STRUCTURAL ANALYSIS OF STE20-RELATED PROLINE-ALANINE-RICH KINASE (SPAK)

APPROVED BY SUPERVISORY COMMITTEE

Melanie H. Cobb, Ph.D.

Elizabeth J. Goldsmith, Ph.D

Philip J. Thomas, Ph.D

Kevin H. Gardner, Ph.D

DEDICATED TO MY FAMILY

STRUCTURAL ANALYSIS OF STE20-RELATED PROLINE-ALANINE-RICH KINASE (SPAK)

by

YU-CHI JUANG

DISSERTATION

Presented to the Faculty of the Graduate School of Biomedical Sciences

The University of Texas Southwestern Medical Center at Dallas

In Partial Fulfillment of the Requirements

For the Degree of

DOCTOR OF PHILOSOPHY

The University of Texas Southwestern Medical Center at Dallas

Dallas, Texas

January, 2009

ACKNOWLEDGMENTS

There are many people to thank for my great experience in the graduate study. First, I would like to express my deepest thanks to my mentor, Dr. Melanie Cobb. Her brilliant guidance and generous support helped me get through my graduate career. Her enthusiasm, knowledge and wisdom always inspire me. I would also like to thank Dr. Betsy Goldsmith, for her vast knowledge of crystallography and indispensable help that assisted me to finish this project. I gratefully acknowledge my committee members, Drs. Kevin Gardner, and Philip Thomas for their valuable suggestions.

I would like to thank all past and current members of the Cobb laboratory for establishing such a nice working environment and for their generous help and support to me. Many thanks to Drs. Radha Akella, Dominika Borek, and Xiaoshan Min for helping me with crystallography. I also thank Drs. Mischa Machius and Diana Tomchick in the Structural Biology Group for their helps on collecting the synchrotron datasets.

Finally, I would like to thank my family and friends for all their support and encouragement during my whole life.

STRUCTURAL ANALYSIS OF STE20-RELATED PROLINE-ALANINE-RICH KINASE (SPAK)

Yu-Chi Juang, Ph.D.

The University of Texas Southwestern Medical Center at Dallas, Graduation 2009

Supervising Professor: Melanie H. Cobb, Ph.D.

Ste20-related proline-alanine-rich kinase (SPAK) is a kinase that regulates ion cotransporters including KCC, NKCC1, NKCC2 and NCC. Recently, SPAK was identified as a target regulated by the WNK (With No lysine (K)) family of protein kinases. Overexpression of WNK1 has been associated to a hereditary form of hypertension. The link between WNK-SPAK-cotransporters provides one of the

functional pathways for WNK to regulate intracellular and extracellular salt and water balance, thus contribute to the control of blood pressure. My study focused on determining the crystal structure of SPAK and providing a molecular basis for understanding the catalytic and regulatory mechanisms of this enzyme.

I have solved the crystal structure of a partially active form of SPAK 63-390 (T243D). The structure contains the kinase domain and part of the C-terminal PF1 domain with a mutation of the WNK phosphorylation site, Thr243, to aspartate to mimic the phosphorylated state. The structure reveals two dimer interfaces. One is from the unique activation loop-swapped structure. The other is formed between the N-terminal domain of one molecule and the C-terminal domain of the second molecule. Comparison the structure with unphosphorylated inactive OSR1 revealed significant conformational changes in the glycine-rich loop, helix αC, and the activation loop. A more striking feature is the reformation of the P+1 pocket in SPAK. The interaction is formed by residues from both molecules of the domain-swapped dimer supporting the conclusion that the domain-swapped dimer is a functional unit. As remodeling of the P+1 pocket is usually associated with kinase activation, the conformational change found between the crystallized forms of OSR1 and SPAK is thus believed to be a regulated event. The unique activation loop-swapped feature suggests a novel regulatory mechanism that could be used in kinases.

TABLE OF CONTENTS

PRIOR PUBLICATIONS

Anselmo AN, Earnest S, Chen W, **Juang YC**, Kim SC, Zhao Y, Cobb MH. 2006. WNK1 and OSR1 regulate the Na+, K+, 2Cl- cotransporter in HeLa cells. *Proc Natl Acad Sci U S A.* 103(29):10883-8.

Wilsbacher JL, **Juang YC**, Khokhlatchev AV, Gallagher E, Binns D, Goldsmith EJ, Cobb MH. 2006. Characterization of mitogen-activated protein kinase (MAPK) dimers. Biochemistry. 45(44):13175-82.

Chen W, Chen Y, Xu BE, **Juang YC**, Stippec S, Zhao Y, Cobb MH. 2008. Regulation of a third conserved phosphorylation site in serum and glucocorticoid-induced protein kinase 1. J Biol Chem. 2008 Dec 9.

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

Angstrom	Å
Adenosine 5′-(β,γ-imido)triphosphate	AMP-PNP
adenosine 5'-triphosphate	ATP
atypical protein kinases	aPK
Ca^{2+}-dependent protein kinase	PKC
calmodulin-dependent kinase	CAMK
cAMP-dependent protein kinase	PKA
casein kinase 1	CK1
Casein kinase II	CKII
Cdc42/Rac1 interactive binding motif	CRIB
cGMP-dependent protein kinase	PKG
charge-coupled device	CCD
checkpoint kinase 2	CHK2
common docking	CD
cyclin-dependent kinase	CDK
cystic fibrosis transmembrane conductance regulator chloride channel	CFTR
death-associated protein kinase 3	DAPK3
DNA-dependent protein kinase	DNA-PK
epidermal growth factor receptor	EGFR
eukaryotic protein kinase superfamily	ePK
germinal center kinase	GCK
glycogen synthase kinase	GSK
isoelectric point	pI
isopropyl-1-thio-β-D- galactopyranoside	IPTG
K^+-Cl^- cotransporter	KCC
localization signal	NLS
log-likelihood gain	LLG
lymophocyte-originated kinase	LOK
MAK2K kinase	MAP3K
mammalian target of rapamycin	mTOR
MAP3K kinase	MAP4K
MAPK kinase	MAP2K
Matthews coefficient	V_M

mitogen-activated kinase	MAPK
mixed-linkage kinase	MLK
molecular replacement	MR
Na^+-Cl^- cotransporter	NCC
Na^+-K^+-$2Cl^-$ cotransporter	NKCC
nitrilotriacetic acid-agarose	Ni^{2+}-NTA
oxidative stress-responsive kinase 1	OSR1
p21-activated protein kinase	Pak
phenylmethanesulphonyl fluoride	PMSF
phosphatidylinositol 3-kinase	PI3K
phosphoinositide-dependent protein kinase	PDK1
phospholipid-dependent kinase	PDK
pithelial sodium channel	ENaC
polyethylene glycol	PEG
protein kinase inhibitor	PKI
pseudohypoaldosteronism type II	PHAII
renal outer medullary potassium channel	ROMK
root mean square	r.m.s.
rotation function Z-score	RFZ
serine threonine kinase-10	STK10
Serum- and glucocorticoid- induced kinase	SGK1
single nucleotide polymorphism	SNP
small angle X-ray scattering	SAXS
Src-homology 2	SH2
Ste20-like kinase	SLK
Ste20-related proline-alanine-rich kinase	SPAK
sterile kinase	STE
tau tubulin kinase	TTBK
the tyrosine kinase-like	TKL
tobacco etch virus	TEV
transforming growth factor beta activated kinase-1	TAK1
translation function Z-score	TFZ
tyrosine kinase	TK
vaccinia-related kinase	VRK
with no lysine (K) kinase	WNK

CHAPTER ONE

INTRODUCTION

I. Protein Kinases

Protein phosphorylation is one of the major mechanisms for regulating protein function in virtually every aspect of cell functions. Phosphorylation was first discovered in the early 20th century by Phoebus Levene (Pawson & Scott, 2005). However, it was not until 1954 that Eugene Kennedy and George Burnett observed that an enzyme activity was involved in phosphorylation (Burnett & Kennedy, 1954). The enzyme responsible was a liver protein that catalyzed the phosphorylation of casein. This enzyme was later known as casein kinase. After that, the role of phosphorylation has become recognized as more important as Edwin Krebs and Edward Fisher showed that through addition and removal of a phosphate, the enzyme activities in glycogen metabolism can be regulated. They further identified the first protein kinase, phosphorylase kinase, for which they received the Nobel Prize in 1992 (Fischer & Krebs, 1955).

During the process of protein phosphorylation, the protein kinase binds both ATP and its protein substrate and directly transfers the γ-phosphoryl group of ATP to the hydroxyl acceptor groups of Ser, Thr or Tyr residues on the substrate. Depending on the residues they phosphorylate, protein kinases can be separated into three groups: Ser/Thr, Tyr, and dual specificity kinases. While Ser/Thr kinases and Tyr kinases only transfer

phosphoryl groups to either Ser/Thr or Tyr, dual specificity kinases are able to phosphorylate Ser, Thr and Tyr residues in a limited member of substrates.

To date, more than 500 human protein kinases have been identified (Manning et al, 2002). They make up one of the largest gene families in humans. The majority of protein kinases belong to the eukaryotic protein kinase superfamily (ePK) which is responsible for phosphorylation of Ser, Thr and Tyr residues. Members of this superfamily are found in all eukaryotic organisms. The superfamily of eukaryotic protein kinases has been intensively studied. Protein kinases are major regulators of signaling pathways involved in almost every cellular process, including cell growth, proliferation, differentiation, metabolism, gene transcription, apoptosis, and homeostasis. Studies on members of this superfamily will continue to provide understanding of fundamental mechanisms underlying cellular behavior. A small number of human protein kinases are categorized as atypical protein kinases (aPK), such as pyruvate dehydrogenase kinase and phosphatidylinositol phosphate kinase. Although they possess phosphotransferase activity, the amino acid sequences and mechanisms used in this transfer are very different from those of the eukaryotic protein kinase family (Steussy et al, 2001; Walker et al, 1999; Yamaguchi et al, 2001).

Phosphorylation regulates substrate proteins by altering their enzymatic activity, subcellular localization, protein-protein interactions and stability, and thus leading to a wide variety of changes in cellular processes (Hunter, 2000). Therefore, precise regulation of kinase activity is required for normal cell function. Mutations and aberrant regulation of protein kinases have been associated with many diseases such as cancer,

diabetes, and hypertension (Blume-Jensen & Hunter, 2001; Cohen, 2002). Indeed, results from human genomic analysis revealed that 164 kinases map to regions often associated with human tumors and 80 kinases map to loci of other human diseases (Manning et al, 2002). Accordingly, protein kinases have gained attentions as targets for developing drugs for different diseases. Several drugs developed to inhibit kinases are currently in clinical use, including Imatinib (known as Gleevec; an ABL kinase inhibitor) (Nagar et al, 2002) and Gefitinib (known as Tarceva; an EGFR tyrosine kinase inhibitor) (Schindler et al, 2000).

II. Classification of Protein Kinases

Protein kinases are characterized by the presence of a catalytic kinase domain consisting of approximately 250 amino acid residues (Hanks et al, 1988). Most protein kinases also contain N-terminal and C-terminal extensions from the kinase domain. These non-catalytic regions usually possess domains involved in regulation of kinase activity, interactions with other protein partners, and control of subcellular localization. Alignment of all protein kinase sequences reveals that kinases share relatively low sequence similarity. The average sequence identity is about 20% in the kinase domain; while there is less similarity outside the kinase domain (Farnum et al, 2003).

Based on sequence similarity within the kinase domain, along with functional and structural information from the accessory domains, eukaryotic protein kinases are

now classified into seven major groups, including six Ser/Thr kinase families and one Tyr kinase family (Manning et al, 2002). They are as follows:

1) The AGC family consists of a group of kinases regulated by second-messengers such as cAMP-dependent protein kinase (PKA), Ca^{2+}-dependent protein kinase (PKC), cGMP-dependent protein kinase (PKG), and phospholipid-dependent kinase (PDK); 2) the CMGC family includes cyclin-dependent kinase (CDK), mitogen-activated kinase (MAPK), glycogen synthase kinase (GSK), and Casein kinase II (CKII); 3) the calmodulin-dependent kinase (CAMK) family; 4) the tyrosine kinase (TK) family includes both receptor protein kinases and cytoplasmic protein kinases; 5) the tyrosine kinase-like (TKL) family, which are similar to tyrosine kinases in protein sequence but are in fact serine/threonine protein kinases, consists of mixed-linkage kinase (MLK), Raf, and others; 6) the sterile (STE) kinases family includes members that act upstream of MAP kinase pathways, such as MAPK kinases (MAP2Ks), MAK2K kinases (MAP3Ks), and MAP3K kinases (MAP4Ks); and 7) the casein kinase (1CK1) family includes CK1, tau tubulin kinase (TTBK), and vaccinia-related kinase (VRK) subfamilies.

III. Structure of protein kinases

1. Overview of kinase structure

cAMP-dependent protein kinase A (PKA) was the first protein kinase to be crystallized (Bossemeyer et al, 1993; Knighton et al, 1991). This structure consisted of the protein kinase domain of PKA in complex with a pseudosubstrate peptide inhibitor

derived from the protein kinase inhibitor (PKI) and was in the active conformation. Since then, more than fifty eukaryotic protein kinase structures have been solved in either active or inactive states (Goldsmith et al, 2007). This structural information has provided understanding of the catalytic and regulatory mechanisms of protein kinases. Comparison of available structures reveals that all kinase domains share a common architecture despite differences in sequence, regulation, and function (Nolen et al, 2004).

The kinase domain can be viewed as a two-domain structure consisting of a small N-terminal domain composed primarily of an antiparallel β-sheet and a larger primarily helical C-terminal domain. The N-terminal domain contains five β strands (strands 1-5) and one α helix (helix C) (Knighton et al, 1991). The C-terminal domain is comprised of six large helices (helices D-I) and two β strands and it is responsible for protein substrate binding and phosphotransfer. The two domains are connected by a flexible linker (between β strand 5 and helix D), which functions as a hinge allowing movements between the two domains. The active site falls in the interface between the two domains. ATP and Mg^{2+} ions bind in the interior of the active site and protein substrate sits on the outside (Figure 1-1).

Several highly conserved sequences are found in the kinase domain that play crucial roles in catalysis. The consensus sequence GxGxxG, known as the glycine-rich loop, lies between β strands 1 and 2 of the small domain. This loop provides structural flexibility that is necessary for its function as an ATP anchor. The highly conserved lysine (Lys72 in PKA) in β strand 3 forms a salt bridge to a nearly-invariant glutamate

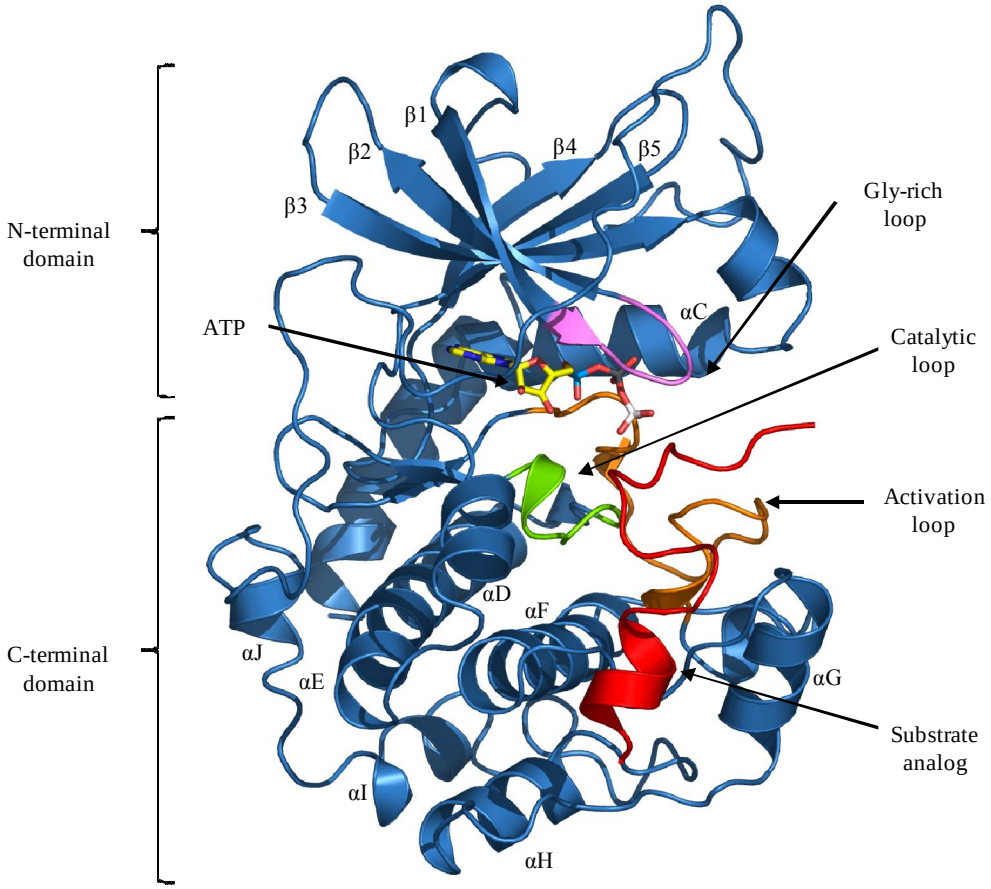


Figure 1-1 Protein kinase structure

Ribbon representation of the catalytic subunit of cAMP-dependent protein kinase A (PKA, pdb code 1atp, Knighton et al, 1991). The activation loop is colored in orange, the catalytic loop is in green, the glycine-rich loop is in pink, and the substrate peptide inhibitor is in red. ATP molecule is represented in sticks.

(Glu91 in PKA) in α-helix C (Knighton et al, 1991). This interaction ensures the proper orientation of lysine that further coordinates α- and β- phosphates of ATP (Figure 1-2).

The C-terminal domain contains both the conserved catalytic and activation loop. These two structures play critical roles in the phosphoryl transfer reaction (Figure 1-3). The catalytic loop refers to a segment between β strands 6 and 7 (Asp166-Asn171 in PKA) with the consensus sequence HRDLKxxN. Three residues (Asp166, Lys168, and Asn171 in PKA) are involved in the catalytic reaction. Asp166 functions as a catalytic base, which accepts a proton and increases the nucleophilicity of the hydroxyl group on the substrate phosphorylation. This process facilitates the transfer of the γ- phosphoryl group from ATP to protein substrate. Lys168 helps to stabilize the transition state of the catalytic reaction by partially neutralizing the negative charge of the γ- phosphate. Asn171 binds the secondary Mg^{2+} ion, which coordinates with α- and γ- phosphates of the ATP; therefore, may help to stabilize the active site organization.

The activation loop is a segment between two highly conserved sequences Asp-Phe-Gly (DFG) and Ala-Pro-Glu (APE). In PKA, the activation loop is located between residues 184-208. The length of the activation loop varies among kinases and is usually 20-30 amino acids long (Liu et al, 2006; Taylor et al, 1993). Asp184, in the DFG motif, coordinates with the primary Mg^{2+} ion that bonds with the β- and γ- phosphates of ATP. This interaction ensures the proper orientation of ATP.

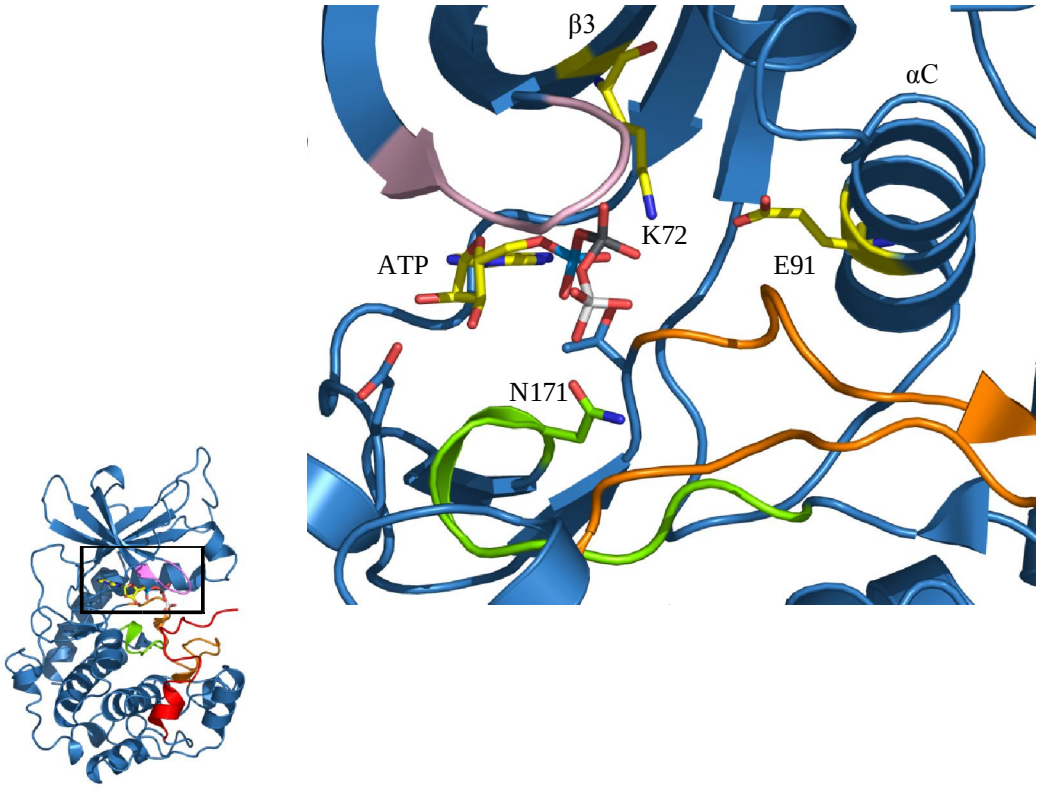


Figure 1-2 The ATP binding site in PKA

The ATP molecule and side chains of the residues involved in ATP binding are shown as sticks (pdb code 1atp, Knighton et al, 1991).

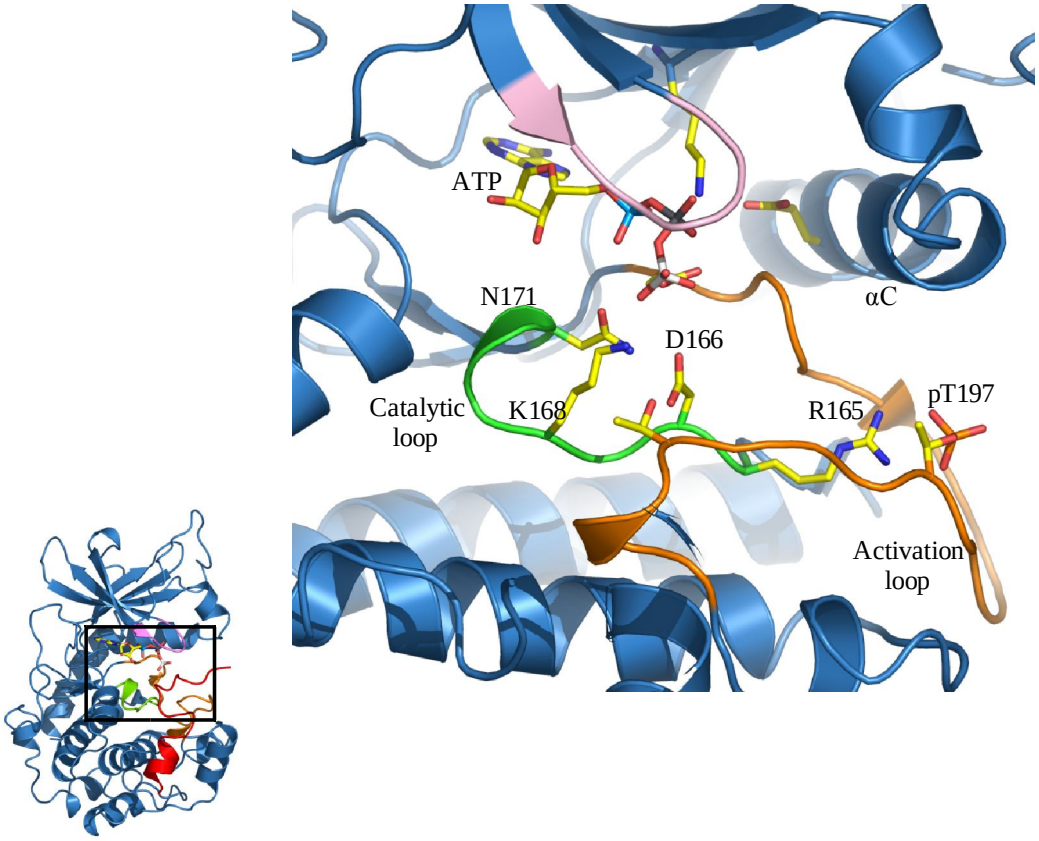


Figure 1-3 The catalytic residues of PKA

Residues involved in enzyme and substrate binding are shown as sticks (pdb code 1atp, Knighton et al, 1991).

A cluster of 7-10 residues (198-205 in PKA) located immediately upstream of the APE motif in the activation loop are important for determining the substrate specificity (Figure 1-4). These residues form a substrate-binding pocket, known as the P+1 pocket, which contacts the residue on the C-terminal side of the phosphorylation site (P+1) in the substrate. The crystal structure of PKA with the PKI substrate inhibitor peptide (TTYADFIASGRTGRRNAIHD, where the underlined Ala would be a phosphoacceptor Ser in protein substrates) revealed that the hydrophobic residue, Ile, at the P+1 position in the substrate is buried in the hydrophobic pocket formed by the side chains of Leu198, Pro202, and Leu205 from the kinase domain.

The P+1 pocket is the most conserved mechanism used in determining substrate specificity (Nolen et al, 2004). Residues within the P+1 pocket select different types of amino acids following the phosphorylation site in the substrate. For example, AGC, CAMK and STE kinases prefer a hydrophobic residue at the P+1 site in the substrate, while MAPKs and CDKs have a shallow depression at this place and thus prefer a residue that has short side chain such as proline (Kannan & Neuwald, 2004). The P+1 pocket is also a regulatory site. Refolding of this pocket has been shown in several kinases during kinase activation (Nolen et al, 2004). Comparison of available structures reveals that P+1 pocket is very flexible as it adopts multiple conformations. This corresponds to its role in controlling the binding, positioning and releasing of substrate. The P+1 pocket is formed by highly conserved interactions mediated by threonine from the activation loop (Thr201 in PKA), and two residues aspartate (Asp166 in PKA) and lysine (Lys168 in PKA) from the catalytic loop.

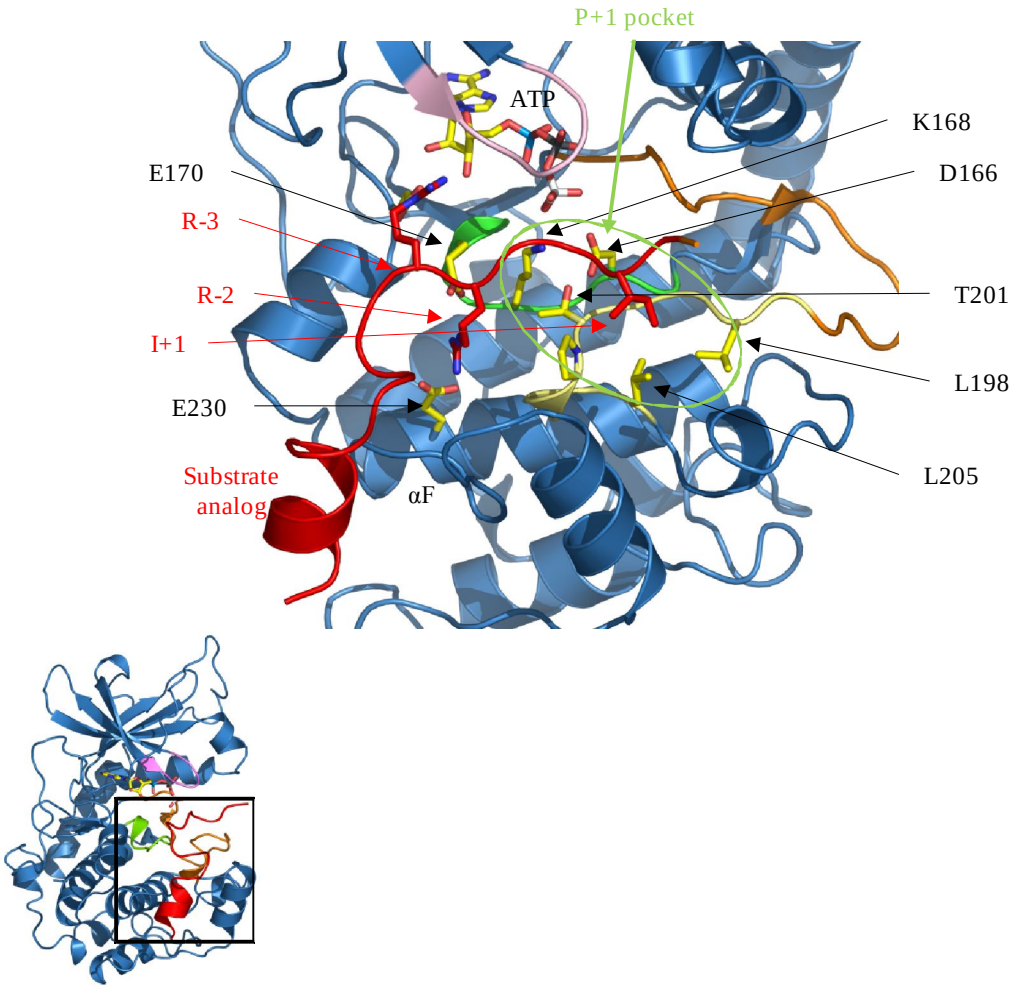


Figure 1-4 The substrate binding pocket in PKA

Residues involved in enzyme and substrate binding are shown as sticks (pdb code 1atp, Knighton et al, 1991).

The P+1 pocket is not the only region determining substrate specificity. Structural information shows that substrate and enzyme binding is mediated by multiple interactions (Brown et al, 1999; Lowe et al, 1997; Madhusudan et al, 1994). On the substrate side, residues before and after the phosphorylation site have all been found to contribute to the interaction (Brinkworth et al, 2003; Zhu et al, 2005). The consensus sequence of PKA substrates is RRxS/Tϕ (Madhusudan et al, 1994). The two arginines at P-2 and P-3 form hydrogen bonds with Glu127 in helix D, Glu170 in catalytic loop, and Glu230 in helix F. Several less specific interactions are mediated by large helix F in PKA and the N-terminus of peptide substrate. Most of them are through hydrophobic interactions.

The control of substrate binding to kinase is usually very complicated. Interaction with the active site, as described above, is usually not enough for kinase to achieve high specificity. Usually, additional docking sites are involved. These docking sites can be from either inside or outside the kinase domain. MAPK, for example, has an additional substrate docking site, known as common docking (CD) domain, located on the back surface of the kinase domain opposite from the active site (Tanoue et al, 2001). The CD domain is characterized by a cluster of negatively charged residues and a hydrophobic groove. It specifically interacts with a consensus sequence $(R/K)_{2-3}$-X_{1-6}-Φ-X-Φ (Φ, hydrophobic residue), which is found in many MAPK interacting proteins, including upstream kinases such as MEKs and downstream substrates like RSK and Elk. Several studies have suggested that docking interactions through binding sites outside of the active site may modulate kinase activity. Crystal structures of p38 and ERK2 in

complex with docking peptides showed that binding of substrate to the CD domain induces an allosteric conformational change which increases the flexibility of the activation loop leading to adopt a more active-like conformation (Chang et al, 2002; Zhou et al, 2006).

As mentioned earlier, many protein kinases possess extra sequences outside the kinase domain. Many of these sequences mediate interactions with other proteins and therefore may control the accessibility of kinase to substrate. For example, the Src-homology 2 (SH2) domain, a common kinase-associated domain, binds specifically to phospho-tyrosine residues. This sort of interactions provides an extra mechanism to control substrate selection. Moreover, the relative positions of these additional domains a may be important for substrate recognition; for example, a change in the arrangement of SH2 and SH3 domains in the Src tyrosine kinase alters its kinase activity toward certain substrates (Yadav & Miller, 2008).

2. Kinase activation mechanisms

Protein kinases exist in inactive (off) and active (on) forms. The switch between these two states is tightly controlled and therefore allows kinases to regulate signaling pathways efficiently. A comparison of currently available protein kinase structures suggests that active kinases adopt conformations that are somewhat similar, while inactive kinases display higher conformational diversity (Huse & Kuriyan, 2002). The resemblance among the active conformations may arise from the fact that all kinases

perform the same catalytic reaction, phosphorylation. The conformational differences in the inactive states suggest that various mechanisms may be used to switch kinases from inactive to active forms. It also implies that kinases may have developed multiple ways to be restricted to the inactive conformation.

Structures of several protein kinases are available in both active and inactive forms, including ERK2, p38, CDKs, and PKB/Akt (Bellon et al, 1999; Canagarajah et al, 1997; De Bondt et al, 1993; Huang et al, 2003; Jeffrey et al, 1995; Wang et al, 1997; Yang et al, 2002; Zhang et al, 1994). These snapshots of multiple states provide insights into the details of the activation mechanisms. For the majority of kinases, activation requires phosphorylation of the activation loop (Gibbs & Zoller, 1991; Johnson et al, 1996; Skamnaki et al, 1999). In the active Ser/Thr kinase structures, the phospho-residue hydrogen bonds with the conserved arginine in the catalytic loop (Knighton et al, 1991). This interaction helps to stabilize the conformation of the activation loop. Activation loops are highly flexible and are often not captured by crystal structures of the inactive states. Activation loop conformation is closely associated with kinase activity; part is because this loop plays a central role in coordinating ATP and substrate binding. The activation loop also provides interactions with the helix αC the N-terminal domain. This interaction is believed to help repositioning the helix αC during kinase activation and leading to the formation of an important salt bridge between glutamate in the helix αC and the ATP-binding lysine in the β strand 3. The interaction between activation loop and C-helix reveals another connection between N- and C- domains besides the hinge. This interaction suggests that domain movements generally come with kinase activation.

In fact, kinases usually adopt open conformations in the inactive state, and closed conformations when activated. These conformational changes during kinase activation permit the correct binding of ATP and substrate, and assure optimal phosphoryl transfer.

In the MAPKs, double phosphorylation in the activation loop is required for full activation. Structural studies on ERK2 showed that phosphorylation on Thr183 and Thr185 in ERK2 cause a large conformational change in the activation loop (Figure 1-5) (Canagarajah et al, 1997). Compared to the active state, the activation loop of ERK2 in the inactive state is comparably more flexible, with higher B-factors. Although the positions of the highly conserved DFG and APE motifs in the activation loop are well superimposed between the active and inactive conformations, the region in between is very different. In the inactive structure, the activation loop folds downward and interacts with the MAPK insertion, a 30 residue insert between α helices G and H (Zhang et al, 1994). On the contrary, in active ERK2, the activation loop is positioned closer to the N-terminal domain (Canagarajah et al, 1997). This conformational change is associated with the formation of the substrate binding pocket which allows subsequent substrate binding and phosphoryl transfer.

In addition to phosphorylation in the activation loop, the most commonly used mechanism to activate a kinase, several other mechanisms have also been proposed. These include interaction with other proteins, removal of an autoinhibitory subunit, and reformation part of structures (e.g. α helix C) in the kinase domain (Lei et al, 2000; Russo et al, 1996; Yang et al, 2002; Zhang et al, 2006). Activation of a kinase though protein-

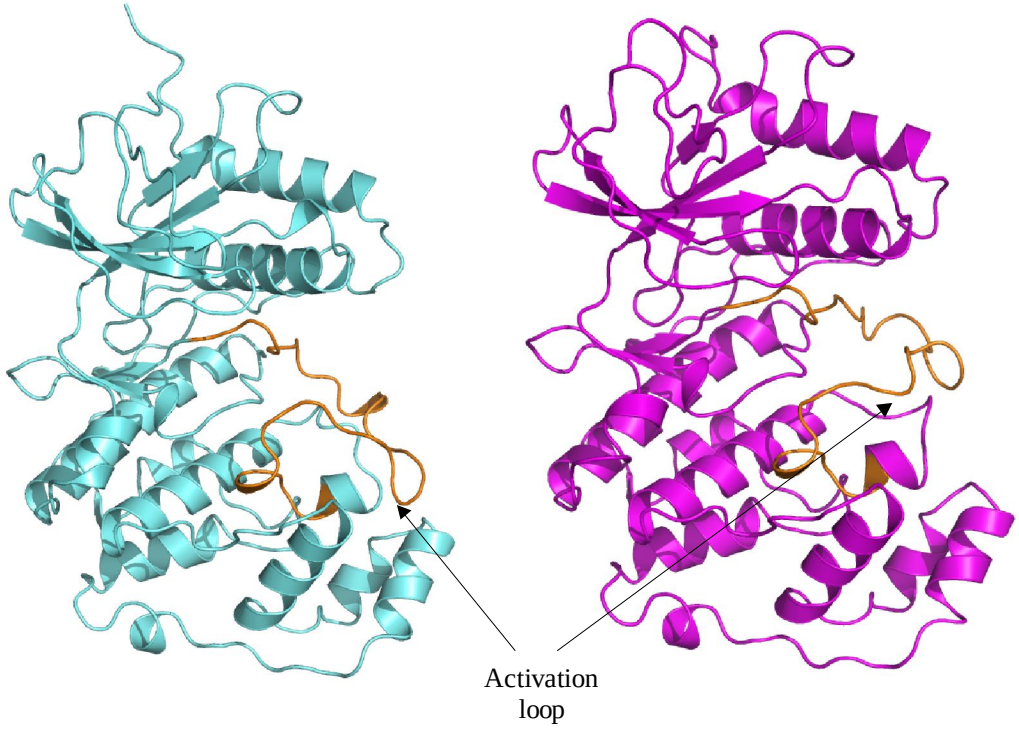


Figure 1-5 Structures of the inactive (unphosphorylated) and active (phsophorylated) ERK2

Inactive ERK2 (pdb code 1erk, Zhang et al, 1994) is colored in cyan and active ERK2 (pdb code 2erk, Canagarajah et al, 1997) is colored in magenta. The activation loops of both structures are colored in orange.

protein interactions has been shown in cyclin-dependent kinase 2 (CDK2) (Russo et al, 1996). The full activation of CDK2 requires phosphorylation of the activation loop, dephosphorylation of two residues in the glycine-rich loop, and association with cyclin (Nurse, 1990). The crystal structure of CDK2/cyclin A complex revealed that cyclin A activates CDK2 by interacting with the C-helix in the N-terminal domain (Figure 1-6). The interaction forces reorientation of the helix αC and moves it towards β strand 3, and allows Glu51 to pair with Lys33. As noted already, this ionic interaction is conserved in active kinases and essential for correct orientation of ATP. Although the interaction interface is restricted in the N-terminal domain, repositioning the C-helix also causes a conformational change in the C-terminal activation loop. The activation loop adopts a more active conformation compared to that in cyclin-free CDK2 (Jeffrey et al, 1995).

A similar mechanism has been shown to be used in certain other protein kinases. The epidermal growth factor receptor (EGFR) tyrosine kinase is one such example. However, in this case, the interaction is between two identical monomeric proteins. EGFR is a transmembrane protein with an extracellular ligand-binding domain and an intracellular kinase domain. Ligand-binding triggers receptor dimerization, or higher order oligomerization, leading to the activation of EGFR (Mohammadi et al, 1993). Unlike most protein kinases, activation of the EGFR does not require phosphorylation of the activation loop. In the crystal structure of active EGFR, the C-terminal domain of one EGFR kinase domain activates the second kinase domain by interacting with its helix αC, which facilitates salt bridge formation and proper organization of the activation loop required for catalysis. Interaction through this dimer interface is proposed to activate the

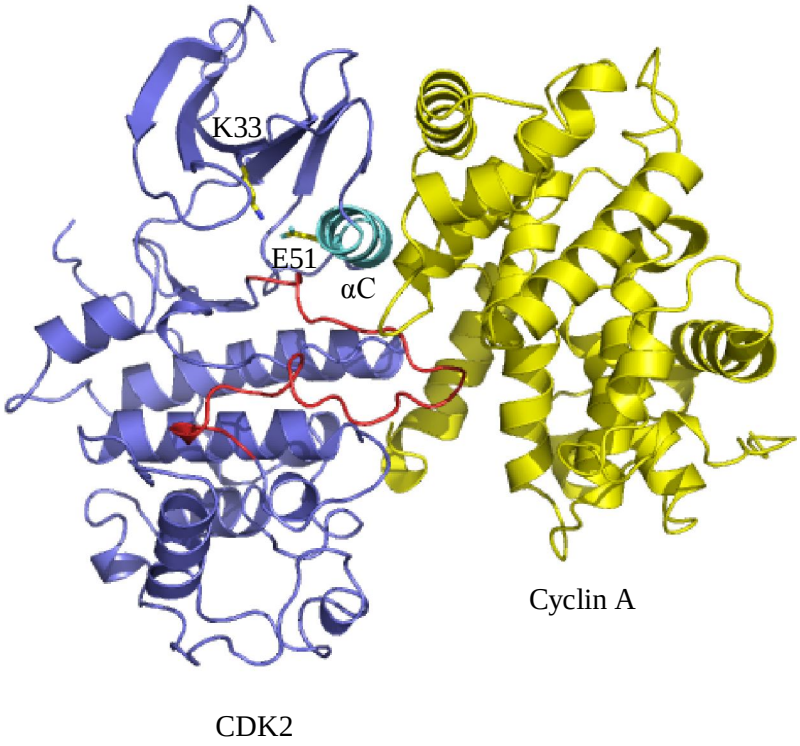


Figure 1-6 Activation of a kinase though movement of α-helix C

Ribbon representation of the structure of CDK2/cyclin A (pdb code 1fin, (Russo et al, 1996). CDK is colored in blue and cyclin is in yellow. The helix αC in the kinase is shown in cyan, and the activation loop is in red. Side chains of the lysine and glutamate involved in salt bridge formation are shown

kinase, and indeed disruption of this interface by mutagenesis significantly affects EGFR activation (Zhang et al, 2006).

Many protein kinases use autoinhibition to regulate their kinase activity. These kinases contain an autoinhibitory region outside the catalytic domain, which suppresses kinase activity. Several mechanisms have been found. Autoinhibition is believed to prevent random activation of a kinase in the absence of proper stimulation. The p21-activated protein kinase 1 (Pak1) is one of the examples (Lei et al, 2000). The activation of Pak1 requires both phosphorylation of the activation loop and binding of small G proteins, Cdc42 or Rac1 (Daniels & Bokoch, 1999; Manser et al, 1994). The crystal structure of Pak1 revealed a trans-autoinhibited dimer in which the N-terminal inhibitory domain of one Pak1 molecule of the dimer binds and inhibits the catalytic domain of the other. A portion of the inhibitory domain forms a pseudosubstrate that positions directly into the active site of the kinase domain. This causes a displacement of the C-helix and results in the breakdown of catalytic interactions (Figure 1-7). Binding of small G proteins to the inhibitory domain are proposed to release this autoinhibition and allow subsequent autophosphorylation in the activation loop.

The crystal structures of active and inactive PKB (protein kinase B) revealed another way to regulate kinase activity: through refolding part of structure (Figure 1-8). PKB has two phosphorylation sites, one located in the activation loop and the other in the hydrophobic motif located in the C-terminus of the kinase (Alessi et al, 1996). Phosphorylation of both residues is required for full activity. In the active structure, the

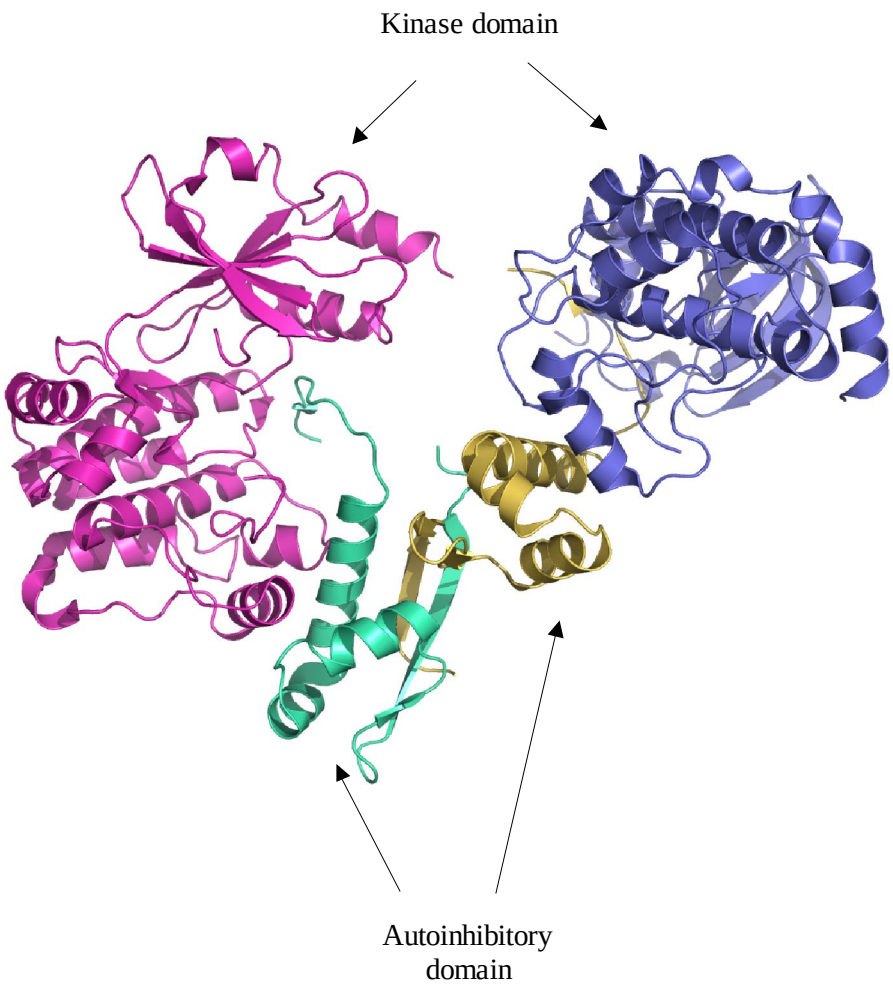


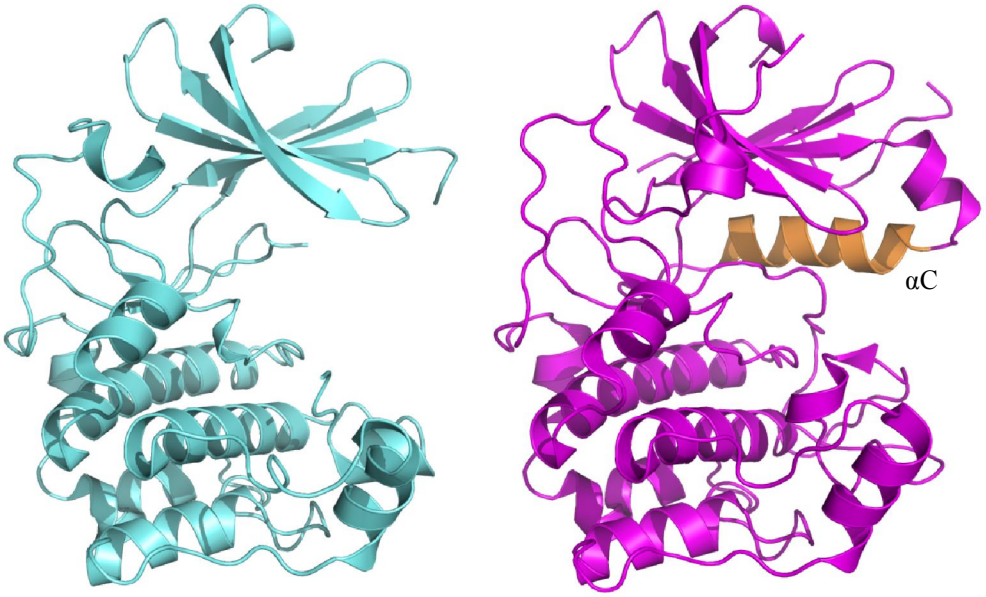
Figure 1-7 Structure of inactive PAK1

Ribbon representation of the kinase domain and the autoinhibitory domain of PAK1 (pdb code 1f3m, (Lei et al, 2000). The kinase domains are colored in magenta and blue. The autoinhibitory domains are in green and yellow.

phosphorylated hydrophobic motif folds back on to a hydrophobic groove in the N-terminal domain (Yang et al, 2002). Comparing with its inactive structure, this hydrophobic interaction is suggested to stabilize the C-helix, which is required for salt bridge formation and proper alignment of catalytic residues as discussed earlier.

Above are structural-based mechanisms for the control of kinase activity. Regulation of kinases in the context of signaling pathways is more complicated. The signal output can be affected by the subcellular localization, post-translational modification, and the complexes associated with the kinases (Quatela & Philips, 2006; Ramos, 2008). Subcellular redistribution of kinases is usually accompanied with changes in their activities. Relocalization allows kinases to phosphorylate proteins in multiple cell compartments, for example, in the nucleus phosphorylation of nuclear transcription factors controls gene expression that is required for cells to respond to the environment. The localization of protein kinases is tightly regulated; they are regulated by components of the transport machinery, including members of the importin and exportin families, cytoskeletal proteins, or through interactions between kinases and other nuclear or cytoplasmic proteins (Adachi et al, 2000; Whitehurst et al, 2004).

Many studies have demonstrated that kinase activities are regulated by post-translational modifications. As mentioned above, phosphorylation is required for the activation of most kinases. Phosphorylation is also involved in inhibiting kinase activities or modulating their functions through regulating protein-protein interactions.



Inactive PKB Active PKB

Figure 1-8 Structure of inactive and active PKB/Akt

Inactive PKB (pdb code 1mrv, Huang et al, 2003) is colored in cyan and active PKB (pdb code 1o6k, Yang et al, 2002) is colored in magenta. The α helix C in the active structure is colored in orange, whereas it is missing in the inactive PKB.

In Src kinase, phosphorylation in the C-terminal tail outside the kinase domain locks the kinase in the inactive state by intramolecular interactions and only after it is dephosphorylated can the kinase become active (Xu et al, 1997). Thus, dephosphorylation maybe equally important in determining the magnitude and duration of a kinase. Three types of phosphatases have been identified, including Ser/Thr phosphatases, Tyr phosphatases, and dual specificity phosphatases. Although protein kinase and phosphatases simply play opposite roles, how they coordinate with each other to achieve the fine tuning of signal transduction requires further investigation. Other types of modification, such as acetylation, methylation, and ubiquitination also have been reported about their roles in modulating kinase activity. For example, MAP2Ks were shown to be acetylated at the catalytic loop which prevents enzyme function (Mukherjee et al, 2006). Polyubiquitination has been found on transforming growth factor beta activated kinase-1 (TAK1), which results in the increase of TAK1 kinase activity (Adhikari et al, 2007).

Another approach to control kinase signaling is thorough regulating their associations with upstream and downstream effectors. Scaffolding proteins play essential roles in assembling components in signaling modules which enhances the fidelity and efficiency of action. The first studied scaffolding protein was in the yeast MAPK pathway. MAPK cascade is usually comprised of three sequentially activated kinases: MAP3K which activates MAP2K, which subsequently activates MAPK. In yeast, two MAPK pathways, the mating pathway and the osmolarity pathway, share the same upstream activators at the MAP4K (Ste20) and MAP3K (Ste11) levels. Signaling

specificity is achieved through scaffolding proteins. Signaling specificity is achieved through scaffolding proteins. In the mating pathway, Ste5p assembles the Ste11p (MAP3K) /Ste7p (MAP2K) /Fus3 (MAPK) module, while the osmolarity pathway scaffold Pbs2 is the MAP2K and recruits Ste11p (MAP3K) and Hog1 (MAPK) (Elion, 2001).

IV. Ste20 kinases

The sterile 20 (Ste20p) protein kinase is a Ser/Thr kinase first identified as a yeast MAP4K protein that regulates the MAPK pathway in response to mating pheromones (a factor and α factor) (Leberer et al, 1992). Ste20p is activated by the heterotrimeric G-protein β and γ subunits (Gβγ) when cells are exposed to mating pheromone. Active Ste20p then activates the MAPK cascade: MAP3K, MAP2K, and MAPK leading to increase the expression of a set of genes required for the mating process, such as those involved in G1 cell cycle arrest. Aside from the pheromone-induced mating pathway, yeast Ste20 kinases also participate in the filamentation and the osmotic responsive pathways (Madhani et al, 1997; Raitt et al, 2000).

There are more than thirty Ste20 protein kinases in the mammalian genome. Ste20 kinases are characterized by the presence of diverse regulatory domains fused to the kinase domain. Such diversity allows them to function in numerous signaling pathways. Based on the arrangement of their regulatory domains, Ste20 kinases are divided into two subfamilies, the p21-activated kinases (PAKs), with C-terminal kinase

domains, and the germinal center kinases (GCKs), with N-terminal kinase domains. These two subfamilies are further separated into 10 groups: PAK-I and II and GCK-I to VIII. PAKs are distinguished by containing a kinase domain at C-terminus and an N-terminal regulatory domain responsible for small GTPase binding. Although PAKs have been shown to function as MAP4Ks in ERK, JNK and p38 pathways, most studies have focused on their roles in cytoskeletal rearrangement, cell motility, and apoptosis (Chaudhary et al, 2000; Daniels & Bokoch, 1999; Deacon et al, 2003; Pandey et al, 2002). In contrast, GCKs have an N-terminal kinase domain, and their C-termini consist of a wide variety of sequences usually used for protein-protein interaction. Functions for GCKs are very diverse. Several GCKs participate in MAP kinase signal cascades and they can be either MAP4Ks or in the case of TAO1-3 MAP3Ks (Dan et al, 2001). They are involved in apoptosis, cell migration, proliferation and osmotic stress response (Anselmo et al, 2006; Ponce-Coria et al, 2008). Several GCK proteins remain poorly characterized.

V. Ste20-related proline-alanine-rich kinase (SPAK)

1. Overview of SPAK

a. Discovery of SPAK

Ste20-related proline-alanine-rich kinase (SPAK), also called PASK in rat and STK39 in GeneBank, was first isolated from a transformed rat pancreatic beta cell line (RIN-5AH) as a kinase involved in beta-cell development (Johnston et al, 2000; Ushiro et

al, 1998). The kinase was cloned by PCR amplification using degenerate oligonucleotide primers that recognize kinase subdomains VIb and IX (DeAizpurua et al, 1997).

Sequence homology analysis of the SPAK catalytic domain classified it as a member of the Ste20 kinase family; thus raising the possibility that SPAK may function as a MAP4K (Johnston et al, 2000). SPAK is most similar to OSR1 (oxidative stress-responsive kinase 1), showing 73% sequence identity overall and almost 90% identity in the kinase domain. As the protein sequence and domain organization of SPAK and OSR1 display modest similarities with other Ste20 kinases, SPAK and OSR comprise a distinct subfamily, the GCK (germinal center kinase)-VI subfamily. SPAK/OSR1 orthologs have been found in organisms from worm to human but none homologs have been found in unicellular organisms or plants. Sequence analysis of all SPAK/OSR1 orthologs indicated that OSR1 is the prototypical kinase while SPAK has evolved by gene duplication, as OSR1 is conserved from worm to human but SPAK is only present in vertebrates (Delpire & Gagnon, 2006).

b. Tissue distribution and subcellular localization of SPAK

Both SPAK and OSR1 are widely expressed in tissues and cell lines as determined by Northern and Western analyses, suggesting roles in many organs (Johnston et al, 2000; Piechotta et al, 2003; Ushiro et al, 1998). A comparison of mRNA expression levels between these two homologs reveals that they have different tissue distributions. For example, SPAK is highly expressed in brain, salivary gland, pancreas,

adrenal gland and testis, and less in heart, lung, kidney, stomach, intestine, ovary, thymus, spleen, and skeletal muscle. OSR has highest expression in heart and skeletal muscle, followed by liver, small intestine and colon. This difference suggested that SPAK and OSR1 might play distinct roles in the body. However, this tissue distribution difference is based on mRNA expression. A direct comparison between SPAK and OSR1 expression at the protein level is difficult as antibodies are not yet available that to distinguish these two highly related proteins.

Immunohistochemical analysis confirmed that SPAK/OSR1 have wide tissue distribution but appear to be expressed only in a subset of cells within a particular tissue (Ushiro et al, 1998). For example, in the kidney, SPAK/OSR1 are highly expressed in principal cells of the collecting duct, epithelial cells of the distal tubule, but not detected in intercalated cells of the collecting duct, epithelial cells of the proximal tubule and glomerulus. In the brain, SPAK and OSR1 are rich in neurons and in transporting epithelial cells, which transport ions and water to regulate the ionic composition of cerebrospinal fluid. Additionally, SPAK/OSR1 proteins are also abundantly expressed in other transporting epithelia in different tissues such as epithelial cells of the sublingual gland, and parietal cells of the stomach. The association of SPAK and OSR1 with transporting epithelial cells provided the first evidence that they maybe involved in the regulation of electrolyte and fluid homeostasis.

Early studies on SPAK were focused on its involvement in pancreas development. In situ hybridization analysis detected the expression of SPAK mRNA in gut and pancreatic epithelia as early as embryonic day 12 (E12). By embryonic day 15

(E15), mRNA transcripts are predominantly found in cells that will eventually become exocrine acinar cells (Miao et al, 2000). However, no further biochemical and functional studies have been done on SPAK in this aspect.

At the cellular level, immunofluorescence analysis demonstrated that endogenous SPAK/OSR1 proteins are mostly expressed in the cytoplasm of epithelial cells at rest, but are translocated to nucleus under some circumstances, such as loss of cell-cell contact, and chronic exposure to osmotic stress (unpublished data). These results suggested that the kinase is translocated from cytosol to nucleus in response to cellular stress; nevertheless, stimuli and mechanisms underlying this relocalization require further investigation. Subcellular fractionation studies also showed that SPAK/OSR1 proteins exist mostly in the cytosol and less in the Triton X-100-insoluble cytoskeletal fraction under normal conditions, whereas their presence in the insoluble fraction is largely increased under osmotic, oxidative and high temperature stress (Ushiro et al, 1998).

2. Functions of SPAK

a. Stimulus of SPAK

In order to understand SPAK/OSR1 signaling, different agents were tested to determine their ability to activate SPAK/OSR1. The activity of immunoprecipitated endogenous OSR1 was determined in immune complexes using autophosphorylation or substrate phosphorylation after HeLa cells were treated with sorbitol, NaCl, anisomycin, okadaic acid, serum, nocodazole, Taxol, H_2O_2, phorbol ester, and epidermal growth

factor (Chen et al, 2004). Among these stimuli, only sorbitol or NaCl activated osmotic stress were able to activate SPAK/OSR1, which may imply functions in regulating ion homeostasis in response to changes in cell volume.

b. Regulation of ion transport

SPAK was identified from yeast two-hybrid screening with a cation-chloride cotransporter KCC3 (K^+-Cl^- cotransporter) (Piechotta et al, 2002). Based on its interactions with several ion cotransporters such as KCC, NKCC1 (Na^+-K^+-$2Cl^-$ cotransporter), NKCC2 and NCC (Na^+-Cl^- cotransporter), it is believed that a fundamental function of SPAK/OSR1 is to regulate ion balance (Figure 1-9). SPAK/OSR1 has received considerable attention after it was identified as a downstream target for the hypertension related kinases WNK1 and WNK4 (Anselmo et al, 2006). WNK (with no lysine (K)) kinases are Ste20-related Ser/Thr protein kinases that are also activated by osmotic stress (Lenertz et al, 2005). To date, four WNK kinases, WNK1-4, have been identified in the human genome (Verissimo & Jordan, 2001; Xu et al, 2000). All WNK kinases have a unique catalytic site organization, with the ATP binding lysine located in β2 rather than in β3 as it is in all other protein kinases. Genetic linkage analysis showed that mutations in two WNK kinases, WNK1 and WNK4, were associated with a heritable form of hypertension, pseudohypoaldosteronism type II (PHAII) (Wilson et al, 2001). PHAII, also known as Gordon's syndrome, is characterized by hyperkalemia (high serum K^+), hyperchloremia (high serum Cl^-), suppressed plasma

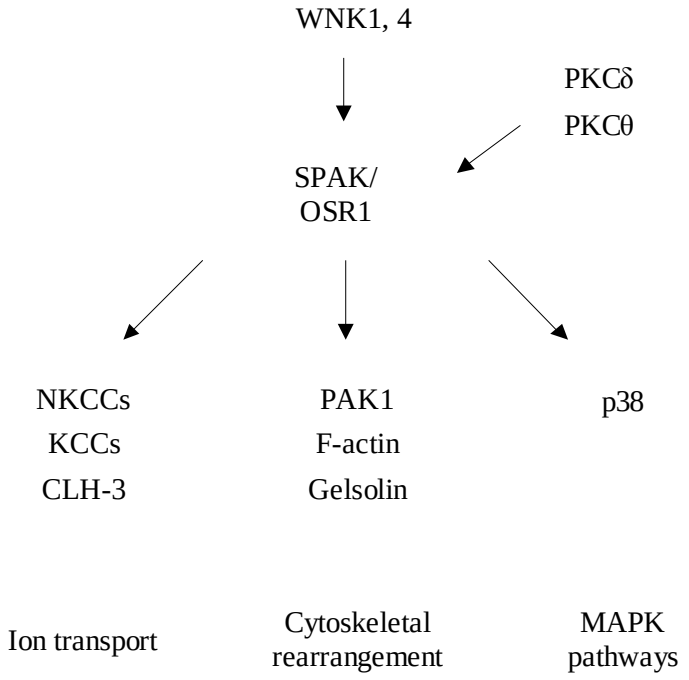


Figure 1-9 Functions of SPAK and OSR1

SPAK and OSR1 are activated by WNK kinases and regulate ion transport through Na^+-K^+-$2Cl^-$ cotransporters, K^+-Cl^- cotransporters, and Cl^- channels. Other SPAK and OSR1 interacting proteins that have been reported include PKCθ, PKCδ, PAK1, F-actin, gelsolin, and p38. See text for references.

renin activity, and reduced bicarbonate. The WNK1 mutations associated with disease are large intronic deletions, believed to lead to overexpression of this gene. On the other hand, WNK4 mutations were all caused by point mutations in the coding sequence and affected the normal function of this protein; however, their underlying mechanisms are not clear.

Hypertension, or high blood pressure, is a major health problem that leads to higher chance for stroke, heart attack and kidney failure. The primary function of the kidney is to maintain the volume and electrolyte composition of body fluids by filtration, reabsorption, and secretion. Adult kidneys filter 150 to 180 liters of blood in a day and reabsorb more than 99% of the filtered sodium. While the major sodium reabsorption is carried out in the proximal tubule, the fine regulation of sodium reabsorption takes place in the distal nephron and collecting duct where ion channels and cotransporters play important roles. Excessive salt reabsorption in this segment due to malfunction of channels or cotransporter was suggested to increase extracellular fluid accumulation (as a result of maintaining plasma osmolarity) and ultimately increase blood pressure. The specific causes of hypertension are usually not known. It is likely that both genetic and environmental factors are involved. The kidney plays a critical role in the long-term regulation of blood pressure. Genetic studies have also linked several renal disorders to hypertension. These include Liddle's syndrome, Gittleman's syndrome, and Gordon's syndrome (Rossier, 2003).

One mechanism for WNK regulation of blood pressure is mediated by SPAK/OSR1. The current model is that WNK1 and WNK4 interact with and

phosphorylate SPAK/OSR1 leading to its activation. Activated SPAK/OSR1 phosphorylates and subsequently regulates activities of ion cotransporters.

A recent study from Cobb lab uncovered the interaction between SPAK/OSR1 and WNK1 by yeast two-hybrid screening and confirmed the existence of this interaction in cells by co-immunoprecipitation and mass spectrometry (Anselmo et al, 2006). In this study, they also showed that the C-terminal portion of WNK1 contains several SPAK/OSR1-binding motifs (R/K)FX(V/I) that are responsible for this interaction. In vitro kinase assays demonstrated a direct phosphorylation and activation of OSR1 by WNK1. Both RNAi depletion and overexpression of catalytically inactive WNK1 in cells resulted in significant decreases in OSR1 activity. The physiological link between WNK1 and NKCC was provided by the reduction of NKCC activity in HeLa cells in which WNK1 and/or OSR1 were knocked down by RNAi. In the case of WNK4, several groups also demonstrated that WNK4 uses same mechanism as WNK1 to regulate SPAK/OSR1. Reconstitution studies in *Xenopus* oocytes showed that coexpression of SPAK and WNK4 is necessary to activate NKCC under both isosmotic (basal) and hyperosmotic (stimulated) conditions. Furthermore, expression of kinase dead WNK4 or SPAK or a SPAK/OSR1-binding motif mutant of WNK4 resulted in a decrease of NKCC stimulation, suggesting that both kinases activities and protein-protein interaction are required for NKCC activation (Gagnon et al, 2006a).

As mentioned earlier, activation of SPAK/OSR1 is through phosphorylation by WNK. Analysis of WNK phosphorylation sites by mutagenesis has been conducted in our lab (Anselmo et al, 2006). Since phosphorylation of residues within the kinase

activation loop is usually responsible for activation of protein kinases, our phosphorylation site mapping was focused in this segment. There are three Thr residues, 173, 178 and 185, in the OSR1 activation loop. In mouse SPAK, these residues are Thr231, Thr236, and Thr243. To identify residues responsible for WNK phosphorylation and activation, these three sites were mutated individually or together. Mutation of Thr185 resulted in a significant decrease in OSR1 kinase activity while mutations in Thr173 and Thr178 only caused slight reductions. This indicated that Thr185 is essential for OSR1 activation. However, the roles of Thr173 and 178 are less clear. Phosphorylation of Thr173 and 178 may be required for full activation of OSR1, but we also cannot rule out the possibility that this small reduction is due to unexpected conformational changes from mutations. In addition, mutation of Thr185 also reduced, but not abolished, WNK1 phosphorylation, suggesting that WNK1 may phosphorylate multiple sites in OSR1.

Using mass spectrometry, a different group identified two WNK1 phosphorylation sites in mouse SPAK, Thr243 within the activation loop and S383 in the C-terminal non-catalytic region (Vitari et al, 2005). Thr243 is common in both studies, and mutation of Thr243 to alanine also prevented SPAK activation. In contrast, mutation of S383 did not affect WNK1-specific activation of SPAK. The importance of Thr243 phosphorylation in this pathway was further supported by *Xenopus* oocyte reconstitution experiments showing that coexpression of WNK4 with SPAK (T243A) reduced NKCC1 activation in comparison to the wildtype (Gagnon et al, 2006a).

Like many protein kinases, autophosphorylation in SPAK/OSR1 has been described. However, autophosphorylation is not enough for SPAK/OSR1 to reach the maximum catalytic activity for its substrate. One study reported that two residues within the activation loop of SPAK, Thr243 and Thr247, are autophosphorylation sites (Gagnon et al, 2006a). Using in vitro kinase assays, they demonstrated that mutation of Thr243 or Thr247 to alanine decreased the level of autophosphorylation. Nevertheless, both residues have roles in catalysis. Thr243, as described above, is required for kinase activation. Thr247 is conserved in the protein kinase family and participates in forming the substrate-binding pocket. It is not clear whether the reduction of autophosphorylation found in these two mutants is due to loss of kinase catalytic activity. The identities of autophosphorylation sites and their function remain to be determined.

Regulation of Sodium Potassium Chloride Cotransporters

As mentioned above, interactions between SPAK/OSR1 and NKCC1 and NKCC2 were identified by yeast two-hybrid screening, and the SPAK/OSR1-binding sequence (R/K)FX(V/I), located in the N-terminal cytoplasmic region of the cotransporters, is required for the binding (Piechotta et al, 2002). Both NKCCs are sodium potassium chloride cotransporters that transport one sodium, one potassium, and two chlorides into the cell. NKCC1 is widely distributed in tissues and cells and has many functions, including cell volume regulation, chloride secretion in secretory tissues, and potassium and chloride homeostasis (Haas & Forbush, 2000). In contrast, NKCC2 expression is more restricted, and is found predominantly in the thick ascending limb of

Henle (TAL) of the renal tubule (Lytle et al, 1995). NKCC2 a major component is responsible for sodium reabsorption. Loss-of-function mutations in NKCC2 were linked to Bartter's syndrome, featuring hypokalemia, hypercalciuria, and low blood pressure (Simon et al, 1996).

Activation of NKCC1 and 2 are regulated by phosphorylation. Using mass spectrometry and Edman sequencing analyses, three sites corresponding to Thr206, Thr211, and Thr224 in mouse NKCC1 were shown to be phosphorylated in the shark NKCC1 (Thr184, Thr189, and Thr202) when shark rectal gland tubules were treated with forskolin (Darman & Forbush, 2002). Identification of SPAK/OSR1 as one of the upstream kinases of NKCCs was done by another group who used mass spectrometry to show that three residues Thr197, Thr201, and Thr206 in mouse NKCC1 were phosphorylated in vitro by SPAK/OSR1 (Vitari et al, 2006). Thr206 is the only residue common in the two studies. Subsequent functional studies suggested that T206 and T211 are major phosphorylation sites responsible for cotransporter activation (Gagnon et al, 2007). Phosphorylation of these two sites increases NKCCs activity. However, whether or not T211 can be phosphorylated by SPAK/OSR1 has not been determined.

Regulation of Potassium Chloride Cotransporters

Compared to NKCC1 and 2, the regulation of potassium chloride cotransporters KCC activity by SPAK and OSR1 has been less studied. KCCs are that transport one potassium and one chloride out of cells. There are four mammalian KCC homologs,

KCC1-KCC4 (Mercado et al, 2004). A recent report using *Xenopus* oocyte reconstitution assays demonstrated that SPAK deactivates KCC2 in a kinase-dependent manner. Evidence was based on the increase of KCC2 activity when coexpressed with kinase-dead, but not the wildtype, SPAK (Gagnon et al, 2006b).

Regulation of Chloride Channels

GCK-3 is a SPAK/OSR1 homolog in *Caenorhabditis elegans*. A study reported that GCK-3 interacts with CLH-3, a cell-cycle dependent and cell volume-regulated chloride channel, via the C-terminal (R/K)FX(V/I) binding motif. GCK-3 was also shown to phosphorylate and inhibit CLH-3 activity (Denton et al, 2005). Furthermore, the single *C. elegans* WNK kinase, WNK-1, was shown to interact with GCK-3 and these two proteins function in a same signaling pathway. These studies suggested a conserved mechanism used in regulation of cell volume.

c. SPAK as upstream of MAPKs

Because SPAK/OSR1 are members of the Ste20 kinase family, we were interested to understand the functional link between SPAK/OSR1 and MAPK pathways. Knocking down the *Drosophila* homolog of mammalian SPAK/OSR1, Fray, showed a significant decrease in sorbitol-induced JNK activation in *Drosophila* S2 cells as determined by phospho-JNK immunoblotting (Chen et al, 2004). In the contrast,

mammalian OSR1 was not able to activate any MAPK pathways, including ERK2, JNK, p38, and ERK5. Studies were done by overexpressing OSR1 with individual MAPKs in HEK cells and measuring the ability of immunoprecipitated MAPKs to phosphorylate their substrates. In a separate study, overexpression of SPAK was found to activate the p38 pathway in COS7 cells (Johnston et al, 2000). Thus, it remains unclear whether SPAK or OSR1 participate in MAPK signaling. This discrepancy suggested that SPAK and OSR1 might each have distinct functions in this regard. On the other hand, this difference might due to the difference in cell type, conditions, mode of activation, or the presence of accessory proteins required for the activation.

d. SPAK and cytoskeletal rearrangement

PAK1 (p21-activated kinase 1) is an SPAK/OSR1 interacting protein identified by yeast two-hybrid screening (Chen et al, 2004). OSR1 was shown to phosphorylate Thr84 of PAK1 located in the N-terminal regulatory domain containing a Cdc42/Rac1 interactive binding motif (CRIB). Association of GTP-bound Cdc42 or Rac1 with the PAK1 CRIB motif releases the PAK1 autoinhibitory conformation, which is formed by interaction between the PAK1 kinase domain and its N-terminal regulatory domain (Lei et al, 2000). The binding of small G proteins leads to activation of PAK1. The authors further showed that phosphorylation of PAK1 by OSR1 desensitizes the ability of PAK1 to be activated by small G proteins. The conclusions were based on the observation that mutation of Thr84 to glutamate significantly reduced PAK1 activation by Cdc42, while substitution to alanine had no effect (Chen et al, 2004). As PAK1 has crucial roles in

cytoskeletal rearrangement, phosphorylation of PAK1 by SPAK/OSR1suggests that SPAK/OSR1 might participate in cytoskeletal regulation.

Several interactors of SPAK/OSR1, such as F-actin and gelsolin, also have roles in cytoskeletal rearrangement (Delpire & Gagnon, 2006; Tsutsumi et al, 2000). F-actin is the key component of microfilaments of the cytoskeleton. Gelsolin is an actin binding protein, which regulates actin filament assembly and disassembly. However, activity of SPAK/OSR1 on the cytoskeleton has not been identified and mechanisms underlying any such activation of cytoskeletal rearrangement in response to osmotic stress still need to be examined.

e. SPAK/OSR1 knockout model

Knocking out *Drosophila* Fray, the homolog of mammalian SPAK/OSR1, has results in larval lethality during early development (Leiserson et al, 2000). These null fray larvae carried a phenotype of nerve swelling and axonal defasciculation that is usually associated with abnormal ensheathment of axons by glial cells. This finding implies a functional role of Fray in nerve development. Moreover, the authors were able to rescue the phenotype by introducing the rat SPAK homolog (PASK), indicating functional conservation between the mammalian protein and its *Drosophila* counterpart.

Both OSR1 and SPAK knockout mice have been generated (Delpire & Gagnon, 2008). Similar to the finding from *Drosophila*, OSR1 knockout mice died early in development. Surprisingly, the SPAK-deficient mice were born normally and have no

obvious phenotype; however, whether they are defective in responses to stress, such as high osmolarity, is not known. This observation suggests that OSR1 certainly plays a much more important role during mouse development. The fact that SPAK could not compensate for the loss of OSR1 implies that SPAK and OSR1 have distinct roles at least during development, despite the fact that the two share high sequence similarity. Differences between SPAK and OSR1 in tissue distribution and relative abundance might account for this opposite outcome, if SPAK is not yet expressed when embryos die. However, this opposite result might also be due to different gene deletions in the SPAK and OSR1 knockout mice. Using the gene trapping technique, SPAK knockout mice lack the kinase domain, while the OSR1 knockout mice only lack the C-terminal interacting domain. As a result, the OSR1 kinase domain is intact and expressed. Therefore, it might be able to phosphorylate substrate but in an unregulated way, leading to serious developmental defects.

f. Other functions of SPAK

SPAK and OSR1 were also identified from yeast two-hybrid screening with Protein kinase C-theta (PKCθ) (Li et al, 2004) and Protein kinase C-delta (PKCδ) (Smith et al, 2008). Both PKCθ and PKCδ were suggested to activate SPAK through phosphorylation. Ser321 located in SPAK kinase domain (in the loop connecting helix G and helix H) but not within the activation loop was identified as a phosphorylation site by in vitro kinase assay. However, how does phosphorylation of this atypical site lead to activation of SPAK is not known. The study also showed that SPAK activity is required

for subsequent activation of AP-1 transcription factor during T-cell activation. Both knocking down SPAK and expressing kinase-dead SPAK reduced PKCθ and TCR/CD28-induced AP-1 activation (Li et al, 2004). Aside from the observation that SPAK/OSR1 translocates to the nucleus in response to cellular stresses, this finding further supports the idea that SPAK/OSR may play roles in regulating gene transcription.

A recent report showed that a common single nucleotide polymorphism (SNP) in SPAK is associated with autism (Ramoz et al, 2008). This result infers that SPAK might have roles in brain development. Nevertheless, there are no further biochemical and functional insights into this connection.

VI. Description of the thesis

This dissertation is focused on determining the crystal structure of SPAK, a substrate of the hypertension-associated WNK kinase family. Although a substantial amount of studies on protein kinase structure have been conducted, little was known about the structure of the Ste20 kinases, particularly the GCK subfamily which SPAK belongs to. No GCK-VI structures were available. Here, we solved the crystal structures of two SPAK mutants: SPAK 63-390 (Thr243Asp) and SPAK 63-390 (Val235Cys/Thr243Asp). Both structures contain a mutation on WNK phosphorylation site (Thr243) to mimic the phosphorylated state, and the latter one has an additional mutation in the activation (Val235Cys) to improve the stability. All together, these results provide structural insights into the activation and regulation mechanisms of

SPAK. The structural information will also provide aids in identifying specific drugs for hypertension. The dissertation contains following chapters:

Chapter 1 Introduction; includes the general information about protein kinases and the background of SPAK.

Chapter 2 X-Ray Crystal Structure Determination of the Mouse SPAK 63-390 (T243D); describes the structure of SPAK 63-390 (T243D).

Chapter 3 Biochemical and Structural Characterization of Mutations in the Activation Loop of SPAK, including biochemical analysis of different activation loop mutants and reveals the structure of SPAK 63-390 (V235C/T243D).

Chapter 4 Expression, Purification and Crystallization of Mouse SGK1.

Chapter 5 Expression, Purification and Crystallization of ERK2 (E320K).

Chapter 6 Conclusions and Future Directions.

CHAPTER TWO

X-RAY CRYSTAL STRUCTURE DETERMINATION OF THE MOUSE SPAK 63-390 (T243D)

Introduction

SPAK is a member of the Ste20 GCK-VI subfamily of protein kinases. A notable feature of Ste20 kinases is the presence of structurally diverse noncatalytic domains in addition to a conserved kinase domain (Dan et al, 2001). SPAK encodes a protein of 556 amino acids. Domain prediction on SPAK only revealed a kinase domain located between amino acids 75-350 (Figure 2-1). Nevertheless, sequence alignment of SPAK and other Ste20 GCK-VI members, including mammalian OSR1, *Drosophila* Fray, and *C. elegans* GCK-3 homologs, identified two other conserved regions outside the kinase domain, referred to as PF1 and PF2 domains (Figure 2-2) (Leiserson et al, 2000). (PF stands for "PASK" and "Fray"). PF1 in mouse SPAK is located within residues 350-402, while PF2 is within residues 463-556. Further conservation in SPAK and OSR1 is found around the PF1 domain and extends from residues 350 to 421. As such, the PF1 domain appears to be part of the kinase domain of these kinases, as deletion of the PF1 domain results in a significant decrease in kinase activity towards both MBP and physiological substrates, and a loss of protein stability (Chen et al, 2008). The PF2 domain is involved in protein-protein interactions. Proteins binding to the PF2 domain include upstream kinases, i.e. WNK1 and WNK4, and downstream substrates such as

NKCC1, NKCC2, and KCC3 (Anselmo et al, 2006; Piechotta et al, 2002). Sequence alignment of binding proteins identified by yeast two-hybrid or co-immunoprecipitation experiments revealed a consensus motif [(R/K)FX(V/I)]. Mutagenesis combined with *in vitro* binding experiments indicated that this motif binds the PF2 domain.

Genome-wide searching did not reveal homology of either the PF1 or PF2 domains outside of the Ste20 GCK-VI kinase subfamily. Nevertheless, the WNK auto-inhibitory domain, a stretch of 50 amino acids located C-terminal to the WNK kinase domain, shows approximately 20% sequence identity to the N-terminal segment of the PF2 domain (amino acids 490-519) (Delpire & Gagnon, 2008). Sequence conservation between these two regions implies that SPAK might be able to regulate WNK kinase activity. This possibility suggests that there may be another function of SPAK. However, this possibility remains to be examined.

SPAK and its mammalian homolog OSR1 share about 73% sequence identity, but possess more than 90% sequence identity within their N-terminal kinase domains. Unlike OSR1, which only has 17 residues in front of the kinase domain, SPAK has 74 residues. This N-terminal extension in SPAK is rich in alanine and proline and contains several proline-alanine repeats; this region is also called a PAPA box (Johnston et al, 2000). Biochemical studies suggested that this region is not required for kinase activity or substrate phosphorylation. Indeed, an alternative splice variant lacking the PAPA box was identified from inflamed intestinal tissues, and this isoform retained its kinase activity (Yan et al, 2007).

A stretch of basic residues, RAKKVRR, similar to the canonical nuclear localization signal (NLS) of the SV40 T antigen is found in the PF1. However, functional studies on this putative NLS have not yet been reported. A putative caspase-3 cleavage site (DEXD) is present at the end of the PF1 domain. Studies have demonstrated that a mutant corresponding to caspase-cleaved SPAK, lacking the C-terminus of SPAK, is expressed predominantly in the nucleus (Ushiro et al, 1998). Furthermore, this cleavage is induced by TRAIL, suggesting a possible role for SPAK in apoptosis (Polek et al, 2006).

SPAK is phosphorylated and activated by WNK family protein kinases. Two WNK family members, WNK1 and WNK4, have been associated with a form of hypertension. Thus, studies done on SPAK will gain insights into the mechanisms of blood pressure regulation. Solving the structure of SPAK will provide further understanding of the catalytic and regulatory mechanisms of this kinase. Here, we have determined the crystal structure of a partially active form of SPAK 63-390 (T243D) at a 2.5 Å resolution. The structure reveals a unique domain-swapped feature which suggests a novel regulatory mechanism that could be used in SPAK. With the availability of the inactive OSR1 structure, we further compare the structural differences between SPAK and OSR1. Such analysis allows us to understand the conformational changes associated with the activation of SPAK. This three-dimensional structural information on SPAK may also provide aids in the rational design of specific inhibitors for hypertension.

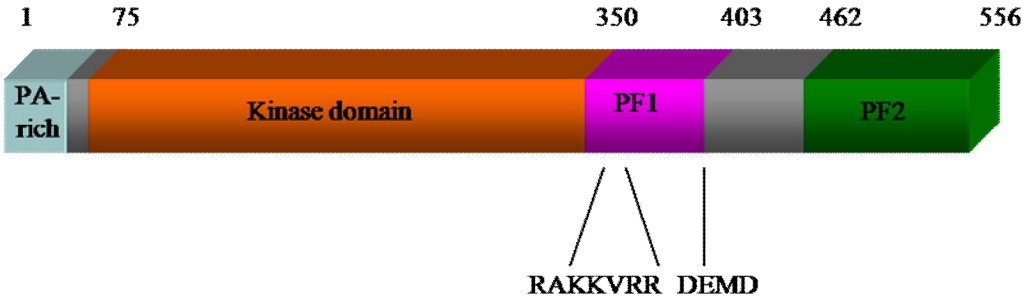
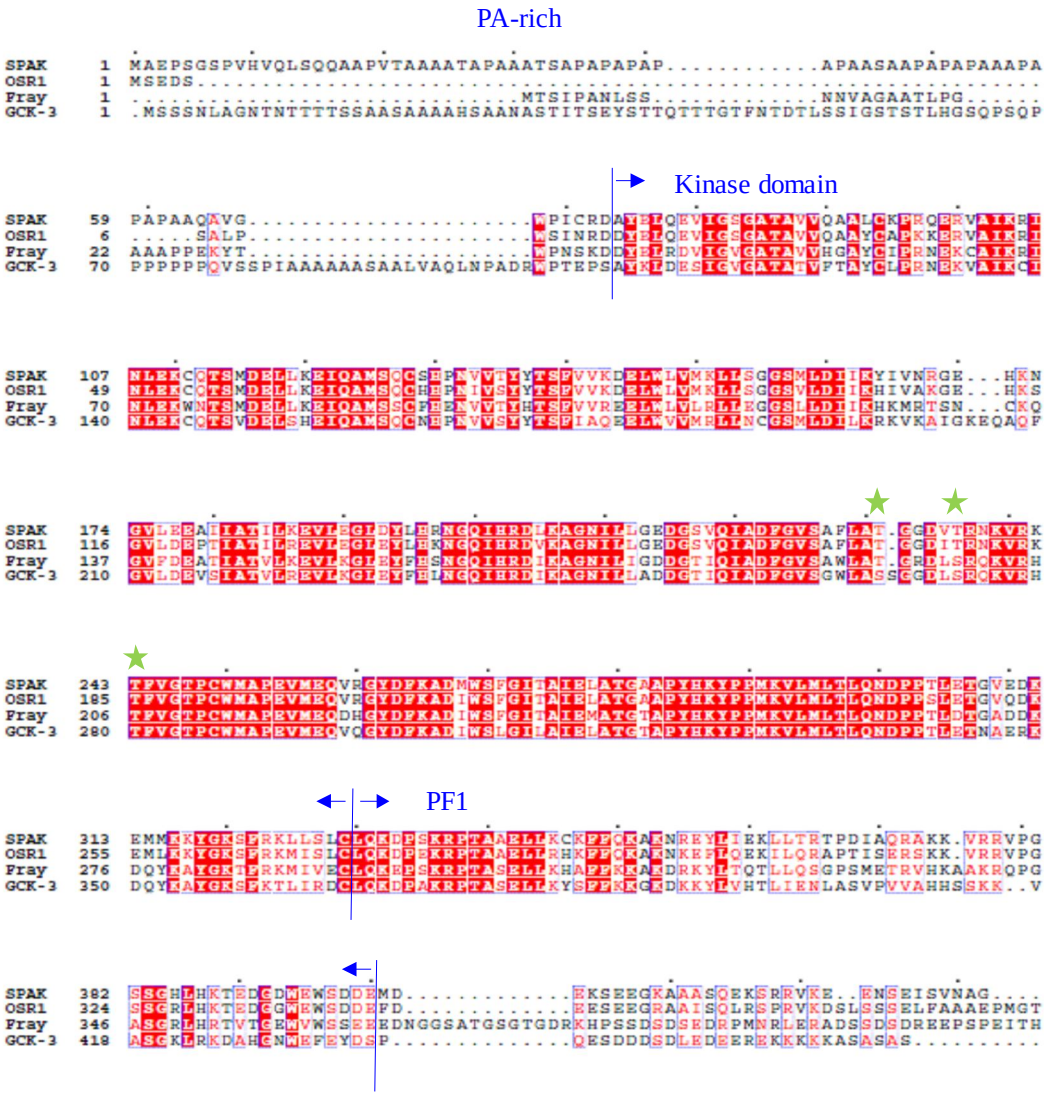


Figure 2-1 Domain organization of mouse SPAK.

The kinase domain (residue 75-350) is shown in orange. The PF1 domain (residue 350-403) is shown in magenta. The PF2 domain (residue 462-556) is shown in green.


PA-rich
Kinase domain
PF1
SPAK
OSR1
Fray
GCK-3

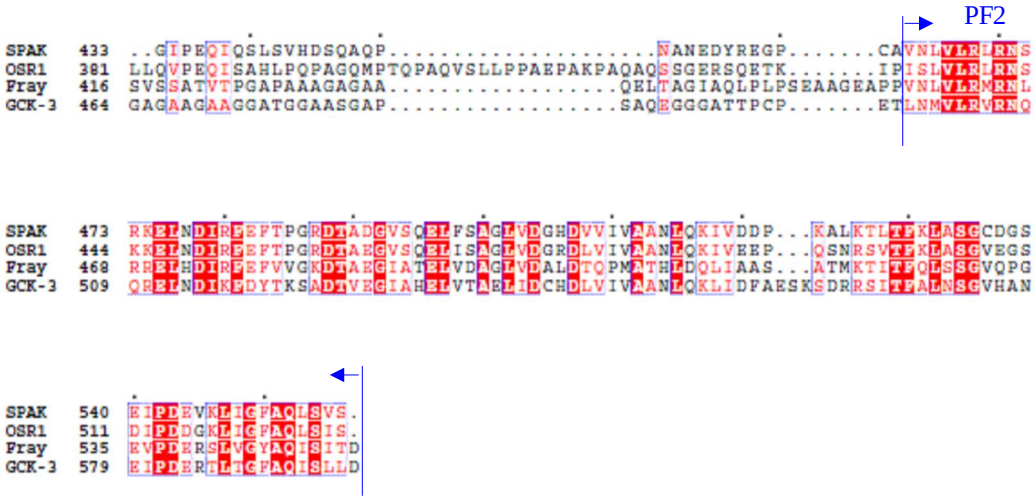


Figure 2-2 Sequence alignment of SPAK and Ste20 GCK-VI kinases

Four Ste20 GCK-VI kinases are aligned. The accession numbers for each protein are: mouse SPAK NP_058562; mouse OSR1 Q6P9R2; *Drosophila* Fray AAO24953; and *C. elegans* GCK-3 NP_507517. Identical residues are in red boxes and conserved residues are in red letters. Three potential phosphorylation sites in the activation loop (Thr231, Thr236, and Thr243 in SPAK) are indicated by green stars.

Materials and Methods

Construction of Expression Plasmid

Full-length and a series of truncated SPAK constructs were PCR amplified from the mouse cDNA clone 6843981 (purchased from ATCC). The resulting PCR products were digested with EcoR1 and Xho1 and cloned into the EcoR1/Xho1 region of bacterial expression vector pHisParallel (Sheffield et al, 1999), containing a T7 promoter, N-terminal His6-tag, and a tobacco etch virus (TEV) protease cleavage site.

Protein Expression and Purification

The pHisParallel-SPAK plasmid was transformed into *E. coli* strain Rosetta (DE3) cells (Novagen) for protein expression. The starter culture was made by inoculating 50 ml liquid Luria-Bertani (LB) medium containing 50 μg/ml ampicillin and 34 μg/ml chloramphenicol with a single colony. The starter culture was first grown at 37°C until OD_{600}=0.6, and then 10 ml of it was used to inoculate 1 l LB medium containing 50 μg/ml ampicillin and 34 μg/ml chloramphenicol. The 1 l culture was grown at 37°C until OD_{600} reached 0.4-0.6 and was cooled to 16°C before inducing with 0.5 mM isopropyl-1-thio-β-D- galactopyranoside (IPTG). The culture was grown at 16°C overnight. The cells were harvested by centrifugation and resuspended in buffer containing 50 mM HEPES, pH 8.0, 300 mM NaCl, 10% glycerol, and supplemented with

protease inhibitors, 1 mM phenylmethanesulphonyl fluoride (PMSF), 0.2 μg/ml leupeptin, 10 mM benzamidine, and 2 μg/ml aprotinin). The lysates were frozen at –80 ºC before purified. To break the cells, *E. coli* cells were frozen and thawed for several times, then incubated with 1 mg/ml lysozyme on ice for 30 minutes, and followed by sonication. Cell debris was removed by centrifugation at 20,000 rpm for 1 hour.

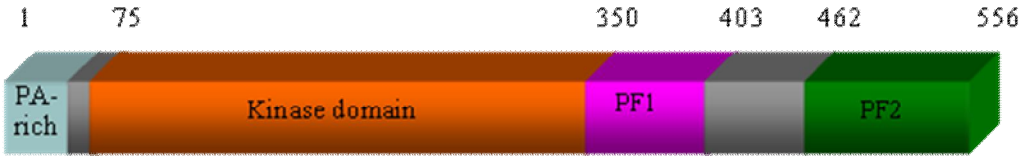
To obtain stable and well-expressed proteins suitable for crystallization, different constructs, expression hosts, and expression conditions were examined. Based on sequence alignment, domain prediction, and secondary structure prediction, a series of constructs consisting of different lengths of SPAK were generated. Three N-terminal starting sites were chosen: 1 (the beginning of protein), 63 (the beginning of region conserved between SPAK and its homologues), and 75 (the beginning of the kinase domain). Five C-terminal ending sites: 350 (the end of kinase domain), 390 (the beginning of a long loop), 403 (the end of the region conserved between SPAK and homologs), 421(the end of the region conserved between SPAK and OSR), and 556 (the end of protein). Generally, constructs consisting of the kinase domain plus either the whole (residues 350-403) or part of (residues 350-390) PF1 were highly stable and well expressed. Full-length SPAK was found to have significant degradation products, and removing the region after the PF1 solved this problem. Limited tryptic protease digestion showed a protein of approximately 40 kDa, which roughly corresponds to the piece consisting of the kinase domain plus the PF1, which was resistant to protease digestion. Unexpectedly, pieces of SPAK with N-terminus starting at the beginning of kinase domain (amino acid residue 75) were insoluble, and inclusion of additional 12 amino

acids in front of the kinase domain stabilized the protein. The outcome of protein expression screening is summarized in Table 2-1.

Initially, *E. coli* strain BL21 (DE3) was used to express SPAK, but the expression level was extremely low. Switching to Rosetta (DE3) gave a significantly higher expression compared to BL21 (DE3). The *E. coli* Rosetta strain includes extra copies of tRNA genes which recognize codons rarely used in *E. coli* but frequently used in higher organisms (Novy et al, 2001). This modification helps to resolve low expression problem due to codon usage bias.

Induction temperature and duration are two other crucial factors for successful protein expression. After examining different induction temperatures, 16°C, 30°C, and 37°C, and different induction durations, 4 hours and overnight, SPAK was found best expressed at 16°C overnight.

Based on protein expression screening, we focused the structural study on SPAK 63-390, one of the most stable and well-expressed constructs that had been tested. Since we were interested in understanding the activation mechanism of SPAK, we also mutated the WNK phosphorylation site Thr243 to aspartate to mimic the partially active form. SPAK 63-390 (T243D) was expressed equally well as wildtype discussed above in *E. coli* Rosetta (DE3).



Construct	Start	End	Host	Expression
Spak-full length	1	556	BL21	No
Spak-full length	1	556	Rosetta	Expressed but not stable
Spak75-350	75	350	Rosetta	Not soluble
Spak75-556	75	556	Rosetta	Not soluble
Spak63-350	63	350	Rosetta	Not soluble
Spak63-421	63	421	Rosetta	Expressed but low
Spak63-421	63	421	Sf9	Expressed
Spak63-390	63	390	Rosetta	Expressed
Spak63-403	63	403	Rosetta	Expressed
Spak63-556	63	556	Rosetta	Expressed but not stable

Table 2-1 Summary of protein expression screening of mouse SPAK

SPAK 63-390 (T243D) has an N-terminal 6xHis tag, followed by a TEV protease cleavage site, and a calculated isoelectric point (pI) of 8.9. Therefore, Ni-NTA agarose (Qiagen), Mono S (Pharmacia), and Superdex75 (Pharmacia) columns were used sequentially to purify protein.

The cell extract after clarification by centrifugation was applied to a nickel-nitrilotriacetic acid-agarose (Ni^{2+}-NTA) column (Qiagen), pre-equilibrated with buffer containing 50 mM HEPES, pH 8.0, 300 mM NaCl, and 10% glycerol. The column was washed with 10 column volumes of the equilibrium buffer and another 10 column volumes of same buffer containing 20 mM imidazole. Protein was eluted in a single step with 50 mM HEPES, pH 8.0, 300 mM NaCl, 10% glycerol, and 250 mM imidazole. The Coomassie blue stained gel of proteins after the nickel column is shown in Figure 2-3. Typically, a 6-liter culture gave a yield of approximately 30 mg of protein at this step. The protein eluted from the nickel column was dialyzed against 50 mM HEPES, pH 7.5, and 50 mM NaCl at 4°C for overnight; at the same time, the N-terminal 6xHis tag was removed by treatment with TEV protease at a ratio of 50:1 (protein: protease). To separate untagged protein from 6xHis-tagged protein, the cleaved protein was passed through a second nickel column. Protein that passed through the column was collected.

Fractions containing SPAK protein were further purified on a Mono S HR5/5 column (Pharmacia) after passed through a 0.22 μm syringe filter. After washing the column with 5 column volumes of Mono S buffer A (50 mM HEPES, pH 7.5, 50 mM NaCl, 1 mM DTT, and 1 mM EDTA), protein was eluted with a gradient of 50 mM to 1 M NaCl in Mono S buffer. Figure 2-4 shows the elution profile of proteins from Mono S.

Usually, two peaks were found in the Mono S chromatogram. The first and larger peak eluted at a salt concentration of approximately 200-250 mM NaCl. The second and smaller peak eluted with a higher salt concentration, 450-700 mM NaCl. Both peaks contained the protein of interest as shown by SDS-PAGE.

Each peak containing SPAK protein from Mono S was further purified by a Superdex75 16/60 gel filtration column (Pharmacia), equilibrated in 50 mM HEPES, pH 7.0, 50 mM NaCl, 1 mM DTT, and 1 mM EDTA. The elution profile and Coomassie blue stained gel of proteins from Superdex75 are shown in Figure 2-5. Proteins from the first Mono S peak were eluted at elution volume of 60-70 ml corresponding to monomers, while proteins from the second Mono S peak were eluted mainly at elution volume of 50-60 ml that matches dimers. All proteins from Superdex75 gave a homogeneous 37 kDa band on SDS-PAGE. Rechromatography of proteins from the monomer fraction gave rise to dimer. Similarly, rechromatography of the dimer fraction gave rise to monomer. Thus, SPAK 63-390 (T243D) proteins are suggested to exist in monomer-dimer equilibrium in solution.

To further verify the homogeneity of the protein, the final purification product was examined by electrospray ionization mass spectrometry (ESI-MS) analysis. ESI-MS provides a more accurate molecular weight measurement. A single peak at molecular weight of 37604 Da was found and it matched the correct molecular weight of the protein.

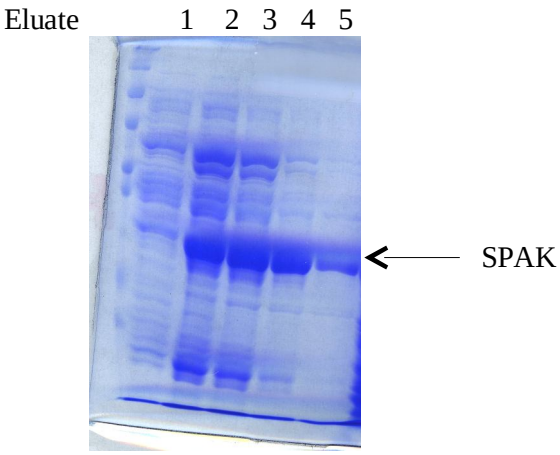


Figure 2-3 Purification of SPAK by nickel affinity column
Commassie blue stained gel of protein from nickel column

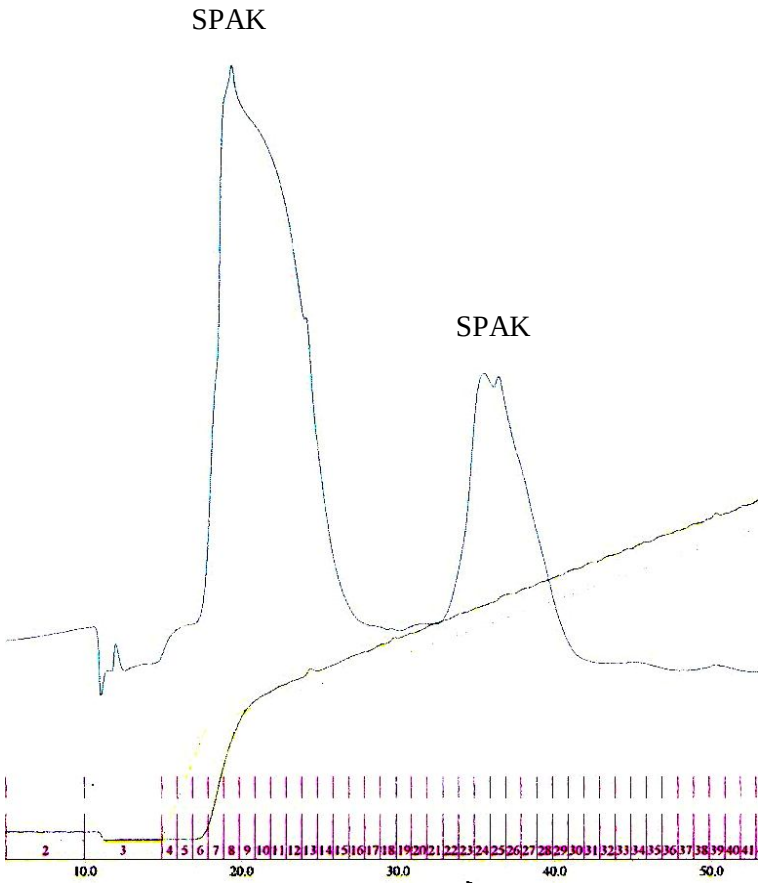


Figure 2-4 Purification of SPAK by Mono S cation exchange column

A chromatogram (OD_{260}) of SPAK on the Mono S column.

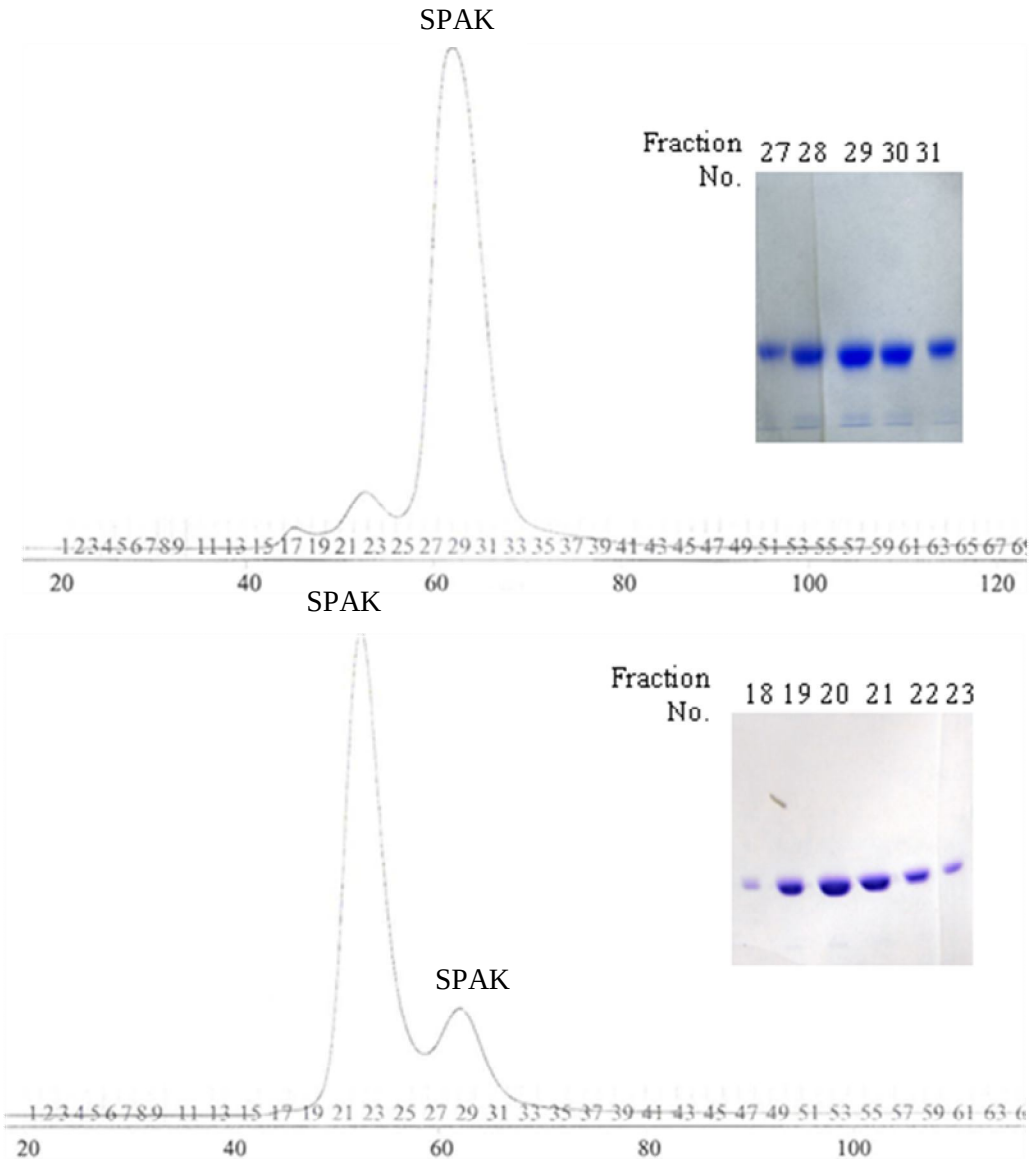


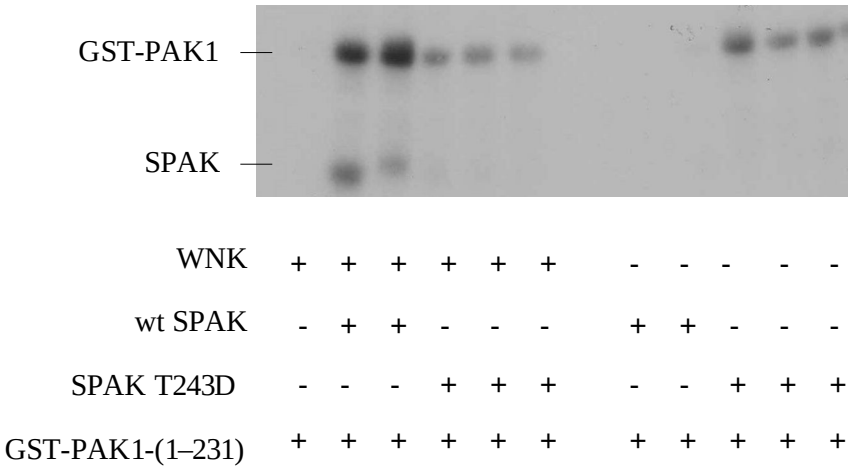
Figure 2-5 Purification of SPAK by Superdex75 gel filtration column

Chromatogram (OD_{260}) and Coomassie blue stained gel of SPAK on the Superdex75 column. Top panel shows SPAK from the first Mono S peak. Bottom panel shows SPAK from the second Mono S peak.

Kinase assay

To determine the kinase activity of SPAK 63-390 (T243D), in vitro kinase assays were performed. 0.5-1 μg of SPAK was added to a 30 μl kinase reaction in 20 mM Hepes, pH 7.6, 10 μM ATP, 10 mM $MgCl_2$, 10 mM β-glycerophosphate, 1 mM DTT, 1 mM benzamidine, and 10 μCi [γ-^{32}P]ATP with 5 μg GST-PAK1 1-230 as substrate. The reaction was carried out at 30°C for 30 min. The results showed that SPAK 63-390 (T243D) has greater (about 5-fold) catalytic activity than that of unphosphorylated wildtype SPAK. Compared to SPAK that was phosphorylated by WNK1, the activity of SPAK 63-390 (T243D) was still low (10-fold). This suggested that mutation of the activation loop residue Thr243 to Asp only causes a partial activation of this protein (Figure 2-6). Additionally, the phosphorylation of SPAK protein by WNK1 was greatly reduced in the T243D mutant and the activity of this mutant kinase was no longer increased by WNK1. This supported the conclusion that WNK1 phosphorylates SPAK on residue Thr243. Notably, autophosphorylation was still observed in the SPAK 63-390 (T243D) protein, indicating that autophosphorylation occurs on other additional sites.

A



B

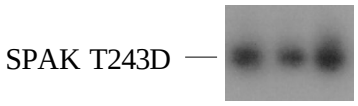


Figure 2-6 Kinase activity of SPAK 63-390 (T243D)

A. Comparison of kinase activity of SPAK 63-390 (T243D) and wild type SPAK 63-390. 0.5 µg of wt SPAK or SPAK T243D were added to a 30 µl kinase reaction with and 5 µg GST-PAK1 1-230 as substrate and with or without 0.2 µg GST-WNK 160-490 as activating kinase. The reaction was carried out at 30°C for 30 min.

B. Autophosphorylation of SPAK T243D. 1 µg of SPAK T243D was added to a 30 µl kinase reaction . The reaction was carried out at 30°C for 30 min.

Protein Crystallization

Extensive crystallization screens were performed on all the stable SPAK proteins that were made, including SPAK 63-390, SPAK 63-403, SPAK 63-421, and SPAK 63-390 (T243D). The results are summarized in Table 2-2. Proteins were concentrated to approximately 10 mg/ml for initial crystallization screening. Screening was carried out at 20°C using the hanging drop vapor diffusion method. Protein was mixed with an equal volume of screening solution. Several commercially available kits were used, including Crystal Screen and Crystal Screen 2 from Hampton Research and Wizard 1 and Wizard 2 from Emerald Biosystems. Crystallization screens on proteins complexed with adenosine derivatives such as ADP, ATP and AMP-PNP were also performed. Protein complexes were made by preincubating the protein with 5 mM of each compound and 5 mM $MgCl_2$ for 30 min on ice before setting up the screens.

Among all the screens, only SPAK 63-390 (T243D) in complex with either ATP or AMP-PNP crystallized and both crystals were found in the same condition in Hampton crystal screen kit 1 which consisted of 0.2 M magnesium acetate, 0.1 M sodium cacodylate pH 6.5, and 20% PEG8000. Although SPAK 63-390 (T243D) can exist as both monomer and dimer in solution, proteins from both states crystallized under the same conditions, and their crystals looked similar. This suggested that in spite of the existence of a monomer-dimer equilibrium in SPAK 63-390 (T243D), the crystallization condition might drive the equilibrium toward one state.

Source		Screenings	Crystal
SPAK 63-390: 13mg/ml	-	Hampton crystal I, II Wizard I, II	No
	AMP-PNP/$MgCl_2$	Hampton crystal I, II Wizard I, II	No
SPAK 63-403: 8mg/ml	-	Hampton crystal I, II Wizard I, II	No
SPAK 63-421 (from Sf9): 8mg/ml	-	Hampton crystal I, II Wizard I, II	No
SPAK 63-390 (T243D): 13mg/ml	-	Hampton crystal I, II Wizard I, II	No
	ADP/$MgCl_2$	Hampton crystal I, II Wizard I, II	No
	ATP/$MgCl_2$	Hampton crystal I, II Wizard I, II	Yes
	AMP-PNP/$MgCl_2$	Hampton crystal I, II Wizard I, II	Yes
	Staurosporine	Hampton crystal I, II Wizard I, II	No

Table 2-2 Summary of crystallization screening of SPAK

Crystals from the initial screening appeared about 2 to 3 hours after plate set up. These crystals were very tiny and needle-like in shape. The longest dimension was only about 30 μm (Figure 2-7). These crystals diffracted poorly to less than 20 Å (Figure 2-8). To improve the quality of the crystals, numerous changes were examined. The crystallization conditions were optimized by changing the protein concentration, crystallization temperature, precipitant species, buffer type and pH of the well solution, and with the inclusion of different additives. Crystals were found to grow better with a higher protein concentration, usually above 20 mg/ml, and also at a higher temperature. (20 °C was usually better than 16 °C or 4 °C.) The well solution composition was extensively modified. Higher quality crystals were finally obtained in a well solution consisting of 0.2 M Mg acetate$_2$, 0.1 M Tris, pH 8.5, 16~18% PEG3350, 0.01 M Na cacodylate, and 0.01 M $CaCl_2$ (Figure 2-7) .

As protein size is a crucial factor that affects crystal diffraction, different lengths of proteins were tried. I both added and removed three to five residues from the C-terminus of SPAK 63-390 (T243D) but did not find any significant improvement was found. I also tested an even longer construct, SPAK 63-403 (T243D), but the added length only lowered the quality of the crystals. Adding a 6xHis-tag or other fusion at the N-terminus is another trick that was suggested, but in this case the tag only made the crystals smaller with greater mosaicity.

All diffraction data were collected under a stream of nitrogen to minimize radiation damage. To prevent ice forming during this process, crystals were soaked in a series of well solutions containing increasing amounts of cryoprotectant before freezing.

Different cryoprotectants were examined, including glycerol, ethylene glycol, polyethylene glycol 200 (PEG-200), polyethylene glycol 400 (PEG-400), and 2-methy-2,4-pentanediol (MPD). Among these, glycerol gave the best result. Other cryoprotectants either cracked or increased the mosaicity of the crystal.

After optimization, crystals grew up to 300-400 μm and typically appeared within three days. Crystals were flash frozen in liquid nitrogen after sequentially soaking in well solution containing additionally 5%, 10%, and 15% glycerol as cryoprotectant. Crystals diffracted to 3 Å using the home X-ray source and further improved to 2.5 Å at the synchrotron beamline (Figure 2-9).

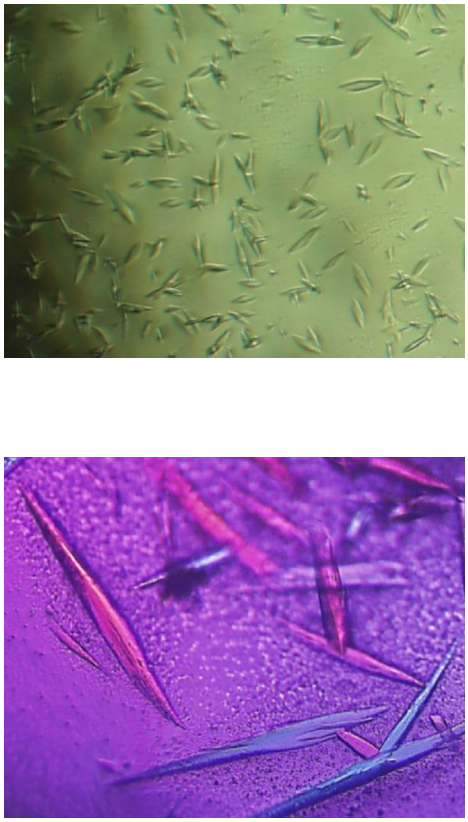


Figure 2-7 Crystals of SPAK 63-390 (T243D)

Crystals from the initial screening were very tiny with the longest dimension about 30 μm(top). Crystals after optimization increased the size to 300-400 μm (bottom).

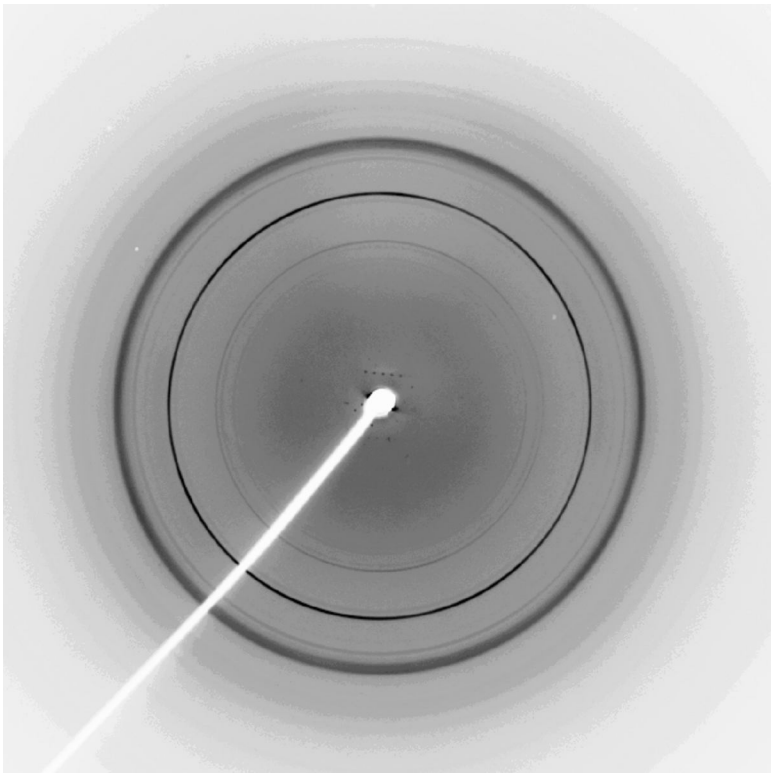


Figure 2-8 X-ray diffraction pattern of a initial SPAK 63-390 (T243D) crystal

Diffraction data was collected in-house on an RAXIS-IV image plate detector equipped with a Rigaku FR-E SuperBright X-ray generator. Resolution is about 20 Å

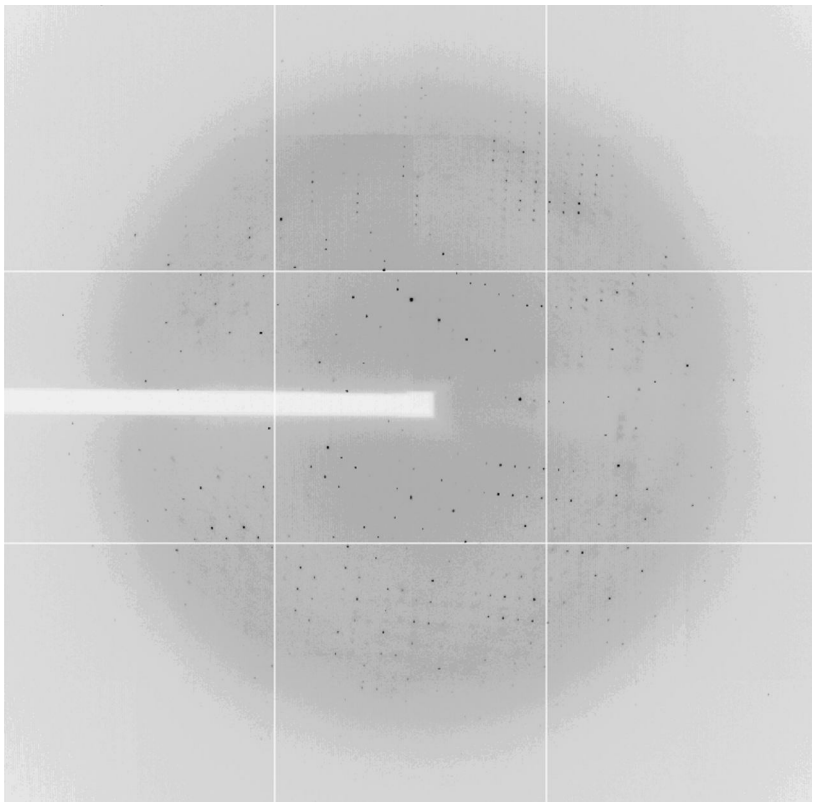


Figure 2-9 X-ray diffraction pattern of an optimized SPAK 63-390 (T243D) crystal

Diffraction data was collected at beamline 19-ID at the Advance Photon Source, Argonne National Laboratory (Argonne, IL).

Resolution is 2.5 Å.

Unit cell is a=66.44 Å, b=101.75 Å, c=104.08 Å, and $\alpha = \beta = \gamma = 90°$.

Data Collection and Processing

The first diffraction dataset for SPAK 63-390 (T243D) was collected at the home source X-ray beam (λ=1.54 Å) and RAXIS-IV image plate detector. About 130 images were collected, with an exposure time of 10 minutes per image and one degree rotation between each image. Unfortunately, the crystal seriously decayed after the long data collection procedure. A higher resolution dataset was collected at beamline 19-ID (λ=0.9794 Å) at the Advanced Photon Source (Argonne National Laboratory, Argonne, IL), and this dataset was used in this study. Diffraction data were from a single crystal collected over 150° with a 0.5° rotation per image recorded by a charge-coupled device (CCD) detector. Data were indexed, integrated, and scaled using HKL2000 program package (Z. Otwinowski and W. Minor, 1997). The crystals of SPAK 63-390 (T243D) are orthorhombic with unit cell parameters of a =66.44 Å, b =101.75 Å, c =104.08 Å, $\alpha=\beta=\gamma$= 90°. Based on the scaling output, odd indexed reflections were systematically absent along all three perpendicular axes indicating the space group to be $P2_12_12_1$.

The number of molecules in the asymmetric unit was estimated by the Matthews coefficient. The Matthews coefficient (V_M) defines the unit cell volume per dalton, which is calculated by the formula: $V_M = V/(nM)$, where V is volume of the unit cell, n is the number monomers in the asymmetric unit, and M is the molecular weight of the protein. For protein crystals, the value of V_M is between 1.77 and 3.54 Å^3 per dalton. Assuming the molecular weight of SPAK 63-390 (T243D) is 35,000 daltons, the calculated Matthews coefficient for two molecules per asymmetric unit will be 2.51, an acceptable value, whereas the Matthews coefficient will be out of the range, 5.03 and

1.68, if each asymmetric unit is occupied by one or three molecules. For the two-molecule asymmetric unit, the corresponding solvent content is about 50%.

Structure Determination

Initial phases were obtained by the molecular replacement (MR) method. Molecular replacement is a technique which relies on the phase information from the known protein models to calculate the phases for a new protein. In the process, the known phasing model is placed in the unit cell of a new protein, and through extensive rotational and translational searches, the most likely phases for the new protein will be estimated. The success of molecular replacement is largely dependent on the phasing models. Usually, if the model is fairly complete and shares at least 40% sequence similarity with the new protein, molecular replacement can be straightforward (Tickle & Driessen, 1996). Less completeness, low homology, and conformational differences can increase the difficulty of molecular replacement.

To identify suitable models for molecular replacement, I used PSI-BLAST to search for structures that share high sequence similarity with SPAK. Ste20 protein kinases, PAK1 (pdb code 1f3m), PAK4 (pdb code 2bva), and PAK6 (pdb code 2c30) were the available structures with sequences closest to SPAK at that time. Initially, these models were used directly for phase searching by the program Phaser in the CCP4 package (McCoy et al, 2005), and using only data between resolutions 4 and 20 Å in the calculation. Even higher resolution data were not included as they are usually very

distinct from the search models. Unfortunately, none of these three models I used generated a good solution. I further tested many other protein kinases including Tao2, MEK, and PKA. Still no solution was found. One possibility for our failure was that the sequence identity between SPAK and each model was too low; all were less than 35%. To increase the similarity, all unconserved side chains in the models were modified to alanine using the program Chainsaw (Schwarzenbacher et al, 2004). Moreover, based on an understanding of kinase structures, the activation loop in the kinase domain usually adopts a different conformation in each kinase. Therefore, I removed the activation loop from all search models. These changes allowed me to obtain phases from PAK1, PAK4, and PAK6, all of which gave the same solution. PAK6 showed the best score. The rotation function Z-score (RFZ) was 4.0, the translation function Z-score (TFZ) was 14.1, and the log-likelihood gain (LLG) was 405. For Phaser, TFZ and LLG are two major indicators for the correct solutions. A correct solution normally has a TFZ over 5 and is clearly separated from other solutions. As for LLG, the value should be increased after each run. Rigid body and restrained refinements were subsequently carried out by Refmac5 (Murshudov et al, 1997) in the CCP4 package. The R_{work} and R_{free} were 40 % and 50 %, respectively, which are unacceptably high. The resulting electron density map was marginally fit to the structure and obviously needed a lot of improvement.

Several approaches were employed in order to improve the phases of the structure factors further, but there was no significant improvement that resulted. First, numerous available structures were used simultaneously as phasing models in a single search. Given that different models can be better matched to different areas, a higher

accuracy of prediction may be achieved. However, in our case, the quality of the electron density map remained the same. Next, based on other studies of kinase structures, the relative orientations of the N- and C-terminal domains of each kinase may differ. They adopt open or closed conformations and also have different rotations relative to each other. Therefore, I decided to separate the N- and C-terminal domains of each kinase model, and use them independently to search the phase. Different combinations of the N- and C- terminal domains from various models were used in searches, but still there was no improvement. Lastly, experimental phasing was also tried. Phase information from experiments may help to solve the molecular replacement problem. In fact, several studies suggested that the combination of these two techniques is particularly useful when the molecular replacement model is too poor to find a solution and the experimental phases are too ambiguous to provide an interpretable map (Mccoy et al, 2007). Two methods were used here. First, I generated the selenomethionine-substituted SPAK 63-390 (T243D) protein. The selenomethionyl variant was crystallized in a condition similar to its native counterpart; however, the crystals were much smaller and diffracted poorly. At best, these crystals diffracted to 3.5 Å under the synchrotron X-ray beamline, but decayed before I could collect a full dataset. Second, I made heavy-atom incorporated crystals by soaking native crystals with heavy-atom derivatives. A large number of heavy-atom compounds were tested, including these mercury derivatives ethyl mercuric phosphate and 2-chloromercuri-4-nutrophenol; several platinum derivatives, including potassium hexabromoplatinate, potassium tetranitroplatinate, and diamminedichloroplatinum; the gold derivative, gold cyanide; and others samarium acetate, uranyl acetate, and potassium iodide. Most of the heavy-atom compounds

caused severe damage to the crystals even after I modified the soaking concentration and duration. For those crystals that survived after soaking, I collected diffraction data. I concluded that heavy atoms were not incorporated into any of these because the diffraction patterns were the same as that of the original crystals.

Model Building and Refinement

As I was unable to improve the phases further, I began the model building process. The initial electron density map was calculated from the molecular replacement with PAK6 and its subsequent refinement procedures. Model building was carried out manually using the program Coot (Emsley & Cowtan, 2004). Automatic building using software such as ARP/wARP did not work well in this case, likely due to the low resolution and poor phases. I started the rebuilding from areas that were better matched to the density map, such as α-helices E and F in the C-terminal domain. Several criteria were used during the modeling, including the bond lengths and bond angles. For example, the distance between two Cα atoms in consecutive residues should be about 3.8Å, and the length for a hydrogen bond should be about 2.7 Å.

After numerous cycles of rebuilding and refinement, the electron density map was significantly improved. The R_{work} and R_{free} dropped to 25 % and 31 %, respectively; however, these were not sufficient to guarantee a quality model. At this point, the density for AMP-PNP was revealed in the interface between the N- and C-terminal domains, which is the expected ATP-binding location in protein kinases. AMP-PNP was placed in

each protein molecule. I further improved the R factors by the TLS (translation, libration and screw-rotation) refinement and anisotropic scaling. This improvement resulted in a clearer electron density map and allowed me to build the model more properly.

The TLS refinement was carried out in Refmac5. This refinement method is a type of B-factor (temperature factors) refinement which is particularly useful if the structure has a high B-factor. The refinement uses a set of parameters that divide the protein model into several segments and each segment is viewed as an individual rigid body undergoing translation, libration and screw-rotation displacements (Winn et al, 2001).

Anisotropic scaling was performed by the Diffraction Anisotropic Server at UCLA (Strong et al, 2006). This helped to solve the mild diffraction anisotropy problem found in my dataset. Diffraction anisotropy is commonly seen in protein crystals which is attributed to uneven crystal packing. In other words, crystal packing interactions are more uniform in one direction than another and thus a higher resolution is observed in one direction than the other two. The problem for diffraction anisotropy is that the regular refinement programs are unable to choose a suitable resolution boundary. If the resolution cutoff is set low based on the weakly diffracting direction, high resolution data in the better diffracting directions will be lost. On the other hand, if the cutoff is set too high, all the bad reflections in the weakly diffracting direction will be forced to be included. Through anisotropic scaling, the resolution limits along the three cell directions are determined and the data are truncated by ellipsoidal limits.

After a series of model adjustments and refinements, and the addition of water molecules into the model, the current model was refined to a resolution of 2.43 Å with R_{work} = 21 % and R_{free} = 26%. Two molecules were found in the crystallographic asymmetric unit. The model comprises residues 64-230 and 244-365 of the residues 63-390 of molecule A, and residues 65-233 and 244-365 of molecule B. There are a total of 580 protein residues, 2 AMP-PNP molecules and 104 water molecules in the model. Finally, the geometry of the model was validated by the MolProbity web server to make sure all the bond angles and bond lengths are within acceptable limits. The crystal data and refinement statistics are summarized in Table 2-3

Table 2-3 Summary of crystal data and refinement statistics

Data	SPAK 63-390 (T243D)
Spacegroup	$P\,2_1\,2_1\,2_1$
Cell dimensions (Å)	a= 66.44, b= 101.74, c= 104.08
α, β, γ	90 °, 90 °, 90 °
Resolution (Å)	50-2.5
Observed reflection	308327
Unique reflection	27322
Completeness (outer shell)	93 (80.7)
R_{merge} (outer shell)[1]	0.043 (0.41)
I/> (outer shell)	31.4 (1.98)
Refinement	
R_{work}[2]	0.21
R_{free}[3]	0.26
R.m.s.d. bond length	0.009 Å
R.m.s.d. bond angle	1.34 °
Average B factor (Å^2)	46.2

[1] $R_{merge} = \Sigma|(I_{hkl}) - \langle I \rangle|/\Sigma(I_{hkl})$, where I_{hkl} is the integrated intensity of a given reflection

[2] $R_{work} = \Sigma_{hkl}|Fo - Fc|/\Sigma_{hkl}|Fo|$, where Fo and Fc are the observed and calculated structure factors, respectively.

[3] Five percent of the reflections were used in the calculation of R_{free}.

Results and Discussion

The overall structure of SPAK 63-390 (T243D)

SPAK 63-390 (T243D) is the protein we used for crystallization. It comprises the N-terminal extension (19 residues upstream of the kinase domain, 12 of the residues are from SPAK, residues 63-74, and 7 additional residues are from the vector), the kinase domain (residues 75-350), part of the PF1 domain (residues 351-390), and a substitution of Thr243 to Asp in the activation loop to mimic the phosphorylated state. The crystal structure of SPAK 63-390 (T243D) adopts a canonical two-domain kinase fold consisting of a small N-terminal domain (residues 63-155) connected to the larger C-terminal domain (residues 156-365) through a hinge region. The hinge is located between the two domains which forms part of the active site where AMP-PNP binds. The N-terminal domain contains five anti-parallel β strands and a single α helix (helix αC) that connects β strands 3 and 4. The C-terminal domain consists of mostly α helices and two β strands 7 and 8 (Figure 2-10).

The overall topology of the SPAK kinase domain is most similar to other Ste20 family kinases. A search for similar structural folds by the program DALI (Holm & Sander, 1995) reveals that OSR1 (pdb code 2vwi), MST3 (pdb code 3ckw), STK-10 (pdb code 2j7t) and PAK 7 (pdb code 2f57) are among the top results, with an overall root mean square (r.m.s.) deviation for all main chain atoms ranging from 2.2 to 2.5 Å.

The N-terminal domain of the structure is mostly well-ordered except for the most N-terminal end (residues -7 to -1 from the vector and residues 63-64 from SPAK)

and the loop segment connecting β strand 3 and α helix C (residues 112-114). This loop segment has been reported to be flexible in several kinase structures. It may function as an adaptor enabling the repositioning of α helix C during kinase activation. Following the loop is α helix C (residues 115-121) which is also relatively flexible in the crystal structure. Although its main-chain can be clearly identified, the electron density for the side chains of α helix C is diffuse. Inspecting the residue information reveals that all residues of α helix C contain high temperature factors (the average B-factor here is 60 $Å^2$ versus 46 $Å^2$ for the protein). In protein crystallography, high temperature factors are primarily caused by multiple local conformations, which are usually associated with structural disorder. This suggests that α-helix C might undergo some conformational changes. As α-helix C plays an essential role in kinase activation, by providing a salt bridge interaction between highly conserved lysine and glutamate, the conformational changes found in α-helix C might allow the kinase to switch from the inactive to active conformation.

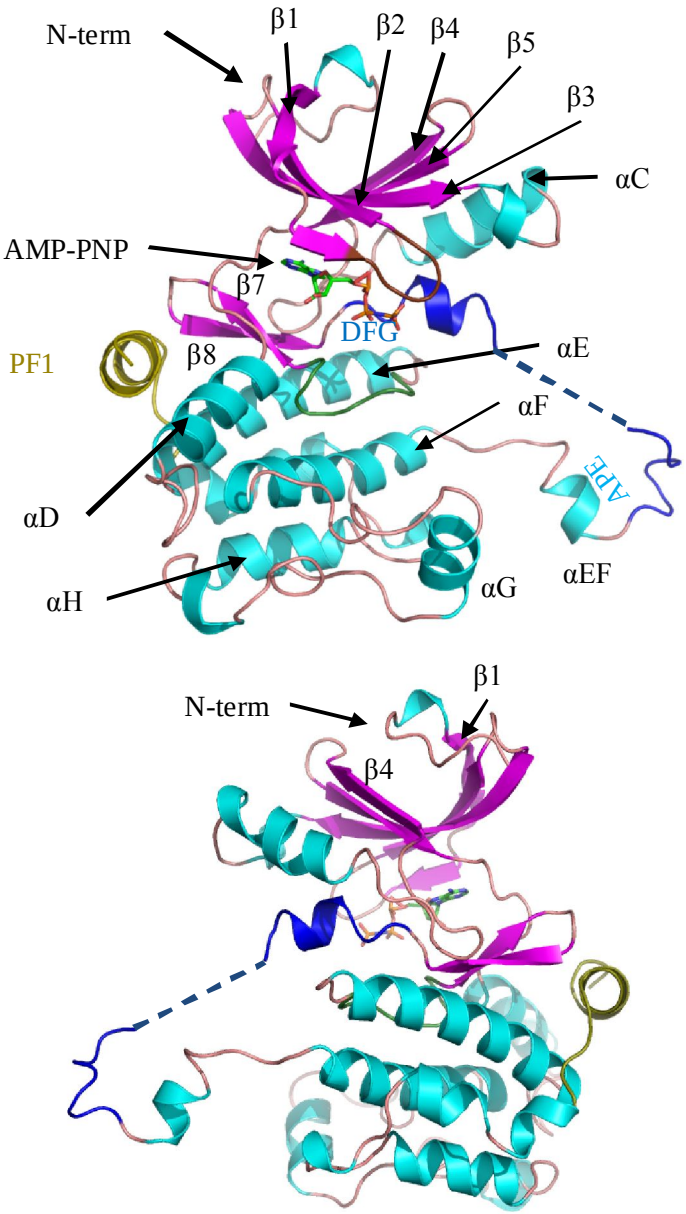


Figure 2-10 General topology of SPAK 63-390 (T243D)

Ribbon diagram of a monomeric SPAK structure. Top panel shows the front view, and bottom panel shows the back view. α helices are shown in cyan and β strands are shown in magenta. The glycine-rich loop is colored in brown, catalytic loop is in green, activation loop is in blue, and PF1 domain is in yellow. AMP-PNP is shown in a stick representation.

The N-terminal extension to the kinase core was believed to be required for proper folding because SPAK protein became insoluble when this segment was removed. In the crystal structure, this N-terminal extended segment (residues 65-74) forms a loop sitting nearly parallel to β strand 4, and a 3_{10} helix close to β strand 1. The contact between the loop and β stand 4 is mainly through hydrophobic interactions. This extended region also interacts with the C-terminal domain of another molecule in the asymmetric unit. All these interactions may contribute to protein stabilization.

The C-terminal domain of the kinase comprises the activation loop and the substrate-binding site. The crystal structure of SPAK 63-390 (T243D) contains several disordered regions in the C-terminal domain, including a portion of the activation loop (residues 231-243 in molecule A; and residues 234-243 in molecule B) and the C-terminus of the PF1 domain (residues 366-390). The WNK phosphorylation sites Thr243 and Ser383 are both missing in the structure. Most of the structures in the C-terminal domain are located at their typical positions. However, one unique feature is the protrusion of the activation loop. This extended activation loop interacts with a symmetry-related molecule in the neighboring unit, and is placed in the position where the activation loop from the interacting molecule should be. Similarly, the activation loop from the second molecule is located in the place where the activation loop of the first molecule is supposed to be (Figure 2-11A). Such exchange of identical structural elements between two molecules is called domain swapping (Bennett et al, 1994).

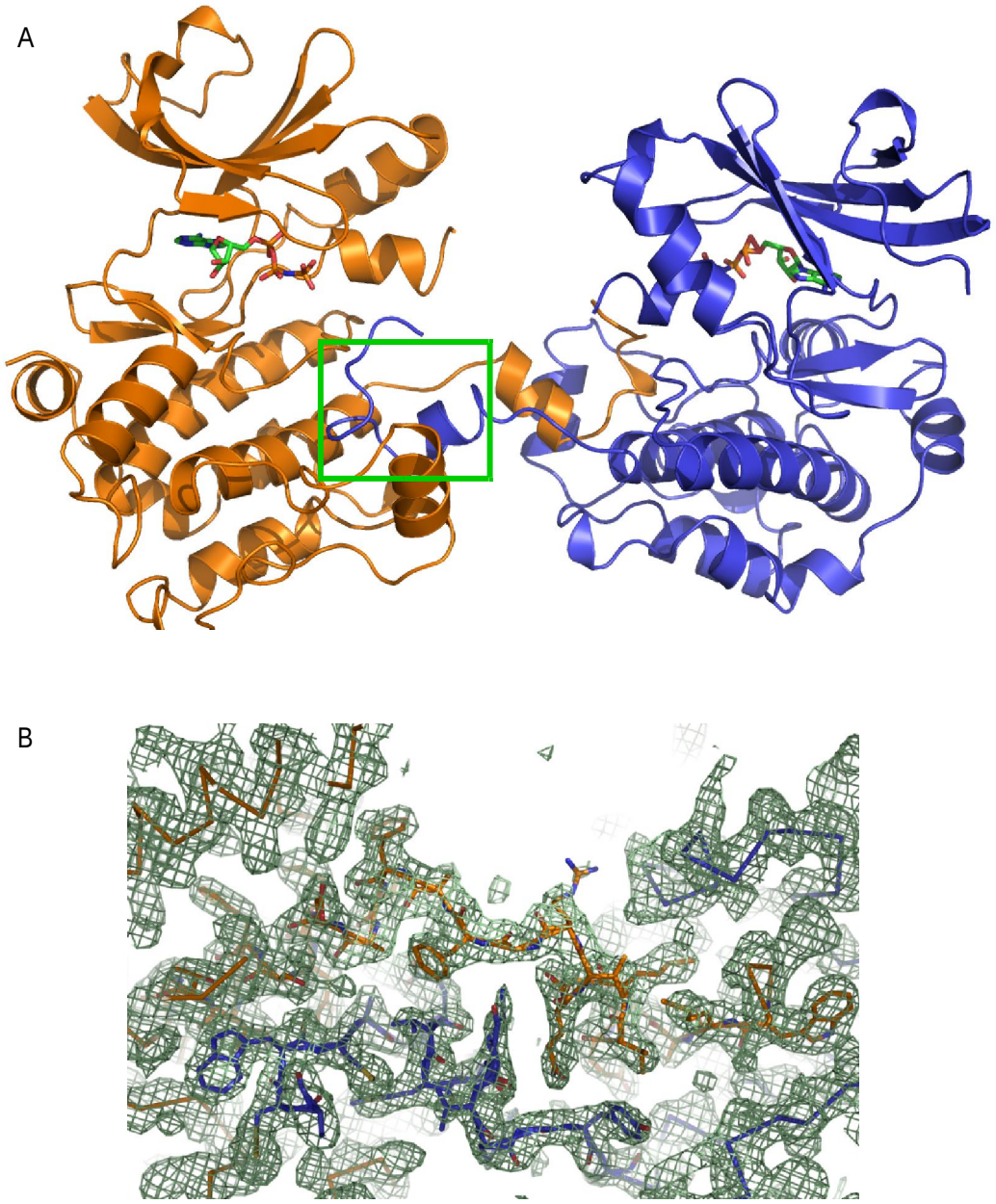


Figure 2-11 Domain-swapped dimer of SPAK 63-390 (T243D)

A. The two monomers are colored in orange and blue, respectively. AMP-PNP in each molecule is shown in a stick representation.

B. A section of 2Fo – Fc electron density map of the domain-swapped region (green box), contoured at 1σ.

Structure of the SPAK 63-390 (T243D) domain swapped dimer

The swapped domain in the structure of SPAK 63-390 (T243D) includes the region containing almost the whole activation loop and the segment following the activation loop, which comprises a total of 30 residues. The N-terminus of the exchanged region starts from several residues after the DFG motif. The DFG motif is not switched in the structure. The exact N-terminal switch point for the swapping is not known, as part of activation loop is absent in the structure. Residues within the missing region (residues 234-243) may perform this switch function. The C-terminus of the swapped region ends in the loop that connects helices αEF (a short helix following the activation loop) and αF. The domain exchange is almost certain because the electron density for the swapped activation loop is clear, especially around the C-terminal hinge region (Figure 2-11B).

The swapping of the activation loop creates an extensive dimer interface between the two monomers. A surface representation of a domain swapped dimer is shown in Figure 2-12. This dimer interface buries about an 1880 $Å^2$ surface area accounting for 11.8 % of total surface area (calculated by the Protein Interfaces, Surfaces and Assemblies, PISA, server. http://www.ebi.ac.uk/msd-srv/prot_int/pistart.html (Krissinel & Henrick, 2007)) and is almost entirely hydrophobic in nature. Residues involved in this contact includes Trp269 and Ile273 in α helix F, and Val293, Leu294, Leu298, and Arg337 in α helix G of one molecule, and Trp250, Pro253, Glu254, Val255, and Met256 in α helix EF of the other molecule (Figure 2-13).

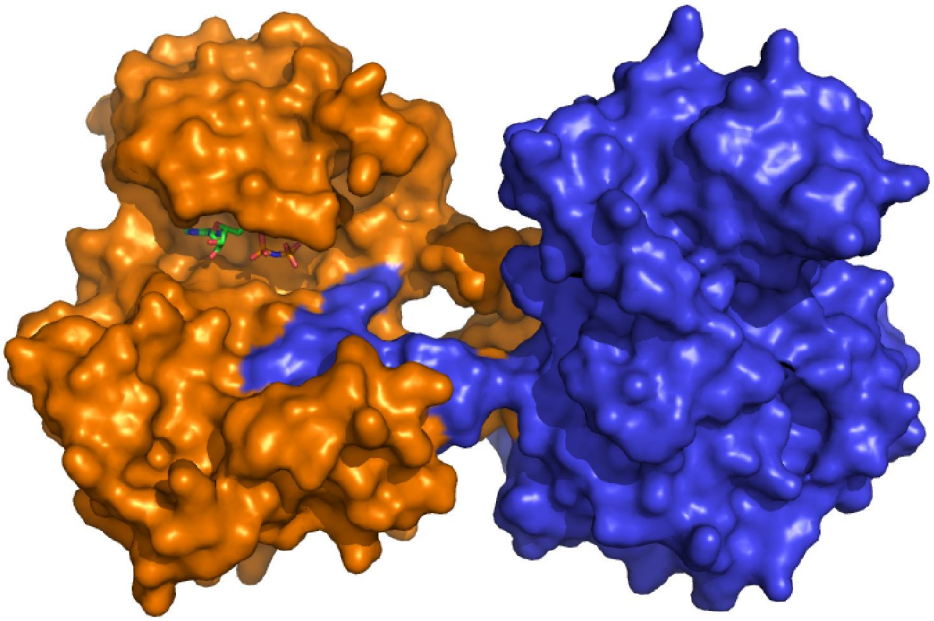


Figure 2-12 Surface representation of a domain-swapped dimer

The two monomers are colored in orange and blue, respectively. AMP-PNP is marked in a stick representation.

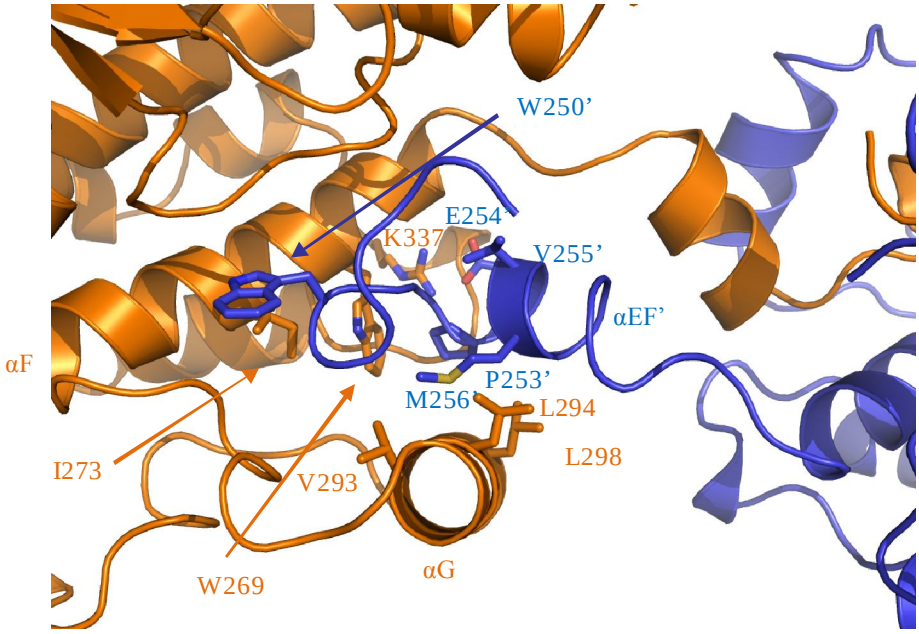


Figure 2-13 Interface of the domain-swapped dimer

The two SPAK molecules are colored in orange and blue, respectively. Residues involved in the contact are shown as sticks

Several other members of the protein kinase family also form dimers through activation loop swapping. These include Ste20-like kinase (SLK), lymophocyte-originated kinase (LOK), death-associated protein kinase 3 (DAPK3), checkpoint kinase 2 (CHK2) serine threonine kinase-10 (STK10), and oxidative stress-responsive kinase-1 (OSR1) (Oliver et al, 2006; Pike et al, 2008). Most of them are Ste20 kinases. The exchanged activation loop has been suggested to allow trans-autophosphorylation to occur, as the activation loop from one kinase docks into the catalytic cleft formed by the other kinase (Pike et al, 2008). Superimposition of SPAK with other domain swapped kinases shows that the domain swapped region adopts various orientations (Figure 2-14). However, a similar feature was found in these activation loop-swapped structures. A comparison of domain swapped kinases and non domain swapped kinases, such as TAO2 (pdb code 1u5r) and PAK6 (pdb code 2c30), reveals a major conformational difference exists in the αEF/αF loop (Figure 2-15). In the non swapped structures, the αEF/αF loop folds back towards the kinase core. As a consequence, the helix αEF docks between α helices F and G, and the P+1 loop interacts with its own kinase domain and forms the substrate binding pocket. Whereas in a domain swapping situation, the αEF/αF loop moves towards the other molecule and thus helix α EF and the P+1 loop are buried inside in the opposite kinase.

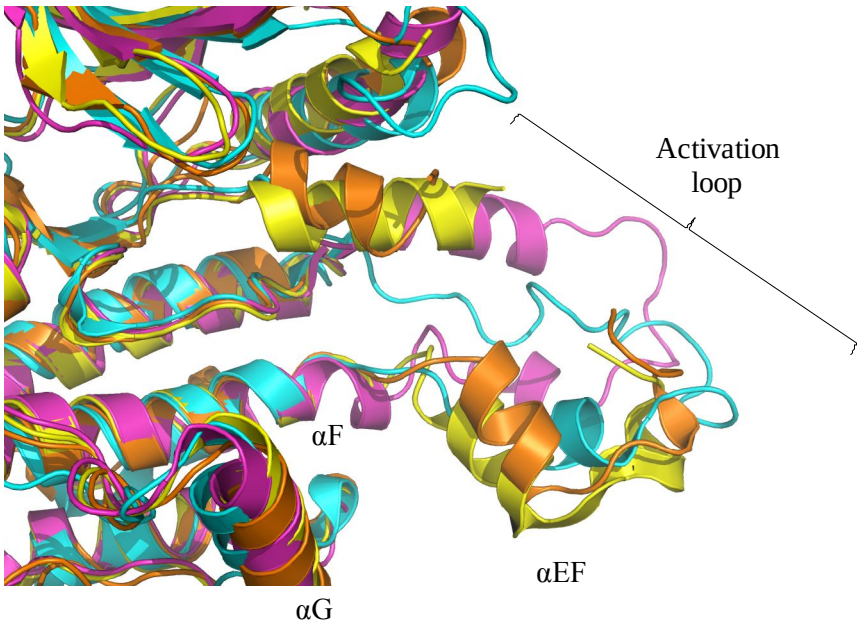


Figure 2-14 Superimposition of SPAK and other domain-swapped kinases in the activation loop

SPAK is colored in orange, SLK (pdb code 2jfm) is in magenta, STK10 (pdb code 2j7t) is in yellow, and DAPK3 (pdb code 2j90) is in cyan.

Inspection of the αEF/αF loop conformations of SPAK and PAK6 shows a significant difference in the backbone conformation between residues Arg260/Gly261 in SPAK and Ser576/Leu577 in PAK6, suggesting that this might be the switch point for domain swapping (Figure 2-15). A structure-based multiple sequence alignment of SPAK and relevant kinases in the domain swapped segment shows that all the domain-swapped kinases contain either a glycine or a proline at the position corresponding to Gly261 in SPAK, while the non-swapped kinases PAK6 and TAO2 have leucine and glutamine instead (Figure2-16). Glycine in this region might increase the flexibility of the backbone, and thus facilitate the domain swapping. As for proline, it is frequently found in the hinge loops of domain swapped proteins. Although proline rigidifies the hinge region, as it prefers an extended conformation, it might favor oligomerization and prevent folding of the hinge loop to the core structure (Rousseau et al, 2003).

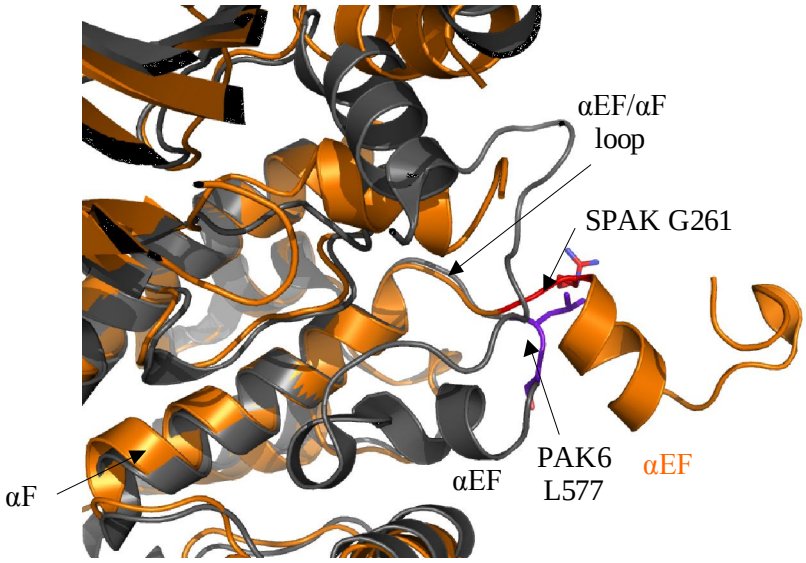


Figure 2-15 Superimposition of SPAK and PAK6 (pdb code 2c30) reveals the difference in the αEF/αF loop

SPAK is colored in orange and PAK6 is in gray. The switch point R260/G261 in SPAK and S576/L577 in PAK6 are represented in sticks and colored in red and purple, respectively.

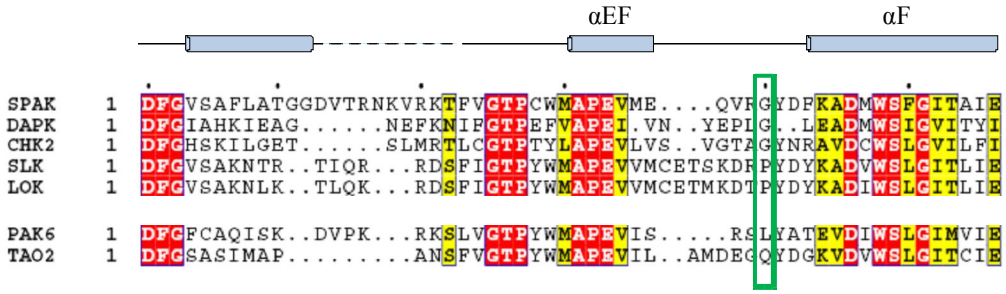


Figure 2-16 Multiple sequence alignment of representative sequences of domain swapped kinases SPAK, DAPK3, CHK2, SLK, LOK, and non domain swapped kinases PAK6 and TAO2. The position corresponding to Gly261 in SPAK is in green box.

The dimeric SPAK 63-390 (T243D) protein is also present in solution. The gel filtration profile of SPAK 63-390 (T243D) indicates that the protein has a monomer-dimer equilibrium (Figure 2-5). However, there are two dimer interfaces found in the structure. One is through the activation loop swapped region which forms interaction with a symmetry-related molecule, as discussed above. The other is found in the asymmetric unit in which the N-terminal domain of one molecule contacts the C-terminal domain of the other molecule, as will be discussed in the next section. Because higher order oligomers were not observed by size exclusion chromatography, it is not clear which kind of dimer is present in solution. Although the interface created by domain swapping is considerably larger than the other interface, 1880 $Å^2$ versus 650 $Å^2$ buried surface area, I still cannot rule out the possibility that this domain-swapped dimer interaction is caused by crystal packing or is a result of protein truncation. Nonetheless, as the number of structures showing domain-swapped kinases continues to increase, it becomes less likely that domain swapping is simply an artifact. In addition, in cells, local protein concentrations could under some circumstances be higher than in solution. This potentially high concentration along with the binding of accessory proteins may all facilitate domain swapping. Moreover, as will be discussed later, the conserved P+1 substrate binding pocket is formed by residues from both monomers of the domain-swapped dimer, which further supports the idea that the domain-swapped dimer is a functional unit for SPAK and probably for other domain-swapped kinases as well.

Domain swapping is found in many other proteins, such as prions (Knaus et al, 2001), cystatin C (Janowski et al, 2001), and transthyretin (Serag et al, 2002) which are

known to form amyloid fibers in vivo, also RNase A (Liu et al, 1998) and yeast cyclin-dependent kinase subunit ($p13^{suc1}$CKS) (Rousseau et al, 2001). Physiological roles are associated with the domain swapping of all of the protein above. The swapped region can be anywhere in the protein, the N-terminus, C-terminus, or in the middle. Based on the available structural evidence, no specific sequence motif or secondary structure is found among these swapped regions. In addition, interactions within the interface may vary. Several studies suggested that domain swapping in proteins can provide an alternative control of the enzyme activity, produce a new substrate binding interface, and create larger protein interaction networks (Liu & Eisenberg, 2002).

At present, it is hard to speculate on how domain swapping affects the function of SPAK. There are all kinds of possibilities such as affecting the kinase catalytic activity, substrate recognition, protein-protein interaction, or subcellular localization. In other domain-swapped kinases, it has been proposed that dimerization is necessary for activation loop autophosphorylation (Pike et al, 2008). However, in SPAK, the role of autophosphorylation is not clear. The activity of SPAK is mainly regulated by the phosphorylation of its activation loop by the WNK kinases (Anselmo et al, 2006). Although autophosphorylation does occur, such phosphorylation is not sufficient for SPAK activation. Clearly, further functional studies are required to understand the role of domain swapping in SPAK fully. Investigation of the underlying mechanisms of domain swapping in SPAK may assist the functional characterization. Mutagenesis of the candidate switch point residues may be evaluated to determine if these affect dimerization, protein kinase activity or substrate preference.

The dimer interface in the asymmetric unit

As mentioned previously, there are two SPAK 63-390 (T243D) molecules in the asymmetric unit. These two molecules form a dimer through the N-terminal domain of one molecule and the C-terminal domain of the second molecule. Two regions in each molecule are involved in this interaction. One is formed between the N-terminal extended loop and β2/β3 loop of the first molecule and the αF/αG loop of the second molecule. The other is from β strand 4 and α-helix C of the first molecule and α-helix D of the second molecule (Figure 2-17). Residues within the interface include Val66, Ile70, Asp73, Pro96, Glu117, Leu119, Ser139, Phe140, and Val141, of the first molecule and Ile162, Val163, Asn167, Arg168, His286, Lys287, Pro289, Pro290, and Glu310 of the second molecule. Both hydrophobic and ionic interactions are involved in this contact. The conformations of each monomer in the dimer are very similar, with an average r.m.s. deviation in Cα positions of 0.45 Å.

Asymmetric dimerization is also found in many other protein kinases. In most cases, this was suggested to occur as a result of crystal packing. However, in the crystal structure of the epidermal growth factor receptor (EGFR) kinase domain (pdb code 2gs6), the interaction between the N-terminal domain of one molecule and the C-terminal domain of the second molecule was demonstrated as a mechanism for kinase activation (Zhang et al, 2006). The interaction ensures a correctly positioned α-helix C which allows the formation of a

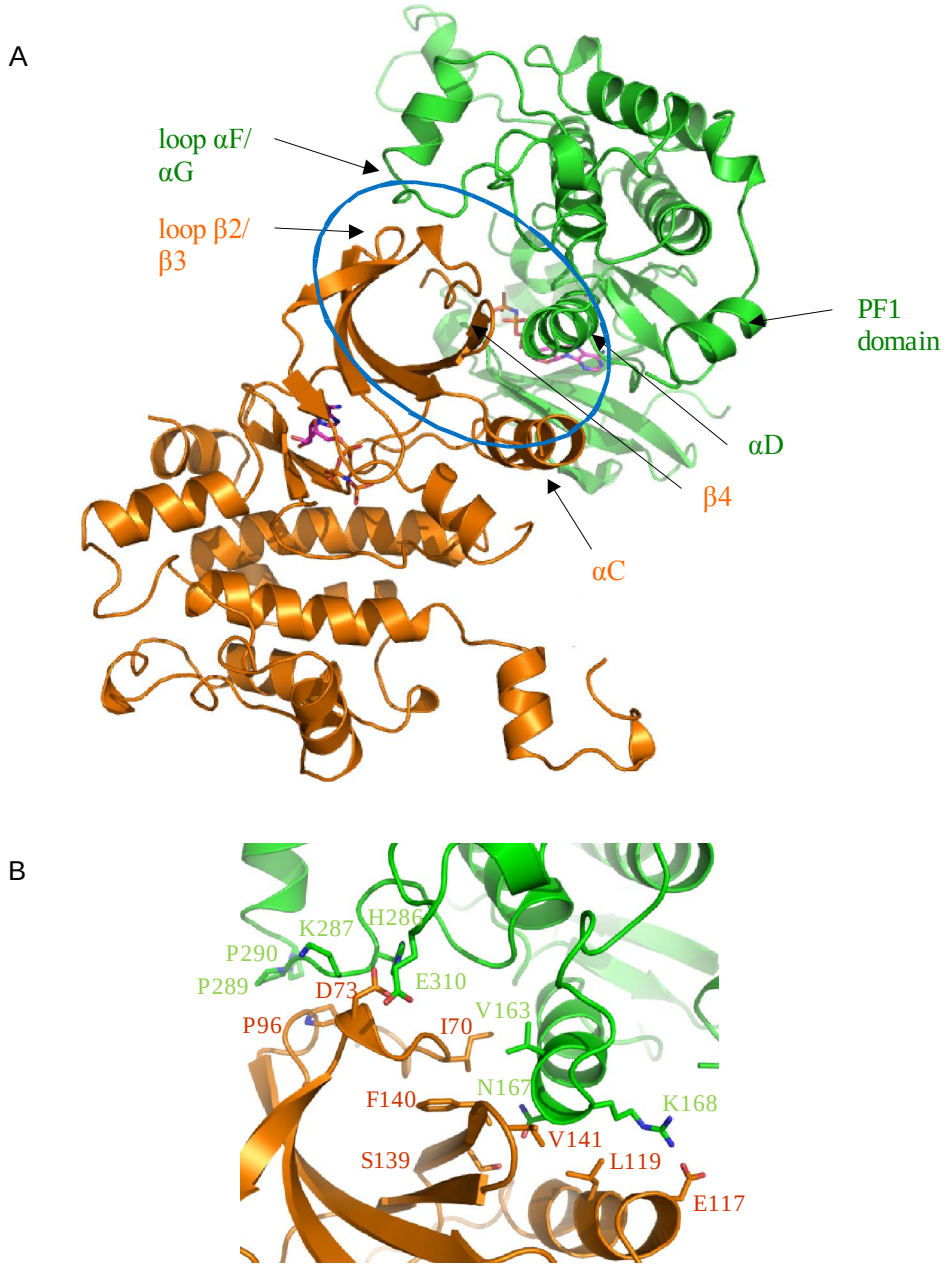


Figure 2-17 Dimer interface between the two asymmetric molecules

A. The two monomers are colored in orange and green, respectively. AMP-PNP in each molecule is shown in a stick representation.

B. The interaction interface is illustrated in close up. Residues involved in the contact are shown as sticks.

salt bridge between the conserved α-helix C glutamate and the ATP-binding lysine. Nonetheless, the details of this interface differ, and the relative orientations of the two monomers are also different between SPAK and the EGFR.

Currently, it is not clear whether this contact found in the asymmetric unit is associated with the regulation of SPAK kinase activity or any other functions. The interaction appears to cause a displacement of α-helix C away from its kinase core when compared with PAK6, and this binding also partially blocks the substrate docking groove. Thus, perhaps the asymmetric dimerization found in SPAK may serve some kind of negative regulatory function. To investigate the precise roles of asymmetric dimerization in SPAK, structure-based mutagenesis of the dimer interface residues will be carried out and followed by characterization of dimer formation and kinase activity.

The active site of SPAK 63-390(T243D)

The kinase domain active site contains the ATP and substrate binding sites and is surrounded by highly conserved residues involved in catalysis. These conserved residues include glycines 82, 84 and 87 in the glycine-rich loop (residues 82-87), Lys104 in β strand 3, Glu121 in α-helix C, the consensus sequence HRDLKxxN (residues 202-209) in the catalytic loop, as well as the DFG motif (residues 222-224) and the P+1 pocket (residues 244-251) in the activation loop. Inspection of each of the above structures will help us to gain a comprehensive understanding of the structure of SPAK.

The ATP binding site

The structure of SPAK 63-390 (T243D) was determined in complex with AMP-PNP, an ATP analog. AMP-PNP binds in the typical position in the interface of the two domains. Figure 2-18 shows the ATP binding pocket. The interactions between AMP-PNP and SPAK are very similar to those observed in other protein kinases. Three hydrogen bonds are involved in the binding. One of them is a direct bonding between the 6-amino and 1-imido group of AMP-PNP with the main-chain carbonyl group of hinge residue Lys151 in SPAK which is a highly conserved interaction in kinases. The adenosine ring of AMP-PNP is buried in a hydrophobic pocket formed by residues Ile81, Val89, Ala102, Val134, Met150, Leu153, and Leu211 from five N-terminal β strands, the hinge region, and β strand 7 in the C-terminal domain. The charged phosphate groups form another two hydrogen bonds with the main-chain amide group of residue Ala85 and the side chain of residue Lys104. Ala85 sits within the glycine-rich loop which adopts a closed conformation and interacts with AMP–PNP. Lys104 located in β strand 3 is an essential residue in protein kinases. Although Lys104 still interacts with AMP-PNP, its side chain is a little shifted toward the α-phosphate of AMP-PNP rather than bonding with the β-phosphate. In structures of active kinases, this ATP-binding lysine forms an ionic interaction with a glutamate in helix αC. The corresponding Glu121 in SPAK, however, is displaced away from this preferred position (Figure 2-19). Although the side chain of Glu121 is not visible in the electron density, the main chain is about 6 Å away from the expected position. Therefore, the salt bridge between the catalytically important lysine and glutamate is not formed in this SPAK structure.

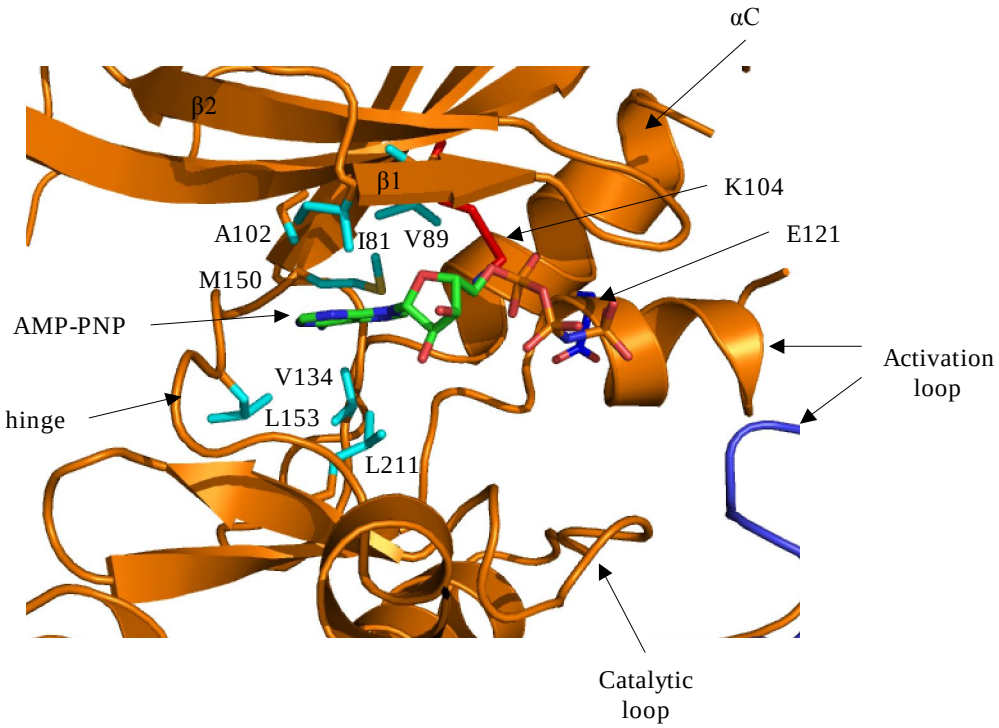


Figure 2-18 The ATP binding pocket in SPAK

Hydrophobic residues surrounding the ATP pocket are shown as cyan sticks. The catalytic Lys104 and Glu121 are represented in sticks and colored in red and blue, respectively. AMP-PNP is shown in a stick representation.

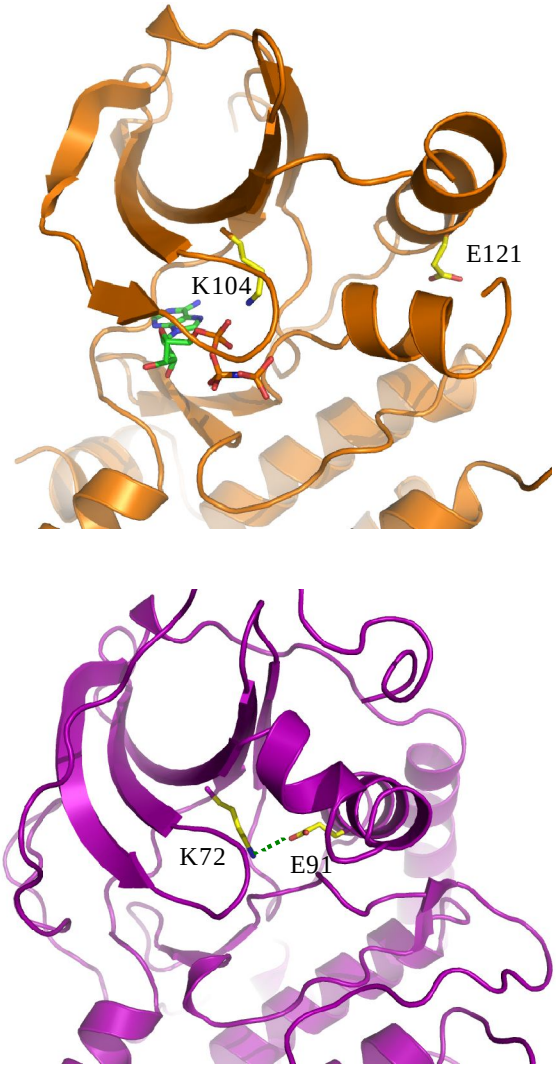


Figure 2-19 Comparison of SPAK and PKA in the ATP binding region

The ionic interaction between highly-conserved lysine and glutamate residues is lost in SPAK. SPAK is colored in orange (upper panel), and PKA is in purple (bottom panel). Both lysine and glutamate residues are marked in a stick representation.

The activation loop

The N-terminus of the activation loop adopts a unique short helix in the structure of SPAK in contrast to a β strand normally found in this region of an active kinase. As a result, the side chain of the catalytic Lys104 in β strand 3 makes an unusual contact with the invariant Asp222 in the DFG motif of the activation loop (Figure 2-20). The activation loop helix forms extensive contacts with α-helix C that are primarily hydrophobic. The involved residues include Leu118, Ile122, and Met125 in α-helix C, and Val225 and Phe228 in the activation loop. Short helical activation loop inserts have been reported in several low activity kinase structures such as cyclin-dependent kinase 2 (CDK2) and with-no-lysine [K] 1 (WNK1) (De Bondt et al, 1993; Min et al, 2004). This activation loop helix appears to displace the α-helix C from the active site, blocking the interaction between α helix C and the middle segment of the activation loop, and destroying the catalytic lysine-glutamate ion pair. This short helical structure also causes a rotation of the preceding DFG motif (residues 222-224) and particularly affects the position of Asp222. In the active state, this highly conserved Asp222 will interact with all three ATP phosphates directly or via coordination with magnesium. In our case, Asp222 rotates about 100° away from the catalytic site, even though it is still in near the AMP-PNP.

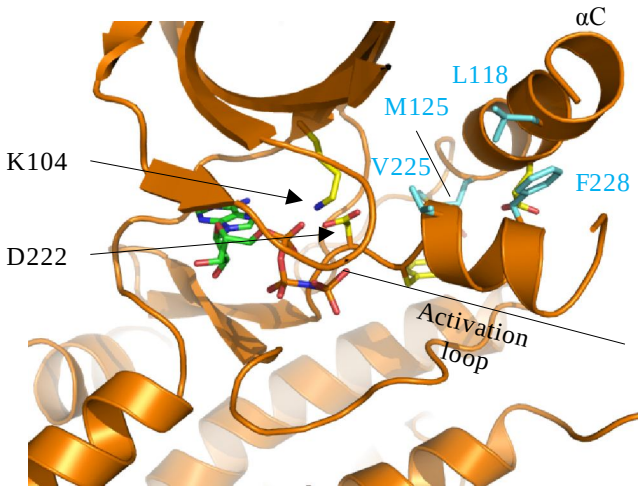


Figure 2-20 Interactions between the activation loop and the N-terminal domain of SPAK

Lys104 from β strand 3 forms contact with Asp222 in the activation loop. α-helix C also interact with the N-terminal helical region of the activation loop. Hydrophobic residues involved in the interaction are represented in cyan sticks.

The catalytic loop

The catalytic loop of SPAK, HRDLKxxN (residues 202-209), is well-ordered and superimposed well with that of other kinases. Residue Asn209 forms a series of expected interactions with the AMP-PNP phosphates. In addition controlling of the arrangement of ATP, the catalytic loop was also described to form important interactions with the activation loop, which is characterized by the arginine in the HRD motif. In the active state, a salt bridge which will form between the arginine and the negatively charged moiety from phosphates of the activation loop phosphorylation site is suggested to realign important residues in active site and support the configuration of the activation loop. This arginine, corresponding to Arg203 in SPAK, however, does not interact with the WNK phosphorylation site Thr243, which was replaced with aspartate in this SPAK mutant structure. Despite the fact that Asp243 is not clear in the structure, based on the location of its following residue Phe244, I estimate that Asp243 is 13 Å away from the expected position (Figure 2-21). In the domain swapped situation, although Asp243 is closer to Arg203 of the second molecule, it is still about 8 Å away. Instead, another potential phosphorylation site in the activation loop, Thr231, sits closer to Arg203, about 5 Å away. This is inconsistent with the present data suggesting that Thr243 is the key activation site in SPAK. As the crystal structure is in the low activity state, it is not possible to determine which residue in the activation loop is mainly responsible for kinase activation. Whether multiple phosphorylations are required for activation remains an open question. Structures of phosphorylated SPAK and proteins with different phosphorylation sites mutants might help to solve this problem.

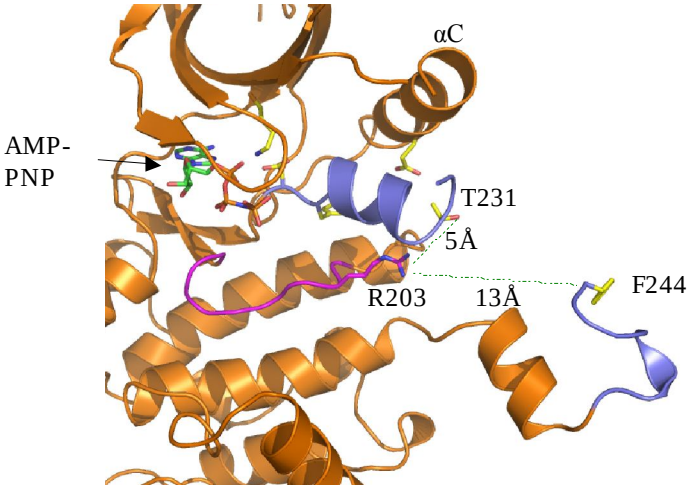


Figure 2-21 The active site conformation of SPAK

Lack of interaction between the catalytic loop Arg203 and the phosphorylation site in the activation loop. Arg203 is represented in magenta sticks. T231 and Phe244 are shown as yellow sticks. The catalytic loop and the activation loop are colored in magenta and blue, respectively.

The P+1 loop

The C-terminal region of the activation loop is known as the P+1 loop which in part forms the P+1 pocket for docking the substrate residue immediately following the phosphorylation site (P+1 site). One striking feature of SPAK is the formation of the P+1 substrate binding pocket in a domain-swapped fashion. Thr247 from the first P+1 loop is hydrogen-bonded to two essential catalytic loop residues, Asp204 and Lys206, in the second molecule (Figure 2-22a). Asp204 is the catalytic base, and Lys206 is believed to stabilize the transition state of the γ-phosphate. The interactions between these three residues, threonine, aspartate and lysine, are highly conserved in active kinases but, in contrast to this structure, they are all through residues from the same molecule.

The P+1 loop of SPAK contains the sequence FVGTPCWM (residues 244-251). Overall, the conformation is very similar to other active kinases such as PKA, except for the F244 in the beginning of this loop which is translated 7.6 Å away from the expected position (Figure 2-22b). This relocation is probably due to the lack of interactions between the preceding phosphorylation site-mimic residue Asp243 and the catalytic loop Arg203. The P+1 pocket of SPAK is hydrophobic and is formed by residues Phe244, Pro248, and Met251 from the P+1 loop and Leu294 from α-helix G. This is consistent with the presence of hydrophobic residues in the P+1 sites in substrates including NKCCs and PAK1.

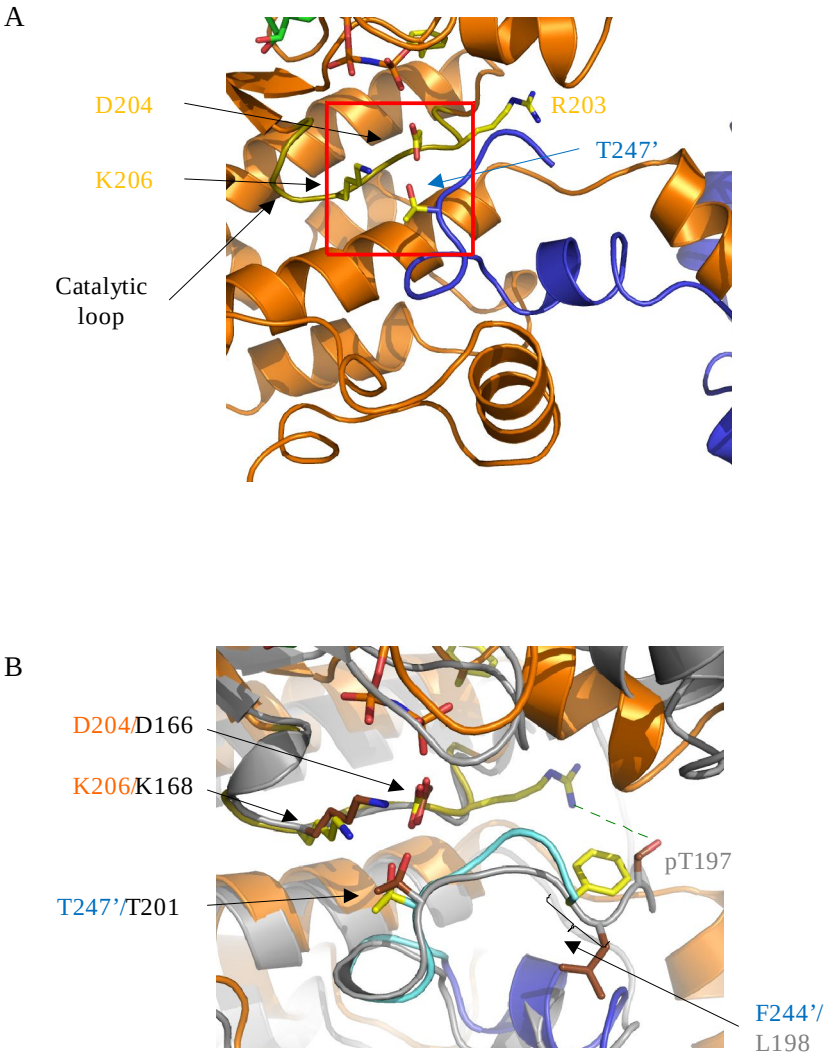


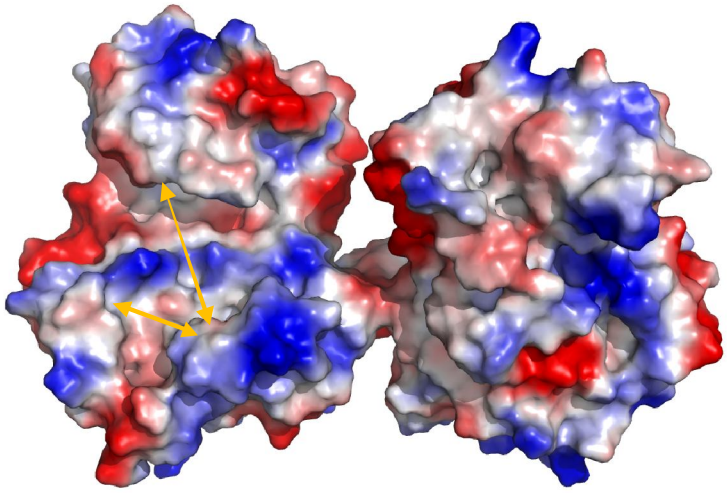
Figure 2-22 The P+1 pocket

(A) The P+1 pocket is formed in the interface of the domain-swapped dimer. The two monomers are colored in orange and blue, respectively. Residues involved in the interaction are shown as sticks.

(B) Comparison of SPAK and PKA in the P+1 pocket. The two SPAK monomers are colored in orange and blue, and PKA is in gray. Residues in the P+1 pocket are shown as sticks.

To understand more about the substrate specificity of SPAK, I examined the surface area of the substrate binding groove which lies in front of the P+1 pocket. Comparison of SPAK with PKA reveals that SPAK has a much broader substrate binding groove (Figure 2-23). Indeed, several SPAK substrates consist of several bulky hydrophobic residues N-terminal to the phosphorylation site. There is no distinguishable charged cluster found on the surface, implying that the substrate specificity of SPAK might not be determined by charge interactions. Although several small charged patches surround the groove, including a negatively charged patch near the P-3 position, a preference for positively charged residues at the P-3 position has not been obvious appeared in SPAK substrates. This inconsistency might be due to the limited number of known substrates. The substrates identified so far appear to have no obvious sequence similarities. Clearly, more substrates will be helpful to elucidate the substrate specificity of SPAK. The lack of consensus sequence around the phosphorylation site may also suggest that the substrate binding groove is not the only region for determining substrate specificity. The PF2 domain of SPAK also provides an extra specificity determinant by binding substrates which have a consensus sequence [(R/K)FX(V/I)]. Other substrate docking sites may also exist on the surface of the kinase domain. For example, when examining the surface of SPAK, an extended groove on the backside of the kinase domain between α-helix C and β-strand 4 was found (Figure 2-24). The functions of this groove require further biochemical and structural investigation.

A



NKCC1 YYD**T**HTN**T**YYLR**T**FGHNTMDA
NKCC2 AYD**S**HTN**T**YYLQ**T**FGHNTMDA
NCC SHL**T**HSSTFCMR**T**FGYN**T**IDV

B

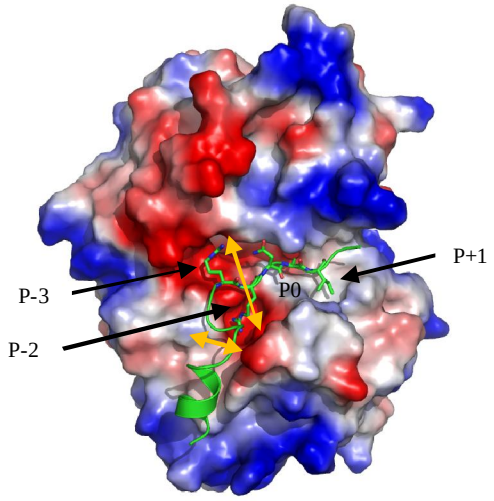


Figure 2-23 Surface comparison of SPAK and PKA at the substrate binding groove

Panel A shows the surface of SPAK domain-swapped dimer (top). SPAK phosphorylation sites in NKCC1, 2 and NCC are aligned in the bottom. Panel B shows the surface of PKA in complex with substrate represented in ribbon diagram. Negatively charged surface are colored in red and positively charged surface are in blue.

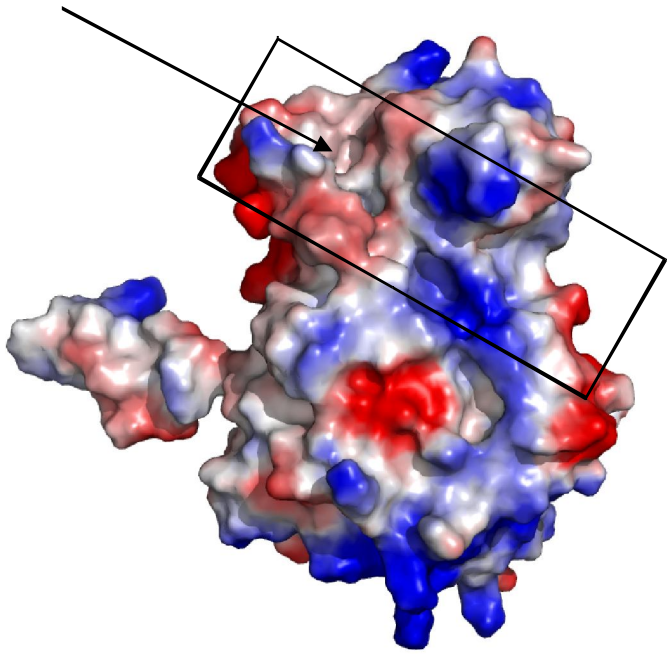


Figure 2-24 Surface representation of the backside of SPAK

An elongated groove indicated by the arrow is located between α-helix C and β strand 4. The approximate dimensions of the groove is 35 Å long and 8 Å wide.

The PF1 domain

The PF1 domain of SPAK consists of residues 351-402. The majority of the PF1 domain in the crystal structure of SPAK 63-390 (T243D) is missing. Only the N-terminal 15 residues (residues 351-365) are present. The N-terminal part of the PF1 domain forms an extra helix sitting on the groove at the back of the kinase close to α-helix D and β-strands 7 and 8. This binding of PF1 to the kinase domain is similar to the common docking (CD) interaction found in MAPKs (Chang et al, 2002). The CD domain functions as a substrate binding site. Several structural studies have demonstrated that binding of peptides to this region of MAPKs caused allosteric conformational changes and could lead to kinase activation (Zhou et al, 2006).

To determine if the PF1 domain uses the same mechanism as the MAPK docking motif to activate SPAK, I compared the kinase activity of three SPAK/OSR1 fragments: SPAK 63-390, SPAK 63-370 (contains the kinase domain plus the N-terminal helical region of the PF1 domain), and OSR1 1-295 (which only contains the kinase domain corresponding to SPAK63-350) (Figure 2-25). After phosphorylation by WNK1, although SPAK 63-370 is about 3-fold more active than OSR1 1-295, its activity is significantly lower than that of SPAK 63-390. This suggested that the C-terminal segment of PF1 containing residues 371-390 is essential for SPAK kinase activity whereas the N-terminal helical region (residues 351-370) can only induce marginal activation.

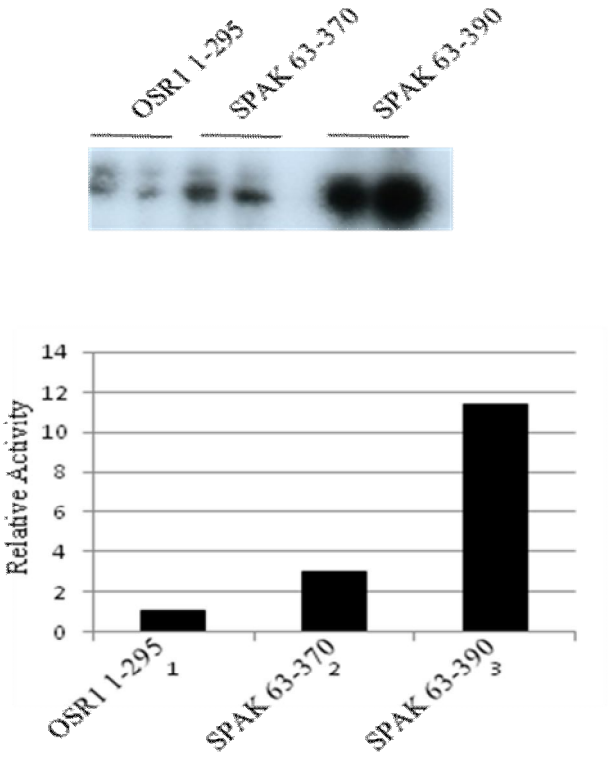


Figure 2-25 Kinase activity of WNK1 phosphorylated wildtype OSR1 1-295, SPAK 63-370, and SPAK 63-390

0.5 µg of the indicated enzymes were added to a 30 µl kinase reaction in 20 mM Hepes, pH 7.6, 10 µM ATP, 10 mM $MgCl_2$, 10 mM β-glycerophosphate, 1 mM DTT, 1 mM benzamidine, and 10 µCi [γ-^{32}P]ATP with 0.2 µg GST-WNK 160-490 as activating kinase and 5 µg GST-PAK1 1-230 as substrate. The reaction was carried out at 30°C for 30 min. Kinase activity of SPAK/OSR1 was determined by the phosphorylation of SPAK/OSR1 substrate GST-PAK1-231.

So far, it remains unclear why the PF1 domain is required for SPAK kinase activity. There are several possibilities. The PF1 domain might be necessary to induce some conformational changes during kinase activation. For instance, the PF1 domain might reposition α-helix C and allow ionic interactions to form between the ATP-binding Lys and Glu in the N-terminal domain. The C-terminal end of the PF1 domain in the structure projects toward α-helix C of another molecule. Although the distance between the C-terminus of the PF1 domain of the first molecule and α helix C of the second molecule is still about 10Å, as 25 residues are missing in the structure, it is possible that the PF1 domain may be used to reposition α-helix C. Considering that the conformation is relatively flexible in α-helix C in the SPAK structure, interaction of α helix C with PF1 may help stabilize the kinase. Such speculation is also supported by a separate experiment using co-immunoprecipitation assays, suggesting that the PF1 domain is required for self-oligomerization (unpublished data). The results showed that protein with the PF1 domain was able to bind to itself while protein that only has the kinase domain failed to do so.

The PF1 domain might function as an anchor as well. The interaction could be with its upstream kinase, downstream substrates, or scaffolding proteins. The binding might help to position these accessory proteins appropriately and thus facilitate the enzyme's catalytic activity. It might also trigger some conformational changes that are required for SPAK activation. Furthermore, interactions with other proteins might help to stabilize the PF1 domain. This might partially explain why we were unable to see the PF1 domain in our structure.

Comparison of SPAK 63-390(T243D) and OSR1 1-295

To understand the activation mechanism of SPAK, we compared the structure of SPAK 63-390 (T243D) in the partially active state and OSR1 in the inactive state. OSR1 is a homolog of SPAK and their kinase domains share more than 90 % sequence identity. The structure of OSR1 was recently solved in the unphosphorylated state and only contains the kinase domain (Lee et al, 2008). This truncated OSR1 also forms a domain swapped dimer.

The overall structure of SPAK 63-390 (T243D) is similar to OSR1 1-295 with an r.m.s. deviation for the main chain atoms of about 2.0 Å. A main chain superimposition of SPAK and OSR1 is shown in Figure 2-26. Notably, there are several regions showing significant differences between the two structures. First, the N-terminus of α-helix C with the preceding loop β3/αC shift upward. The difference is about 4.7 Å at the tip of α-helix C. Second, a more remarkable movement was found in the activation loop. The whole activation loop and the following α-helix EF move closer to the N-terminal domain. This rearrangement allows formation of the P+1 pocket in SPAK. Comparison of the P+1 pocket of SPAK and the corresponding region in OSR1 shows that the highly conserved P+1 loop threonine residues, Thr247 in SPAK and Thr189 in OSR1, are separated by 11 Å (Figure 2-26). Instead of interacting with the catalytic loop arginine, Thr189 in OSR1 is located near α-helix G of the second molecule. The conformational change in the activation loop also results in a broader substrate binding groove in SPAK, which becomes more accessible for substrate binding. A surface representation of SPAK

and OSR1 are shown in Figure 2-27. Although both SPAK and OSR1 form domain-swapped dimers, the orientation between two monomers is different. In SPAK, the two monomers are more separated than in OSR1 (Figure 2-27). As remodeling of the P+1 pocket has been reported in several kinases during kinase activation, we think that the structure of SPAK 63-390 (T243D) is in an intermediate state of activation. Finally, the glycine-rich loop also reveals some changes. In SPAK, the glycine-rich loop falls closer to the ATP binding pocket than it does in OSR1; it also adopts a closed conformation, resembling other active kinases.

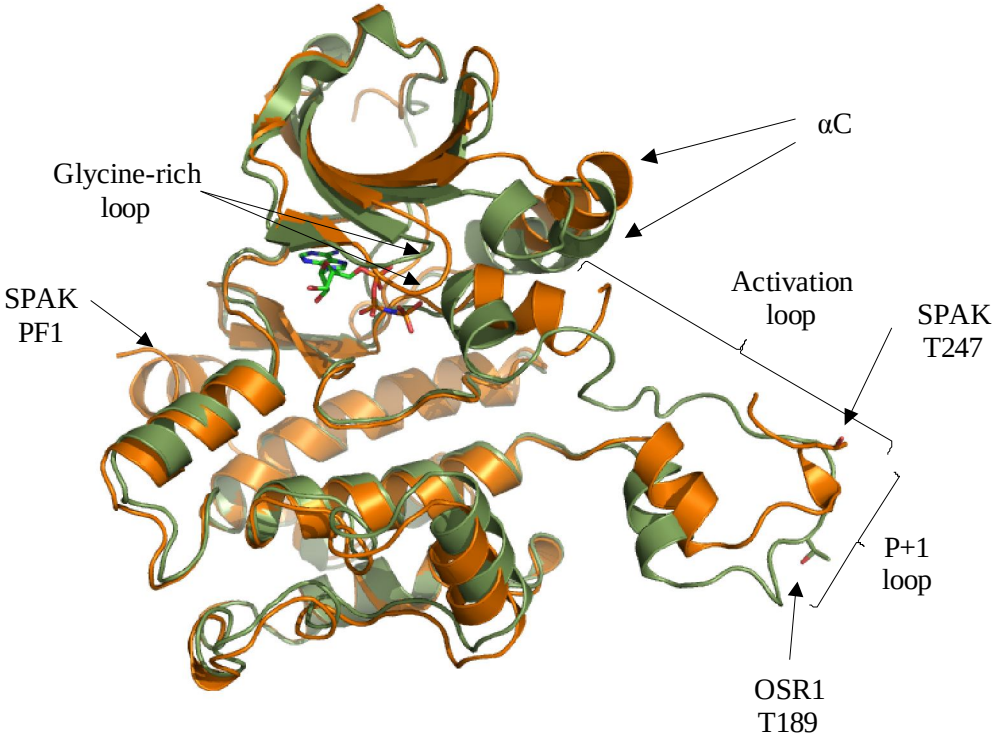


Figure 2-26 Superimposition of SPAK and OSR1

Ribbon diagrams of SPAK (orange) and OSR1 (green). AMP-PNP is shown as sticks.

SPAK OSR1

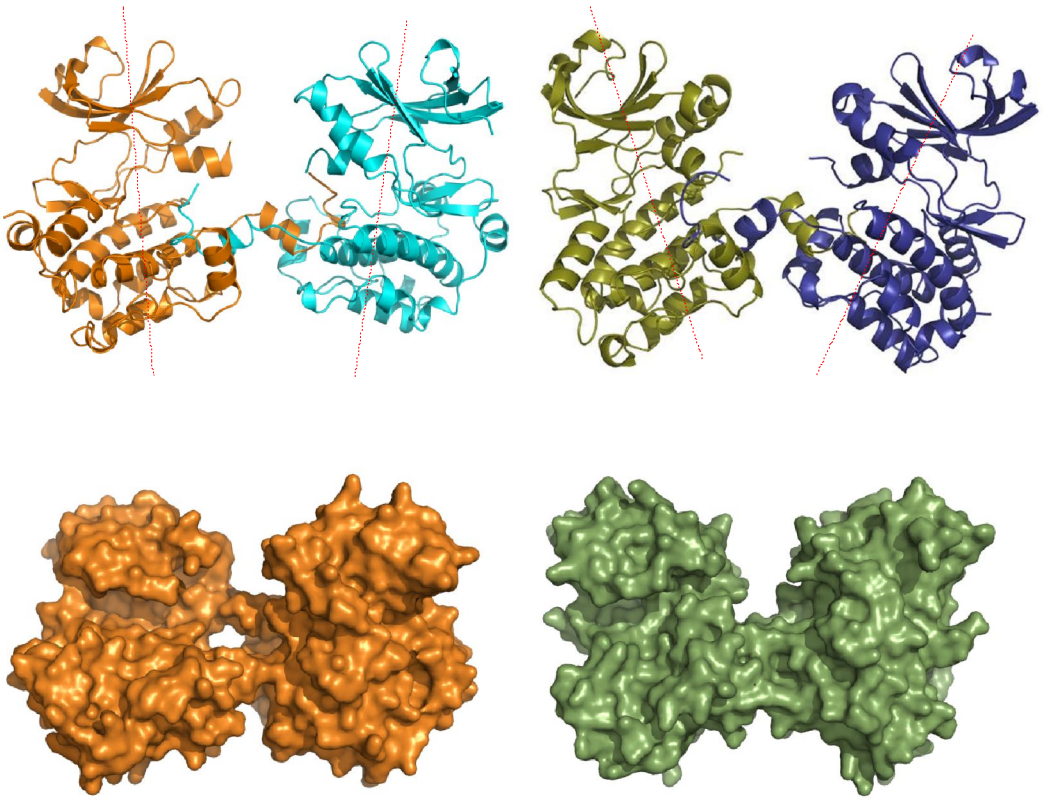


Figure 2-27 Comparison of domain-swapped dimer of SPAK and OSR1

The top panel shows the ribbon diagram of the domain-swapped dimers. The bottom panel represents the surface of the domain-swapped dimers. SPAK is shown in the left panel and OSR1 is in the right panel.

At present, it is not clear whether these conformational differences between SPAK and OSR1 are due to the T243D mutation or the inclusion of the PF1 domain. Further structural studies are required to distinguish between possibilities. For example, comparison of the results with a structure of SPAK T243D kinase domain but in the absence of the PF1 domain might provide insights into whether the PF1 can induce the conformational changes or it simply acts as a part of the structure that is required for stabilizing the kinase fold. On the other hand, by comparing with a structure of wildtype SPAK including the PF1 domain, we might understand the effect of the substitution of the phosphorylation site in the activation loop.

Based on biochemical studies, phosphorylation is the main regulatory mechanism for SPAK kinase activity (Anselmo et al, 2006; Vitari et al, 2005). Residues Thr243 and Ser383 have been shown to be phosphorylated by WNK kinases. While phosphorylation on Thr243 within the activation loop controls the kinase activity, phosphorylation on Ser383 residing in the PF1 domain appears to have no regulatory effect. Neither Thr243 nor Ser383 is visible in the electron density. Thr243 is close to the end of an ordered region which allows us to predict its location. As the adjacent residue Phe244 is translocated from the expected position, phosphorylation of Thr243 may induce a conformational change required for the subsequent rearrangement of the activation loop. As for Ser383, it is difficult to speculate any functional association with its phosphorylation. It should be noted that several other kinases also possess an additional phosphorylation site beyond the kinase domain (Yang et al, 2002). Some are required for kinase activity and are suggested to reorder a certain part of the structure. Further

structural studies are required to fully understand the function of Ser383 phosphorylation. I have attempted to crystallize SPAK63-390 (T243D/S383E) protein. The crystals grew under the same conditions and appeared to have no differences from SPAK63-390 (T243D), and further, the unit cell had the same organization as well. Thus, I do not expect SPAK63-390 (T243D/S383E) will adopt a different conformation from our current structure. As this is in the background of the T243D mutation, I do not know what conformational changes might be associated with S383E in unphosphorylated or phosphorylated proteins. Structural information from wildtype and phosphorylated SPAK with the presence of S383E mutation might provide insights into the consequence of Ser383 phosphorylation.

Another phosphorylation site found on SPAK is Ser321 which is phosphorylated by PKCθ and δ (Li et al, 2004; Smith et al, 2008). Although phosphorylation of Ser321 has been shown to activate the kinase, the underlying mechanism is not clear. In the crystal structure, Ser321 is located at the tip of α-helix H near the base of the C-terminal domain. This residue is in close proximity to a pocket formed by helices αE, αF, and αH. A structural study on c-Abl revealed that this is the binding pocket for the myristoyl group and the binding of the myristoylated N-terminus of c-Abl into this location induces a conformational change necessary for the subsequent binding of the SH2 and SH3 domains (Nagar et al, 2003). Such a feature implies flexibility at the base of the C-terminal domain of the kinase. Thus, phosphorylation of Ser321 in SPAK might also influence the conformation of this C-terminal region and lead to changes in kinase activity.

The activation of SPAK kinase is mediated by multiple complicated and tightly regulated processes. The structure shown here is in an intermediate state of activation. It only displays a few of the features usually found in active kinases. To understand the full activation mechanism, the crystal structure of phosphorylated SPAK is required. I have tried to co-express SPAK with WNK1 in bacteria. However, the phosphate incorporation was too low to facilitate purification of a homogeneously phosphorylated species, despite the fact that some protein was phosphorylated. The protein was crystallized in the same condition as SPAK 63-390 (T243D), but the quality of the crystals was much worse. *In vitro* phosphorylation is another way to generate the phospho-protein. This will be attempted in the future. Including a suitable binding protein during crystallization may be useful for stabilizing the conformation of the PF1 domain.

CHAPTER THREE

BIOCHEMICAL AND STRUCTURAL CHARACTERIZATION OF MUTATIONS IN THE ACTIVATION LOOP OF SPAK

Introduction

Previous structural studies of SPAK 63-390 (T243D) revealed an activation loop-swapped structure. Compared with the unphosphorylated, inactive OSR1 structure, the SPAK structure demonstrated dramatic conformational differences and suggested that SPAK 63-390 (T243D) adopts a conformation of intermediate activity. However, the structure contained several disordered regions including part of the activation loop and part of the PF1 domain. Because introducing mutations in flexible regions sometimes can reduce conformational entropy and thus help promote crystallization, I made several mutations within the disordered activation loop, including V235C, N238C, K239M, and Δ232-236, in hopes of better understanding the configuration of the activation loop. Biochemical properties, including protein dimerization and kinase activities, of these mutants were analyzed. We further crystallized all of them. Among all, SPAK 63-390 (V235C/T243D) produced high-quality crystals and thus allowed a high-resolution structural analysis. The crystal structure of SPAK 63-390 (V235C/T243D) was solved at a 2.5 Å resolution. Similar to the first SPAK structure, this structure adopts an activation loop-swapped conformation. However, unlike SPAK 63-390 (T243D), the activation loop is fully modeled in SPAK 63-390 (V235C/T243D) providing insights into the

interaction network of the pseudo-phosphorylation site, T243D. One striking feature of SPAK 63-390 (V235C/T243D) is that the two molecules in the dimer exhibit significant differences, particularly in the activation loop. Compared with SPAK 63-390 (T243D) and inactive OSR1, this structure reveals that one molecule adopts a conformation more similar to that of SPAK 63-390 (T243D) while the other is more like an inactive OSR1.

Materials and Methods

Generation of Mutants

Site-directed mutagenesis was carried out on the pHisParallel SPAK 63-390 (T243D) construct used in the previous study. Mutations containing V235C, N238C K239M, and Δ232-236 were made with the QuikChange (Stratagene) mutagenesis kit. All mutants were confirmed by sequencing.

Protein Expression and Purification

The expression and purification of all the SPAK mutation variants were carried out by the same procedures as described in Chapter Two. Proteins were produced in *E. coli* strain Rosetta (DE3). Induction was carried out by adding 0.5 mM IPTG once *E. coli* cells reached an OD_{600} of 0.5-0.6. The cells were then grown at 16°C overnight and harvested by centrifugation. The cells were lysed, sonicated, and insoluble material was

sedimented. Soluble material was loaded onto a nickel column (Qiagen). After washing, protein was eluted in a single step with 250 mM imidazole. The N-terminal 6xHis tag was removed with TEV protease, and the cleaved protein was separated from the tag on another nickel column. Protein was further purified on a Mono S cation exchange column (Pharmacia) followed by a Superdex75 16/60 gel filtration column (Pharmacia). The final protein product was stored in buffer containing 50 mM Hepes at pH 7.0, 50 mM NaCl, 1 mM EDTA, and 5 mM DTT.

The SPAK 63-390 (V235C/T243D) protein was expressed to incorporate selenomethionine. The expression of selenomethionine-substituted protein was carried out by the methionine pathway inhibition method (Doublie, 1997). Protein was expressed in Rosetta (DE3) cells using M9 minimal medium supplemented with required nutrients. As OD_{600} reached 0.7, 0.5 mM IPTG, 60 mg selenomethionine, 100 mg each of threonine, lysine, and phenylalanine, and 50 mg each of leucine, isoleucine, and valine were added. After induction overnight at 16 °C, cells were harvested and lysed. Purification of the selenomethionine derivative protein was performed using the same protocol as described above except 5 mM DTT were added in every solution to prevent selenium oxidation.

Protein Crystallization and Data Collection

Prior to crystallization, all proteins containing different activation loop mutations were concentrated to approximately 25 mg/ml and incubated with 5 mM AMP-PNP for

30 min. Crystals were grown at 20ºC in hanging drops by mixing the protein with an equal volume of the well solution previously identified for SPAK 63-390 (T243D), which is 0.2 M Mg acetate$_2$, 0.1 M Tris pH 8.5, 16~18% PEG3350, 0.01 M Na cacodylate, and 0.01 M $CaCl_2$. All crystallization conditions were optimized by changing the buffer pH and precipitant concentrations. Crystal quality was further improved using additive screening. The SeMet-substituted protein was crystallized in a condition similar to the native protein, but the crystals were much smaller.

Heavy atom soaking experiments were also carried out on SPAK 63-390 (V235C/T243D) crystals. Different heavy atom compounds were tried including mercury, gold, and platinum derivatives. Only by soaking the crystals with 2 mM $AuK(CN)_2$ for 10 min were gold atoms stably incorporated in the crystals.

Diffraction data for native, gold, and selenomethionine derivative crystals were collected at the beamline 19-ID (λ=0.9794 Å) at the Advanced Photon Source (Argonne National Laboratory, Argonne, IL). Data were indexed, integrated, and scaled using the HKL2000 program package (Z. Otwinowski and W. Minor, 1997). Crystals of SPAK 63-390 (V235C/T243D) are orthorhombic with unit cell parameters of a =54.68 Å, b =69.94 Å, c =180.41 Å, $\alpha=\beta=\gamma=$ 90°, and a space group of $P2_12_12_1$. The unit cell is different from the previous SPAK 63-390 (T243D) crystals. There are two molecules in an asymmetric unit as determined by Matthews coefficient.

The structure of SPAK 63-390 (V235C/T243D) was solved by molecular replacement using the program Phaser in the CCP4 package. SPAK 63-390 (T243D) was used as the search model, followed by rigid body and restrained refinements using

Refmac5 (Murshudov et al, 1997). Phases from the partially refined model were used to calculate anomalous difference Fourier maps to locate the positions of selenomethionine residues and gold atoms in the structure. The locations of these heavy atoms were used for the subsequent model building. The structure was manually built with the graphic program Coot (Emsley & Cowtan, 2004), and was refined by TLS and restrained refinements in Refmac5. The bond angle and bond length of the final model were confirmed by the MolProbity web server.

Protein kinase assay

Activities of mutant SPAK proteins were determined by in vitro kinase assay using GST-PAK (1-231) as substrate. Assays were performed in 30 µl of kinase buffer (20 mM Hepes at pH7.6, 10 mM $MgCl_2$, 1 mM benzamidine, 1 mM DTT, and 10 µM ATP) with 10 µCi [γ-^{32}P]ATP, 5 µg PAK, and 0.5 µg SPAK at 30°C for 30 min. Reactions were stopped by the addition of 5x SDS sample buffer and followed by SDS-PAGE and autoradiography.

Results and Discussion

Characterization of dimerization and kinase activity of SPAK activation loop mutants

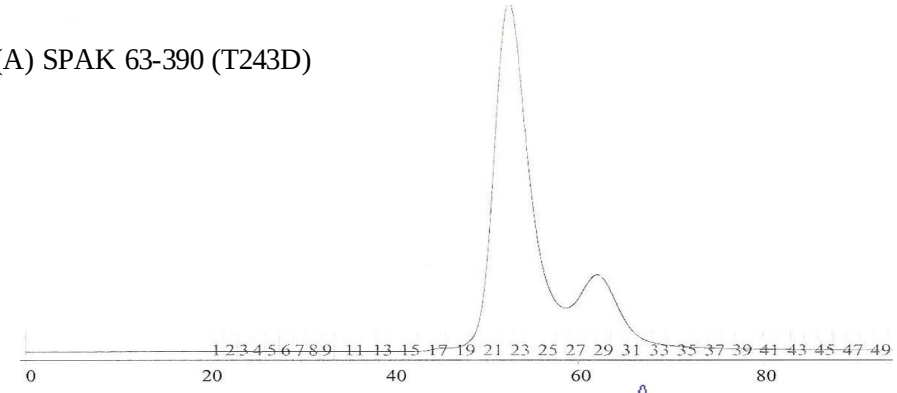
Previously, the crystal structure of SPAK 63-390 (T243D) showed that residues 234-243 within the activation loop were missing in the structure (Chapter 2). As the activation loop is one of the most flexible regions in the kinase domain, we randomly mutated some of these omitted residues to examine if mutations can stabilize the activation loop conformation. The mutations including V235C, N238C, K239M, and Δ232-236 (deleting residues 232 to 236) were added to the original T243D construct. Cysteine and methionine were used to create an extra heavy atom binding site, because the sulfur atoms in these residues generally react with heavy atom compounds. Increasing the number of heavy atom sites might improve the anomalous signal for phasing and might also stabilize the configuration upon heavy atom binding.

All the mutant variants displayed equivalent expression; however, they showed differences in the ability to dimerize. The SPAK 63-390 (T243D) protein previous studied was demonstrated to form a monomer-dimer equilibrium in solution, as determined by size exclusion chromatography profile (Figure 3-1). The monomer and dimer fractions can be even separated in an earlier Mono S ion exchange column. Protein eluted from first Mono S peak (200-250 mM NaCl) was eluted at around 70 ml in the Superdex75 column which corresponds to a monomer, while protein eluted from the second Mono S peak (450-700 mM NaCl) was mainly eluted at about 60 ml in the Superdex75 which corresponds to a dimer. The ratio of monomer to dimer was not affected by increasing the concentration of NaCl from 50 mM to 200 mM. Three

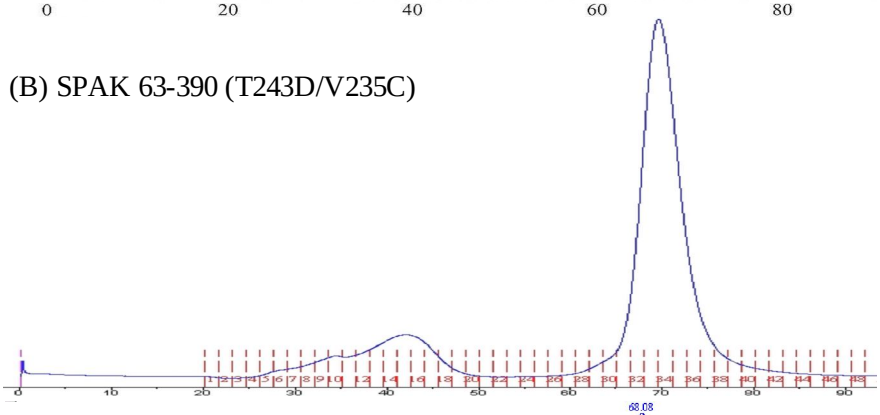
mutants, V235C, N238C, and K239M, showed similar Mono S profiles as SPAK 63-390 (T243D). Nonetheless, when further purified by gel filtration, only K239M retained the dimer population. V235C and N238C proteins were almost completely monomers (Figure 3-1). Proteins from both Mono S peaks were eluted in the monomeric range (around 70 ml) in the Superdex75 column; however, a peak containing higher-order oligomers (around 40 ml) was also found. As for the K239M protein, the first peak eluted from the Mono S remained a monomer. Protein from the second Mono S peak was eluted as two unresolved peaks, corresponding to a dimer and a monomer (Figure 3-1). The ratio of dimer to monomer is about 1:1 and the dimer fraction is less than in SPAK 63-390 (T243D).

To determine whether these mutations in the activation loop affect SPAK activity, we performed in vitro kinase assays. GST-PAK1 (1-231) was used as the substrate for SPAK. Compared with SPAK 63-390 (T243D), the V235C mutant showed less activity toward PAK1, but the difference was typically less than two fold. The SPAK mutant Δ232-236 was consistently inactive (Figure 3-2).

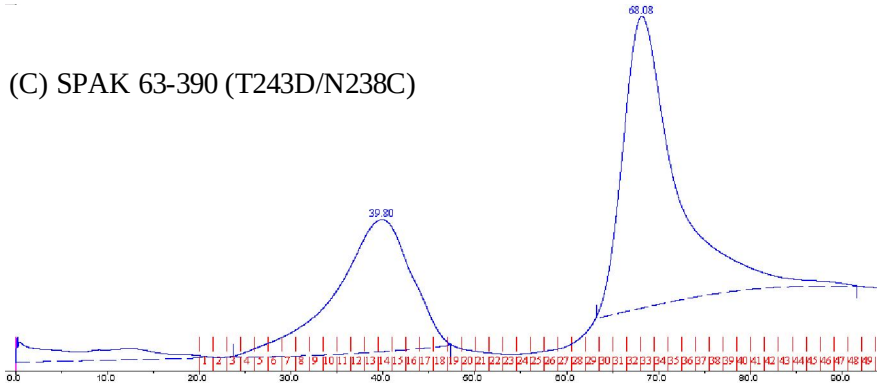
(A) SPAK 63-390 (T243D)



(B) SPAK 63-390 (T243D/V235C)



(C) SPAK 63-390 (T243D/N238C)



(D) SPAK 63-390 (T243D/K239M)

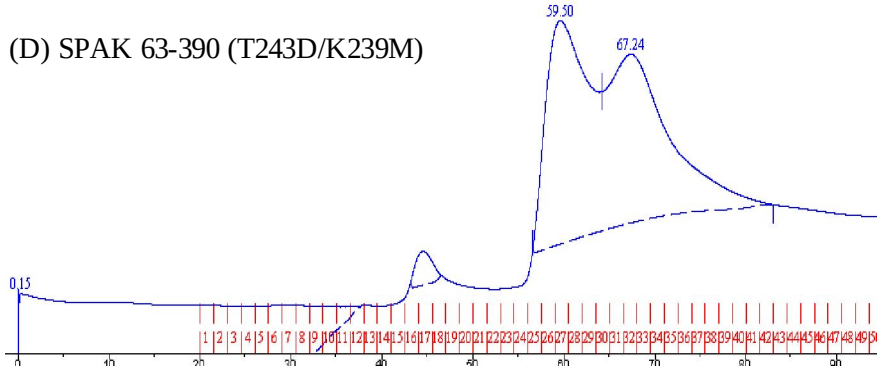


Figure 3-1 Gel filtration of different SPAK activation loop mutants

Proteins SPAK 63-390 (T243D) (A), SPAK 63-390 (T243D/V235C) (B), SPAK 63-390 (T243D/N238C) (C), and SPAK 63-390 (T243D/K239M) were applied to a Superdex75 16/60 gel filtration column. Protein elution was detected by OD_{260}.

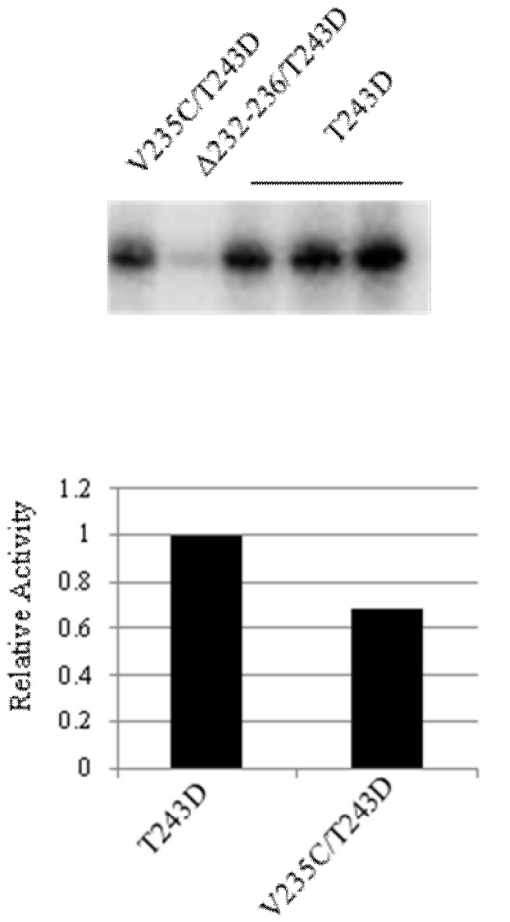


Figure 3-2 Kinase activity of different SPAK activation loop mutants

The kinase activity of different SPAK mutants was determined by the phosphorylation of its substrate GST-PAK1-231.

Crystallization of SPAK activation loop mutants

To examine the structures of these activation loop mutants, I crystallized all the mutant proteins. All mutant variants can be crystallized in the same conditions as their prototype protein SPAK 63-390 (T243D). However, the quality of the crystals differed. Mutant K239M formed exactly the same type crystals as the starting SPAK 63-390 (T243D). The crystals diffracted to the same extent and had the same unit cell parameters. It was not further pursued. Mutant N238C formed tiny crystals, with the longest dimension not exceeding 100 μm even after optimization. The mutant SPAK Δ232-236 T243D gave crystals that were plate-like clusters, and no single crystals were obtained.

Among the four activation loop mutants, SPAK 63-390 (V235C/T243D) showed the best crystals. We further optimized its crystallization condition and solved the structure. Using the condition mentioned above, SPAK 63-390 (V235C/T243D) protein formed large plate-shaped crystals (>500 μm) which were distinct from those of SPAK 63-390 (T243D). After optimization and replacement of 0.01M $CaCl_2$ to 3% aminocaproic acid, identified from the Hampton Additive Screen kit, higher quality and three-dimensional crystals were generated (Figure 3-3). The crystals were diffracted to 2.5 Å at the synchrotron beamline 19-ID (Figure 3-4).

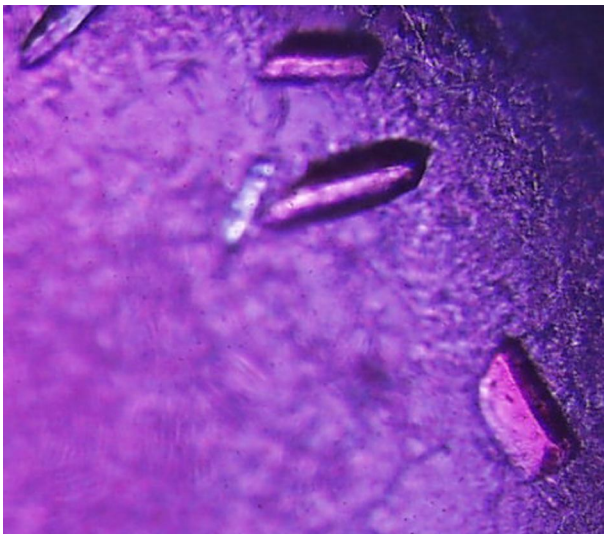


Figure 3-3 Crystals of native SPAK 63-390 (V235C/T243D)

The crystals were grown using the hanging drop vapor diffusion method at 20°C. The well solution contains 0.2 M magnesium acetate, 0.1 M Tris pH 8.5, 16% PEG3350, 0.01 M sodium cacodylate, and 3% aminocaproic acid. The protein concentration was 25 mg/ml.

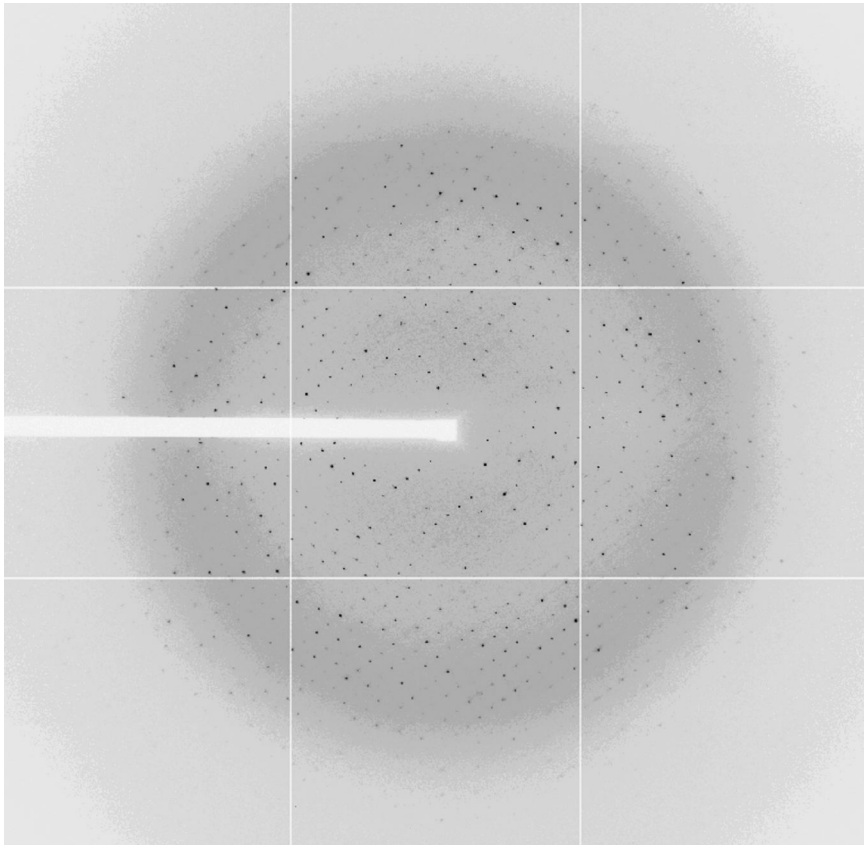


Figure 3-4 X-ray diffraction pattern of a native SPAK 63-390 (V235C/T243D) crystal

Diffraction data was collected at beamline 19-ID at the Advance Photon Source, Argonne National Laboratory (Argonne, IL).

Resolution is 2.5 Å.

Unit cell is a =54.68 Å, b =69.94 Å, c =180.41 Å, $\alpha=\beta=\gamma= 90°$, and the space group is $P2_12_12_1$.

The structure of SPAK 63-390 (V235C/T243D)

The crystal structure of SPAK 63-390 (V235C/T243D) was refined to a 2.5 Å resolution with R_{work} = 20 % and R_{free} = 25.4 %. The final model contains two monomers in the asymmetric unit (labeled A and B), in the form of a domain-swapped dimer (Figure 3-6A). Molecule A contains residues 66-111 and 114-364, while molecule B contains residues 59-364. There are a total of 603 amino acid residues and 2 AMP-PNP molecules in the structure. Residues missing from the structure are as follows: molecule A: 62-65, 112-113 and 365-390; molecule B: 365-390. Although a cysteine residue was introduced in this mutant protein, there is no disulfide bond formed in the structure. The crystal data and refinement statistics are summarized in Table 3-1.

Table 3-1 Summary of crystal data and refinement statistics

Data	SPAK 63-390 (V235C/T243D)
Space group	P 21 21 21
Cell dimensions (Å)	a= 54.68 b= 69.94 c= 180.41
α, β, γ	90°, 90°, 90°
Resolution (Å)	50-2.5
Observed reflection	506176
Unique reflection	24737
Completeness (outer shell)	97.5 (99.7)
R_{merge} (outer shell)[1]	0.07 (0.66)
I/> (outer shell)	22.76 (2.34)
Refinement	
R_{work}[2]	0.20
R_{free}[3]	0.25
R.m.s.d. bond length	0.011 Å
R.m.s.d. bond angle	1.312°
Average B factor (Å^2)	29.54

[1] $R_{merge} = \Sigma|(I_{hkl}) - \langle I \rangle|/\Sigma(I_{hkl})$, where I_{hkl} is the integrated intensity of a given reflection

[2] $R_{work} = \Sigma_{hkl}|Fo - Fc|/\Sigma_{hkl}|Fo|$, where Fo and Fc are the observed and calculated structure factors, respectively.

[3] Five percent of the reflections were used in the calculation of R_{free}.

The overall structure of SPAK 63-390 (V235C/T243D)

Overall, the structure of SPAK 63-390 (V235C/T243D) is similar to SPAK 63-390 (T243D) with an average r.m.s. deviation of 1.2 Å (Figure 3-5). The active site including the ATP-binding pocket, the catalytic loop, and the P+1 loop all adopt similar features between the two structures. Furthermore, the two dimer interfaces identified in SPAK 63-390 (T243D), one between the two domain-swapped molecules and the other between an N-terminal domain and a C-terminal domain, are also conserved in SPAK 63-390 (V235C/T243D) (Figure 3-6). This suggests that mutating residue Val235 does not inhibit activation loop swapping or dimerization occurring between the N- and C-domains of two molecules, which is inconsistent with the monomeric status we observed in the gel filtration column. It is not clear how the dimer interfaces found in the crystal structure relate to the behavior in solution. However, it is possible that the V235C mutation does not completely prevent dimerization but instead increases the dissociation constant (K_d). With the presence of the high protein concentration during crystallization, the dimerization may be facilitated in the crystal.

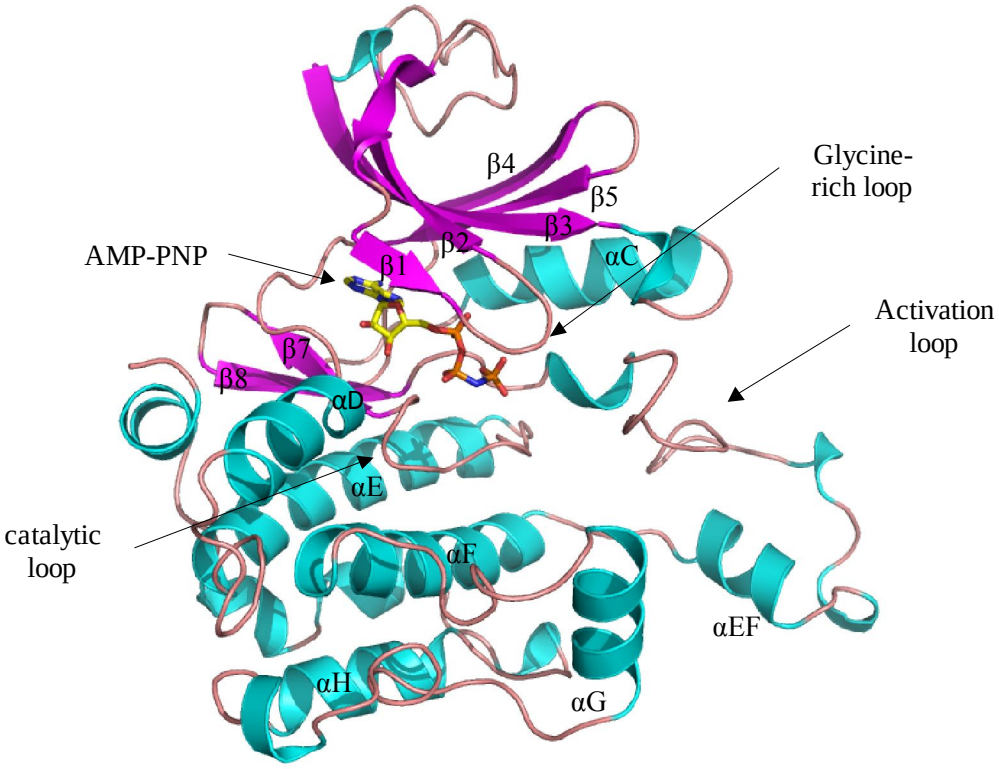
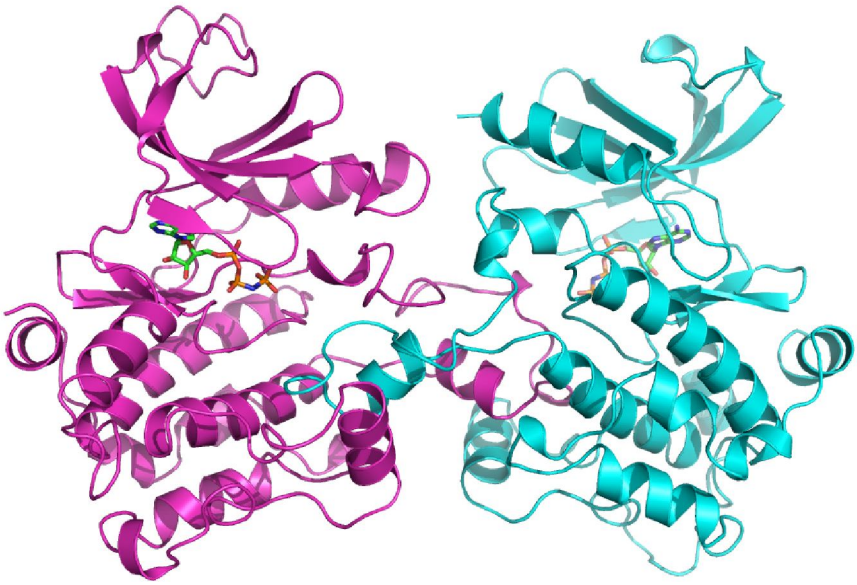


Figure 3-5 General topology of SPAK 63-390 (V235C/T243D)

Ribbon diagram of a monomeric SPAK structure. α helices are shown in cyan and β strands are shown in magenta. AMP-PNP is shown in sticks.

A



B

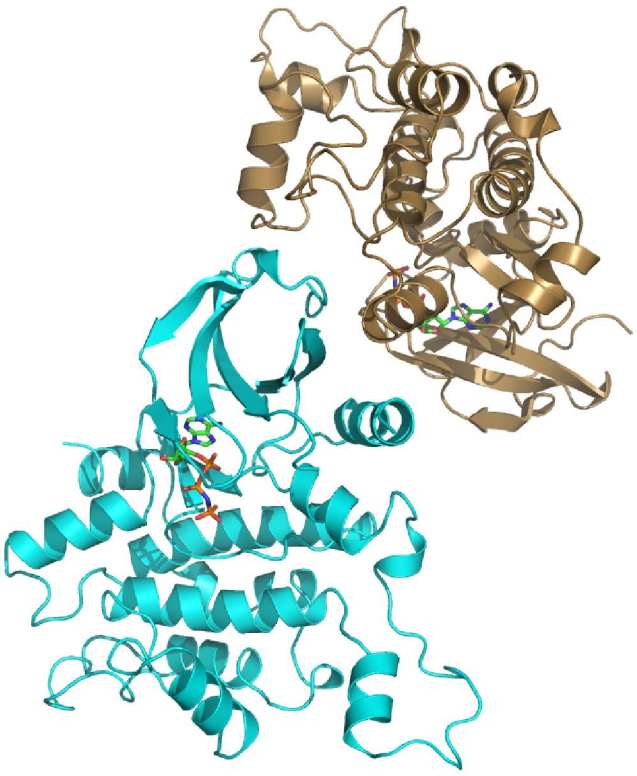


Figure 3-6 Ribbon diagram of SPAK 63-390 (V235C/T243D) dimer

The domain swapped dimer is shown in (A) with two monomers colored in cyan and magenta, respectively. The second dimer interface is formed by the interaction between an N-terminal domain with another C-terminal domain, as shown in (B). The two monomers are colored in cyan and tan, respectively. AMP-PNPs are shown in sticks.

Comparison of two monomers in the asymmetric unit

One unique feature of the structure of SPAK 63-390 (V235C/T243D) is that the two molecules in the asymmetric unit show some structural differences, even though the overall r.m.s. deviation between the two structures is not exceedingly high (1.8 Å between the superimposed Cα atoms). A superimposition of two molecules is shown in Figure 3-7. Major differences are located at the glycine-rich loop, α-helix C, and the activation loop. One molecule adopts a closed conformation in which the glycine-rich loop is closer to the ATP-binding site, while the other adopts a more open conformation. Nevertheless, when compared with the inactive OSR1 structure, both of the two molecules in the SPAK mutant still display a more closed conformation, which is consistent with the higher kinase activity than observed in unphosphorylated OSR1. As the mutant residue V235C does not directly contact the kinase core or the ATP-binding site, the exact structural basis of these movements is not clear. It seems likely to be a result of an allosteric effect from the activation loop.

The activation loop also shows significant differences. The most different area begins from the segment after the highly conserved DFG motif in the activation loop to the end of α helix EF (residues 230-260) (Figure 3-7). The P+1 loop and α helix EF are translated approximately 4 Å away from the two molecules. In one molecule, this region is positioned closer to the base of the C-terminal domain. This movement results in the loss of contact between Thr247 in the P+1 loop and Asp204 and Lys206 in the catalytic loop of the second molecule. The interactions between Thr247, Asp204 and Lys206 are conserved in active kinase structures which are required for the formation of the P+1

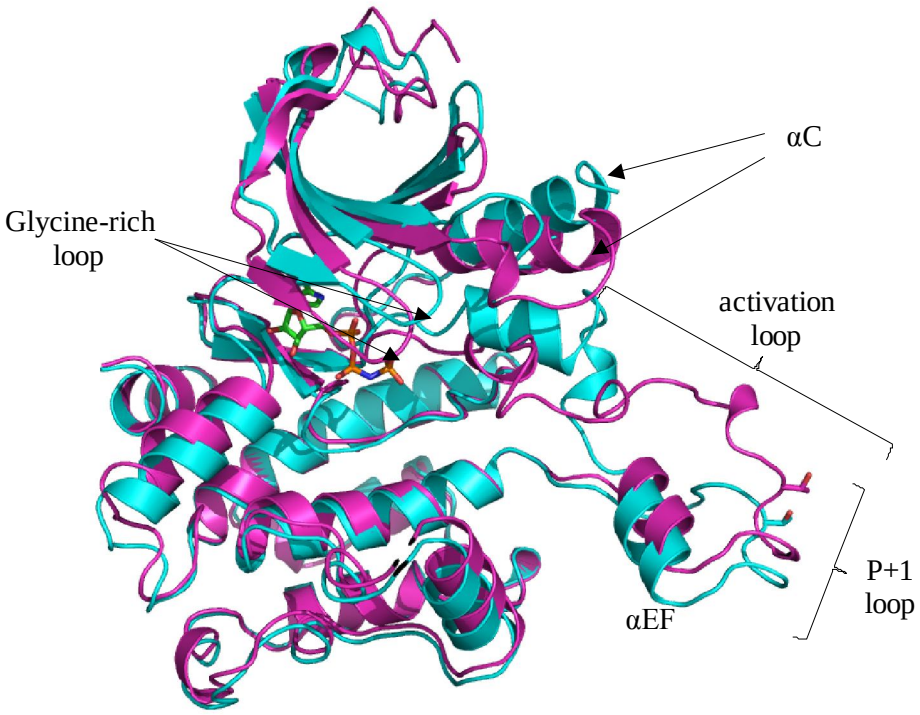


Figure 3-7 Superimposition of the two SPAK 63-390 (V235C/T243D) molecules

The two molecules are colored in cyan and magenta. The conserved Thr247 in the P+1 loop is represented in sticks. AMP-PNP is also shown in sticks.

pocket used for docking the P+1 site in the substrate. Comparison of the P+1 loop of SPAK 63-390 (V235C/T243D) with that of inactive OSR1 and SPAK 63-390 (T243D) reveals that SPAK 63-390 (V235C/T243D) is in a conformation intermediate between those of OSR1 and SPAK 63-390 (T243D). One molecule is more similar to SPAK 63-390 (T243D), while the other one, with the displacement in the P+1 loop and helix αEF, adopts a conformation closer to inactive OSR1 (Figure 3-8). The intermediate conformations found in SPAK 63-390 (V235C/T243D) might explain its lower kinase activity when compared to SPAK 63-390 (T243D). The structural changes in the P+1 site suggest that the P+1 loop is flexible. This flexibility allows the refolding of the P+1 pocket that is associated with SPAK activation.

Within the activation loop, the segment between the DFG motif and the P+1 loop shows a drastic difference between the two molecules in the SPAK 63-390 (V235C/T243D) structure (Figure 3-7). Notably, the B-factors of this segment are higher than that of neighboring amino acid residues (45 $Å^2$ versus 25 $Å^2$), indicating partial occupancy or disordering of this region. Indeed, this same region is missing in the structure of SPAK 63-390 (T243D). Nevertheless, the electron density in SPAK 63-390 (V235C/T243D) is clear enough to allow unambiguous assignment of both main chains and side chains of the residues. Currently, it is not clear whether this structural difference in the activation loop is due to crystal packing or reflects a conformational switch between inactive and active states. Structural information from phosphorylated SPAK as opposed to phosphomimic mutants is required to solve this problem. However, it is clear

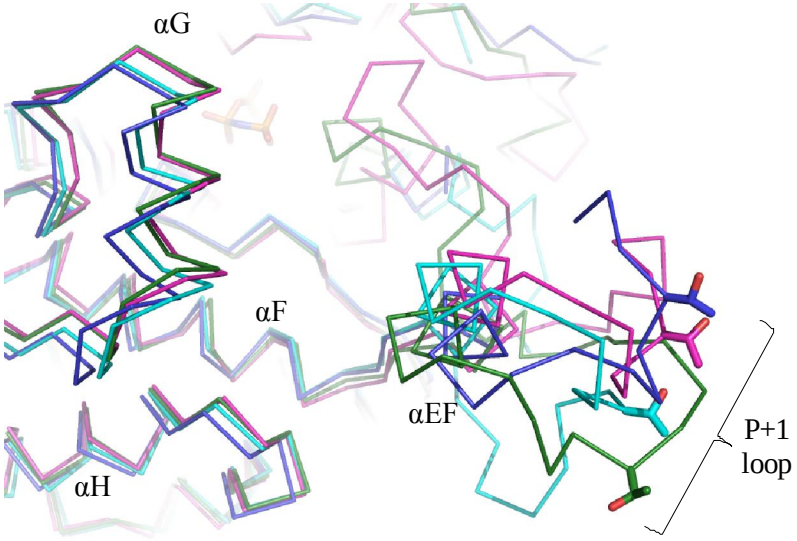


Figure 3-8 Superimposition of the P+1 loop in the two SPAK 63-390 (V235C/T243D) molecules in cyan and magenta, SPAK 63-390 (T243D) in dark blue, and inactive OSR1 1-295 in dark green.

Only Cα traces of proteins are shown. The conserved threonine residues, Thr247 in SPAK and Thr 189 in OSR1, are shown in sticks.

that the activation loop is comparably more flexible than other parts of the kinase domain and is undergoing some structural rearrangements during activation.

Interaction network of the phosphorylated residue

As the activation loop is fully modeled in the structure of SPAK 63-390 (V235C/T243D), we inspected the interactions formed through the WNK phosphorylation site, Thr243, within the activation loop to understand the activation mechanism. For most kinases, phosphorylation in the activation loop is required for kinase activation. The phosphate forms ion pairs with a basic residue in the activation loop and a highly conserved arginine (Arg203 in SPAK) situated in the catalytic loop (Figure 3-9A). These interactions stabilize the activation loop in a conformation that allows catalysis and substrate binding. In the crystal structure of SPAK 63-390 (V235C/T243D), the WNK phosphorylation site Thr243 is mutated to aspartate to mimic the phosphorylated state. Although the region around Asp243 exhibits a different conformation in each molecule of the SPAK dimer, Asp243 in both molecules forms an intramolecular ionic interaction with Arg241 within the activation loop (Figure 3-9B). This bonding is similar to that observed in the monomeric active kinases such as PKA and PAKs, suggesting a conserved mechanism used in both monomeric and domain-swapped kinases. In contrast, Thr185 in OSR1, corresponding to Thr243 in SPAK, is solvent exposed and does not contact any protein residue. Thus, the ion-pairing interaction between phospho-mimic T243D and basic Arg241 in SPAK is likely to be important for the massive conformational changes between SPAK and OSR1. However,

the canonical interaction between the catalytic loop Arg203 and the phospho-mimic Asp243 is not observed in the structure. This might explain why mutation of Thr243 to Asp could not fully activate the kinase.

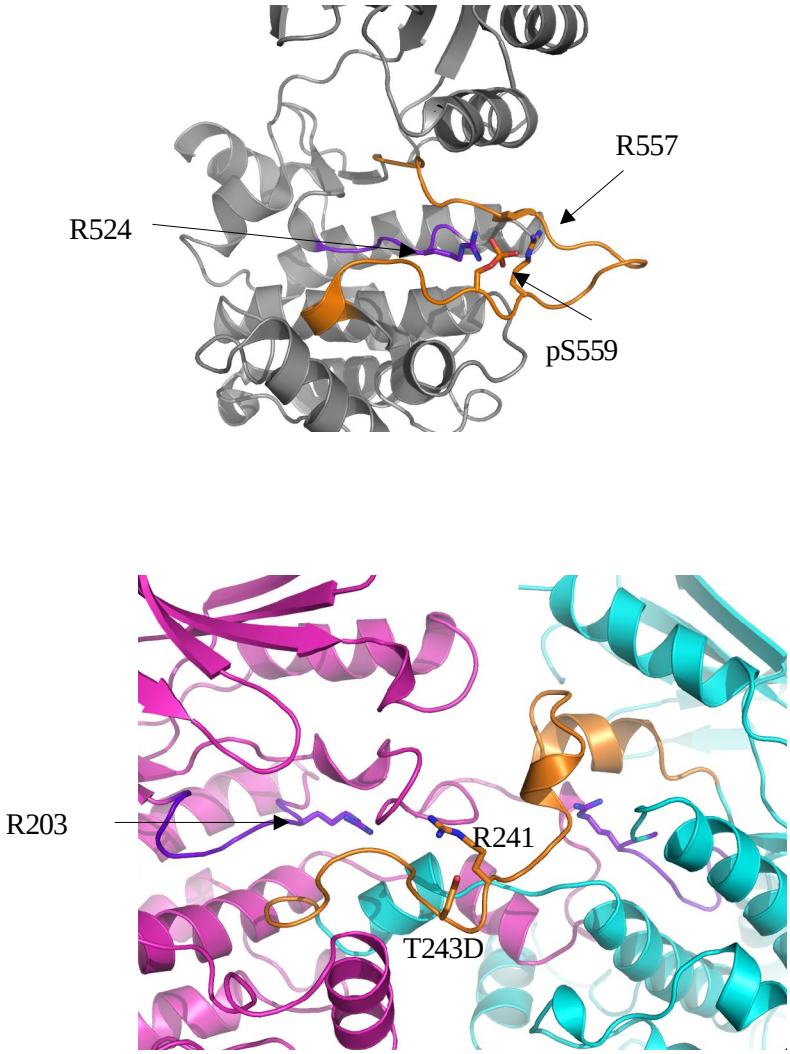


Figure 3-9 Comparison of SPAK and PAK6 around the phosphorylated activation loop

The activation loop and catalytic loop are colored in orange and purple, respectively. The top panel shows the interactions formed by the phosphorylated residue in PAK6 (pdb code 2c30). The phosphorylated residue Ser559 forms a hydrogen bond network with Arg557 in the activation loop and Arg524 in the catalytic loop. The bottom panel shows the interaction between T243D and Arg241 in the activation loop. The catalytic loop Arg203 is shown in purple sticks.

CHAPTER FOUR

EXPRESSION, PURIFICATION AND CRYSTALLIZATION OF MOUSE SGK1

Introduction

Serum- and glucocorticoid- induced kinase 1 (SGK1), a serine/threonine protein kinase, is a member of the AGC protein kinase subfamily, which includes protein kinases A, G and C (Webster et al, 1993b). As its name suggests, SGK1 was first identified as an immediate early gene whose transcription was increased by serum or glucocorticoid stimulation in epithelial cells (Webster et al, 1993a). Most functional studies on SGK1 have focused on its role in the regulation of ion channels, salt homeostasis, and moreover, the control of blood pressure. Numerous ion channels have been reported to be regulated by SGK1, including the epithelial sodium channel (ENaC), the renal outer medullary potassium channel (ROMK), and the cystic fibrosis transmembrane conductance regulator chloride channel (CFTR) (Lang et al, 2003; Naray-Fejes-Toth et al, 1999; Yoo et al, 2003). One of the mechanisms for SGK1 to regulate these ion channels is through phosphorylating Nedd4-2, a HECT domain E3 ubiquitin-protein ligase which adds ubiquitin to the channels and triggers their endocytosis (Debonneville et al, 2001). SGK1 phosphorylates Nedd4-2 and reduces the interaction between Nedd4-2 and ENaC, leading to an increased amount of ENaC on the cell surface. The colocalization of SGK1 with ENaC in the renal distal tubule and collecting duct suggests an important role for SGK1

in the regulation of sodium balance. Indeed, SGK1-knockout mice display an abnormality in sodium retention which results in low blood pressure.

SGK1 encodes a protein of 432 amino acids. The kinase domain is located between amino acids 97-346 and shares about 54% sequence identity with that of Akt/PKB. Similar to other AGC kinases, SGK1 contains a C-terminal extension (amino acids 347-432) following the kinase domain. This C-terminal extension is required for kinase activity and has a conserved hydrophobic motif (F-X-X-F-S/T-F/Y) near the end of the protein (amino acids 418-423) (Figure 4-1). The kinase activity of SGK1 is regulated by phosphorylation. Several phosphorylation sites have been identified, including Ser74, Ser78, Thr256, Ser397, Ser401, and Ser422 (Chen et al, 2008; Zhao et al, 2007). Among these phosphorylation sites, phosphorylation on both Thr256 and Ser422 is required for full SGK1 activation. Thr256 within the activation loop is a conserved phosphorylation site throughout the serine/threonine kinases. Phosphorylation of this residue results in a conformational change in the activation loop, which in turn rearranges the catalytic residues as well as the substrate binding site. Like closely-related kinase, Akt/PKB, Thr256 in SGK1 is phosphorylated by phosphoinositide-dependent protein kinase (PDK1) (Kobayashi & Cohen, 1999; Park et al, 1999). This phosphorylation is triggered by insulin and growth factors through a phosphatidylinositol 3-kinase (PI3K) dependent pathway. As for Ser422, which is located in the hydrophobic motif, the kinase responsible for its phosphorylation is less clear. Several candidates have been reported, including mammalian target of rapamycin (mTOR) and DNA-dependent protein kinase (DNA-PK) (Feng et al, 2004; Sarbassov et al, 2005). Based the

structural studies on Akt/PKB, phosphorylation of this residue can stabilize the conformation of α helix C and allow the subsequent interaction between the conserved α-helix C glutamate and the ATP-binding lysine (Huang et al, 2003; Yang et al, 2002). In contrast to

Figure 4-1 Domain organization of mouse SGK1.

The kinase domain (residues 97-346) is shown in orange. Residues 1-96 (N) and 347-432 (C) are the terminal extensions and are shown in cyan and green, respectively. Major phosphorylation sites Thr256 and Ser422 are indicated by arrows.

the well characterized Thr256 and Ser422, the functional and physiological roles of other SGK1 phosphorylation sites are less clear.

SGK1 was also identified as a WNK1 associated protein from our lab using yeast two-hybrid screening with WNK1 (1-555) as the bait (Xu et al, 2005a; Xu et al, 2005b). Biochemical analysis demonstrated that WNK1 not only binds to, but also stimulates the kinase activity of, SGK1 through a WNK1-catalytic independent manner. In other words, the N-terminus of WNK1, without the presence of the catalytic domain, can activate SGK1. The activation of SGK1 by WNK1 was supported by RNAi experiments showing that the depletion of WNK1 significantly decreases the activity of SGK1 after IGF-1 stimulation. Furthermore, phosphorylation of Thr58 on WNK1 by Akt is essential for this activation. These data raise the question of how the WNK1 N-terminal segment activates SGK1. One possibility is that binding of WNK1 phospho-Thr58 to SGK1 might induce some conformational changes that either directly activate the enzyme or facilitate the phosphorylation of Thr256 or Ser422. To understand the underlying mechanism, we decided to determine the crystal structure of SGK1, and if possible, in complex with a WNK1 peptide.

The crystal structure of SGK1 was recently reported by the Schnackenberg's group (Zhao et al, 2007). The protein used to crystallization contained residues 60-431and carried multiple mutations, including S74A, S78A, R192A, S397A, S401A, and S422D to prevent heterogeneous phosphorylation during protein expression. The structure adopted an inactive conformation. Although the structure still maintained the typical kinase fold, the entire active site was rearranged. The activation loop was

extended into the N-terminal domain and formed an antiparallel β-sheet with a β-strand from part of the α-helix C. The rest of α helix C together with the C-terminal kinase extension were disordered in the structure. This unusual structure is likely due to the unnecessary mutations. Certain residues mutated residues might need to be for proper structural folding.

Materials and Methods

Construction

Mouse SGK1 constructs 1-432, Δ1-60, Δ1-60 (S422D), 97-432, and 97-432 (S422D) were made by PCR amplification using Flag-SGK1 or Flag-SGK1 (S422D) as templates (Xu et al, 2005a). PCR products were then cloned into the bacterial expression vector pHisParallel(Sheffield et al, 1999) or the insect cell expression vector pFastBacHT (Invitrogen).

Expression of 6xHis-SGK1 protein in *E. coli*

pHisParallel-SGK1 plasmids were transformed into *E. coli* strain BL21 (DE3) or Rosetta (DE3) (Novagen). Cells were grown in LB medium containing 50 μg/ml ampicillin (and 34 μg/ml chloramphenicol if using Rosetta) and incubated at 37°C until OD_{600} reached 0.5-0.6. A final concentration of 0.5 mM IPTG was added to induce protein expression. The culture was grown at 37°C for an additional 4 hrs or at 20°C

overnight. Cells were harvested by centrifugation and resuspended in Ni-buffer A containing 50 mM HEPES, pH 8.0, 0.3 M NaCl, 10% glycerol, and protease inhibitors. Cells were lysed by incubating with 1 mg/ml lysozyme on ice for 30 min, followed by sonication. Cell debris was removed by centrifugation at 20,000 rpm for 1 hour.

Expression of 6xHis-SGK1 protein in Sf9 insect cells

Expression was performed by using the Bac-to-Bac® system. Baculoviruses were generated as described (Invitrogen). Sf9 cells were grown at 27°C in Sf-900 II SFM medium (Invitrogen). Typically, at a density of 1.5 x 10^6 cells/ml, Sf9 cells were infected with recombinant baculoviruses expressing different pieces of SGK1 with a 50:1 ratio. Virus-infected Sf9 cells were harvested 48-60 hours after infection and collected by centrifugation at 4000g for 10 minutes. The cell pellet was then resuspended in lysis buffer containing 50 mM HEPES, pH 8.0, 0.3 M NaCl, 10% glycerol, 0.03% Brij-35, and protease inhibitors. Cells were lysed by a Dounce homogenizer and were clarified by centrifugation at 45000 rpm for 1 hour.

Purification of 6xHis-SGK1 protein

The clarified cell extract was loaded onto a pre-equilibrated Ni-NTA agarose column (Qiagen), followed by washing with 0 mM and 20 mM imidazole buffer. Protein was eluted in a single step with 250 mM imidazole. Eluted protein was dialyzed at 4ºC

overnight against buffer containing 50 mM Tris, pH8.0, and 50 mM NaCl. Further purification was carried out on a Mono Q HR 5/5 column (Pharmacia). Protein was eluted with a gradient of NaCl (50-1000 mM) in buffer. Fractions containing SGK proteins were combined and applied to a Superdex75 column (Pharmacia) for final purification. The purity of protein was checked at each step by SDS-PAGE.

Crystallization of SGK1

Crystallization screens were carried out using the hanging drop vapor diffusion method at 20ºC. Protein was concentrated to 5-8 mg/ml and mixed with an equal volume of screening solution. Several commercially available kits were used, including Crystal Screen and Crystal Screen 2 from Hampton Research and Wizard 1 and Wizard 2 from Emerald Biosystems. Crystallization screen on proteins in complex with AMP-PNP were also performed. The protein complexes were made by preincubating the protein with 5 mM AMP-PNP and 5 mM $MgCl_2$ on ice for 30 min before setting up the screens.

Results and Discussion

Summary of SGK1 protein expression and purification

SGK1 protein expression

To obtain the most stable and well-expressed protein suitable for crystallization, several different expression constructs were tested. Based on biochemical studies and sequence alignment, a series of constructs consisting of different segments of SGK1 were generated. Three different N-terminal starting sites, residues 1, 61 and 97, were selected. Residues 97 is the beginning of the kinase domain. The design of constructs starting at residue 61 is based on biochemical studies showing that the segment comprising the N-terminal 60 amino acids can be ubiquitinated, which leads to proteasome-dependent degradation of SGK1 (Brickley et al, 2002). . In fact, SGK1 lacking this segment was demonstrated to result in higher expression than the full-length protein in epithelial cells. To increase the stability of the protein, a Ser422 to aspartate mutation was made. Ser422 is located in the hydrophobic motif, and its mutation to aspartate mimics the phosphorylated state. Similar mutations were used in solving the crystal structure of Akt/PKB and revealed a more stable conformation relative to the structure lacking the hydrophobic motif.

In addition to protein fragments and mutations, numerous efforts were made to express SGK1. Protein was first expressed in bacteria. Two *E. coli* strains BL21 (DE3) and Rosetta (DE3) were utilized. Different bacterial expression conditions were also tried, including varying the inducing temperatures (37ºC and 20ºC), duration (4 hours and overnight), and the amount of IPTG (0.2, 0.5, 1 mM). However, in all of these conditions, protein was either not expressed or became insoluble. Similarly, expression of GST-tagged SGK1 in bacteria was pursued, but the expression level was also very low. Expression of mammalian recombinant proteins can often be improved by using Sf9

insect cells due to the codon usage are similar to that in mammals and better protein folding machinery. Therefore, different SGK1 constructs were expressed in Sf9 cells. To improve protein folding further, co-expression with Cdc37, a chaperone protein, was carried out. No improvement was observed. Among all the tests, only SGK1 Δ1-60 (S422D) gave good expression and remained stable during the following purification process. The results of protein expression screening are summarized in Table 4-1.

Purification of SGK1 Δ1-60 (S422D) from Sf9 cells

With a 6xHis-tag at the N-terminus SGK1 Δ1-60 (S422D) has a pI 6.7; thus, the protein was purified on Ni-NTA agarose (Qiagen), Mono Q anion exchange (Pharmacia), and Superdex75 size exclusion (Pharmacia). Figure 4-2 shows Coomassie blue staining of protein after the nickel column. The eluted SGK1 protein was still contaminated with numerous proteins endogenous to Sf9 cells. Mono Q HR 5/5 was used for subsequent purification. Figure 4-3 shows a chromatogram and the Coomassie blue staining of protein in fractions from the Mono Q column. A major peak containing SGK1 protein was eluted in the broad peak of around 400-600 mM NaCl. Based on the Coomassie blue staining, the protein was about 80-90 % pure. Typically, a 1-liter culture gave a yield of approximately 3.5 mg of protein at this step. Final purification was carried out on a Superdex75 gel filtration column. Figure 4-4 shows a chromatogram and Coomassie blue staining of protein in fractions from the Superdex75 column. Usually, more than two peaks were found in the chromatogram. The major peak of SGK1 was eluted at 60-75 ml, a smaller peak at 50-60 ml, and an even smaller and broader peak before 50 ml. Despite the fact that only the major peak matches the correct molecular weight; all peaks

gave a homogeneous 40 kDa band on SDS-PAGE, suggesting that SGK1 forms a monomer, a dimer, and higher order oligomers in solution. The exact order of these high molecular weight oligomers is not clear, but it appears to be a mixture due to the broad peak.

To further examine the homogeneity of the protein, the final purified SGK1 Δ1-60 (S422D) was sent for electrospray ionization mass spectrometry (ESI-MS) analysis. Multiple peaks were found, indicating that the protein is heterogeneously modified (Figure 4-5). The predicted molecular weight of the protein is 45512 Da. The two major peaks from mass spectrometry, 45512 Da and 45592 Da, were separated by 80 Da which corresponds to the molecular weight of a phosphate.

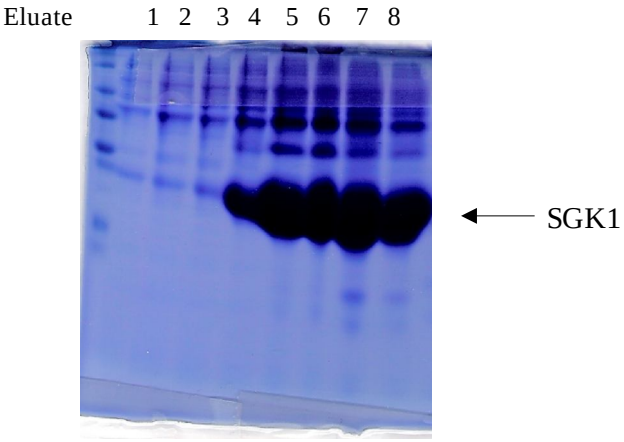
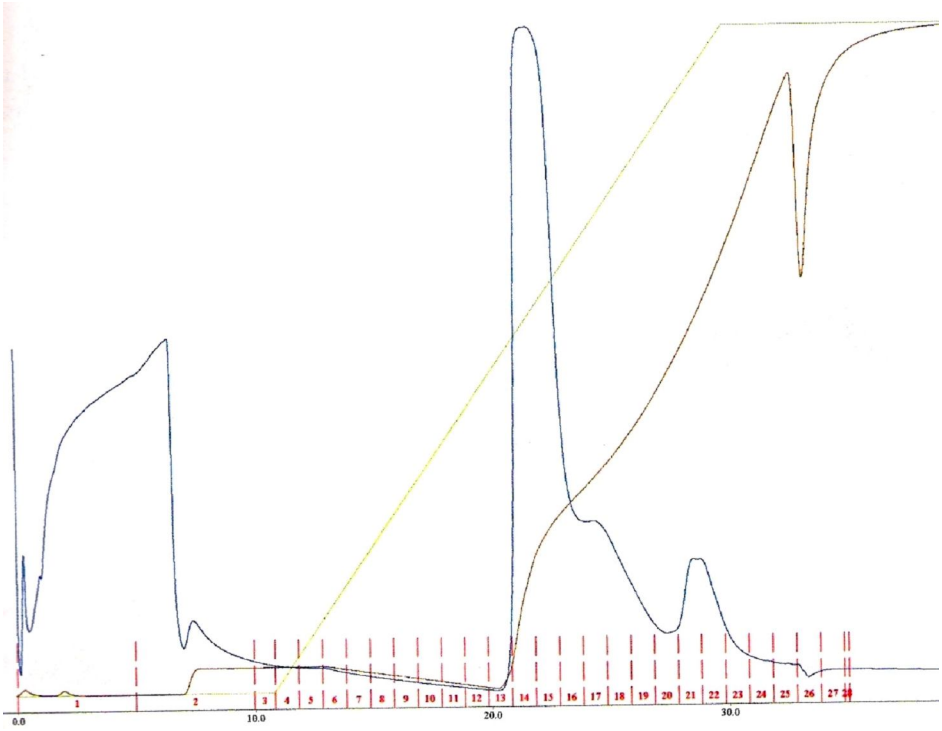


Figure 4-2 Purification of SGK1 by nickel affinity column

Commassie blue stained gel of protein from each fraction of the nickel column



Fraction No. 14 15 16 17

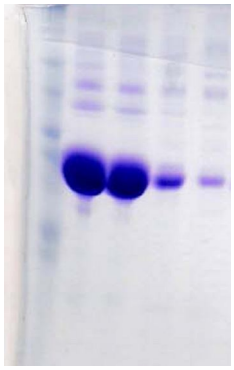
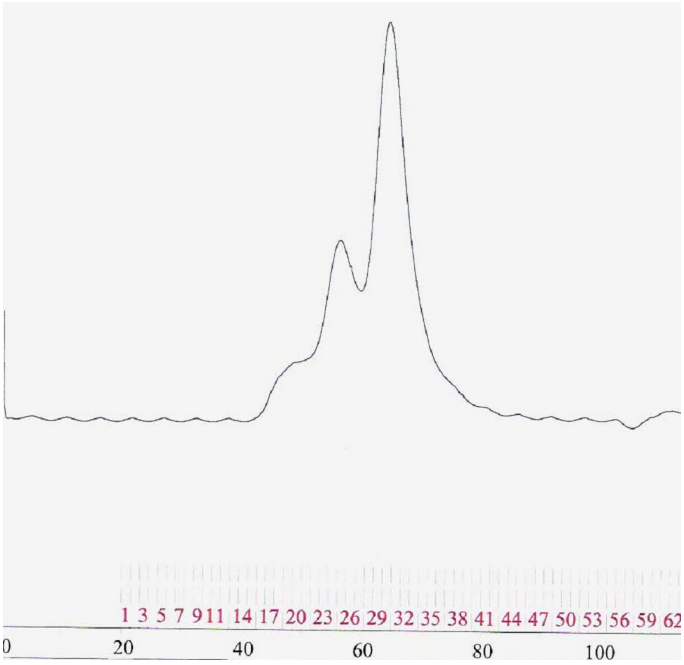


Figure 4-3 Purification of SGK by Mono Q ion exchange column

The top panel shows a chromatogram (OD_{260}) of SGK1 on the Mono Q column. The bottom panel shows a Coomassie blue stained gel of protein from some fractions of Mono Q.



Fraction No. 23 24 25 26 27 28 29 30 31 32 33 34 35

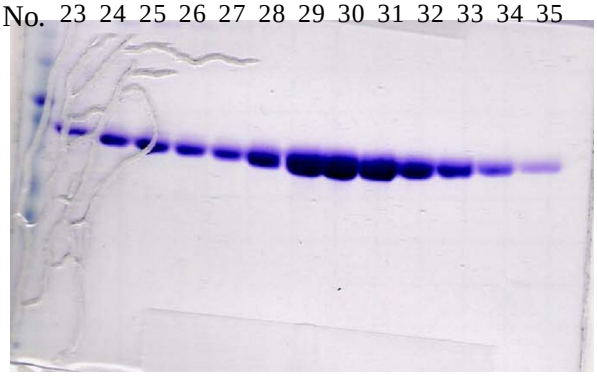


Figure 4-4 Purification of SGK by Superdex75 16/60 gel filtration column

The top panel shows a chromatogram (OD_{260}) of SGK1 on the Superdex75 column. The bottom panel shows Coomassie blue stained gel of protein from each fraction of Superdex75 column.

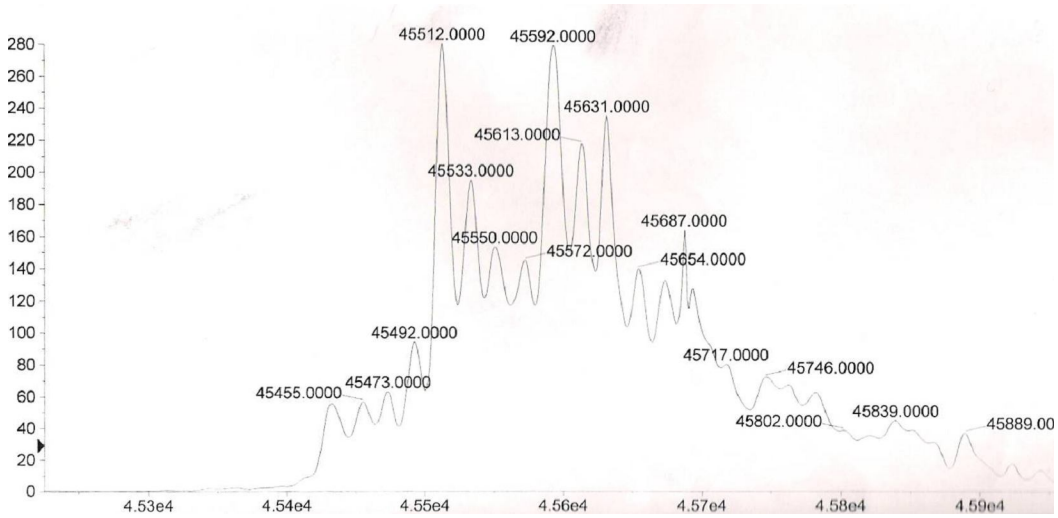


Figure 4-5 Electrospray ionization mass spectrometry (ESI-MS) analysis of SGK1 Δ1-60 (S422D)

Crystallization of SGK1 Δ1-60 (S422D)

Extensive crystallization screens were performed on both SGK1 Δ1-60 (S422D) protein alone and in complex with AMP-PNP. Only SGK1 Δ1-60 (S422D) in complex with AMP-PNP crystallized after two to three months in a single condition consisting of 12% PEG20,000 and 0.1 M MES, pH6.5 (identified from reagent No. 24 Crystal Screen II, Hampton research). The crystals are needle-shaped, with the longest dimension about 150 nm (Figure 4-6). However, these crystals could not be reproduced. The cause for this failure to reproduce SGK1 crystals was not clear. There are several possibilities. The major issue is that the protein was not homogeneous. Based on the mass spectrometry results, the protein purified from Sf9 cells is heterogeneously phosphorylated. Such phosphorylation events might occur in an unregulated manner. Thus, proteins from different purification batches might contain different mixtures of phosphorylated proteins. This variation might affect the reproducibility of crystallization. Moreover, as demonstrated by the gel filtration chromatography profile, SGK1 Δ1-60 (S422D) forms monomers as well as higher order oligomers in solution. Although only the protein from monomeric fractions was used for crystallization, the monomeric proteins still have the potential to form different types of oligomers. This might increase the heterogeneity and impede crystallization.

To obtain a homogenous SGK1 protein, the construct used here needs to be modified. As the protein was heterogeneously phosphorylated, mutation of the phosphorylation sites would be one approach. However, generating mutations should be carefully to prevent protein misfolding. This problem was evident in the inactive SGK1

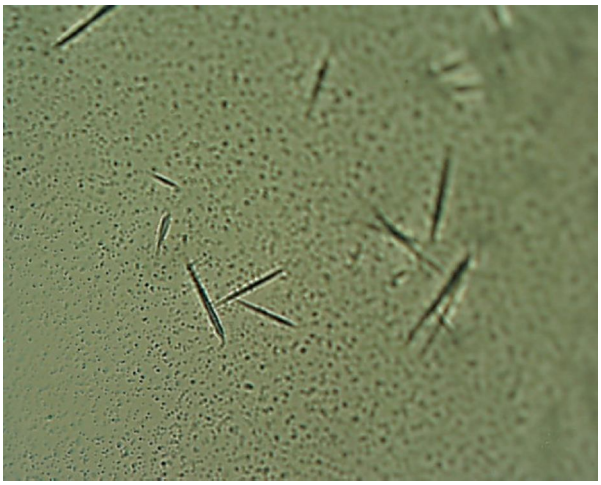


Figure 4-6 Crystals of SGK1 Δ1-60 (S422D)

The crystals were grown using the hanging drop vapor diffusion method at 20°C. The well solution contained 12% PEG20,000 and 0.1 M MES, pH6.5. The protein concentration was 8 mg/ml and in complex with 5 mM AMP-PNP. Crystals were about 150 nm long.

structure solved by Schnackenberg's group (Zhao et al, 2007). The key portions of the protein, perhaps one quarter of the included residues were disordered. Instead of using multiple mutations at the same time, a single mutation could be first tried and in combination with biochemical characterization to verify if the mutant protein still maintains its activity. The SGK1 structure that was solved is unphosphorylated at the Thr256 in the activation loop. Phosphorylation of this residue is required for kinase activation and is mediated by PDK1. An active protein might help for crystallization. Active SGK1 protein can be generated by either co-expressing SGK1 with PDK1 in Sf9 cells or by in vitro activation by purified PDK1 protein. Buffers used for protein purification might require further optimizations. Protein oligomerization is often affected by buffer ionic strength. Thus, changing the composition of solutions might make the protein more homogeneous.

Another concern is that the crystal we obtained here is from a protein degradation product. As the original crystals were formed months after the protein was placed in the hanging drop, sporadic protein degradation may have occurred during this long period. Mass spectrometry might be used to analyze the composition of these crystals. The result might provide insights into which part of protein is most stable and suitable for crystallization and thus help for future construct modification.

The process of crystal nucleation occurs spontaneously and randomly, and usually is affected by numerous factors, including temperature, humidity, and the presence of nucleation-promoting factors. Some proteins are extremely sensitive to these changes and can only be crystallized in a restricted set of conditions. Thus, repeating the

screening might be necessary. Different protein concentrations or the ratio of protein to well solution can also be tried. Further extending the crystallization conditions might help to identify a better crystallization condition. The conditions can be modified not only by changing the buffer pH and precipitant concentration, but also by using different types of buffer, precipitant, and adding salts and other additives.

The difficulty of crystallizing SGK1 might be also due to the intrinsic flexibility and instability of the protein. Instability was observed during protein purification. The protein was prone to precipitate when it was concentrated over 15 mg/ml or repeatedly frozen and thawed. Co-crystallizing with an interacting protein, such as WNK, might help to solve this problem. In addition, two other mammalian relatives of SGK1, SGK2 and SGK3, have been identified (Kobayashi et al, 1999). Their catalytic domains share more than 80% sequence identify with SGK1. Furthermore, the activities of both SGK2 and SGK3 were also stimulated by the N-terminus of WNK1, suggesting a high similarity with SGK1. Thus, these two isoforms might be tested as alternatives for crystallization trials.

CHAPTER FIVE

EXPRESSION, PURIFICATION AND CRYSTALLIZATION OF ERK2 (E320K)

Introduction

MAP kinases (MAPKs) play critical roles in a wide array of cellular processes including cell growth, cell differentiation, cell death, and metabolic homeostasis. Currently, there are three major subgroups of MAPKs: ERK, JNK, and p38 (Chang & Karin, 2001). Each subgroup is activated by specific stimuli and participates in different functions. For example, ERK is stimulated by growth factors and phorbol esters, and is associated with cell proliferation and differentiation. MAPKs are activated by a three-layered signaling cascade, in which an activated MAPK kinase kinase (MAP3K) phosphorylates and activates a MAPK kinase (MAP2K) that in turn phosphorylates and activates a MAPK (English et al, 1999). The activated MAPK further phosphorylates its substrates leading to various responses. Although there are multiple upstream activators and downstream targets, each MAPK is only recognized by specific MAP2Ks and substrates, suggesting that interactions between each component within a signaling pathway are tightly regulated.

Several mechanisms have been identified to assure this high enzyme-substrate specificity. For instance, MAPKs usually phosphorylate substrates on a Ser or a Thr residue followed by a Pro, termed the Ser/Thr-Pro motif (Lewis et al, 1998). However, such specificity is too simple to select a unique substrate. Indeed, numerous proteins

possess this minimal phosphorylation recognition motif, but only a subset of these are MAPK substrates. Thus, additional protein-protein interactions are required to maintain pathway fidelity. Among the known MAPK protein interaction mechanisms, interactions through the MAPK common docking (CD) domain are best studied (Tanoue et al, 2000; Tanoue et al, 2001). The CD domain is located on the backside of the kinase domain and is characterized by a cluster of negatively charged residues including Asp316, Asp319, and Glu320 in mouse ERK2. It specifically interacts with a consensus sequence $(R/K)_{2-3}$-X_{1-6}-Φ-X-Φ (Φ hydrophobic residue). In the crystal structure of ERK2 in complex with a CD domain docking peptide derived from the hematopoietic tyrosine phosphatase (HePTP), Asp319 forms direct ionic interactions with two Arg residues in the peptide, while Asp316 and Glu320 appear to stabilize the structure (Figure 5-1) (Zhou et al, 2006).

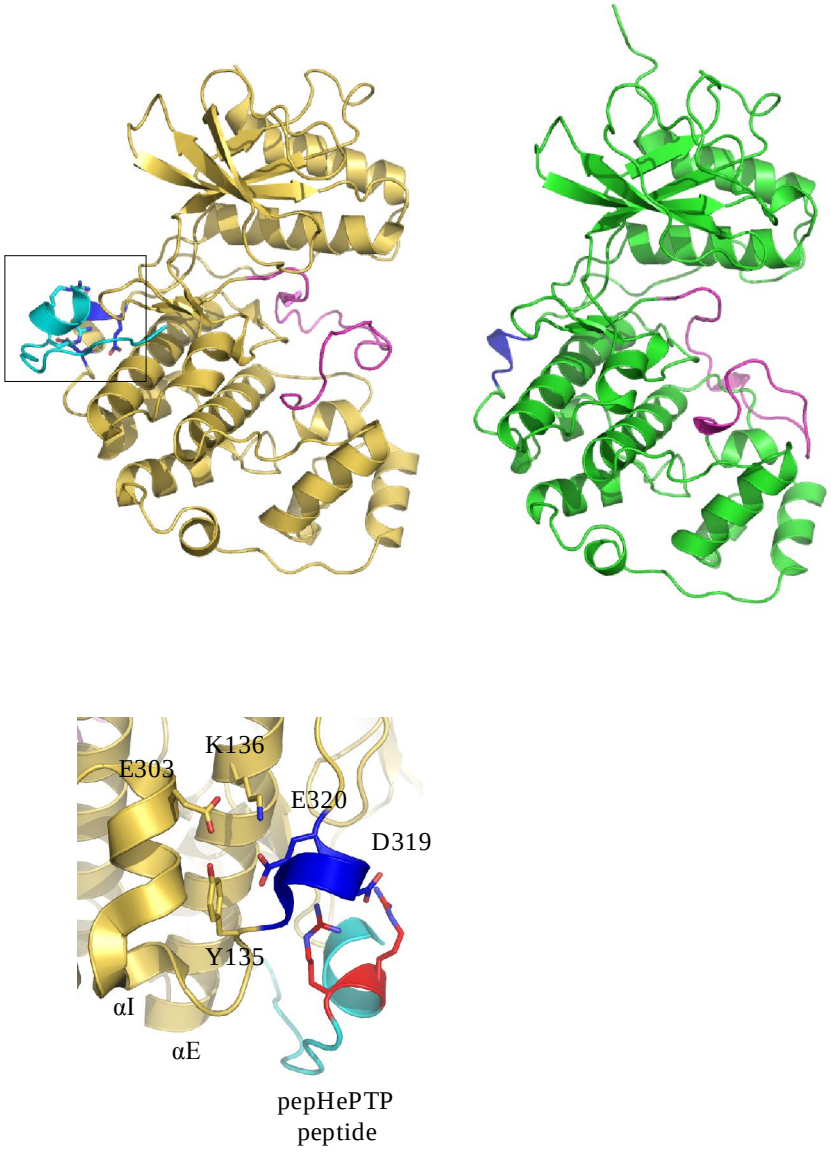


Figure 5-1 The MAPK CD domain

The top panel shows the structural comparison of ERK2 in complex with a CD docking peptide, pepHePTP (left; pdb code 2gph), and the unphosphorylated ERK2 (right; pdb code 1erk). The activation loop and CD domain in the kinase domain are colored in magenta and blue, respectively; and the peptide is in cyan. The bottom panel highlights the interface of the CD domain and the docking peptide.

More interestingly, the residue Glu322 in human ERK2, corresponding to Glu320 in mouse, is mutated to lysine in the oral squamous carcinoma cell line, HSC6 (Arvind et al, 2005). This mutation resulted in a faster migration of protein on SDS-PAGE. In addition, mutated protein expressed from both mammalian 293 cells and bacteria displayed higher kinase activity than the wild-type which was suggested to enhance tumorigenicity. The question becomes how mutation in the kinase CD domain increases the catalytic activity. Structural studies on MAPKs and docking peptides suggested that binding of peptide to the CD domain induces allosteric conformational changes particularly in the activation loop such that the activation loop exhibits a conformation more similar to the active kinase (Zhou et al, 2006). The CD domain mutation identified from the cancer cell line might cause a local disturbance in the CD domain that further increases the flexibility in the activation loop leading to a higher kinase activity. Thus, determining the crystal structure of the mutant mouse ERK2 (E320K) will help to understand the underlying mechanism of ERK2 regulation.

Materials and Methods

Protein Expression and Purification

ERK2 (E320K) was expressed in *E. coli* BL21 (DE3) cells as described previously (Zhang et al, 1994) with a 6xHis-tag at the N terminus. Protein was purified by nickel agarose (Qiagen), Mono Q HR 5/5 (Pharmacia), and Superdex75 16/60

(Pharmacia) columns, and stored in buffer consisting of 25 mM Tris, pH 7.5, 100 mM NaCl, 1 mM EDTA, and 1 mM DTT.

Crystallization

Crystallization screenings were carried out using the hanging drop vapor diffusion method at 20ºC. Protein alone or in complex with 5 mM AMP-PNP was first concentrated to approximately 10 mg/ml and then mixed with an equal volume of screening solution. Commercially available kits including Crystal Screen 1 and 2, SaltRx 1 and 2, Index 1 and 2 (Hampton Research), and Wizard 1 and Wizard 2 (Emerald Biosystems) were used for screening.

Results and Discussion

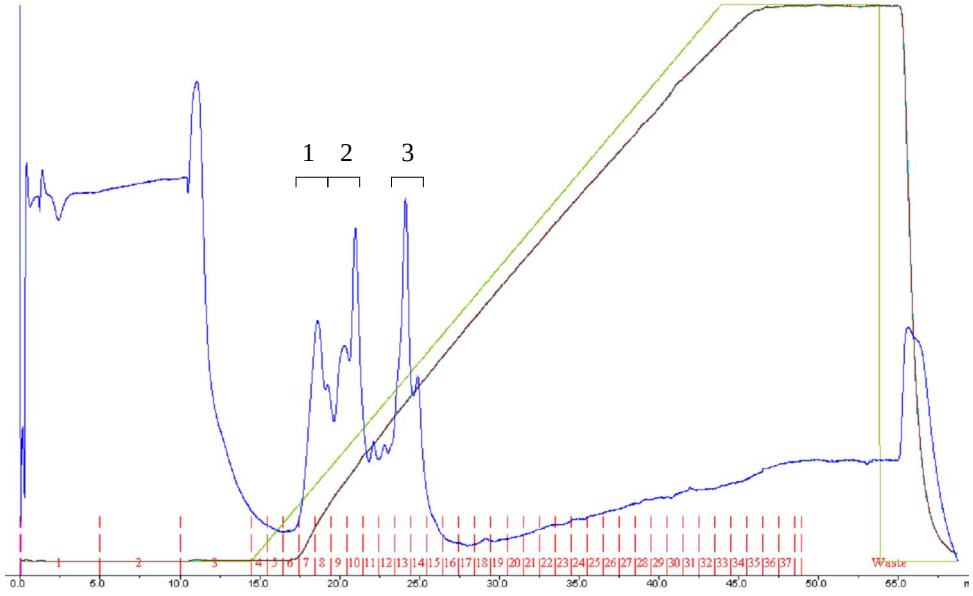
Expression and purification of ERK2 (E320K)

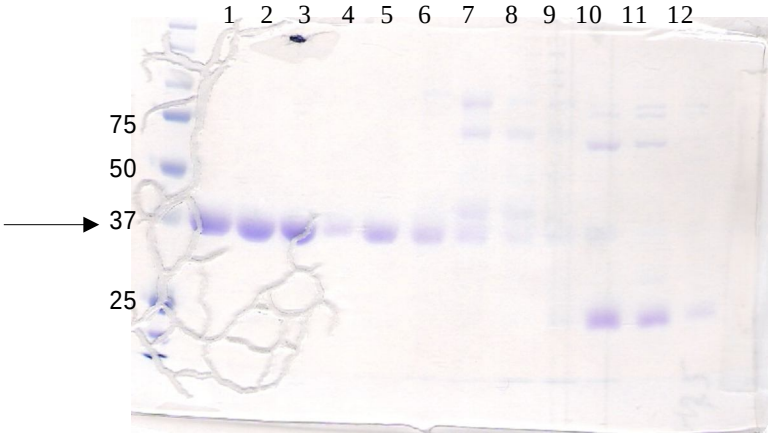
ERK2 (E320K) was well expressed in *E. coli* strain BL21. Figure 5-2 shows the Coomassie blue stained gel of protein after the nickel column. The protein was estimated to be more than 90 % pure at this stage. A Mono Q ion exchange column was used for the next step purification. ERK2 (E320K) eluted at three regions of ionic strength (Figure 5-3). The majority of ERK2 was in the flow-through at a NaCl concentration of 50 mM. Multiple peaks were observed in the gradient elution area, but three major peaks

dominate this profile. The first two peaks (peaks 1 and 2) were eluted at a NaCl concentration 200-300 mM and 300-400 mM NaCl, respectively, and both peaks contained ERK2. The third peak eluted at 450-600 mM NaCl was mostly due to contaminating proteins. Protein obtained from both flow-though and the first peak gave a single homogeneous band at 37 kDa on SDS-PAGE, while protein from the second peak resolved into multiple bands suggesting more heterogeneity (Figure 5-3). Proteins derived from different regions were stored separately, and the flow-through and peak 1 proteins were further purified by Superdex75 gel filtration column. Both proteins eluted mainly at approximately 70 ml that corresponds to a monomer. A comparably small portion of the protein eluted at around 60 ml that corresponds to a dimer (Figure 5-4).

Figure 5-2 Purification of ERK2 (E320K) by nickel affinity column

Coomassie blue stained gel of protein from each fraction of a nickel column





Lane
1. Fraction no. 1
2. Fraction no. 2
3. Fraction no. 3
4. Fraction no. 7
5. Fraction no. 8
6. Fraction no. 9
7. Fraction no. 10
8. Fraction no. 11
9. Fraction no. 12
10. Fraction no. 13
11. Fraction no. 14
12. Fraction no. 15

Figure 5-3 Purification of ERK2 (E320K) by Mono Q ion exchange column

The top panel shows a chromatogram of Mono Q column. The bottom panel shows a Coomassie blue stained gel of protein from each fraction of Mono Q.

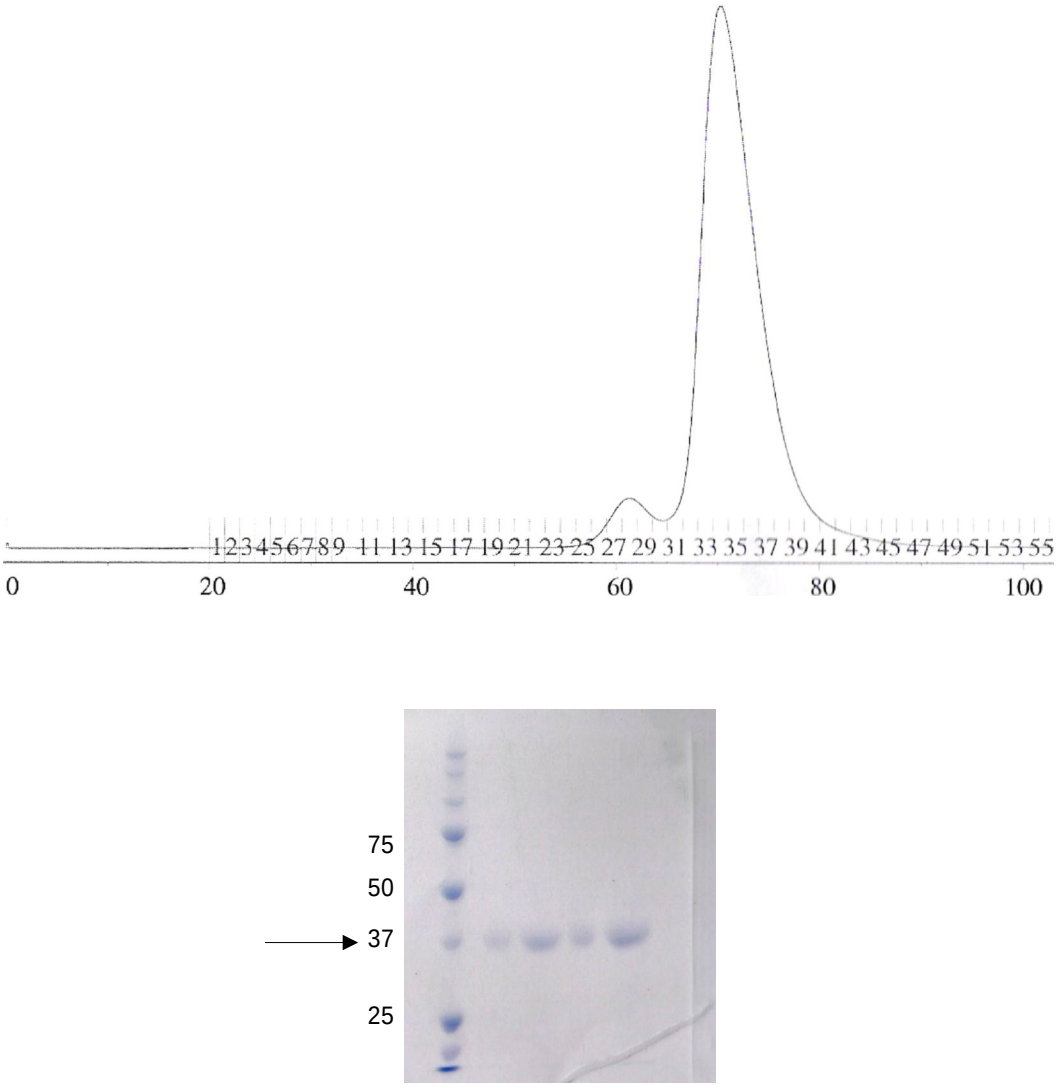


Figure 5-4 Purification of ERK2 (E320K) by Superdex75 16/60 gel filtration column

The top panel shows a chromatogram of Superdex75 column. The bottom panel shows Coomassie blue staining of protein after Superdex75 column.

Crystallization screening of ERK2 (E320K)

Numerous crystallization conditions have been tried on both ERK2 (E320K) protein alone and in complex with AMP-PNP, but there was no crystal. I also examined if this ERK2 (E320K) mutant could be crystallized under conditions that were identified for unphospho- and phospho- ERK2s, and still no crystals were found. The inability to generate ERK2 (E320K) crystals under those conditions implies that the conformation of ERK2 (E320K) mutant might be different from either that of the inactive or active states.

The failure to crystallize ERK2 (E320K) is probably due to the flexibility of the protein. As mentioned earlier, Glu320 is involved in the structural arrangement in the CD domain and the CD domain is associated with allosteric conformational changes in the kinase activation loop (Zhou et al, 2006). Mutation of Glu320 to lysine is expected to increase the flexibility not only in the CD domain but also elsewhere in the molecule. Therefore, methods used to stabilize the conformation of proteins should be helpful for crystallization. For example, ERK2 (E320K) could possibly be co-crystallized with an interacting peptide or protein such as a substrate or a scaffold protein. Furthermore, kinase inhibitors usually form extensive interactions with proteins which might stabilize protein conformation. Thus, ERK2 (E320K) in complex with inhibitors like staurosporine can be tried. Currently, all the crystal screens were performed at 20 °C. Screenings at 4 °C or 16 °C can be carried out to see if it can lower the conformational entropy in the protein.

CHAPTER SIX

CONCLUSIONS AND FUTURE DIRECTIONS

SPAK was first identified as a kinase that regulates ion cotransporters including KCC, NKCC1, NKCC2 and NCC (Piechotta et al, 2002). Recently, SPAK has gained increasing attention due to its function as a substrate of WNK kinases whose mutations cause a hereditary form of hypertension (Anselmo et al, 2006; Wilson et al, 2001). The link between WNK-SPAK-cotransporters provides one of the functional pathways for WNK to regulate intracellular and extracellular salt and water balance, thus contribute to the control of blood pressure. The research here is focused on determining the crystal structure of SPAK and providing a molecular basis for understanding the catalytic and regulatory mechanisms of this enzyme.

We first solved the crystal structure of SPAK 63-390 (T243D) at 2.5 Å resolution. The structure contains the kinase domain and part of the PF1 domain with a mutation of the WNK phosphorylation site, Thr243, to aspartate to mimic the phosphorylated state. The PF1 structure that was detected forms an extra helix and packs on the groove formed by α helices D and E and β stands 7 and 8. Two dimer interfaces were found in the structure. One is from the unique activation loop-swapped structure. The other is formed between the N-terminal domain of one molecule and the C-terminal domain of the second molecule. Comparison the structure with unphosphorylated and inactive OSR1 (Lee et al, 2008) revealed significant conformational changes in the

glycine-rich loop, helix αC, and the activation loop. A more striking feature is the reformation of the P+1 pocket in SPAK. The interaction is formed by residues from both molecules of the domain-swapped dimer supporting the conclusion that the domain-swapped dimer is a functional unit. As remodeling of the P+1 pocket is usually associated with kinase activation, the conformational change found between the crystallized forms of OSR1 and SPAK is thus believed to be a regulated event.

The activation loop is disordered in SPAK 63-390 (T243D). In order to gain more insights into the conformation of the activation loop, we also generated several activation loop mutants and solved one of the structures, SPAK 63-390 (V235C/T243D). With the additional mutation, the activation loop is fully modeled. Similar to other active state kinases, the phospho-mimic Thr243Asp forms an ion-pairing interaction with a basic residue within the activation loop and is believed to stabilize the activation loop conformation. Regardless of the differences in ability to dimerize in solution, the overall structure of SPAK 63-390 (V235C/T243D) is similar to that of SPAK 63-390 (T243D). However, one unexpected result is that the two molecules in the asymmetric unit show considerable structural differences, including in the glycine-rich loop, helix αC, and particularly in the activation loop. Comparative structural analysis revealed that one molecule adopts a more active conformation than the other.

Despite the availability of these partially active structures, it remains unclear how a fully active SPAK should look and why the PF1 domain is required for kinase activity. The present structure adopts a few features of active kinases but still several interactions important for catalytic activity are not formed, including the interaction of β strand 3

Lys104 with α helix C Glu121, and the catalytic loop Arg203 with phospho-mimic Thr243Asp. As demonstrated by the *in vitro* kinase assays, without these interactions residues in the active site are not properly arranged, thereby affecting the catalytic reaction. To understand the full activation mechanism, the crystal structure of the phosphorylated SPAK will be required. As co-expression of SPAK with WNK1 in bacteria was not successful in generating phosphorylated SPAK, *in vitro* phosphorylation of wild-type SPAK by purified WNK1 protein will be used. In the fully active structure, the PF1 might be more stable given that it is required for the kinase activity, and thus its structure is likely to help unravel the underlying mechanisms. If not, co-crystallization with a substrate (e.g. a peptide from NKCC) or a binding protein (e.g. WNK1 kinase domain) will be tried to see if either can stabilize the PF1 domain.

As mentioned above, two types of dimer were found in the structure. I will use mutagenesis that should disrupt the interface to determine if these crystallographically observed dimers are also presented in solution. Mutation of interface residues will be generated followed by gel filtration to analyze dimer formation or small angle X-ray scattering (SAXS) to analyze the surface shape of the protein. The interaction between the domain-swapped dimer is largely hydrophobic, except for one ion pair formed by Glu254 and Arg337. We will first mutate either Glu254 to arginine, or Arg337 to glutamate to introduce repulsion between the two molecules. If this is not successful, we will randomly mutate hydrophobic residues that are involved in the contact, including Trp250, Pro253, Glu254, Val255, Met256, Trp269, Ile273, Val293, Leu294, and Leu298, to alanines to disrupt all the contacts formed between side chains. Based on the structure,

residues Arg260/Gly261 are located at the switch point of domain swapping. These two residues will be exchanged for a sequence found in non-swapped kinases such as Ser/Leu in PAK6 to determine if Arg260/Gly261 are permissive for swapping. As for the interface between N- and C-terminal domains, charged residues, such as Asp73, Glu117, Arg168, Lys287, and Glu310, will be mutated to amino acids of the opposite charged to disrupt the dimer. If the dimer is still intact, mutation of the following hydrophobic residues to alanine will be performed: Val66, Ile70, Pro96, Leu119, Ser139, Phe140, Val141, Ile162, Val163, Asn167, His286, Pro289, and Pro290.

The next question is whether dimerization affects the function of SPAK. Dimerization might stabilize a conformation that is suitable for the catalytic reaction or protein-protein interactions. Additional studies demonstrated that mutations of several specific residues in the activation loop of SPAK converted the protein into a complete monomer in solution, suggesting a coupling between dimerization and the activation loop. As the activation loop plays critical roles in kinase activation, the kinase activity of dimerization-incompetent mutants will be tested first by *in vitro* kinase assays. It has been reported that dimerization might have different effects on the catalytic activity toward different substrates; therefore, both monomeric (e.g. MBP) and dimeric or multimeric proteins (e.g. NKCC, PAK) will be used as substrates to analyze the detailed kinetics of the enzymes. We will also examine if dimerization affects autophosphorylation using *in vitro* kinase assays as well. Second, *in vitro* binding assays will be carried out to investigate the connection between dimerization and protein-protein interaction. Dimerization-incompetent SPAK mutants will be analyzed in their abilities

to bind interacting proteins including WNK and NKCC. Next, to determine the effect of dimerization on SPAK function in cells, various functional assays will be performed using the dimerization-incompetent mutants. For example, subcellular localization can be inspected by immunofluorescence studies and the impact of the activities of cotransporters such as NKCC can be measured using transport assays as described (Anselmo et al, 2006).

Currently, the substrate specificity of SPAK is still not clear. More substrates of SPAK need to be identified and interacting regions mapped. Based on the crystal structure, the substrate binding site does not contain a distinguished charged surface cluster, implying less stringent selectivity at this site than for many other kinases. It is highly possible that other substrate docking sites exist on the surface of the kinase domain. For example, an elongated groove was found on the backside of the kinase domain between α helix C and β strand 4. To determine if this groove is involved in substrate binding, mutagenesis of the surrounded residues will be carried out followed by *in vitro* kinase assays and binding assays. Candidate mutation sites are as follows: residues with small side-chains such as Ser126, Ser129, T135, and S139 will be changed to bulky amino acids; and residues with charged side-chains such as Asp116, Glu117, and Lys120 will be changed to alanine.

In contrast to the active kinases that usually adopt a very similar structural conformation required for the catalytic reaction—transfer of the γ-phosphate of ATP to the hydroxyl group of the substrate, inactive kinases can adopt many distinct conformations. This structural variation allows the design of specific kinase inhibitor.

The structure of SPAK 63-390 (T243D) reveals a unique inactive conformation in the ATP-binding site, in which helix αC is rotated outward, thus shifting the conserved glutamate (Glu121) out of the active site and disrupting the interaction with the lysine (Lys104) that coordinates ATP. Furthermore, the N-terminal portion of the activation loop forms a helix that forces the invariant Asp222 in the DFG motif, situated at the base of the activation loop, to rotate away from the ATP site. Instead, Asp222 forms a rare contact with Lys104. This distinct conformation might be useful in developing SPAK inhibitors that might be further used to treat hypertension. With the availability of the crystal structure, *in silico* screening can be conducted. Screening of this sort is based on computational algorithms to identify small molecules that have low-energy binding to the active site of a macromolecule. Several software packages are available, including DOCK, FlexX, GLIDE, and OpenEye to perform this type of screen (Ewing et al, 2001; Hawkins et al, 2007; Venkatachalam et al, 2003).

REFERENCE

Flex X, version 1.13.5; Saint Augustin, Germany, BioSolveIT GmbH.

Adachi M, Fukuda M, Nishida E (2000) Nuclear export of MAP kinase (ERK) involves a MAP kinase kinase (MEK)-dependent active transport mechanism. *J Cell Biol* 148(5): 849-856

Adhikari A, Xu M, Chen ZJ (2007) Ubiquitin-mediated activation of TAK1 and IKK. *Oncogene* 26(22): 3214-3226

Alessi DR, Andjelkovic M, Caudwell B, Cron P, Morrice N, Cohen P, Hemmings BA (1996) Mechanism of activation of protein kinase B by insulin and IGF-1. *EMBO J* 15(23): 6541-6551

Anselmo AN, Earnest S, Chen W, Juang YC, Kim SC, Zhao Y, Cobb MH (2006) WNK1 and OSR1 regulate the Na+, K+, 2Cl- cotransporter in HeLa cells. *Proc Natl Acad Sci U S A* 103(29): 10883-10888

Arvind R, Shimamoto H, Momose F, Amagasa T, Omura K, Tsuchida N (2005) A mutation in the common docking domain of ERK2 in a human cancer cell line, which was associated with its constitutive phosphorylation. *Int J Oncol* 27(6): 1499-1504

Bellon S, Fitzgibbon MJ, Fox T, Hsiao HM, Wilson KP (1999) The structure of phosphorylated p38gamma is monomeric and reveals a conserved activation-loop conformation. *Structure* 7(9): 1057-1065

Bennett MJ, Choe S, Eisenberg D (1994) Domain swapping: entangling alliances between proteins. *Proc Natl Acad Sci U S A* 91(8): 3127-3131

Blume-Jensen P, Hunter T (2001) Oncogenic kinase signalling. *Nature* 411(6835): 355-365

Bossemeyer D, Engh RA, Kinzel V, Ponstingl H, Huber R (1993) Phosphotransferase and substrate binding mechanism of the cAMP-dependent protein kinase catalytic subunit from porcine heart as deduced from the 2.0 A structure of the complex with Mn2+ adenylyl imidodiphosphate and inhibitor peptide PKI(5-24). *EMBO J* 12(3): 849-859

Brickley DR, Mikosz CA, Hagan CR, Conzen SD (2002) Ubiquitin modification of serum and glucocorticoid-induced protein kinase-1 (SGK-1). *J Biol Chem* 277(45): 43064-43070

Brinkworth RI, Breinl RA, Kobe B (2003) Structural basis and prediction of substrate specificity in protein serine/threonine kinases. *Proc Natl Acad Sci U S A* 100(1): 74-79

Brown NR, Noble ME, Endicott JA, Johnson LN (1999) The structural basis for specificity of substrate and recruitment peptides for cyclin-dependent kinases. *Nat Cell Biol* 1(7): 438-443

Burnett G, Kennedy EP (1954) The enzymatic phosphorylation of proteins. *J Biol Chem* 211(2): 969-980

Canagarajah BJ, Khokhlatchev A, Cobb MH, Goldsmith EJ (1997) Activation mechanism of the MAP kinase ERK2 by dual phosphorylation. *Cell* 90(5): 859-869

Chang CI, Xu BE, Akella R, Cobb MH, Goldsmith EJ (2002) Crystal structures of MAP kinase p38 complexed to the docking sites on its nuclear substrate MEF2A and activator MKK3b. *Mol Cell* 9(6): 1241-1249

Chang L, Karin M (2001) Mammalian MAP kinase signalling cascades. *Nature* 410(6824): 37-40

Chaudhary A, King WG, Mattaliano MD, Frost JA, Diaz B, Morrison DK, Cobb MH, Marshall MS, Brugge JS (2000) Phosphatidylinositol 3-kinase regulates Raf1 through Pak phosphorylation of serine 338. *Curr Biol* 10(9): 551-554

Chen W, Chen Y, Xu BE, Juang YC, Stippec S, Zhao Y, Cobb MH (2008) Regulation of a third conserved phosphorylation site in serum and glucocorticoid-induced protein kinase 1. *J Biol Chem*

Chen W, Yazicioglu M, Cobb MH (2004) Characterization of OSR1, a member of the mammalian Ste20p/germinal center kinase subfamily. *J Biol Chem* 279(12): 11129-11136

Cohen P (2002) Protein kinases--the major drug targets of the twenty-first century? *Nat Rev Drug Discov* 1(4): 309-315

Dan I, Watanabe NM, Kusumi A (2001) The Ste20 group kinases as regulators of MAP kinase cascades. *Trends Cell Biol* 11(5): 220-230

Daniels RH, Bokoch GM (1999) p21-activated protein kinase: a crucial component of morphological signaling? *Trends Biochem Sci* 24(9): 350-355

Darman RB, Forbush B (2002) A regulatory locus of phosphorylation in the N terminus of the Na-K-Cl cotransporter, NKCC1. *J Biol Chem* 277(40): 37542-37550

De Bondt HL, Rosenblatt J, Jancarik J, Jones HD, Morgan DO, Kim SH (1993) Crystal structure of cyclin-dependent kinase 2. *Nature* 363(6430): 595-602

Deacon K, Mistry P, Chernoff J, Blank JL, Patel R (2003) p38 Mitogen-activated protein kinase mediates cell death and p21-activated kinase mediates cell survival during chemotherapeutic drug-induced mitotic arrest. *Mol Biol Cell* 14(5): 2071-2087

DeAizpurua HJ, Cram DS, Naselli G, Devereux L, Dorow DS (1997) Expression of mixed lineage kinase-1 in pancreatic beta-cell lines at different stages of maturation and during embryonic pancreas development. *J Biol Chem* 272(26): 16364-16373

Debonneville C, Flores SY, Kamynina E, Plant PJ, Tauxe C, Thomas MA, Munster C, Chraibi A, Pratt JH, Horisberger JD, Pearce D, Loffing J, Staub O (2001) Phosphorylation of Nedd4-2 by Sgk1 regulates epithelial Na(+) channel cell surface expression. *EMBO J* 20(24): 7052-7059

Delpire E, Gagnon KB (2006) SPAK and OSR1, key kinases involved in the regulation of chloride transport. *Acta Physiol (Oxf)* 187(1-2): 103-113

Delpire E, Gagnon KB (2008) SPAK and OSR1: STE20 kinases involved in the regulation of ion homoeostasis and volume control in mammalian cells. *Biochem J* 409(2): 321-331

Denton J, Nehrke K, Yin X, Morrison R, Strange K (2005) GCK-3, a newly identified Ste20 kinase, binds to and regulates the activity of a cell cycle-dependent ClC anion channel. *J Gen Physiol* 125(2): 113-125

Doublie S (1997) Preparation of selenomethionyl proteins for phase determination. *Methods Enzymol* 276: 523-530

Elion EA (2001) The Ste5p scaffold. *J Cell Sci* 114(Pt 22): 3967-3978

Emsley P, Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60(Pt 12 Pt 1): 2126-2132

English J, Pearson G, Wilsbacher J, Swantek J, Karandikar M, Xu S, Cobb MH (1999) New insights into the control of MAP kinase pathways. *Exp Cell Res* 253(1): 255-270

Ewing TJ, Makino S, Skillman AG, Kuntz ID (2001) DOCK 4.0: search strategies for automated molecular docking of flexible molecule databases. *J Comput Aided Mol Des* 15(5): 411-428

Farnum MA, Xu H, Agrafiotis DK (2003) Exploring the nonlinear geometry of protein homology. *Protein Sci* 12(8): 1604-1612

Feng J, Park J, Cron P, Hess D, Hemmings BA (2004) Identification of a PKB/Akt hydrophobic motif Ser-473 kinase as DNA-dependent protein kinase. *J Biol Chem* 279(39): 41189-41196

Fischer EH, Krebs EG (1955) Conversion of phosphorylase b to phosphorylase a in muscle extracts. *J Biol Chem* 216(1): 121-132

Gagnon KB, England R, Delpire E (2006a) Characterization of SPAK and OSR1, regulatory kinases of the Na-K-2Cl cotransporter. *Mol Cell Biol* 26(2): 689-698

Gagnon KB, England R, Delpire E (2006b) Volume sensitivity of cation-Cl-cotransporters is modulated by the interaction of two kinases: Ste20-related proline-alanine-rich kinase and WNK4. *Am J Physiol Cell Physiol* 290(1): C134-142

Gagnon KB, England R, Delpire E (2007) A single binding motif is required for SPAK activation of the Na-K-2Cl cotransporter. *Cell Physiol Biochem* 20(1-4): 131-142

Gibbs CS, Zoller MJ (1991) Identification of electrostatic interactions that determine the phosphorylation site specificity of the cAMP-dependent protein kinase. *Biochemistry* 30(22): 5329-5334

Goldsmith EJ, Akella R, Min X, Zhou T, Humphreys JM (2007) Substrate and docking interactions in serine/threonine protein kinases. *Chem Rev* 107(11): 5065-5081

Haas M, Forbush B, 3rd (2000) The Na-K-Cl cotransporter of secretory epithelia. *Annu Rev Physiol* 62: 515-534

Hanks SK, Quinn AM, Hunter T (1988) The protein kinase family: conserved features and deduced phylogeny of the catalytic domains. *Science* 241(4861): 42-52

Hawkins PC, Skillman AG, Nicholls A (2007) Comparison of shape-matching and docking as virtual screening tools. *J Med Chem* 50(1): 74-82

Holm L, Sander C (1995) Dali: a network tool for protein structure comparison. *Trends Biochem Sci* 20(11): 478-480

Huang X, Begley M, Morgenstern KA, Gu Y, Rose P, Zhao H, Zhu X (2003) Crystal structure of an inactive Akt2 kinase domain. *Structure* 11(1): 21-30

Hunter T (2000) Signaling--2000 and beyond. *Cell* 100(1): 113-127

Huse M, Kuriyan J (2002) The conformational plasticity of protein kinases. *Cell* 109(3): 275-282

Janowski R, Kozak M, Jankowska E, Grzonka Z, Grubb A, Abrahamson M, Jaskolski M (2001) Human cystatin C, an amyloidogenic protein, dimerizes through three-dimensional domain swapping. *Nat Struct Biol* 8(4): 316-320

Jeffrey PD, Russo AA, Polyak K, Gibbs E, Hurwitz J, Massague J, Pavletich NP (1995) Mechanism of CDK activation revealed by the structure of a cyclinA-CDK2 complex. *Nature* 376(6538): 313-320

Johnson LN, Noble ME, Owen DJ (1996) Active and inactive protein kinases: structural basis for regulation. *Cell* 85(2): 149-158

Johnston AM, Naselli G, Gonez LJ, Martin RM, Harrison LC, DeAizpurua HJ (2000) SPAK, a STE20/SPS1-related kinase that activates the p38 pathway. *Oncogene* 19(37): 4290-4297

Kannan N, Neuwald AF (2004) Evolutionary constraints associated with functional specificity of the CMGC protein kinases MAPK, CDK, GSK, SRPK, DYRK, and CK2alpha. *Protein Sci* 13(8): 2059-2077

Knaus KJ, Morillas M, Swietnicki W, Malone M, Surewicz WK, Yee VC (2001) Crystal structure of the human prion protein reveals a mechanism for oligomerization. *Nat Struct Biol* 8(9): 770-774

Knighton DR, Zheng JH, Ten Eyck LF, Ashford VA, Xuong NH, Taylor SS, Sowadski JM (1991) Crystal structure of the catalytic subunit of cyclic adenosine monophosphate-dependent protein kinase. *Science* 253(5018): 407-414

Kobayashi T, Cohen P (1999) Activation of serum- and glucocorticoid-regulated protein kinase by agonists that activate phosphatidylinositide 3-kinase is mediated by 3-phosphoinositide-dependent protein kinase-1 (PDK1) and PDK2. *Biochem J* 339 (Pt 2): 319-328

Kobayashi T, Deak M, Morrice N, Cohen P (1999) Characterization of the structure and regulation of two novel isoforms of serum- and glucocorticoid-induced protein kinase. *Biochem J* 344 Pt 1: 189-197

Krissinel E, Henrick K (2007) Inference of macromolecular assemblies from crystalline state. *J Mol Biol* 372(3): 774-797

Lang F, Henke G, Embark HM, Waldegger S, Palmada M, Bohmer C, Vallon V (2003) Regulation of channels by the serum and glucocorticoid-inducible kinase - implications for transport, excitability and cell proliferation. *Cell Physiol Biochem* 13(1): 41-50

Leberer E, Dignard D, Harcus D, Thomas DY, Whiteway M (1992) The protein kinase homologue Ste20p is required to link the yeast pheromone response G-protein beta gamma subunits to downstream signalling components. *EMBO J* 11(13): 4815-4824

Lee SJ, Cobb MH, Goldsmith EJ (2008) Crystal structure of domain-swapped STE20 OSR1 kinase domain. *Protein Sci* 18(2): 304-313

Lei M, Lu W, Meng W, Parrini MC, Eck MJ, Mayer BJ, Harrison SC (2000) Structure of PAK1 in an autoinhibited conformation reveals a multistage activation switch. *Cell* 102(3): 387-397

Leiserson WM, Harkins EW, Keshishian H (2000) Fray, a Drosophila serine/threonine kinase homologous to mammalian PASK, is required for axonal ensheathment. *Neuron* 28(3): 793-806

Lenertz LY, Lee BH, Min X, Xu BE, Wedin K, Earnest S, Goldsmith EJ, Cobb MH (2005) Properties of WNK1 and implications for other family members. *J Biol Chem* 280(29): 26653-26658

Lewis TS, Shapiro PS, Ahn NG (1998) Signal transduction through MAP kinase cascades. *Adv Cancer Res* 74: 49-139

Li Y, Hu J, Vita R, Sun B, Tabata H, Altman A (2004) SPAK kinase is a substrate and target of PKCtheta in T-cell receptor-induced AP-1 activation pathway. *EMBO J* 23(5): 1112-1122

Liu Y, Belkina NV, Graham C, Shaw S (2006) Independence of protein kinase C-delta activity from activation loop phosphorylation: structural basis and altered functions in cells. *J Biol Chem* 281(17): 12102-12111

Liu Y, Eisenberg D (2002) 3D domain swapping: as domains continue to swap. *Protein Sci* 11(6): 1285-1299

Liu Y, Hart PJ, Schlunegger MP, Eisenberg D (1998) The crystal structure of a 3D domain-swapped dimer of RNase A at a 2.1-A resolution. *Proc Natl Acad Sci U S A* 95(7): 3437-3442

Lowe ED, Noble ME, Skamnaki VT, Oikonomakos NG, Owen DJ, Johnson LN (1997) The crystal structure of a phosphorylase kinase peptide substrate complex: kinase substrate recognition. *EMBO J* 16(22): 6646-6658

Lytle C, Xu JC, Biemesderfer D, Forbush B, 3rd (1995) Distribution and diversity of Na-K-Cl cotransport proteins: a study with monoclonal antibodies. *Am J Physiol* 269(6 Pt 1): C1496-1505

Madhani HD, Styles CA, Fink GR (1997) MAP kinases with distinct inhibitory functions impart signaling specificity during yeast differentiation. *Cell* 91(5): 673-684

Madhusudan, Trafny EA, Xuong NH, Adams JA, Ten Eyck LF, Taylor SS, Sowadski JM (1994) cAMP-dependent protein kinase: crystallographic insights into substrate recognition and phosphotransfer. *Protein Sci* 3(2): 176-187

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

Manser E, Leung T, Salihuddin H, Zhao ZS, Lim L (1994) A brain serine/threonine protein kinase activated by Cdc42 and Rac1. *Nature* 367(6458): 40-46

Mccoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ (2007) Phaser crystallographic software. *Journal of Applied Crystallography* 40: 658-674

McCoy AJ, Grosse-Kunstleve RW, Storoni LC, Read RJ (2005) Likelihood-enhanced fast translation functions. *Acta Crystallogr D Biol Crystallogr* 61(Pt 4): 458-464

Mercado A, Mount DB, Gamba G (2004) Electroneutral cation-chloride cotransporters in the central nervous system. *Neurochem Res* 29(1): 17-25

Miao N, Fung B, Sanchez R, Lydon J, Barker D, Pang K (2000) Isolation and expression of PASK, a serine/threonine kinase, during rat embryonic development, with special emphasis on the pancreas. *J Histochem Cytochem* 48(10): 1391-1400

Min X, Lee BH, Cobb MH, Goldsmith EJ (2004) Crystal structure of the kinase domain of WNK1, a kinase that causes a hereditary form of hypertension. *Structure* 12(7): 1303-1311

Mohammadi M, Honegger A, Sorokin A, Ullrich A, Schlessinger J, Hurwitz DR (1993) Aggregation-induced activation of the epidermal growth factor receptor protein tyrosine kinase. *Biochemistry* 32(34): 8742-8748

Mukherjee S, Keitany G, Li Y, Wang Y, Ball HL, Goldsmith EJ, Orth K (2006) Yersinia YopJ acetylates and inhibits kinase activation by blocking phosphorylation. *Science* 312(5777): 1211-1214

Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* 53(Pt 3): 240-255

Nagar B, Bornmann WG, Pellicena P, Schindler T, Veach DR, Miller WT, Clarkson B, Kuriyan J (2002) Crystal structures of the kinase domain of c-Abl in complex with the

small molecule inhibitors PD173955 and imatinib (STI-571). *Cancer Res* 62(15): 4236-4243

Nagar B, Hantschel O, Young MA, Scheffzek K, Veach D, Bornmann W, Clarkson B, Superti-Furga G, Kuriyan J (2003) Structural basis for the autoinhibition of c-Abl tyrosine kinase. *Cell* 112(6): 859-871

Naray-Fejes-Toth A, Canessa C, Cleaveland ES, Aldrich G, Fejes-Toth G (1999) sgk is an aldosterone-induced kinase in the renal collecting duct. Effects on epithelial na+ channels. *J Biol Chem* 274(24): 16973-16978

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

Novy R, Drott D, Yaeger K, Mierendorf R (2001) Overcoming the codon bias of E. coli for enhanced protein expression. *inNovations* 12: 1-3

Nurse P (1990) Universal control mechanism regulating onset of M-phase. *Nature* 344(6266): 503-508

Oliver AW, Paul A, Boxall KJ, Barrie SE, Aherne GW, Garrett MD, Mittnacht S, Pearl LH (2006) Trans-activation of the DNA-damage signalling protein kinase Chk2 by T-loop exchange. *EMBO J* 25(13): 3179-3190

Pandey A, Dan I, Kristiansen TZ, Watanabe NM, Voldby J, Kajikawa E, Khosravi-Far R, Blagoev B, Mann M (2002) Cloning and characterization of PAK5, a novel member of mammalian p21-activated kinase-II subfamily that is predominantly expressed in brain. *Oncogene* 21(24): 3939-3948

Park J, Leong ML, Buse P, Maiyar AC, Firestone GL, Hemmings BA (1999) Serum and glucocorticoid-inducible kinase (SGK) is a target of the PI 3-kinase-stimulated signaling pathway. *EMBO J* 18(11): 3024-3033

Pawson T, Scott JD (2005) Protein phosphorylation in signaling--50 years and counting. *Trends Biochem Sci* 30(6): 286-290

Piechotta K, Garbarini N, England R, Delpire E (2003) Characterization of the interaction of the stress kinase SPAK with the Na+-K+-2Cl- cotransporter in the nervous system: evidence for a scaffolding role of the kinase. *J Biol Chem* 278(52): 52848-52856

Piechotta K, Lu J, Delpire E (2002) Cation chloride cotransporters interact with the stress-related kinases Ste20-related proline-alanine-rich kinase (SPAK) and oxidative stress response 1 (OSR1). *J Biol Chem* 277(52): 50812-50819

Pike AC, Rellos P, Niesen FH, Turnbull A, Oliver AW, Parker SA, Turk BE, Pearl LH, Knapp S (2008) Activation segment dimerization: a mechanism for kinase autophosphorylation of non-consensus sites. *EMBO J* 27(4): 704-714

Polek TC, Talpaz M, Spivak-Kroizman TR (2006) TRAIL-induced cleavage and inactivation of SPAK sensitizes cells to apoptosis. *Biochem Biophys Res Commun* 349(3): 1016-1024

Ponce-Coria J, San-Cristobal P, Kahle KT, Vazquez N, Pacheco-Alvarez D, de Los Heros P, Juarez P, Munoz E, Michel G, Bobadilla NA, Gimenez I, Lifton RP, Hebert SC, Gamba G (2008) Regulation of NKCC2 by a chloride-sensing mechanism involving the WNK3 and SPAK kinases. *Proc Natl Acad Sci U S A* 105(24): 8458-8463

Quatela SE, Philips MR (2006) Ras signaling on the Golgi. *Curr Opin Cell Biol* 18(2): 162-167

Raitt DC, Posas F, Saito H (2000) Yeast Cdc42 GTPase and Ste20 PAK-like kinase regulate Sho1-dependent activation of the Hog1 MAPK pathway. *EMBO J* 19(17): 4623-4631

Ramos JW (2008) The regulation of extracellular signal-regulated kinase (ERK) in mammalian cells. *Int J Biochem Cell Biol* 40(12): 2707-2719

Ramoz N, Cai G, Reichert JG, Silverman JM, Buxbaum JD (2008) An analysis of candidate autism loci on chromosome 2q24-q33: evidence for association to the STK39 gene. *Am J Med Genet B Neuropsychiatr Genet* 147B(7): 1152-1158

Rossier BC (2003) Negative regulators of sodium transport in the kidney: key factors in understanding salt-sensitive hypertension? *J Clin Invest* 111(7): 947-950

Rousseau F, Schymkowitz JW, Itzhaki LS (2003) The unfolding story of three-dimensional domain swapping. *Structure* 11(3): 243-251

Rousseau F, Schymkowitz JW, Wilkinson HR, Itzhaki LS (2001) Three-dimensional domain swapping in p13suc1 occurs in the unfolded state and is controlled by conserved proline residues. *Proc Natl Acad Sci U S A* 98(10): 5596-5601

Russo AA, Jeffrey PD, Pavletich NP (1996) Structural basis of cyclin-dependent kinase activation by phosphorylation. *Nat Struct Biol* 3(8): 696-700

Sarbassov DD, Guertin DA, Ali SM, Sabatini DM (2005) Phosphorylation and regulation of Akt/PKB by the rictor-mTOR complex. *Science* 307(5712): 1098-1101

Schindler T, Bornmann W, Pellicena P, Miller WT, Clarkson B, Kuriyan J (2000) Structural mechanism for STI-571 inhibition of abelson tyrosine kinase. *Science* 289(5486): 1938-1942

Schwarzenbacher R, Godzik A, Grzechnik SK, Jaroszewski L (2004) The importance of alignment accuracy for molecular replacement. *Acta Crystallogr D Biol Crystallogr* 60(Pt 7): 1229-1236

Serag AA, Altenbach C, Gingery M, Hubbell WL, Yeates TO (2002) Arrangement of subunits and ordering of beta-strands in an amyloid sheet. *Nat Struct Biol* 9(10): 734-739

Sheffield P, Garrard S, Derewenda Z (1999) Overcoming expression and purification problems of RhoGDI using a family of "parallel" expression vectors. *Protein Expr Purif* 15(1): 34-39

Simon DB, Karet FE, Hamdan JM, DiPietro A, Sanjad SA, Lifton RP (1996) Bartter's syndrome, hypokalaemic alkalosis with hypercalciuria, is caused by mutations in the Na-K-2Cl cotransporter NKCC2. *Nat Genet* 13(2): 183-188

Skamnaki VT, Owen DJ, Noble ME, Lowe ED, Lowe G, Oikonomakos NG, Johnson LN (1999) Catalytic mechanism of phosphorylase kinase probed by mutational studies. *Biochemistry* 38(44): 14718-14730

Smith L, Smallwood N, Altman A, Liedtke CM (2008) PKCdelta acts upstream of SPAK in the activation of NKCC1 by hyperosmotic stress in human airway epithelial cells. *J Biol Chem* 283(32): 22147-22156

Steussy CN, Popov KM, Bowker-Kinley MM, Sloan RB, Jr., Harris RA, Hamilton JA (2001) Structure of pyruvate dehydrogenase kinase. Novel folding pattern for a serine protein kinase. *J Biol Chem* 276(40): 37443-37450

Strong M, Sawaya MR, Wang S, Phillips M, Cascio D, Eisenberg D (2006) Toward the structural genomics of complexes: crystal structure of a PE/PPE protein complex from Mycobacterium tuberculosis. *Proc Natl Acad Sci U S A* 103(21): 8060-8065

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

Tanoue T, Maeda R, Adachi M, Nishida E (2001) Identification of a docking groove on ERK and p38 MAP kinases that regulates the specificity of docking interactions. *EMBO J* 20(3): 466-479

Taylor SS, Zheng J, Radzio-Andzelm E, Knighton DR, Ten Eyck LF, Sowadski JM, Herberg FW, Yonemoto WM (1993) cAMP-dependent protein kinase defines a family of enzymes. *Philos Trans R Soc Lond B Biol Sci* 340(1293): 315-324

Tickle IJ, Driessen HP (1996) Molecular replacement using known structural information. *Methods Mol Biol* 56: 173-203

Tsutsumi T, Ushiro H, Kosaka T, Kayahara T, Nakano K (2000) Proline- and alanine-rich Ste20-related kinase associates with F-actin and translocates from the cytosol to cytoskeleton upon cellular stresses. *J Biol Chem* 275(13): 9157-9162

Ushiro H, Tsutsumi T, Suzuki K, Kayahara T, Nakano K (1998) Molecular cloning and characterization of a novel Ste20-related protein kinase enriched in neurons and transporting epithelia. *Arch Biochem Biophys* 355(2): 233-240

Venkatachalam CM, Jiang X, Oldfield T, Waldman M (2003) LigandFit: a novel method for the shape-directed rapid docking of ligands to protein active sites. *J Mol Graph Model* 21(4): 289-307

Verissimo F, Jordan P (2001) WNK kinases, a novel protein kinase subfamily in multi-cellular organisms. *Oncogene* 20(39): 5562-5569

Vitari AC, Deak M, Morrice NA, Alessi DR (2005) The WNK1 and WNK4 protein kinases that are mutated in Gordon's hypertension syndrome phosphorylate and activate SPAK and OSR1 protein kinases. *Biochem J* 391(Pt 1): 17-24

Vitari AC, Thastrup J, Rafiqi FH, Deak M, Morrice NA, Karlsson HK, Alessi DR (2006) Functional interactions of the SPAK/OSR1 kinases with their upstream activator WNK1 and downstream substrate NKCC1. *Biochem J* 397(1): 223-231

Walker EH, Perisic O, Ried C, Stephens L, Williams RL (1999) Structural insights into phosphoinositide 3-kinase catalysis and signalling. *Nature* 402(6759): 313-320

Wang Z, Harkins PC, Ulevitch RJ, Han J, Cobb MH, Goldsmith EJ (1997) The structure of mitogen-activated protein kinase p38 at 2.1-A resolution. *Proc Natl Acad Sci U S A* 94(6): 2327-2332

Webster MK, Goya L, Firestone GL (1993a) Immediate-early transcriptional regulation and rapid mRNA turnover of a putative serine/threonine protein kinase. *J Biol Chem* 268(16): 11482-11485

Webster MK, Goya L, Ge Y, Maiyar AC, Firestone GL (1993b) Characterization of sgk, a novel member of the serine/threonine protein kinase gene family which is transcriptionally induced by glucocorticoids and serum. *Mol Cell Biol* 13(4): 2031-2040

Whitehurst AW, Robinson FL, Moore MS, Cobb MH (2004) The death effector domain protein PEA-15 prevents nuclear entry of ERK2 by inhibiting required interactions. *J Biol Chem* 279(13): 12840-12847

Wilson FH, Disse-Nicodeme S, Choate KA, Ishikawa K, Nelson-Williams C, Desitter I, Gunel M, Milford DV, Lipkin GW, Achard JM, Feely MP, Dussol B, Berland Y, Unwin RJ, Mayan H, Simon DB, Farfel Z, Jeunemaitre X, Lifton RP (2001) Human hypertension caused by mutations in WNK kinases. *Science* 293(5532): 1107-1112

Winn MD, Isupov MN, Murshudov GN (2001) Use of TLS parameters to model anisotropic displacements in macromolecular refinement. *Acta Crystallogr D Biol Crystallogr* 57(Pt 1): 122-133

Xu B, English JM, Wilsbacher JL, Stippec S, Goldsmith EJ, Cobb MH (2000) WNK1, a novel mammalian serine/threonine protein kinase lacking the catalytic lysine in subdomain II. *J Biol Chem* 275(22): 16795-16801

Xu BE, Stippec S, Chu PY, Lazrak A, Li XJ, Lee BH, English JM, Ortega B, Huang CL, Cobb MH (2005a) WNK1 activates SGK1 to regulate the epithelial sodium channel. *Proc Natl Acad Sci U S A* 102(29): 10315-10320

Xu BE, Stippec S, Lazrak A, Huang CL, Cobb MH (2005b) WNK1 activates SGK1 by a phosphatidylinositol 3-kinase-dependent and non-catalytic mechanism. *J Biol Chem* 280(40): 34218-34223

Xu W, Harrison SC, Eck MJ (1997) Three-dimensional structure of the tyrosine kinase c-Src. *Nature* 385(6617): 595-602

Yadav SS, Miller WT (2008) The evolutionarily conserved arrangement of domains in SRC family kinases is important for substrate recognition. *Biochemistry* 47(41): 10871-10880

Yamaguchi H, Matsushita M, Nairn AC, Kuriyan J (2001) Crystal structure of the atypical protein kinase domain of a TRP channel with phosphotransferase activity. *Mol Cell* 7(5): 1047-1057

Yan Y, Nguyen H, Dalmasso G, Sitaraman SV, Merlin D (2007) Cloning and characterization of a new intestinal inflammation-associated colonic epithelial Ste20-related protein kinase isoform. *Biochim Biophys Acta* 1769(2): 106-116

Yang J, Cron P, Good VM, Thompson V, Hemmings BA, Barford D (2002) Crystal structure of an activated Akt/protein kinase B ternary complex with GSK3-peptide and AMP-PNP. *Nat Struct Biol* 9(12): 940-944

Yoo D, Kim BY, Campo C, Nance L, King A, Maouyo D, Welling PA (2003) Cell surface expression of the ROMK (Kir 1.1) channel is regulated by the aldosterone-induced kinase, SGK-1, and protein kinase A. *J Biol Chem* 278(25): 23066-23075

Z. Otwinowski and W. Minor PoX-rDDCiOM, Methods in Enzymology, Volume 276: Macromolecular Crystallography, part A, p.307‑326, 1997, C. W. Carter, Jr. & R. M. Sweet, Eds., Academic Press (New York). (1997) Processing of X-ray Diffraction Data Collected in Oscillation Mode. *Methods in Enzymology* 276

Zhang F, Strand A, Robbins D, Cobb MH, Goldsmith EJ (1994) Atomic structure of the MAP kinase ERK2 at 2.3 A resolution. *Nature* 367(6465): 704-711

Zhang X, Gureasko J, Shen K, Cole PA, Kuriyan J (2006) An allosteric mechanism for activation of the kinase domain of epidermal growth factor receptor. *Cell* 125(6): 1137-1149

Zhao B, Lehr R, Smallwood AM, Ho TF, Maley K, Randall T, Head MS, Koretke KK, Schnackenberg CG (2007) Crystal structure of the kinase domain of serum and glucocorticoid-regulated kinase 1 in complex with AMP PNP. *Protein Sci* 16(12): 2761-2769

Zhou T, Sun L, Humphreys J, Goldsmith EJ (2006) Docking interactions induce exposure of activation loop in the MAP kinase ERK2. *Structure* 14(6): 1011-1019

Zhu G, Fujii K, Liu Y, Codrea V, Herrero J, Shaw S (2005) A single pair of acidic residues in the kinase major groove mediates strong substrate preference for P-2 or P-5 arginine in the AGC, CAMK, and STE kinase families. *J Biol Chem* 280(43): 36372-36379