe of the kinase core (Gould, Kannan et al. 2009). This interaction exemplifies how the AGC tail can serve in a *trans* regulatory fashion. The AST, which is highly dynamic and contributes to ATP binding at the active site, contains a conserved $FD(X)_{1-2}F/Y$ motif, also known as an Ade motif, that is required for binding of ATP in the active kinase (Narayana, Cox et al. 1997). This site is also the "priming site" for binding of PDK1 (Romano, Kannan et al. 2009), the so called AGC 'master kinase' (Mora, Komander et al. 2004). Finally, the C-terminus of the AGC tail contains the N-Lobe Tether (NLT) that wraps around and stabilizes the N-lobe by binding to the αC Helix (Narayana, Cox et al. 1997, Batkin, Schvartz et al. 2000). The NLT contains the

conserved turn motif (TM) and hydrophobic motif (HM), both of which are phosphosites. The turn motif contributes to thermal stability of the kinase and in some cases contributes to the catalytic activity (Newton 2003). The HM, FXXFS, is usually a phosphosite, and fills an N-lobe pocket comprised of the αC Helix and the β4 sheet, known as the PDK1-interacting fragment (PIF) pocket (Fig 1.3B). PDK1, the AGC kinase responsible for phosphorylating the activation loops of most AGC kinases, lacks its own HM, thus PDK1 is thought to bind to the phosphorylated HM of substrate kinase thus positioning itelf to phosphorylate, in *trans*, their activation loops (Batkin, Schvartz et al. 2000, Biondi, Cheung et al. 2000, Gao, Toker et al. 2001). Recently, a Tor interaction motif (TIM) was identified in the AST of PKCβII, and in this case it controls HM phosphorylation (Baffi, Lordén et al. 2021). TIM sites are present in other AGC kinases, including RSK (Chou, Rajput et al. 2020), although their purpose and function in other kinases is not yet understood.

The CTK is a member of the CaMK superfamily and includes CaMKs I and II. The CaMK family is characterized by the presence of an autoinhibitory C-terminal helix that blocks the active site from nucleotides and substrates, suppressing basal kinase activity (Temmerman, Simon et al. 2013) (Fig 1.4A). For most CaMK members, activation loop phosphorylation does not drive formation of the active conformation. In CAMKs, they are maintained in closed conformations with intact R spines, and properly positioned αC helices. Rather, Ca^{2+}/Camodulin (CaM) or phosphorylation (in the case of the CTK) by an activating kinase causes a conformational change that releases the autoinhibitory helix from the active site,

unleashing catalytic activity (Rellos, Pike et al. 2010) (Malakhova, Tereshko et al. 2008)(Fig 1.4B).

The kinase domains of RSK are both flanked by IDRs that are embedded with conserved SLiMs, are important for RSK regulation and complex assembly. The N-terminal tail is poorly conserved among the RSK isoforms (Romeo, Zhang et al. 2012). The homologous N-tail in AGC kinases is important for thermal stability of the enzyme, and in PKA, this tail contains a myristylation site that targets PKA to the membrane as well as a stable helix that flanks the C-Lobe of the kinase core (Taylor, Søberg et al. 2021). Similarly, RSK2 N-terminal truncations are unstable (Malakhova, Kurinov et al. 2009), and RSK3 contains a predicted nuclear localization signal (NLS) (Zhao, Bjørbaek et al. 1995). Thus, the N-terminal tails in RSK likely have multiple functions similar to other AGC kinase N-tails. The linker that connects the two kinases contains the AGC C-tail and conserved motifs like the AST/Ade motif, turn, TIM, and HM (Anjum and Blenis 2008, Romeo, Zhang et al. 2012, Gógl, Kornev et al. 2019) (Fig 1.2B). Beyond the AGC C-tail, the rest of the linker is poorly conserved and contains few bulky hydrophobic residues, typical for IDRs. Little is known about its function. The RSK C-terminal tail, comprises the auto-inhibitory αL helix that suppresses CTK basal activity by blocking the active site and preventing access of subtrates to the P+1 loop that enables substrate binding (Malakhova, Tereshko et al. 2008). The rest of the C-terminal tail contains SLiMs that mediate protein:protein interactions and complex formation. This includes a conserved MAPK docking sequence, or revD motif, that enables high affinity binding to the activating kinase, ERK2 (Gavin, Ni Ainle et al. 1999, Roux, Richards et al. 2003). A type I PDZ domain

binding motif is located at the C-terminus. While there are a few known RSK PDZ domain binding partners, their specific interaction sites are relatively poorly understood. The few known PDZ domain partners are also known RSK substrates hinting that the PDZ motif might provide access to a RSK-specific class of substrates (Gógl, Biri-Kovács et al. 2018, Gógl, Biri-Kovács et al. 2019).

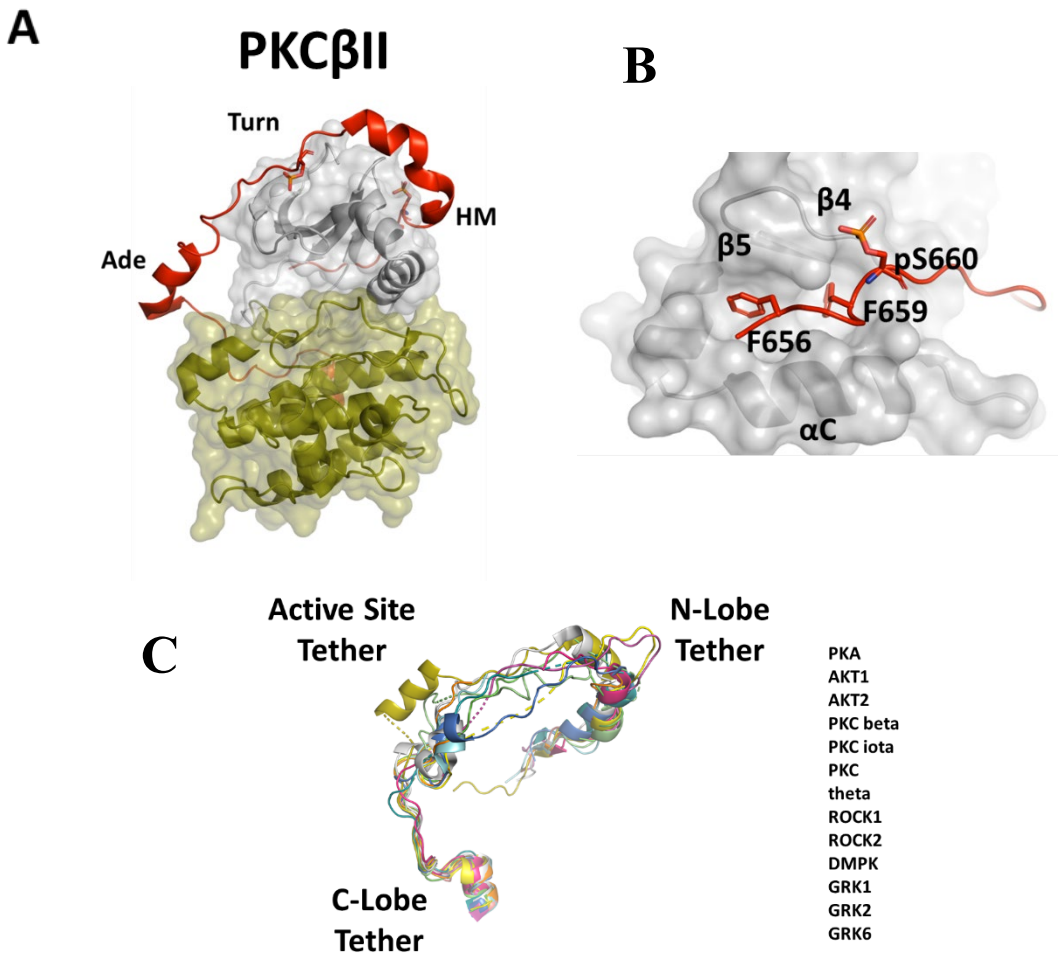


Figure 1.3. The C terminal extension is a hallmark feature of AGC kinases.
(A) Crystal structure of PKCβII, AGC tail is shown in red, as it wraps around the C-lobe, active site, and N-lobe.
(B) The hydrophobic motif (FXXFS/T) fills the so called PIF pocket in AGC pockets, which is comprised of mainly, the αC helix, and β4 strand. Phenylalanines fill a hydrophobic pocket, while the phosphorylated serine makes electrostatic interactions with the N-lobe
(C) Overlay of multiple AGC kinase C terminal tails, with the sites of the C-lobe tether (CLT), Active site tether (AST), and N-lobe tether (NLT).

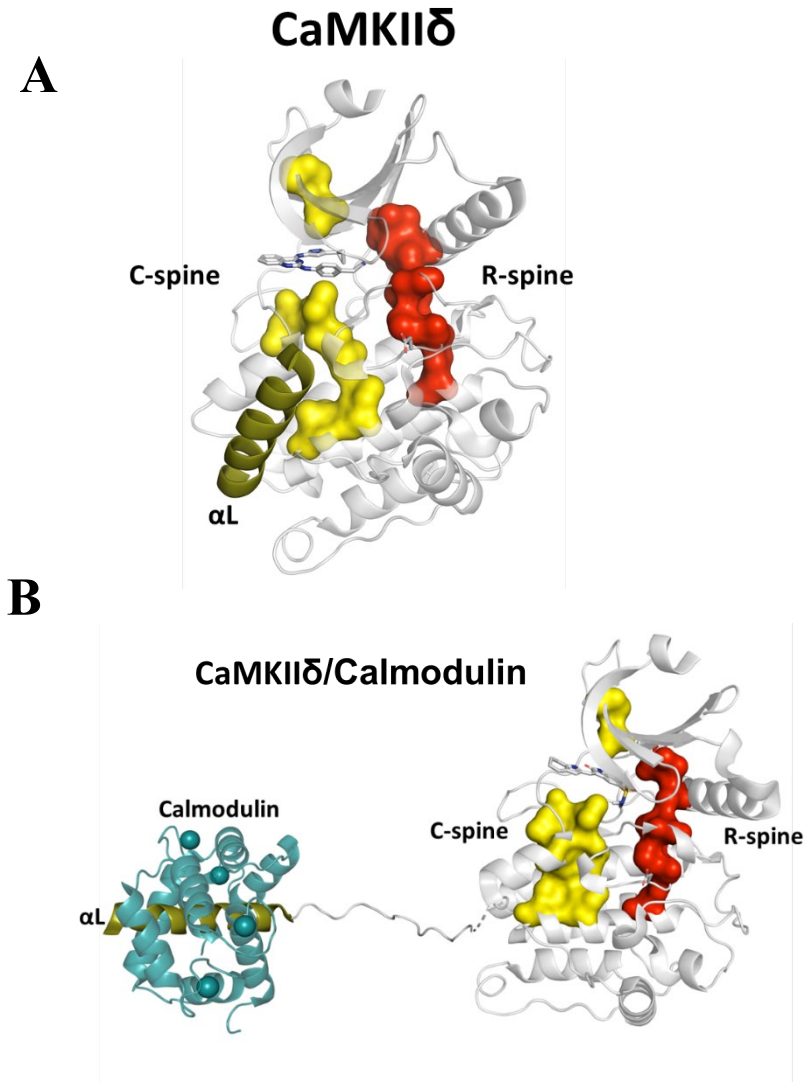


Figure 1.4. Regulation of CaMK is controlled by its C-terminal autoinhibitory helix.
(A) Crystal structure of CaMKIIδ; catalytic and regulatory spines are represented as yellow and red surfaces, respectively. The autoinhibitory helix that blocks substrate/nucleotide access is shown in green
(B) Crystal structure of CaMKIIδ with bound Ca^{2+}/calmodulin: catalytic and regulatory spines are represented as yellow and red surfaces, respectively. Ca^{2+} and calmodulin bind the autoinhibitory helix and cause a large conformational change that exposes the active site for nucleotide and substrate binding.

D. RSK activation mechanism

The RSK activation mechanism is a multi-step process that involves the action of multiple IDRs and kinase domains from different kinase families (Fig 1.5A). RSK activation is initiated following stimulation by EGF, which activates the MAPK pathway. Active ERK2, binds the revD motif in the Rsk C-terminal tail (Roux, Richards et al. 2003, Garai, Zeke et al. 2012, Alexa, Gógl et al. 2015), and phosphorylates the CTK Activation Loop (Sutherland, Campbell et al. 1993, Smith, Poteet-Smith et al. 1999), which subsequently phosphorylates the HM (Dalby, Morrice et al. 1998). The phosphorylated HM then binds PDK1, enabling phosphorylation of the NTK Activation Loop (Jensen, Buch et al. 1999) (Frödin, Jensen et al. 2000, Frödin, Antal et al. 2002). ERK2 is also suspected to phosphorylate the NTK turn motif, although when this occurs with respect to the other phosphorylation events is unclear (Dalby, Morrice et al. 1998, Richards, Dreisbach et al. 2001). An active NTK is then able to phosphorylate downstream substrates with a consensus sequence of RXRXXS/T (Leighton, Dalby et al. 1995). The NTK can then also autophosphorylate its own C-terminal tail, resulting in a decrease in ERK2 binding affinity, and this final step serves as a mechanism for negative feedback (Akamine, Madhusudan et al. 2003, Roux, Richards et al. 2003, Gógl, Biri-Kovács et al. 2018).

This multistep activation mechanism stimulated our interest in RSK and its activation. Assembly of an active kinase requires the coordinated action of four kinase domains from 3 distinct kinase families (Manning, Whyte et al. 2002) (the NTK and PDK1 belong to the AGC family, the CTK is a CAMK, and ERK2 is part of the

CMGC family) (Fig 1.5B) and IDRS. RSK forms a stable complex with ERK2, its activating complex, which is rare among substrate:activator kinase pairs, providing an opportunity to better understand how a kinase phosphorylates another kinase.

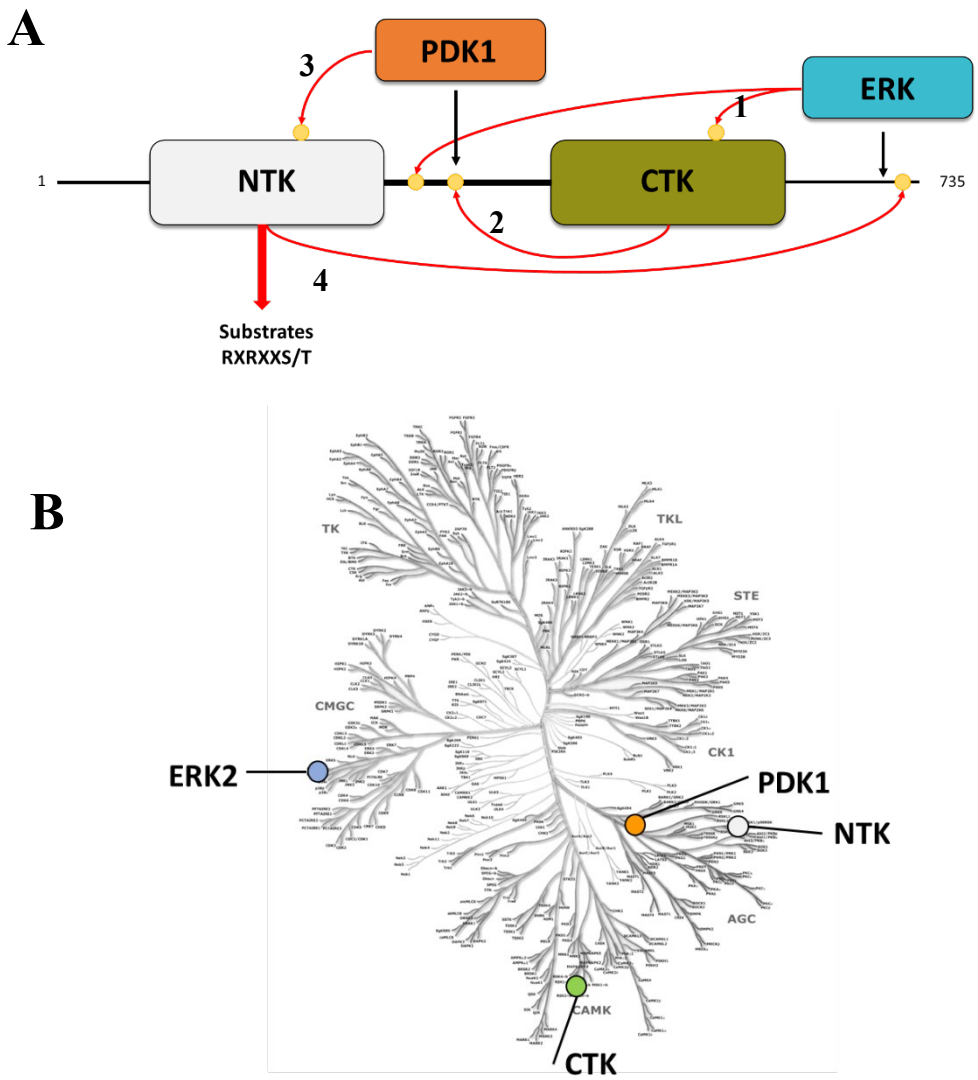


Figure 1.5. RSK activation is a multistep, multiple kinase process.
(A) Summary of the steps required for RSK activation. Black arrows represent formation of protein protein interactions and red arrows represent phosphorylation steps. The combined action of ERK2, CTK, and PDK1 result in activation of the NTK, allowing it to phosphorylate downstream substrates.
(B) Kinome tree indicating kinase families for the RSK NTK, CTK, ERK2, and PDK1. The NTK and PDK1 are AGC kinases, the CTK is a CaMK , and ERK2 is a CMGC member. Figure is adapted from Manning, G., et al. *Science*, **298**, 1912-1934 (2002).

E. RSK perturbation in disease and structural information

Misregulation of RSK is associated with a number of diseases. Mutations to RSK2 are responsible for the neurological disorder Coffin Lowry Syndrome (CLS) (Trivier, De Cesare et al. 1996) (Fig 1.6A). This disorder presents in children, and causes mental retardation, and is characterized by digit abnormalities. RSK2 CLS mutations are distributed throughout the protein, primarily in the kinase domains, and sparingly in the HM of the linker, and revD motif of the C-terminal tail. Some of these mutations alter conserved residues known to be necessary for kinase catalytic function such as the HRD motif, activation loops, and αC helices, suggesting that CLS mutations inhibit RSK2 activity. RSK mutations (Tate, Bamford et al. 2018)and differential expression patterns, the various RSK isoforms are found in various forms of cancer (Sulzmaier and Ramos 2013, Houles and Roux 2018)(Fig 1.6B). In particular RSK1/2 are particularly important for triple negative breast cancer (TNBC) (Stratford, Reipas et al. 2012).

Although there are no structures of any FL RSK isoform, crystal structures of the isolated kinase domains have been solved. Structures of the NTK lack the N-terminal tail, contain truncated AGC tails, and are bound to non-hydrolzable ATP analogues or other inhibitors. Inhibitors trap the NTK in near closed, active conformations but mostly in more drastic, distorted inactive conformations (Ikuta, Kornienko et al. 2007, Malakhova, Kurinov et al. 2009, Utepbergenov, Derewenda et al. 2012, Derewenda, Artamonov et al. 2013, Aronchik, Appleton et al. 2014, Costales, Mathur et al. 2014, Jain, Mathur et al. 2015). The CTK of RSK1/2 has been

solved with the autoinhibitory αL helix blocking the active site, while the rest of the C-terminal tail is disordered and not visible (Malakhova, Tereshko et al. 2008, Li, Fu et al. 2012, Serafimova, Pufall et al. 2012, Miller, Paavilainen et al. 2013, Krishnan, Miller et al. 2014, London, Miller et al. 2014, Andersen, Gesser et al. 2018). A major step in understanding RSK function occurred when the structure of the CTK-ERK2 heterodimer was solved. The structure of ERK2 bound to a peptide of the RSK C-terminal tail revD motif identified the primary mechanism of interaction (Garai, Zeke et al. 2012). The ERK2 D-recognition site is comprised of hydrophobic pockets and an acidic common docking (CD) surface. Hydrophobic residues in the revD fill the pockets in the DRS while, revD Args establish electrostatic interactions with the CD surface, enabling specificity and high affinity binding. The revD motif is visible in the CTK-ERK2 heterodimer structure, and confirmed this primary interaction. This complex identified a secondary interaction surface between the CTK APE motif, and the ERK2 Glycine-rich Loop, which positions the CTK activation loop to face the catalytic center of ERK2 (Alexa, Gógl et al. 2015). This structure provides a reasonable mechanism as to how one kinase phosphorylates another kinase. With the success of this structure, the follow up question remains, what is the structure of FL RSK, and in particular, what is the structure of FL RSK bound to ERK2, and how are all three kinase domains oriented relative to one another, to mediate the downstream steps of RSK activation.

Here, we have chosen to further characterize RSK and its activation because it is an intriguing model system to better understand how kinases and IDRs interact and mediate assembly of kinase complexes. RSK activation requires a variety of IDR

interactions with folded kinase domains. The RSK-ERK complex is also the rare example of a stable activator:substrate kinase pair, therefore further characterizing this complex can provide more details as to how one kinase activates another kinase. Here our goals are to further analyze the CTK-ERK2 heterodimer to determine how the interfaces in this complex evolve over time, and to use a network analysis framework to identify additional important, underappreciated residues. A final goal is to characterize, biophysically, the FL RSK-ERK complex and attempt to determine its structure by cryoEM or xray crystallography. A long-term goal is to perform biochemical analysis to further elucidate the molecular mechanism for the remaining steps in RSK activation, in particular for activation of RSK by PDK1.

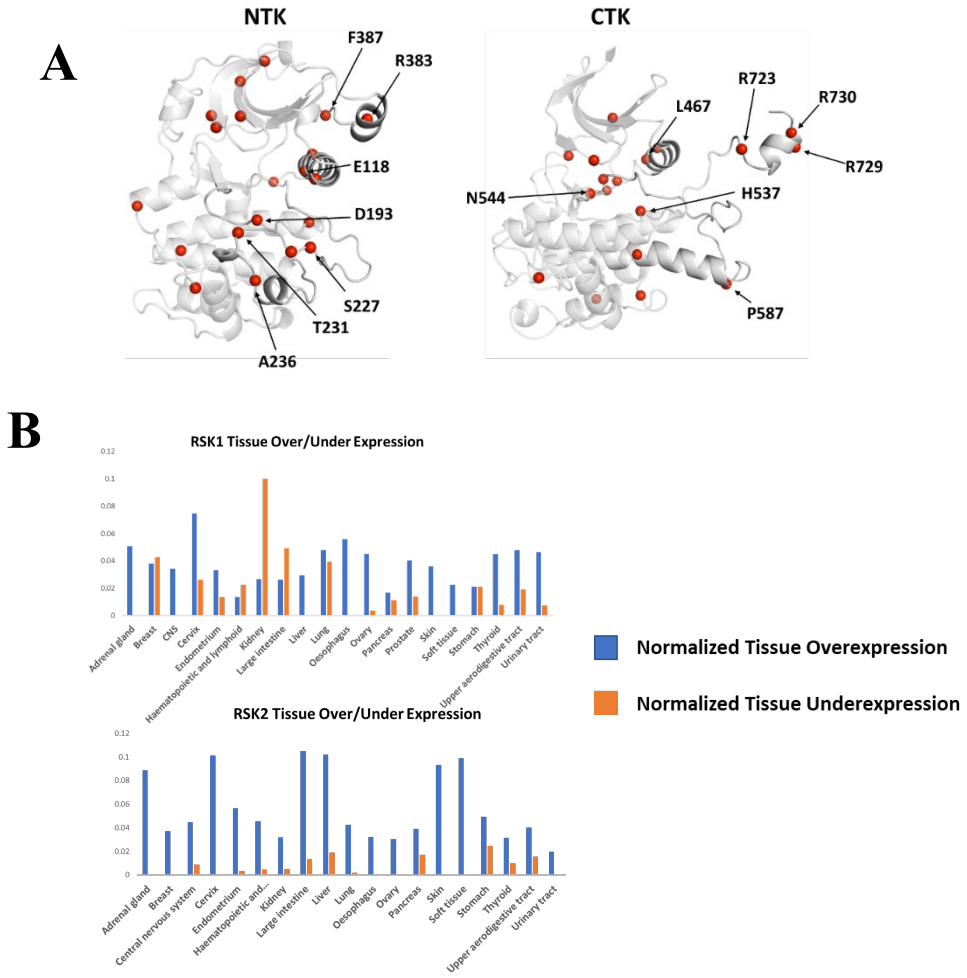


Figure 1.6. Mutations in RSK2 drive Coffin Lowry Syndrome (CLS) and perturbed expression levels are found in many forms of cancer.
(A) CLS mutations are overlaid onto AF generated models of the NTK through the HM, and the CTK through the C-terminal tail. Red spheres indicate positions of Cα of mutated residues. CLS mutations are distributed throughout both kinases, and mutations are found to highly conserved residues known to be important for kinase activity.
(B) Frequency of over or underexpression of RSK1 and RSK2 in tumors of various tissue types. Data is from COSMIC database, which defines RSK1/2 as being over or under expressed in tumor cells from various tissue types. Total counts of over/underexpressed are normalized by the maximum number of samples found per tissue type.

F. References

C Akamine, P., Madhusudan, J. Wu, N. H. Xuong, L. F. Ten Eyck and S. S. Taylor (2003). "Dynamic features of cAMP-dependent protein kinase revealed by apoenzyme crystal structure." J Mol Biol **327**(1): 159-171.

Alexa, A., G. Gógl, G. Glatz, Á. Garai, A. Zeke, J. Varga, E. Dudás, N. Jeszenői, A. Bodor, C. Hetényi and A. Reményi (2015). "Structural assembly of the signaling competent ERK2-RSK1 heterodimeric protein kinase complex." Proc Natl Acad Sci U S A **112**(9): 2711-2716.

Andersen, J. L., B. Gesser, E. D. Funder, C. J. F. Nielsen, H. Gotfred-Rasmussen, M. K. Rasmussen, R. Toth, K. V. Gothelf, J. S. C. Arthur, L. Iversen and P. Nissen (2018). "Dimethyl fumarate is an allosteric covalent inhibitor of the p90 ribosomal S6 kinases." Nat Commun **9**(1): 4344.

Anjum, R. and J. Blenis (2008). "The RSK family of kinases: emerging roles in cellular signalling." Nature Reviews Molecular Cell Biology **9**(10): 747-758.

Aronchik, I., B. A. Appleton, S. E. Basham, K. Crawford, M. Del Rosario, L. V. Doyle, W. F. Estacio, J. Lan, M. K. Lindvall, C. A. Luu, E. Ornelas, E. Venetsanakos, C. M. Shafer and A. B. Jefferson (2014). "Novel potent and selective inhibitors of p90 ribosomal S6 kinase reveal the heterogeneity of RSK function in MAPK-driven cancers." Mol Cancer Res **12**(5): 803-812.

Ascierto, P. A., J. M. Kirkwood, J. J. Grob, E. Simeone, A. M. Grimaldi, M. Maio, G. Palmieri, A. Testori, F. M. Marincola and N. Mozzillo (2012). "The role of BRAF V600 mutation in melanoma." J Transl Med **10**: 85.

Baffi, T. R., G. Lordén, J. M. Wozniak, A. Feichtner, W. Yeung, A. P. Kornev, C. C. King, J. C. Del Rio, A. J. Limaye, J. Bogomolovas, C. M. Gould, J. Chen, E. J. Kennedy, N. Kannan, D. J. Gonzalez, E. Stefan, S. S. Taylor and A. C. Newton (2021). "mTORC2 controls the activity of PKC and Akt by phosphorylating a conserved TOR interaction motif." Sci Signal **14**(678).

Barker, W. C. and M. O. Dayhoff (1982). "Viral src gene products are related to the catalytic chain of mammalian cAMP-dependent protein kinase." Proc Natl Acad Sci U S A **79**(9): 2836-2839.

Bastidas, A. C., M. S. Deal, J. M. Steichen, Y. Guo, J. Wu and S. S. Taylor (2013). "Phosphoryl transfer by protein kinase A is captured in a crystal lattice." J Am Chem Soc **135**(12): 4788-4798.

Batkin, M., I. Schvartz and S. Shaltiel (2000). "Snapping of the carboxyl terminal tail of the catalytic subunit of PKA onto its core: characterization of the sites by mutagenesis." Biochemistry **39**(18): 5366-5373.

Biondi, R. M., P. C. Cheung, A. Casamayor, M. Deak, R. A. Currie and D. R. Alessi (2000). "Identification of a pocket in the PDK1 kinase domain that interacts with PIF and the C-terminal residues of PKA." Embo j **19**(5): 979-988.

Bjørbaek, C., Y. Zhao and D. E. Moller (1995). "Divergent functional roles for p90rsk kinase domains." J Biol Chem **270**(32): 18848-18852.

Chou, P. C., S. Rajput, X. Zhao, C. Patel, D. Albaciete, W. J. Oh, H. Q. Daguplo, N. Patel, B. Su, G. Werlen and E. Jacinto (2020). "mTORC2 Is Involved in the

Induction of RSK Phosphorylation by Serum or Nutrient Starvation." Cells **9**(7).
Cohen, P. (2002). "The origins of protein phosphorylation." Nat Cell Biol **4**(5): E127-130.
Costales, A., M. Mathur, S. Ramurthy, J. Lan, S. Subramanian, R. Jain, G. Atallah, L. Setti, M. Lindvall, B. A. Appleton, E. Ornelas, P. Feucht, B. Warne, L. Doyle, S. E. Basham, I. Aronchik, A. B. Jefferson and C. M. Shafer (2014). "2-Amino-7-substituted benzoxazole analogs as potent RSK2 inhibitors." Bioorg Med Chem Lett **24**(6): 1592-1596.
Dalby, K. N., N. Morrice, F. B. Caudwell, J. Avruch and P. Cohen (1998). "Identification of regulatory phosphorylation sites in mitogen-activated protein kinase (MAPK)-activated protein kinase-1a/p90rsk that are inducible by MAPK." J Biol Chem **273**(3): 1496-1505.
Derewenda, U., M. Artamonov, G. Szukalska, D. Utepbergenov, N. Olekhnovich, H. I. Parikh, G. E. Kellogg, A. V. Somlyo and Z. S. Derewenda (2013). "Identification of quercitrin as an inhibitor of the p90 S6 ribosomal kinase (RSK): structure of its complex with the N-terminal domain of RSK2 at 1.8 Å resolution." Acta Crystallogr D Biol Crystallogr **69**(Pt 2): 266-275.
Dümmler, B. A., C. Hauge, J. Silber, H. G. Yntema, L. S. Kruse, B. Kofoed, B. A. Hemmings, D. R. Alessi and M. Frödin (2005). "Functional characterization of human RSK4, a new 90-kDa ribosomal S6 kinase, reveals constitutive activation in most cell types." J Biol Chem **280**(14): 13304-13314.
Fisher, T. L. and J. Blenis (1996). "Evidence for two catalytically active kinase domains in pp90rsk." Mol Cell Biol **16**(3): 1212-1219.
Frödin, M., T. L. Antal, B. A. Dümmler, C. J. Jensen, M. Deak, S. Gammeltoft and R. M. Biondi (2002). "A phosphoserine/threonine-binding pocket in AGC kinases and PDK1 mediates activation by hydrophobic motif phosphorylation." The EMBO Journal **21**(20): 5396-5407.
Frödin, M., C. J. Jensen, K. Merienne and S. Gammeltoft (2000). "A phosphoserine-regulated docking site in the protein kinase RSK2 that recruits and activates PDK1." Embo j **19**(12): 2924-2934.
Gao, T., A. Toker and A. C. Newton (2001). "The carboxyl terminus of protein kinase c provides a switch to regulate its interaction with the phosphoinositide-dependent kinase, PDK-1." J Biol Chem **276**(22): 19588-19596.
Garai, Á., A. Zeke, G. Gógl, I. Törő, F. Fördős, H. Blankenburg, T. Bárkai, J. Varga, A. Alexa, D. Emig, M. Albrecht and A. Reményi (2012). "Specificity of Linear Motifs That Bind to a Common Mitogen-Activated Protein Kinase Docking Groove." Science Signaling **5**(245): ra74-ra74.
Gavin, A. C., A. Ni Ainle, E. Chierici, M. Jones and A. R. Nebreda (1999). "A p90(rsk) mutant constitutively interacting with MAP kinase uncouples MAP kinase from p34(cdc2)/cyclin B activation in Xenopus oocytes." Mol Biol Cell **10**(9): 2971-2986.
Gibbs, C. S. and M. J. Zoller (1991). "Rational scanning mutagenesis of a protein kinase identifies functional regions involved in catalysis and substrate interactions." Journal of Biological Chemistry **266**(14): 8923-8931.

Gimple, R. C. and X. Wang (2019). "RAS: Striking at the Core of the Oncogenic Circuitry." Frontiers in Oncology **9**.

Gógl, G., B. Biri-Kovács, F. Durbesson, P. Jane, Y. Nomine, C. Kostmann, V. Bilics, M. Simon, A. Reményi, R. Vincentelli, G. Trave and L. Nyitray (2019). "Rewiring of RSK-PDZ Interactome by Linear Motif Phosphorylation." J Mol Biol **431**(6): 1234-1249.

Gógl, G., B. Biri-Kovács, L. Póti Á, H. Vadászi, B. Szeder, A. Bodor, G. Schlosser, A. Ács, L. Turiák, L. Buday, A. Alexa, L. Nyitray and A. Reményi (2018). "Dynamic control of RSK complexes by phosphoswitch-based regulation." Febs j **285**(1): 46-71.

Gógl, G., A. P. Kornev, A. Reményi and S. S. Taylor (2019). "Disordered Protein Kinase Regions in Regulation of Kinase Domain Cores." Trends Biochem Sci **44**(4): 300-311.

Gould, C. M., N. Kannan, S. S. Taylor and A. C. Newton (2009). "The chaperones Hsp90 and Cdc37 mediate the maturation and stabilization of protein kinase C through a conserved PXXP motif in the C-terminal tail." J Biol Chem **284**(8): 4921-4935.

Hanks, S. K. and T. Hunter (1995). "Protein kinases 6. The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification." Faseb j **9**(8): 576-596.

Hari, S. B., E. A. Merritt and D. J. Maly (2013). "Sequence determinants of a specific inactive protein kinase conformation." Chem Biol **20**(6): 806-815.

Houles, T. and P. P. Roux (2018). "Defining the role of the RSK isoforms in cancer." Seminars in Cancer Biology **48**: 53-61.

Huang, L., Z. Guo, F. Wang and L. Fu (2021). "KRAS mutation: from undruggable to druggable in cancer." Signal Transduction and Targeted Therapy **6**(1): 386.

Ikuta, M., M. Kornienko, N. Byrne, J. C. Reid, S. Mizuarai, H. Kotani and S. K. Munshi (2007). "Crystal structures of the N-terminal kinase domain of human RSK1 bound to three different ligands: Implications for the design of RSK1 specific inhibitors." Protein Sci **16**(12): 2626-2635.

Jain, R., M. Mathur, J. Lan, A. Costales, G. Atallah, S. Ramurthy, S. Subramanian, L. Setti, P. Feucht, B. Warne, L. Doyle, S. Basham, A. B. Jefferson, M. Lindvall, B. A. Appleton and C. M. Shafer (2015). "Discovery of Potent and Selective RSK Inhibitors as Biological Probes." J Med Chem **58**(17): 6766-6783.

Jensen, C. J., M. B. Buch, T. O. Krag, B. A. Hemmings, S. Gammeltoft and M. Frödin (1999). "90-kDa ribosomal S6 kinase is phosphorylated and activated by 3-phosphoinositide-dependent protein kinase-1." J Biol Chem **274**(38): 27168-27176.

Kannan, N., N. Haste, S. S. Taylor and A. F. Neuwald (2007). "The hallmark of AGC kinase functional divergence is its C-terminal tail, a cis-acting regulatory module." Proceedings of the National Academy of Sciences **104**(4): 1272-1277.

Kholodenko, B. N., J. F. Hancock and W. Kolch (2010). "Signalling ballet in space and time." Nat Rev Mol Cell Biol **11**(6): 414-426.

Knighton, D. R., J. H. Zheng, L. F. Ten Eyck, V. A. Ashford, N. H. Xuong, S. S. Taylor and J. M. Sowadski (1991). "Crystal structure of the catalytic subunit of cyclic adenosine monophosphate-dependent protein kinase." Science **253**(5018): 407-414.

Knighton, D. R., J. H. Zheng, L. F. Ten Eyck, N. H. Xuong, S. S. Taylor and J. M. Sowadski (1991). "Structure of a peptide inhibitor bound to the catalytic subunit of cyclic adenosine monophosphate-dependent protein kinase." Science **253**(5018): 414-420.

Kornev, A. P., N. M. Haste, S. S. Taylor and L. F. Eyck (2006). "Surface comparison of active and inactive protein kinases identifies a conserved activation mechanism." Proc Natl Acad Sci U S A **103**(47): 17783-17788.

Kornev, A. P. and S. S. Taylor (2015). "Dynamics-Driven Allostery in Protein Kinases." Trends Biochem Sci **40**(11): 628-647.

Kornev, A. P., S. S. Taylor and L. F. Ten Eyck (2008). "A helix scaffold for the assembly of active protein kinases." Proc Natl Acad Sci U S A **105**(38): 14377-14382.

Krebs, E. G. (1998). "An accidental biochemist." Annu Rev Biochem **67**: xii-xxxii.

Krishnan, S., R. M. Miller, B. Tian, R. D. Mullins, M. P. Jacobson and J. Taunton (2014). "Design of reversible, cysteine-targeted Michael acceptors guided by kinetic and computational analysis." J Am Chem Soc **136**(36): 12624-12630.

Leighton, I. A., K. N. Dalby, F. B. Caudwell, P. T. Cohen and P. Cohen (1995). "Comparison of the specificities of p70 S6 kinase and MAPKAP kinase-1 identifies a relatively specific substrate for p70 S6 kinase: the N-terminal kinase domain of MAPKAP kinase-1 is essential for peptide phosphorylation." FEBS Lett **375**(3): 289-293.

Li, D., T. M. Fu, J. Nan, C. Liu, L. F. Li and X. D. Su (2012). "Structural basis for the autoinhibition of the C-terminal kinase domain of human RSK1." Acta Crystallogr D Biol Crystallogr **68**(Pt 6): 680-685.

Liao, J. J.-L. (2007). "Molecular Recognition of Protein Kinase Binding Pockets for Design of Potent and Selective Kinase Inhibitors." Journal of Medicinal Chemistry **50**(3): 409-424.

London, N., R. M. Miller, S. Krishnan, K. Uchida, J. J. Irwin, O. Eidam, L. Gibold, P. Cimermančič, R. Bonnet, B. K. Shoichet and J. Taunton (2014). "Covalent docking of large libraries for the discovery of chemical probes." Nat Chem Biol **10**(12): 1066-1072.

Malakhova, M., I. Kurinov, K. Liu, D. Zheng, I. D'Angelo, J. H. Shim, V. Steinman, A. M. Bode and Z. Dong (2009). "Structural diversity of the active N-terminal kinase domain of p90 ribosomal S6 kinase 2." PLoS One **4**(11): e8044.

Malakhova, M., V. Tereshko, S. Y. Lee, K. Yao, Y. Y. Cho, A. Bode and Z. Dong (2008). "Structural basis for activation of the autoinhibitory C-terminal kinase domain of p90 RSK2." Nat Struct Mol Biol **15**(1): 112-113.

Manning, G., D. B. Whyte, R. Martinez, T. Hunter and S. Sudarsanam (2002). "The protein kinase complement of the human genome." Science **298**(5600): 1912-1934.

Masterson, L. R., A. Cembran, L. Shi and G. Veglia (2012). "Allostery and binding cooperativity of the catalytic subunit of protein kinase A by NMR spectroscopy and molecular dynamics simulations." Adv Protein Chem Struct Biol **87**: 363-389.

Masterson, L. R., C. Cheng, T. Yu, M. Tonelli, A. Kornev, S. S. Taylor and G. Veglia (2010). "Dynamics connect substrate recognition to catalysis in protein kinase A." Nat Chem Biol **6**(11): 821-828.

Masterson, L. R., L. Shi, E. Metcalfe, J. Gao, S. S. Taylor and G. Veglia (2011). "Dynamically committed, uncommitted, and quenched states encoded in protein kinase A revealed by NMR spectroscopy." Proc Natl Acad Sci U S A **108**(17): 6969-6974.

Miller, R. M., V. O. Paavilainen, S. Krishnan, I. M. Serafimova and J. Taunton (2013). "Electrophilic fragment-based design of reversible covalent kinase inhibitors." J Am Chem Soc **135**(14): 5298-5301.

Möbitz, H. (2015). "The ABC of protein kinase conformations." Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics **1854**(10, Part B): 1555-1566.

Mora, A., D. Komander, D. M. van Aalten and D. R. Alessi (2004). "PDK1, the master regulator of AGC kinase signal transduction." Semin Cell Dev Biol **15**(2): 161-170.

Narayana, N., S. Cox, X. Nguyen-huu, L. F. Ten Eyck and S. S. Taylor (1997). "A binary complex of the catalytic subunit of cAMP-dependent protein kinase and adenosine further defines conformational flexibility." Structure **5**(7): 921-935.

Narayana, N., S. Cox, S. Shaltiel, S. S. Taylor and N.-h. Xuong (1997). "Crystal Structure of a Polyhistidine-Tagged Recombinant Catalytic Subunit of cAMP-Dependent Protein Kinase Complexed with the Peptide Inhibitor PKI(5−24) and Adenosine." Biochemistry **36**(15): 4438-4448.

Newton, A. C. (2003). "Regulation of the ABC kinases by phosphorylation: protein kinase C as a paradigm." Biochem J **370**(Pt 2): 361-371.

Newton, A. C. (2018). "Protein kinase C: perfectly balanced." Crit Rev Biochem Mol Biol **53**(2): 208-230.

Nolen, B., S. Taylor and G. Ghosh (2004). "Regulation of protein kinases; controlling activity through activation segment conformation." Mol Cell **15**(5): 661-675.

Pearce, L. R., D. Komander and D. R. Alessi (2010). "The nuts and bolts of AGC protein kinases." Nat Rev Mol Cell Biol **11**(1): 9-22.

Rellos, P., A. C. Pike, F. H. Niesen, E. Salah, W. H. Lee, F. von Delft and S. Knapp (2010). "Structure of the CaMKIIdelta/calmodulin complex reveals the molecular mechanism of CaMKII kinase activation." PLoS Biol **8**(7): e1000426.

Richards, S. A., V. C. Dreisbach, L. O. Murphy and J. Blenis (2001). "Characterization of regulatory events associated with membrane targeting of p90 ribosomal S6 kinase 1." Mol Cell Biol **21**(21): 7470-7480.

Romano, R. A., N. Kannan, A. P. Kornev, C. J. Allison and S. S. Taylor (2009). "A chimeric mechanism for polyvalent trans-phosphorylation of PKA by PDK1." Protein Science **18**(7): 1486-1497.

Romeo, Y., X. Zhang and P. P. Roux (2012). "Regulation and function of the RSK family of protein kinases." Biochem J **441**(2): 553-569.
Roskoski, R., Jr. (2021). "Properties of FDA-approved small molecule protein kinase inhibitors: A 2021 update." Pharmacol Res **165**: 105463.
Roux, P. P., S. A. Richards and J. Blenis (2003). "Phosphorylation of p90 ribosomal S6 kinase (RSK) regulates extracellular signal-regulated kinase docking and RSK activity." Mol Cell Biol **23**(14): 4796-4804.
Serafimova, I. M., M. A. Pufall, S. Krishnan, K. Duda, M. S. Cohen, R. L. Maglathlin, J. M. McFarland, R. M. Miller, M. Frödin and J. Taunton (2012). "Reversible targeting of noncatalytic cysteines with chemically tuned electrophiles." Nat Chem Biol **8**(5): 471-476.
Smith, J. A., C. E. Poteet-Smith, K. Malarkey and T. W. Sturgill (1999). "Identification of an extracellular signal-regulated kinase (ERK) docking site in ribosomal S6 kinase, a sequence critical for activation by ERK in vivo." J Biol Chem **274**(5): 2893-2898.
Srivastava, A. K., L. R. McDonald, A. Cembran, J. Kim, L. R. Masterson, C. L. McClendon, S. S. Taylor and G. Veglia (2014). "Synchronous opening and closing motions are essential for cAMP-dependent protein kinase A signaling." Structure **22**(12): 1735-1743.
Steichen, J. M., G. H. Iyer, S. Li, S. A. Saldanha, M. S. Deal, V. L. Woods, Jr. and S. S. Taylor (2010). "Global consequences of activation loop phosphorylation on protein kinase A." J Biol Chem **285**(6): 3825-3832.
Steichen, J. M., M. Kuchinskas, M. M. Keshwani, J. Yang, J. A. Adams and S. S. Taylor (2012). "Structural basis for the regulation of protein kinase A by activation loop phosphorylation." J Biol Chem **287**(18): 14672-14680.
Stratford, A. L., K. Reipas, K. Hu, A. Fotovati, R. Brough, J. Frankum, M. Takhar, P. Watson, A. Ashworth, C. J. Lord, A. Lasham, C. G. Print and S. E. Dunn (2012). "Targeting p90 ribosomal S6 kinase eliminates tumor-initiating cells by inactivating Y-box binding protein-1 in triple-negative breast cancers." Stem cells (Dayton, Ohio) **30**(7): 1338-1348.
Sulzmaier, F. J. and J. W. Ramos (2013). "RSK isoforms in cancer cell invasion and metastasis." Cancer Res **73**(20): 6099-6105.
Sutherland, C., D. G. Campbell and P. Cohen (1993). "Identification of insulin-stimulated protein kinase-1 as the rabbit equivalent of rskmo-2. Identification of two threonines phosphorylated during activation by mitogen-activated protein kinase." Eur J Biochem **212**(2): 581-588.
Tate, J. G., S. Bamford, H. C. Jubb, Z. Sondka, D. M. Beare, N. Bindal, H. Boutselakis, C. G. Cole, C. Creatore, E. Dawson, P. Fish, B. Harsha, C. Hathaway, S. C. Jupe, C. Y. Kok, K. Noble, L. Ponting, C. C. Ramshaw, C. E. Rye, H. E. Speedy, R. Stefancsik, S. L. Thompson, S. Wang, S. Ward, P. J. Campbell and S. A. Forbes (2018). "COSMIC: the Catalogue Of Somatic Mutations In Cancer." Nucleic Acids Research **47**(D1): D941-D947.
Taylor, S. S., R. Ilouz, P. Zhang and A. P. Kornev (2012). "Assembly of allosteric macromolecular switches: lessons from PKA." Nat Rev Mol Cell Biol **13**(10): 646-658.

Taylor, S. S., M. M. Keshwani, J. M. Steichen and A. P. Kornev (2012). "Evolution of the eukaryotic protein kinases as dynamic molecular switches." Philosophical Transactions of the Royal Society B: Biological Sciences **367**(1602): 2517-2528.

Taylor, S. S. and A. P. Kornev (2011). "Protein kinases: evolution of dynamic regulatory proteins." Trends Biochem Sci **36**(2): 65-77.

Taylor, S. S., A. S. Shaw, N. Kannan and A. P. Kornev (2015). "Integration of signaling in the kinome: Architecture and regulation of the αC Helix." Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics **1854**(10, Part B): 1567-1574.

Taylor, S. S., K. Søberg, E. Kobori, J. Wu, S. Pautz, F. W. Herberg and B. S. Skålhegg (2021). "The tails of PKA." Molecular Pharmacology: MOLPHARM-MR-2021-000315.

Taylor, S. S., J. Yang, J. Wu, N. M. Haste, E. Radzio-Andzelm and G. Anand (2004). "PKA: a portrait of protein kinase dynamics." Biochim Biophys Acta **1697**(1-2): 259-269.

Temmerman, K., B. Simon and M. Wilmanns (2013). "Structural and functional diversity in the activity and regulation of DAPK-related protein kinases." Febs j **280**(21): 5533-5550.

Trivier, E., D. De Cesare, S. Jacquot, S. Pannetier, E. Zackai, I. Young, J. L. Mandel, P. Sassone-Corsi and A. Hanauer (1996). "Mutations in the kinase Rsk-2 associated with Coffin-Lowry syndrome." Nature **384**(6609): 567-570.

Utepbergenov, D., U. Derewenda, N. Olekhnovich, G. Szukalska, B. Banerjee, M. K. Hilinski, D. A. Lannigan, P. T. Stukenberg and Z. S. Derewenda (2012). "Insights into the inhibition of the p90 ribosomal S6 kinase (RSK) by the flavonol glycoside SL0101 from the 1.5 Å crystal structure of the N-terminal domain of RSK2 with bound inhibitor." Biochemistry **51**(33): 6499-6510.

Walsh, D. A., J. P. Perkins and E. G. Krebs (1968). "An adenosine 3',5'-monophosphate-dependant protein kinase from rabbit skeletal muscle." J Biol Chem **243**(13): 3763-3765.

Wong, L., P. A. Jennings and J. A. Adams (2004). "Communication pathways between the nucleotide pocket and distal regulatory sites in protein kinases." Acc Chem Res **37**(5): 304-311.

Wu, P., T. E. Nielsen and M. H. Clausen (2015). "FDA-approved small-molecule kinase inhibitors." Trends Pharmacol Sci **36**(7): 422-439.

Yang, J., S. M. Garrod, M. S. Deal, G. S. Anand, V. L. Woods, Jr. and S. Taylor (2005). "Allosteric network of cAMP-dependent protein kinase revealed by mutation of Tyr204 in the P+1 loop." J Mol Biol **346**(1): 191-201.

Zaman, A., W. Wu and T. G. Bivona (2019). "Targeting Oncogenic BRAF: Past, Present, and Future." Cancers (Basel) **11**(8).

Zeniou, M., T. Ding, E. Trivier and A. Hanauer (2002). "Expression analysis of RSK gene family members: the RSK2 gene, mutated in Coffin-Lowry syndrome, is prominently expressed in brain structures essential for cognitive function and learning." Hum Mol Genet **11**(23): 2929-2940.

Zhao, Y., C. Bjørbaek, S. Weremowicz, C. C. Morton and D. E. Moller (1995). "RSK3 encodes a novel pp90rsk isoform with a unique N-terminal sequence: growth factor-stimulated kinase function and nuclear translocation." Mol Cell Biol **15**(8): 4353-4363.

Zheng, J., D. R. Knighton, N. H. Xuong, S. S. Taylor, J. M. Sowadski and L. F. Ten Eyck (1993). "Crystal structures of the myristylated catalytic subunit of cAMP-dependent protein kinase reveal open and closed conformations." Protein Sci **2**(10): 1559-1573.

Chapter II

Network Analysis of the RSK1

CTK:ERK2

A. Introduction

Protein-protein interfaces (PPI) are ubiquitous in biology and play fundamental roles in a variety of processes (Yan, Wu et al. 2008, Esmaielbeiki, Krawczyk et al. 2016). Estimates of the size of the human interactome place the number of human PPI's around 650,000 (Stumpf, Thorne et al. 2008). Genetic analysis of disease mutations indicate that mutations tend to perturb PPIs rather than affect overall protein folding and stability (Cheng, Zhao et al. 2021). Thus, further analysis of PPI's, their dynamics, and modulation remain of great interest to both academic and drug discovery pipelines. These interfaces tend to be flat with large surface areas, and current PPI targeting small molecule docking sites are bigger and more hydrophobic with lower solubility. Therefore, the potential of allosteric modulators of PPI's remains a "holy grail" for rational drug design, as allosteric sites are typically more amenable for small molecules(Rask-Andersen, Almén et al. 2011, Ni, Liu et al. 2019, Ni, Lu et al. 2019, Shen, Yuan et al. 2019). Allostery is a fundamental regulatory mechanism that has been recognized for decades, but the molecular mechanism of how an allosteric signal is communicated within a protein remains unclear. Traditionally, allostery is thought to involve significant structural or conformational changes, driven by enthalpic changes to the protein (Gunasekaran, Ma et al. 2004, Gautier, Laursen et al. 2019, Wodak, Paci et al. 2019). However, it was also shown that allostery can be an entropy-driven process that transmits allosteric signals through modulation of a protein's thermal vibrations, without significant conformational changes (Kornev and Taylor 2015, Chopra, Wales et al. 2016, Ahuja, Kornev et al.

2017, Gourinchas, Etzl et al. 2017, Kornev 2018). One of the major goals of this project is, as a proof-of-principle, to allosterically modulate a PPI and alter protein complex binding affinities via redistribution of the protein's internal thermal vibrations.

Graph theory and network analysis approaches have been used to analyze networks across a variety of different fields (Christley, Pinchbeck et al. 2005, Boccaletti, Latora et al. 2006, Yu, Kim et al. 2007, Reggiani, Nijkamp et al. 2015), including the analysis of different proteins (Dokholyan, Li et al. 2002, Amitai, Shemesh et al. 2004, Blacklock and Verkhivker 2014, Lee, Choi et al. 2014, Kapetis, Sassone et al. 2017, Negre, Morzan et al. 2018). In these types of networks, graphs of proteins consist of nodes which represent individual amino acids and edges connect "interacting" amino acids or nodes (Kannan and Vishveshwara 1999, Vendruscolo, Paci et al. 2001, Sethi, Eargle et al. 2009, Tiberti, Invernizzi et al. 2014). Recently, a methodology for the construction of these networks and recommendations for centrality metrics used to determine important residues was proposed using Protein Kinase A (PKA) as a model system. A local spatial alignment (LSP) algorithm (Kornev, Haste et al. 2006, Kornev, Taylor et al. 2008) connected nodes with correlated motions that behave as semi-rigid bodies over the course of a molecular dynamics (MD) simulation. There are multiple centrality metrics to detect critical network elements, and out of all metrics tested, degree centrality (DC), and betweenness centrality (BC), most efficiently detected highly conserved residues in PKA (Kornev, Aoto et al. 2022). The DC of a node is a local measure, and it is the number of weighted connections to other nodes. BC is a global parameter that

represents the number of instances a node lies in the shortest path between any other two nodes (Boccaletti, Latora et al. 2006). BC can be thought of as a measure of information flow through a given node. In various network studies of yeast interactome datasets, important elements tend to be classified as “party-hubs”, nodes with high DC and low BC, “non-hub bottlenecks” nodes with high BC, low DC, or “date-hubs” or “hub bottlenecks” nodes with high DC and high BC. Disruption of these important “party-hub” or “bottleneck” nodes have significant functional consequences for the network (Estrada 2006, Yu, Kim et al. 2007, Pandini, Fornili et al. 2012). In agreement with previous studies, for PKA, central, highly connected N and C-lobe residues that are not involved in the communication transfer in the network have high DC but low to moderate BC. In contrast, highly conserved residues such as the DFG motif aspartate, K72 (N-lobe), and Y/HRD motifs had both high DC and high BC, indicating that these highly conserved, functionally important residues are major community forming residues that are critical for the flow of information within the network (Kornev, Aoto et al. 2022).

For this project, we focus on the RSK1 CTK-ERK2 heterodimeric complex as a model system (Alexa, Gógl et al. 2015). P90 ribosomal S6 kinase (RSK1-4) is a member of the MAPKAPK family and is specifically activated by the MAPK ERK1/2. RSK is a multi-kinase domain protein that contains an N-terminal kinase (NTK) which is a member of the AGC kinase family, and a C-terminal kinase (CTK) which belongs to the CAMK family (Anjum and Blenis 2008, Romeo, Zhang et al. 2011). The C-terminus of RSK also contains a conserved D motif (712-730), or so-called “revD motif” that binds to the D-recruitment site (DRS) of ERK with specificity and high

affinity (Tanoue, Adachi et al. 2000, Roux, Richards et al. 2003, Garai, Zeke et al. 2012). Once bound to RSK, ERK phosphorylates the activation loop of the CTK ($T573^{RSK1}$) (Sutherland, Campbell et al. 1993, Smith, Poteet-Smith et al. 1999), which initiates a series of phosphorylation events that result in the activation of the NTK, enabling phosphorylation of downstream substrates. RSK phosphorylates a wide range of substrates that are involved in a variety of processes such as cell growth, proliferation, and cell division (Anjum and Blenis 2008, Romeo, Zhang et al. 2011). The two kinase domains of p90rSK are both misregulated in a number of diseases, and atypical ERK1/2 signaling is implicated in a variety of diseases including various forms of cancer, long-term potentiation and epithelium maintenance, as well as a potential role in diabetes (Lawrence, Jivan et al. 2008). Mutations to RSK2 cause Coffin Lowry Syndrome (CLS), and there is growing evidence that RSK mutations and perturbations in RSK expression levels have a significant role in various cancers (Trivier, De Cesare et al. 1996, Romeo and Roux 2011). The crystal structures of CTK-ERK2 heterodimer complexes bound to a CTK D-motif peptide were solved and these structures not only elucidated the main interfaces but also demonstrated how ERK2 is positioned to phosphorylate T573 on the the CTK activation loop (Garai, Zeke et al. 2012, Alexa, Gógl et al. 2015). It is also the rare example in which a substrate kinase is stably bound to its activating kinase, providing a mechanism for how one kinase activates another.

The CTK-ERK2 heterodimer is held together primarily by two distinct interfaces (Figure 2.1A,B) (Garai, Zeke et al. 2012, Alexa, Gógl et al. 2015). The flexible CTK C-terminus revD motif (~712-730) binds with high affinity to the DRS

of ERK2 (Interface 1). L714 lies in a hydrophobic groove of the DRS, and Args 725,726 form electrostatic interactions with the acidic common docking (CD) surface of the DRS (D318, D321, E322). The CTK-ERK2 complex contains a second interface formed by an extended APE-αF helix, and the Glycine-rich Loop (Interface 2). This positions the CTK Activation Loop so that it faces the active site of ERK2, enabling efficient phosphorylation of the Activation Loop. Disruption of Interface 2 via mutation of the CTK APE motif does not cause a significant change in binding affinity, but the CTK becomes a less efficient substrate, while mutation of Interface 1 at either the electrostatic or hydrophobic components prevents DRS site binding. Here, we further analyze the CTK-ERK2 complex and its interfaces using a variety of computational techniques. We use free energy calculations and LSP-based protein residue networks to analyze the CTK-ERK2 complex and its interfaces. We also wish to identify high BC "non-hub bottleneck" or high DC/BC "bottleneck"/ "date-hub" (Estrada 2006, Yu, Kim et al. 2007, Pandini, Fornili et al. 2012) residues that are critical to the information flow of the complex, that when mutated will cause a change in the complex binding affinity and/or function. Moreover, we predict that mutation of these important high BC residues will redistribute the thermal vibrations of the complex and potentially lead to an allosteric change. Unsurprisingly, we found that Interface 1 makes the greatest contribution to free energy in the complex. The LSP-based network analysis of the complex enabled the detection of functionally key conserved residues as well as Interface residues. Finally, these networks revealed residues with large DC/BC and unknown function. For these we predict how the mutants will affect complex stability, and validate these experimentally. We focus on

two high DC/BC residues E593 and R588 of the extended αF helix, and a third modest DC/BC residue E33 in the ERK2 glycine rich loop. R588 lies at Interface 2, has high BC, and when mutated causes the CTK to be a more efficient substrate. E33 also lies at Interface 2, and when mutated, strengthens the affinity of the complex. Surprisingly, E593 has among the highest DC and BC scores in the entire complex, and despite not forming any contacts with ERK2, when mutated, weakens the affinity of the complex, suggesting this occurs in an indirect, allosteric fashion. These experimental results suggest while our LSP networks can identify novel important residues. Furthermore, our predictions agreed with experimental results, however, experimental validation is still required for most analysis.

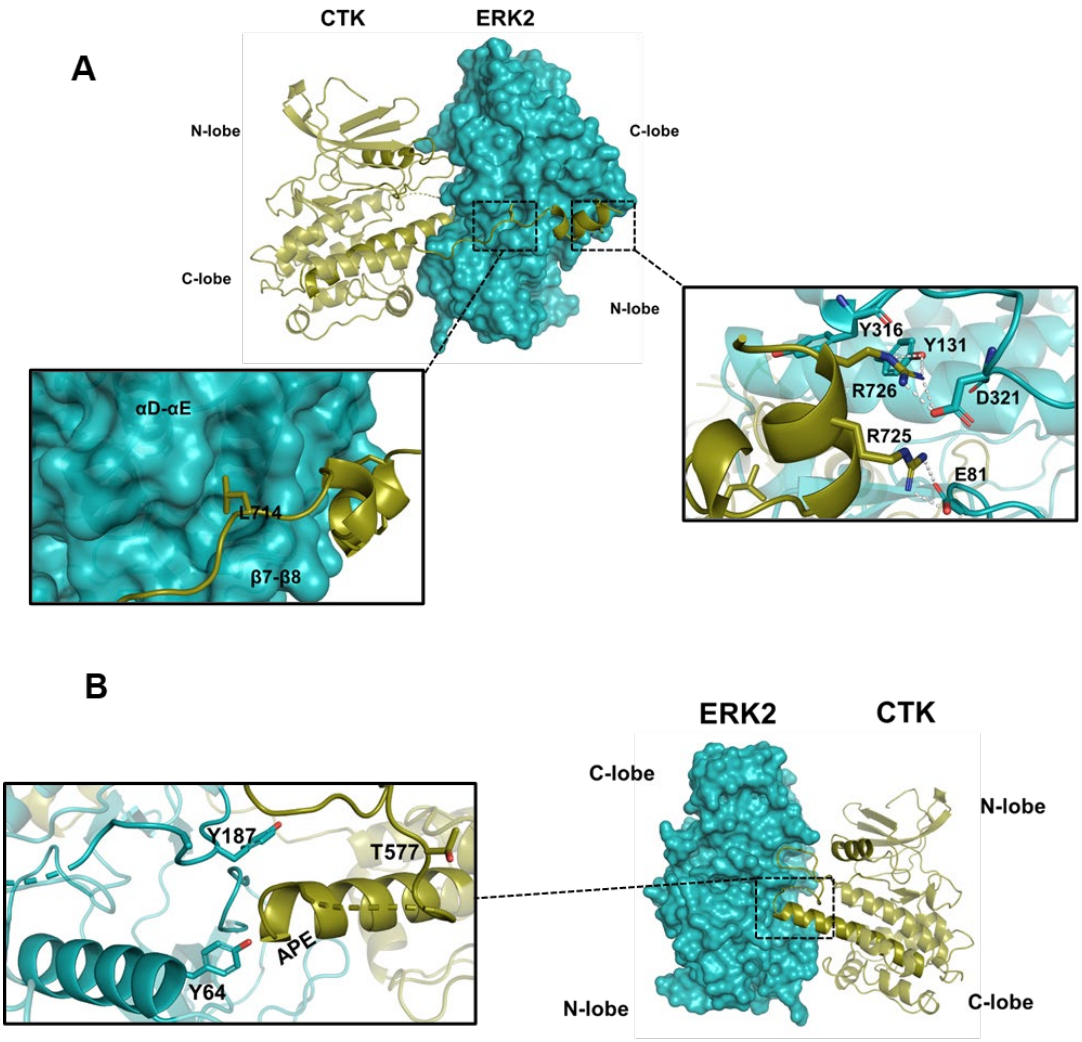


Figure 2.1. CTK-ERK2 crystal structure.
(A) displays Interface 1 of the CTK(green)-ERK2(teal) crystal structure (pdb: 4NIF). Inset shows detailed view of hydrophobic and electrostatic interactions that comprise CTK D motif and ERK2 DRS hydrophobic groove and CD surface. (B) displays Interface 2 of the CTK-ERK2 crystal structure. Inset shows the APE motif of the CTK forming a primarily hydrophobic interaction with the glycine rich loop of ERK2.

B. RESULTS

1. MD/MMGBSA

The CTK-ERK2 crystal structure was used as an initial model for a 1.6 us MD simulation, and we asessed how the complex and its interfaces evolved over time. For the initial model, we used CTK-ERK2 with Mg/ATP bound to ERK2. The two main interfaces in the complex, Interface 1 and Interface 2 remained quite stable over the time course of the simulation. The CTK activation loop is also quite flexible, however we did not observe the activation loop enter the active site of ERK2 throughout the simulation, as was seen in Alexa et al. (2015) (Alexa, Gógl et al. 2015). To quantify the extent in which each interface contributes to the stability of the complex on a per residue basis, we also performed MMGBSA (Genheden and Ryde 2015)to calculate the change in free energy upon complex formation. As expected, Interface 1 residues made the greatest contribution to the change in free energy in the complex (Figures 2.2, 2.S3C). The free energy in ERK2 is distributed throughout the protein over the DRS binding surface. The acidic CD surface has the greatest change in free energy. The free energy change in the CTK is concentrated mostly in the C-terminal tail among the revD motif consensus sequence. Interface 2 residue free energies are quite modest in comparison to Interface 1. Since genetic diseases tend to affect protein:protein interfaces as opposed to overall protein function and stability (Cheng, Zhao et al. 2021), we also mined cancer databases to assess the change in free energy among cancer mutations in the CTK-ERK2 complex (Dingerdissen, Torcivia-Rodriguez et al. 2018, Tate, Bamford et al. 2018). Overall, cancer mutation residues

make greater contributions to the free energy of the complex than non-cancer mutation residues (Figure 2.S3A).

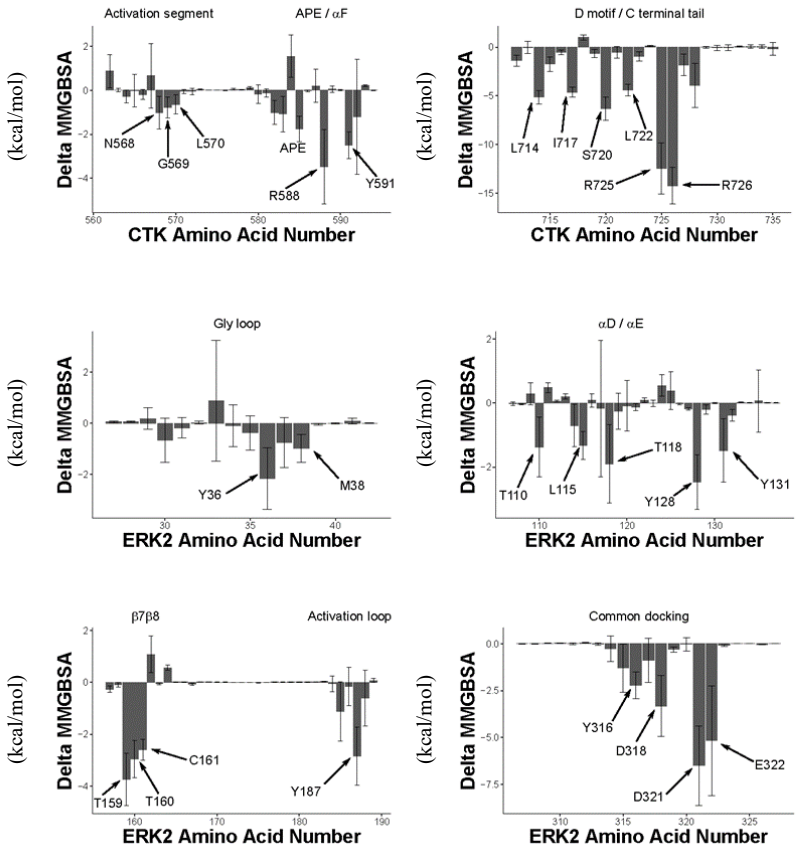


Figure 2.2. CTK-ERK2 MMGBSA.
Bar plots for the change in free energy per residue for complex formation, for various sections of the complex. Standard deviation of the calculation is represented by error bars for each residue. Residues with a favorable change in free energy are indicated by arrows and text labels. The activation loop, APE motif, through the extended aF helix have favorable change in free energy. CTK D motif has the greatest change in free energy among the complex at R725 and 726. Favorable change in free energy are observed throughout ERK2 at the Gly rich loop, and the DRS: aD – aE helices, b7b8 loop, and the acidic CD surface.

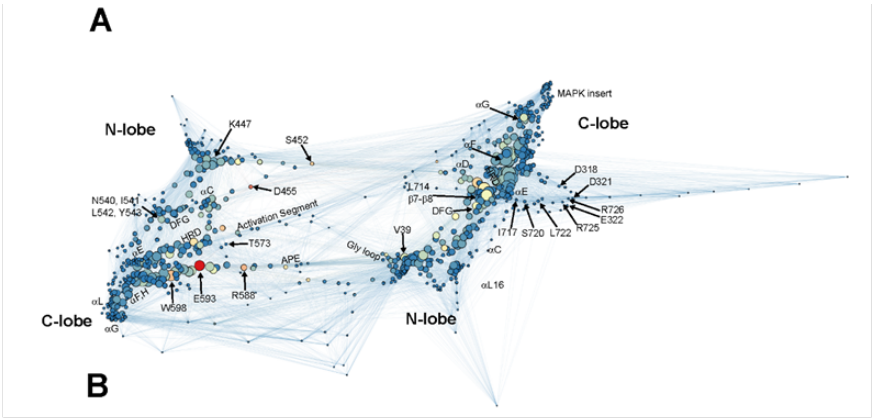
2. LSP network analysis

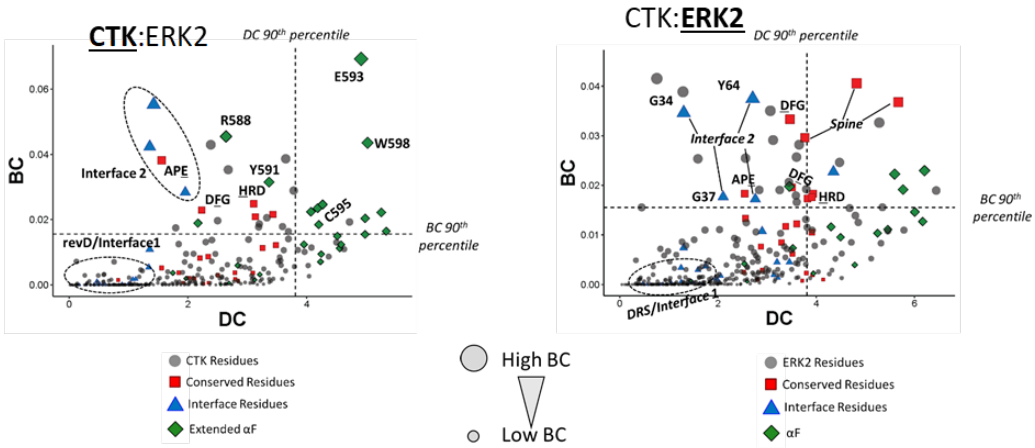
To further analyze the CTK-ERK2 complex and explore which residues play important roles in complex formation and stability, we took a network analysis approach, and constructed an LSP based network of the complex. The corresponding graph was laid out using Gephi's ForceAtlas2 algorithm (Bastian, Heymann et al. 2009, Jacomy, Venturini et al. 2014), in which highly interconnected nodes form dense clusters that in the LSP-based networks represent groups of residues that move as cohesive semi-rigid bodies. The WT CTK-ERK2 graph is organized such that the bilobal architecture for both the CTK and ERK2, is apparent (Figure 2.3A). In both kinases, N-lobe residues and C-lobe residues are clustered densely together on the tops and bottoms of each kinase, with regulatory and catalytic residues in the middle. The CTK activation segment, β3-αC loop (D450-D455), and the CTK APE motif occupy the middle of the graph between the two kinases. The revD motif of the CTK C-terminal tail passes through a dense cluster of ERK2 nodes, and the final ~ 10 nodes of the C-terminal tail protrude out as a long extension. L714 is positioned in the heart of a dense ERK2 cluster, partially comprised of the hydrophobic groove of the DRS (β7-β8). And basic residues R725 and R726 are on the periphery of this ERK2 cluster and next to the acidic common docking surface nodes.

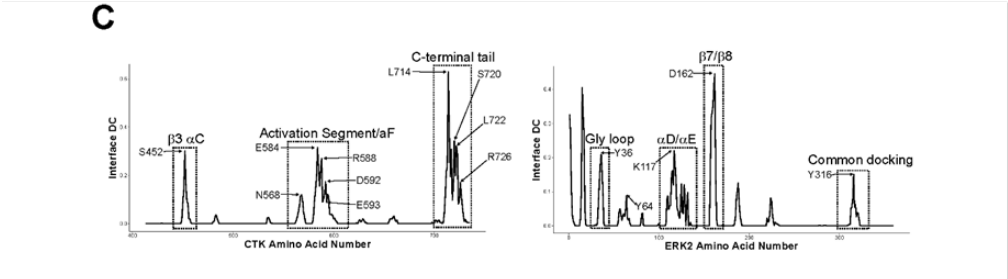
To evaluate the importance of individual amino acids in the CTK-ERK2 complex, we calculated degree centrality and betweenness centrality for each residue. Scatterplots of both DC and BC identify important nodes in the CTK-ERK2 complex (Figure 2.3B). Highly conserved, key residues necessary for kinase function and

regulation overall have significantly greater DC and BC than the rest of the complex residues (Figure 2.S3B). Interface residues have mixed results (Figure 2.3B). Interface 2 residues and $\beta 3\alpha C^{CTK}$:αD^{ERK2} have significantly higher BC than both Interface 1 and non-interface residues (Figure 2.S3C**)**. Interface 1 residues surprisingly have neither strong DC or BC. Since Interface 1 contains the unstructured revD that is removed from the CTK kinase core, and more a part of ERK2 nodes, the low DC/BC scores of this site may be due to a lack of intra-molecular edges or connections with the CTK. Therefore, we considered only intermolecular edges, those that connect CTK and ERK2 nodes, and calculated a corresponding interface DC score for each residue that provides a measure of interface stability. This approach clearly distinguishes Interfaces 1 and 2 of the CTK-ERK2 complex (Figure 2.3C). In particular, the hydrophobic groove, $L714^{CTK}$ and ERK2 β7-β8 loop, have the strongest, most stable intermolecular connections in the complex. This interface DC does resemble MMGBSA of the complex, but MMGBSA predicts that the electrostatic interactions of Interface 1 have the greatest impact on complex binding affinity, not the hydrophobic interaction (Figure 2.2). Both are critical for complex affinity, as mutation of both the electrostatic and hydrophobic components are drivers of different forms of cancer and completely abolish DRS-based interactions (Tanoue, Adachi et al. 2000, Alexa, Gógl et al. 2015).

Figure 2.3. Network analysis of the CTK-ERK2 complex.
(A) Gephi generated network with ForceAtlas2 layout. Nodes represent individual amino acids in the complex and edges connect nodes with correlated motions over MD simulation. Tightly clustered nodes behave as semi rigid bodies. The size of the node is directly proportional to degree centrality (DC). Node betweenness centrality (BC) is represented by the color of the node, over a blue to red gradient. Lower BC values have more blue colors, while hotter, more red colors have high BC. The majority of the CTK nodes are on the left of the graph, and most of the ERK2 nodes are on the right. For each kinase the bilobal architecture is visible, N and C- Lobes. The CTK β3αC loop, activation loop, and APE motif are found in the middle of the graph, between the CTK and ERK2. The CTK D motif passes through the C-lobe of ERK2 and extend out to the right of ERK2.
(B) Scatterplots of DC vs BC of the CTK-ERK2 complex. The left plot is the CTK scatterplot, the right plot is the ERK2 scatterplot. Point color and shape vary according to if a residue is a "key" conserved, Interface (Interface 1 or Interface 2) residue, or any other residue. Red triangles, blue squares, and grey points represent "key", Interface, or other residues, respectively.
(C) Line plot of Interface DC for CTK and ERK2 in the complex. Interface DC is the sum of only intermolecular edges per residue. Known interaction sites are indicated by dashed boxes, and interface DC peaks are indicated by arrows and labels. Hydrophobic components of Interface 1 have the greatest interface DC in the complex.


A
N-lobe
C-lobe
K447
S452
D455
T573
E593
R588
W598
MAPK insert
D318
D321
R726
E322
R725
I717
S720
L722
V39
αL16
B


CTK:ERK2
DC 90th percentile
BC 90th percentile
BC
DC
E593
R588
W598
Y591
C595
Interface 2
APE
HRD
DFG
revD/Interface1
CTK:ERK2
G34
Y64
DFG
Spine
Interface 2
G37
APE
HRD
DRS/Interface 1
CTK Residues
Conserved Residues
Interface Residues
Extended αF
High BC
Low BC
ERK2 Residues
Conserved Residues
Interface Residues
αF


C
C-terminal tail
β3 αC
Activation Segment/αF
β7/β8
Gly loop
αD/αE
Common docking
Interface DC
CTK Amino Acid Number
ERK2 Amino Acid Number

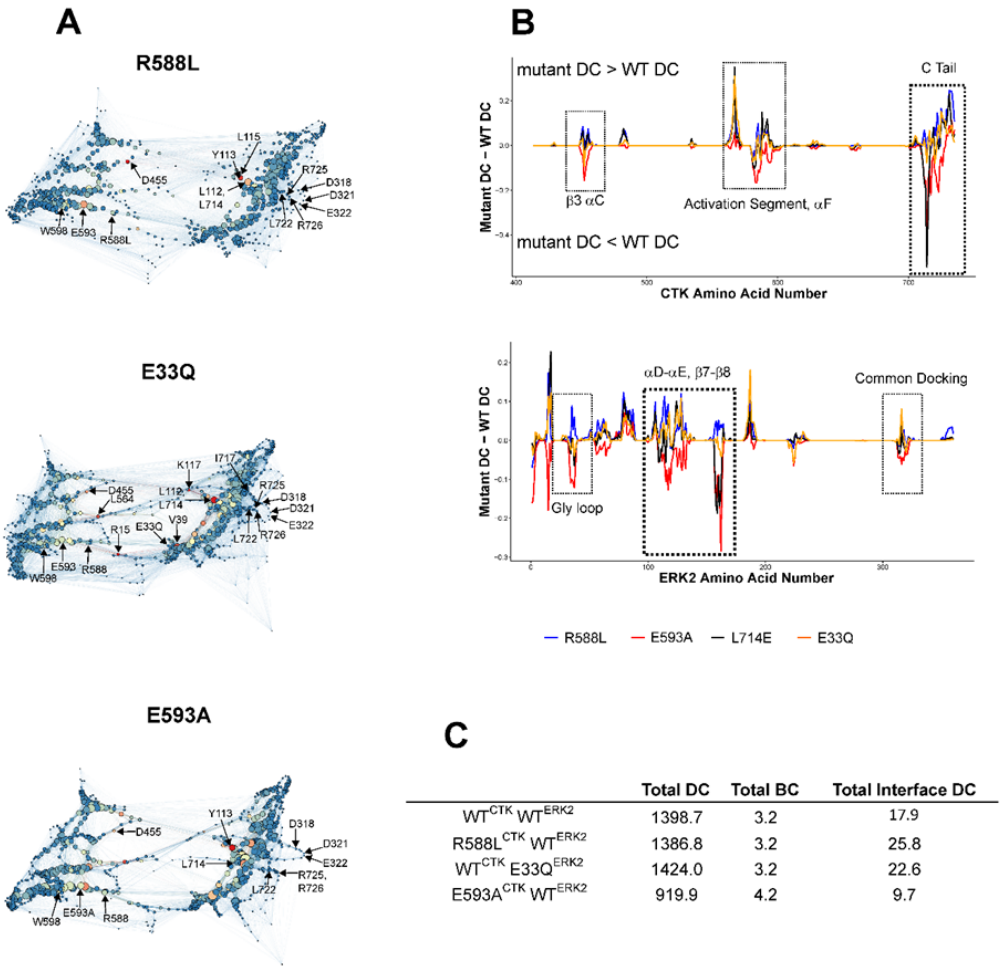
3. LSP network mutational analysis

There are other critical residues in the CTK-ERK2 network and corresponding DC/BC scatterplots whose functions are unknown. We identified a set of residues with strong DC/BC scores, placing an emphasis on residues from the CTK as it is not as well characterized as ERK2. Here we focus on two residues from the CTK extended αF helix, R588 and E593 which have two of the highest combined DC/BC scores (Figures 2.2A-B, 2.S4), and a third modest DC/BC residue, E33. R588 lies at Interface 2, has one of the strongest BC, moderate DC score, and energetically favors the CTK-ERK2 complex. Over the course of our simulation, R588 forms an electrostatic interaction with E33 of the ERK2 G-Loop. Both R588, and E33 also happen to be known cancer mutation sites (R588L, and E33Q). E33 has modest interface DC, slightly weaker than R588, effectively no change in free energy upon complex formation, and yet has weak DC/BC. E593 has the highest combined DC and BC score (Figure 2.3A,B), indicating that it is one of the most important residues in the CTK-ERK2 complex, serving as a major community forming hub that mediates network communication throughout the CTK. Intriguingly, E593 has essentially no change in free energy upon complex formation according to MMGBSA, and does not establish any direct contacts to any ERK2 residues (Figures 2.2, 2.S4). Therefore, we were intrigued by the potential of E593 mutation to allosterically modulate the CTK-ERK2 complex.

To predict the functional relevance of R588, E33 (that is the functional relevance of the R588-E33 electrostatic interaction), and E593 we generated *in silico*

mutations to the CTK-ERK2 complex: R588L^{CTK}, E33Q^{ERK2}, and E593A^{CTK}. MD simulations were performed for each mutant complex, and corresponding LSP-based networks were constructed (Figure 2.4A). Each mutant network organizes into roughly the same layout as WT, with N-lobe nodes and C-lobe nodes forming distinct clusters, and regulatory and catalytic nodes in the middle. The total number of edges or connections, and the total DC and total BC remained relatively unchanged for both R588L and E33Q complexes (Figure 2.4C). However, the network for E593A has significantly fewer connections and thus total lower DC than the other networks. Intriguingly, this reduction in connections is compensated by an increase in total BC, indicating that although E593A disrupts the connectivity of the CTK-ERK network, making a more sparse network, there are more high BC "bottlenecks" throughout the network.

We calculated the difference of interface DC between mutant and WT to predict the effect of the mutations on the stability of the CTK-ERK2 complex interfaces (Figure 2.4B). Both R588L and E33Q complexes causes an increase in total interface DC specifically at Interfaces 1 and 2. Therefore, we would expect that both R588L and E33Q enhance the stability of the complex relative to WT. For the E593A complex, the total interface DC decreased at all of the main interfaces relative to WT, suggesting a weakening of the CTK-ERK2 complex affinity.



	Total DC	Total BC	Total Interface DC
WT^{CTK} WT^{ERK2}	1398.7	3.2	17.9
$R588L^{CTK}$ WT^{ERK2}	1386.8	3.2	25.8
WT^{CTK} $E33Q^{ERK2}$	1424.0	3.2	22.6
$E593A^{CTK}$ WT^{ERK2}	919.9	4.2	9.7

Figure 2.4. Mutant CTK-ERK2 network analysis.
(A) Gephi generate networks for each mutant CTK-ERK2 complex, laid out using ForceAtlas2 algorithm. Node size is represented by DC, and color indicates BC on a blue to red gradient.
(B) Lineplot of the difference between each mutant interface DC and WT interface DC. Positive values indicate residues that have interface DC greater than the corresponding residue in the WT complex. Negative values indicate residues that have interface DC less than the corresponding residues interface DC in the WT complex. R588L, E33Q, E593A, and L714E are shown by blue, orange, red, and black lines, respectively. E593A has lower interface DC at all of the known interaction sites in the complex.
(C) Summary statistics of the Total DC, Total BC, and Total Interface DC for each mutant complex. R588L and E33Q complexes have similar total DC and total BC, but higher interface DC compared to WT. E593A has lower total and interface DC, but higher total BC compared to WT.

4. FEP predictions and validation

We utilized free energy perturbation (FEP) calculations to predict the relative change in free energy for CTK-ERK2 complex formation for each mutation (Figure 2.5A) (Jorgensen and Thomas 2008). Predictions of the complex binding affinity, based on network analysis, is a qualitative assessment, while FEP provides a more quantitative prediction of relative binding affinity. FEP accurately predicts that the negative control, L714E does not bind to ERK2 in agreement with previous experimental results. Network analysis based predictions and FEP are in agreement for R588L and E593A. Both predict strengthened binding for R588L by a factor of 4-5, and ~ 5 fold reduction in binding affinity for E593A. However, network analysis and FEP dramatically disagree with one another when predicting the effect of E33Q. While network analysis predicts enhanced stability of the complex, FEP predicts a dramatic reduction in binding affinity. (Figure 2.4).

To validate the predictive power of our network analysis and FEP calculations, we grew and purified WT and mutant CTK/ERK2, and performed SPR to measure the affinity of the CTK-ERK2 complexes (Figure 2.5B). Results in the literature have indicated that the phosphorylation state of the Activation Loops of MAPKs such as ERK2 contributes to high affinity binding of proteins via DRS based interactions (Lee, Hoofnagle et al. 2005, Piserchio, Warthaka et al. 2011, Tokunaga, Takeuchi et al. 2014). Supporting this, $Y187^{ERK2}$, the second of the phosphorylation sites of the ERK2 Activation Loop, is energetically favored in the complex (Figure 2.2). L714 and most of the hydrophobic groove are in the same cluster of closely connected residues as the ERK2 DFG aspartate (Figure 2.3A), whose conformation is dependent on the

phosphorylation state of the kinase. Therefore, we also tested the ability of these mutants to bind to the phosphorylated (2p) or non-phosphorylated (0p) forms of ERK2. Equilibrium based Kd measurements indicate that dual phosphorylation of the activation loop of ERK2 strengthens the binding affinity 4-5 fold. Despite FEP and network analysis both predicting an increase in binding affinity, R588L did not significantly change the affinity of the complex in either 0P or 2P forms of ERK2. E33Q enhances CTK affinity in all combinations, which is in clear disagreement with FEP predictions. Conversely, in agreement with both LSP network analysis and FEP calculations, $E593A^{CTK}$ bound to ERK2 0P, and 2P, with weaker affinity than WT, suggesting that E593A potentially allosterically weakens the binding affinity of the complex

As the CTK is a substrate of ERK2, we also monitored the ability of ERK2 to phosphorylate the CTK over time by mass spectrometry (Figure 2.5C), as we reasoned that these mutations might have an effect on the ability of ERK2 to efficiently phosphorylate the CTK. Phosphorylation by E33Q was not considered as any changes in the phosphorylation of the CTK could be due to impaired ATP binding/positioning as E33 is in the middle of the ERK2 Glycine-rich Loop, rather than altered CTK binding affinity. The negative control, L714E, is essentially unable to be efficiently phosphorylated by ERK2, presumably due to a lack of ERK2 binding. Despite, not observing a significant change in the binding affinity, R588L appears to be a better substrate compared to WT. Additionally, E593A appears to be a slightly weaker substrate as WT. Collectively these mutations appear to have an opposite effect,

R588L makes the CTK a better substrate, while E593A makes the substrate a slightly weaker substrate.

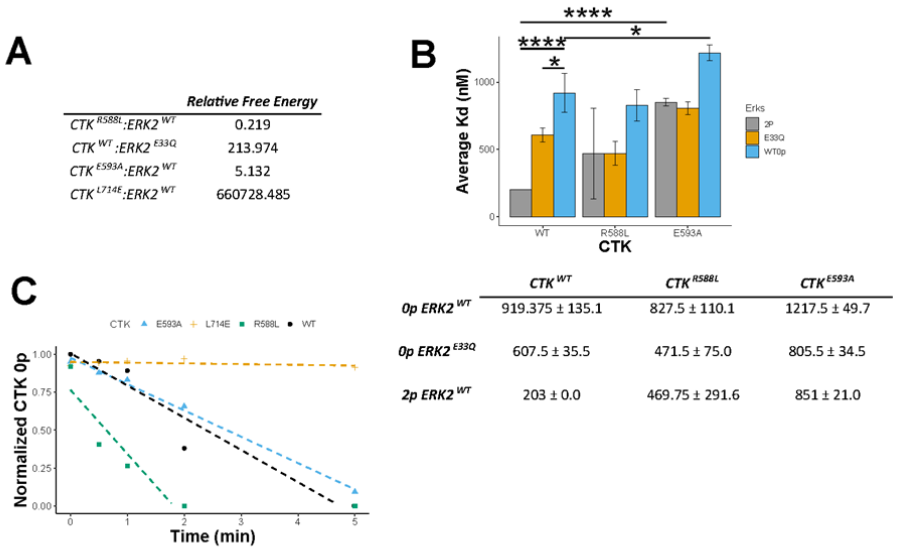


Figure 2.5. Free energy perturbation (FEP) prediction of mutant complexes and experimental validation

(A) FEP calculations of the relative free energies of each mutant complex relative to WT. R588L is expected to enhance complex stability, while E593A, E33Q, and L714E are all predicted to weaken or completely prevent binding of the complex (for L714E).

(B) Bar plot of the equilibrium based average Kd as measured by SPR for each CTK and ERK2 combination. CTK constructs tested were WT, R588L, and E593A. ERK2 constructs tested were WT 0p (inactive), E33Q 0p (inactive), and WT 2p (active). Bar plot error bars show standard deviation for each combination. Average Kd for each combination with standard deviation is summarized in the table below. Students t-test, ns: not significant, * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$, **** $p \leq 0.0001$

(C) Plot of the depletion of nonphosphorylated 0p CTK as measured by mass spectrometry. For multiple time points, the amount of 0p CTK relative to the total CTK, as determined by MS, is plotted. Over time the disappearance of the 0p signal corresponds to complete phosphorylation of the CTK activation loop.

C. Discussion

The analysis of protein:protein interfaces and the prediction of potential allosteric modulators at protein:protein interfaces remain of great interest to both academic and pharmaceutical institutions(Rask-Andersen, Almén et al. 2011, Ni, Liu et al. 2019, Ni, Lu et al. 2019, Shen, Yuan et al. 2019). We considered entropy driven allostery in which allosteric changes are driven by a redistribution of thermal vibrations of a protein or of protein complexes rather than large conformational changes (Kornev and Taylor 2015, Chopra, Wales et al. 2016, Ahuja, Kornev et al. 2017, Gourinchas, Etzl et al. 2017, Kornev 2018). We, therefore, search for residues that are major mediators of communication within the complex, that is residues with high DC/BC, where mutation of these residues could potentially redistribute internal thermal vibrations, leading to functional and potentially allosteric changes. Here we used free energy calculations and network analysis to analyze the CTK-ERK2 heterodimeric kinase complex and its interfaces. We verified that conserved key kinase residues have strong DC/BC scores and found that the main interfaces are the energetic drivers of complex formation and control stable intermolecular communication via high BC and high interface DC. And we identified a novel set of residues with strong DC/BC, whose functional importance would not be obvious just by inspection of the structure. We predicted and then experimentally validated the effect of mutating these important unknown sites. Mutation of R588 to the cancer mutation R588L (Dingerdissen, Torcivia-Rodriguez et al. 2018, Tate, Bamford et al. 2018) does not affect the complex binding affinity, however, it appears to make the CTK a more efficient ERK2 substrate.

R588 forms a salt bridge with E33 of ERK2, which surprisingly has quite modest DC/BC scores. Mutation of E33 to the cancer mutation glutamine (E33Q) (Dingerdissen, Torcivia-Rodriguez et al. 2018, Tate, Bamford et al. 2018)strengthened the affinity of the complex despite the relatively weak DC/BC. And mutation of E593 to alanine allosterically reduces the complex binding affinity.

In our analysis of the complex, we were particularly intrigued by E593, whose high combined DC and BC coupled with a lack of direct contacts with ERK2 encouraged us that our LSP network analysis identified a potential allosteric modulator. E593, when mutated to alanine, creates a more disconnected CTK-ERK2 complex with fewer overall connections, and weakens the CTK-ERK2 complex, causing a modest reduction in both binding affinity and ERK2 phosphorylation efficiency. The LSP-based network was also able to identify a different known allosteric site, C595. This MAPKAPK conserved Cys, when covalently linked to the inhibitor dimethyl fumarate (DMF), stabilizes the autoinhibitory αL helix, preventing activation of the CTK, which allosterically inhibits the downstream activity of RSK (Andersen, Gesser et al. 2018). Furthermore, the RSK2 homologous site is a Coffin Lowry Syndrome mutation (Trivier, De Cesare et al. 1996). Similar to E593, which has a similar DC/BC score, C595 does not contribute to complex formation (Figure 2.3B). Thus, our LSP network analysis is capable of identifying known allosteric sites, and this section of the extended αF helix may be an important allosteric site in the complex. Furthermore, we find the long αF helix particularly intriguing overall. This unusual long helix is observed in CTK structures of both RSK1/2 even in the absence of ERK2, and helps suppress basal activity by stabilizing the autoinhibitory αL helix via a hydrogen bond between S599

and Y702 (Malakhova, Tereshko et al. 2008, Li, Fu et al. 2012, Alexa, Gógl et al. 2015). Moreover, this αF helix blocks formation of the P+1 loop necessary for canonoical substrate binding. Mutation of this αL tyrosine to alanine, breaking this interaction, results in a constitutively active CTK. Thus, our network analysis highlights the importance of the extended αF helix in the CTK as it serves as a hydrophobic scaffold of the kinase, and contributes to CTK auto-inhibition and IF2 ERK2 binding.

How would we assess the ability of our LSP network analysis to identify novel, functionally important residues, and the reliability of our network/FEP predictions? Both were moderately successful but face limitations that must be taken into consideration. Residues with outstanding DC/BC scores, such as R588 and E593, when mutated had significant experimental consequences. However, our LSP network analysis also failed to identify the functionally important E33, as it was assigned low DC/BC values. As E33 is an interaction partner of R588 and is also part of Interface 2, it has a similar interface DC value; however, we were surprised that the BC of E33 was significantly less than of R588. We initially expected mutation of E33 would behave similarly to R588 given their interaction. But clearly these two mutations have different effects on Interface 2 and the complex. Furthermore, given the magnitude of the DC/BC score of E593, one might have expected a larger, more dramatic functional effect. We hypothesize that the magnitude of the change in binding affinity might be greater with a more severe mutation than alanine. Moreover, there may be some other function beside complex binding affinity or CTK activation loop phosphorylation, in which the effect of the E593A mutation is more dramatic. Lastly, the change in complex affinity caused by E593A is a biologically relevant magnitude of change. Following full RSK

activation after ERK2 phosphorylation of the CTK, RSK will auto-phosphorylate its own C-terminal tail, as a negative feedback step to release ERK2 from the complex (Gavin, Ni Ainle et al. 1999, Roux, Richards et al. 2003). Auto-phosphorylation reduces the affinity of the C-terminal tail for ERK2 by a factor of ~ 4-5 (Gógl, Biri-Kovács et al. 2018), which is a similar reduction in binding affinity for E593A binding to active 2p ERK2 relative to CKT WT.

Network analysis and FEP based predictions of the generated mutant complexes predicted the stability of the CTK-ERK2 complex reasonably well, with a few exceptions. With both techniques E593A and R588L were both predicted weaken and strengthen the interfaces of the complex, respectively. SPR measurements confirm that E593A weakens the affinity of the complex, however R588L does not affect complex affinity, but it is a more efficient substrate suggesting that this mutant complex may have enhanced interfaces that mediate efficient phosphorylation. Network based predictions of E33Q are in agreement with SPR, the mutation strengthens the complex interfaces and thus enhances the complex binding affinity. However, FEP predictions of E33Q are dramatic outliers, predicting that this mutation completely abolishes complex formation. This raises the question, why is the FEP prediction of E33Q so dramatically different than the experimentally determined SPR values? This is unclear at this time, but overall discrepancies in FEP and SPR may be due to large scale, slow timescale conformational changes that are not captured by microsecond scale MD simulations.

Here we have used a novel approach, LSP analysis of DC/BC, to analyze a protein kinase complex. We implemented this newly developed network analysis

approach to the CTK-ERK2 heterodimeric kinase complex. In this way we were able to not only identify key conserved residues and known interface residues, but we also identified previously unknown, yet important residues for the CTK-ERK2 complex. We predicted the effect of mutating these identified, important residues and validated them experimentally. Overall, our predictions based both on network analysis and FEP reasonably matched experimental values, with few exceptions. Therefore, these computational and network analysis predictions should still be experimentally validated.

D. Methods

1. Molecular dynamics simulations

The CTK-ERK2 heterodimer was prepared for all atom molecular dynamic simulations using the crystal structure of the CTK-ERK2 complex with AMP-PNP bound to ERK2 (PDB: 4nif). AMP-PNP was replaced with ATP and two Mg ions were modeled in following superposition of ERK2 with the crystal structure of PKA in complex with ATP/Mg and small protein kinase inhibitor, PKI (PDB: 1atp). Models were processed in Maestro (Schrodinger). Protein Preparation Wizard was used to build missing sidechains and model charge states of ionizable residues at neutral pH. Hydrogens and counter ions were added and the model solvated in a cubic box of TIP4P-EW (Horn, Swope et al. 2004) and 150 mM KCl with a 10Å buffer in AMBER tools (Song, Lee et al. 2019). Parameters from the Bryce AMBER Parameter Database were used for ATP (Meagher, Redman et al. 2003). AMBER16 was used for energy minimization, heating, and equilibration steps. Systems were minimized by 1000 steps

of hydrogen-only minimization, 2000 steps of solvent minimization, 2000 steps of ligand minimization, 2000 steps of side-chain minimization, and 5000 steps of all-atom minimization. Systems were heated from 0°K to 300°K linearly over 250 ps with 2 fs time-steps and 10.0 kcal-1 mol-1 Å position restraints on protein. Temperature was maintained by the Langevin thermostat. Constant pressure equilibration with a 10 Å non-bonded cut-off with particle mesh Ewald was performed with 300 ps of protein and ligand restraints followed by 300 ps of unrestrained equilibration. Hydrogen mass repartition was implemented to achieve a 4 fs time-step for production runs (Hopkins, Le Grand et al. 2015). Production simulations were performed on GPU enabled AMBER16 (Le Grand, Götz et al. 2013, Salomon-Ferrer, Götz et al. 2013) as above in triplicate for an aggregate of 1.2ms.

2. Network creation and centrality calculation

LSP alignment was performed using previously created software (Kornev, Haste et al. 2006, Kornev, Taylor et al. 2008) adapted for molecular dynamics simulation. Graphs were generated, as described earlier, using coordinates of C_α and C_β atoms for all residues with the exception of glycine (C_α and N) and ATP molecule (N_1,C_8). In brief: protein structure is represented by a graph with residues as nodes. Links are formed if the distance between C_α atoms of the corresponding residues are within a predefined cutoff level $\Delta C_{\alpha\alpha}$. Each link carries information about mutual orientation of the corresponding $C_\alpha C_\beta$ vectors: three distances ($C_{\alpha 1}$-$C_{\alpha 2}$, $C_{\alpha 1}$-$C_{\beta 2}$, $C_{\beta 1}$-$C_{\alpha 2}$) and the dihedral angle θ ($C_{\beta 1}$-$C_{\alpha 1}$-$C_{\alpha 2}$-$C_{\beta 2}$). Comparison of two protein structures is presented as a graph with residues as nodes and links created only if all the three

distances and the dihedral angles are similar, i.e., within predefined cutoff levels: $\Delta C_{\alpha 1\alpha 2}$<0.2Å, $\Delta C_{\alpha 1\beta 2}$<0.45Å, $\Delta\theta$<10°. Weights for the links are calculated using the following formula:

$$W = \frac{1}{4} * \left(\left(1 - \frac{\delta_{\alpha 1\alpha 2}}{\Delta C_{\alpha 1\alpha 2}}\right) + \left(1 - \frac{\delta_{\alpha 1\beta 2}}{\Delta C_{\alpha 1\beta 2}}\right) + \left(1 - \frac{\delta_{\beta 1\alpha 2}}{\Delta C_{\alpha 1\beta 2}}\right) + \left(1 - \frac{\delta\theta}{\Delta\theta}\right) \right)$$

where $\delta_{\alpha 1\alpha 2}$, $\delta_{\alpha 1\beta 2}$, $\delta_{\beta 1\alpha 2}$ and $\delta\theta$ are the corresponding differences between two $C_\alpha C_\beta$ vectors. Weights for the not matching links are assigned zero values.

For LSP-alignment of multiple structures all comparisons were made to a single reference structure and the detected weights were averaged. The reference structure was determined by finding a structure within the working set that has the smallest C_α RMSD to average C_α coordinates for the set.

LSP alignment binaries for Linux and MS Windows OS are available free of charge to academic non-profit institutions upon request. For commercial entities please contact UCSD Office of Innovation and Commercialization at innovation@ucsd.edu.

Normalized centralities were calculated using igraph R library (version 1.2.5) (Csárdi and Nepusz 2006). To calculate betweenness and closeness centralities weights were converted to distances using the following formula: $D = -\log W$. "Strength" function was used to calculate weighted degree centrality. Interface degree centrality was calculated by taking the weighted sum of intermolecular edges.

3. Protein expression and purification.

The CTK of RSK1, RPS6KA1 was a gift from Nicola Burgess-Brown (Addgene plasmid # 39021; http://n2t.net/addgene:39021 ; RRID:Addgene_39021), and was cloned into a pGEX vector via Gibson assembly to include an N-terminal GST tag. CTK was transformed into E.coli BL21 Rosetta pLysS, and an overnight starter culture was grown overnight. Overnight starter culture was inoculated into fresh LB media with Kanamycin and Chloramphenicol, cultures were grown at 37° until OD: 0.8, then the temperature was dropped to 18°, IPTG was added to 0.5mM. Expression continued overnight, and the next morning, pellets were collected and frozen. ERK2 was coexpressed with either lambda-phosphatase for producing non-phosphorylated ERK2 or with a constitutively active MKK1 construct (MKK1_4D; S218D, M219D, N221D,S22D) for active ERK2. ERK2 plasmids were transformed and grown in LB with Ampicillin and Chloramphenicol until OD: 0.6, then IPTG was added to 0.1mM and expression was continued at 18° overnight.

GST tagged CTK pellets were resuspended in lysis buffer consisting of 50mM Tris, 200mM NaCl, 5mM DTT, 5% glycerol, 5mM Bezamidine, 0.1mM AEBSF, 50 μM TPCL/TLCK. Resuspended pellets were lysed using a microfluidizer, and the soluble fraction of centrifuge clarified lysate was applied to equilibrated glutathione resin. Immobilized resin was washed, and protein was eluted by addition of lysis buffer supplemented with 10mM reduced glutathione. Purified protein was dialyzed into 20mM Tris pH 8, 150mM NaCl, 2mM DTT, 10% glycerol and frozen for storage.

His tagged ERK2 was purified as described previously (Alexa, Gógl et al. 2015). Briefly, ERK2 was lysed by microfluidizer, lysate clarified, and ERK2 was purified by Ni resin. To cleave the N-terminal His tag, tev protease was added to

purified ERK2 (1:100 w/w) , and the protein was dialyzed overnight to remove the imidazole. The cleaved protein was further purified by anion exchange (using hitrap Q column) both non-phosphorylated ERK2, and phosphorylated ERK2 were then polished by gel filtration using a S75 column. The phosphorylation state of active ERK2 was verified by mass spectrometry.

4. SPR and kinase reactions

We performed SPR to measure the binding affinity between CTK constructs and ERK2. An Anti GST Ab immobilized CM5 chip captured GST-CTK, and ERK2 was flowed over the chip at 30µl/min. Assays were performed using 10mM HEPES pH 7.5, 150mM NaCl, 0.005% Tween 20, 10% glycerol, 1mM TCEP, 1% BSA. For kinase reactions, GST-CTK constructs were incubated with active ERK2 and ATP/Mg2+ for 0.5, 1, 2, or 5 minutes. Reactions were quenched by addition of 50mM EDTA, and samples were immediately flash frozen. The phosphorylation state of ERK2 treated CTK was determined by intact time of flight mass spectrometry.

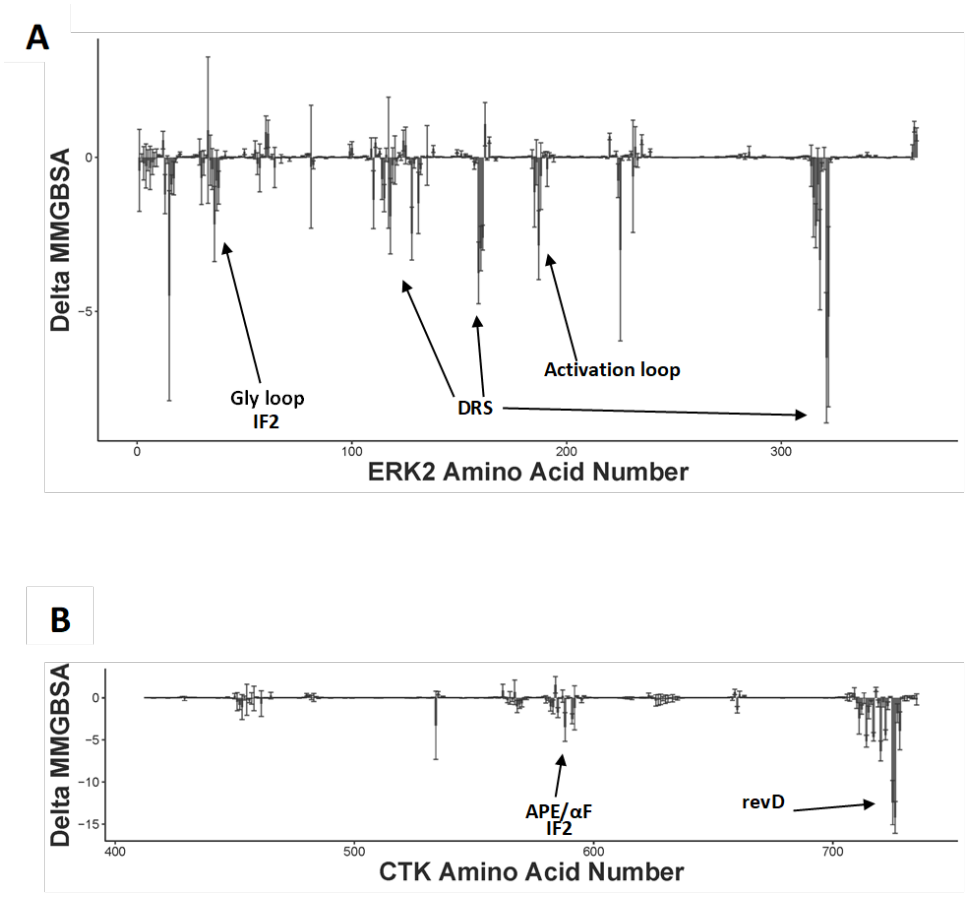


Figure 2.S1. Delta MMGBSA of CTK-ERK2 complex showing per residue change in free energy upon complex formation.
A) ERK2 and B) CTK. Arrows highlight large changes in free energy are observed in Interface 1 and Interface 2.

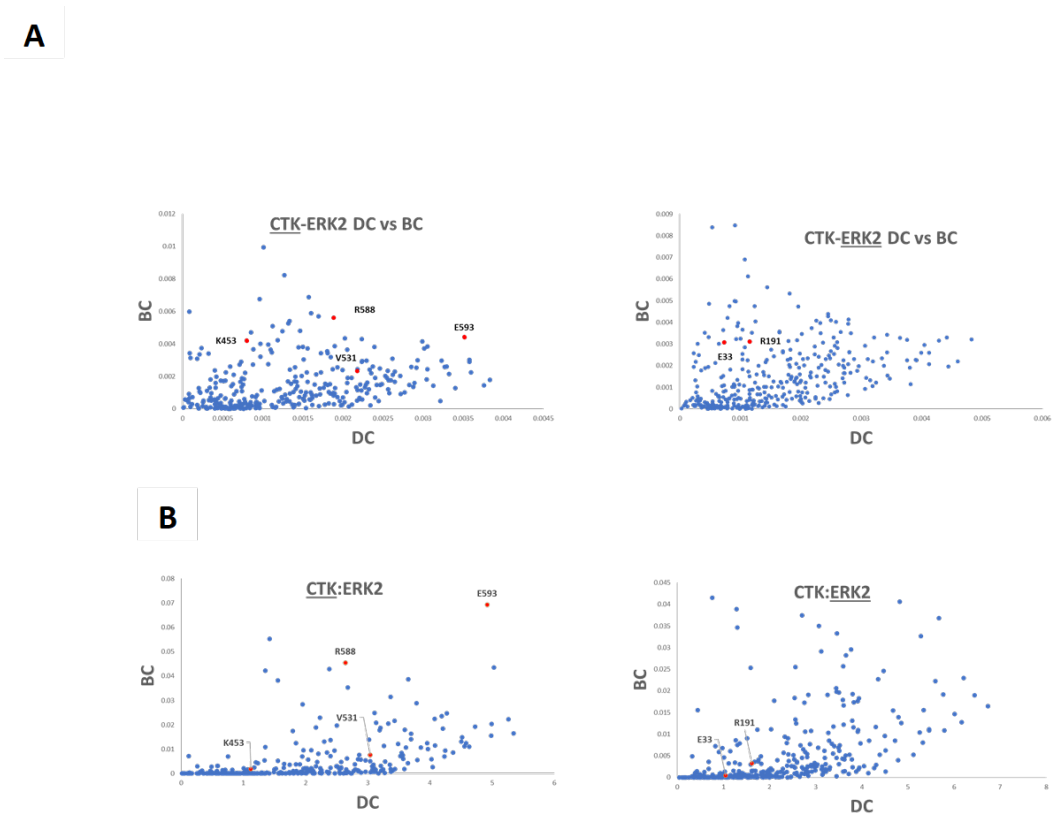
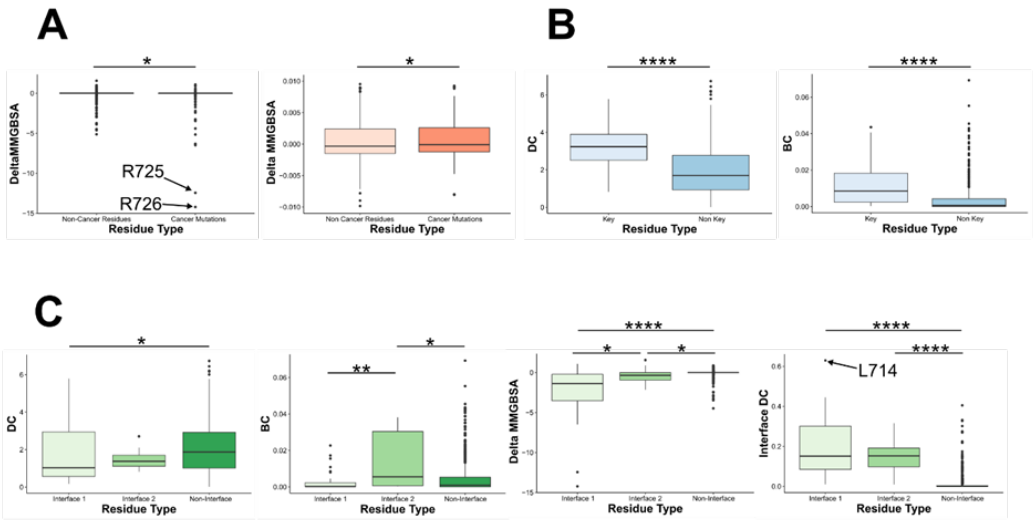


Figure 2.S2. Development of LSP network DC vs BC scatterplots.
A) DC/BC scatterplots of CTK (left) and ERK2 (right) in CTK-ERK2 complex during initial LSP iteration, using non-weighted graphs. DC and BC were calculated using non-weighted graphs. Selected residues for further analysis are labeled and colored red. B) DC/BC scatterplots of CTK (left) and ERK2 (right) in CTK-ERK2 complex using later development LSP networks. DC and BC were calculated using weighted graphs. Analyzed residues are labeled and colored red.

Figure 2.S3. Free energy and network centrality metrics for various residue types.

Non-parametric Wilcoxon rank sum tests were performed for all significance tests as every combination of samples contained non-normally distributed data. ns: not significant, * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$, **** $p \leq 0.0001$. (A) Left is a boxplot of Delta MMGBSA for cancer mutations compared to non-cancer mutations. The median Delta MMGBSA for cancer mutations is less than the median for non-cancer mutations. Right is a zoomed in view of the same boxplot to show the distributions of each sample.
(B) Boxplots of DC and BC for key or non-key residues in the complex. Key residues have significantly greater DC and BC than non-key residues.
(C) Boxplots of DC, BC, Delta MMGBSA, and interface DC for Interface 1, Interface 2, and non-interface residues. The non-interface median DC is significantly greater than Interface 1 median DC. Interface 2 has greater BC than both non-interface and Interface 1 residues. Interface 1 has the most favorable Delta MMGBSA compared to Interface 2 and non-interface residues. Both Interfaces 1 and 2 have greater interface DC than non-interface residues.
(D) Table summarizing significance results for other combinations of residues in the complex. The combined αC, αE, and αF helices from both the CTK and ERK2 have greater DC and BC (αF has greater interface DC as well). While the CTK β3αC loop and ERK2 αD helix are communication interfaces as they both have significantly greater BC and interface DC.



D

	DC	*BC*	*Delta MMGBSA*	*Interface DC*
CTK, ERK2 αF : Remaining Residues	****	****	ns	***
CTK, ERK2 αE : Remaining Residues	****	****	ns	ns
CTK, ERK2 αC : Remaining Residues	***	****	ns	ns
CTK β3αC : Remaining Residues	ns	**	*	****
ERK2 αD : Remaining Residues	ns	***	ns	****

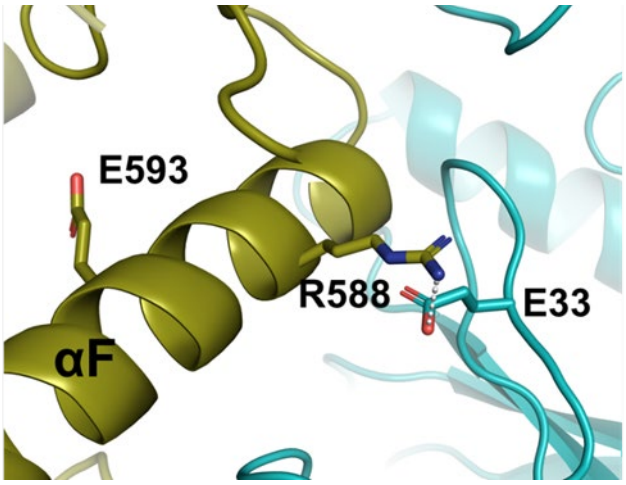


Figure 2.S4. $R588^{CTK}$, $E593^{CTK}$, and $E33^{ERK2}$ in CTK-ERK2 heterodimer complex.
R588 faces E33 in the crystal structure and forms electrostatic interaction. E593 faces away from the CTK-ERK2 interface.

Figure 2.S5. Gephi FA network with local edges for select nodes in the CTK-ERK2 network.

(A) CTK-ERK2 network is filtered to show only nodes that share an edge with L714. L714 is surrounded by ERK2 residues, making connections with aD, aE, b7-b8 loop residues, ERK2 C spine residue: L112, and b3aC loop residue S452.
(B) CTK-ERK2 network is filtered to show only nodes that share an edge with R588. R588 forms connections with the ERK2 glycine rich loop, CTK APE motif, the aF helix of the CTK, the CTK activation loop, H533 and R534 of the HRD motif.
(C) CTK-ERK2 network is filtered to show only nodes that share an edge with E33. E33 completes edges with the N-lobe of ERK2, as well as the DFG aspartate and catalytic asparagine. Connections are also made with the CTK activation segment, APE motif, and extended aF helix.
(D) CTK-ERK2 network is filtered to show only nodes that share an edge with E593. E593 is the center of a dense set of edges connecting to other CTK C-lobe (aF) including D596, W598, the HRD motif, and DFG pheylalanine. Connection with R15 and some other N-terminal tail residues of ERK2 are the extent of the edges to ERK2 nodes.

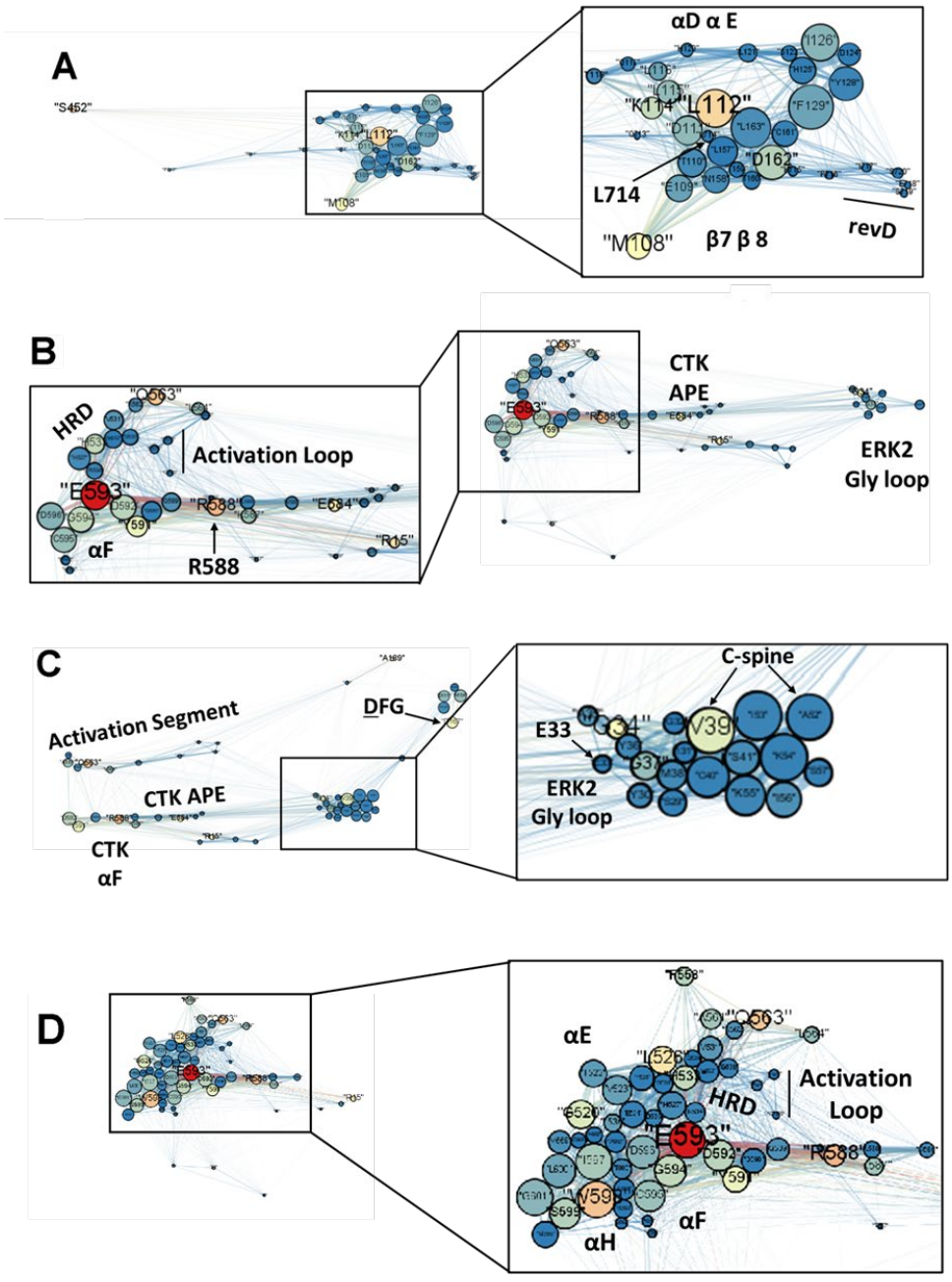
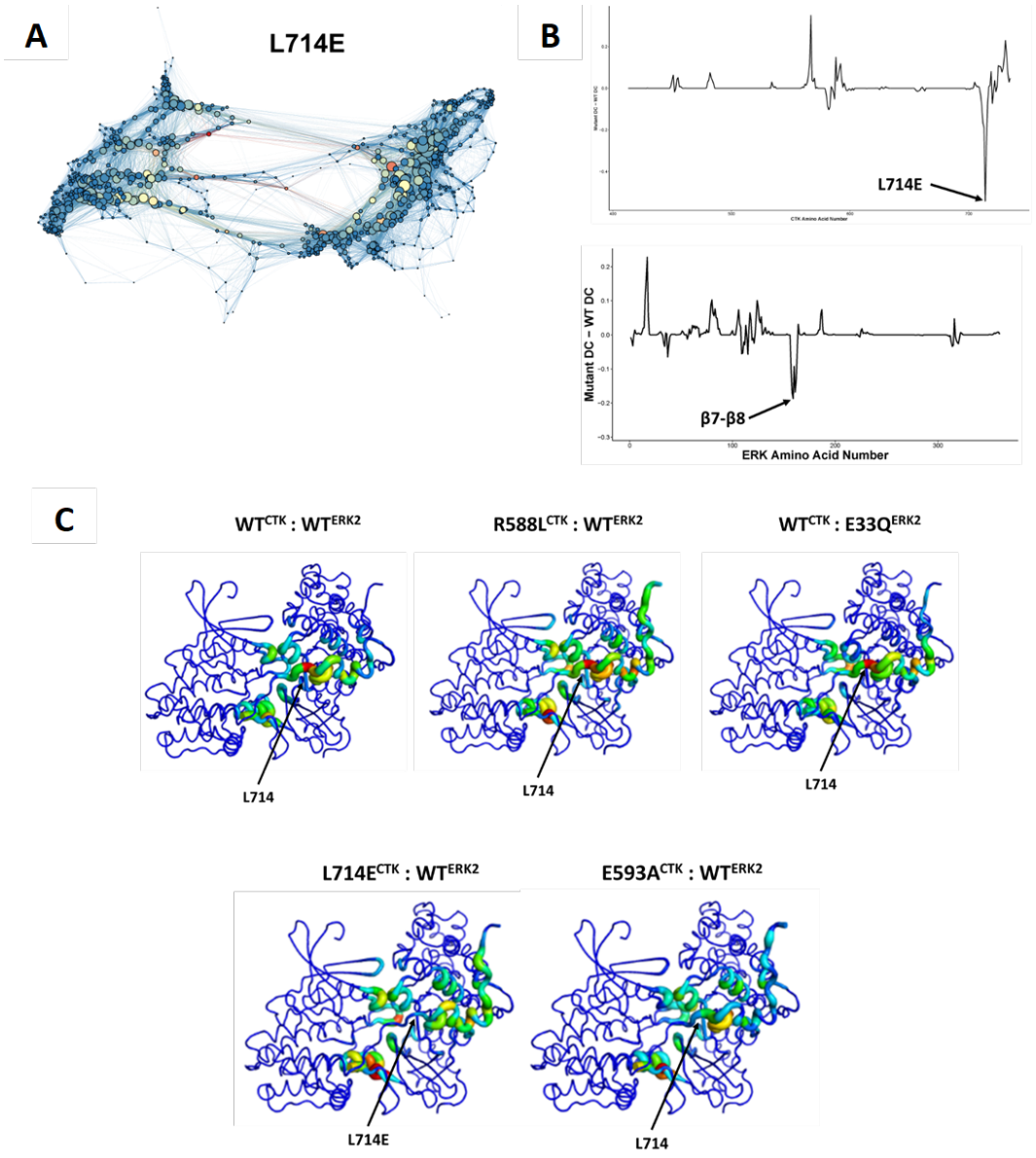
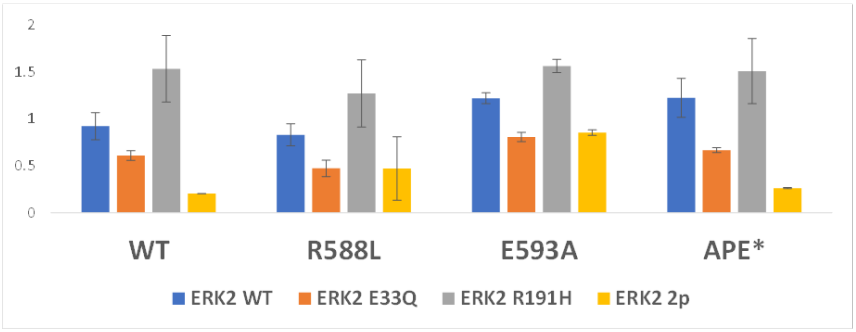

A
"S452"
αD α E
"K114" "L112"
"D162"
L714
"M108"
β7 β 8
revD
B
HRD
Activation Loop
"E593"
αF
R588
CTK
APE
ERK2
Gly loop
C
Activation Segment
DFG
CTK APE
CTK
αF
C-spine
E33
ERK2
Gly loop
D
αE
HRD
Activation
Loop
αF
αH

Figure 2.S6. LSP network of L714E and interface DC of CTK-ERK2 complex WT and mutants.

(A) Gephi generated network with ForceAtlas 2 layout for CTK^{L714E} – $ERK2^{WT}$ complex. Majority of CTK nodes are on the left and majority of ERK2 nodes are on the right. Overall layout of L714E complex is similar to WT network. (B) Lineplot for the difference in L714E mutant complex relative to WT. Positive values indicate that the mutation causes an increase in interface DC relative to WT. Negative values indicate that the mutation causes a decrease in interface DC relative to WT. CTK (top) and ERK2 (bottom). Arrows indicate that L714E and b7-b8 in ERK2, which comprise the major hydrophobic component of the DRS, displays a large decrease in Interface DC relative to WT. (C) Interface DC for the CTK-ERK2 complex is displayed as B-factors on the CTK-ERK2 structure (PDB:4NIF) for WT and all tested mutants. The arrow denotes the position for the Interface DC for L714 in each complex. Interface DC decreases for L714 in L714E and E593A complexes.


A
L714E
B
L714E
Mutant DC – WT DC
β7-β8
ERK Amino Acid Number
C
WT^CTK : WT^ERK2
R588L^CTK : WT^ERK2
WT^CTK : E33Q^ERK2
L714
L714
L714
L714E^CTK : WT^ERK2
E593A^CTK : WT^ERK2
L714E
L714

A



B

	$CTK^{WT}:ERK2^{WT}$
$CTK^{WT}:ERK2^{E33Q}$	**
$CTK^{WT}:ERK2^{R191H}$	ns
$CTK^{WT}:ERK2^{2p}$	****

	$CTK^{R588L}:ERK2^{WT}$
$CTK^{R588L}:ERK2^{E33Q}$	***
$CTK^{R588L}:ERK2^{R191H}$	ns
$CTK^{R588L}:ERK2^{2p}$	ns

	$CTK^{E593A}:ERK2^{WT}$
$CTK^{E593A}:ERK2^{E33Q}$	*
$CTK^{E593A}:ERK2^{R191H}$	*
$CTK^{E593A}:ERK2^{2p}$	***

	$CTK^{APE*}:ERK2^{WT}$
$CTK^{APE*}:ERK2^{E33Q}$	*
$CTK^{APE*}:ERK2^{R191H}$	ns
$CTK^{APE*}:ERK2^{2p}$	**

	$CTK^{WT}:ERK2^{WT}$
$CTK^{R588L}:ERK2^{WT}$	ns
$CTK^{E593A}:ERK2^{WT}$	*
$CTK^{APE*}:ERK2^{WT}$	*

	$CTK^{WT}:ERK2^{E33Q}$
$CTK^{R588L}:ERK2^{E33Q}$	ns
$CTK^{E593A}:ERK2^{E33Q}$	ns
$CTK^{APE*}:ERK2^{E33Q}$	ns

	$CTK^{WT}:ERK2^{R191H}$
$CTK^{R588L}:ERK2^{R191H}$	ns
$CTK^{E593A}:ERK2^{R191H}$	ns
$CTK^{APE*}:ERK2^{R191H}$	ns

	$CTK^{WT}:ERK2^{2p}$
$CTK^{R588L}:ERK2^{2p}$	ns
$CTK^{E593A}:ERK2^{2p}$	*
$CTK^{APE*}:ERK2^{2p}$	*

C

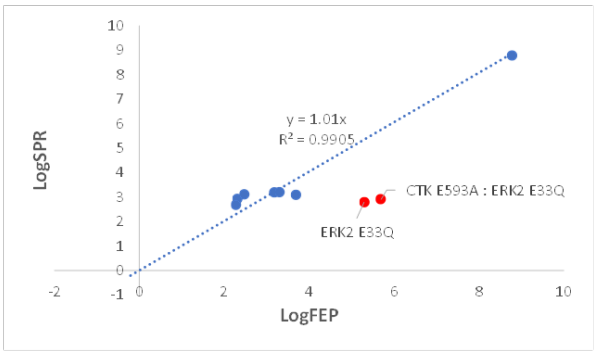


Figure 2.S7. Experimental validation and FEP prediction comparison of CTK-ERK complexes.

(A) SPR equilibrium measured K_D in uM for various WT and mutant combinations of CTK-ERK2 complex. (B) Tabulated significance levels using students T-test for various WT/mutant combinations of CTK-ERK2.

ns: not significant, * $p \leq 0.05$, ** $p \leq 0.01$, ***$p \leq 0.001$, **** $p \leq 0.001$. (C) Plot of logFEP prediction vs logSPR affinity for various CTK-ERK2 complex combinations. Linear regression fit to points excluding E33Q and E593A/E33Q.

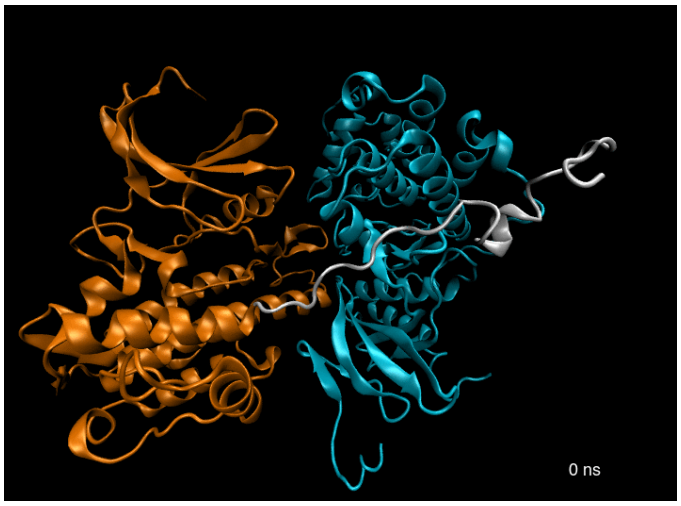


Fig 2.SM1. MD simulation of CTK-ERK2 WT heterodimer complex.
CTK is shown in orange and ERK2 is shown in cyan. The complex remains stable and together over the course of the simulation.

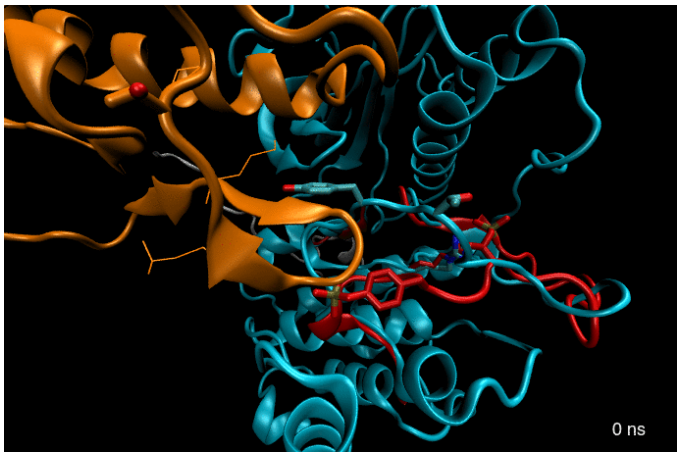


Fig 2.SM2. MD simulation of CTK-ERK2 WT active site.
Showing the activation loop of the CTK and active site of ERK2. The activation loop of active ERK2 is overlaid in red.

Table 2.S1. Summarized FEP affinity predictions for CTK-ERK2 WT and mutant complexes.
FEP free binding affinities shown as nM.

Complex	FEP (nM)
CTK^{K447A}:$ERK2^{WT}$	922.75
CTK^{K453A}:$ERK2^{WT}$	806.81
CTK^{V531F}:$ERK2^{WT}$	5466.96
CTK^{R588L}:$ERK2^{WT}$	200.34
CTK^{E593A}:$ERK2^{WT}$	4700.49
CTK^{L714E}:$ERK2^{WT}$	605227292.8
CTK^{WT}:$ERK2^{R191H}$	1501.32
CTK^{WT}:$ERK2^{E33Q}$	196000
CTK^{WT}:$ERK2^{H147Y}$	61.87
CTK^{WT}:$ERK2^{2pR191H}$	209.93
CTK^{WT}:$ERK2^{2pE33Q}$	15947.96
CTK^{WT}:$ERK2^{2pH147Y}$	5.32
CTK^{WT}:$ERK2^{R191H(Apo)}$	2701.45
CTK^{WT}:$ERK2^{E33Q(Apo)}$	3060551.15
CTK^{R588L}:$ERK2^{E33Q}$	195.56
CTK^{E593A}:$ERK2^{E33Q}$	496000
CTK^{R588L}:$ERK2^{R191H}$	307
CTK^{E593A}:$ERK2^{R191H}$	2084

Table 2.S2. Summarized SPR values for CTK-ERK2 WT and mutant complexes.

SPR values are shown in nM ± standard deviation. Table below shows number of observations for each WT/mutant combination.

	WT^{CTK}	$R588L^{CTK}$	$E593A^{CTK}$	$APE*^{CTK}$
$0pWT^{ERK2}$	919.4 ± 144.5	827.5 ± 117.7	1217.5 ± 57.4	1222.3 ± 206.9
$E33Q^{ERK2}$	607.5 ± 50.2	471.5 ± 86.6	805.5 ± 48.8	665.0 ± 25.5
$R191H^{ERK2}$	1530.0 ±353.6	1267.5 ± 357.3	1560.0 ± 70.7	1505.0 ± 346.5
$2pWT^{ERK2}$	203 ± 0.0	469.8 ± 336.7	851.0 ± 29.7	261.0 ± 5.7

	WT^{CTK}	$R588L^{CTK}$	$E593A^{CTK}$	$APE*^{CTK}$
$0pWT^{ERK2}$	8	8	4	4
$E33Q^{ERK2}$	2	4	2	2
$R191H^{ERK2}$	2	4	2	2
$2pWT^{ERK2}$	2	4	2	2

E. References

Ahuja, L. G., A. P. Kornev, C. L. McClendon, G. Veglia and S. S. Taylor (2017). "Mutation of a kinase allosteric node uncouples dynamics linked to phosphotransfer." Proc Natl Acad Sci U S A **114**(6): E931-e940.

Alexa, A., G. Gógl, G. Glatz, Á. Garai, A. Zeke, J. Varga, E. Dudás, N. Jeszenői, A. Bodor, C. Hetényi and A. Reményi (2015). "Structural assembly of the signaling competent ERK2-RSK1 heterodimeric protein kinase complex." Proc Natl Acad Sci U S A **112**(9): 2711-2716.

Amitai, G., A. Shemesh, E. Sitbon, M. Shklar, D. Netanely, I. Venger and S. Pietrokovski (2004). "Network analysis of protein structures identifies functional residues." J Mol Biol **344**(4): 1135-1146.

Andersen, J. L., B. Gesser, E. D. Funder, C. J. F. Nielsen, H. Gotfred-Rasmussen, M. K. Rasmussen, R. Toth, K. V. Gothelf, J. S. C. Arthur, L. Iversen and P. Nissen (2018). "Dimethyl fumarate is an allosteric covalent inhibitor of the p90 ribosomal S6 kinases." Nature Communications **9**(1): 4344.

Anjum, R. and J. Blenis (2008). "The RSK family of kinases: emerging roles in cellular signalling." Nature Reviews Molecular Cell Biology **9**(10): 747-758.

Bastian, M., S. Heymann and M. Jacomy (2009). "Gephi: An Open Source Software for Exploring and Manipulating Networks." Proceedings of the International AAAI Conference on Web and Social Media **3**(1): 361-362.

Blacklock, K. and G. M. Verkhivker (2014). "Computational Modeling of Allosteric Regulation in the Hsp90 Chaperones: A Statistical Ensemble Analysis of Protein Structure Networks and Allosteric Communications." PLOS Computational Biology **10**(6): e1003679.

Boccaletti, S., V. Latora, Y. Moreno, M. Chavez and D. U. Hwang (2006). "Complex networks: Structure and dynamics." Physics Reports **424**(4): 175-308.

Cheng, F., J. Zhao, Y. Wang, W. Lu, Z. Liu, Y. Zhou, W. R. Martin, R. Wang, J. Huang, T. Hao, H. Yue, J. Ma, Y. Hou, J. A. Castrillon, J. Fang, J. D. Lathia, R. A. Keri, F. C. Lightstone, E. M. Antman, R. Rabadan, D. E. Hill, C. Eng, M. Vidal and J. Loscalzo (2021). "Comprehensive characterization of protein–protein interactions perturbed by disease mutations." Nature Genetics **53**(3): 342-353.

Chopra, N., T. E. Wales, R. E. Joseph, S. E. Boyken, J. R. Engen, R. L. Jernigan and A. H. Andreotti (2016). "Dynamic Allostery Mediated by a Conserved Tryptophan in the Tec Family Kinases." PLOS Computational Biology **12**(3): e1004826.

Christley, R. M., G. L. Pinchbeck, R. G. Bowers, D. Clancy, N. P. French, R. Bennett and J. Turner (2005). "Infection in social networks: using network analysis to identify high-risk individuals." Am J Epidemiol **162**(10): 1024-1031.

Csárdi, G. and T. Nepusz (2006). The igraph software package for complex network research.

Dingerdissen, H. M., J. Torcivia-Rodriguez, Y. Hu, T. C. Chang, R. Mazumder and R. Kahsay (2018). "BioMuta and BioXpress: mutation and expression

knowledgebases for cancer biomarker discovery." Nucleic Acids Res **46**(D1): D1128-d1136.
Dokholyan, N. V., L. Li, F. Ding and E. I. Shakhnovich (2002). "Topological determinants of protein folding." Proceedings of the National Academy of Sciences **99**(13): 8637-8641.
Esmaielbeiki, R., K. Krawczyk, B. Knapp, J. C. Nebel and C. M. Deane (2016). "Progress and challenges in predicting protein interfaces." Brief Bioinform **17**(1): 117-131.
Estrada, E. (2006). "Spectral scaling and good expansion properties in complex networks." Europhysics Letters **73**(4): 649.
Garai, Á., A. Zeke, G. Gógl, I. Törő, F. Fördős, H. Blankenburg, T. Bárkai, J. Varga, A. Alexa, D. Emig, M. Albrecht and A. Reményi (2012). "Specificity of linear motifs that bind to a common mitogen-activated protein kinase docking groove." Sci Signal **5**(245): ra74.
Gautier, C., L. Laursen, P. Jemth and S. Gianni (2019). "Seeking allosteric networks in PDZ domains." Protein Engineering, Design and Selection **31**(10): 367-373.
Gavin, A. C., A. Ni Ainle, E. Chierici, M. Jones and A. R. Nebreda (1999). "A p90(rsk) mutant constitutively interacting with MAP kinase uncouples MAP kinase from p34(cdc2)/cyclin B activation in Xenopus oocytes." Mol Biol Cell **10**(9): 2971-2986.
Genheden, S. and U. Ryde (2015). "The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities." Expert Opin Drug Discov **10**(5): 449-461.
Gógl, G., B. Biri-Kovács, Á. L. Póti, H. Vadászi, B. Szeder, A. Bodor, G. Schlosser, A. Ács, L. Turiák, L. Buday, A. Alexa, L. Nyitray and A. Reményi (2018). "Dynamic control of RSK complexes by phosphoswitch-based regulation." The FEBS Journal **285**(1): 46-71.
Gourinchas, G., S. Etzl, C. Göbl, U. Vide, T. Madl and A. Winkler (2017). "Long-range allosteric signaling in red light–regulated diguanylyl cyclases." Science Advances **3**(3): e1602498.
Gunasekaran, K., B. Ma and R. Nussinov (2004). "Is allostery an intrinsic property of all dynamic proteins?" Proteins **57**(3): 433-443.
Hopkins, C. W., S. Le Grand, R. C. Walker and A. E. Roitberg (2015). "Long-Time-Step Molecular Dynamics through Hydrogen Mass Repartitioning." Journal of Chemical Theory and Computation **11**(4): 1864-1874.
Horn, H. W., W. C. Swope, J. W. Pitera, J. D. Madura, T. J. Dick, G. L. Hura and T. Head-Gordon (2004). "Development of an improved four-site water model for biomolecular simulations: TIP4P-Ew." J Chem Phys **120**(20): 9665-9678.
Jacomy, M., T. Venturini, S. Heymann and M. Bastian (2014). "ForceAtlas2, a Continuous Graph Layout Algorithm for Handy Network Visualization Designed for the Gephi Software." PLOS ONE **9**(6): e98679.
Jorgensen, W. L. and L. L. Thomas (2008). "Perspective on Free-Energy Perturbation Calculations for Chemical Equilibria." J Chem Theory Comput **4**(6): 869-876.
Kannan, N. and S. Vishveshwara (1999). "Identification of side-chain clusters in protein structures by a graph spectral method." J Mol Biol **292**(2): 441-464.

Kapetis, D., J. Sassone, Y. Yang, B. Galbardi, M. N. Xenakis, R. L. Westra, R. Szklarczyk, P. Lindsey, C. G. Faber, M. Gerrits, I. S. J. Merkies, S. D. Dib-Hajj, M. Mantegazza, S. G. Waxman, G. Lauria, M. Taiana, M. Marchi, R. Lombardi, D. Cazzato, F. M. Boneschi, A. Zauli, F. Clarelli, S. Santoro, I. Lopez, A. Quattrini, J. Hoeijmakers, M. Sopacua, B. de Greef, H. J. M. Smeets, R. Al Momani, J. M. Vanoevelen, I. Eijkenboom, S. Cestèle, O. Chever, R. Malik, M. Tavakoli, D. Ziegler and P. S. G. on behalf of the (2017). "Network topology of NaV1.7 mutations in sodium channel-related painful disorders." BMC Systems Biology **11**(1): 28.

Kornev, A. P. (2018). "Self-organization, entropy and allostery." Biochem Soc Trans **46**(3): 587-597.

Kornev, A. P., P. C. Aoto and S. S. Taylor (2022). "Calculation of centralities in protein kinase A." Proceedings of the National Academy of Sciences **119**(47): e2215420119.

Kornev, A. P., N. M. Haste, S. S. Taylor and L. F. Eyck (2006). "Surface comparison of active and inactive protein kinases identifies a conserved activation mechanism." Proc Natl Acad Sci U S A **103**(47): 17783-17788.

Kornev, A. P. and S. S. Taylor (2015). "Dynamics-Driven Allostery in Protein Kinases." Trends Biochem Sci **40**(11): 628-647.

Kornev, A. P., S. S. Taylor and L. F. Ten Eyck (2008). "A helix scaffold for the assembly of active protein kinases." Proc Natl Acad Sci U S A **105**(38): 14377-14382.

Lawrence, M. C., A. Jivan, C. Shao, L. Duan, D. Goad, E. Zaganjor, J. Osborne, K. McGlynn, S. Stippec, S. Earnest, W. Chen and M. H. Cobb (2008). "The roles of MAPKs in disease." Cell Research **18**(4): 436-442.

Le Grand, S., A. W. Götz and R. C. Walker (2013). "SPFP: Speed without compromise—A mixed precision model for GPU accelerated molecular dynamics simulations." Computer Physics Communications **184**: 374-380.

Lee, T., A. N. Hoofnagle, K. A. Resing and N. G. Ahn (2005). "Hydrogen exchange solvent protection by an ATP analogue reveals conformational changes in ERK2 upon activation." J Mol Biol **353**(3): 600-612.

Lee, Y., S. Choi and C. Hyeon (2014). "Mapping the intramolecular signal transduction of G-protein coupled receptors." Proteins **82**(5): 727-743.

Li, D., T. M. Fu, J. Nan, C. Liu, L. F. Li and X. D. Su (2012). "Structural basis for the autoinhibition of the C-terminal kinase domain of human RSK1." Acta Crystallogr D Biol Crystallogr **68**(Pt 6): 680-685.

Malakhova, M., V. Tereshko, S. Y. Lee, K. Yao, Y. Y. Cho, A. Bode and Z. Dong (2008). "Structural basis for activation of the autoinhibitory C-terminal kinase domain of p90 RSK2." Nat Struct Mol Biol **15**(1): 112-113.

Meagher, K. L., L. T. Redman and H. A. Carlson (2003). "Development of polyphosphate parameters for use with the AMBER force field." J Comput Chem **24**(9): 1016-1025.

Negre, C. F. A., U. N. Morzan, H. P. Hendrickson, R. Pal, G. P. Lisi, J. P. Loria, I. Rivalta, J. Ho and V. S. Batista (2018). "Eigenvector centrality for

characterization of protein allosteric pathways." Proc Natl Acad Sci U S A **115**(52): E12201-e12208.
Ni, D., N. Liu and C. Sheng (2019). "Allosteric Modulators of Protein-Protein Interactions (PPIs)." Adv Exp Med Biol **1163**: 313-334.
Ni, D., S. Lu and J. Zhang (2019). "Emerging roles of allosteric modulators in the regulation of protein-protein interactions (PPIs): A new paradigm for PPI drug discovery." Med Res Rev **39**(6): 2314-2342.
Pandini, A., A. Fornili, F. Fraternali and J. Kleinjung (2012). "Detection of allosteric signal transmission by information-theoretic analysis of protein dynamics." Faseb j **26**(2): 868-881.
Piserchio, A., M. Warthaka, A. K. Devkota, T. S. Kaoud, S. Lee, O. Abramczyk, P. Ren, K. N. Dalby and R. Ghose (2011). "Solution NMR Insights into Docking Interactions Involving Inactive ERK2." Biochemistry **50**(18): 3660-3672.
Rask-Andersen, M., M. S. Almén and H. B. Schiöth (2011). "Trends in the exploitation of novel drug targets." Nat Rev Drug Discov **10**(8): 579-590.
Reggiani, A., P. Nijkamp and D. Lanzi (2015). "Transport resilience and vulnerability: The role of connectivity." Transportation Research Part A: Policy and Practice **81**: 4-15.
Romeo, Y. and P. P. Roux (2011). "Paving the way for targeting RSK in cancer." Expert Opin Ther Targets **15**(1): 5-9.
Romeo, Y., X. Zhang and Philippe P. Roux (2011). "Regulation and function of the RSK family of protein kinases." Biochemical Journal **441**(2): 553-569.
Roux, P. P., S. A. Richards and J. Blenis (2003). "Phosphorylation of p90 ribosomal S6 kinase (RSK) regulates extracellular signal-regulated kinase docking and RSK activity." Mol Cell Biol **23**(14): 4796-4804.
Salomon-Ferrer, R., A. W. Götz, D. Poole, S. Le Grand and R. C. Walker (2013). "Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 2. Explicit Solvent Particle Mesh Ewald." Journal of Chemical Theory and Computation **9**(9): 3878-3888.
Sethi, A., J. Eargle, A. A. Black and Z. Luthey-Schulten (2009). "Dynamical networks in tRNA:protein complexes." Proceedings of the National Academy of Sciences **106**(16): 6620-6625.
Shen, L., Y. Yuan, Y. Guo, M. Li, C. Li and X. Pu (2019). "Probing the Druggablility on the Interface of the Protein–Protein Interaction and Its Allosteric Regulation Mechanism on the Drug Screening for the CXCR4 Homodimer." Frontiers in Pharmacology **10**.
Smith, J. A., C. E. Poteet-Smith, K. Malarkey and T. W. Sturgill (1999). "Identification of an extracellular signal-regulated kinase (ERK) docking site in ribosomal S6 kinase, a sequence critical for activation by ERK in vivo." J Biol Chem **274**(5): 2893-2898.
Song, L. F., T. S. Lee, C. Zhu, D. M. York and K. M. Merz, Jr. (2019). "Using AMBER18 for Relative Free Energy Calculations." J Chem Inf Model **59**(7): 3128-3135.

Stumpf, M. P. H., T. Thorne, E. de Silva, R. Stewart, H. J. An, M. Lappe and C. Wiuf (2008). "Estimating the size of the human interactome." Proceedings of the National Academy of Sciences **105**(19): 6959-6964.

Sutherland, C., D. G. Campbell and P. Cohen (1993). "Identification of insulin-stimulated protein kinase-1 as the rabbit equivalent of rskmo-2. Identification of two threonines phosphorylated during activation by mitogen-activated protein kinase." Eur J Biochem **212**(2): 581-588.

Tanoue, T., M. Adachi, T. Moriguchi and E. Nishida (2000). "A conserved docking motif in MAP kinases common to substrates, activators and regulators." Nat Cell Biol **2**(2): 110-116.

Tate, J. G., S. Bamford, H. C. Jubb, Z. Sondka, D. M. Beare, N. Bindal, H. Boutselakis, C. G. Cole, C. Creatore, E. Dawson, P. Fish, B. Harsha, C. Hathaway, S. C. Jupe, C. Y. Kok, K. Noble, L. Ponting, C. C. Ramshaw, C. E. Rye, H. E. Speedy, R. Stefancsik, S. L. Thompson, S. Wang, S. Ward, P. J. Campbell and S. A. Forbes (2018). "COSMIC: the Catalogue Of Somatic Mutations In Cancer." Nucleic Acids Research **47**(D1): D941-D947.

Tiberti, M., G. Invernizzi, M. Lambrughi, Y. Inbar, G. Schreiber and E. Papaleo (2014). "PyInteraph: a framework for the analysis of interaction networks in structural ensembles of proteins." J Chem Inf Model **54**(5): 1537-1551.

Tokunaga, Y., K. Takeuchi, H. Takahashi and I. Shimada (2014). "Allosteric enhancement of MAP kinase p38α's activity and substrate selectivity by docking interactions." Nature Structural & Molecular Biology **21**(8): 704-711.

Trivier, E., D. De Cesare, S. Jacquot, S. Pannetier, E. Zackai, I. Young, J. L. Mandel, P. Sassone-Corsi and A. Hanauer (1996). "Mutations in the kinase Rsk-2 associated with Coffin-Lowry syndrome." Nature **384**(6609): 567-570.

Vendruscolo, M., E. Paci, C. M. Dobson and M. Karplus (2001). "Three key residues form a critical contact network in a protein folding transition state." Nature **409**(6820): 641-645.

Wodak, S. J., E. Paci, N. V. Dokholyan, I. N. Berezovsky, A. Horovitz, J. Li, V. J. Hilser, I. Bahar, J. Karanicolas, G. Stock, P. Hamm, R. H. Stote, J. Eberhardt, Y. Chebaro, A. Dejaegere, M. Cecchini, J. P. Changeux, P. G. Bolhuis, J. Vreede, P. Faccioli, S. Orioli, R. Ravasio, L. Yan, C. Brito, M. Wyart, P. Gkeka, I. Rivalta, G. Palermo, J. A. McCammon, J. Panecka-Hofman, R. C. Wade, A. Di Pizio, M. Y. Niv, R. Nussinov, C. J. Tsai, H. Jang, D. Padhorny, D. Kozakov and T. McLeish (2019). "Allostery in Its Many Disguises: From Theory to Applications." Structure **27**(4): 566-578.

Yan, C., F. Wu, R. L. Jernigan, D. Dobbs and V. Honavar (2008). "Characterization of protein-protein interfaces." Protein J **27**(1): 59-70.

Yu, H., P. M. Kim, E. Sprecher, V. Trifonov and M. Gerstein (2007). "The Importance of Bottlenecks in Protein Networks: Correlation with Gene Essentiality and Expression Dynamics." PLOS Computational Biology **3**(4): e59.

Chapter III

Hydrogen deuterium exchange mass spectrometry of the RSK2:ERK2 complex

A. Introduction

Intrinsically disordered regions (IDRs) are embedded within and flank kinase domains, playing critical roles in kinase function and regulation (Gógl, Kornev et al. 2019). Kinases are flexible and dynamic enzymes that sample a variety of conformations and typically transition between regulated inactive and active states following phosphorylation by upstream kinases and/or *cis* auto-phosphorylation (Kornev and Taylor 2015). Kinases are bilobal, and the larger C-lobe is more stable, while the smaller N-lobe is more dynamic and flexible. Overall the N- and C-lobes should be considered as independent rigid bodies (Tsigelny, Greenberg et al. 1999). Phosphorylation of the activation loop, an internal IDR, establishes an interaction network that links the kinase active site with the N and C-lobes of the kinase domain, and this typically, stabilizes and enables the formation of an active kinase (Taylor, Keshwani et al. 2012, Gógl, Kornev et al. 2019). Highly conserved small linear motifs (SLiMs) (Davey, Cyert et al. 2015, Wright and Dyson 2015)within IDRs are promiscuous binding partners capable of a variety of intra and intermolecular interactions (Tompa, Davey et al. 2014, Gouw, Michael et al. 2018). Interactions between kinase domains and SLiMs also promote the assembly of substrate:activator kinase complexes. The combined action of kinase cores and IDRs/SLiMs is essential for function and assembly of an active enzyme.

The MAPK activated protein kinase (MAPKAPK) p90 RSK1-4 is an unusual protein that contains two distinct kinase domains in a single chain and is one of the rare kinases that forms a stable complex with its activating kinase ERK1/2 (Anjum

and Blenis 2008, Romeo, Zhang et al. 2012). RSK contains an N-terminal kinase (NTK) and C-terminal kinase (CTK), which are members of the AGC and CaMK families, respectively (Jones, Erikson et al. 1988, Fisher and Blenis 1996). Each kinase domain is flanked by IDRs embedded with SLiMs critical for RSK activation and function. Following the CTK, a C-terminal tail contains an autoinhibitory αL helix that suppresses CTK basal activity (Malakhova, Tereshko et al. 2008), and a revD motif that mediates high affinity binding to ERK (Roux, Richards et al. 2003, Garai, Zeke et al. 2012). The Linker that connects the NTK and CTK contains the conserved AGC C-terminal extension, which is a hallmark feature of AGC kinases (Kannan, Haste et al. 2007), and includes the conserved Ade, Turn, and Hydrophobic motifs (HM). The Ade motif completes the ATP binding site, the phosphorylatable Turn motif enhances kinase thermal stability, and the HM fills a so-called PDK1-interacting fragment (PIF) pocket made up of the αC helix and the β4 strand of the N-lobe, anchoring the AGC extension to the N-lobe (Dalby, Morrice et al. 1998, Biondi, Cheung et al. 2000, Pearce, Komander et al. 2010). This site is often phosphorylated or contains a negatively charged phosphomimetic residue that forms additional interactions with the N-lobe of the kinase that aid in the stabilization of the αC in an active conformation. PDK1 lacks its own C-terminal extension and is thought to utilize the Ade/HM motifs of their substrate kinases to facilitate substrate activation loop phosphorylation (Mora, Komander et al. 2004, Romano, Kannan et al. 2009, Pearce, Komander et al. 2010). Canonical RSK activation requires the input of four different kinases, and the Linker/C-terminal tail. First ERK binds the revD motif and phosphorylates the activation loop of the CTK, enabling CTK phosphorylation of the HM in the

Linker/AGC extension (Anjum and Blenis 2008, Romeo, Zhang et al. 2012). The phosphorylated HM (pHM) then binds the PIF pocket in PDK1 (Mora, Komander et al. 2004) to facilitate phosphorylation of the NTK activation loop, allowing phosphorylation of downstream substrates (Biondi, Cheung et al. 2000). The multistep activation mechanism involves numerous binding events between IDRs/SLiMs and ordered kinase domains from different kinase families (Manning, Whyte et al. 2002). Thus, analysis of RSK and its activation provides an opportunity to better understand the interplay of ordered kinase domains and IDRs (Gógl, Kornev et al. 2019).

The crystal structure of the CTK-ERK2 heterodimer reveals a mechanism for ERK2 binding and phosphorylation of the CTK activation loop. There are no structures involving FL RSK, and structures of the isolated NTK and CTK (Ikuta, Kornienko et al. 2007, Malakhova, Tereshko et al. 2008, Malakhova, Kurinov et al. 2009, Utepbergenov, Derewenda et al. 2012, Derewenda, Artamonov et al. 2013, Aronchik, Appleton et al. 2014, Costales, Mathur et al. 2014, Jain, Mathur et al. 2015) lack the AGC extension and remaining Linker. Thus, the precise molecular details that define the remaining steps in canonical RSK activation are not well understood. Following ERK2 phosphorylation, how does the CTK engage and phosphorylate the HM? Furthermore, how then does PDK1 bind the Linker/NTK, particularly the pHM, to phosphorylate the NTK activation loop?

Here we perform hydrogen deuterium exchange mass spectrometry (HDXMS) of the RSK2-ERK2 complex to obtain information on the solvent accessibility of all three kinase domains and IDRs. We find that the intrinsically disordered regions (IDRs) of RSK that flank the kinase domains: N-terminal tail, Linker, C-terminal tail

have the greatest solvent accessibility in the complex. Among the kinase domains, the N-lobe of the NTK displays the greatest deuterium uptake indicating that this kinase domain is capable of sampling a greater conformational space than ERK2 or the CTK. By analyzing solved NTK crystal structures we identified major conformational states of the NTK that varied from completely inactive, to 'active-like conformations'. We also investigated the mechanism whereby the strongly, solvent accessible RSK2 Linker mediates RSK activation following ERK2 phosphorylation of the CTK. Peptide arrays identified CTK and PDK1 binding sites in the RSK2 Linker that enable HM and NTK activation loop phosphorylation, respectively. Because the HM resembles the CTK αL helix and a canonical CaMK substrate, we predict that the HM engages the active site of the CTK similarly as CaMK:substrate complexes. Thus, a conserved P-3 Arg and P+1 Phe in the HM may form electrostatic interactions with acidic residues in the CTK αD helix and hydrophobic interactions with the P+1 CTK loop, respectively. PDK1 binds to the NTK AGC extension at both the Ade motif and HM. Surprisingly, FL RSK2 forms a stable complex with PDK1 only when the HM is not phosphorylated. As HM phosphorylation is required for PDK1 enabled NTK phosphorylation (Frödin, Jensen et al. 2000), we may have captured RSK2 and PDK1 in an intermediate conformation that is still incapable of NTK phosphorylation.

B. RESULTS

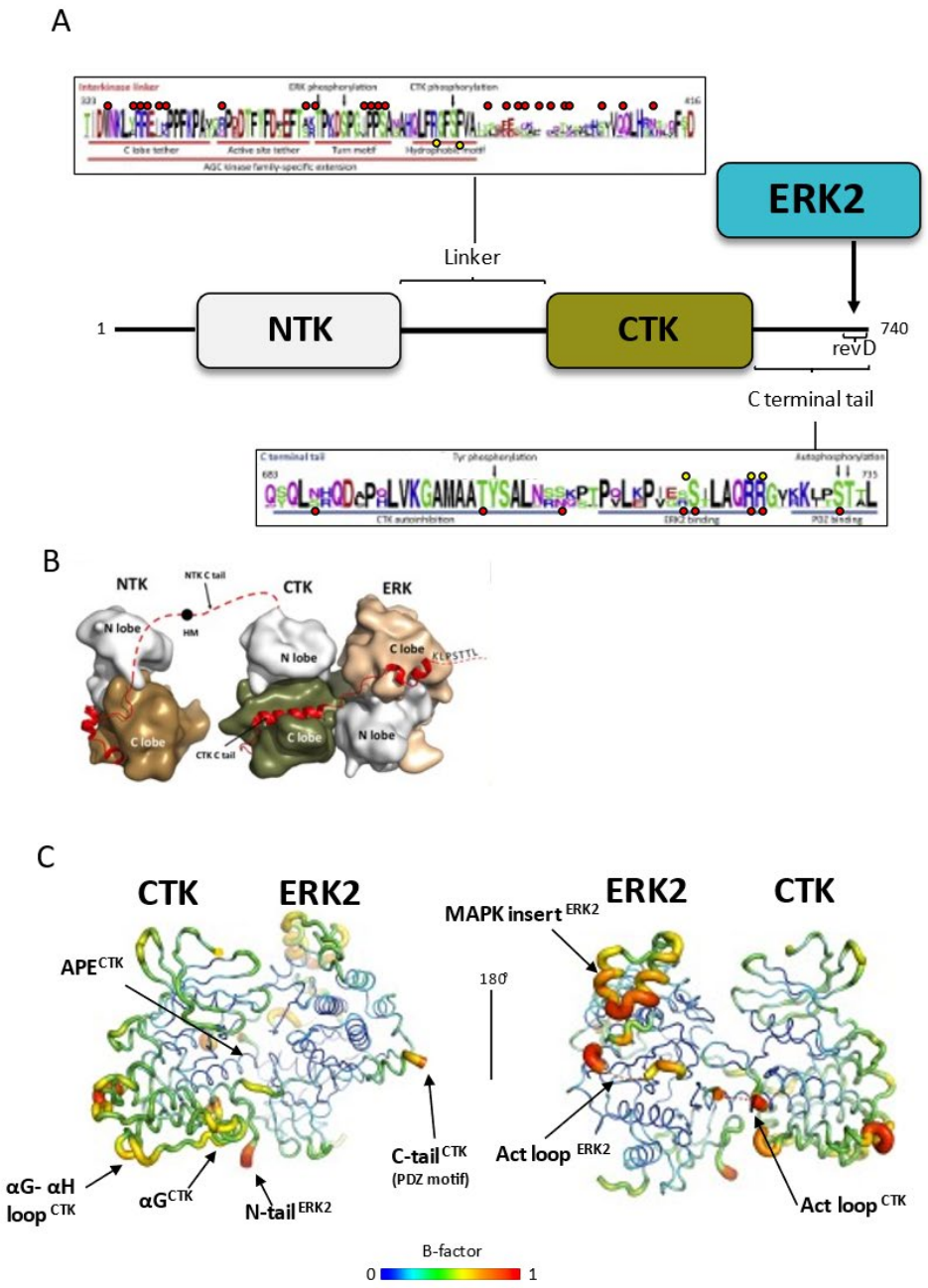
1. RSK2:ERK2 HDXMS

We performed HDXMS on the FL RSK2-ERK2 signaling complex (Fig. 3.1A-B) to obtain solvent accessibility/dynamics information over a 5-minute time course to identify dynamic or unstructured regions. Here we analyzed non-phosphorylated/inactive RSK2 and ERK2 in the absence of ATP/Mg, a precatalytic apo complex. We achieved excellent sequence coverage for the entire complex, 99% and 98% for RSK2 and ERK2, respectively, with multiple overlapping peptides (Fig. 3.2). The unstructured, partially intrinsically disordered tails flanking all three kinase domains have the greatest solvent exposure (Fig. 3.2, 3.3). Thus, the RSK2 N-tail, Linker, and C-tail are the most flexible or unstructured regions in the RSK2-ERK2 complex. HDXMS was recently performed on RSK4, in the absence of ERK2, and appears to have a similar pattern of solvent exposure; significant solvent exposure at the N and C termini, as well as the linker connecting the NTK and CTK (Chrysostomou, Roy et al. 2021). Inspection of B-factors, and MD simulation of the CTK-ERK2 heterodimeric complex indicates that the tails of both CTK and ERK2 are dynamic (Fig 3.1C) (Alexa, Gógl et al. 2015). MD simulation of a FL RSK2-ERK2 model further verifies the RSK2 N-tail, Linker, and C-terminal tail are flexible and disordered. The CTK revD motif binding to ERK is stable throughout the MD simulation, suggesting that the strong solvent exposure of this region may be due to flexibility in the rest of the C-tail, particularly the PDZ domain binding motif at the tails C-terminus. Additionally, the APE/extended αF that establishes a second

interface with ERK2, properly positioning the CTK for efficient ERK2 phosphorylation, has significant solvent exposure as well. Although the B factors of this region are quite low, MD simulations show that there is some flexibility around the main, stable APE^{CTK} – $Gly\ loop^{ERK2}$ pivot point (Alexa, Gógl et al. 2015), suggesting that the surrounding extended αF helix is more flexible than the crystal structure would indicate.

The solvent exposure/protection patterns of all three kinases of the RSK2-ERK2 apo complex are similar to HDXMS analysis of other kinases like PKA and ERK2 (Yang, Garrod et al. 2005, Steichen, Iyer et al. 2010). For all three kinases, the core hydrophobic helices of the C-lobes (αE,F,H) and the catalytic loops are well shielded from solvent, but the αG helix and its connected loops to the αF and αH helices are exposed. Additionally, in ERK2, the MAPK insert is solvent exposed and dynamic (Fig. 3.1C). Both the αG and MAPK insert contribute to protein:protein or protein:substrate binding interactions. (Taylor and Kornev 2011, Roskoski 2012, Taylor, Keshwani et al. 2012). Unsurprisingly, solvent exposure is observed in the activation segments and glycine rich loops of all three kinases, indicating that they are flexible and dynamic. Both are known to sample multiple conformations in the absence of phosphorylation or bound nucleotide (Madhusudan, Akamine et al. 2002, Steichen, Kuchinskas et al. 2012, Bastidas, Wu et al. 2015)..

Figure 3.1. Domain organization of the RSK2-ERK2 complex.
(A) Domain organization of RSK2 and schematic of the RSK2-ERK2 complex. Arrows denote binding events, ERK2 binds to the revD motif in the C-terminal tail. Sequence logos of the RSK Linker and C-terminal tail, with conserved motifs and phosphorylation sites denoted. The sequence logos are adapted from Gogl, et al., TIBS, 2019. Red dots indicate the positions of cancer mutations (COSMIC database) and yellow dots indicate the positions of Coffin Lowry Syndrome mutations.
(B) Model of the FL RSK2-ERK2 signaling complex. The model is adapted from Gogl, et al., TIBS, 2019, and is built using the CTK-ERK2 heterodimer crystal structure and an isolated NTK crystal structure. Dashed red lines indicate potential locations of flexible regions, for which structural information is not available.
(C) The crystal structure of the CTK-ERK2 heterodimer with the color and thickness of the cartoon representative of the B-factors of the complex.


A
ERK2
Linker
1
NTK
CTK
740
revD
C terminal tail
B
NTK
NTK C tail
CTK
ERK
N lobe
HM
C lobe
N lobe
C lobe
CTK C tail
C lobe
N lobe
C
CTK
ERK2
APE CTK
αG- αH loop CTK
αG CTK
N-tail ERK2
C-tail CTK (PDZ motif)
MAPK insert ERK2
ERK2
CTK
180°
Act loop ERK2
Act loop CTK
B-factor
0
1

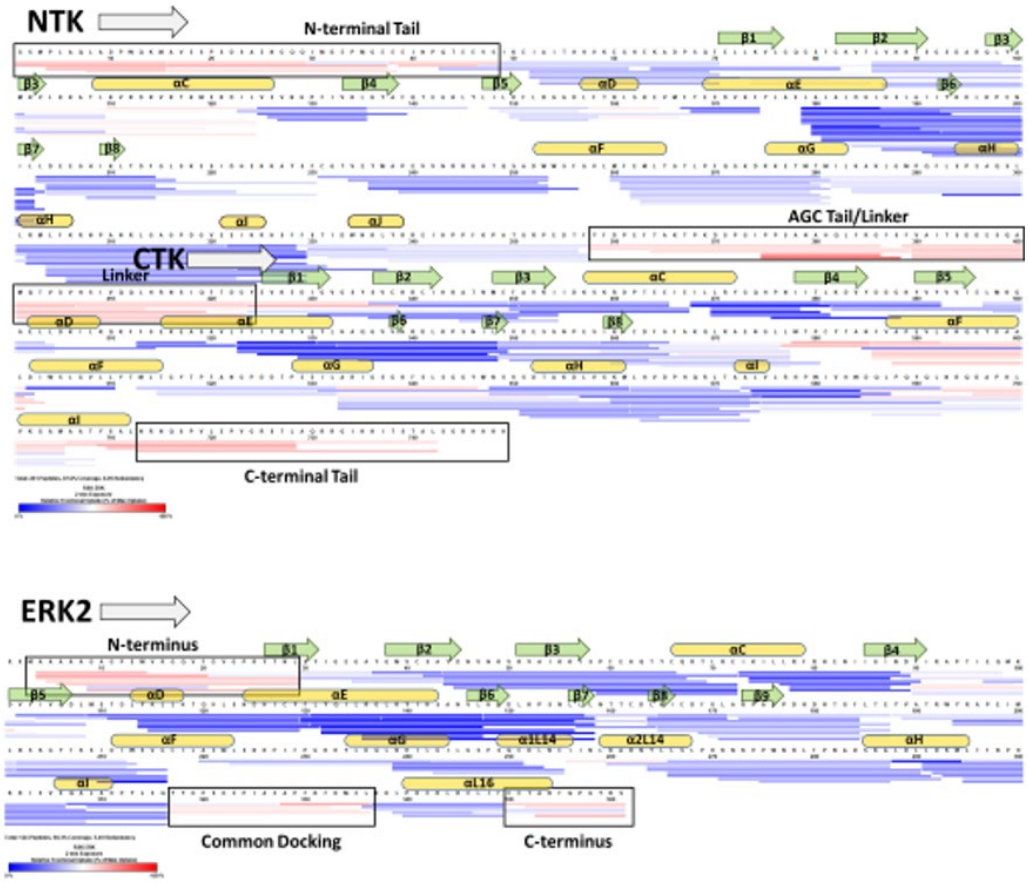


Figure 3.2. Coverage maps of RSK2 and ERK2 in the RSK2-ERK2 complex.

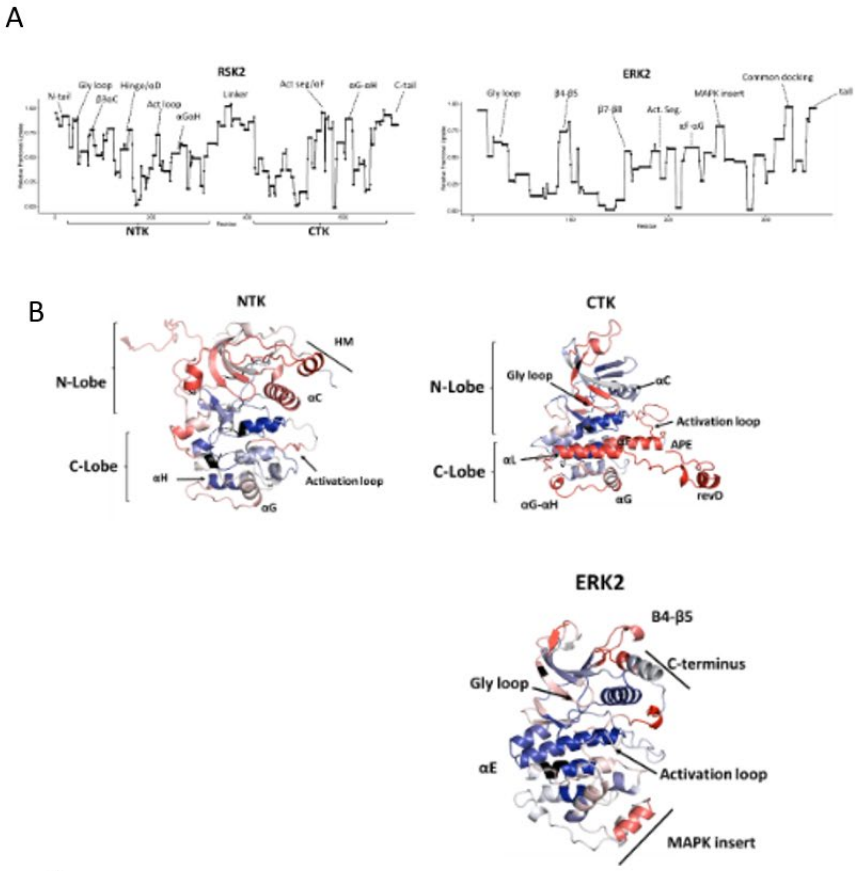
Top and bottom, are the coverage maps of RSK2 and ERK2 in the RSK2-ERK2 complex, respectively, at the 2-minute timepoint. Blue sticks representing solvent shielded peptides and red sticks representing solvent exposed peptides. Predicted secondary structure is represented above the sequence, yellow cylinder and green arrow cartoons represent α helices and β sheets, respectively.

Figure 3.3. Deuterium uptake of RSK2, ERK2 in the RSK2-ERK2 complex.

(A) Relative fractional uptake plot for RSK2 and ERK2, regions with relatively strong solvent exposure are denoted in the plots.

(B) Relative fractional uptake is represented as B-factors for all three kinase domains in the RSK2-ERK2 complex. Red represents solvent exposure and blue represents solvent protection. B-factors are applied to the crystal structure of ERK2 (pdb: 2erk) and to an AlphaFold model of RSK2 (45-740).

(C) B-factors of representative NTK structures. Cooler colors represent low b-factors, while warmer colors indicate greater b-factors and therefore greater flexibility within the crystal structure


A
RSK2
NTK
CTK
ERK2
B
NTK
N-Lobe
C-Lobe
HM
αC
αH
αG
Activation loop
CTK
Gly loop
αL
APE
revD
αG-αH
ERK2
B4-β5
C-terminus
αE
MAPK insert
C
NTK+SL0101
NTK+LJH685
NTK+AMP -PCP
N-lobe
Activation loop
αG

2. The conformational space of the NTK

The N-lobe of the NTK is considerably more solvent exposed than the corresponding N-lobes of the CTK and ERK2 (Fig 3.2, 3.3A). In addition to the Gly loop and ATP binding hinge region, the β3-αC loop, αC, β5, and αD helix are all solvent exposed, suggesting that the N-lobe is more dynamic and conformationally malleable. To explore the conformational dynamics of the NTK, we compared all of the publicly available crystal structures of the NTK bound to various nucleotides and inhibitors (Ikuta, Kornienko et al. 2007, Malakhova, Tereshko et al. 2008, Utepbergenov, Derewenda et al. 2012, Aronchik, Appleton et al. 2014, Costales, Mathur et al. 2014, Jain, Mathur et al. 2015). The C-lobes of all of these structures are well aligned regardless of isoform or inhibitor, while the activation segments and N-lobes are poorly aligned and dynamic (Fig. 3.4A). These structures all lack the AGC extension, and we identified that the NTK N-lobes adopt one of two major conformational states, Conformations A and B, which can be distinguished by the secondary structure and position of their αC helices (Fig. 3.4A, bottom). The αC helices are known to be mobile, capable of moving between 'in' and 'out' conformations depending on the activation state of the kinase (Johnson, Akamine et al. 2001, Taylor, Shaw et al. 2015). Conformation A structures feature a helical, intact αC helix in more of a canonical, 'in' position. RSK1 NTK structures are considered as Conformation A despite an unfolded or disordered αC (Ikuta, Kornienko et al. 2007), since no other residues occupy this position. Some structures contain rather extreme secondary structure perturbations to the αC helix that result in the formation an

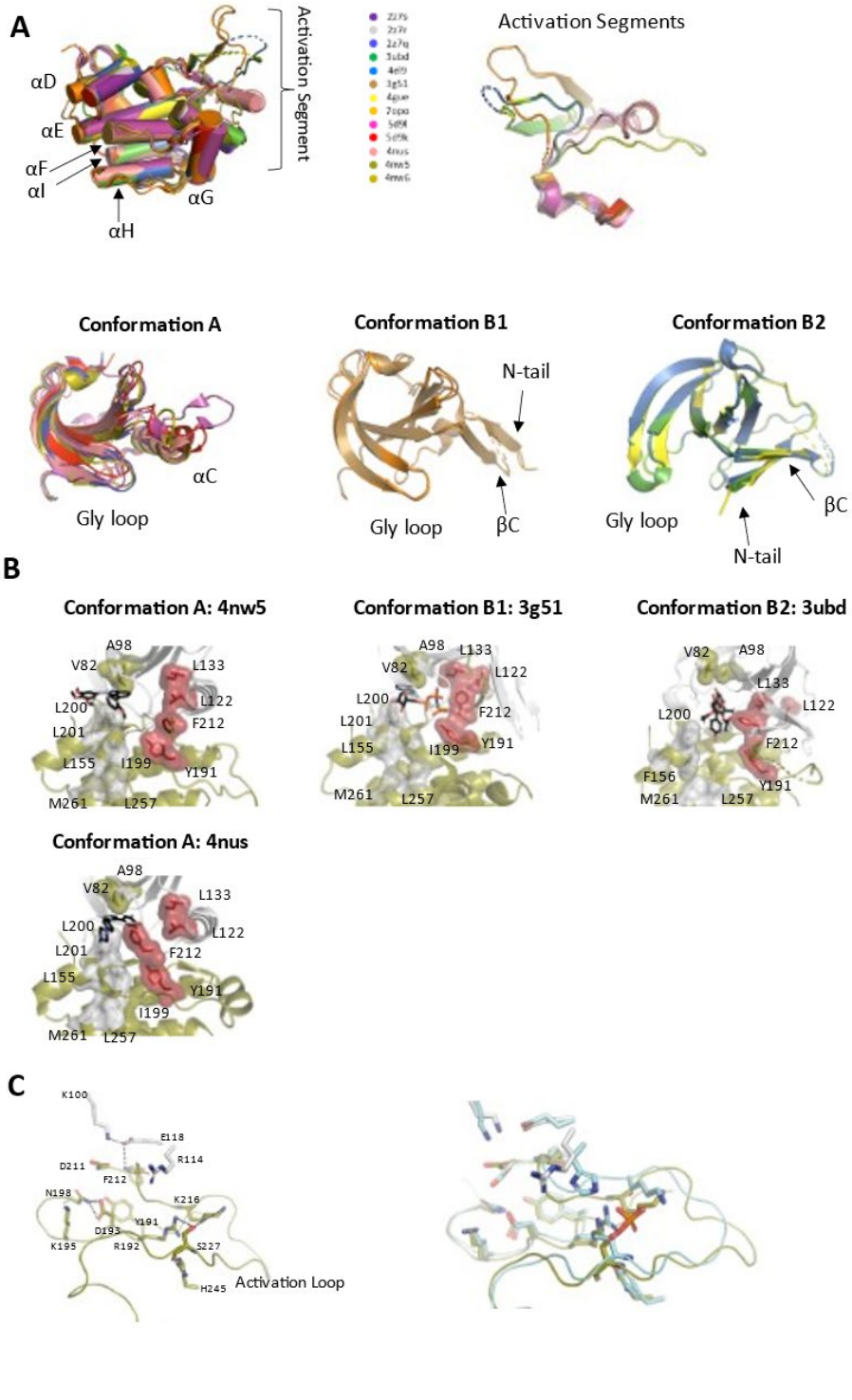
unusual β sheet (hereafter referred to as βC) that replaces the αC Helix.(Malakhova, Kurinov et al. 2009, Utepbergenov, Derewenda et al. 2012). We designate these structures as Conformation B with two subtypes, B1 and B2. The β9 strand of the activation segment along with the βC replace the position of the canonical αC helix in B1 structures, while in B2 structures, the βC, in concert with an N-terminal β strand, replace the αC helix. The B2 conformation is also characterized by a distortion of the glycine rich loop, where it is rotated out away from the active site cleft.

These conformations also adopt unique internal hydrophobic architectures, which can be appreciated by inspecting their hydrophobic spines (Fig. 3.4B). All kinases contain two sets of non-contiguous hydrophobic residues that span the N and C-lobes. The catalytic (C) spine is kept intact and is completed by the adenine ring of ATP, while the regulatory (R) spine is dynamically assembled. The R spine is broken for inactive kinases and assembled for active kinases (Kornev, Haste et al. 2006, Kornev, Taylor et al. 2008). The C spines of all NTK structures are fairly similar across all conformations, while the R spines are more dynamic. In the B2 conformation the R spine is broken, indicating that this is an inactive conformation, completely incapable of catalytic activity. The B1 R spine is more of an intermediate case. While the spine is contiguous, it experiences a branching at RS3 (L122) and RS4 (L133) that is not typical for canonical active AGC kinases, suggesting that this may be some form of an intermediate state. The R spines are variable in conformation A, broken when the RSK2 NTK is bound to the inhibitor LJH685, but intact when bound to a benzoxazole compound (Costales, Mathur et al. 2014) (pdb: 4nw5, 4nw6). Thus, some Conformation A structures better resemble an active-like AGC kinase

conformation. Furthermore, the activation segment of the NTK bound to the benzoxazole compound, Conformation A, is quite similar to the activation segment of an active kinase (Fig. 3.4C). However, the absence of an activation loop phosphate prevents coupling to basic residues in the αC helix and HRD motif of the catalytic loop, which are required for an active AGC kinase (Taylor and Kornev 2011, Taylor, Keshwani et al. 2012, Kornev and Taylor 2015). Thus, the activation loop phospho-site is positioned to establish conserved interactions upon activation loop phosphorylation, which would promote formation of a closed active kinase. This analysis of NTK structures and conformations in combination with the HDXMS demonstrates the flexibility and malleability of the NTK. The flexible nature of the NTK allows various inhibitors to stabilize the NTK in more inactive (B2), intermediate (B1), and active-like (A) conformations and highlight its extreme malleability. It should be emphasized that none of these NTK structures has the N-terminal linker extension wrapped around the N-lobe which is a conserved feature of all active AGC kinases (Newton 2003, Kannan, Haste et al. 2007).

Figure 3.4. Conformational diversity of the NTK

(A) Top displays all of the C-lobes of RSK1/2 NTK crystal structures, aligned to pdb: 4nus, and the alignment of all of the activation segments. Each color corresponds to a unique crystal structure. The bottom displays all of the N-lobes of the NTK crystal structures aligned to the pdb: 4nus. The N-lobes partition into two main conformation types, A and B. Conformation B has two subtypes, B1 and B2.
(B) Hydrophobic spines of a sampling of NTK structures from Conformations A and B. Regulatory (R) spines are shown in red, while catalytic (C) spines are shown in grey, for C-lobe C spines, or green for N-lobe C spines.
(C) Left, an overview of highly conserved regulatory and catalytic residues in the β3, αC, catalytic loop, and activation segment for Conformation A, pdb: 4nw5. Black dashed lines represent hydrogen bonding or electrostatic interactions. Right is an overlay of the conserved regulatory and catalytic residues from an active conformation of PKA, pdb: 1atp (cyan), and an NTK in Conformation A, pdb:4nw5 (green).


A
αD
αE
αF
αI
αH
αG
Activation Segment
Activation Segments
Conformation A
αC
Gly loop
Conformation B1
N-tail
Gly loop
βC
Conformation B2
Gly loop
βC
N-tail
B
Conformation A: 4nw5
A98
V82
L133
L122
L200
F212
L201
L155
I199
Y191
M261
L257
Conformation B1: 3g51
A98
L133
V82
L122
L200
F212
L201
L155
I199
Y191
M261
L257
Conformation B2: 3ubd
V82
A98
L133
L200
L122
F212
F156
Y191
M261
L257
Conformation A: 4nus
A98
V82
L133
L200
L122
L201
F212
L155
Y191
I199
M261
L257
C
K100
E118
D211
R114
F212
N198
K216
Y191
D193
K195
R192
S227
H245
Activation Loop

3. The CTK binds to the HM of the Linker

The strong solvent accessibility of the linker is not entirely unexpected, due to this region's importance for mediating additional steps in RSK activation. The linker must remain accessible to facilitate CTK phosphorylation of the HM, and PDK1 binding to enable NTK activation loop phosphorylation. Following ERK2 binding and phosphorylation of the CTK, the molecular details that define the remaining steps in RSK activation are not well understood. We thus sought to better understand how the CTK and PDK1 engage the linker to enable HM and NTK activation loop phosphorylation, respectively.

To further investigate how the CTK, following ERK2 phosphorylation, is able to bind and phosphorylate the HM (FXXFS/TF), we tested the ability of phosphorylated CTK to bind peptide array that covers the entire Linker (Fig. 3.5A). As expected, the CTK binds HM-containing peptides, and this interaction appears to be dependent on the P-3 R383, and several residues upstream of the HM. Moreover, the CTK also interacts with the C-terminal end of the linker (400-416) and is dependent on L411 and H412. Phosphorylation of the HM weakens this interaction, while phosphorylation of other conserved sites such as the turn motif do not appear to significantly alter binding (Fig. 3.5B). The reliance of the HM phosphorylation state for CTK binding is not surprising because following phosphorylation by the CTK, the pHM needs to be accessible to interact with PDK1 to facilitate NTK activation loop phosphorylation.

Upon closer inspection of the HM, we noticed that the HM resembles the autoinhibitory αL helix that lies in the substrate binding groove of the CTK, for all

RSK isoforms (Fig. 3.6A). The αL helix occupies the CTK substrate binding groove, and is stabilized by an electrostatic interaction between K700 and E500/D503 in the αD helix, and hydrogen bonding of Y706 and S603 of the αF (Malakhova, Tereshko et al. 2008, Li, Fu et al. 2012). Furthermore, the HM of the CTK is also a CaMKII recognition motif (White, Kwon et al. 1998). CaMKII phosphorylates substrates with an Arg in the P-3 position and hydrophobic residue (φ) in the P+1 position. The P-3 Arg interacts with acidic residues in the αD helix, and the P+1 hydrophobic residue fills a hydrophobic P+1 loop pocket, positioning the phospho-site toward the catalytic center (Fig. 3.6B). Mutation of both the P-3 Arg, and P+1 hydrophobic residue significantly weakens the affinity of CaMKII and its substrate (Chao, Pellicena et al. 2010, Castro-Rodrigues, Zhao et al. 2018). Therefore, we predict that R383 interacts with acidic αD residues E500/D503 to position the phospho-site of the HM, S386, in the active site of the CTK. Intriguingly, an AlphaFold (AF) prediction of the HM through CTK, truncated to remove the αL, positions the HM in the active site, with R383 facing E500 and D503 of the αD (Fig. 3.6C). Furthermore, the AF model also predicts L411 and H412 interact with the N-lobe of the CTK, in agreement with previous peptide arrays. To validate these models, we generated mutations to disrupt electrostatic interactions between R383 and the CTK αD (Fig. 3.6D). Mutation of each residue in the HM, block HM and NTK activation loop phosphorylation. Crystal structures of PDK1 bound to HM containing peptides, and a modelling of the RSK2 HM bound to PDK1 indicate that the P-3 residue does not contribute to PDK1 binding (Frödin, Antal et al. 2002), therefore the loss of NTK activation loop phosphorylation for R383A is likely due to a reduction in HM phosphorylation, rather than altered

PDK1 binding. E500R was completely unable to phosphorylate the HM. E500 homologues in other CAMKs and in PKA, coordinate ribose hydroxyls of ATP (Narayana, Cox et al. 1997, Narayana, Cox et al. 1997), so lack of HM phosphorylation may be due to defective ATP binding rather than perturbed HM substrate binding. Charge swap mutation for D503R was not tested due to a lack of expression. We hypothesize that following ERK2 phosphorylation of the CTK, the αL inhibitor helix is displaced from the substrate docking site to facilitate HM binding. R383 establishes electrostatic interactions with the αD helix and F387 fills the P+1 loop to sufficiently bind the CTK until phosphorylation of S386, followed by pHM release.

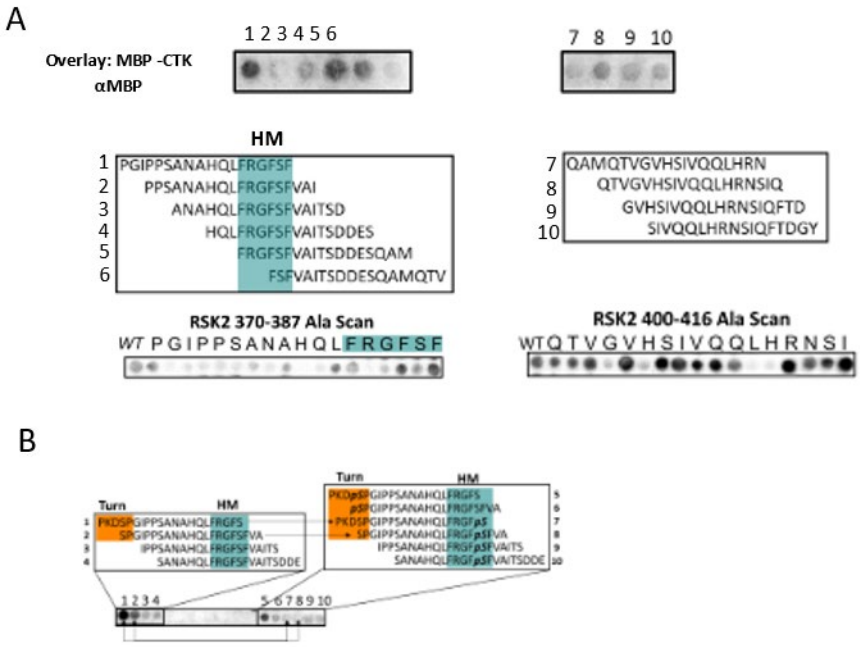


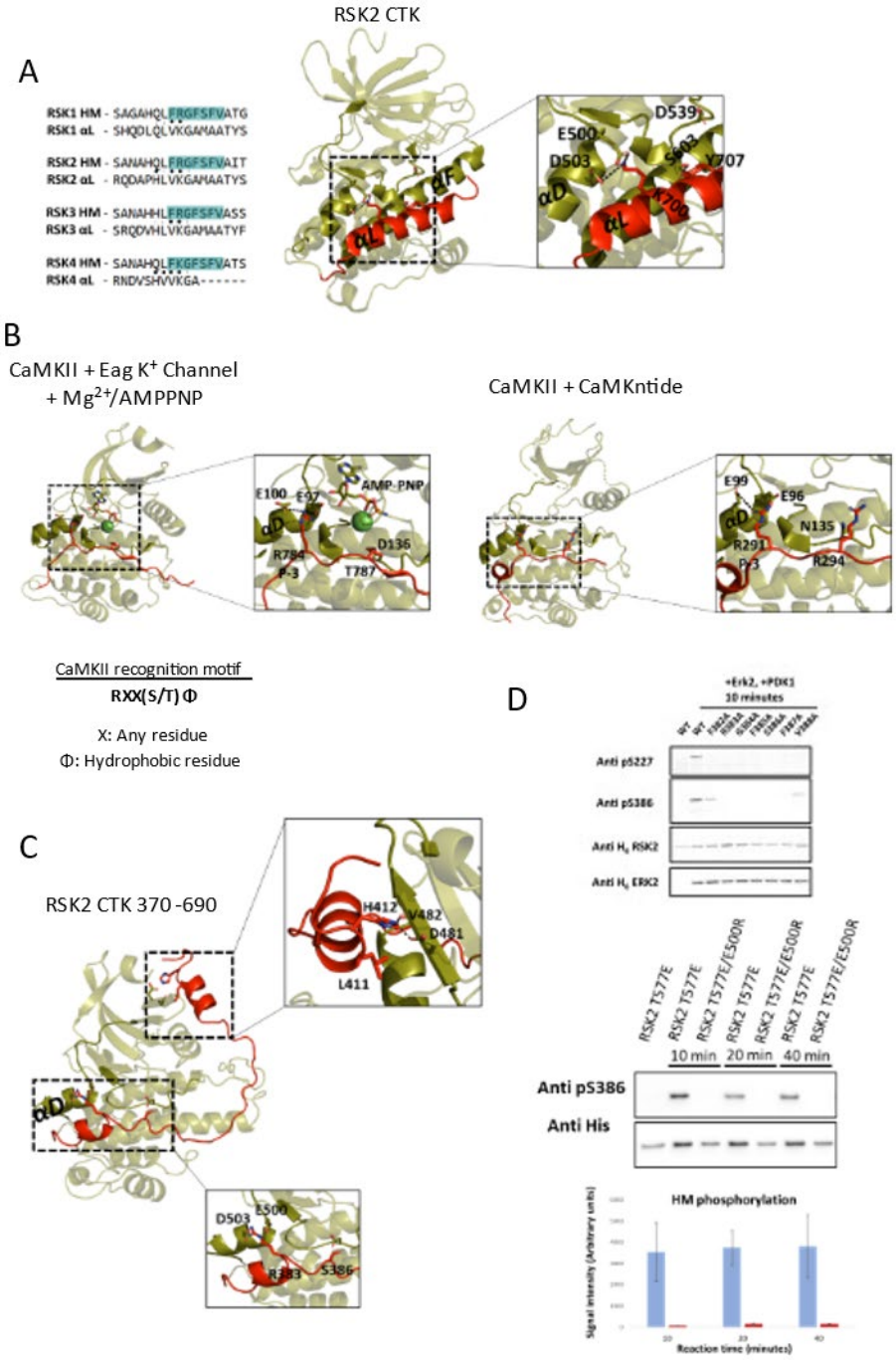
Figure 3.5. The active, phosphorylated CTK binds to Linker and HM peptides.

(A) Peptide array of the NTK C-lobe through linker, overlaid with phosphorylated (pT577) MBP-CTK^{RSK2}-H$_6$. Boxed regions display specific peptide sequences that reproducibly bound pT577 MBP-CTK^{RSK2}-H$_6$. Ala scan peptide arrays were performed for peptides #1 and 8.

(B) Peptide array of NTK HM containing phosphopeptides, overlaid with pT577 MBP-CTK^{RSK2}-H$_6$. HM containing peptides contain no phosphorylation, phosphor-serine, or phosphor-threonine at the NTK turn motif or HM. Arrows connect pairs of peptides that differ only by S386 (HM) phosphorylation.

(C) Electrostatic potential of RSK2 CTK (pdb:2qr8) boxed regions indicate locations of αD helix and active site.

Figure 3.6. CaMKII substrate binding resembles CTK:αL binding.
(A) Left, sequence comparison of each RSK isoform HM and autoinhibitory αL helix. In each RSK HM isoform, a P-3 Arg and P-4 Phe are similar to a Lys and Val, respectively. Right, in RSK CTK (pdb:2qr8) crystal structure, K700 forms electrostatic contacts with E500 and D503 of the αD helix, and Y707 forms an hydrogren bond with S603 in the extended αF helix.
(B) Left, crystal structure of CaMKIIc bound to Eag potassium channel substrate fragment in the presence of Mg/AMPPN (pdb:5h9b). Inset shows detailed view of substrate binding. P-3 Arg forms ionic interactions with glutamates in αD helix, and the P site faces HRD Asp. Right, crystal structure of CaMKII bound with CaMKIINtide inhibitor. Inset shows detailed view of inhibitor fragment binding. P-3 Arg forms ionic interactions with glutamates in αD helix, and the P site pseudosubstrate faces the active site. Bottom, minimum consensus CaMKII substrate consensus sequence.
(C) AlphaFold2 model of RSK2 (370-690) HM-CTK ΔαL. HM through the linker is colored red, and the kinase domain is colored dark green. Insets highlights ionic interaction between R383 (P-3) and acidic αD, and S386 (P-site) pointing toward HRD Asp. Other inset indicates the linker at L411 and H412 interact with CTK N-lobe at αCβ4.
(D) Top, In vitro activation of MBP-RSK2-H_6 for HM mutants, immunoblotted against phospho-specific antibodies for the NTK activation loop or HM. Bottom, western blot of kinase reaction of MBP-RSK2 with CTK phosphomimetic mutation and/or E500R mutation for various lengths of time, immunoblotted with phospho HM antibody. Blot is quantified following normalization below.


A
RSK2 CTK
RSK1 HM - SAGAHQLFRGFSFVATG
RSK1 αL - SHQDLQLVKGAMAATYS
RSK2 HM - SANAHQLFRGFSFVAIT
RSK2 αL - RQDAPHLVKGAMAATYS
RSK3 HM - SANAHHLFRGFSFVASS
RSK3 αL - SRQDVHLVKGAMAATYF
RSK4 HM - SANAHQLFKGFSFVATS
RSK4 αL - RNDVSHVVKGA------
αF
αL
αD
D539
E500
D503
S363
Y707
T708
B
CaMKII + Eag K+ Channel
+ Mg2+/AMPPNP
CaMKII + CaMKntide
E100
E97
AMP-PNP
αD
R784
P-3
T787
D136
E99
E96
N135
R291
P-3
R294
CaMKII recognition motif
RXX(S/T)Φ
X: Any residue
Φ: Hydrophobic residue
C
RSK2 CTK 370 -690
H412
V482
D481
L411
αD
D503
E500
R343
S386
D
+Erk2, +PDK1
10 minutes
WT
WT
T577A
S580A
T577A
S386A
T577A
Y704A
Anti p5227
Anti pS386
Anti H6 RSK2
Anti H6 ERK2
RSK2 T577E
RSK2 T577E
RSK2 T577E/E500R
RSK2 T577E
RSK2 T577E/E500R
RSK2 T577E
RSK2 T577E/E500R
10 min
20 min
40 min
Anti pS386
Anti His
HM phosphorylation
Signal intensity (Arbitrary units)
Reaction time (minutes)

4. PDK1 binds the HM and Ade motif of the Linker

Next, we explored the mechanism of PDK1 binding to the Linker (Fig. 3.7A), and as expected, PDK1 interacts with hydrophobic motif peptides. In previous papers, modulation of the P-1 or P-4 Phe in other AGC kinases greatly weakens PDK1 binding and catalytic efficiency (Etchebehere, Van Bemmelen et al. 1997, Batkin, Schvartz et al. 2000, Biondi, Cheung et al. 2000, Gao, Toker et al. 2001). Here, however, we find that mutation of the P+1 F387, significantly weakened PDK1 HM peptide binding (Fig. 3.7B). A phosphorylated HM (pHM) is required for NTK activation loop phosphorylation (Fig. 3.6D), as S386A abolishes NTK activation loop phosphorylation, and has been demonstrated to be necessary for high affinity binding to PDK1 in other kinases (Frödin, Jensen et al. 2000). Therefore, we were surprised to find that PDK1 only bound to non-phosphorylated HM peptides in our peptide arrays (Fig. 3.7B). Phosphorylation of other sites such as the turn motif or the newly identified Tor interaction motif (TIM), did not appear to affect PDK1 binding (Baffi, Lordén et al. 2021). We verified this dependence on the phosphorylation state of the pHM using full length (FL) purified proteins (Fig. 3.7C). PDK1 bait successfully pulls down phosphatase treated RSK2, while non-λ phosphatase treated RSK2 that contains detectable pS386 phosphorylation, failed to interact with PDK1 bait.

We also found that the Ade motif of the active site tether (AST) can interacted with PDK1 as well (Fig. 3.8A). The Ade motif typically contributes to binding of the ATP adenine ring. PDK1 lacks its own Ade motif, and the corresponding Ade motifs of PKA, AKT, and PRK2 have been shown to bind to PDK1. Thus, PDK1 may utilize

the Ade motif of substrate kinases to complete its own ATP binding site (Narayana, Cox et al. 1997, Romano, Kannan et al. 2009). Here, in RSK2, mutation of F355, Y356, F357, D358, and F361 inhibits Ade peptide binding to PDK1. In FL proteins however, only Y356A weakens NTK activation loop phosphorylation. Similar to PKA, the RSK2 Ade motif may complete the ATP binding pocket of PDK1 to facilitate phosphorylation of the NTK activation loop (Fig. 3.8B).

In our analysis of NTK crystal structures, we also noticed in Conformation A structures, the Ade motif partially fills the PIF pocket of a crystallographically generated symmetry mate (Jain, Mathur et al. 2015) (Fig. 3.8C). In RSK2, for example, F355 fills the hydrophobic pocket ordinarily occupied by the P-1 Phe of the HM, while Y356 fills the corresponding position of the P site (S386). We thus reasoned that the Ade motif of RSK2 may partially fill the PIF pocket of PDK1 which introduces an alternative mechanism for Ade motif binding to PDK1. The ability of the RSK2 Ade motif to partially fill the PIF pocket, also raises the question of whether the HM also can partially complete the ATP binding pocket? We aligned an HM peptide to the Ade fragment in our model of Ade bound to the ATP binding site of PDK1, and found that the P+1 Phe, F387 (HM), is superimposed with F357 (Ade) to complete the ATP binding site (Fig. 3.8D). HM binding to the ATP binding site of PDK1 would also explain the dependence on the P+1 Phe, F387 (Fig. 3.7B). It has been well established that phosphorylation of the HM is required for high affinity PDK1 binding, and activation loop phosphorylation in RSK and other AGC kinases (Etchebehere, Van Bemmelen et al. 1997, Batkin, Schvartz et al. 2000, Biondi, Cheung et al. 2000, Frödin, Jensen et al. 2000, Gao, Toker et al. 2001, Frödin, Antal et

al. 2002). Therefore, our observed binding of PDK1 to only non-phosphorylated RSK2 suggests that we may have captured an intermediate conformation that is incapable of NTK activation loop phosphorylation. Perhaps, the RSK2 Ade motif fills the PDK1 PIF pocket, while the RSK2 HM completes the ATP binding site of PDK1. HM phosphorylation may cause a conformational change that enables the pHM to fill the PDK1 PIF pocket and the Ade motif to complete the ATP binding site and facilitate NTK activation loop phosphorylation.

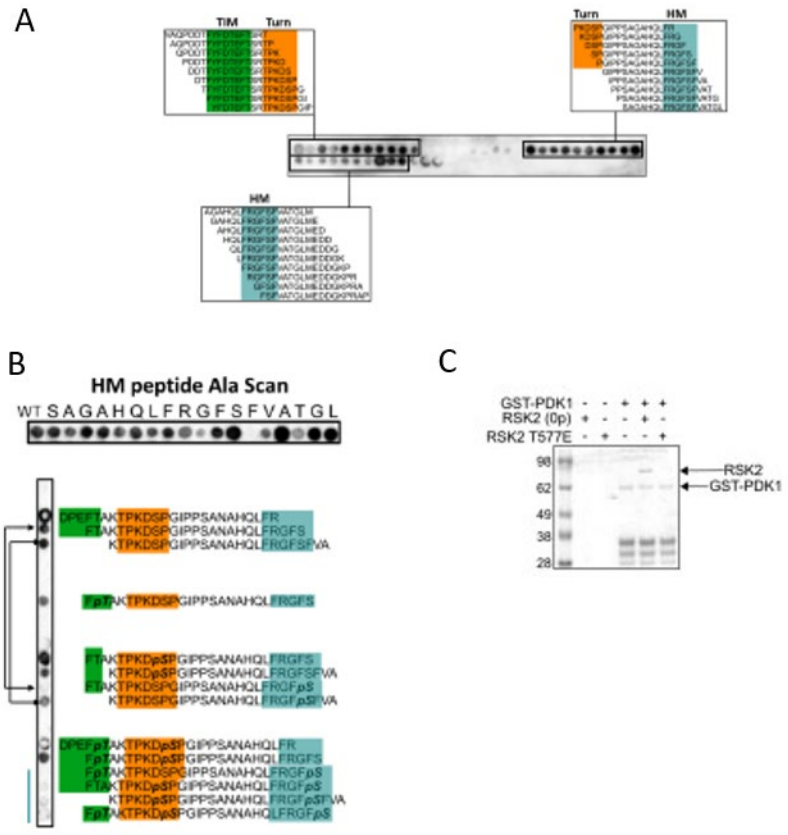


Figure 3.7. PDK1 binds to non-phosphorylated HM in peptides and full length RSK2.

(A) Peptide array of NTK^{RSK1} AGC tail, overlaid with PDK1 cell lysate immunoblotted against Myc tag. Insets display specific binding peptide sequences. (B) Top, peptide array ala scan of HM peptide overlaid with GST-PDK1, immunoblotted against GST tag. Bottom, peptide array of NTK HM containing peptides with or without phosphopeptides, overlaid with GST-PDK1, immunoblotted against GST tag. Green, orange, and cyan highlights Ade/AST, turn, and HM, respectively. Arrows connect pairs of peptides that differ by HM phosphorylation. (C) In vitro RSK2 pulldown using GST-PDK1 as bait. GST-PDK1 is immobilized on glutathione resin, RSK2 T577E is basally autophosphorylates its HM, RSK2 (0P) is treated with λ phosphatase to generate a completely non-phosphorylated protein.

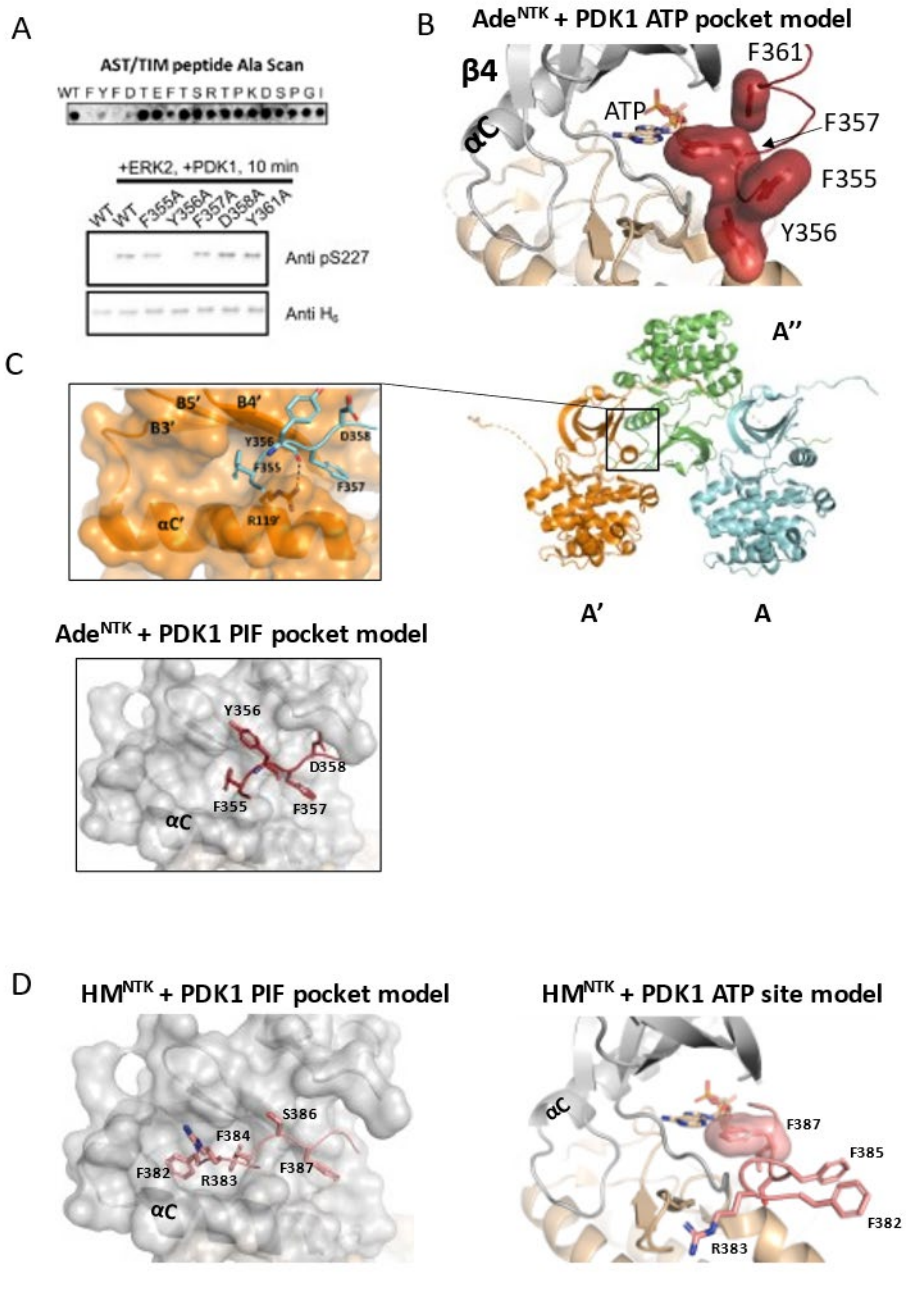
Figure 3.8. PDK1 binds RSK2 Ade motif and this may engage PDK1 ATP binding site or PIF pocket.

(A) Top, peptide array ala scan of AST/Ade motif peptide overlaid with GST-PDK1, immunoblotted against GST tag. Bottom, western blot of kinase reaction of MBP-RSK2-H_6 Ade mutations, incubated with ERK2 and PDK1, immunoblotted against NTK activation loop phospho-antibody.

(B) Model of the RSK2 Ade motif filling the PDK1 ATP binding pocket. The Rsk2 Ade motif is colored red.

(C) Crystallographically generated symmetry mates of the NTK in pdb: 5d9k. Inset reveals how the RSK2 Ade motif partially fills the PIF pocket of a symmetry mate. Beneath the inset is a model of the RSK2 Ade motif partially filling the PIF pocket of PDK1.

(D) Left, model of the RSK2 HM docked into the PIF pocket of PDK1, the HM is colored in salmon. Right, model of the RSK2 HM filling the PDK1 ATP pocket


A
AST/TIM peptide Ala Scan
WT F Y F D T E F T S R T P K D S P G I
+ERK2, +PDK1, 10 min
WT
WT
F355A
Y356A
F357A
D358A
Y361A
Anti pS227
Anti H6
B
Ade^NTK + PDK1 ATP pocket model
β4
αC
ATP
F361
F357
F355
Y356
C
β5'
β4'
β3'
Y356
D358
F355
F357
αC'
R119'
A''
A'
A
Ade^NTK + PDK1 PIF pocket model
Y356
D358
F355
F357
αC
D
HM^NTK + PDK1 PIF pocket model
S386
F384
F382
R383
F387
αC
HM^NTK + PDK1 ATP site model
αC
F387
F385
F382
R383

C. Discussion

The p90 RSK kinases with their multistep activation pathway are defined by their two distinct kinase domains, three IDRs with critical SLiMs (Anjum and Blenis 2008, Romeo, Zhang et al. 2012). The concerted action of the NTK, CTK, ERK2, and PDK1, the RSK Linker and C-terminal tail coordinate the generation of an active NTK. Thus, further characterization of this pathway can provide a greater understanding of how IDRs and kinase domains integrate to regulate kinase function (Gógl, Kornev et al. 2019). Here, we performed HDXMS on the stable pre-catalytic FL RSK2-ERK2 kinase complex to explore the dynamic or unstructured regions of the complex. This allowed us to analyze the conformational malleability of the NTK, and explore how the Linker/AGC extension engage a phosphorylated CTK and PDK1.

HDXMS reveals that the tails flanking all three kinases, in particular the N-terminal tail, Linker, and C-terminal tail of RSK2 display the greatest solvent exposure in the complex, indicating that these are the most dynamic or disordered regions. The N-terminal tail of RSK is poorly conserved among the four RSK isoforms and has unknown function. Similar N-terminal tails of other AGC kinases such as PKA aid in kinase thermal stability and cellular targeting (Taylor, Søberg et al. 2021) while in other AGC kinases such as PKC the N-terminus of the kinase domain is fused to other folded domains such as the C1 and C2 domains that facilitate localization by binding to second messengers like calcium and diacyl glycerol (Newton 1995, Newton 2018). The N-terminus of RSK3 contains a putative nuclear localization signal (NLS) (Zhao, Bjørbaek et al. 1995), and RSK2 N-terminal deletion constructs are more susceptible to

precipitation (Malakhova, Kurinov et al. 2009). Thus, the N-terminal tail in RSK isoforms likely contains other SLiMs, and may have similar functions in localization and/or thermal stability. Variable N-terminal tails are used in other kinase families to provide isoform specificity. The N-terminal tails of PKA Cα and various Cβ isoforms are poorly conserved, and it is believed that sequence differences in its tail may explain isoform functional diversity and subcellular localization patterns (Taylor, Wallbott et al. 2021). The strong solvent exposure and flexibility of the Linker and AGC extension, particularly for the HM, is necessary to facilitate the various steps in RSK activation. The HM serves as a good example of a versatile SLiM capable of different interactions depending on the biochemical and cellular environment (Tompa, Davey et al. 2014, Gouw, Michael et al. 2018, Gógl, Kornev et al. 2019). The HM is a substrate for the CTK, a docking site and a *trans* activator of PDK1, and the HM stabilizes the NTK active conformation by filling the NTK PIF pocket, anchoring the AGC extension to the N-lobe (Kannan, Haste et al. 2007, Anjum and Blenis 2008, Romeo, Zhang et al. 2012, Newton 2018). The RSK C-terminal tail is strongly solvent exposed as well. Structures of the isolated CTK, indicate the C-terminal tail, following the αL helix is completely disordered and not visible. ERK2 binding orders the C-terminal tail such that only the very C-terminus is disordered (Malakhova, Tereshko et al. 2008, Alexa, Gógl et al. 2015), and MD simulations confirms that this interface is quite stable. Therefore, the observed, strong solvent exposure of the CTK C-terminal tail may be due to the lack of secondary structure and flexibility of residues flanking the revD motif.

All three kinases displayed solvent protection/exposure patterns similar to other analyzed kinases. The modest solvent exposure observed in the C-lobe, particularly the

αGαHαI helices for all three kinases is indicative of their role in mediating protein:protein interactions. PKA utilizes this surface for additional interactions with its various Regulatory (RI/RII) subunits (Taylor, Ilouz et al. 2012) as well as for its interactions with the heat stable protein kinase inhibitor (PKI) (Knighton, Zheng et al. 1991, Olivieri, Wang et al. 2020). Moreover, the αH-αI loop can also serve as a distal tethering site for other substrates (Deminoff, Ramachandran et al. 2009). Similarly, the MAPK insert of ERK2 mediates nucleic acid binding among other binding partners (Roskoski 2012). ERK2 and the CTK are better shielded from solvent than the NTK, suggesting that these kinases sample fewer conformational states and are maintained in a more active-like conformation. In apo, non-phosphorylated conditions, both kinases have intact regulatory (R) spines, and their αC helices are maintained in active conformations. ERK2 activity is lacking due to a non-phosphorylated activation loop and incomplete P+1 Loop for substrate binding (Canagarajah, Khokhlatchev et al. 1997), and CTK activity is dampened by the autoinhibitory αL helix blocking access to the substrate docking groove . The CTK is an intriguing kinase due to its highly specialized nature and unusual long, extended APE/αF helix that is observed in both the presence and absence of ERK2. The HM is the only known substrate of the CTK, thus following CTK activation loop phosphorylation, only displacement of αL and rearrangement of the extended APE/αF helix are needed to allow access to the active site and arrangement of a P+1 substrate docking loop (Malakhova, Tereshko et al. 2008, Li, Fu et al. 2012).

The N-lobe of the NTK is malleable, and capable of sampling a wide conformational space. Recently, an analysis of the conformational space of the LRRK2

kinase domain indicates that different inhibitors can stabilize the kinase into distinct conformational states (Weng, Aoto et al. 2022). We made similar observations here, we analyzed crystal structures of the NTK bound to various inhibitors, and observed two major conformational states (A and B), that are differentiated by the secondary structure and position of the αC helix. Mobility of the αC helix is well characterized in other kinases as is its unfolding and disorder. However, we are not aware of any other kinase that refolds its αC into a βC strand as in NTK Conformation B structures. Because these NTK structures all lack the AGC extension or an ordered phosphorylated activation loop, they highlight the conformational diversity and/or malleability of inactive kinases. Most of the NTK structures feature broken R spines and activation loops that are either disordered or in inactive conformations. However, a few Conformation A (pdb:4nw5, 4nw6) structures are in 'active-like states" with intact R spines and mostly properly positioned catalytic and regulatory residues. These active-like structures suggest that assembly of an R spine does not necessarily require activation loop phosphorylation nor an AGC extension. Moreover, activation loop phosphorylation and AGC extension anchoring appear to be the final steps for the assembly of an active kinase.

To better understand how the highly solvent accessible Linker and HM bind the CTK, we recognized similarities between the HM, the autoinhibitory αL, and a CAMK consensus substrate, allowing us to construct a model proposing how the HM docks the ERK2 phosphorylated CTK (Chao, Pellicena et al. 2010, Castro-Rodrigues, Zhao et al. 2018). We propose that electrostatic interactions between R383(P-3) and F387(P+1) of the HM mediate electrostatic and hydrophobic interactions, respectively, to allow HM phosphorylation. Intriguingly, both R383 and F387 are known Coffin Lowry Syndrome

(CLS) mutations (Trivier, De Cesare et al. 1996); CLS mutations are thought to inhibit RSK2 catalytic activity, and mutation of both abolished NTK activation loop phosphorylation. Mutation of E500R abolished HM phosphorylation, and this mutation may also disrupt CTK ATP binding. It is therefore important to determine the effect of the D503R mutation. Since this mutant protein did not express, more fine tuning of the expression conditions is needed to test the effect of disrupting the potential electrostatic interaction between R383 and the acidic αD helix. Ca^{2+}/CaM binding to CaMK results in a dramatic conformational change that causes the autoinhibitory helix to be completely displaced and rotated out of the active site, providing substrates and nucleotides access to the active site (Rellos, Pike et al. 2010). Does the CTK experience as large of a conformational change following phosphorylation by ERK2? Does this displacement of the αL autoinhibitory helix and formation of the P+1 loop reposition ERK2 relative to the rest of RSK? These are outstanding questions in elucidating RSK activation. For future work, utilizing both low and high resolution structural techniques such as SAXS and crystallography would be important to determine if a large conformational change occurs following ERK2 phosphorylation, and, if so, what does this conformational change entail?

We also sought to better understand PDK1 engagement and phosphorylation of the NTK activation loop. We verified that PDK1 binds to the NTK HM, and that HM phosphorylation is required for activation loop phosphorylation. Furthermore, the Ade motif interacts with PDK1 similarly to how the Ade motifs of PKA, AKT, and PRK2 engage PDK1 (Romano, Kannan et al. 2009). In PKA, the activation segment, APE motif, and αH-αI loop also contribute to PDK1 binding (Moore, Kanter et al. 2002,

Yang, Wu et al. 2012). In future work, we would want to determine if the corresponding sites of the NTK similarly contribute to PDK1 binding and activation loop phosphorylation. We made the unexpected finding that PDK1 strongly binds and forms a stable complex with non phosphorylated RSK2. The importance of HM phosphorylation for PDK1 enabled NTK activation loop phosphorylation suggests that we may have captured RSK2 and PDK1 in a catalytically incompetent conformation, with potentially the Ade motif partially filling the PDK1 PIF pocket and the HM filling the PDK1 ATP binding site. The biological relevance of this potential conformation is unclear, but to our knowledge, there are no structures of PDK1:substrate kinase complexes. Biophysical and structural characterization of this non-phosphorylated RSK2:PDK1 complex to identify the interfaces and the size and shape of the complex could provide valuable insight into how PDK1 engages a substrate kinase.

There are additional aspects of RSK signaling and activation that remain poorly understood. Recently, a Tor interaction motif (TIM) site was identified in PKC, and it was found that its phosphorylation by mTORC controls HM phosphorylation. Newly translated PKC is thought to form a nascent TIM-mediated dimer, and mTORC phosphorylation of this site breaks this dimer interface, allowing maturation of PKC (Baffi, Lordén et al. 2021). Based on sequence alignments a TIM site is present in all RSK isoforms, although very little is known about this site's function, or potential impact on the other conserved AGC extension motifs. A recent study found that optimal HM phosphorylation requires intact mTORC2, but mTOR inhibition did not affect HM phosphorylation suggesting that mTORC2 may act as more of a scaffold (Chou, Rajput

et al. 2020). RSK2 also forms low affinity, transient dimers (Utepbergenov, Hennig et al. 2016), but it is unknown if these dimers utilize a TIM mediated surface like PKC.

Here we have provided solvent accessibility information on the RSK2-ERK2 complex to identify the most dynamic and flexible regions of the complex. The IDRs that flank the kinase domains, the N-terminal tail, Linker, and C-terminal tail, have the greatest solvent accessibility. The solvent accessibility of the NTK indicates that it is quite dynamic and malleable. Going forward, additional biochemical experiments and combinations of low and high resolution structural techniques will be needed to determine the molecular details that control the additional steps and conformational changes associated with RSK activation. Furthermore, additional analysis of the RSK2:ERK2 complex and RSK activation are valuable to better understand how folded kinase domains interact with critical IDRs.

D. References

Alexa, A., G. Gógl, G. Glatz, Á. Garai, A. Zeke, J. Varga, E. Dudás, N. Jeszenői, A. Bodor, C. Hetényi and A. Reményi (2015). "Structural assembly of the signaling competent ERK2-RSK1 heterodimeric protein kinase complex." Proc Natl Acad Sci U S A **112**(9): 2711-2716.

Anjum, R. and J. Blenis (2008). "The RSK family of kinases: emerging roles in cellular signalling." Nature Reviews Molecular Cell Biology **9**(10): 747-758.

Aronchik, I., B. A. Appleton, S. E. Basham, K. Crawford, M. Del Rosario, L. V. Doyle, W. F. Estacio, J. Lan, M. K. Lindvall, C. A. Luu, E. Ornelas, E. Venetsanakos, C. M. Shafer and A. B. Jefferson (2014). "Novel potent and selective inhibitors of p90 ribosomal S6 kinase reveal the heterogeneity of RSK function in MAPK-driven cancers." Mol Cancer Res **12**(5): 803-812.

Baffi, T. R., G. Lordén, J. M. Wozniak, A. Feichtner, W. Yeung, A. P. Kornev, C. C. King, J. C. Del Rio, A. J. Limaye, J. Bogomolovas, C. M. Gould, J. Chen, E. J. Kennedy, N. Kannan, D. J. Gonzalez, E. Stefan, S. S. Taylor and A. C. Newton (2021). "mTORC2 controls the activity of PKC and Akt by phosphorylating a conserved TOR interaction motif." Sci Signal **14**(678).

Bastidas, A. C., J. Wu and S. S. Taylor (2015). "Molecular features of product release for the PKA catalytic cycle." Biochemistry **54**(1): 2-10.

Batkin, M., I. Schvartz and S. Shaltiel (2000). "Snapping of the carboxyl terminal tail of the catalytic subunit of PKA onto its core: characterization of the sites by mutagenesis." Biochemistry **39**(18): 5366-5373.

Biondi, R. M., P. C. Cheung, A. Casamayor, M. Deak, R. A. Currie and D. R. Alessi (2000). "Identification of a pocket in the PDK1 kinase domain that interacts with PIF and the C-terminal residues of PKA." Embo j **19**(5): 979-988.

Canagarajah, B. J., A. Khokhlatchev, M. H. Cobb and E. J. Goldsmith (1997). "Activation mechanism of the MAP kinase ERK2 by dual phosphorylation." Cell **90**(5): 859-869.

Castro-Rodrigues, A. F., Y. Zhao, F. Fonseca, G. Gabant, M. Cadene, G. A. Robertson and J. H. Morais-Cabral (2018). "The Interaction between the Drosophila EAG Potassium Channel and the Protein Kinase CaMKII Involves an Extensive Interface at the Active Site of the Kinase." J Mol Biol **430**(24): 5029-5049.

Chao, L. H., P. Pellicena, S. Deindl, L. A. Barclay, H. Schulman and J. Kuriyan (2010). "Intersubunit capture of regulatory segments is a component of cooperative CaMKII activation." Nat Struct Mol Biol **17**(3): 264-272.

Chou, P. C., S. Rajput, X. Zhao, C. Patel, D. Albaciete, W. J. Oh, H. Q. Daguplo, N. Patel, B. Su, G. Werlen and E. Jacinto (2020). "mTORC2 Is Involved in the Induction of RSK Phosphorylation by Serum or Nutrient Starvation." Cells **9**(7).

Chrysostomou, S., R. Roy, F. Prischi, L. Thamlikitkul, K. L. Chapman, U. Mufti, R. Peach, L. Ding, D. Hancock, C. Moore, M. Molina-Arcas, F. Mauri, D. J. Pinato, J. M. Abrahams, S. Ottaviani, L. Castellano, G. Giamas, J. Pascoe, D. Moonamale, S. Pirrie, C. Gaunt, L. Billingham, N. M. Steven, M. Cullen, D. Hrouda, M. Winkler, J. Post, P. Cohen, S. J. Salpeter, V. Bar, A. Zundelevich,

S. Golan, D. Leibovici, R. Lara, D. R. Klug, S. N. Yaliraki, M. Barahona, Y. Wang, J. Downward, J. M. Skehel, M. M. U. Ali, M. J. Seckl and O. E. Pardo (2021). "Repurposed floxacins targeting RSK4 prevent chemoresistance and metastasis in lung and bladder cancer." Science Translational Medicine **13**(602): eaba4627.

Costales, A., M. Mathur, S. Ramurthy, J. Lan, S. Subramanian, R. Jain, G. Atallah, L. Setti, M. Lindvall, B. A. Appleton, E. Ornelas, P. Feucht, B. Warne, L. Doyle, S. E. Basham, I. Aronchik, A. B. Jefferson and C. M. Shafer (2014). "2-Amino-7-substituted benzoxazole analogs as potent RSK2 inhibitors." Bioorg Med Chem Lett **24**(6): 1592-1596.

Dalby, K. N., N. Morrice, F. B. Caudwell, J. Avruch and P. Cohen (1998). "Identification of regulatory phosphorylation sites in mitogen-activated protein kinase (MAPK)-activated protein kinase-1a/p90rsk that are inducible by MAPK." J Biol Chem **273**(3): 1496-1505.

Davey, N. E., M. S. Cyert and A. M. Moses (2015). "Short linear motifs – ex nihilo evolution of protein regulation." Cell Communication and Signaling **13**(1): 43.

Deminoff, S. J., V. Ramachandran and P. K. Herman (2009). "Distal recognition sites in substrates are required for efficient phosphorylation by the cAMP-dependent protein kinase." Genetics **182**(2): 529-539.

Derewenda, U., M. Artamonov, G. Szukalska, D. Utepbergenov, N. Olekhnovich, H. I. Parikh, G. E. Kellogg, A. V. Somlyo and Z. S. Derewenda (2013). "Identification of quercitrin as an inhibitor of the p90 S6 ribosomal kinase (RSK): structure of its complex with the N-terminal domain of RSK2 at 1.8 Å resolution." Acta Crystallogr D Biol Crystallogr **69**(Pt 2): 266-275.

Etchebehere, L. C., M. X. Van Bemmelen, C. Anjard, F. Traincard, K. Assemat, C. Reymond and M. Véron (1997). "The catalytic subunit of Dictyostelium cAMP-dependent protein kinase -- role of the N-terminal domain and of the C-terminal residues in catalytic activity and stability." Eur J Biochem **248**(3): 820-826.

Fisher, T. L. and J. Blenis (1996). "Evidence for two catalytically active kinase domains in pp90rsk." Mol Cell Biol **16**(3): 1212-1219.

Frödin, M., T. L. Antal, B. A. Dümmler, C. J. Jensen, M. Deak, S. Gammeltoft and R. M. Biondi (2002). "A phosphoserine/threonine-binding pocket in AGC kinases and PDK1 mediates activation by hydrophobic motif phosphorylation." The EMBO Journal **21**(20): 5396-5407.

Frödin, M., C. J. Jensen, K. Merienne and S. Gammeltoft (2000). "A phosphoserine-regulated docking site in the protein kinase RSK2 that recruits and activates PDK1." Embo j **19**(12): 2924-2934.

Gao, T., A. Toker and A. C. Newton (2001). "The carboxyl terminus of protein kinase c provides a switch to regulate its interaction with the phosphoinositide-dependent kinase, PDK-1." J Biol Chem **276**(22): 19588-19596.

Garai, Á., A. Zeke, G. Gógl, I. Törő, F. Fördős, H. Blankenburg, T. Bárkai, J. Varga, A. Alexa, D. Emig, M. Albrecht and A. Reményi (2012). "Specificity of Linear Motifs That Bind to a Common Mitogen-Activated Protein Kinase Docking Groove." Science Signaling **5**(245): ra74-ra74.

Gógl, G., A. P. Kornev, A. Reményi and S. S. Taylor (2019). "Disordered Protein Kinase Regions in Regulation of Kinase Domain Cores." Trends Biochem Sci **44**(4): 300-311.

Gouw, M., S. Michael, H. Sámano-Sánchez, M. Kumar, A. Zeke, B. Lang, B. Bely, L. B. Chemes, N. E. Davey, Z. Deng, F. Diella, C. M. Gürth, A. K. Huber, S. Kleinsorg, L. S. Schlegel, N. Palopoli, K. V. Roey, B. Altenberg, A. Reményi, H. Dinkel and T. J. Gibson (2018). "The eukaryotic linear motif resource - 2018 update." Nucleic Acids Res **46**(D1): D428-d434.

Ikuta, M., M. Kornienko, N. Byrne, J. C. Reid, S. Mizuarai, H. Kotani and S. K. Munshi (2007). "Crystal structures of the N-terminal kinase domain of human RSK1 bound to three different ligands: Implications for the design of RSK1 specific inhibitors." Protein Sci **16**(12): 2626-2635.

Jain, R., M. Mathur, J. Lan, A. Costales, G. Atallah, S. Ramurthy, S. Subramanian, L. Setti, P. Feucht, B. Warne, L. Doyle, S. Basham, A. B. Jefferson, M. Lindvall, B. A. Appleton and C. M. Shafer (2015). "Discovery of Potent and Selective RSK Inhibitors as Biological Probes." J Med Chem **58**(17): 6766-6783.

Johnson, D. A., P. Akamine, E. Radzio-Andzelm, M. Madhusudan and S. S. Taylor (2001). "Dynamics of cAMP-dependent protein kinase." Chem Rev **101**(8): 2243-2270.

Jones, S. W., E. Erikson, J. Blenis, J. L. Maller and R. L. Erikson (1988). "A Xenopus ribosomal protein S6 kinase has two apparent kinase domains that are each similar to distinct protein kinases." Proc Natl Acad Sci U S A **85**(10): 3377-3381.

Kannan, N., N. Haste, S. S. Taylor and A. F. Neuwald (2007). "The hallmark of AGC kinase functional divergence is its C-terminal tail, a cis-acting regulatory module." Proceedings of the National Academy of Sciences **104**(4): 1272-1277.

Knighton, D. R., J. H. Zheng, L. F. Ten Eyck, N. H. Xuong, S. S. Taylor and J. M. Sowadski (1991). "Structure of a peptide inhibitor bound to the catalytic subunit of cyclic adenosine monophosphate-dependent protein kinase." Science **253**(5018): 414-420.

Kornev, A. P., N. M. Haste, S. S. Taylor and L. F. Eyck (2006). "Surface comparison of active and inactive protein kinases identifies a conserved activation mechanism." Proc Natl Acad Sci U S A **103**(47): 17783-17788.

Kornev, A. P. and S. S. Taylor (2015). "Dynamics-Driven Allostery in Protein Kinases." Trends Biochem Sci **40**(11): 628-647.

Kornev, A. P., S. S. Taylor and L. F. Ten Eyck (2008). "A helix scaffold for the assembly of active protein kinases." Proc Natl Acad Sci U S A **105**(38): 14377-14382.

Li, D., T. M. Fu, J. Nan, C. Liu, L. F. Li and X. D. Su (2012). "Structural basis for the autoinhibition of the C-terminal kinase domain of human RSK1." Acta Crystallogr D Biol Crystallogr **68**(Pt 6): 680-685.

Madhusudan, P. Akamine, N. H. Xuong and S. S. Taylor (2002). "Crystal structure of a transition state mimic of the catalytic subunit of cAMP-dependent protein kinase." Nat Struct Biol **9**(4): 273-277.

Malakhova, M., I. Kurinov, K. Liu, D. Zheng, I. D'Angelo, J. H. Shim, V. Steinman, A. M. Bode and Z. Dong (2009). "Structural diversity of the active N-terminal kinase domain of p90 ribosomal S6 kinase 2." PLoS One **4**(11): e8044.

Malakhova, M., V. Tereshko, S. Y. Lee, K. Yao, Y. Y. Cho, A. Bode and Z. Dong (2008). "Structural basis for activation of the autoinhibitory C-terminal kinase domain of p90 RSK2." Nat Struct Mol Biol **15**(1): 112-113.

Manning, G., D. B. Whyte, R. Martinez, T. Hunter and S. Sudarsanam (2002). "The protein kinase complement of the human genome." Science **298**(5600): 1912-1934.

Moore, M. J., J. R. Kanter, K. C. Jones and S. S. Taylor (2002). "Phosphorylation of the catalytic subunit of protein kinase A. Autophosphorylation versus phosphorylation by phosphoinositide-dependent kinase-1." J Biol Chem **277**(49): 47878-47884.

Mora, A., D. Komander, D. M. van Aalten and D. R. Alessi (2004). "PDK1, the master regulator of AGC kinase signal transduction." Semin Cell Dev Biol **15**(2): 161-170.

Narayana, N., S. Cox, X. Nguyen-huu, L. F. Ten Eyck and S. S. Taylor (1997). "A binary complex of the catalytic subunit of cAMP-dependent protein kinase and adenosine further defines conformational flexibility." Structure **5**(7): 921-935.

Narayana, N., S. Cox, S. Shaltiel, S. S. Taylor and N.-h. Xuong (1997). "Crystal Structure of a Polyhistidine-Tagged Recombinant Catalytic Subunit of cAMP-Dependent Protein Kinase Complexed with the Peptide Inhibitor PKI(5−24) and Adenosine." Biochemistry **36**(15): 4438-4448.

Newton, A. C. (1995). "Protein Kinase C: Structure, Function, and Regulation (*)." Journal of Biological Chemistry **270**(48): 28495-28498.

Newton, A. C. (2003). "Regulation of the ABC kinases by phosphorylation: protein kinase C as a paradigm." Biochem J **370**(Pt 2): 361-371.

Newton, A. C. (2018). "Protein kinase C: perfectly balanced." Crit Rev Biochem Mol Biol **53**(2): 208-230.

Olivieri, C., Y. Wang, G. C. Li, M. V S, J. Kim, B. R. Stultz, M. Neibergall, F. Porcelli, J. M. Muretta, D. D. T. Thomas, J. Gao, D. K. Blumenthal, S. S. Taylor and G. Veglia (2020). "Multi-state recognition pathway of the intrinsically disordered protein kinase inhibitor by protein kinase A." eLife **9**: e55607.

Pearce, L. R., D. Komander and D. R. Alessi (2010). "The nuts and bolts of AGC protein kinases." Nat Rev Mol Cell Biol **11**(1): 9-22.

Rellos, P., A. C. Pike, F. H. Niesen, E. Salah, W. H. Lee, F. von Delft and S. Knapp (2010). "Structure of the CaMKIIdelta/calmodulin complex reveals the molecular mechanism of CaMKII kinase activation." PLoS Biol **8**(7): e1000426.

Romano, R. A., N. Kannan, A. P. Kornev, C. J. Allison and S. S. Taylor (2009). "A chimeric mechanism for polyvalent trans-phosphorylation of PKA by PDK1." Protein Science **18**(7): 1486-1497.

Romeo, Y., X. Zhang and P. P. Roux (2012). "Regulation and function of the RSK family of protein kinases." Biochem J **441**(2): 553-569.

Roskoski, R., Jr. (2012). "ERK1/2 MAP kinases: structure, function, and regulation." Pharmacol Res **66**(2): 105-143.
Roux, P. P., S. A. Richards and J. Blenis (2003). "Phosphorylation of p90 ribosomal S6 kinase (RSK) regulates extracellular signal-regulated kinase docking and RSK activity." Mol Cell Biol **23**(14): 4796-4804.
Steichen, J. M., G. H. Iyer, S. Li, S. A. Saldanha, M. S. Deal, V. L. Woods, Jr. and S. S. Taylor (2010). "Global consequences of activation loop phosphorylation on protein kinase A." J Biol Chem **285**(6): 3825-3832.
Steichen, J. M., M. Kuchinskas, M. M. Keshwani, J. Yang, J. A. Adams and S. S. Taylor (2012). "Structural basis for the regulation of protein kinase A by activation loop phosphorylation." J Biol Chem **287**(18): 14672-14680.
Taylor, S. S., R. Ilouz, P. Zhang and A. P. Kornev (2012). "Assembly of allosteric macromolecular switches: lessons from PKA." Nat Rev Mol Cell Biol **13**(10): 646-658.
Taylor, S. S., M. M. Keshwani, J. M. Steichen and A. P. Kornev (2012). "Evolution of the eukaryotic protein kinases as dynamic molecular switches." Philosophical Transactions of the Royal Society B: Biological Sciences **367**(1602): 2517-2528.
Taylor, S. S. and A. P. Kornev (2011). "Protein kinases: evolution of dynamic regulatory proteins." Trends Biochem Sci **36**(2): 65-77.
Taylor, S. S., A. S. Shaw, N. Kannan and A. P. Kornev (2015). "Integration of signaling in the kinome: Architecture and regulation of the αC Helix." Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics **1854**(10, Part B): 1567-1574.
Taylor, S. S., K. Søberg, E. Kobori, J. Wu, S. Pautz, F. W. Herberg and B. S. Skålhegg (2021). "The tails of PKA." Molecular Pharmacology: MOLPHARM-MR-2021-000315.
Taylor, S. S., M. Wallbott, E. M. F. Machal, K. Søberg, F. Ahmed, J. Bruystens, L. Vu, B. Baker, J. Wu, F. Raimondi, E. M. Ongeri, F. W. Herberg and B. S. Skålhegg (2021). "PKA Cβ: a forgotten catalytic subunit of cAMP-dependent protein kinase opens new windows for PKA signaling and disease pathologies." Biochemical Journal **478**(11): 2101-2119.
Tompa, P., N. E. Davey, T. J. Gibson and M. M. Babu (2014). "A million peptide motifs for the molecular biologist." Mol Cell **55**(2): 161-169.
Trivier, E., D. De Cesare, S. Jacquot, S. Pannetier, E. Zackai, I. Young, J. L. Mandel, P. Sassone-Corsi and A. Hanauer (1996). "Mutations in the kinase Rsk-2 associated with Coffin-Lowry syndrome." Nature **384**(6609): 567-570.
Tsigelny, I., J. P. Greenberg, S. Cox, W. L. Nichols, S. S. Taylor and L. F. Ten Eyck (1999). "600 ps molecular dynamics reveals stable substructures and flexible hinge points in cAMP dependent protein kinase." Biopolymers **50**(5): 513-524.
Utepbergenov, D., U. Derewenda, N. Olekhnovich, G. Szukalska, B. Banerjee, M. K. Hilinski, D. A. Lannigan, P. T. Stukenberg and Z. S. Derewenda (2012). "Insights into the inhibition of the p90 ribosomal S6 kinase (RSK) by the flavonol glycoside SL0101 from the 1.5 Å crystal structure of the N-terminal domain of RSK2 with bound inhibitor." Biochemistry **51**(33): 6499-6510.

Utepbergenov, D., P. M. Hennig, U. Derewenda, M. V. Artamonov, A. V. Somlyo and Z. S. Derewenda (2016). "Bacterial Expression, Purification and In Vitro Phosphorylation of Full-Length Ribosomal S6 Kinase 2 (RSK2)." PLOS ONE **11**(10): e0164343.

Weng, J.-H., P. C. Aoto, R. Lorenz, J. Wu, S. H. Schmidt, J. T. Manschwetus, P. Kaila-Sharma, S. Silletti, S. Mathea, D. Chatterjee, S. Knapp, F. W. Herberg and S. S. Taylor (2022). "LRRK2 dynamics analysis identifies allosteric control of the crosstalk between its catalytic domains." PLOS Biology **20**(2): e3001427.

White, R. R., Y. G. Kwon, M. Taing, D. S. Lawrence and A. M. Edelman (1998). "Definition of optimal substrate recognition motifs of Ca2+-calmodulin-dependent protein kinases IV and II reveals shared and distinctive features." J Biol Chem **273**(6): 3166-3172.

Wright, P. E. and H. J. Dyson (2015). "Intrinsically disordered proteins in cellular signalling and regulation." Nat Rev Mol Cell Biol **16**(1): 18-29.

Yang, J., S. M. Garrod, M. S. Deal, G. S. Anand, V. L. Woods, Jr. and S. Taylor (2005). "Allosteric network of cAMP-dependent protein kinase revealed by mutation of Tyr204 in the P+1 loop." J Mol Biol **346**(1): 191-201.

Yang, J., J. Wu, J. M. Steichen, A. P. Kornev, M. S. Deal, S. Li, B. Sankaran, V. L. Woods, Jr. and S. S. Taylor (2012). "A conserved Glu-Arg salt bridge connects coevolved motifs that define the eukaryotic protein kinase fold." J Mol Biol **415**(4): 666-679.

Zhao, Y., C. Bjørbaek, S. Weremowicz, C. C. Morton and D. E. Moller (1995). "RSK3 encodes a novel pp90rsk isoform with a unique N-terminal sequence: growth factor-stimulated kinase function and nuclear translocation." Mol Cell Biol **15**(8): 4353-4363.

Chapter IV

Identification of the RSK2 ORF45 binding site

A. Introduction

Kaposi's sarcoma-associated herpesvirus (KSHV) is a cancer causing virus that alter host cellular signaling pathways causing prolonged cellular life and replication (Davey, Travé et al. 2011). ORF45 is a non-catalytic viral product of (KSHV), and causes a global upregulation of MAPK signaling by forming a stable complex with p90 ribosomal S6 kinase (RSK1/2) and ERK1/2. ORF45 has been previously shown to simultaneously bind both RSK and ERK stabilizing the complex and potentially protecting the two kinases from phosphatases (Kuang, Tang et al. 2008, Kuang, Wu et al. 2009). A short N-terminal fragment of ORF45 (~10-80 amino acids) is sufficient for RSK and ERK2 binding. This ORF45 fragment engages the so-called F-site (FRS) of ERK2 via an N-terminal FXFP motif. ORF45 also binds RSK through a DDVF sequence (hereafter referred to as VF motif), and mutation of F66 to alanine abolishes the ability of ORF45 to enhance sustained RSK/ERK activation and diminishes the ability of the virus to generate infectious progeny (Fu, Kuang et al. 2015, Li, Huang et al. 2019, Alexa, Sok et al. 2022).

RSK and ERK are downstream effectors of the well characterized Ras/MAPK signaling pathway and the two kinases form a substrate:activator kinase complex (Anjum and Blenis 2008, Garai, Zeke et al. 2012, Romeo, Zhang et al. 2012, Gógl, Kornev et al. 2019, Alexa, Sok et al. 2022). RSK contains two kinase domains separated by a flexible linker, an N-terminal kinase (NTK) and a C-terminal kinase (CTK). ERK binds to a C-terminal D motif in RSK before phosphorylating the CTK. The active CTK then phosphorylates the flexible linker at the hydrophobic motif (HM)

(Pearce, Komander et al. 2010) allowing binding phosphoinositide dependent kinase 1 (PDK1) and phosphorylation of the NTK activation loop. An active NTK is then able to phosphorylate various downstream substrates.

To better understand how ORF45 binds and stabilizes a complex with both RSK and ERK we utilized hydrogen deuterium exchange mass spectrometry (HDXMS) to determine the ORF45 binding site in the RSK2-ERK2 complex. By measuring changes in deuterium uptake in the presence and absence of a binding partner, HDXMS can identify direct protein protein interfaces and more indirect allosteric changes associated with partner protein binding (Englander, Del Mar et al. 2003, Chalmers, Busby et al. 2011). We identified that ORF45 strongly engages the C-lobe of the NTK, and indirectly enhances the CTK-ERK2 interface. A conserved VF motif in ORF45 binds a deep hydrophobic pocket in the C-lobe of the NTK, and we demonstrate that this pocket is also utilized by other VF motifs in unstructured regions of other viral proteins, endogenous proteins, RSK substrates, and phosphatases. Thus, this newly identified NTK surface may provide RSK with an entirely new class of substrates and regulation that are independent from classical RSK substrates.

B. RESULTS

1. HDXMS to determine RSK2 ORF45 binding site

To determine how ORF45 binds to the RSK2-ERK2 complex we utilized HDXMS to identify regions of altered solvent accessibility in the complex (Fig. 4.2). Comparison of deuterium uptake between the RSK2-ERK2-ORF45 and RSK2-ERK2 complexes was unable to identify any observable ORF45 ERK2 protection. The affinity of FXFP motifs/F-site based interactions is strongly dependent on the phosphorylation state of the activation loop (Lee, Hoofnagle et al. 2004), the affinity of ORF45 for ERK2 and (2p)ERK2 is ~ 20uM and 1uM, respectively. Our analysis utilized non-phosphorylated ERK2, and presumably is the cause for the lack of observed solvent protection on ERK2.

However, we were able to identify a handful of regions of significant protection on RSK2 in the C-lobe of the NTK and in the CTK at the CTK/ERK2 Interface 2 (Figs. 4.3A). The greatest and most extensive solvent protection is observed in the C-lobe of the NTK at the αF-αG loop, αG, αH, and αH-αI loop. Additional protection is observed at the αD/αD-αE loop of the NTK. And, at the CTK-ERK2 Interface 2, the CTK activation loop through the APE/αF helix, is protected upon ORF45 binding as well (Alexa, Sok et al. 2022). To validate the observed solvent protection, we generated an array of NTK 18mer peptides that span the observed sites of protection, and tested their ability to bind to an ORF45 fragment. In agreement with the HDXMS, ORF45 directly binds peptides of the αG, αH, αH-αI loop, and αD/αD-αE loop (Fig. 4.3B). The N-terminal ORF45 FXFP motif also

interacts with the NTK αD/αD-αE loop (Alexa, Sok et al. 2022). Ordinarily, the ORF45 FXFP interacts with the F-site of ERK2 F-site binder when ERK2 is doubly phosphorylated, thus the crystallographic interaction and HDX protection observed at the hinge/αD helix may be artifactual due to the lack of ERK2 in the crystal structure and use of the non-phosphorylated ERK2 in HDX analysis. The observed HDXMS ORF45 interface with the NTK, and indirect protection of the CTK activation loop and APE through αF helix was further validated by a RSK2-ORF45 crystal structure, SPR affinity measurements, and SAXS analysis (Alexa, Sok et al. 2022).

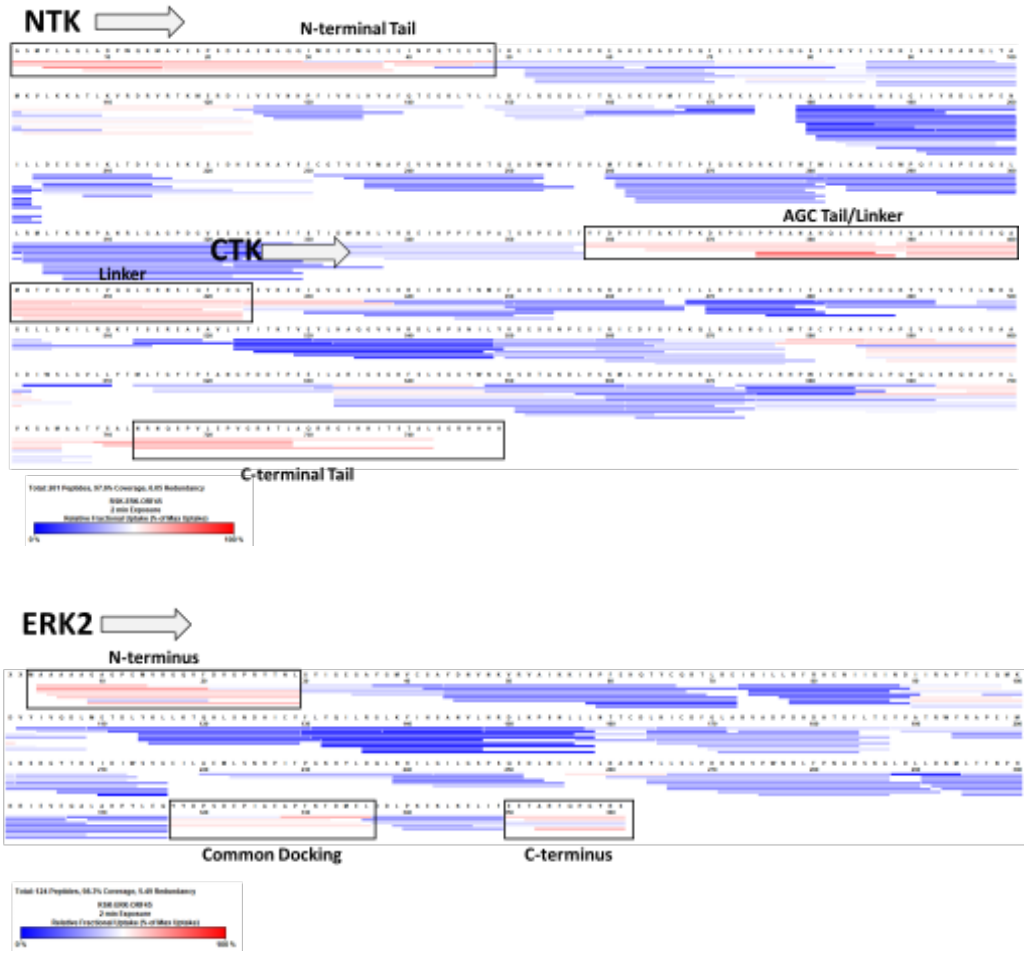


Figure 4.1. Coverage maps of RSK2 and ERK2 in the RSK2-ERK2-ORF45 complex.

Top and bottom, are the coverage maps of RSK2 and ERK2 in the RSK2-ERK2-ORF45 complex, respectively, at the 2-minute timepoint. Blue sticks representing solvent shielded peptides and red sticks representing solvent exposed peptides.

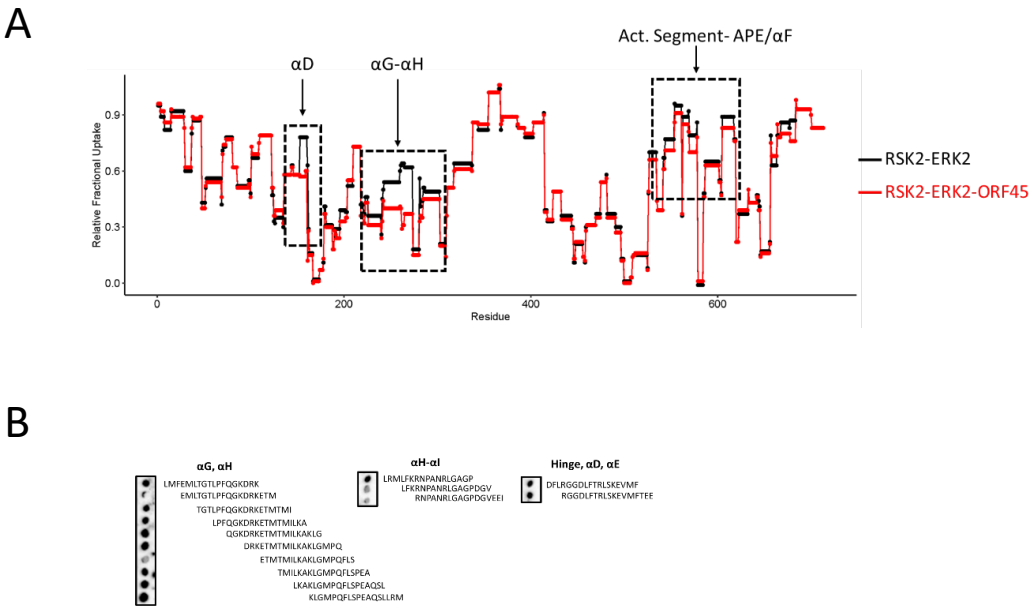


Figure 4.2. ORF45 binds to a hydrophobic C-lobe surface in the NTK.

(A) Relative fractional uptake plot of RSK2 in the RSK2-ERK2 complex (black) or RSK2-ERK2-ORF45 complex (red).

(B) Peptide arrays of NTK peptides that correspond to HDXMS protected regions upon ORF45 binding, tested for ORF45 (16-76) binding.

The identification of this NTK C-lobe αG-αH hydrophobic pocket raised the hypothesis that perhaps this site is utilized by other unstructured VF-motifs found in other endogenous or pathogenic proteins. Multiple VF-motifs were identified in other pathogenic proteins, phosphatases, and other endogenous proteins, some of which are known RSK interacting partners and substrates (Krystkowiak and Davey 2017, Alexa, Sok et al. 2022) . To determine if these VF-motif containing proteins do indeed interact with RSK2, we generated an array of VF motif containing peptides and tested the ability of the peptides to bind RSK2 (Fig. 4.5A). Numerous VF motif-containing peptides did indeed interact with RSK2. And binding of these novel VF motifs to RSK2 is dependent on the phenylalanine (Fig. 4.5B). In all cases, the mutation of phenylalanine to alanine abolished RSK2 binding. A large number of these VF motif containing binding partners also feature a RSK substrate consensus sequence nearby. VF motif containing substrates were less efficiently phosphorylated in the presence of ORF45, and classical RSK substrates that lack VF motifs were unaffected by addition of ORF45. Therefore, this αG-αH hydrophobic pocket provides an additional binding site that VF containing motifs can utilize for specific phosphorylation (Alexa, Sok et al. 2022).

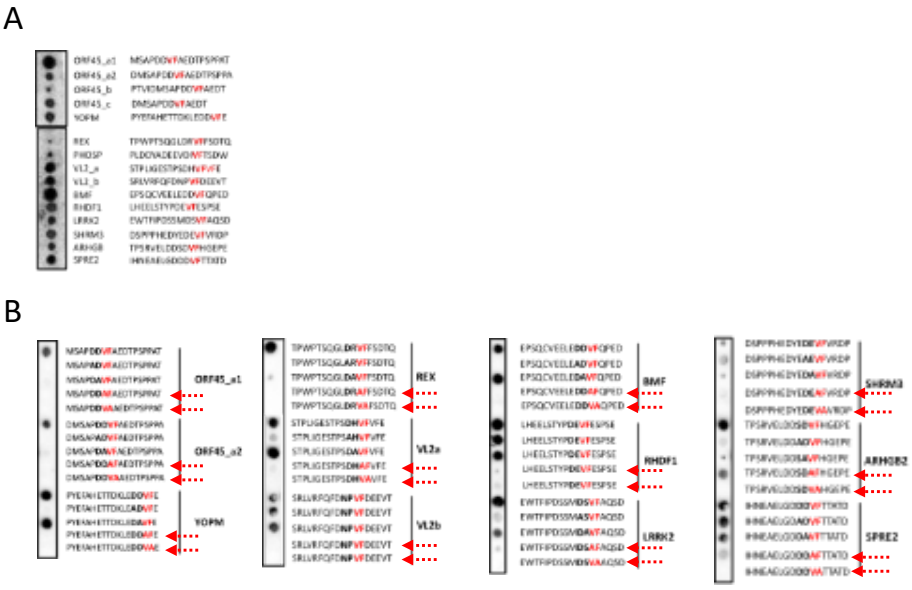


Figure 4.3. NTK C-lobe hydrophobic surface is utilized by other endogenous proteins containing VF motifs in unstructured regions.
(A) Peptide arrays of various VF motif containing peptides tested for binding to FL RSK2
(B) Ala scan peptide arrays of various VF motif RSK2 binders. Peptides were synthesized for alanine at each of a XXVF motif site and tested for RSK2 binding.

C. Discussion

KSHV alters host cellular signaling networks for its own viral replication and progeny production in part by altering protein:protein interactions (Davey, Travé et al. 2011). Here, with HDXMS we identified a novel hydrophobic pocket in the NTK C-lobe utilized by the non-catalytic viral protein ORF45 to bind and stabilize the RSK2-ERK2 complex. This novel NTK C-lobe pocket is used by endogenous proteins with VF motifs in unstructured regions to enable interaction with multiple novel endogenous proteins of a new class of RSK substrates (Alexa, Sok et al. 2022). Many of these confirmed VF motif binders are known or predicted RSK substrates (Lei and Davis 2003, Saha, Carriere et al. 2012, Düsterhöft, Babendreyer et al. 2019) and contain a putative NTK consensus sequence K/RXXS/T (Hauge and Frödin 2006, Galan, Geraghty et al. 2014). These VF motif sequences may enhance the efficiency of phosphorylation of a nearby target sequence, similarly to MAPKs docking motifs enhanced substrate phosphorylation (Tanoue, Adachi et al. 2000, Kirsch, Zeke et al. 2020).

The NTK is a promiscuous kinase as it shares the target consensus sequence is with other AGC kinases like AKT, and p70 S6 kinase. Recently, it was demonstrated that a RSK C-terminal PDZ domain binding motif can efficiently bind certain PDZ domain containing RSK substrates. Thus, the RSK PDZ domain binding motif may serve as a docking motif to enable a RSK specific group of substrates (Gógl, Biri-Kovács et al. 2018, Gógl, Biri-Kovács et al. 2019). Similarly, these VF motifs adjacent to a RSK

target substrate sequence may provide RSK a highly group of substrates with greater specificity than other classical RSK substrates.

D. References

Alexa, A., P. Sok, F. Gross, K. Albert, E. Kobori, Á. L. Póti, G. Gógl, I. Bento, E. Kuang, S. S. Taylor, F. Zhu, A. Ciliberto and A. Reményi (2022). "A non-catalytic herpesviral protein reconfigures ERK-RSK signaling by targeting kinase docking systems in the host." Nature Communications **13**(1): 472.

Anjum, R. and J. Blenis (2008). "The RSK family of kinases: emerging roles in cellular signalling." Nature Reviews Molecular Cell Biology **9**(10): 747-758.

Chalmers, M. J., S. A. Busby, B. D. Pascal, G. M. West and P. R. Griffin (2011). "Differential hydrogen/deuterium exchange mass spectrometry analysis of protein–ligand interactions." Expert Review of Proteomics **8**(1): 43-59.

Davey, N. E., G. Travé and T. J. Gibson (2011). "How viruses hijack cell regulation." Trends Biochem Sci **36**(3): 159-169.

Düsterhöft, S., A. Babendreyer, A. A. Giese, C. Flasshove and A. Ludwig (2019). "Status update on iRhom and ADAM17: It's still complicated." Biochim Biophys Acta Mol Cell Res **1866**(10): 1567-1583.

Englander, J. J., C. Del Mar, W. Li, S. W. Englander, J. S. Kim, D. D. Stranz, Y. Hamuro and V. L. Woods (2003). "Protein structure change studied by hydrogen-deuterium exchange, functional labeling, and mass spectrometry." Proceedings of the National Academy of Sciences **100**(12): 7057-7062.

Fu, B., E. Kuang, W. Li, D. Avey, X. Li, Z. Turpin, A. Valdes, K. Brulois, J. Myoung and F. Zhu (2015). "Activation of p90 ribosomal S6 kinases by ORF45 of Kaposi's sarcoma-associated herpesvirus is critical for optimal production of infectious viruses." J Virol **89**(1): 195-207.

Galan, J. A., K. M. Geraghty, G. Lavoie, E. Kanshin, J. Tcherkezian, V. Calabrese, G. R. Jeschke, B. E. Turk, B. A. Ballif, J. Blenis, P. Thibault and P. P. Roux (2014). "Phosphoproteomic analysis identifies the tumor suppressor PDCD4 as a RSK substrate negatively regulated by 14-3-3." Proc Natl Acad Sci U S A **111**(29): E2918-2927.

Garai, Á., A. Zeke, G. Gógl, I. Törő, F. Fördős, H. Blankenburg, T. Bárkai, J. Varga, A. Alexa, D. Emig, M. Albrecht and A. Reményi (2012). "Specificity of Linear Motifs That Bind to a Common Mitogen-Activated Protein Kinase Docking Groove." Science Signaling **5**(245): ra74-ra74.

Gógl, G., B. Biri-Kovács, F. Durbesson, P. Jane, Y. Nomine, C. Kostmann, V. Bilics, M. Simon, A. Reményi, R. Vincentelli, G. Trave and L. Nyitray (2019). "Rewiring of RSK-PDZ Interactome by Linear Motif Phosphorylation." J Mol Biol **431**(6): 1234-1249.

Gógl, G., B. Biri-Kovács, L. Póti Á, H. Vadászi, B. Szeder, A. Bodor, G. Schlosser, A. Ács, L. Turiák, L. Buday, A. Alexa, L. Nyitray and A. Reményi (2018). "Dynamic control of RSK complexes by phosphoswitch-based regulation." Febs j **285**(1): 46-71.

Gógl, G., A. P. Kornev, A. Reményi and S. S. Taylor (2019). "Disordered Protein Kinase Regions in Regulation of Kinase Domain Cores." Trends Biochem Sci **44**(4): 300-311.

Hauge, C. and M. Frödin (2006). "RSK and MSK in MAP kinase signalling." J Cell Sci **119**(Pt 15): 3021-3023.

Kirsch, K., A. Zeke, O. Tőke, P. Sok, A. Sethi, A. Sebő, G. S. Kumar, P. Egri, Á. L. Póti, P. Gooley, W. Peti, I. Bento, A. Alexa and A. Reményi (2020). "Co-regulation of the transcription controlling ATF2 phosphoswitch by JNK and p38." Nature Communications **11**(1): 5769.

Krystkowiak, I. and N. E. Davey (2017). "SLiMSearch: a framework for proteome-wide discovery and annotation of functional modules in intrinsically disordered regions." Nucleic Acids Res **45**(W1): W464-w469.

Kuang, E., Q. Tang, G. G. Maul and F. Zhu (2008). "Activation of p90 ribosomal S6 kinase by ORF45 of Kaposi's sarcoma-associated herpesvirus and its role in viral lytic replication." J Virol **82**(4): 1838-1850.

Kuang, E., F. Wu and F. Zhu (2009). "Mechanism of sustained activation of ribosomal S6 kinase (RSK) and ERK by kaposi sarcoma-associated herpesvirus ORF45: multiprotein complexes retain active phosphorylated ERK AND RSK and protect them from dephosphorylation." J Biol Chem **284**(20): 13958-13968.

Lee, T., A. N. Hoofnagle, Y. Kabuyama, J. Stroud, X. Min, E. J. Goldsmith, L. Chen, K. A. Resing and N. G. Ahn (2004). "Docking motif interactions in MAP kinases revealed by hydrogen exchange mass spectrometry." Mol Cell **14**(1): 43-55.

Lei, K. and R. J. Davis (2003). "JNK phosphorylation of Bim-related members of the Bcl2 family induces Bax-dependent apoptosis." Proc Natl Acad Sci U S A **100**(5): 2432-2437.

Li, X., L. Huang, Y. Xiao, X. Yao, X. Long, F. Zhu and E. Kuang (2019). "Development of an ORF45-Derived Peptide To Inhibit the Sustained RSK Activation and Lytic Replication of Kaposi's Sarcoma-Associated Herpesvirus." J Virol **93**(10).

Pearce, L. R., D. Komander and D. R. Alessi (2010). "The nuts and bolts of AGC protein kinases." Nat Rev Mol Cell Biol **11**(1): 9-22.

Romeo, Y., X. Zhang and P. P. Roux (2012). "Regulation and function of the RSK family of protein kinases." Biochem J **441**(2): 553-569.

Saha, M., A. Carriere, M. Cheerathodi, X. Zhang, G. Lavoie, J. Rush, P. P. Roux and B. A. Ballif (2012). "RSK phosphorylates SOS1 creating 14-3-3-docking sites and negatively regulating MAPK activation." Biochem J **447**(1): 159-166.

Tanoue, T., M. Adachi, T. Moriguchi and E. Nishida (2000). "A conserved docking motif in MAP kinases common to substrates, activators and regulators." Nat Cell Biol **2**(2): 110-116.

Chapter V

CryoEM of RSK2:ERK2:ORF45, RSK2, and ERK2

A. Introduction

Protein kinases are enzymes that are dynamically regulated between active and inactive conformations (Taylor, Keshwani et al. 2012, Kornev and Taylor 2015). In most cases, kinase activation is achieved following phosphorylation by an upstream activating kinase. These activator-substrate kinase complexes are mediated by both domain:domain interactions, in addition to interactions between small linear motifs (SLiMs) in intrinsically disordered regions (IDRs) and kinase cores (Gógl, Kornev et al. 2019). IDRs typically flank kinase domains, and SLiMs are short flexible sequences that mediate medium affinity interactions, allowing multiple binding partners, depending on the cellular context (Davey, Cyert et al. 2015, Wright and Dyson 2015). Moreover, these are usually conserved motifs that are critical for kinase regulation, and can be altered by post translational modifications that adjust their binding properties. SLiM/IDR interactions are dynamic and typically short lived, making the study of activator-substrate kinase complexes challenging. A meta-analysis of various disease mutations found that mutations tend to perturb protein:protein interfaces rather than alter protein folding or stability (Cheng, Zhao et al. 2021). Thus, the study of these dynamic and short-lived kinase complexes and their interfaces is critical. A greater structural understanding of kinase complexes, in particular, how kinases interact with their flanking IDRs, is needed to better understand how one kinase phosphorylates a substrate kinase.

The MAP kinase activated protein kinase (MAPKAPK) p90 RSK is a unique protein because it contains two kinase domains in a single polypeptide, an N-terminal

kinase (NTK) and a C-terminal kinase (CTK) (Jones, Erikson et al. 1988, Fisher and Blenis 1996, Anjum and Blenis 2008). RSK is also a rare example of a kinase that is stably bound to its activating kinase ERK1/2, making it a good model system to better understand how one kinase activates another (3). IDRs with conserved SLiMs flank each kinase, and an interkinase linker connects the NTK and CTK (Fig 1A). Part of the linker contains the conserved AGC tail (Pearce, Komander et al. 2010), the hallmark of AGC kinases (Kannan, Haste et al. 2007), and conserved SLiMs like the hydrophobic motif (HM), enable activation of the NTK (Dalby, Morrice et al. 1998, Biondi, Cheung et al. 2000, Mora, Komander et al. 2004). An active NTK phosphorylates downstream substrates, while the CTK intramolecularly phosphorylates the HM to promote NTK activation (Bjørbaek, Zhao et al. 1995, Fisher and Blenis 1996). Thus, the linker mediates communication between the two kinases. A disordered C-terminal tail follows the CTK, and contains an autoinhibitory helix that suppresses basal CTK activity (Temmerman, Simon et al. 2013), and a conserved revD motif that enables binding to the activating kinase ERK1/2 (Roux, Richards et al. 2003, Garai, Zeke et al. 2012). The perturbation of the C-terminal tail in a variety of diseases demonstrates the importance of this region. Mutations to the revD motif are found in a variety of cancers (Tate, Bamford et al. 2019), RSK2 revD sites cause Coffin Lowry Syndrome (CLS) (Trivier, De Cesare et al. 1996), and in melanoma, the melanoma marker S100B binds the entire C-terminal tail affecting CTK basal activity (Hartman, Vitolo et al. 2014, Gógl, Alexa et al. 2016).

Despite the challenges in obtaining structures of substrate-activator kinase complexes, due to their typically transient nature, some kinase dimeric pairs have been

trapped crystallographically. These provide insight into how kinases position themselves to achieve substrate kinase phosphorylation (Gógl, Kornev et al. 2019). KSR/MEK1 heterodimers and PAK1 homodimers, assemble in a parallel orientation (their N-lobes are in the same direction) mediated by interactions between their N-lobes, αG helices, and activation segments, to position each kinase active site toward the other (Brennan, Dar et al. 2011, Wang, Wu et al. 2011). Homodimers of Chk2, Ste20-like kinase (SLK), serine/threonine kinase 10 (STK10), and death associated protein kinase 3 (DAPK3), among others, display activation segment exchange in which each kinase swaps activation segments with its pair to auto-phosphorylate its own activation loop and promote its active conformation (Oliver, Knapp et al. 2007, Marcotte, Rushe et al. 2017). The heterodimer of the CTK of p90 RSK and ERK2, adopts an antiparallel orientation with the N-lobe of one kinase facing the C-lobe of the other (Garai, Zeke et al. 2012, Alexa, Gógl et al. 2015). The revD motif of the CTK binds the D-recognition site (DRS) of ERK2 with high affinity and is the primary interface. A secondary interface between a CTK extended αF helix capped by the APE motif and the glycine rich loop of ERK2, properly positions the activation loop near the active site of ERK2, and provides a mechanism for CTK phosphorylation. Recently, the structure of a p38-MK2 complex, two kinases similar to ERK2 and the CTK, respectively, was solved and the orientation of these kinases depends on the activation state of p38 (Sok, Gógl et al. 2020). Non-phosphorylated p38 adopts a parallel oriented complex with MK2 that is incompatible with MK2 activation loop phosphorylation. However, solution-based method like SAXS and NMR indicate that active, phosphorylated p38 adopts an antiparallel complex with

MK2 and is compatible with MK2 activation loop phosphorylation. This suggests that an antiparallel kinase complex may be a more general mechanism for how kinases phosphorylate other kinases.

Here we strive to structurally characterize the FL RSK2-ERK2 complex to better understand how ERK2 phosphorylates the CTK, and more generally how kinase complexes assemble to phosphorylate downstream kinase substrates. Our goal is to determine how all three kinases are oriented relative to one another, and further explore how the NTK, CTK, and ERK2 interact with the Linker and C-terminal tail. Currently, there are no structures of FL RSK or of the isolated NTK/CTK that include the Linker. Thus, characterization of FL RSK or its RSK2-ERK2 complex, can provide valuable structural information on this critical IDR to hypothesize if the linker is positioned in a way to facilitate downstream phosphorylation steps required for RSK activation. Moreover, analysis of the FL RSK2-ERK2 complex may also identify additional interactions between the NTK, CTK, and ERK2.

To aid in our structural studies, we also utilize a peptide of ORF45, the viral protein that simultaneously binds both RSK and ERK2, stabilizing the RSK-ERK complex (Kuang, Tang et al. 2008, Kuang, Wu et al. 2009) (Fig 1A). Recently, an ORF45 peptide was found to bind the NTK with high affinity, and ORF45 peptide binding to the RSK2-ERK2 complex enhanced solvent protection of the CTK APE-ERK2 glycine rich loop, thus stabilizing this interface. Moreover, SAXS has demonstrated that FL RSK2 has an extended shape, the RSK2-ERK2 complex is more compact, and a ternary RSK2-ERK2-ORF45 complex has the most compact shape

(Alexa, Sok et al. 2022). Thus, ORF45 is an intriguing tool that we use to aid in our structural studies of the RSK2-ERK2 complex.

Here, we use cryoEM to analyze the RSK2-ERK2-ORF45 complex. Although our analysis is limited to low resolution, we are able to determine the overall size and shape of ERK2, FL RSK2, and the RSK2-ERK2-ORF45 complex. Class averages of ERK2 match corresponding theoretical projections of the solved ERK2 isolated crystal structure. In agreement with SAXS data, the RSK2-ERK2-ORF45 complex is larger, and more globular in shape while FL RSK2 has a more extended shape. SAXS or AlphaFold (AF) (Jumper, Evans et al. 2021, Varadi, Anyango et al. 2021, Mirdita, Schütze et al. 2022) generated models of RSK2 or the RSK2-ERK2-ORF45 complex, reasonably fit our low-resolution reconstructions. The low resolution of the reconstructions limits our interpretation of these structures, higher resolution is required to determine the orientation of all of the kinases, potential interactions among the NTK and the CTK or ERK2, and how the linker and C-terminal tail interact with all three kinases. Thus, higher resolution is needed to advance our understanding of substrate:activator kinase complex phosphorylation. Representative 2D class averages of the RSK2-ERK2-ORF45 complex have strong features and display secondary structure, therefore we believe our resolution is limited by a smaller complex population size, and collecting more data to significantly increase our complex particle population can significantly improve the resolution of the RSK2-ERK2-ORF45 complex.

B. RESULTS

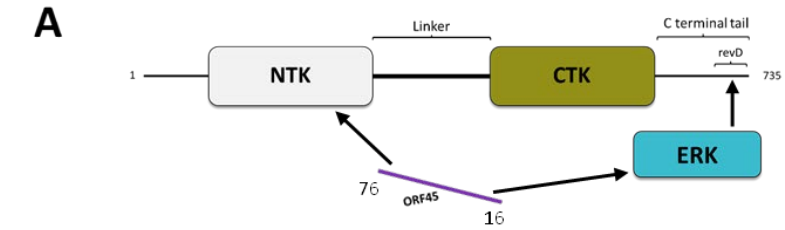
1. RSK2:ERK2:ORF45 negative stain EM

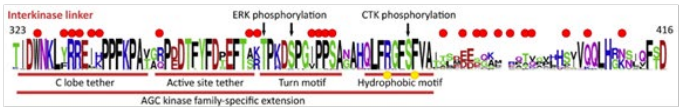
To obtain structural information on the RSK2-ERK2 complex, and to gain a better understanding on the orientation of all three kinase domains relative to one another, we initially utilized negative stain electron microscopy to identify an optimal sample for cryoEM. First, we compared the RSK2-ERK2 binary and RSK2-ERK2-ORF45 ternary complexes (Figs 5.1B,C). Representative micrographs of the complex are monodisperse, and corresponding 2D class averages are rather amorphous without any recognizable features. Next, we tested the effect of a chemical crosslinker to further stabilize the complex. Crosslinking enhances RSK2-ERK2-ORF45 complex class averages such that they are more featureful, have sharper edges, and are more indicative of a three-domain complex (Fig 5.1D,E). Recent hydrogen deuterium exchange mass spectrometry of the RSK2-ERK2 complex reveals the N-lobe of the NTK is solvent exposed and dynamic, leading us to believe that further stabilization of the NTK may aid as well (Figs 3.2, 3.3). Addition of an NTK inhibitor, SL0101, further improves the class averages of the crosslinked RSK2-ERK2-ORF complex. This flavonoid compound binds to the active site of the NTK and stabilizes a rather extreme inhibited conformation, and thermally stabilizes FL RSK2 (Utepbergenov, Derewenda et al. 2012, Utepbergenov, Hennig et al. 2016). Class averages of the crosslinked complex with SL0101 are more representative of a three-kinase domain complex, are more featureful, and display sharper particle edges and cavities. Based

on the negative staining, the crosslinked RSK-ERK2-ORF45 complex is the most promising sample for cryoEM analysis.

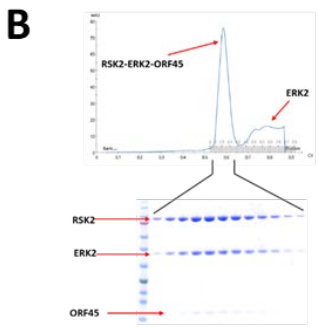
Figure 5.1. Negative stain analysis of the RSK2:ERK2:ORF45 complexes.

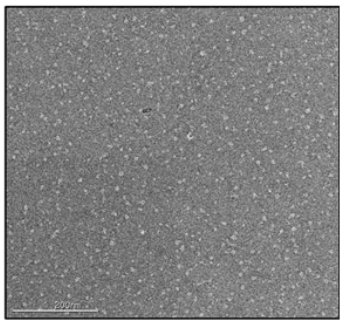
(A) Domain organization of RSK2 and schematic of the RSK2-ERK2-ORF45 complex. Arrows denote binding events, ERK2 binds to the revD motif in the C-terminal tail, and ORF45 binds to the NTK and ERK2. Sequence logos of the RSK Linker and C-terminal tail, with conserved motifs and phosphorylation sites denoted. The sequence logos are adapted from (Gogl, et al., TIBS, 2019). Red dots indicate the positions of cancer mutations (COSMIC database) and yellow dots indicate the positions of Coffin Lowry Syndrome mutations.
(B) Representative size exclusion chromatogram of the RSK2-ERK2-ORF45 complex and SDS-PAGE of the corresponding fractions.
(C) Representative negative stain micrograph of the RSK2-ERK2-ORF45 complex and representative 2D class averages of RSK2-ERK2 and RSK2-ERK2-ORF45 complexes.
(D) SDS-PAGE of the crosslinking reaction of RSK2-ERK2-ORF45 complex stained with Sypro ruby. Arrows point to bands representing RSK2, ERK2, ORF45 and crosslinked RSK2-ERK2-ORF45 complex.
(E) Representative 2D class averages of crosslinked RSK2-ERK2-ORF45 complex ± SL0101.

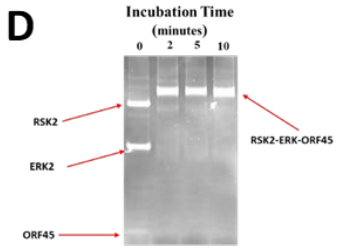

A
Linker
C terminal tail
revD
1
NTK
CTK
735
ERK
76
ORF45
16

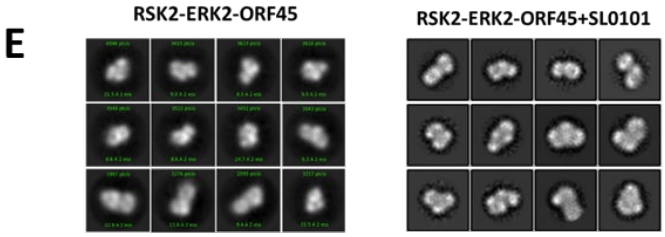

Interkinase linker
323
ERK phosphorylation
CTK phosphorylation
416
C lobe tether
Active site tether
Turn motif
Hydrophobic motif
AGC kinase family-specific extension


C terminal tail
683
Tyr phosphorylation
Autophosphorylation
735
CTK autoinhibition
ERK2 binding
PDZ binding


B
RSK2-ERK2-ORF45
ERK2
RSK2
ERK2
ORF45


C
RSK2-ERK2
RSK2-ERK2-ORF45


D
Incubation Time
(minutes)
0
2
5
10
RSK2
ERK2
ORF45
RSK2-ERK-ORF45


E
RSK2-ERK2-ORF45
RSK2-ERK2-ORF45+SL0101

2. CryoEM analysis of RSK2-ERK2-ORF45

An initial cryoEM analysis of the crosslinked RSK2-ERK2-ORF45 complex under apo conditions was able to generate a low-resolution 3D reconstruction that reveals the overall size and shape of the crosslinked complex (Fig 5.2A,B). Both AF and SAXS predicted models of the RSK2-ERK2-ORF45 complex fit this low-resolution density quite well. 2D class averages display particles with a compact/globular shape with three apparent kinase domains. We collected additional data to try and improve the resolution of the reconstruction (Fig 5.2C). While corresponding 2D class averages were more featureful, they also display signs of artifactual over-fitting, and 3D reconstructions are still low resolution.

Next, we tested the crosslinked RSK2-ERK2-ORF45 complex with the RSK inhibitor, SL0101 (Fig 5.3A). We hypothesized that further stabilization of the NTK may lead to an improvement in resolution. Following multiple rounds of classification, we could identify featureful class averages of free ERK2, FL RSK2, and the RSK2-ERK2-ORF45 complex. This compositional heterogeneity was not appreciated in smaller datasets of apo complex. These class averages have observable secondary structure without the artifactual streaking observed in earlier datasets. Reconstructions were still limited to low resolution, however, density for each apparent kinase domain is more distinguishable, allowing clear placement of three kinase domains, furthermore 2D projections of the SAXS model bear resemblance to the observed 2D class averages (Fig 5.3B,C). However, the low resolution of the reconstruction still prevents knowing the exact placement of the NTK, CTK and ERK2.

To verify that chemical crosslinking does not dramatically alter the size and shape of the RSK2-ERK2-ORF45 complex, we collected data on the non-crosslinked complex with inhibitor SL0101 (Fig 5.4A). We can identify class averages that correspond to free ERK2, RSK2, and RSK2-ERK2-ORF45 complex. The class averages have fewer promising secondary structure features, but these appear to be similar in size and shape to class averages from the crosslinked complex, indicating that crosslinking does not appear to dramatically alter the conformations of the complex.

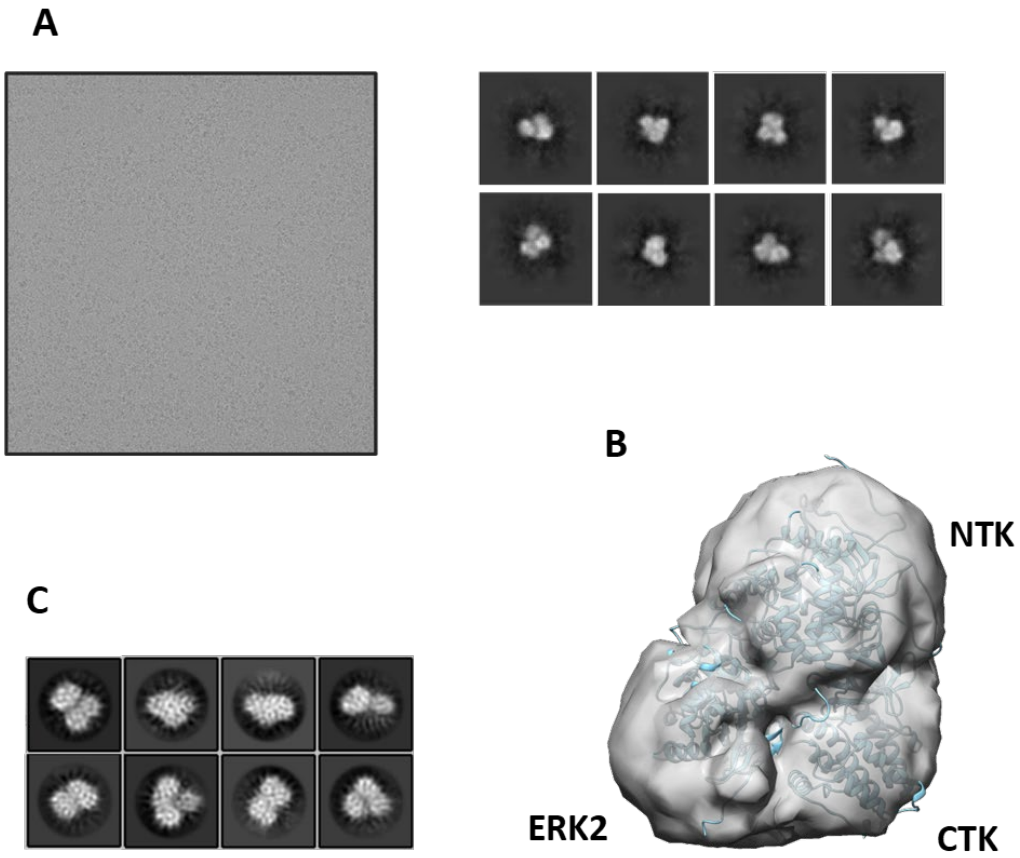


Figure 5.2. Initial cryoEM analysis of crosslinked RSK2-ERK2-ORF45 complex.

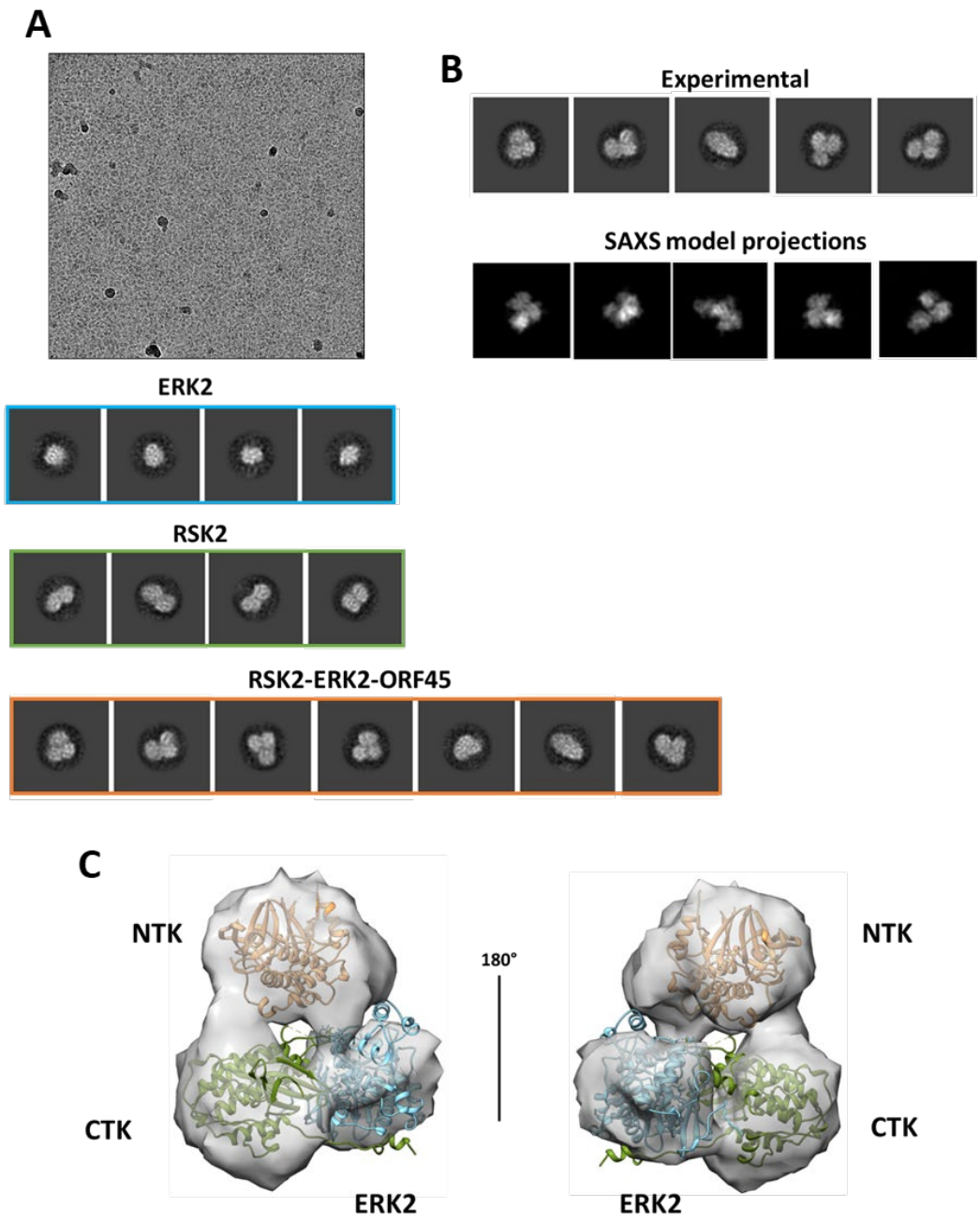
(A) Representative micrograph of initial cryoEM dataset of crosslinked RSK2-ERK2-ORF45 complex imaged at 200 keV, and corresponding 2D class averages.
(B) Corresponding 3D reconstruction, and an AlphaFold generated model of the RSK2-ERK2 complex is fit in the density.
(C) 2D class averages of the crosslinked RSK2-ERK2-ORF45 apo complex imaged at 300 keV. Class averages have visible secondary structure, but radial streaking in the background is indicative of overfitting.

Figure 5.3. CryoEM analysis of crosslinked RSK2-ERK2-ORF45+SL0101.

(A) Representative micrograph of crosslinked RSK2-ERK2-ORF45 complex + SL0101. This dataset includes 2D class averages that can correspond to free ERK2, RSK2, and crosslinked RSK2-ERK2-ORF45. Suspected 2D class averages for each species are displayed.

(B) Selected class averages of crosslinked RSK2-ERK2-ORF45 are compared to theoretical 2D projections of the model of RSK2-ERK2-ORF45 based on SAXS analysis.

(C) Ab initio reconstruction of the crosslinked RSK2-ERK2-ORF45+SL0101 complex, fit with the RSK2 NTK and CTK-ERK2 heterodimer. The SAXS model was initially fit into the density, and then the orientation of the NTK was repositioned to better fit the density. The NTK is colored orange, CTK is green, and ERK2 is cyan.


A
B
Experimental
SAXS model projections
ERK2
RSK2
RSK2-ERK2-ORF45
C
NTK
CTK
ERK2
180°
NTK
CTK
ERK2

3. CryoEM of free RSK2 and free ERK2

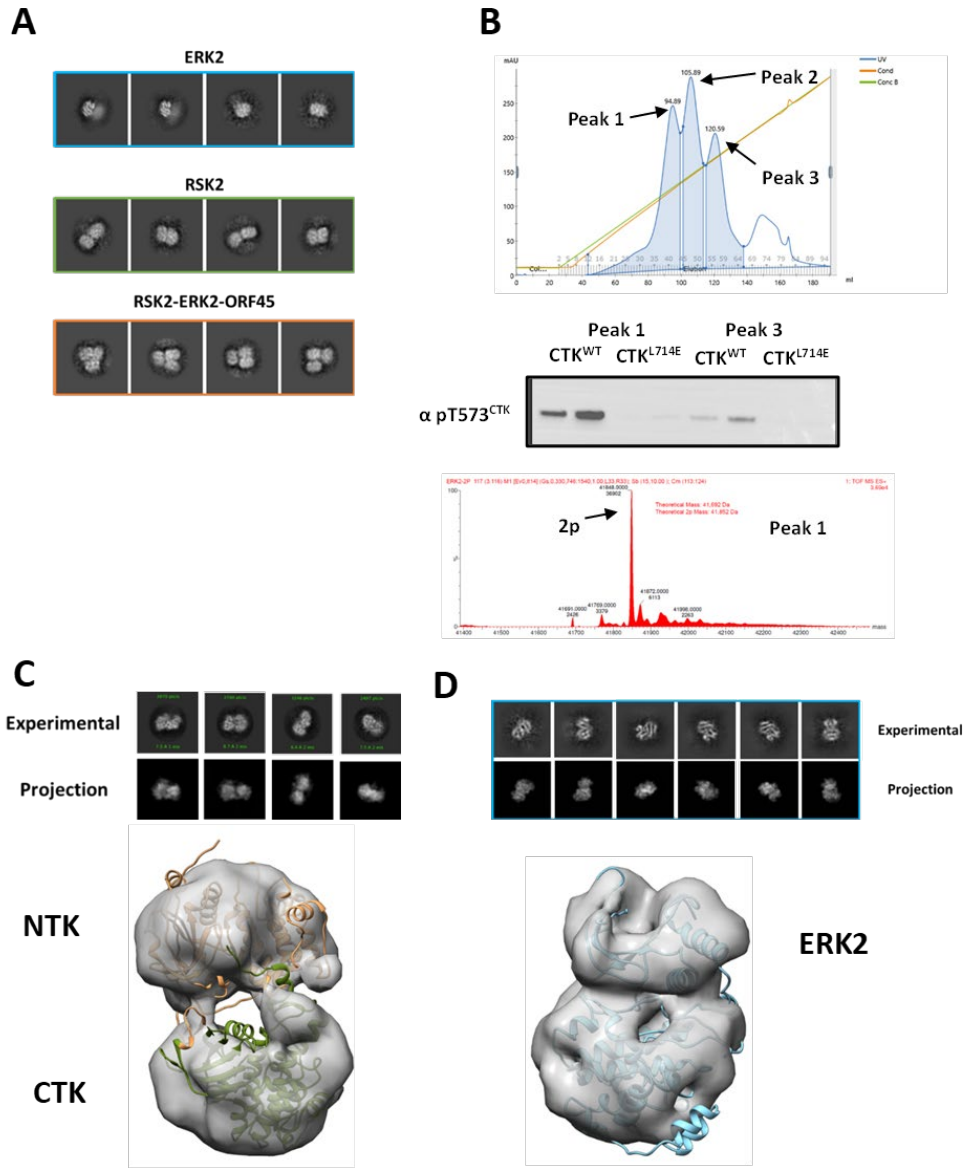
Our biochemical and biophysical characterization of ORF45 and its binding to the RSK2-ERK2 complex revealed that the phosphorylation state of ERK2 significantly strengthens the binding affinity of ERK2 to both ORF45 and RSK2 (Alexa, Sok et al. 2022). The activation loop of ERK2 has two phosphorylation sites (Canagarajah, Khokhlatchev et al. 1997), and phosphorylation of both sites causes solvent protection throughout the C-lobe of the kinase (Hoofnagle, Resing et al. 2001), enhanced affinity for RSK (Fig 2.5), and orders ERK2's F docking site to better accommodate F-site binders like ORF45 (Lee, Hoofnagle et al. 2004). Prior datasets utilized non-phosphorylated, 0p ERK2, so we tested the crosslinked RSK2-2pERK2-ORF45 complex using active, doubly phosphorylated ERK2 (2pERK2), with the hope that this would make a more stable complex and thus achieve higher resolution (Fig 5.4B). Surprisingly, no meaningful improvements in 2D class averages were observed, nor were improvements made to 3D reconstructions. However, subpopulations of free RSK2 and 2pERK2, were still identified. Class averages of RSK2 are similar to 2D projections of an AF predicted model of free RSK2 (truncated to remove the disordered highly solvent exposed N-terminal 45 amino acids) (Fig 5.4C). Moreover, the AF model nicely fits a low-resolution reconstruction of RSK2. And the overall shape of the RSK2 reconstruction is more extended than the complex. Class averages of 2p ERK2 are quite promising, as they display clear secondary structure, and match theoretical 2D projections of the solved ERK2 crystal structure (Fig 5.4D). The 2pERK2 reconstruction was limited to low resolution as well, as the total number of particles used was quite small, but it accommodates an ERK2 crystal structure quite

well. This reconstruction also provides the overall dimensions of a kinase domain (Herzik, Wu et al. 2019), and these were used to verify free RSK2, and the RSK2-ERK2-ORF45 complex contain two and three kinases, respectively (Fig 5.5A,B). Furthermore, comparison of all three reconstructions confirms that, as expected, the complex is the largest, and that all three species have a different overall shape.

As an alternative strategy to cryoEM, we also attempted to crystallize numerous samples of FL RSK2, RSK2-ERK2, and RSK2-ERK2-ORF45. All complexes were tested in the presence of Mg/AMP-PNP, and RSK NTK inhibitors BI-D1870 and SL0101. None of these conditions were successful. All of the published crystal structures of the NTK truncate the disordered NTK N-terminal tail, suggesting that it may be necessary to truncate the N-terminal tail to induce crystallization of FL RSK2 or its complexes. However, truncations of various lengths were unsuccessful as well. These truncated RSK2 constructs are less stable and more prone to aggregation than FL RSK2, suggesting that the N-terminal tail of RSK2 may play a role in kinase stability. Together these suggest that cryoEM may be better equipped to handle the challenges of FL RSK2 and the RSK2-ERK2-ORF45 complex than crystallography.

Figure 5.4. CryoEM analysis of non-crosslinked RSK2-ERK2-ORF45+SL0101 and free 2pERK2, RSK2.

(A) Corresponding 2D class averages of small dataset of non-crosslinked RSK2-ERK2-ORF45 complex. Class averages of free ERK2, RSK2, and RSK2-ERK2-ORF45 complex were identified in this dataset as well, and appear similar to class averages from prior datasets.
(B) Anion exchange chromatogram of ERK2 coexpressed with constitutively active MEK1 mutant. Three distinct ERK2 peaks elute off an anion exchange column. Western blot of in vitro kinase reaction of different ERK2 anion exchange peaks phosphorylating purified GST-CTK WT or GST-CTK L714E (negative control), probed for phosphospecific CTK activation loop antibody (pT573). Intact time-of-flight mass spectrometry of Peak 1 ERK2 confirms that the vast majority of this protein is 2pERK2.
(C) 2D class averages of free RSK2; particles were picked from a dataset of crosslinked RSK2-2pERK2-ORF45 +Mg/AMPPNP. Theoretical 2D projections of AlphaFold RSK2 (46-740) model reasonably match 2D class averages of free RSK2 from the crosslinked RSK2-2pERK2-ORF45 dataset. The corresponding 3D reconstruction is fit by the RSK2 AF model.
(D) 2D class averages of free 2pERK2. Particles were picked from a dataset of crosslinked RSK2-2pERK2-ORF45 +Mg/AMPPNP. The experimental class averages match theoretical 2D projections of the 2pERK2 crystal structure (pdb:2erk). The corresponding ab inito reconstruction of 2pERK2 is fit with a crystal structure of 2pERK2 (pdb:2erk).


A
ERK2
RSK2
RSK2-ERK2-ORF45
B
Peak 1
Peak 2
Peak 3
Peak 1
Peak 3
CTK^WT
CTK^L714E
CTK^WT
CTK^L714E
α pT573^CTK
2p
Peak 1
C
Experimental
Projection
NTK
CTK
D
Experimental
Projection
ERK2

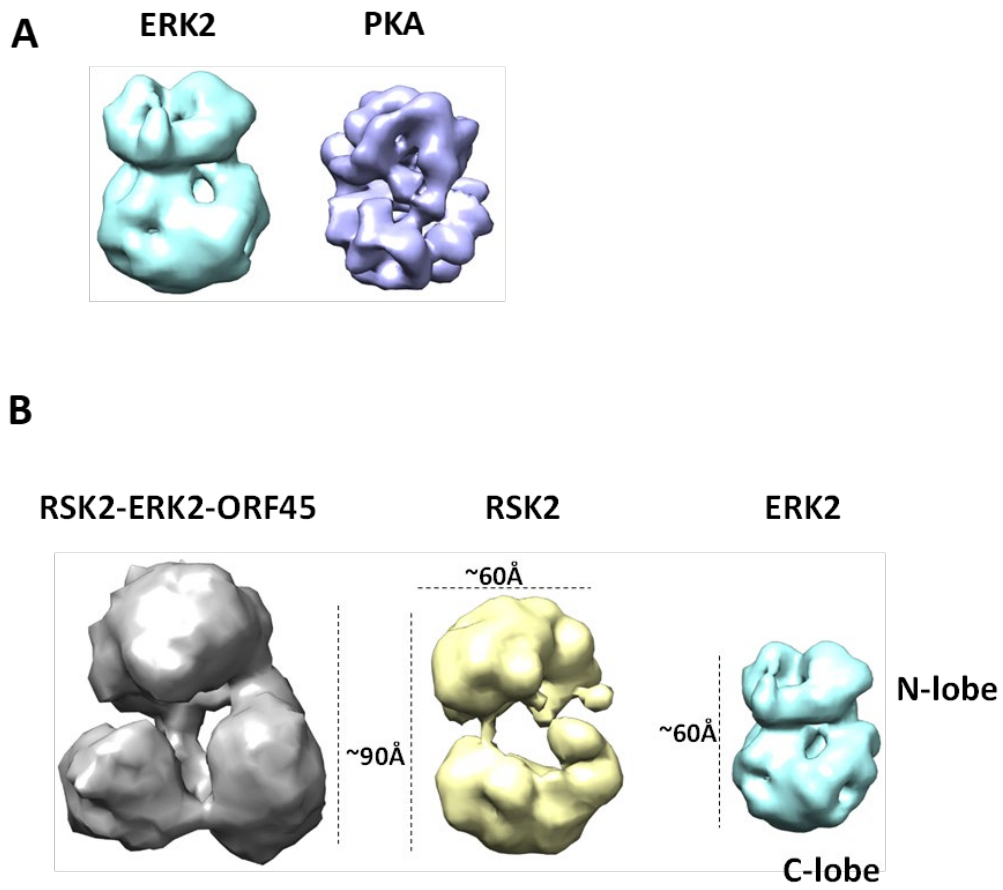


Figure 5.5. Comparison of low resolution densities of ERK2, RSK2, and RSK2-ERK2-ORF45+SL0101.

(A) Comparison of the ab inito reconstruction of ERK2 and ~6Å reconstruction of PKA (EMD-0409).

(B) Comparison of low resolution reconstructions of RSK2-ERK2-ORF45+SL0101 (grey), free RSK2 (yellow), and 2pERK2 (Teal). Dashed lines indicate the approximate size of each reconstruction. The distance from N-lobe to C-lobe in ERK2 is ~ 60Å, and represents the maximum size of a single kinase domain. The reconstructions of free RSK, and RSK2-ERK2-ORF45 have two and three kinase domains, respectively.

C. Discussion

Flanking regions of kinase domains, either unstructured disordered, or folded domains, are critical for kinase regulation and function. Flanking SLiMs/IDRs or regulatory domains can stabilize active or inactive conformations, or the assembly of activator-substrate complexes. The flexibility and formation of medium affinity interactions allow SLiMs to perform a variety of functions depending on the cellular context. The N and C terminal tails of AGC kinases, which contain SLiMs like the turn and hydrophobic motif (HM), wrap around the kinase core and are necessary for the stability and catalytic activity of the kinase (Kannan, Haste et al. 2007, Pearce, Komander et al. 2010, Taylor, Søberg et al. 2021). In RSK, the HM, which is a classic SLiM, serves as a substrate for the CTK, then becomes a binding partner for PDK1, the so-called master regulator of AGC kinases, before finally binding the NTK's own N-lobe, stabilizing the phosphorylated active kinase (Dalby, Morrice et al. 1998, Biondi, Cheung et al. 2000, Frödin, Jensen et al. 2000, Anjum and Blenis 2008). The N and C-terminal tails of MAP2Ks and MAPKAPKs contain D-motif and revD motifs, respectively, that enable binding to substrate or activator MAPKs (Garai, Zeke et al. 2012, Zeke, Bastys et al. 2015). An N-terminal D motif and C-terminal revD motif in MEK and RSK, respectively, drive assembly with ERK2. N-terminal SH2 and SH3 domains of Src suppress basal activity by preventing formation of an active kinase until binding of ligand (Wong, Jennings et al. 2004). Furthermore, kinases tend to assemble in higher order complexes to further regulate their activity and allow them to rapidly respond to second messengers (Gógl, Kornev et al. 2019). PKA forms

tetrameric holoenzymes with two regulatory subunits and two catalytic subunits to suppress catalytic activity until addition of cAMP (Taylor and Kornev 2011, Taylor, Keshwani et al. 2012). mTORC1,2 form mega Dalton complexes that prevents formation of the kinase active conformation until binding of its activator (Liu and Sabatini 2020). CAMKII is known to form dodecameric assemblies to promote activating autophosphorylation in the presence of Ca^{2+} and CaM binding (Temmerman, Simon et al. 2013, Bayer and Schulman 2019).

Since the first structure of PKA was solved over thirty years ago (Knighton, Zheng et al. 1991), the majority of studies focus on the kinase core only, and truncate flanking disordered tails or domains, as opposed to studying FL kinases and their complexes. As these flanking tails and domains have important functional or regulatory roles, analysis of FL kinases and their complexes is needed to better understand function. The size and relative dynamic nature of these FL proteins and complexes, make these challenging targets to study structurally. The developments in cryoEM over the past decade have enabled the study of more dynamic full length kinase complexes/assemblies. CryoEM structures have been determined for the PKA RIIβ holoenzyme (Lu, Aoto et al. 2021), mTORC1/2 complexes (Yang, Wang et al. 2016, Chen, Liu et al. 2018), PI3K complex (Liu, Yang et al. 2021), CAMKII holoenzymes (Myers, Zaegel et al. 2017), BRAF/MEK autoinhibited complex (Park, Rawson et al. 2019), among others.

Kinases also represent one of the most popular and successful drug targets in a variety of diseases (Wu, Nielsen et al. 2015). Most kinase inhibitors competitively target the kinase active sites. Recently, inhibitors that target complexes of RAF, or

RAF/MEK1 were developed or more fully understood. A new RAF inhibitor was shown to specifically inhibit BRAF dimers by disrupting their dimerization surface (Yao, Gao et al. 2019). Allosteric MEK inhibitors were found to be far more potent against the BRAF/MEK complex rather than free MEK, suggesting that further development of allosteric MEK inhibitors should be designed against BRAF bound MEK rather than free MEK (Gonzalez-Del Pino, Li et al. 2021). Thus, a greater understanding of FL kinase complexes may enable the development of novel small molecules that target kinase complexes and their protein protein interfaces needed for their stability and function.

Here we attempt to solve the structure of the FL RSK2-ERK2-ORF45 signaling complex using cryoEM to determine how the three kinase domains in the complex are all oriented relative to one another and better understand the interplay between kinase domains and IDRs. Furthermore, how is the Linker positioned in the complex, and do additional interactions between the NTK and CTK, or ERK2 exist? The rare ability of RSK to stably bind its activator kinase provides the opportunity to study how one kinase phosphorylates a substrate kinase, and the assembly of multi kinase complexes. Our cryoEM analysis of the RSK2-ERK2-ORF45 complex generated promising 2D class averages with visible secondary structure for free ERK2, RSK2, and RSK2-ERK2-ORF45 complex. These corresponding reconstructions are all limited to low resolution, but they are sufficient to determine the overall size and shape of each construct. Our inability, thus far, to achieve higher resolution, reinforces the challenge of structurally characterizing dynamic, flexible kinase complexes. The molecular envelope of ERK2 is fit well by the crystal structure

of ERK2. SAXS estimates of the size and shape of RSK2 and the RSK2-ERK2-ORF45 complex agrees with 3D reconstructions. The RSK2-ERK2-ORF45 density is larger than FL RSK2, and has a different, more compact, less extended shape.

Our most promising class averages and reconstructions of the RSK2-ERK2-ORF45 complex occurred with RSK2 bound to SL0101. The NTK active site inhibitor thermally stabilizes RSK2 (Utepbergenov, Derewenda et al. 2012, Utepbergenov, Hennig et al. 2016), presumably by stabilizing an unusual inactive conformation in which the αC helix is partially unfolded and replaced by the flexible N-terminal tail that adopts a β sheet. This inhibitor also causes significant deformations to the overall kinase fold, but it may be further stabilized by the formation of 'hydrophobic-like spines' (Fig 5.6A,B). These are not the canonical regulatory or catalytic spine (Kornev, Haste et al. 2006, Kornev, Taylor et al. 2008) residues but they are noncontiguous hydrophobic residues that similarly span both lobes of the kinase. Furthermore, the adenine ring in ATP completes the catalytic spine, in a similar fashion, SL0101 may also bridge the two hydrophobic like spines. Thus, this internal hydrophobic core may drive the stabilization of the NTK. And further cryoEM analysis may require SL0101 binding, and formation of this unusual NTK conformation to reach the highest possible resolution. We believe compositional heterogeneity, the presence of both RSK2-ERK2-ORF45 and particularly free RSK, and conformational heterogeneity, due to the flexible nature of kinase complexes, are responsible for limiting the resolution of our reconstructions. Given the secondary structure visible in the 2D class averages, we are optimistic that we can improve the resolution of our reconstructions by collecting more data, to significantly increase our

initial particle population. With a larger initial particle stack we will have a greater ability to isolate a more homogenous particle stack of RSK2-ERK2-ORF45 complex, and potentially obtain a higher resolution reconstruction.

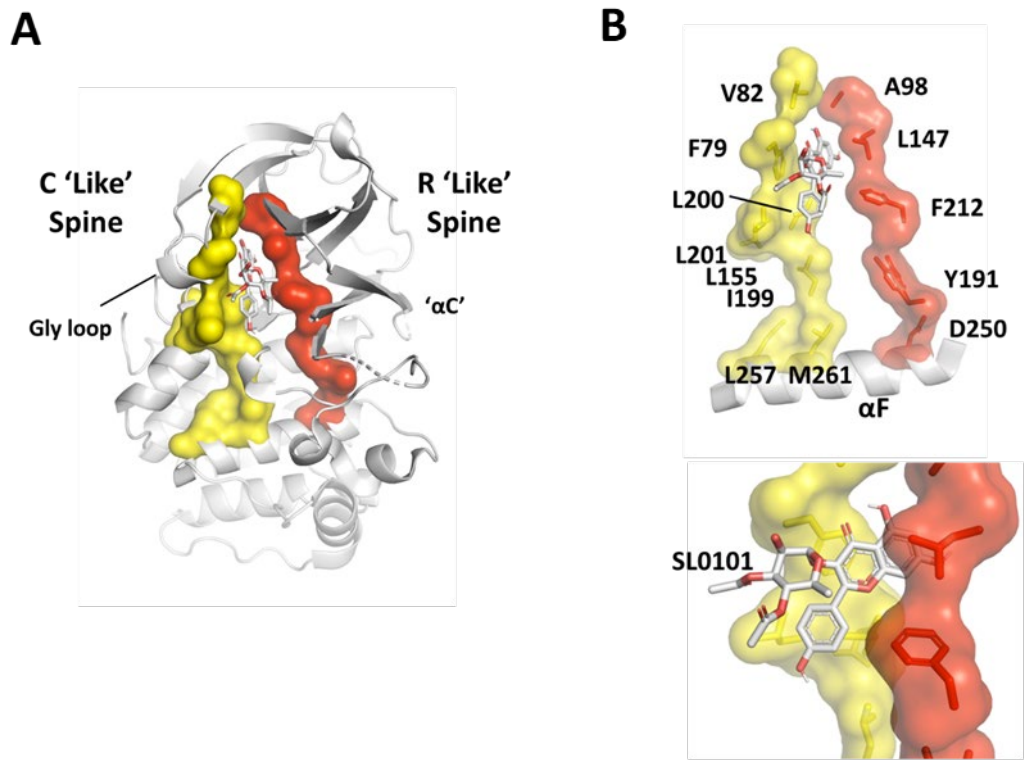


Figure 5.6. The structure of RSK2 NTK+SL0101 reveals 'catalytic-like' and 'regulatory-like' hydrophobic spines.
(A) The crystal structure of the RSK2 NTK+SL0101 features significant deformations to the overall kinase fold. The glycine rich loop is distorted and twisted out. The 'αC' helix partially adopts a β strand (the rest is disordered), and is replaced by a β strand from the flexible, disordered N-terminal tail. Similar to the canonical catalytic and regulatory spines, non-contiguous hydrophobic residues span both N and C lobes. (B) Top, the regulatory-like spine is anchored to the αF helix via D250 ($D220^{PKA}$), and consists of Y191 (YRD), F212 (DFG), L147 from the β5 strand, and A98 from the β3 strand. The 'catalytic-like' spine consists of L257 and M261 of the αF helix, catalytic loop residues I199, L200, and L201, the hinge residue L155, β2 strand V82, and F79 from the Gly loop. Bottom, the inhibitor SL0101 is sandwiched between the 'catalytic-like' and 'regulatory-like' spines, and may bridge the two spines.

D. Methods

1. Protein expression and purification

Proteins were expressed in either E. Coli BL21 Rosetta pLysS or BL21 in pET or pGEX vectors. Human full length (FL) RSK2 and EKR2 were expressed and purified as described previously (Alexa, Gógl et al. 2015, Alexa, Sok et al. 2022). Briefly, RSK2 was expressed with an N-terminal MBP tag and C-terminal His tag. A Tev cleavage site follows the MBP tag. A Tev cleavable His tagged ERK2 was coexpressed with lambda phosphatase for making inactive ERK2, and coexpressed with a constitutively active MEK1 to produce active ERK2. Human PDK1 (52-559) and ORF45 (16-76) was expressed with an N-terminal GST tag. Following transformations, all proteins were grown in LB at 37° until OD:0.6, then the temperature was reduced to 18° and IPTG was added to 0.1mM. Expression continued overnight, and pellets were collected and frozen.

RSK2 was lysed by a microfluidizer and clarified lysate was applied to equilibrated Ni resin, and following resin washing, eluted with high imidazole. Next, the Ni eluted protein was further purified by performing an amylose resin purification. Ni eluted protein was applied to amylose resin, washed, and eluted using 10mM amylose. The MBP tag was cleaved by addition of Tev protease (1:100 w/w) and overnight dialysis. The cleaved tag was removed from the purified RSK2 by anion exchange, using a hitrapQ column. The purified RSK2 was then concentrated and subjected to gel filtration with an S200 column. ERK2 pellets were resuspended and lysed with a microfluidizer, and centrifuged to isolate the soluble fraction. ERK2 was then purified

by Ni resin, and the eluted protein was incubated with Tev overnight to cleave the N-terminal His tag. Following cleavage, ERK2 was further purified by anion exchange (hitrapQ) to isolate non-phosphorylated or dually phosphorylated ERK2. Inactive or active ERK2 were then purified by gel filtration with an S200 column. ORF45 was lysed and purified via GST purification. The tag was cleaved by Tev, and sample was boiled for 5 minutes to precipitate most proteins, centrifuged, and the soluble fraction was applied to an S200 gel filtration column. RSK2-ERK2 or RSK2-ERK2-ORF45 complexes were prepared by mixing purified RSK2 and ERK2 proteins in an 1:1.3, or 1:1.3:1.5 molar ratio, respectively, and then injected onto a S200 gel filtration column. PDK1 pellets were resuspended and lysed by a microfluidizer. The soluble fraction of the lysate was applied to an equilibrated glutathione resin, washed, and eluted with 10mM glutathione.

2. Hydrogen deuterium exchange mass spectrometry

HDXMS was performed at the Biomolecular and Proteomics Mass Spectrometry Facility (BPMSF) of the University California San Diego, using a Waters Synapt G2Si system with HDX technology (Waters Corporation) according to methods previously described (Peacock, Davis et al. 2018). The binary ERK2-RSK2 complex was subjected HDXMS analysis and results of this were compared to that of the ERK2-RSK2-ORF45(16-76) sample. Briefly, deuterium exchange reactions were performed using a Leap HDX PAL autosampler (Leap Technologies, Carrboro, NC). D_2O buffer was prepared by lyophilizing gel filtration buffer, (20mM Hepes pH 7.5, 150mM

NaCl, and 1mM TCEP) initially dissolved in ultrapure water and redissolving the powder in the same volume of 99.96% D_2O (Cambridge Isotope Laboratories, Inc., Andover, MA) immediately before use. Deuterium exchange was measured in triplicate at each time point (0 min, 0.5 min, 1 min, 2 min, 5 min). For each deuteration time point, 4 µL of protein was held at 25 °C for 5 min before being mixed with 56 µL of D_2O buffer. The deuterium exchange was quenched for 1 min at 1°C by combining 50 µL of the deuteration reaction with 50 µL of 3M guanidine hydrochloride, final pH 2.66. The quenched sample (90 µL) was then injected in a 100 µL sample loop, followed by digestion on an in-line pepsin column (Immobilized Pepsin, Pierce) at 15°C. The resulting peptides were captured on a BEH C18 Vanguard precolumn, separated by analytical chromatography (Acquity UPLC BEH C18, 1.7 µm 1.0 × 50 mm, Waters Corporation) using a 7–85% acetonitrile gradient in 0.1% formic acid over 7.5 min, and electrosprayed into the Waters Synapt G2Si quadrupole time-of-flight mass spectrometer. The mass spectrometer was set to collect data in the Mobility, ESI+ mode; mass acquisition range of 200–2000 (m/z); scan time 0.4 s. Continuous lock mass correction was accomplished with infusion of leu-enkephalin (m/z = 556.277) every 30 s (mass accuracy of 1 ppm for calibration standard).

For peptide identification, the mass spectrometer was set to collect data in mobility-enhanced data-independent acquisition (MS^E), mobility ESI+ mode instead. Peptide masses were identified from triplicate analyses and data were analyzed using the ProteinLynx global server (PLGS) version 3.0 (Waters Corporation). Peptide masses were identified using a minimum number of 250 ion counts for low energy peptides

and 50 ion counts for their fragment ions; the peptides also had to be larger than 1,500 Da. The following cutoffs were used to filter peptide sequence matches: minimum products per amino acid of 0.2, minimum score of 7, maximum MH+ error of 5 ppm, and a retention time RSD of 5%. In addition, the peptides had to be present in two of the three ID runs collected. The peptides identified in PLGS were then analyzed using DynamX 3.0 data analysis software (Waters Corporation). Peptides containing mutated residues were manually assigned. The relative deuterium uptake for each peptide was calculated by comparing the centroids of the mass envelopes of the deuterated samples with the undeuterated controls following previously published methods (Wales, Fadgen et al. 2008). The peptides reported on the coverage maps are actually those from which data were obtained. For all HDX-MS data, at least 2 biological replicates were analyzed each with 3 technical replicates. Data are represented as mean values +/- SEM of 3 technical replicates due to processing software limitations, however the LEAP robot provides highly reproducible data for biological replicates. The deuterium uptake was corrected for back-exchange using a global back exchange correction factor (typically ~25%) determined from the average percent exchange measured in disordered termini of various proteins (Ramsey, Dembinski et al. 2017). ANOVA analyses and t tests with a p value cutoff of 0.05 implemented in the program, DECA, were used to determine the significance of differences between HDX data points (Lumpkin and Komives 2019). Deuterium uptake plots were generated in DECA (github.com/komiveslab/DECA) and the data are fitted with an exponential curve for ease of viewing.

3. Peptide arrays

Peptide arrays were synthesized with a MultiPep Flexible Parallel Peptide Synthesizer (Intavis Bioanalytical Instruments, Germany) as previously described (Happ, Arveseth et al. 2022). The blots were probed with GST-PDK1 or MBP-CTK. Unless noted otherwise all steps were carried out at room temperature. The blots were initially soaked with 100% ethanol for 5 minutes followed by 5 times 5-minute washes with water. All subsequent washes were carried out for 5 minutes 5 times with TBST. The membranes were washed with TBST, blocked with 5% BSA in TBST for 1 hour and then soaked with TBST containing 20 µg of protein, 5% BSA, 1 mM TCEP overnight at 4°. The next day the blots were washed before incubation overnight at 4° C with TBST containing 5% BSA and a primary Anti GST (Santa Cruz Biosciences) or Anti MBP (Cell Signaling Technology) antibody. The next day the blots were washed, then incubated with an HRP-conjugated anti-rabbit antibody (PrometheusTM) in 5% BSA TBST for one hour followed by a final round of washes with TBST. The membranes were finally covered with SuperSignal West Pico PLUS Chemiluminescent Substrate for detection of HRP (Thermo Scientific # 34580) and imaged with a ChemiDoc MP Imaging System from BIO RAD.

4. Kinase reactions and pulldowns

MBP and His tagged RSK2 WT and mutants were expressed, lysed by incubating with B-per complete supplemented with 150mM NaCl, 2mM DTT, and 1x Complete EDTA free protease inhibitor. The lysate was centrifuged and the soluble fraction was incubated with purified PDK1, ERK2, ATP/Mg for 10 minutes. Reactions were quenched by addition of SDS loading dye supplemented with 50mM EDTA. MBP

RSK2 with a CTK phosphomimetic mutation (T577E) were incubated only with ATP/Mg.

For pulldown assays, GST-PDK1 was immobilized against glutathione resin, and washed three times with binding buffer (20mM Hepes pH 7.5, 150mM NaCl, 5mM DTT). Purified RSK2 T577E or lambda phosphatase coexpressed RSK2 was the added to the resin/GST-PDK1 mixture and incubated for 30 minutes. Beads were washed and eluted with SDS sample buffer.

5. Crosslinking

RSK2-ERK2-ORF45 complexes were further stabilized by chemical crosslinking by incubation with glutaraldehyde. Glutaraldehyde was diluted in crosslinking buffer (50mM Hepes pH 7.5, 150 mM NaCl, 1mM TCEP) to a 5% stock solution. Purified RSK2-ERK2-ORF45 complex was diluted in crosslinking buffer to 0.1 mg/ml and was crosslinked by addition of diluted glutaraldehyde 0.1 % v/v. The reaction proceeded for 10 minutes at room temperature, and crosslinking was quenched by addition of 100mM Tris pH 8, and the sample was placed on ice. The sample was then concentrated and applied to an S200 gel filtration column to separate crosslinked RSK2-ERK2-ORF45 complex from overcrosslinked higher molecular weight artifacts.

6. Negative stain EM

Purified RSK2-ERK2-ORF45 with and without crosslink was diluted to 100 nM and applied to glow discharged carbon coated copper grids. Grids were stained in 2% uranyl acetate and dried. Samples were imaged on a Tecnai Sphera operating at 120 kEV using a CCD camera.

7. CryoEM

Purified samples that were stabilized by crosslinker or without were applied to plasma cleaned UltrAufoil 1.2/1.3 gold grids. Grids were prepared using a Vitrobot and stored in liquid nitrogen until imaged. Initial samples were collected on a Talos Arctica operating at 200 kV equipped with a K2 summit direct detector using Leginon (Suloway, Pulokas et al. 2005). Images were taken at a dose rate of 5.20 electrons $Å^{-2}s^{-1}$ for 10 seconds with 50 total frames. These images have a pixel size of 1.16 Å.

Samples were also imaged on a Titan Krios operating at 300 kV equipped with a K2 summit direct detector and a Quantum energy filter. Images were taken at a dose rate of 5.80 electrons $Å^{-2}s^{-1}$ for 10 seconds with 50 total frames. These images have a pixel size of 0.88 Å.

All raw movies were processed primarily using Cryosparc (Punjani, Rubinstein et al. 2017). Raw movies were uploaded, aligned, and CTF corrected by patch motion and CTF correction jobs. Micrographs with CTF fits worse than 5 Å were discarded. For ERK2 reconstructions, particles were picked using a template derived from an ERK2 crystal structure (pdb:2erk). Picked particles were extracted, and subjected to multiple rounds of 2d classification. Particles with secondary structure features were utilized for the creation of three ab initio structure, and were subsequently heterogenous refined. Final reconstructions were made using non-uniform refinement job. RSK2 reconstructions were performed by particle picking using cryosparcs blob picker. Following picking and extraction, particles were subjected to multiple rounds of 2d classification to remove junk particles. Cleaned RSK2 particles were used for ab inito

reconstruction and refined by non-uniform refinement. Crosslinked RSK2-ERK2-ORF45 particles were selected using an AI assisted Cryolo particle picking (Wagner, Merino et al. 2019). Following particle picking, extraction, and multiple rounds of 2d classification, classes that correspond only to the RSK2-ERK2-ORF45 complex were chosen for ab inito reconstruction.

E. References

Alexa, A., G. Gógl, G. Glatz, Á. Garai, A. Zeke, J. Varga, E. Dudás, N. Jeszenői, A. Bodor, C. Hetényi and A. Reményi (2015). "Structural assembly of the signaling competent ERK2-RSK1 heterodimeric protein kinase complex." Proc Natl Acad Sci U S A **112**(9): 2711-2716.

Alexa, A., P. Sok, F. Gross, K. Albert, E. Kobori, Á. L. Póti, G. Gógl, I. Bento, E. Kuang, S. S. Taylor, F. Zhu, A. Ciliberto and A. Reményi (2022). "A non-catalytic herpesviral protein reconfigures ERK-RSK signaling by targeting kinase docking systems in the host." Nature Communications **13**(1): 472.

Anjum, R. and J. Blenis (2008). "The RSK family of kinases: emerging roles in cellular signalling." Nature Reviews Molecular Cell Biology **9**(10): 747-758.

Bayer, K. U. and H. Schulman (2019). "CaM Kinase: Still Inspiring at 40." Neuron **103**(3): 380-394.

Biondi, R. M., P. C. Cheung, A. Casamayor, M. Deak, R. A. Currie and D. R. Alessi (2000). "Identification of a pocket in the PDK1 kinase domain that interacts with PIF and the C-terminal residues of PKA." Embo j **19**(5): 979-988.

Bjørbaek, C., Y. Zhao and D. E. Moller (1995). "Divergent functional roles for p90rsk kinase domains." J Biol Chem **270**(32): 18848-18852.

Brennan, D. F., A. C. Dar, N. T. Hertz, W. C. Chao, A. L. Burlingame, K. M. Shokat and D. Barford (2011). "A Raf-induced allosteric transition of KSR stimulates phosphorylation of MEK." Nature **472**(7343): 366-369.

Canagarajah, B. J., A. Khokhlatchev, M. H. Cobb and E. J. Goldsmith (1997). "Activation Mechanism of the MAP Kinase ERK2 by Dual Phosphorylation." Cell **90**(5): 859-869.

Chen, X., M. Liu, Y. Tian, J. Li, Y. Qi, D. Zhao, Z. Wu, M. Huang, C. C. L. Wong, H.-W. Wang, J. Wang, H. Yang and Y. Xu (2018). "Cryo-EM structure of human mTOR complex 2." Cell Research **28**(5): 518-528.

Cheng, F., J. Zhao, Y. Wang, W. Lu, Z. Liu, Y. Zhou, W. R. Martin, R. Wang, J. Huang, T. Hao, H. Yue, J. Ma, Y. Hou, J. A. Castrillon, J. Fang, J. D. Lathia, R. A. Keri, F. C. Lightstone, E. M. Antman, R. Rabadan, D. E. Hill, C. Eng, M. Vidal and J. Loscalzo (2021). "Comprehensive characterization of protein–protein interactions perturbed by disease mutations." Nature Genetics **53**(3): 342-353.

Dalby, K. N., N. Morrice, F. B. Caudwell, J. Avruch and P. Cohen (1998). "Identification of regulatory phosphorylation sites in mitogen-activated protein kinase (MAPK)-activated protein kinase-1a/p90rsk that are inducible by MAPK." J Biol Chem **273**(3): 1496-1505.

Davey, N. E., M. S. Cyert and A. M. Moses (2015). "Short linear motifs – ex nihilo evolution of protein regulation." Cell Communication and Signaling **13**(1): 43.

Fisher, T. L. and J. Blenis (1996). "Evidence for two catalytically active kinase domains in pp90rsk." Mol Cell Biol **16**(3): 1212-1219.

Frödin, M., C. J. Jensen, K. Merienne and S. Gammeltoft (2000). "A phosphoserine-regulated docking site in the protein kinase RSK2 that recruits and activates PDK1." The EMBO Journal **19**(12): 2924-2934.

Garai, Á., A. Zeke, G. Gógl, I. Törő, F. Fördős, H. Blankenburg, T. Bárkai, J. Varga, A. Alexa, D. Emig, M. Albrecht and A. Reményi (2012). "Specificity of Linear Motifs That Bind to a Common Mitogen-Activated Protein Kinase Docking Groove." Science Signaling **5**(245): ra74-ra74.

Gógl, G., A. Alexa, B. Kiss, G. Katona, M. Kovács, A. Bodor, A. Reményi and L. Nyitray (2016). "Structural Basis of Ribosomal S6 Kinase 1 (RSK1) Inhibition by S100B Protein: MODULATION OF THE EXTRACELLULAR SIGNAL-REGULATED KINASE (ERK) SIGNALING CASCADE IN A CALCIUM-DEPENDENT WAY." J Biol Chem **291**(1): 11-27.

Gógl, G., A. P. Kornev, A. Reményi and S. S. Taylor (2019). "Disordered Protein Kinase Regions in Regulation of Kinase Domain Cores." Trends Biochem Sci **44**(4): 300-311.

Gonzalez-Del Pino, G. L., K. Li, E. Park, A. M. Schmoker, B. H. Ha and M. J. Eck (2021). "Allosteric MEK inhibitors act on BRAF/MEK complexes to block MEK activation." Proceedings of the National Academy of Sciences **118**(36): e2107207118.

Happ, J. T., C. D. Arveseth, J. Bruystens, D. Bertinetti, I. B. Nelson, C. Olivieri, J. Zhang, D. S. Hedeen, J.-F. Zhu, J. L. Capener, J. W. Bröckel, L. Vu, C. C. King, V. L. Ruiz-Perez, X. Ge, G. Veglia, F. W. Herberg, S. S. Taylor and B. R. Myers (2022). "A PKA inhibitor motif within SMOOTHENED controls Hedgehog signal transduction." Nature Structural & Molecular Biology **29**(10): 990-999.

Hartman, K. G., M. I. Vitolo, A. D. Pierce, J. M. Fox, P. Shapiro, S. S. Martin, P. T. Wilder and D. J. Weber (2014). "Complex formation between S100B protein and the p90 ribosomal S6 kinase (RSK) in malignant melanoma is calcium-dependent and inhibits extracellular signal-regulated kinase (ERK)-mediated phosphorylation of RSK." J Biol Chem **289**(18): 12886-12895.

Herzik, M. A., M. Wu and G. C. Lander (2019). "High-resolution structure determination of sub-100 kDa complexes using conventional cryo-EM." Nature Communications **10**(1): 1032.

Hoofnagle, A. N., K. A. Resing, E. J. Goldsmith and N. G. Ahn (2001). "Changes in protein conformational mobility upon activation of extracellular regulated protein kinase-2 as detected by hydrogen exchange." Proc Natl Acad Sci U S A **98**(3): 956-961.

Jones, S. W., E. Erikson, J. Blenis, J. L. Maller and R. L. Erikson (1988). "A Xenopus ribosomal protein S6 kinase has two apparent kinase domains that are each similar to distinct protein kinases." Proc Natl Acad Sci U S A **85**(10): 3377-3381.

Jumper, J., R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko, A. Bridgland, C. Meyer, S. A. A. Kohl, A. J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, S. Bodenstein, D. Silver, O. Vinyals, A. W. Senior, K. Kavukcuoglu, P. Kohli and D. Hassabis (2021).

"Highly accurate protein structure prediction with AlphaFold." Nature **596**(7873): 583-589.
Kannan, N., N. Haste, S. S. Taylor and A. F. Neuwald (2007). "The hallmark of AGC kinase functional divergence is its C-terminal tail, a cis-acting regulatory module." Proceedings of the National Academy of Sciences **104**(4): 1272-1277.
Knighton, D. R., J. H. Zheng, L. F. Ten Eyck, V. A. Ashford, N. H. Xuong, S. S. Taylor and J. M. Sowadski (1991). "Crystal structure of the catalytic subunit of cyclic adenosine monophosphate-dependent protein kinase." Science **253**(5018): 407-414.
Kornev, A. P., N. M. Haste, S. S. Taylor and L. F. Eyck (2006). "Surface comparison of active and inactive protein kinases identifies a conserved activation mechanism." Proc Natl Acad Sci U S A **103**(47): 17783-17788.
Kornev, A. P. and S. S. Taylor (2015). "Dynamics-Driven Allostery in Protein Kinases." Trends Biochem Sci **40**(11): 628-647.
Kornev, A. P., S. S. Taylor and L. F. Ten Eyck (2008). "A helix scaffold for the assembly of active protein kinases." Proc Natl Acad Sci U S A **105**(38): 14377-14382.
Kuang, E., Q. Tang, G. G. Maul and F. Zhu (2008). "Activation of p90 ribosomal S6 kinase by ORF45 of Kaposi's sarcoma-associated herpesvirus and its role in viral lytic replication." J Virol **82**(4): 1838-1850.
Kuang, E., F. Wu and F. Zhu (2009). "Mechanism of sustained activation of ribosomal S6 kinase (RSK) and ERK by kaposi sarcoma-associated herpesvirus ORF45: multiprotein complexes retain active phosphorylated ERK AND RSK and protect them from dephosphorylation." J Biol Chem **284**(20): 13958-13968.
Lee, T., A. N. Hoofnagle, Y. Kabuyama, J. Stroud, X. Min, E. J. Goldsmith, L. Chen, K. A. Resing and N. G. Ahn (2004). "Docking motif interactions in MAP kinases revealed by hydrogen exchange mass spectrometry." Mol Cell **14**(1): 43-55.
Liu, G. Y. and D. M. Sabatini (2020). "mTOR at the nexus of nutrition, growth, ageing and disease." Nature Reviews Molecular Cell Biology **21**(4): 183-203.
Liu, X., S. Yang, J. R. Hart, Y. Xu, X. Zou, H. Zhang, Q. Zhou, T. Xia, Y. Zhang, D. Yang, M. W. Wang and P. K. Vogt (2021). "Cryo-EM structures of PI3Kα reveal conformational changes during inhibition and activation." Proc Natl Acad Sci U S A **118**(45).
Lu, T.-W., P. C. Aoto, J.-H. Weng, C. Nielsen, J. N. Cash, J. Hall, P. Zhang, S. M. Simon, M. A. Cianfrocco and S. S. Taylor (2021). "Structural analyses of the PKA RIIβ holoenzyme containing the oncogenic DnaJB1-PKAc fusion protein reveal protomer asymmetry and fusion-induced allosteric perturbations in fibrolamellar hepatocellular carcinoma." PLOS Biology **18**(12): e3001018.
Lumpkin, R. J. and E. A. Komives (2019). "DECA, A Comprehensive, Automatic Post-processing Program for HDX-MS Data." Mol Cell Proteomics **18**(12): 2516-2523.
Marcotte, D., M. Rushe, M. A. R, C. Lukacs, K. Atkins, X. Sun, K. Little, M. Cullivan, M. Paramasivam, T. A. Patterson, T. Hesson, D. M. T, T. L. May-

Dracka, Z. Xin, A. Bertolotti-Ciarlet, G. R. Bhisetti, J. P. Lyssikatos and L. F. Silvian (2017). "Germinal-center kinase-like kinase co-crystal structure reveals a swapped activation loop and C-terminal extension." Protein Sci **26**(2): 152-162.

Mirdita, M., K. Schütze, Y. Moriwaki, L. Heo, S. Ovchinnikov and M. Steinegger (2022). "ColabFold: making protein folding accessible to all." Nature Methods **19**(6): 679-682.

Mora, A., D. Komander, D. M. van Aalten and D. R. Alessi (2004). "PDK1, the master regulator of AGC kinase signal transduction." Semin Cell Dev Biol **15**(2): 161-170.

Myers, J. B., V. Zaegel, S. J. Coultrap, A. P. Miller, K. U. Bayer and S. L. Reichow (2017). "The CaMKII holoenzyme structure in activation-competent conformations." Nature Communications **8**(1): 15742.

Oliver, A. W., S. Knapp and L. H. Pearl (2007). "Activation segment exchange: a common mechanism of kinase autophosphorylation?" Trends Biochem Sci **32**(8): 351-356.

Park, E., S. Rawson, K. Li, B. W. Kim, S. B. Ficarro, G. G. Pino, H. Sharif, J. A. Marto, H. Jeon and M. J. Eck (2019). "Architecture of autoinhibited and active BRAF-MEK1-14-3-3 complexes." Nature **575**(7783): 545-550.

Peacock, R. B., J. R. Davis, P. R. L. Markwick and E. A. Komives (2018). "Dynamic Consequences of Mutation of Tryptophan 215 in Thrombin." Biochemistry **57**(18): 2694-2703.

Pearce, L. R., D. Komander and D. R. Alessi (2010). "The nuts and bolts of AGC protein kinases." Nat Rev Mol Cell Biol **11**(1): 9-22.

Punjani, A., J. L. Rubinstein, D. J. Fleet and M. A. Brubaker (2017). "cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination." Nat Methods **14**(3): 290-296.

Ramsey, K. M., H. E. Dembinski, W. Chen, C. G. Ricci and E. A. Komives (2017). "DNA and IκBα Both Induce Long-Range Conformational Changes in NFκB." Journal of molecular biology **429**(7): 999-1008.

Roux, P. P., S. A. Richards and J. Blenis (2003). "Phosphorylation of p90 ribosomal S6 kinase (RSK) regulates extracellular signal-regulated kinase docking and RSK activity." Mol Cell Biol **23**(14): 4796-4804.

Sok, P., G. Gógl, G. S. Kumar, A. Alexa, N. Singh, K. Kirsch, A. Sebő, L. Drahos, Z. Gáspári, W. Peti and A. Reményi (2020). "MAP Kinase-Mediated Activation of RSK1 and MK2 Substrate Kinases." Structure **28**(10): 1101-1113.e1105.

Suloway, C., J. Pulokas, D. Fellmann, A. Cheng, F. Guerra, J. Quispe, S. Stagg, C. S. Potter and B. Carragher (2005). "Automated molecular microscopy: the new Leginon system." J Struct Biol **151**(1): 41-60.

Tate, J. G., S. Bamford, H. C. Jubb, Z. Sondka, D. M. Beare, N. Bindal, H. Boutselakis, C. G. Cole, C. Creatore, E. Dawson, P. Fish, B. Harsha, C. Hathaway, S. C. Jupe, C. Y. Kok, K. Noble, L. Ponting, C. C. Ramshaw, C. E. Rye, H. E. Speedy, R. Stefancsik, S. L. Thompson, S. Wang, S. Ward, P. J. Campbell and S. A. Forbes (2019). "COSMIC: the Catalogue Of Somatic Mutations In Cancer." Nucleic Acids Res **47**(D1): D941-d947.

Taylor, S. S., M. M. Keshwani, J. M. Steichen and A. P. Kornev (2012). "Evolution of the eukaryotic protein kinases as dynamic molecular switches." Philosophical Transactions of the Royal Society B: Biological Sciences **367**(1602): 2517-2528.

Taylor, S. S. and A. P. Kornev (2011). "Protein kinases: evolution of dynamic regulatory proteins." Trends Biochem Sci **36**(2): 65-77.

Taylor, S. S., K. Søberg, E. Kobori, J. Wu, S. Pautz, F. W. Herberg and B. S. Skålhegg (2021). "The tails of PKA." Molecular Pharmacology: MOLPHARM-MR-2021-000315.

Temmerman, K., B. Simon and M. Wilmanns (2013). "Structural and functional diversity in the activity and regulation of DAPK-related protein kinases." Febs j **280**(21): 5533-5550.

Trivier, E., D. De Cesare, S. Jacquot, S. Pannetier, E. Zackai, I. Young, J. L. Mandel, P. Sassone-Corsi and A. Hanauer (1996). "Mutations in the kinase Rsk-2 associated with Coffin-Lowry syndrome." Nature **384**(6609): 567-570.

Utepbergenov, D., U. Derewenda, N. Olekhnovich, G. Szukalska, B. Banerjee, M. K. Hilinski, D. A. Lannigan, P. T. Stukenberg and Z. S. Derewenda (2012). "Insights into the inhibition of the p90 ribosomal S6 kinase (RSK) by the flavonol glycoside SL0101 from the 1.5 Å crystal structure of the N-terminal domain of RSK2 with bound inhibitor." Biochemistry **51**(33): 6499-6510.

Utepbergenov, D., P. M. Hennig, U. Derewenda, M. V. Artamonov, A. V. Somlyo and Z. S. Derewenda (2016). "Bacterial Expression, Purification and In Vitro Phosphorylation of Full-Length Ribosomal S6 Kinase 2 (RSK2)." PLOS ONE **11**(10): e0164343.

Varadi, M., S. Anyango, M. Deshpande, S. Nair, C. Natassia, G. Yordanova, D. Yuan, O. Stroe, G. Wood, A. Laydon, A. Žídek, T. Green, K. Tunyasuvunakool, S. Petersen, J. Jumper, E. Clancy, R. Green, A. Vora, M. Lutfi, M. Figurnov, A. Cowie, N. Hobbs, P. Kohli, G. Kleywegt, E. Birney, D. Hassabis and S. Velankar (2021). "AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models." Nucleic Acids Research **50**(D1): D439-D444.

Wagner, T., F. Merino, M. Stabrin, T. Moriya, C. Antoni, A. Apelbaum, P. Hagel, O. Sitsel, T. Raisch, D. Prumbaum, D. Quentin, D. Roderer, S. Tacke, B. Siebolds, E. Schubert, T. R. Shaikh, P. Lill, C. Gatsogiannis and S. Raunser (2019). "SPHIRE-crYOLO is a fast and accurate fully automated particle picker for cryo-EM." Communications Biology **2**(1): 218.

Wales, T. E., K. E. Fadgen, G. C. Gerhardt and J. R. Engen (2008). "High-speed and high-resolution UPLC separation at zero degrees Celsius." Anal Chem **80**(17): 6815-6820.

Wang, J., J. W. Wu and Z. X. Wang (2011). "Structural insights into the autoactivation mechanism of p21-activated protein kinase." Structure **19**(12): 1752-1761.

Wong, L., P. A. Jennings and J. A. Adams (2004). "Communication pathways between the nucleotide pocket and distal regulatory sites in protein kinases." Acc Chem Res **37**(5): 304-311.

Wright, P. E. and H. J. Dyson (2015). "Intrinsically disordered proteins in cellular signalling and regulation." Nat Rev Mol Cell Biol **16**(1): 18-29.
Wu, P., T. E. Nielsen and M. H. Clausen (2015). "FDA-approved small-molecule kinase inhibitors." Trends Pharmacol Sci **36**(7): 422-439.
Yang, H., J. Wang, M. Liu, X. Chen, M. Huang, D. Tan, M. Q. Dong, C. C. Wong, J. Wang, Y. Xu and H. W. Wang (2016). "4.4 Å Resolution Cryo-EM structure of human mTOR Complex 1." Protein Cell **7**(12): 878-887.
Yao, Z., Y. Gao, W. Su, R. Yaeger, J. Tao, N. Na, Y. Zhang, C. Zhang, A. Rymar, A. Tao, N. M. Timaul, R. McGriskin, N. A. Outmezguine, H. Zhao, Q. Chang, B. Qeriqi, M. Barbacid, E. de Stanchina, D. M. Hyman, G. Bollag and N. Rosen (2019). "RAF inhibitor PLX8394 selectively disrupts BRAF dimers and RAS-independent BRAF-mutant-driven signaling." Nat Med **25**(2): 284-291.
Zeke, A., T. Bastys, A. Alexa, Á. Garai, B. Mészáros, K. Kirsch, Z. Dosztányi, O. V. Kalinina and A. Reményi (2015). "Systematic discovery of linear binding motifs targeting an ancient protein interaction surface on MAP kinases." Mol Syst Biol **11**(11): 837.