

# UC San Diego

## UC San Diego Electronic Theses and Dissertations

### Title

SR protein kinase 1 : conformation, substrate recognition and catalysis

### Permalink

<https://escholarship.org/uc/item/5kz9s49j>

### Author

Ngo, Jacky Chi Ki

### Publication Date

2006

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA, SAN DIEGO

**SR Protein Kinase 1: Conformation, Substrate Recognition and Catalysis**

A dissertation submitted in partial satisfaction of the  
requirements for the degree Doctor of Philosophy

in

Chemistry

by

Jacky Chi Ki Ngo

Committee in charge:

Professor Gourisankar Ghosh, Chair  
Professor Joseph A. Adams  
Professor Xiang-Dong Fu  
Professor James Andrew McCammon  
Professor Susan Taylor

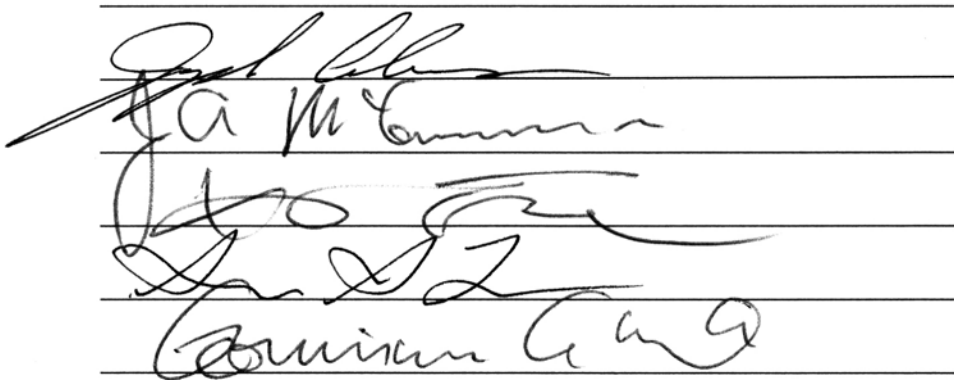
2006

Copyright

Jacky Chi Ki Ngo, 2006

All Rights Reserved.

The dissertation of Jacky Chi Ki Ngo is approved, and it is acceptable in quality and form for publication on microfilm:



The image shows five horizontal lines with handwritten signatures written across them. From top to bottom, the signatures are: 1. A signature that appears to be 'Paul H. ...'. 2. A signature that appears to be 'Ja M ...'. 3. A signature that appears to be 'Jo ...'. 4. A signature that appears to be 'Jan ...'. 5. A signature that appears to be 'Gavin ...'.

Chair

University of California San Diego

2006

## TABLE OF CONTENTS

Signature Page.....	iii
Table of Contents.....	iv
List of Figures.....	viii
List of Tables.....	xii
Acknowledgements.....	xiii
Vita and Publications.....	xvi
Abstract.....	xvii
I. Introduction.....	1
A. mRNA Processing.....	2
B. SR Protein Family.....	3
C. Functional Roles of SR Proteins in Pre-mRNA Splicing.....	7
D. Phosphorylation States of SR Proteins .....	8
E. Expanded Roles of SR Proteins.....	10
F. SR Protein Kinases.....	12
G. Other Functions of SRPK1.....	15
H. SRPK1 and Diseases.....	16
I. Protein Kinases.....	17
J. Mechanisms of Kinase Regulation.....	18
K. Substrate Recognition and Specificity in Protein Kinases....	21
L. Focus of Study.....	23
II. Materials and Methods.....	26
A. Preparation of Recombinant DNA.....	27
1. SRPK1 Expression Plasmids.....	27
2. Mutagenesis of SRPK1.....	28
3. ASF/SF2 Expression Plasmids.....	30
B. Expression and Purification of Recombinant Protein.....	31
C. Limited Proteolysis.....	36

D.	Measuring Kinase Activity.....	36
E.	Testing Kinase Stability.....	37
F.	Start-delay-trap Assay.....	37
G.	Formation of SRPK1 Apo-enzyme Crystal and Data Collection.....	38
H.	Formation of Seleno-methionine Derivative Crystals of SRPK1 $\Delta$ NS1.....	39
	1. Purification of Seleno-methionine Derivative SRPK1 $\Delta$ NS1.....	39
	2. Crystallization and Data Collection.....	40
I.	Structure Solution of SRPK1 $\Delta$ NS1 crystal.....	41
J.	Formation of SRPK1 $\Delta$ NS1/Peptide/ADP Ternary Complex Crystal and Data Collection.....	42
K.	Structure Solution of SRPK1 $\Delta$ NS1/Peptide/ADP Ternary Complex.....	43
L.	Formation of SRPK1 $\Delta$ N1S1:ASF/SF2(BD):AMP-PNP Complex Crystal.....	43
	1. Recombinant SRPK1:ASF/SF2 Complex Formation...	43
	2. Crystallization of SRPK1 $\Delta$ N1S1:ASF/SF2(BD): AMP-PNP Complex.....	44
M.	Data Collection of SRPK1 $\Delta$ N1S1:ASF/SF2(BD):AMP-PNP Complex Crystal.....	45
N.	Structure solution of SRPK1 $\Delta$ N1S1:ASF/SF2(BD) : AMP-PNP Complex.....	45
III.	Identification of a Crystallizable Construct of SRPK1 and its Crystallization.....	47
A.	Introduction.....	48
B.	Results.....	51
	1. Identification of an Active Fragment of SRPK1.....	51
	2. Crystallization of an Active Fragment of SRPK1.....	59
	3. Data Collection of Apo- and Seleno-methionine Derivative Crystals of SRPK1 $\Delta$ NS1.....	62
C.	Discussion.....	65
IV.	Insights into the Constitutive Activity from the X-ray Crystal Structure of SRPK1 Apoenzyme.....	68

A.	Introduction.....	69
B.	Results.....	72
1.	Structure Solution of SRPK1 $\Delta$ NS1.....	72
2.	Overall Architecture of SRPK1 $\Delta$ NS1.....	74
3.	Orientation of Helix $\alpha$ C and Active Conformation.....	80
4.	Spacer Domain of SRPK1.....	82
5.	Activation Segment of SRPK1.....	88
6.	SRPK1-specific Interactions at the Activation Loop.....	90
7.	SRPK1 shows Unusual Resistance to Inactivation.....	92
C.	Discussion.....	96
1.	Alternative Interactions at Activation Segment of SRPK1.....	97
2.	Spacer Helices Stabilize Hinge Motion between the Lobes.....	99
3.	Biological Role of the Spacer Domain.....	99
V.	The SRPK1/Peptide/ADP Ternary Complex Structure Reveals a Docking Groove and its Interaction with ASF/SF2.....	102
A.	Introduction.....	102
B.	Results.....	105
1.	Rational Design of SRPK1 Substrate Peptide.....	105
2.	Formation and Data Collection of SRPK1 $\Delta$ NS1/Peptide/ADP Complex Crystals.....	106
3.	Structure Solution of SRPK1 $\Delta$ NS1/Peptide/ADP Ternary Complex.....	109
4.	Overall Structure of SRPK1 $\Delta$ NS1/Peptide/ADP Complex.....	112
5.	Interactions at the Docking Groove of SRPK1.....	116
6.	Substrate Docking Regulates the Mode of Phosphorylation.....	119
7.	Identification of a Docking Motif in ASF/SF2.....	122
8.	Conservation of Docking Groove in SRPKs.....	122
9.	Conservation of Docking Motif in SR Proteins.....	124
10	Functional Role of Helix $\alpha$ G.....	127
C.	Discussion.....	129
1.	The Docking Motif on ASF/SF2 is Responsible for High Affinity Interaction with SRPK1 and its Mode of Phosphorylation by SRPK1.....	129

2.	The Docking Motif of ASF/SF2 Limits its Phosphorylation by SRPK1 and is Important for its Localization to Nuclear Speckles.....	130
3.	Regulation of ASF/SF2 Phosphorylation by SRPK1 and Clk/Sty.....	131
4.	Docking Interactions between Kinases and their Binding Proteins.....	133
5.	Biological Implications of High Affinity Binding and Processive Phosphorylation.....	134
VI.	Insights into Substrate Recognition and Processive Phosphorylation from the SRPK1:ASF/SF2:AMP-PNP Complex Structure.....	136
A.	Introduction.....	137
B.	Results.....	138
1.	Purification and Complex Formation of SRPK1:ASF/SF2.....	138
2.	Design of SRPK1:ASF/SF2 Crystallization Targets.....	141
3.	Identification of Diffraction Quality SRPK1:ASF/SF2 Complex Crystals.....	146
4.	Data Collection of SRPK1:ASF/SF2:AMP-PNP Complex.....	148
5.	Structure Solution and Overall Architecture of the Complex.....	152
6.	SRPK1:ASF/SF2 Binding Interface.....	155
7.	Peptide at the Docking Groove of SRPK1.....	161
C.	Discussion.....	161
VII	Conclusion.....	165
A.	Constitutive Active SRPK1.....	166
B.	Docking Interactions between SRPK1 and ASF/SF2.....	166
C.	SRPK1 Substrate Recognition.....	167
	References.....	169



## LIST OF FIGURES

Figure 1.1	Spliceosome assembly.....	4
Figure 1.2	Domain organization of SR proteins.....	6
Figure 1.3	Phosphorylation-dependent localization of SR proteins.....	11
Figure 1.4	Domain organization of SRPKs.....	14
Figure 1.5	Protein kinase fold.....	19
Figure 1.6	Regulation mechanisms of protein kinases.....	22
Figure 1.7	Consensus sequence motifs of kinase substrates.....	24
Figure 3.1	Sequence alignment of SRPK1s.....	50
Figure 3.2	Purification of recombinant full length human SRPK1..	52
Figure 3.3	Regions of the spacer domain adopt helical structures.....	54
Figure 3.4	Purification of spacer-deleted SRPK1.....	55
Figure 3.5	Purification of SRPK1 $\Delta$ S1.....	57
Figure 3.6	Stable fragment of SRPK1.....	58
Figure 3.7	Purification of SRPK1 $\Delta$ NS1.....	60
Figure 3.8	SRPK1 is monodispersed at pH 6.5.....	61
Figure 3.9	Crystal of SRPK1 $\Delta$ NS1.....	63
Figure 4.1	Kinase domains of SRPKs are conserved.....	70
Figure 4.2	SRPK1 $\Delta$ NS1 is a monomer in solution.....	75

Figure 4.3	Overall architecture of SRPK1 $\Delta$ NS1.....	76
Figure 4.4	Crystal structure of Sky1p.....	77
Figure 4.5	Primary sequence and secondary structures of SRPK1 $\Delta$ NS1.....	79
Figure 4.6	Conformation of $\alpha$ C' insert.....	81
Figure 4.7	Small distortion of SRPK1 small lobe.....	83
Figure 4.8	Spacer helix $\alpha$ S1 resembles activator of Aurora B kinase.....	84
Figure 4.9	Spacer helix $\alpha$ S2 is structural mimic of binding proteins of other kinases.....	86
Figure 4.10	Activation loop conformation of SRPK1 and comparison with those of PKA and Sky1p.....	89
Figure 4.11	Unique SRPK1 interactions at the activation segment.....	91
Figure 4.12	Spacer inserts mediated interactions are important for maintaining kinase activity.....	94
Figure 4.13	Spacer helices are important to stabilize the conformation of small lobe.....	100
Figure 5.1	Domain organization of ASF/SF2.....	107
Figure 5.2	Electrostatic surface potential of SRPK1 $\Delta$ NS1.....	108
Figure 5.3	Phasing of SRPK1 $\Delta$ NS1/peptide/ADP ternary complex.....	113
Figure 5.4	Presence of a peptide and ADP in the complex crystal.....	114

Figure 5.5	Ribbon diagram of SRPK1 $\Delta$ NS1/peptide/ADP ternary complex X-ray structure.....	115
Figure 5.6	Nucleotide binding pocket of SRPK1 $\Delta$ NS1.....	117
Figure 5.7	Docking groove of SRPK1.....	118
Figure 5.8	The docking groove of SRPK1 plays an important role in its interaction with ASF/SF2.....	121
Figure 5.9	Identification of a docking motif in ASF/SF2.....	123
Figure 5.10	Conservation of docking groove in SRPK1.....	125
Figure 5.11	SRPK1 contains extended $\alpha$ F/ $\alpha$ G loop.....	126
Figure 5.12	SR proteins with two RRM s contain docking motifs....	128
Figure 5.13	Proposed model for a role of both SRPK1 and Clk/Sty kinases in phosphorylation and subcellular localization of ASF/SF2.....	132
Figure 6.1	Formation of SRPK1:ASF/SF2 complex.....	140
Figure 6.2	Constructs of ASF/SF2 and SRPK1 used for crystallization.....	142
Figure 6.3	Purification of different SRPK1:ASF/SF2 complexes...	144
Figure 6.4	Purification of crystallizable SRPK1:ASF/SF2 complex.....	145
Figure 6.5	SRPK1:ASF/SF2 complex crystals.....	147
Figure 6.6	Formation of SRPK1 $\Delta$ N1S1:ASF/SF2(BD):AMP-PNP complex crystals.....	149
Figure 6.7	Phasing of SRPK1 $\Delta$ N1S1:ASF/SF2(BD):AMP-PNP ternary complex structure.....	153

Figure 6.8	Electron density of ASF/SF2 and docking groove peptide.....	154
Figure 6.9	Crystal structure of SRPK1:ASF/SF2 complex.....	156
Figure 6.10	Structure of RRM2 of ASF/SF2.....	157
Figure 6.11	Buried surface area in SRPK1:ASF/SF2 complex.....	158
Figure 6.12	Binding interface of SRPK1:ASF/SF2 complex.....	160
Figure 6.13	Docking groove peptide.....	162

## LIST OF TABLES

Table 3.1	Data collection of SRPK1 $\Delta$ NS1 crystal.....	64
Table 4.1	Phasing and refinement statistics of SRPK1 $\Delta$ NS1.....	73
Table 4.2	Interactions at the activation segment of SRPK1.....	98
Table 5.1	Summary of data collection of 9mer soaked crystals...	110
Table 5.2	Data collection and refinement of SRPK1/peptide/ADP complex crystal.....	111
Table 6.1	Summary of SRPK1:ASF/SF2 complex formation and crystallization.....	150
Table 6.2	Data collection and refinement of SRPK1:ASF/SF2:AMP-PNP ternary complex crystal...	151

## ACKNOWLEDGEMENTS

I thank my parents, Richard and Irene, for giving me life, and their unconditioned caring and understanding in the past 29 years and years to come.

I thank my sisters, Becky and Vicky, their husband and fiancé, Philip and Natsuki, for being my family and all their support.

I thank my advisor and soccer buddy, Dr. Gourisankar Ghosh, for providing me the opportunity to study in his lab, being a wonderful mentor and giving me the help and encouragement to succeed as a graduate student. I also deeply appreciate for his confidence in my capabilities.

I thank a past member of the G.Ghosh lab, Dr. Bradley Nolen, for showing me all the experiments and guided me to become an independent researcher.

I thank professor Dr. Tom Huxford, for providing invaluable solutions when my project was stuck. This project would not have finished without his suggestions.

I thank other past and present lab members, Dr. Christopher Phelps, Dr. Anu Moorthy, Dr. Rashmi Talwar, Dr. Eileen Kennedy, Dr. De-bin Huang, and Devin Drew, for providing invaluable support and intellectual contributions.

I thank my friend, Don Vu, for showing me to do my first transformation in the lab, taking care of my cat when I am away and being a good friend.

I thank my grant-writing buddy, Sutapa Chakrabarti, for her significant contribution to the project, providing support during the long night of grant-writing, and listening to all my dramas.

I thank my eating buddy, Jessica Ho, for having great food together so I can survive the long night at lab, and for the crazy stories we shared.

I thank my drinking buddy, Erika Mathes, for showing me how to take my first jello shot and get buzzed.

I thank my friend, Amanda Fusco, for letting me to tag along during the ski trip and took “that” picture.

I thank my undergraduates, Kayla Giang and Melinda Jean Yeh, for their hard work and finishing all my experiments in the lab when I’m writing this long acknowledgement.

I thank the rest of the lab, Anthony Farina, Vivien Wang, Olga Savinova, and others for their friendships and making the lab a fun place to work in.

I thank my friend, Case McNamara, for the discussion and his help in crystallography.

I thank our collaborator, Dr. Joseph A. Adams and his lab, for helping with the kinetics experiments.

I thank our collaborator, Dr Xiang-Dong Fu and his lab, for providing the clones and helping with the immunolocalization experiments.

I thank our collaborator, Dr. J. Andrew McCammon and Dr. Justin Gullingsrud, for conducting the MD simulations, and for all the patience they have with our manuscript.

I thank member of my committee, Dr. Susan Taylor, for taking time and interest in my project, and Dr. Choel Kim from the Taylor lab for all his help with synchrotron time.

I thank my friends, Sit Yiu Fai, Eva Wong, Vivian Wong, Karmen Suen, Grace Mao and Tina Cheng for all their support and listening to my constant complaints.

But most of all, I am deeply grateful to my fiancé, Rosa Cheng, for her faith, love, devotion and patience. Without her support and encouragement, I might have never found the mental strength or endurance to complete graduate school. I also thank her for spending numerous long nights with me in Urey Hall purifying proteins.

Finally, I'd like to acknowledge the NIH/NCI Cancer Training Grant as my funding resource.

The text of Chapter V is, in part, reprints of material as it appears in *Molecular Cell*, 2005, **20**, 77-89, Ngo, J. C., Chakrabarti, S., Ding, J. H., Velazquez-Dones, A., Nolen, B., Aubol, B. E., Adams, J. A., Fu, X. D. and Ghosh, G.. The dissertation author was the primary researcher and author of this publication.



## VITA

2001	B.S., University of California, San Diego
2004	M.S., University of California, San Diego
2006	Ph.D., University of California, San Diego

## PUBLICATIONS

Ngo, J.C.K., Chakrabarti, S., Giang, K., Fu, X-D., Adams, J.A., and Ghosh, G. (2006) Crystal structure of SRPK1:ASF/SF2 complex: insights into regulation of SR protein. (Manuscript in preparation)

Ngo, J.C.K., Gullingsrud, J., Giang, K., Yeh, M. J., Fu, X-D., Adams, J.A., McCammon, J.A. and Ghosh, G. (2006) SR protein kinase 1 is resilient to inactivation. (Manuscript in preparation)

Ngo, J.C.K., Chakrabarti, S., Ding J.H., Velazquez-Dones A., Nolen B., Aubol B.E., Adams, J.A., Fu X.D. and Ghosh, G. (2005) Interplay between SRPK and Clk/Sty kinases in phosphorylation of the splicing factor ASF/SF2 is regulated by a docking motif in ASF/SF2. *Molecular Cell*, **20**, 77-89.

Aubol, B.E., Chakrabarti, S., Ngo, J., Shaffer, J., Nolen, B., Fu, X-D., Ghosh, G., and Adams, J.A. (2003) Processive phosphorylation of alternative splicing factor/splicing factor 2. *Proc. Natl. Acad. Sci. USA*, **100**, 12601-12606.

Nolen, B., Ngo, J., Chakrabarti, S., Vu, D., Adams, J.A., and Ghosh, G. (2003) Nucleotide-induced conformational changes in the *Saccharomyces cerevisiae* SR protein kinase, Sky1p, revealed by X-ray crystallography. *Biochemistry*. **42**, 9575-9585.

ABSTRACT OF THE DISSERTATION

**SR Protein Kinase 1: Conformation, Substrate Recognition and Catalysis**

by

Jacky Chi Ki Ngo

Doctor of Philosophy in Chemistry

University of California, San Diego, 2006

Professor Gourisankar Ghosh, Chair

Serine/arginines-rich (SR) proteins play important roles in constitutive and alternative pre-mRNA splicing, mRNA export and translation. The C-terminal RS domains of these proteins contain repeats of consecutive arginine-serine (RS) dipeptides, which are extensively phosphorylated and mediate protein-protein and protein-RNA interactions. One of the kinase families that phosphorylate SR proteins is the SR protein kinase family. SRPKs are unusual members of the kinase superfamily in that their kinase domains are bifurcated by spacer regions of 250 to 300 residues. Furthermore, they are constitutively active and do not require any post-translational modification or interaction with regulatory factors.

In this study, I employed a combination of X-ray crystallography and biochemical techniques to investigate the molecular basis of constitutive activity, catalytic mechanism and substrate recognition of human SRPK1. An appropriate fragment of SRPK1 was designed for crystallization. The determination of the SRPK1 X-ray structure at 1.73 Å revealed that SRPK1 contains unique secondary structural elements that mimic in-trans activation mechanisms observed in different kinases. Moreover, the global network of interactions allows the activation loop to adopt a catalytically competent conformation without relying upon any direct stabilization mechanism.

SRPK1 binds its SR protein substrate, ASF/SF2, stably and processively phosphorylates only half of the serines of the RS domain. The crystal structure of SRPK1 bound to a peptide and ADP has led to the identification of a docking groove in SRPK1 and a docking motif in ASF/SF2. Detailed biochemical studies of the docking interactions suggested mechanisms for the restriction in the number of phosphorylated sites and the mode of regulation of phosphorylation. Finally, the current model of the X-ray crystal structure of SRPK1 bound to AMP-PNP and a natural substrate, ASF/SF2, has revealed the structural basis for substrate recognition. Unexpectedly, ASF/SF2 engages SRPK1 at both small and large lobes upon binding. Detailed analysis of the complex structure is underway and will provide insights into the substrate specificity and the mechanism of processive phosphorylation.

## **Chapter I: Introduction**

## **A. mRNA Processing**

The protein-coding regions, known as exons, in most pre-mRNA are interrupted by intervening sequences called introns. The assembly of exons into mature mRNA requires a process termed splicing where the introns are eliminated. In higher eukaryotes, the splicing event can occur in two ways: constitutive and alternative splicing. During constitutive splicing, all introns in the primary transcript are spliced out and all exons are joined together in the order they are present in the nascent transcript. In contrast, alternative splicing allows different ways to join the exons of the same pre-mRNA template. Alternative splicing therefore provides cells the possibility to express multiple proteins, usually with distinct functions, from a single gene and results in a greater diversity in an organism's pool of protein products than in its genome (Smith and Valcarcel, 2000). In the human genome, for example, it was estimated that more than 70% of genes have at least two splice variants (Johnson et al., 2003; Lander et al., 2001). In addition, there is growing evidence that transcription, several pre-mRNA processing steps including splicing and the export of mature mRNA to the cytoplasm are extensively coupled to each other, adding a high degree of complexity to the regulation of gene expression (Maniatis and Reed, 2002; Reed and Hurt, 2002). However, knowledge of the molecular mechanisms that regulate pre-mRNA processing or mRNA transport is limited.

Pre-mRNA splicing and mRNA transport are highly dynamic and complex processes that require a series of protein-protein, protein-RNA and

RNA-RNA interactions. Further complexity is added by the differential modifications of these protein and RNA components. Splicing of pre-mRNAs is mediated by a macromolecular complex called the spliceosome. The assembly of the spliceosome is highly coordinated and involves five U small nuclear ribonucleoproteins (snRNPs): U1, U2, U4, U5 and U6 (Figure 1.1). In addition to the U snRNPs, many non-snRNP proteins are also essential for the spliceosome assembly and splicing process. One important family of proteins that is involved in these processes is the serine/arginines-rich (SR) proteins family.

## **B. SR Protein Family**

The SR proteins represent a family of splicing factors that are essential for both constitutive and alternative splicing and are highly conserved in metazoans and plants (Birney et al., 1993; Lazar et al., 1995; Lopato et al., 1996). However, they are not found in all eukaryotes. For example, the budding yeast *Saccharomyces cerevisiae* does not contain any SR protein. To date, more than ten different members of the human SR protein family have been identified: alternative splicing factor/splicing factor 2 (ASF/SF2 / SRp30a), SC35 (SRp30b), 9G8, SRp20, SRp30c, SRp38, SRp40, SRp46, SRp54, SRp55 and SRp75 (Cavaloc et al., 1994; Fu and Maniatis, 1992; Roth et al., 1990; Sreaton et al., 1995; Shin and Manley, 2002; Soret et al., 1998; Zahler et al., 1992; Zhang and Wu, 1996). All SR proteins share a common structural organization: the N-terminal part consists of one or two RNA

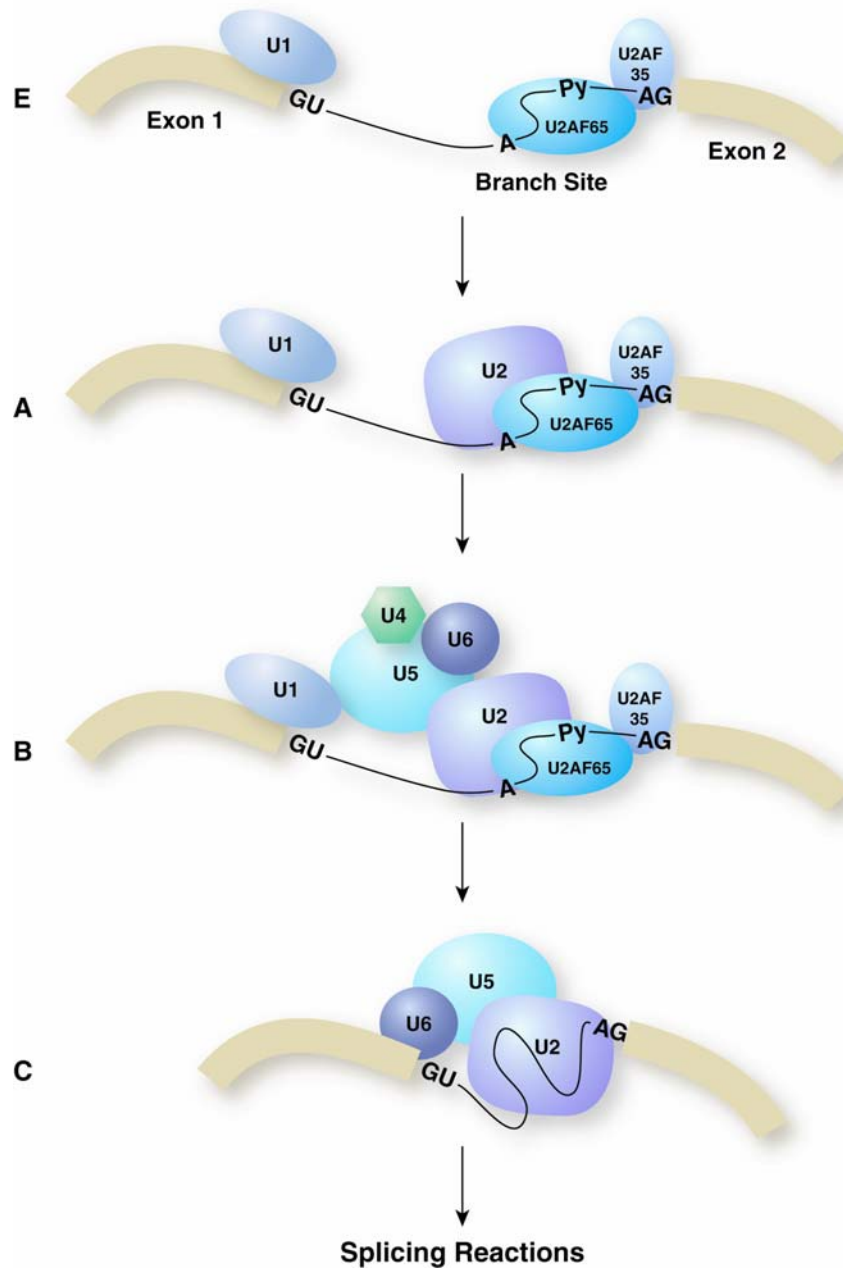


Figure 1.1 Spliceosome assembly. In the first step of spliceosome assembly, U1 snRNP binds to the 5' splice site and U2 auxiliary factor (U2AF) binds to the pyrimidine tract, 3' splice site and contacts the branch point at the same time. This E (or early) complex commits the pre-mRNA to splicing. Next, U2 snRNP is recruited to the branch point and forms the A complex. The A complex is converted to the B complex after the recruitment of the U4, U5, U6 tri-snRNP complex. Finally, the catalytically active C complex is formed after the release of U1 and U4 snRNPs and a rearrangement of the U2, U5 and U6 snRNPs.

recognition motif(s) (RRM(s), also known as RNA binding domain (RBD)), and the C-terminus contains different lengths of arginine/serine (RS) dipeptide repeats, referred to as the RS domain (Figure 1.2). The N-terminal RRMs are responsible for sequence-specific RNA binding whereas the RS domain is believed to function as both a protein-protein and RNA-protein interaction domain, enhancing assembly of the spliceosome (Caceres and Krainer, 1993; Shen and Green, 2004; Shen et al., 2004; Tacke and Manley, 1995; Zuo and Manley, 1993). While the RRM is generally organized into four anti-parallel  $\beta$  strands and two  $\alpha$  helices and are arranged in  $\beta 1-\alpha 1-\beta 2-\beta 3-\alpha 2-\beta 4$  order, the RS domain is predicted to be highly disordered (Haynes and Iakoucheva, 2006; Nagai et al., 1995a; Nagai et al., 1995b). There is also evidence that the RRM and RS domains are individual functional modules as they are interchangeable between different SR proteins (Chandler et al., 1997; Graveley and Maniatis, 1998; Wang et al., 1998b).

SR proteins are localized predominantly in the nucleus in steady-state cells, however a subset of SR proteins, like ASF/SF2, 9G8 and SRp20, shuttle continuously between the cytoplasm and the nucleus (Caceres et al., 1998). Although these proteins do not have a classical nuclear localization signal (NLS), the non-shuttling SR proteins have been shown to contain a nuclear retention signal (NRS) (Cazalla et al., 2002). In the interphase nucleus, SR proteins and many other splicing factors are concentrated in nuclear structures called interchromatin granule clusters (IGCs) or speckles. During transcription, the SR proteins are recruited to the site of pre-mRNA synthesis to perform



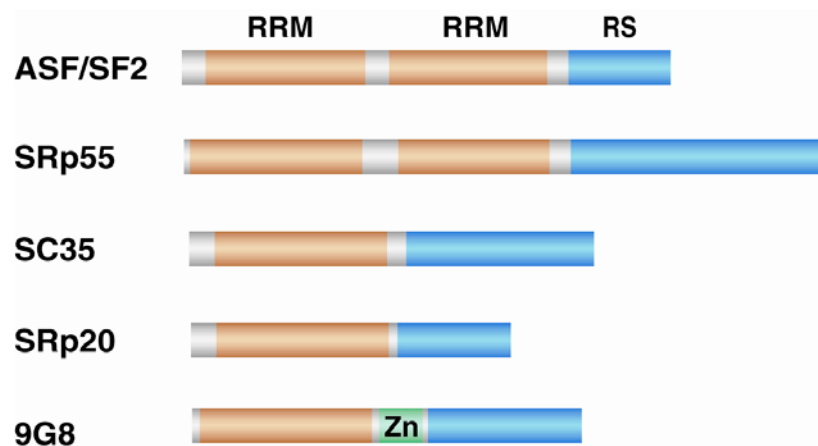


Figure 1.2 Domain organization of SR proteins. Schematic diagram of several human SR proteins. SR proteins contain either one or two RNA recognition motif(s) (RRM) at the N-terminal regions, followed by RS domains rich in RS dipeptide repeats. The SR protein 9G8 contains a zinc knuckle between the RRM and RS domain.

their roles in pre-mRNA processing. It appears that the RS domain, which could undergo phosphorylation, is essential in this recruitment process (Misteli et al., 1998).

### **C. Functional Roles of SR Proteins in Pre-mRNA Splicing**

SR proteins are required for both 5' and 3' splice sites selection in constitutive and alternative splicing. They bind directly to exon sequences and during alternative splicing of certain introns, also bind to additional positive cis-acting elements called exonic splicing enhancers (ESEs). The combination of SR proteins bound at a particular splice site and ESE determines the strength or probability whether an exon will be included or excluded during the splicing process (Graveley, 2000; Graveley et al., 1998; Schaal and Maniatis, 1999). Surprisingly, despite the fact that different SR proteins do display distinct RNA-binding specificities, the consensus RNA sequences they recognize are degenerate (Graveley, 2000). In some cases, some SR proteins are shown to recognize binding site for a different SR protein (Liu et al., 1998). Despite the apparent redundancy in their functions, SR proteins are not completely exchangeable in their roles in alternative splicing. For example, recent gene knockout experiments in mice have shown that in the absence of SC35, T-cell development is affected due to a defect in alternative splicing of CD45, a receptor tyrosine phosphatase known to be regulated by differential splicing during thymocyte development (Wang et al., 2001). On the other hand, cardiomyocytes lacking ASF/SF2 display defects in hypercontraction due to

deviant alternative splicing of the  $\text{Ca}^{2+}$ /calmodulin-dependent kinase II $\delta$  transcript (Xu et al., 2005).

During the splicing event, SR proteins bind to the pre-mRNA and recruit elements of the splicing machinery through RS domain mediated protein-protein interactions. SR proteins act at the stage of E-complex assembly by recruiting U1 snRNP and U2AF to the 5' and 3' splice site respectively (Reed, 1996; Staknis and Reed, 1994). Furthermore, SR proteins also help to recruit U2 snRNP and function during the transition from A to B complex, where the U4/U6.U5 tri-snRNP complex is incorporated into the spliceosome (Rosciigno and Garcia-Blanco, 1995; Tarn and Steitz, 1995). Finally, in addition to mediating protein-protein interactions, the RS domains of SR proteins can also contact the pre-mRNA branch point and 5' splice site directly to promote mature spliceosome assembly (Shen and Green, 2004; Shen et al., 2004).

#### **D. Phosphorylation States of SR Proteins**

SR proteins are extensively phosphorylated at multiple serines in their RS domains (Fu, 1995; Mermoud et al., 1994; Roth et al., 1990). It was suggested that different phosphorylation states of SR proteins could modulate their protein-interacting and RNA-binding properties. For example, the phosphorylation of ASF/SF2 enhances its interactions with U1 70K, an RS domain-containing component of U1 snRNP, and diminishes non-specific RNA binding of the RS domain (Xiao and Manley, 1997). Furthermore, it has been well documented that reversible phosphorylation of SR proteins is required

during the splicing event (Cao et al., 1997; Kanopka et al., 1998). Indeed, it was shown that both hypo- and hyperphosphorylation of the RS domains could inhibit the splicing activities of SR proteins, suggesting that appropriate regulation of phosphorylation is critical to the splicing reaction (Prasad et al., 1999).

Besides their functions in the splicing event, the localization and intracellular trafficking of SR proteins are also dependent on their phosphorylation states (Colwill et al., 1996; Gui et al., 1994a; Yeakley et al., 1999). It is generally believed that SR proteins predominantly reside in the nuclear speckles and hyperphosphorylation of these proteins redistribute them in the nucleoplasm and recruit them to the site of splicing (Colwill et al., 1996; Misteli et al., 1998). Furthermore, the nuclear import of SR proteins mediated by transportin-SR, a member of the importin- $\beta$  family, also requires the phosphorylation of SR proteins (Lai et al., 2000; Lai et al., 2001; Yun et al., 2003). It is apparent that the biological function of SR proteins is highly dependent of their phosphorylation; however, the extensive number of serines present within the RS domains adds another level of complexity when one has to distinguish the precise phosphorylation states, namely hypo- and hyperphosphorylation, of the SR protein in these processes. Therefore, in order to understand the specific role of SR proteins in regulation of pre-mRNA splicing, it is important to study how SR proteins are recognized by the kinases and which serines within their RS domains are phosphorylated. Figure 1.3 summarizes the effects of phosphorylation on SR proteins localization.

## **E. Expanded Roles of SR Proteins**

There is growing evidence that SR proteins also serve additional roles in post-transcriptional gene expression including mRNA export. The subset of shuttling SR proteins including 9G8, SRp20 and ASF/SF2 has recently been shown to function as adapter proteins by recruiting the general mRNA export receptor TAP/NXF1 to mRNAs (Huang et al., 2003; Huang and Steitz, 2001). Moreover, the export receptor TAP/NXF1 preferentially binds mature mRNA and hypophosphorylated SR proteins, suggesting that the SR proteins have to undergo dephosphorylation after their recruitment to the site of splicing in order to be involved in the export of the spliced mRNA (Huang et al., 2004; Lai and Tarn, 2004) (Figure 1.3). In agreement with this, a recent study has shown that shuttling and non-shuttling SR proteins undergo distinct recycling pathway during mRNP maturation in a dephosphorylation – dependent manner (Lin et al., 2005).

Another surprising post-splicing function of SR proteins is that they also take part in translation regulation. ASF/SF2 was found to associate with polyribosomes in cytoplasmic extracts and enhance the translation of a luciferase reporter mRNA both in vivo and in vitro (Sanford et al., 2004). Furthermore, endogenous ASF/SF2 that is associated with the translation machinery was hypophosphorylated and showed increased mRNA binding and activity in translation, once again demonstrating the importance of the phosphorylation status of SR proteins in their subcellular localization and functions (Sanford et al., 2005). Finally, SR proteins have also been shown to

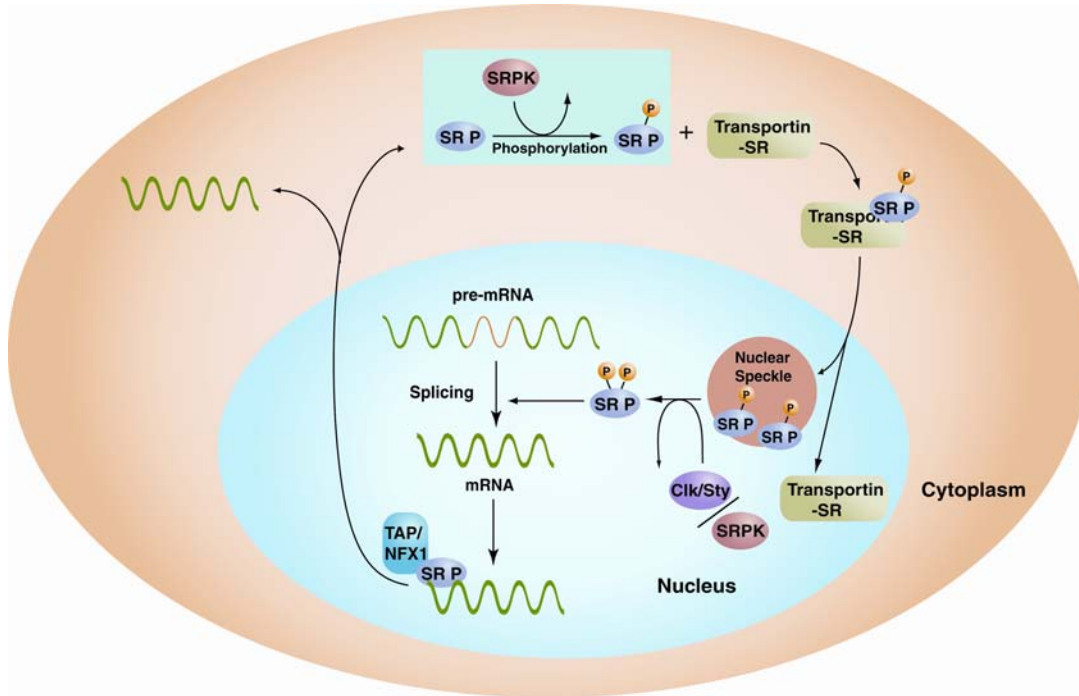


Figure 1.3 Phosphorylation-dependent localization of SR proteins. Hypophosphorylated SR proteins are imported into the nucleus through interactions with transportin-SR proteins (Hubbard, 1997; Lai et al., 2000; Lai et al., 2001; Lowe et al., 1997). The imported SR proteins are stored in nuclear speckles and recruited to the site of splicing upon hyperphosphorylation (Gui et al., 1994a; Yang et al., 2002). Dephosphorylated SR proteins served as adaptor protein between the mRNA export receptor TAP/NFX1 and mRNA and facilitates mRNA export (Huang et al., 2004; Lai and Tarn, 2004).

affect genomic stability, the cytoplasmic mRNA stability of specific transcript, and involved in nonsense-mediated mRNA decay (Lemaire et al., 2002; Li and Manley, 2005; Zhang and Krainer, 2004).

## **F. SR Protein Kinases**

Several protein kinases have been reported to phosphorylate the RS domains of SR proteins. These include members of the SR protein kinase (SRPK) family, Cdc2-like kinase/Ser-Thr-Tyr (Clk/Sty) kinase family, DNA topoisomerase I and Akt (Blaustein et al., 2005; Colwill et al., 1996; Gui et al., 1994b; Rossi et al., 1996). SRPKs are highly RS-specific kinases and only phosphorylate serine (not threonine) in the presence of arginine (not lysine) in their substrates (Gui et al., 1994b; Wang et al., 1998a). In particular, SRPK1 processively phosphorylates only one half of the RS dipeptides within the RS domain of ASF/SF2 *in vitro* (Aubol et al., 2003). However, the mechanism and the biological implication of processive and restrictive phosphorylation are yet to be elucidated.

To date, three members of the SRPK family are known to be expressed in mammalian cells: SRPK1, SRPK1a (an alternative spliced form of SRPK1) and SRPK2. SRPK1 was the first discovered SR protein-specific kinase and was identified based on its activity to induce the disassembly of nuclear speckles during the cell cycle (Gui et al., 1994a). Like SRPK1, SRPK1a and SRPK2 share the same substrate specificity and phosphorylate the RS domains of SR proteins (Nikolakaki et al., 2001; Wang et al., 1998a).

Moreover, all three SRPKs are found predominantly in the cytoplasm and could translocate into the nucleus. The subcellular localization of SR proteins was shown to be regulated by their phosphorylation by SRPKs, implying the importance of SRPKs in the regulation of splicing reactions (Gui et al., 1994a; Koizumi et al., 1999). However, the precise role of SRPKs in the regulation of SR proteins still remains unanswered. The expression levels of SRPKs are highly tissue specific. While SRPK1a is only found in testis, SRPK1 is highly expressed in both testis and pancreas, and SRPK2 is highly expressed in brain, suggesting that different SRPKs might contribute to splicing regulation in different tissues and might not be able to compensate the function of each other (Hayes et al., 2006).

SRPK1 and other members of the SRPK family are constitutively active kinases; i.e. they do not require post-translational modification or interaction with other factors to be activated *in vitro*. The sequences of SRPKs also reveal that they are unusual members of the protein kinase superfamily as their kinase domains are bifurcated by spacer regions of ~250-300 residues in length (Figure 1.4). The spacer regions of SRPKs have been shown to anchor the kinases in the cytoplasm during interphase and their deletions resulted in exclusive localization of the kinases to the nucleus (Ding et al., 2006; Siebel et al., 1999; Wang et al., 1998a). Members of the SRPK family also contain non-homologous N- and C-terminal extensions of various lengths. Although the roles of these extensions are not clear, it has been shown that the C-terminal



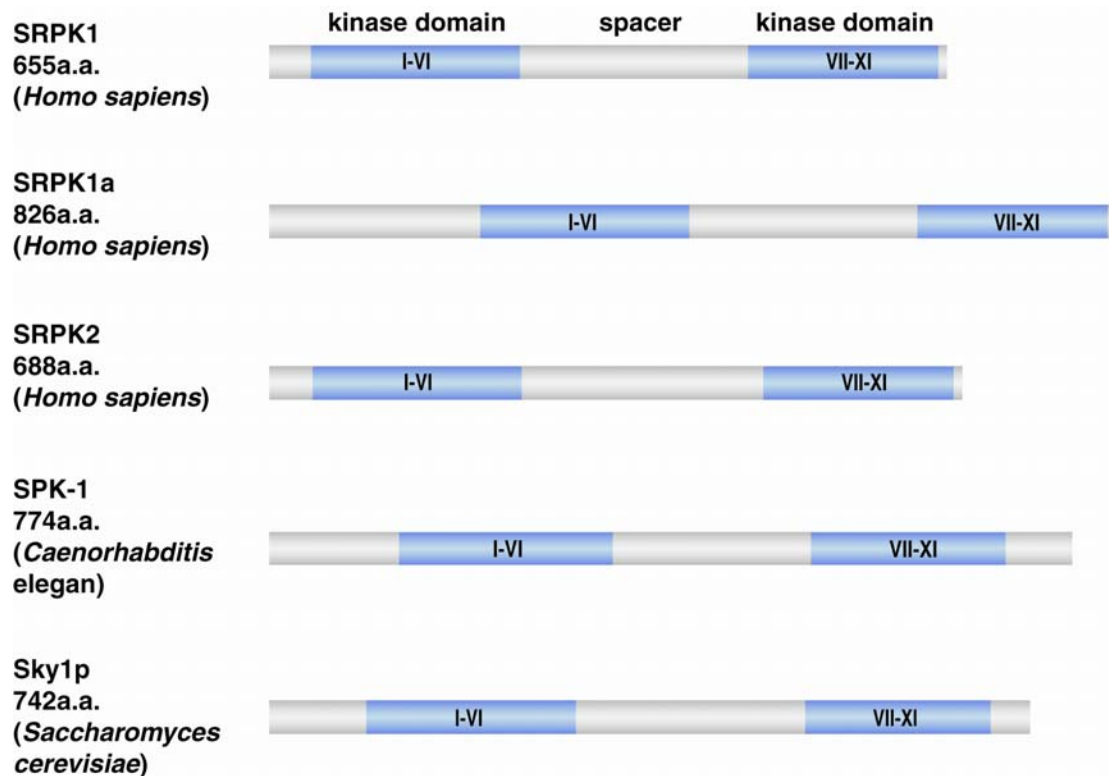


Figure 1.4 Domain organization of SRPKs. The conserved kinase subdomains (blue) defined by Hanks et. al are bifurcated by spacers ~250-300 residues in length (Hanks and Quinn, 1991). SRPKs also contain non-homologous N- and C-terminal extensions that varied in length.

extension in Sky1p is involved in maintaining the constitutive kinase activity (Nolen et al., 2001).

### **G. Other Functions of SRPKs**

Besides the splicing factors described above, there are other classes of protein that also contain RS dipeptide motifs. Indeed, two non-SR proteins, lamin B receptor (LBR) and protamine 1, have been found to contain RS motifs and are substrates of SRPK1. LBR is an integral protein of the inner nuclear membrane that contains a stretch of RS dipeptides at its N-terminal region and is important for nuclear envelope assembly. Recent studies have shown that SRPK1 resembles the substrate specificity and activity of an LBR-associated kinase, suggesting the SRPK1 is indeed responsible for the phosphorylation of LBR and regulates its binding to chromatin (Papoutsopoulou et al., 1999b; Takano et al., 2004). Protamine 1 belongs to a family of small arginine-rich proteins that replace histones during spermatogenesis, resulting in extreme chromatin condensation. The function of protamines is regulated by the phosphorylation state of its RS motifs. As mentioned above, SRPK1 is highly expressed in testis, and the fact that it phosphorylates protamine 1 suggests SRPK1 might play an important role during the development of mature spermatozoa (Papoutsopoulou et al., 1999a).

SRPKs may also serve yet another function in signaling pathways governing apoptosis. In a recent study, SRPK1 was found to be activated early

during Fas-mediated apoptosis. The kinase is then inactivated by caspase-mediated proteolysis and such cleavage can be inhibited by the overexpression of Bcl-2 and Bcl-X<sub>L</sub>. In agreement to this observation, both SRPK1 and SRPK2 were shown to be *in vivo* and *in vitro* substrates for caspase-8 and -9 respectively (Kamachi et al., 2002). Although the precise function of SRPKs during apoptosis remains unclear, their roles in splicing suggest that the kinases might respond to apoptotic stimuli by regulating relative levels of splice variants of certain apoptosis regulatory molecules.

#### **H. SRPK1 and Diseases**

Multiple lines of evidence have emerged in the last few years implying the importance of SRPK1 in cancer studies. For example, SRPK1 was found to be necessary for sensitivity of two anti-cancer drugs, cisplatin and bleomycin (Sanz et al., 2002; Schenk et al., 2001). In both cases, the inactivation of SRPK1 seems to be responsible for resistance to the anti-cancer drugs. In support of this, the downregulation of SRPK1, which has a high expression level in testicular tissues in mammals, seems to determine the unresponsiveness of testicular germ cell tumor (GCT) patients to platinum-containing chemotherapy (cisplatin) (Schenk et al., 2004). Immunohistochemistry done on tumor samples from pancreas, breast and colon has revealed that SRPK1 is overexpressed in the epithelial neoplasms of these tissues, suggesting SRPK1 expression is abnormal in malignant cells. However, in contrast to the cisplatin unresponsiveness in GCTs patients, down

regulation of SRPK1 by siRNA in pancreatic tumor cell lines resulted in enhanced drug sensitivity of gemcitabine and/or cisplatin (Hayes et al., 2006). Furthermore, SRPK1 was also found to be overexpressed in tumor cell samples from patients of adult T-cell leukaemia (ATL) and chronic myelogenous leukemia (CML) (Hishizawa et al., 2005; Salesse et al., 2004). Together these data support SRPK1 as a potential target for the treatment of multiple forms of cancer and structural studies on the kinase will provide invaluable aid in drug design.

As regulators of splicing, the SRPKs may play important roles in the life cycle of viruses that hijack cellular splicing machinery to process their own transcripts. Indeed, recent work has shown that herpes simplex virus protein ICP27 alters SRPK1's localization and activity to affect the phosphorylation of SR proteins, which results in decreased levels of cytoplasmic mRNA and eventually protein synthesis shut-off (Sciabica et al., 2003). Since these viral-mediated modifications are necessary steps for viral infection, SRPK1 may be an important host cell target for therapeutic intervention. Finally, a recent study shows that SRPK1 and SRPK2 serve negative roles in the regulation of hepatitis B virus replication by suppressing the encapsidation of their pregenomic RNA (Zheng et al., 2005).

## **I. Protein Kinases**

The human genome comprises 518 protein kinases that constitute 1.7% of all human genes (Manning et al., 2002). In view of this, it is not surprising

that protein kinases can have profound effects on eukaryotic cells. Indeed, through phosphorylation of about 30% of human proteins, protein kinases mediate most of the signal transduction in eukaryotic cells and are crucial to metabolism, gene expression, cell growth and other cellular responses. Although protein kinases vary in size, function or cellular distribution, their kinase domain cores are conserved and contain multiple signature sequences and residues (Hanks and Quinn, 1991). The crystal structure of cAMP-dependent protein kinase (PKA), the first X-ray structure of protein kinase, has defined a common fold that can be applied to other members of the protein kinase family (Knighton et al., 1991a). The conserved kinase fold is comprised of two lobes: the N-terminal small lobe consists mainly of  $\beta$ -strands and an  $\alpha$ -helix; and the C-terminal large lobe which is mostly  $\alpha$ -helical (Figure 1.5). The small lobe constitutes most of the important features and residues necessary for  $Mg^{2+}$ /ATP binding, including the glycine rich loop, and the invariant K72 and E91 ionic pair (residues are numbered according to PKA structure), which are important to position the phosphates of ATP for catalysis. Between the two lobes is a cleft where the  $Mg^{2+}$ /ATP binds and the catalysis occurs. The large lobe contains the catalytic loop and the activation segment.

## **J. Mechanisms of Kinase Regulation**

Given their important roles in cells, the understanding of molecular mechanisms of protein kinase regulation is essential for the study of signaling

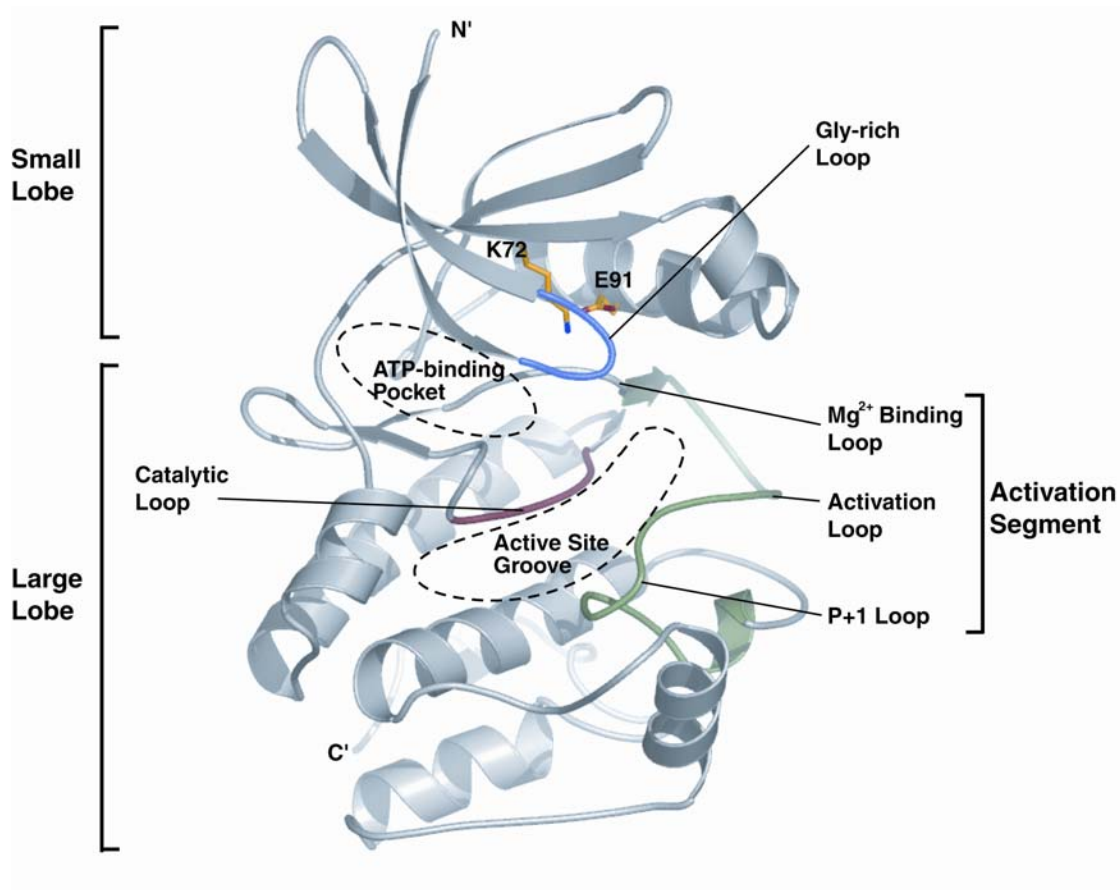


Figure 1.5 Protein kinase fold. Ribbon diagram of PKA (1ATP.pdb) (Knighton et al., 1991a). The N-terminal small lobe consists of mostly B-strands while the C-terminal large lobe is dominated by  $\alpha$ -helices. Important conserved kinase features are indicated.

processes. Kinases are tightly regulated through several mechanisms. The most common kinase regulatory mechanism involves phosphorylation of a loop segment, termed the activation loop, or binding of regulatory domains. The activation segment is defined as the regions spanning the conserved sequence DFG and APE within the large lobe and includes the secondary structural elements, the magnesium binding loop, the activation loop and the P+1 loop (Johnson et al., 1996; Nolen et al., 2004). Phosphorylation of one or more residues in the activation loop usually induces large conformational changes in the loop that switch the kinase between active and inactive states. For example, the activation loop of the insulin receptor kinase (IRK) collapses into the active site, preventing substrate binding in its unphosphorylated state (Hubbard et al., 1994). Upon phosphorylation of three tyrosine residues in its activation loop, it moves away from the active site and adopts a conformation that allows substrate binding (Figure 1.6a)(Hubbard, 1997). Another way to achieve activation is through binding of separate regulatory subunits. One of the best examples is the structure of cyclin-dependent protein kinase 2 (CDK2). In its inactive state, the activation loop of CDK2, like IRK1, is collapsed into the active site and blocks substrate binding. Furthermore, the invariant lysine-glutamate ion pair is disrupted due to the mispositioning of helix  $\alpha$ C, induced by the collapsed activation loop (De Bondt et al., 1993). Upon binding of cyclin, the orientation of helix  $\alpha$ C is changed, restoring the invariant ion pair. Furthermore, the activation loop also adopts a conformation

that relieves the blockage of the active site (Jeffrey et al., 1995). Finally, the phosphorylation of T160 in the activation loop further stabilizes the cyclin binding and fully activates CDK2 (Figure 1.6b)(Russo et al., 1996). In both cases, the conformations of the activation loop have profound effects on the kinases' activity.

### **K. Substrate Recognition and Specificity in Protein Kinases**

Despite the common kinase fold, protein kinases exhibit high specificity toward their substrates. Protein kinases can select serine/threonine (Ser/Thr kinases), tyrosine (Tyr kinases) or both (dual-specificity kinases) residues as the phosphate acceptor. The high substrate specificity usually originates from the recognition of residues flanking the phosphorylatable residue, referred as the consensus sequence motifs (Figure 1.7), although some kinase-substrate pairs possess distal elements that aid in recognition and binding (Biondi and Nebreda, 2003). Several kinase-peptide complexes solved to date have provided insights into the molecular basis of substrate recognition. In the structures of PKA:PKI, PKB:GSK3-peptide, PhK:MC-peptide and CDK2:substrate peptide complexes, the substrate peptides adopt extended structures, leaving the phosphorylatable residues well-exposed for catalysis, and interacts with the kinases' active site grooves (Brown et al., 1999; Knighton et al., 1991b; Lowe et al., 1997; Yang et al., 2002). Besides the ionic interactions between the substrate peptides and charged residues



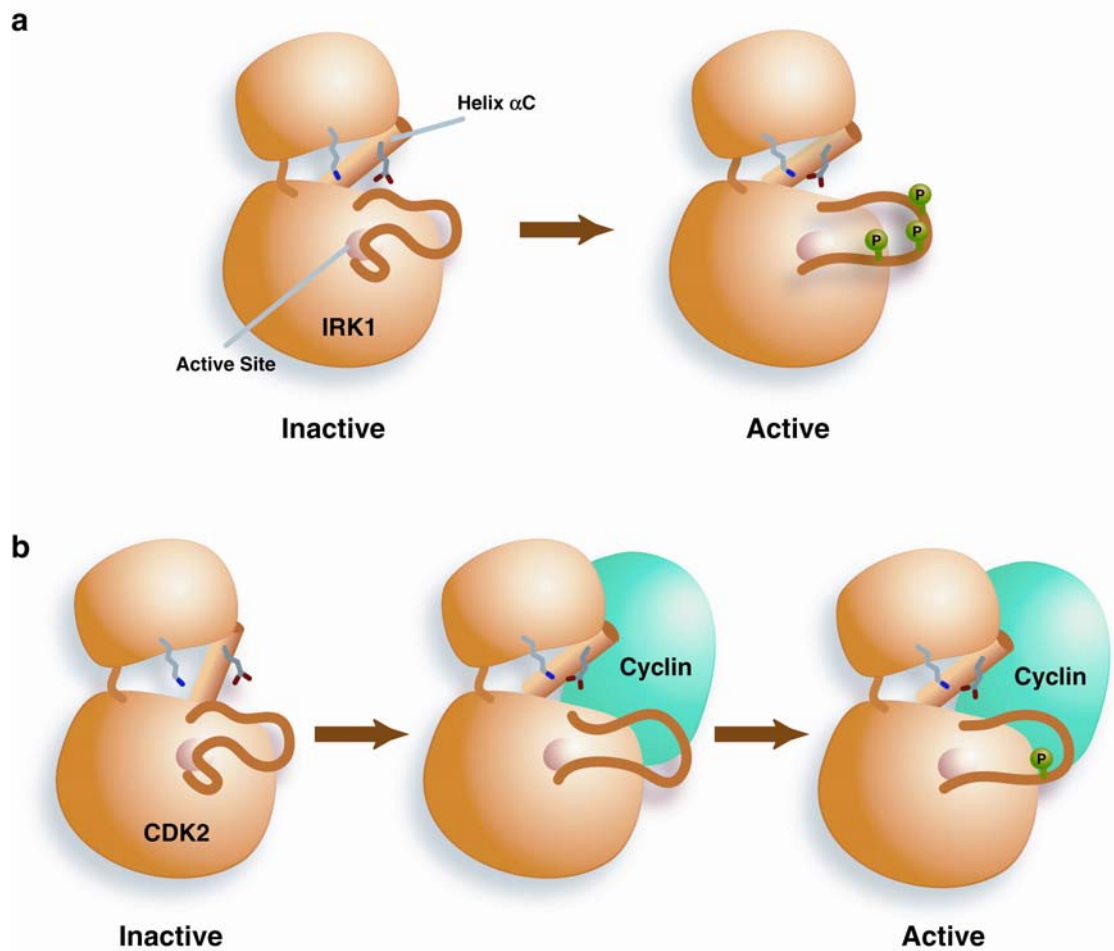


Figure 1.6 Regulation mechanisms of protein kinases. a) Phosphorylation at the activation loop. IRK contains three phosphorylatable tyrosines at its activation loop. In the unphosphorylated state, the activation loop collapses and blocks the active site, and disrupts the conformation of the activation segment. Upon phosphorylation on the tyrosines, the activation loop is reordered and exposes the active site for catalysis (Hubbard et al., 1994). b) Binding of regulatory domain. In the unphosphorylated state of CDK2, the helix  $\alpha$ C is mispositioned and the invariant glutamate is moved away from the lysine. Furthermore, the activation loop adopts a conformation that blocks substrate binding. After the binding of cyclin, the helix  $\alpha$ C is reoriented into the correct position and the invariant lysine-glutamate ion pair is restored. The activation loop now adopts an active conformation, and full activity of CDK2 is achieved upon phosphorylation on the activation loop, which further stabilizes the conformation of the loop and cyclin-binding (Brown et al., 1999; Colwill et al., 1996; De Bondt et al., 1993; Jeffrey et al., 1995; Misteli et al., 1998; Russo et al., 1996).

from the kinases, the non-polar residues within the consensus sequences are usually deeply buried into hydrophobic pockets formed on the kinases' surfaces, suggesting that the hydrophobic residues around the P0 site of the substrates also play significant roles in recognition and specificity (Figure 1.7).

Recently, the first crystal structure of a kinase complexed with a full length protein substrate was solved (Dar et al., 2005). In this PKR:eIF2 $\alpha$  structure, the region of the substrate that contains the phosphorylatable residue is disordered, suggesting that it is in a transient state. This observation illustrates that substrate specificity of some protein kinases could be contributed from regions distal to the active site and the consensus sequence around the P0 site cannot be considered as the only determinants. It also demonstrates the importance of studying the kinase-substrate recognition in the context of the full protein substrate.

#### **L. Focus of study**

Unlike many other protein kinases, the SRPK family members do not appear to require phosphorylation by upstream kinases or interaction with regulatory proteins to be active. This suggests that the cellular mechanisms for regulating SRPK1 may be unique. Moreover, SRPK1 is unique in that it only accepts serine at the phosphorylation site (P0), while most other Ser/Thr kinases will accept serine or threonine. The reason for this remarkable specificity/selectivity is unknown. Finally, it was shown that SRPK1

	P-5	P-4	P-3	P-2	P-1	P0	P+1	P+2	P+3
<b>PKA</b>			R	R/K	X	S/T	Φ		
<b>PKB</b>	R	X	R	X	X	S/T	Φ		
<b>PhK</b>			R/K	X	X	S/T	V/I/F	R	
<b>CDK2</b>						S/T	P	X	K/R

Figure 1.7 Consensus sequence motifs of kinase substrates. The consensus sequences surrounding the phosphorylatable residues (P0, in red) of several kinases are shown. X denotes any amino acid, Φ denotes hydrophobic amino acid. In all cases, the hydrophobic residue at the P+1 position is accommodated by a hydrophobic pocket formed on the kinase (Knighton et al., 1991b).

processively phosphorylates only half of the serines present in the RS domain in ASF/SF2. Our aim was to establish the molecular basis of regulation and substrate specificity of SRPK1 by structural and biochemical experiments. Also, by determining the three-dimensional structure of SRPK1 in complex with ASF/SF2, we wished to understand the molecular mechanism of the restricted and processive phosphorylation of ASF/SF2.

## **Chapter II: Materials and Methods**

## **A. Preparation of Recombinant DNA**

### **1. SRPK1 Expression Plasmids**

SRPK1 and SRPK1 $\Delta$ S (residues 1-225 and 487-655): Expression plasmids of His-tagged SRPK1 in pRSET vector (Invitrogen) and His-tagged SRPK1 $\Delta$ S in pET15b (Novagen) was provided by Dr. Xiang-Dong Fu.

SRPK1 $\Delta$ S1 (residues 1-255 and 474-655): Expression plasmid of His-tagged SRPK1 $\Delta$ S1 was prepared by two-step PCR reaction. During the first round of PCR reaction, pET19a-SRPK1 was used as template.

The N-terminal primer containing an Nde I restriction site for fragment 1-257:

5'-GCA TAG CAT ATG GAG CGG AAA GTG CTT-3'

C-terminal primer: 5'-ATT TCC AGC CGT CGC GGG AGC AGT ACT -3'.

N-terminal primers for fragment 474-655: 5'-GCA TAG CAT ATG GAG CGG AAA GTG CTT-3'

C-terminal primer containing BamHI cut site: GAA GGA TCC CTA GGA GTT AAG CCA AGG-3'.

The resulting DNA products for both fragments were then used as template and the two primers containing the restriction cut sites were also used in the second round of PCR reaction. The resulting full length fragment was first gel purified, then double digested with NdeI and BamHI restriction enzymes (NEB), and gel purified again. The fragment was then ligated in pET15b vector at the corresponding cut sites.

SRPK1 $\Delta$ NS1 (residues 42-255 and 474-655): The plasmid of pET15b-SRPK1 $\Delta$ S1 was used as template.

N-terminal primer containing NdeI cut site: 5'-GCA TAC CAT ATG CCA GAGCAG GAA GAG-3'

C-terminal primer containing BamHI cut site was the same used for SRPK1 $\Delta$ S1.

The PCR reaction products was first double digested with the restriction enzymes, gel purified and ligated into pET15b.

SRPK1 $\Delta$ NS (residues 42-225 and 487-655): The preparation of this plasmid was similar to SRPK1 $\Delta$ NS1 except using pET15b-SRPK1 $\Delta$ S as template.

SRPK1 $\Delta$ N1S1 (residues 58 to 255 and 474-655): The preparation of this plasmid was similar to SRPK1 $\Delta$ NS1 except using N-terminal primer: 5'-GCA TAG CAT ATG GAT CCT AAT GAT TAT TGT-3'.

## **2. Mutagenesis of SRPK1**

Activation segment mutants (M1-M7): pET15b-SRPK1 $\Delta$ NS1 mutants containing 7 mutations surrounding the activation segment were all generated using single step PCR by introducing one or two mutation(s) at a time. The DNA oligonucleotide primers listed below were used to amplify the mutated gene construct. The PCR products were first restriction digested by DpnI

enzymes (NEB) to remove the template. The nicked products were then transform into JM 109 *E.coli* cells. Plasmids isolated from transformants were sequenced to verify the presence of mutations.

R208G\_top: 5'-ACC AAG TGC GGT ATC ATC-3'

R208G\_bottom: 5'-GAT GAT ACC GCA CTT GGT-3'

E510N\_top: 5'-CAT TTC ACT AAC GAT ATT CAA ACA AGG-3'

E510N\_bottom: 5'-CCT TGT TTG AAT ATC GTT AGT GAA ATG-3'

D511S\_top: 5'-CAT TTC ACT GAA TCG ATT CAA ACA AGG-3'

D511S\_bottom: 5'-CCT TGT TTG AAT CGA TTC AGT GAA ATG-3'

Y528W, N529G\_top: 5'-GGA TCT GGC TGG GGT ACC CCT GCT-3'

Y528W, N529G\_bottom: 5'-AGC AGG GGT ACC CCA GCC AGA TCC-3'

V504Y, H505D\_top: 5'-GCT TGT TGG TAC GAC AAA CAT TTC-3'

V504Y, H505D\_bottom: 5'-GAA ATG TTT GTC GTA CCA ACA AGC-3'

Docking groove mutant: pET15b-SRPK1 $\Delta$ NS1 mutant containing 4 mutations at the docking groove was generated as described for the activation segment mutant. The following primers were used:

D548A\_top: 5'-ACT CGA GAT GAA GCG CAC ATT GCA TTG-3'

D548A\_bottom: 5'-TTC AAA CAA ATA GGC ACC TGT GGC CAG-3'

D564A\_top: 5'-ACT CGA GAT GAA GCG CAC ATT GCA TTG-3'

D564A\_bottom: 5'-CAA TGC AAT GTG CGC TTC ATC TCG AGT-3'

E571A\_top: 5'-GCA TTG ATC ATA GCG CTT CTG GGG AAG-3'

E571A\_bottom: 5'-CTT CCC CAG AAG CGC TAT GAT CAA TGC-3'



K615A\_top: 5'-GTT CTA GTG GAG GCA TAT GAG TGG TCG-3'

K615A\_bottom: 5'-CGA CCA CTC ATA TGC CTC CAC TAG AAC-3'

Spacer region mutant: pET15b-SRPK1 $\Delta$ NS1 mutant containing 2 mutations at the interface of the spacer helix  $\alpha$ S1 and small lobe of the kinase was generated using the following primers.

V167A\_top: 5'-ATG GTA TTT GAA GCT TTG GGG CAT-3'

V167A\_bottom: 5'-ATG CCC CAA AGC TTC AAA TAC CAT-3'

I228A\_top: 5'-AGCAGT ACG CTC GGA GGC TGG CT-3'

I228A\_bottom: AGC CAG CCT CCG AGC GTA CTG CT-3'

### 3. ASF/SF2 Expression Plasmids

ASF/SF2: His-tagged ASF/SF2 in pET19b (Novagen) was provided by Dr. Xiang-Dong Fu.

ASF/SF2 $\Delta$ RS2 (residues 1-219): GST-tagged ASF/SF2 $\Delta$ RS2 in pGEX4T2 (Amersham) was a gift from Sutapa Chakrabarti.

Different C-terminus truncated constructs of ASF/SF2: The following pET15b-ASF/SF2 constructs with different C-terminal truncation were generated by PCR reaction using pET19a-ASF/SF2 as template, N-terminal primer containing NdeI restriction site:

5'-GAA CTA CAT ATG TCG GGA GGT GGT GTG-3'

and their corresponding C-terminal primers containing BamHI restriction site:

ASF/SF2(1RS)\_C':

5'-GAA GAC GGA TCC CTA GCG GCT ACG GCC ATA-3'

ASF/SF2 $\Delta$ RS2\_C':

5'-ATT GGA TCC CTA GCT TCT GCT ACG GCT TCT-3'

ASF/SF2(BD) (residues 105-219): The plasmid of pET15b-ASF/SF2(BD) was prepared using N-terminal primer: 5'-GAA TTC CAT ATG GGC GGA GCT CCC CGA GGT-3' and C-terminal primer ASF/SF2 $\Delta$ RS2\_C'.

## **B. Expression and Purification of Recombinant Protein**

His-SRPK1: *E.coli* BL21(DE3) cells transformed with pRSET-SRPK1 were grown in 2 L of LB/ampicillin (200  $\mu$ g/mL) broth to an O.D.<sub>600</sub> of 0.2 at 37°C and induced with 0.1 mM isopropylthiogalactoside (IPTG) for 16 hours at room temperature (~25 °C). Cells were harvested by centrifugation and resuspended in 80 ml of lysis buffer containing 20 mM CAPSO pH 9.5, 100 mM KCl, 20% glycerol, 1 mM PMSF, 200  $\mu$ L of Sigma Protease Inhibitor Cocktail (Sigma/Aldrich). The resuspended cells were lysed at 4 °C by sonication at 1 minute burst intervals. The solution was centrifuged to remove insoluble fraction. The supernatant was loaded onto a fast Q-sepharose column. The column was washed and eluted with a 100 mM – 1 M KCl gradient. Peak

fractions were identified by SDS-PAGE. 5 mM imidazole was added to the collected fractions and the mixture was loaded onto a Ni-NTA affinity column. The column was washed with 40 mL of buffer containing 20 mM CAPSO pH 9.5, 500 mM KCl, 20% glycerol and 5 mM imidazole, followed by 40 mL of wash buffer consisted of the same ingredients except 40 mM imidazole. The protein was then eluted with 30 mL of the same buffer containing 250 mM imidazole. The elution was dialysed against 2 L of buffer containing 20 mM CAPSO pH 9.5, 500 mM KCl, 5% glycerol and 10 mM  $\beta$ -ME overnight at 4°C to remove the imidazole. The dialysate was concentrated using AMICON-50 concentrator and loaded onto a Superdex-200 size-exclusion column (Amersham) equilibrated with 20 mM CAPSO pH 9.5, 500 mM KCl, 5% glycerol and 1 mM dithiothreitol (DTT). Pooled peak fractions were concentrated to ~5 mg/mL, aliquoted and flash frozen.

His-SRPK1 $\Delta$ S: *E.coli* BL21(DE3) cells were transformed with pET15b-SRPK1 $\Delta$ S. Large scale culture were grown in 2 L of LB with 200  $\mu$ g/mL ampicillin to an O.D.<sub>600</sub> of 0.2 at 37°C and induced with 0.1 mM IPTG with shaking overnight at room temperature. Pelleted cells were resuspended in 80 ml of lysis buffer containing 20 mM CAPSO pH 9.5, 100 mM KCl, 20% glycerol, 1 mM PMSF, 200  $\mu$ L of Protease Inhibitor Cocktail. The resuspended cells were lysed at 4°C by sonication six times at 1 minute burst intervals. The solution was centrifuged to remove insoluble fraction. The supernatant was

loaded onto a fast Q-sepharose column. The column was washed and eluted with a 50 mM – 1 M KCl gradient. 5 mM imidazole was added to the collected fractions and the concentration of KCl in the mixture was increased to 1 M. The mixture was loaded onto a Ni-NTA affinity column and washed with 40 mL of buffer containing 20 mM CAPSO pH 9.5, 1 M KCl, 20% glycerol and 5 mM imidazole, followed by 40 mL of wash buffer (40 mM imidazole). The protein was then eluted with 30 mL of the same buffer containing 250 mM imidazole. The protein was dialysed overnight against 2 L of buffer containing 20 mM CAPSO pH 9.5, 1 M KCl, 5% glycerol and 10 mM  $\beta$ -ME. The protein was concentrated and loaded onto a Superdex-75 size-exclusion column (Amersham) equilibrated with 20 mM CAPSO pH 9.5, 1 M KCl, 5% glycerol and 1 mM dithiothreitol (DTT). Pooled peak fractions were concentrated to ~7 mg/mL, aliquoted and stored at -80 °C.

His-SRPK1 $\Delta$ S1: Protein expression was similar to His-SRPK1 $\Delta$ S. Cell pellets were lysed by sonication in 80 mL of 20 mM MES pH 6.5, 50 mM KCl, 20% glycerol, 1 mM PMSF, 200  $\mu$ L of Protease Inhibitor Cocktail. The soluble fraction was loaded onto the Q-sepharose column and washed with 150 mL of lysis buffer. The flow through sample was combined with the wash and the concentration of KCl was brought up to 500 mM followed by an addition of 5 mM imidazole. The mixture was loaded on to a Ni<sup>2+</sup> affinity column washed with 40 mL of buffer containing 20 mM MES pH 6.5, 500 mM KCl, 20%

glycerol and 5 mM imidazole, followed by 40 mL of wash buffer with 40 mM imidazole. The protein was then eluted with 30 mL of buffer containing 250 mM imidazole. The elution was dialysed overnight in two changes of 1 L buffer containing 20 mM MES pH 6.5, 500 mM KCl, 5% glycerol and 10 mM  $\beta$ -ME. The protein was concentrated by AMICON-50 concentrator to ~ 6 mg/mL. The His-tag of the recombinant protein was then removed by adding 0.2 NIH units (1 cleavage unit) of thrombin (Sigma/Aldrich) per mg of His-SRPK1 $\Delta$ S1. The reaction was allowed to perform at room temperature for 16 hours and quenched by adding 1 mM EDTA. The protein was then filtered and loaded on to a Superdex 75 size exclusion column equilibrated with 20 mM MES pH 6.5, 500 mM KCl, 5% glycerol and 1 mM DTT. The peak eluted fractions were combined, concentrated to ~10 mg/mL, aliquoted and flash frozen.

His-SRPK1 $\Delta$ NS1: Protein expression and cell lysis were the same as His-SRPK1 $\Delta$ S1 (same lysis buffer). The solution fraction was loaded on to Q- and SP-sepharose columns connected in tandem. The column was washed with 200 mL of lysis buffer and eluted with a 100 mM – 500 mM KCl gradient. The peak fractions were loaded onto the Ni<sup>2+</sup> affinity, washed and eluted as described for SRPK1 $\Delta$ S1. The protein was dialysed overnight in 2 X 1L of buffer consisted of 20 mM MES pH 6.5, 250 mM KCl, 5% glycerol and 10 mM  $\beta$ -ME to remove the imidazole, concentrated and cleaved with thrombin as described for His-SRPK1 $\Delta$ S1. Cleaved protein was loaded onto a Superdex

75 size exclusion column equilibrated with 20 mM MES pH 6.5, 250 mM KCl, 5% glycerol and 1 mM DTT. The pooled peak fractions were concentrated to ~12 mg/mL, aliquoted and stored in -80 °C.

His-SRPK1 $\Delta$ N1S1 and different His-SRPK1 $\Delta$ NS1 mutants: Protein expression and purification were identical to His-SRPK1 $\Delta$ NS1.

GST-ASF/SF2 $\Delta$ RS2: BL21(DE3) *E.coli* cells were transformed with pGEX4T2-ASF/SF2 $\Delta$ RS2. Large scale culture were grown in 2L of LB with 200  $\mu$ g/mL ampicillin to an O.D.<sub>600</sub> of 0.4 at 37°C and induced with 0.1 mM IPTG with shaking overnight at room temperature. Pelleted cells were resuspended in 80 ml of lysis buffer containing 20 mM Tris-HCl pH 7.5, 100 mM KCl, 20% glycerol, 1 mM DTT, 1 mM PMSF, 200  $\mu$ L of Protease Inhibitor Cocktail. The resuspended cells were lysed at 4 °C by sonication six times at 1 minute burst intervals. The solution was centrifuged to remove insoluble fraction. The supernatant was loaded onto a fast Q-sepharose column and washed with 150 mL of lysis buffer. The flow through and wash from the Q-column were combined and loaded onto a GST-column (Stratagene) overnight at 4 °C. Bound protein was washed three times with 20 mL of buffer containing 20 mM Tris pH 7.5, 100 mM KCl, 20% glycerol and 1 mM DTT, and eluted with buffer containing 10 mM glutathione. Fractions were analyzed by SDS-PAGE.

His-ASF/SF2(1RS), His-ASF/SF2 $\Delta$ RS2 and His-ASF/SF2(BD) (residues 105-219): BL21(DE3) *E. coli* cells transformed with corresponding pET15b-ASF/SF2 construct were grown in 2 L LB/Amp to O.D.<sub>600</sub> of 0.3 and induced with 0.1mM IPTG and shaken at 22 °C overnight. Cells were harvested and resuspended in lysis buffer (20 mM Tris-HCl pH 7.5, 500 mM KCl, 6 M Urea, and 5 mM imidazole). Cells were lysed by sonication and the extract was centrifuged at 15,000 rpm for 20 min. The supernatant was loaded on an Ni-NTA column and washed with lysis buffer containing 40 mM imidazole. Protein was eluted with lysis buffer containing 250 mM imidazole. Purified protein was aliquoted and stored at -80 °C.

### **C. Limited Proteolysis**

To locate flexible regions in the SRPK $\Delta$ S construct, limited proteolysis reactions containing 100  $\mu$ g SRPK1 $\Delta$ S were initiated with either 0.1  $\mu$ g chymotrypsin in the presence of 10 mM CaCl<sub>2</sub> or 0.1  $\mu$ g trypsin. Reactions were performed at room temperature and aliquots of the reaction mix were removed at intervals from 0 to 180 min from the start of the reaction. Reaction was quenched by boiling the aliquot in the presence of SDS buffer for 1 min. The samples were then analyzed by SDS-PAGE. For SRPK1 $\Delta$ S1, only chymotrypsin was used.

### **D. Measuring Kinase Activity**

2 nM of each different kinase construct was incubated with 2  $\mu$ M SR- $\text{I}\kappa\text{B}\alpha$  (a chimeric protein with the C-terminal 13 residues of Npl3 fused to  $\text{I}\kappa\text{B}\alpha$ ) or 2  $\mu$ M GST-ASF/SF2 ( $\Delta$ RS2), 200 mM ATP and 800  $\mu\text{Ci}^{-1}$   $\gamma$ - $^{32}\text{P}$ -ATP in 50 mM Tris-HCl pH 7.5, 10 mM  $\text{MgCl}_2$ , 1 mg/mL BSA and 1 mM dithiothreitol at 22  $^{\circ}\text{C}$ . Reactions were quenched by addition of 4X SDS buffer and subsequent boiling. Reaction mixes were run on a 12.5% SDS gel. The gel was dried and exposed to autoradiography film.

### **E. Testing Kinase Stability**

To test the kinase stability in different ionic strength buffers, SRPK1 $\Delta$ S or SRPK1 $\Delta$ S1 was incubated at room temperature in the presence of 20 mM Tris-HCl pH 7.5, 10% glycerol, 1 mM DTT, 1 mg/mL BSA, and either 30 mM or 500 mM KCl. The kinase reaction was started by adding the enzyme to the reaction mix at intervals from 1 to 60 min from the start of incubation. The reaction mix contained 8  $\mu$ M GST-ASF/SF2 ( $\Delta$ RS2), 200 mM ATP and 800  $\mu\text{Ci}^{-1}$   $\gamma$ - $^{32}\text{P}$ -ATP in 50 mM Tris-HCl pH 7.5, 10 mM  $\text{MgCl}_2$ , 1 mg/mL BSA and 1 mM dithiothreitol. All reactions were quenched after 10 min by addition of 4X SDS buffer and boiled for 30 seconds. Reaction mixes were run on a 12.5% SDS gel. The gel was dried and exposed to autoradiography film.

### **F. Start-delay-trap Assay**



Wild type or mutant SRPK1 $\Delta$ NS1 (0.6  $\mu$ M) were allowed to pre-incubate with 2  $\mu$ M GST-ASF/SF2 $\Delta$ RS2 in 50 mM Tris-Cl pH 7.5, 10 mM MgCl<sub>2</sub>, 1 mg/ml BSA and 1 mM dithiothreitol at 0°C for 2 minutes and at 22°C for 1 minute. 200 mM ATP and 800  $\mu$ Ci<sup>-1</sup>  $\gamma$ -<sup>32</sup>P-ATP (start) were added after the incubation to initiate the reaction. 20mM inhibitor peptide (trap) was then added simultaneously or 1 min after initiation of the reaction (delay). Reactions were quenched with 4X SDS buffer. The reaction was resolved on a 12.5% SDS gel, and the dried gel was exposed to a phosphorimager plate (Molecular Dynamics). The intensity of each band was quantified by using ImageQuant software (Molecular Dynamics, Sunnyvale, CA, USA).

### **G. Formation of SRPK1 $\Delta$ NS1 Apo-enzyme Crystal and Data Collection**

Small crystals were obtained by the hanging drop vapor diffusion at 18°C using the Hampton Screen (Hampton Research). The drop contained 1  $\mu$ l SRPK1 $\Delta$ NS1 protein at 12 mg/ml mixed with 1ul of reservoir solution. Optimal condition for crystallization was found to be 100 mM sodium citrate (pH 5.6), 200 mM ammonium acetate, and 15% PEG 3350 in the reservoir. All crystals were cryo-protected by sequential addition of 10%, 15% and finally 20% ethylene glycol (v/v) in presence of the reservoir solution and flash frozen in liquid nitrogen prior to data collection. SRPK1 $\Delta$ NS1 crystallized in P6<sub>5</sub>22 space group (a = b = 75.11Å, c = 313.33Å,  $\alpha$  =  $\beta$  = 90 and  $\gamma$  = 120) with one molecule per asymmetric unit.

X-ray diffraction data of apo-SRPK $\Delta$ NS1 crystals were collected either at the UCSD home source X-ray facility or using a SBC2 CCD detector at SBC-CAT synchrotron beamline ID-19 of the Advanced Photon Source (APS) at Argonne National Laboratory. All X-ray data was processed using program HKL2000 (Otwinowski and Minor, 1997).

## **H. Formation of Seleno-methionine Derivative Crystals of SRPK1 $\Delta$ NS1**

### **1. Purification of Seleno-methionine Derivative SRPK1 $\Delta$ NS1**

A minimal media was prepared by mixing 7.5 mM AmSO<sub>4</sub>, 8.5 mM NaCl, 55 mM KH<sub>2</sub>PO<sub>4</sub>, 100 mM K<sub>2</sub>HPO<sub>4</sub>, 1 mM MgSO<sub>4</sub>, 20 mM glucose, 1 mg/L CaCl<sub>2</sub>, 1  $\mu$ g/L ZnSO<sub>4</sub>, 1  $\mu$ g/L MnCl<sub>2</sub>, 1 mg/L FeSO<sub>4</sub>, 1  $\mu$ g/L MoO<sub>3</sub>, 20 mg/L d-biotin, 1 mg/L thiamine and 1  $\mu$ g/L CuCl<sub>2</sub> in 1 L of autoclaved water, adjusted to pH 7.4 with addition of NaOH and sterile filtered (buffer 1). Then buffer 2 containing 137 mM NaCl, 2.5 mM KCl, 10 mM Na<sub>2</sub>HPO<sub>4</sub> and 1.76 mM KH<sub>2</sub>PO<sub>4</sub> was adjusted to pH 7.0 and was used to prepare 50 mL of amino acid buffer A, 25 mL of amino acid buffer B, 20 mL of amino acid buffer C and 35 mL of nucleotide base buffer. Amino acid buffer A contained 300 mg alanine, 100 mg cysteine, 200 mg aspartate, 300 mg glutamate, 200 mg glycine, 200 mg histidine, 200 mg isoleucine, 300 mg lysine, 400 mg leucine, 200 mg asparagine, 200 mg proline, 300 mg glutamine, 200 mg threonine and 300 mg valine. Amino acid buffer B was prepared by dissolving 200 mg phenylalanine, 100 mg tryptophan and 100 mg tyrosine in buffer 2 with the addition of 1 M

NaOH. Amino acid buffer C was consisted of 4 g arginine and 4 g serine. Nucleotide base buffer was prepared by dissolving 100 mg each of adenine, guanine, cytosine, thymine and uracil in 35 mL of buffer 2 with the addition of NaOH.

*E.coli* BL21(DE3) cells were transformed with pET15b-SRPK1 $\Delta$ NS1. Small scale culture was grown in 5 mL of LB/ampicillin (200  $\mu$ g/mL) overnight. Immediate before growing large scale culture, all solutions including buffer 1, amino acid buffers A, B and C, nucleotide base buffer, 200  $\mu$ g/mL ampicillin, together with 200 mg of seleno-methionine and 20 mL of 100 X basal medium eagle vitamin solution (Gibco/BRL) were mixed together, brought to 2 L of total volume with the addition of autoclaved water and adjusted to pH 7.4. The small scale culture was pelleted to remove the LB. The cell pellet was then washed three times with 2 mL of minimal media and resuspended in 5 mL of minimal media. The resuspended small culture was then added into the 2 L minimal media and grew to an O.D.<sub>600</sub> of 0.2 at 37°C. The culture was induced with 0.1 mM IPTG with shaking overnight at room temperature. Cell lysis and protein purification were the same as that described for His-SRPK1 $\Delta$ NS1, except all steps were performed in the presence of 1 mM DTT.

## **2. Crystallization and Data Collection**

The derivative crystals were prepared in a similar way as the apo-crystals except the hanging drop vapor diffusion experiment was performed

using seleno-methionine derivative SRPK1 $\Delta$ NS1 protein. Multi-wavelength anomalous dispersion (MAD) data was collected using an ADSC Quantum-4 CCD area detector at ALS synchrotron beamline 5.0.2. of Advanced Light Source (ALS) at Lawrence Berkeley National Laboratory. The Se edge wavelengths were selected from the fluorescence spectrum of the crystal. Independent data sets were collected on the same crystal for 120° at 0.5° oscillations at three different wavelengths 0.97996 Å, 0.97976 Å and 0.9649 Å respectively. The Friedel pairs were recorded by first collecting a small wedge of diffraction data and then rotated the crystal by 180° to collect the same wedge of data. The space group of the tested derivative crystal is the same as the apo-crystal with similar unit cell dimensions ( $a = b = 74.84$  Å,  $c = 312.94$  Å).

#### **H. Structure Solution of SRPK1 Crystal**

Initial phases of apo-SRPK1 $\Delta$ NS1 were obtained by MAD data from a seleno-methionine derivative crystal. The positions of all 5 selenium sites in the asymmetric unit were located by anomalous Patterson and difference Fourier methods using the CNS suite (Brunger et al., 1998). Selenium positions were refined and MAD phases were also calculated using CNS. The initial electron density map was significantly improved after solvent flattening and led to a clear and interpretable map. The MAD electron density map was used to build an initial model using the program Xtalview (McRee, 1999). The model was refined using the program CNS and improved by rebuilding after

recalculation of electron density using weighted combinations of model and MAD phases. The  $R_{\text{cryst}}$  and  $R_{\text{free}}$  for the apo-SRPK1 $\Delta$ NS1 model were 19.72% and 21.76% respectively, for data from 30-1.73 Å. The final model includes residues 63-64, 67-236 and 475-655, 5 molecules of ethylene glycol and 298 water molecules. Analysis of the model by PROCHECK (Laskowski et al., 1993) indicated that over 99.4% of the residues are in most favored or additionally allowed region of the Ramachandaran plot. Q513 was the only residue in the disallowed region and T212 was in a generously allowed region. The coordinates of the final model was deposited into the Protein Data Bank under the accession code 1WAK.

#### **J. Formation of SRPK1 $\Delta$ NS1/ADP/Peptide Ternary Complex Crystal and Data Collection**

Crystals of apo-SRPK1 $\Delta$ NS1 were transferred to a microbridge containing 100 mM sodium citrate (pH 5.6), 200 mM ammonium acetate, 15% PEG 3350, 5 mM 9mer peptide, 5 mM ADP and 10 mM MgCl<sub>2</sub>. Microbridges were sealed in well containing 1 ml of stabilizing solution. The crystals were soaked for 24 hours at 18°C. All crystals were then cryo-protected by sequential addition of 10%, 15% and finally 20% ethylene glycol (v/v) in presence of the soaking condition and flash frozen in liquid nitrogen prior to data collection. X-ray diffraction data of complex crystals were collected using a SBC2 CCD detector at SBC-CAT synchrotron beamline ID-19 of the

Advanced Photo Source (APS) at Argonne National Laboratory. The space group of the complex crystal was identified to  $P6_522$  with  $a = b = 78.68 \text{ \AA}$ ,  $c = 310.54 \text{ \AA}$  and one molecule in the asymmetric unit.

### **K. Structure Solution of SRPK1 $\Delta$ NS1/Peptide/ADP Ternary Complex**

The structure of SRPK1 $\Delta$ NS1/peptide/ADP ternary complex was initially solved by molecular replacement using the program AMoRe (Navaza, 1994) and the structure of apo-SRPK1 $\Delta$ NS1 as the search model, yielding clear rotation function and translation peaks. After rigid body refinement using CNS (Brunger et al., 1998),  $F_o - F_c$  electron density map calculated with the model revealed large positive peaks and readily interpretable density for the peptide and ADP. After model building of the peptide, the structure was refined with several cycles of manual refitting and refinements using CNS. The  $R_{\text{cryst}}$  and  $R_{\text{free}}$  for the SRPK1 $\Delta$ NS1/peptide/ADP model were 22.84% and 24.71% respectively, for data from 30-2.4  $\text{\AA}$ . The final model of the ternary complex includes residues 67-237 and 477-655 of the kinase, residues 3-9 of the peptide, 1 molecule of ADP and 1 molecule of acetate. The coordinates of the final model was deposited into the Protein Data Bank under the accession code 1WBP.

### **L. Formation of SRPK1 $\Delta$ NS1:ASF/SF2(BD):AMP-PNP Complex Crystal**

#### **1. Recombinant SRPK1:ASF/SF2 Complex Formation**

His-tag SRPK1 and his-tag ASF/SF2 were first purified separately. 12 mg of SRPK1 and 16 mg of ASF/SF2 were mixed together in 100 mL buffer containing 500 mM KCl, 20 mM Tris-HCl pH 7.5, 20 % glycerol, 6 M urea and 3 mM DTT. The mixture was mixed by stirring for 30 mins at 4°C. The mixture was then dialysed for 6 hrs against 2 L of buffer containing 500 mM KCl, 20 mM Tris-HCl pH 7.5, 15% glycerol and 3 mM DTT, followed by another 6 hrs dialysis in buffer containing 500 mM KCl, 20 mM Tris-HCl pH 7.5, 10% glycerol and 3 mM DTT. A final dialysis was performed in buffer containing 500 mM KCl, 20 mM CAPSO pH 9.5, 5% glycerol and 1.5 mM DTT. The dialysate was then concentrated by AMICON-50 concentrator. His-tag of the refolded protein was then removed by adding 0.2 NIH units (1 cleavage unit) of thrombin (Sigma/Aldrich) per mg of complex. The reaction was allowed to perform at room temperature for 16 hours and quenched by adding 1 mM EDTA. The untagged complex was then loaded on the Superdex-200 size exclusion column equilibrated in 500 mM KCl, 20 mM CAPSO pH 9.5, 5% glycerol and 1 mM DTT. The pooled peak fractions were concentrated to ~11 mg/mL, aliquoted and stored in -80 °C.

## **2. Crystallization of SRPK1 $\Delta$ N1S1:ASF/SF2(BD):AMP-PNP Complex**

Needle shape crystals were obtained by the hanging drop vapor diffusion at 18°C using the Hampton Screen (Hampton Research). The protein was first prepared by mixing 24  $\mu$ l SRPK1 $\Delta$ N1S1:ASF/SF2(BD) protein at 6

mg/mL with 4 mM AMP-PNP and 8 mM MgCl<sub>2</sub>. 1 μL of the protein was then mixed with 1 μL of reservoir solution. Optimal condition for crystallization was found to be 100 mM sodium citrate (pH 5.6), 200 mM sodium acetate, 5% MeOH or EtOH, and 8% PEG 5000MME in the reservoir. In order to improve the diffraction quality, the crystals were then dehydrated and cryo-protected by dialysis overnight against 100 mM sodium citrate pH 5.6, 200 mM sodium acetate, 5% MeOH of EtOH, 15% PEG 5000MME and 25% (V/V) ethylene glycol. All dialysed crystals were flash frozen in liquid nitrogen prior to data collection.

#### **M. Data Collection of SRPK1ΔN1S1:ASF/SF2(BD):AMP-PNP Complex Crystal**

X-ray diffraction data of the ternary complex crystals were collected using a MAR CCD detector at GM/CA-CAT synchrotron beamline ID-23 of the Advanced Photon Source (APS) at Argonne National Laboratory. Data set of complete 180° was collected at 1° oscillations. The complex crystals belong to orthorhombic space group I222 with unit cell dimensions  $a = 57.406 \text{ \AA}$ ,  $b = 117.525 \text{ \AA}$ ,  $c = 193.554 \text{ \AA}$ ,  $a = b = c = 90^\circ$ . Each asymmetric unit contains one molecule of complex.

#### **N. Structure Solution of SRPK1ΔN1S1:ASF/SF2(BD):AMP-PNP Complex**



The structure of SRPK1 $\Delta$ N1S1:ASF/SF2(BD):AMP-PNP ternary complex was solved by molecular replacement using the program AMoRe and the structure of apo-SRPK1 $\Delta$ NS1 as the search model, yielding clear rotation function and translation peaks. After rigid body refinement using CNS,  $F_o-F_c$  electron density map calculated with the model revealed large positive peaks and readily interpretable density for the ASF/SF2 and AMP-PNP. A poly-alanine backbone of ASF/SF2 was manually built by using the NMR solution structure of ASF/SF2 (RRM2) (PDB accession code 1X4C) as a guide. The two lobes of the SRPK1 and the poly-alanine backbone of ASF/SF2 were allowed to move independently during the first round of rigid body refinement. Side chains of ASF/SF2 and a molecule of AMP-PNP were then fitted in and the complex structure was refined with several cycles of manual refitting and refinements using CNS. The current  $R_{\text{cryst}}$  and  $R_{\text{free}}$  for the SRPK1 $\Delta$ N1S1:ASF/SF2(BD):AMP-PNP model are 29.23% and 38.88% respectively, for data from 50-2.90 Å. The current model of the ternary complex includes residues 68-237 and 477-655 of the kinase, residues 122-194 of the ASF/SF2, a 7 residues peptide and 1 molecule of AMP-PNP.

## **Chapter III: Identification of a Crystallizable Construct of SRPK1 and its Crystallization**

## A. Introduction

Knowledge of the precise molecular structure of biological macromolecules is critical for the study of their cellular functions. X-ray crystallography and nuclear magnetic resonance (NMR) spectroscopy are the two most widely used techniques for structural determination of biological macromolecules at atomic level. While NMR can provide information about the three-dimensional structure and dynamics of the molecules, this method is not suitable for molecules with higher molecular weight. Therefore, X-ray crystallography is usually preferred to study high molecular weight macromolecules. One of the major obstacles in protein X-ray crystallography is obtaining well-diffracting crystals of the target macromolecules. In order to do so, one has to purify the target protein to homogeneity and crystallize it from the aqueous solution. However, many regions within proteins could be flexible and disordered, possibly leading to failure in crystallization of the target protein. To overcome this, ligands or other proteins can be introduced to interact with such regions and stabilize them. Alternatively, flexible regions with no biological significance can be removed to facilitate crystallization.

The protein kinase superfamily share a common catalytic scaffold that can be divided into 12 conserved subdomains (I-VIA, VIB, VII-XI) defined by Hanks et. al (Hanks and Quinn, 1991). The protein kinase fold consists of two lobes: the smaller N-terminal lobe and the larger C-terminal lobe. The small lobe includes regions I-IV and is composed of a five-stranded  $\beta$ -sheet and an  $\alpha$ -

helix, while the large lobe consists of regions VIA-XI and is prevalingly  $\alpha$ -helical.

SRPKs are unusual members of the CMGC group of kinases as their kinase domain cores are bifurcated between subdomains VI and VII by spacers of 250-300 residues in length. The spacer region has been suggested to be involved in subcellular localization of SRPKs by retaining the kinase in the cytoplasm (Ding et al., 2006; Siebel et al., 1999; Wang et al., 1998a). However, their deletion does not affect the kinases' *in vitro* activities and are thus considered to function independently of the kinase domain. SRPKs from different species also contain non-homologous segments N-terminal to the kinase domain core that ranges from 40 to 250 residues in length (Figure 3.1). Some SRPKs also contain non-homologous C-terminal extensions. Indeed, the X-ray crystal structure of Sky1p (a SRPK in yeast), solved earlier in our lab, shows that the removal of the non-homologous N-terminus and the spacer region has no effect on the kinase's *in vitro* catalytic activity and overall fold, while the C-terminal extension is essential for the kinase function (Nolen et al., 2001). In this chapter, we describe the rationale behind and results for the identification of an active fragment of human SRPK1. The proposed constructs for crystallization were designed based on observations made on the Sky1p crystal structure, sequence alignment analysis of SRPK1 among different species and limited proteolysis studies of SRPK1. We also report the crystallization conditions and X-ray diffraction data collection of apo-SRPK1 crystals.

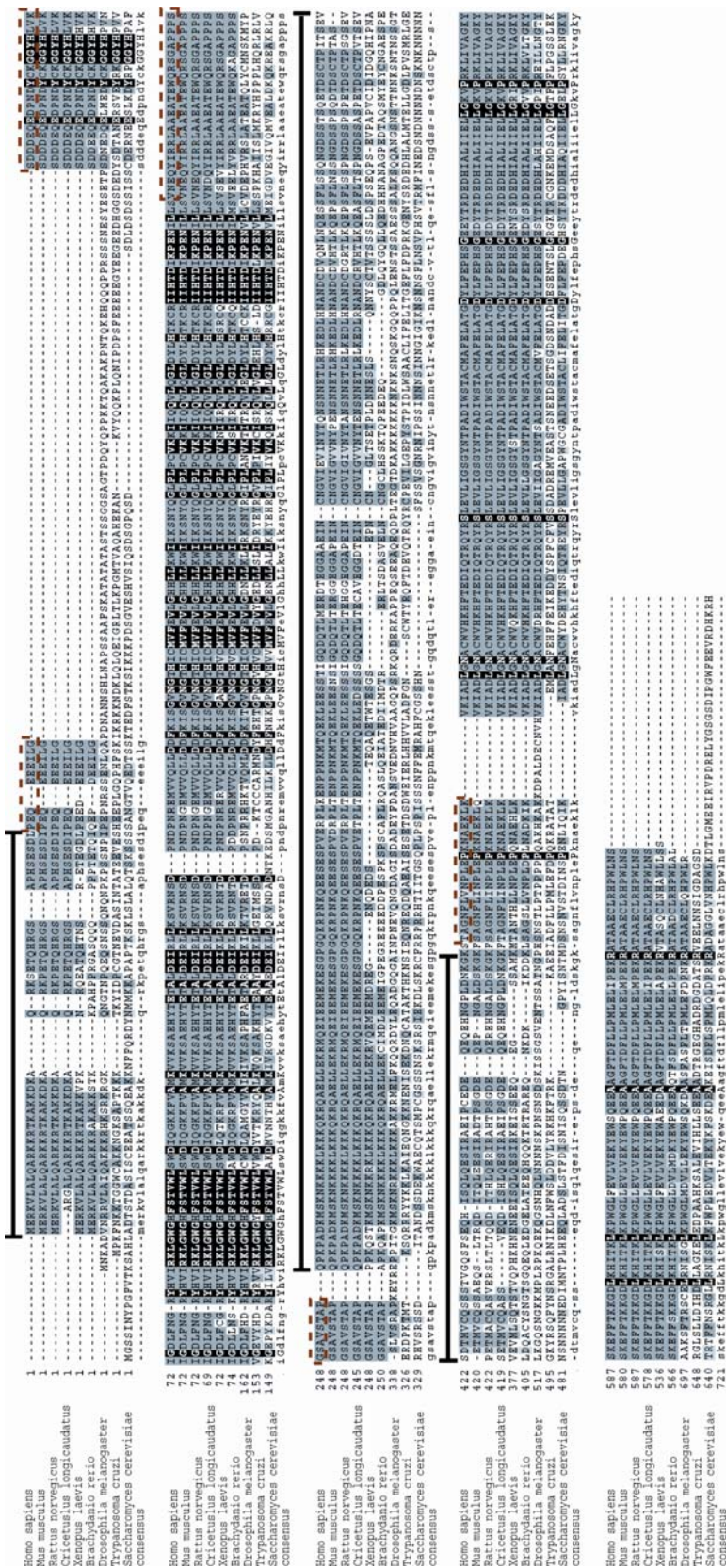


Figure 3.1 Sequence alignment of SRPK1s. SRPK1 sequences from 9 different species were compared. While the kinase domains are highly homologous, the spacer region, N- and C-terminal extension are not conserved. In the crystallizable fragment of SRPK1 (SRPK1ΔNS1), we deleted 217 residues (256-473) of the spacer domain and 41 N-terminal residues of human SRPK1 that bears no sequence homology with other SRPK1. The deleted regions are indicated with a black bar on top. This deletion still retains three small segments outside the conserved kinase core because of their homology among different species: one at the N-terminus and two in the spacer domain (highlighted with red dotted box). Residues that are identical are boxed in black, similar residues are boxed in grey and different residues in light grey.

## B. Results

### 1. Identification of an Active Fragment of SRPK1

We first attempted to purify and crystallize a full-length version of SRPK1 for crystallization studies. A construct of SRPK1(FL) containing a hexa-histidine tag at the N-terminus was obtained from our collaborator, Dr. Fu's lab, and was purified from *E.coli* using multiple chromatographic steps. The full-length proteins were susceptible to proteolytic degradation and yielded multiple degraded fragments as shown in SDS-PAGE analysis (Figure 3.2). The purified full-length protein was estimated to be ~80% pure and crystallization trials were attempted. However, extensive screening of crystallization conditions did not produce suitable crystals for diffraction studies.

The crystal structure of an active fragment of Sky1p was solved previously in our lab. The complete spacer region and 137 amino acids from the divergent N-terminus were removed from Sky1p, resulting in the crystallized construct Sky1p $\Delta$ NS. The model of Sky1p reveals that the spacer region lies between  $\beta$ 7 and  $\beta$ 8 within the large lobe and its deletion would not affect the overall fold and integrity of the kinase core domain. Based on this success and high sequence homology of the kinase domain between the yeast and human SRPKs (Figure 3.1), we decided to crystallize a fragment of SRPK1 with similar truncations that were made in Sky1p. The first major modification to be considered was the removal of the complete spacer region. Using the programs PredictProtein (<http://cubic.bioc.columbia.edu/pp/>), nnPredict (<http://www.cmpharm.ucsf.edu/~nomi/nnpredict.html>) and SSpro

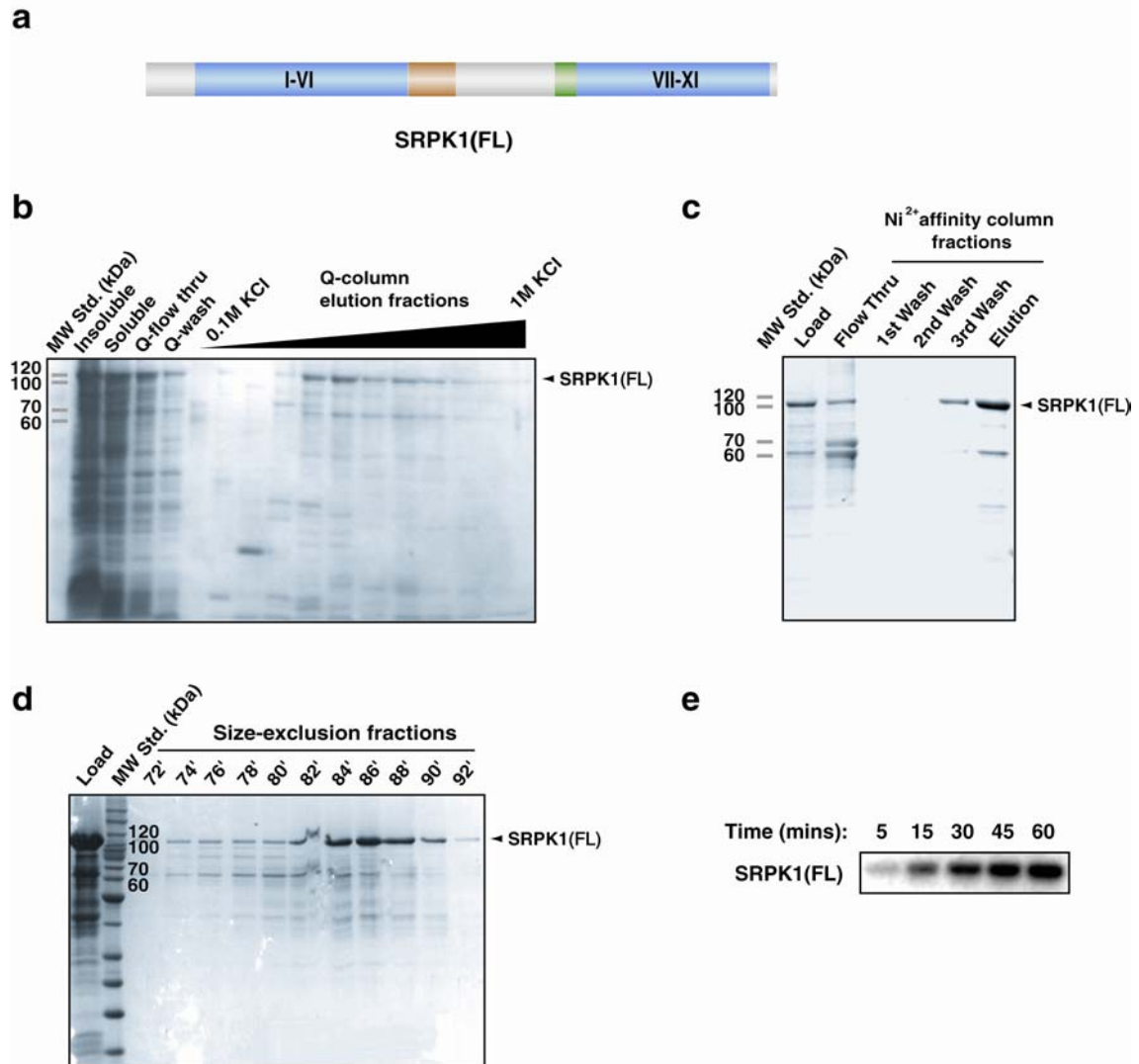


Figure 3.2 Purification of recombinant full length human SRPK1. a) Schematic representation of full length SRPK1. Kinase domain is colored in blue, spacer inserts that are homologous among different species are colored yellow and green. Non-homologous regions are not colored. b) 15% SDS PAGE analysis of crude SRPK1 lysate purified by anion exchange chromatography (Q sepharose). c) 15% SDS PAGE shows that the proteins are about 80 % pure after Ni<sup>2+</sup> affinity chromatography but two degraded products were generated. d) SDS PAGE analysis of peak fractions of SRPK1 from size-exclusion chromatography shows that the full length protein is prone to proteolytic degradation. e) Kinase activity assay using SR-IκBα as substrates shows that the purified kinase is active. SR-IκBα is a fusion protein where the C-terminal 13 residues of Npl3p are fused to a heterologous protein IκBα. The serine of the RS dipeptide at the C-terminus of Npl3p is the physiological target for phosphorylation by Sky1p and can also be phosphorylated by SRPK1.

(<http://www.igb.uci.edu/tools/scratch/>), this region in SRPK1 was predicted to contain a few helical motifs while the rest remains mostly flexible and unstructured (Figure 3.3). As mentioned earlier, although the spacer regions have been implicated to affect the subcellular localization of SRPKs, their deletions do not affect their catalytic activity in vitro. Sequences of these regions in mammalian SRPK1 and Sky1p also share no homology, suggesting that the removal of this region in SRPK1 is likely to have little effect on the kinase activity or kinase fold, as in the case of Sky1p.

The N-terminal His-tagged construct of a spacer-deleted SRPK1 was obtained from our collaborators. In this construct, the 268 residues long spacer region was removed and the amino acid stretch 1-255 was directly connected to the stretch 493-655, and is referred as SRPK1 $\Delta$ S (Figure 3.4a). The protein expressed well in *E.coli* and could be purified to ~85% homogeneity after two steps of chromatography (Figure 3.4b and c). However, the protein was also susceptible to proteolytic degradation over time, yielding two major bands after size exclusion chromatography (Figure 3.4c). Furthermore, this protein has limited solubility and only remains soluble in the presence of at least 1M NaCl. The protein also has a strong tendency to aggregate during concentration and could be concentrated to only ~5mg/ml. The relatively low protein concentration and demand for high salt content were both unfavorable for crystallization and screening trials proved unsuccessful. Activity assays of this construct also showed that the complete removal of the spacer has significantly reduced the kinase activity and will be discussed in more detail in Chapter 4 (Figure 3.4d).



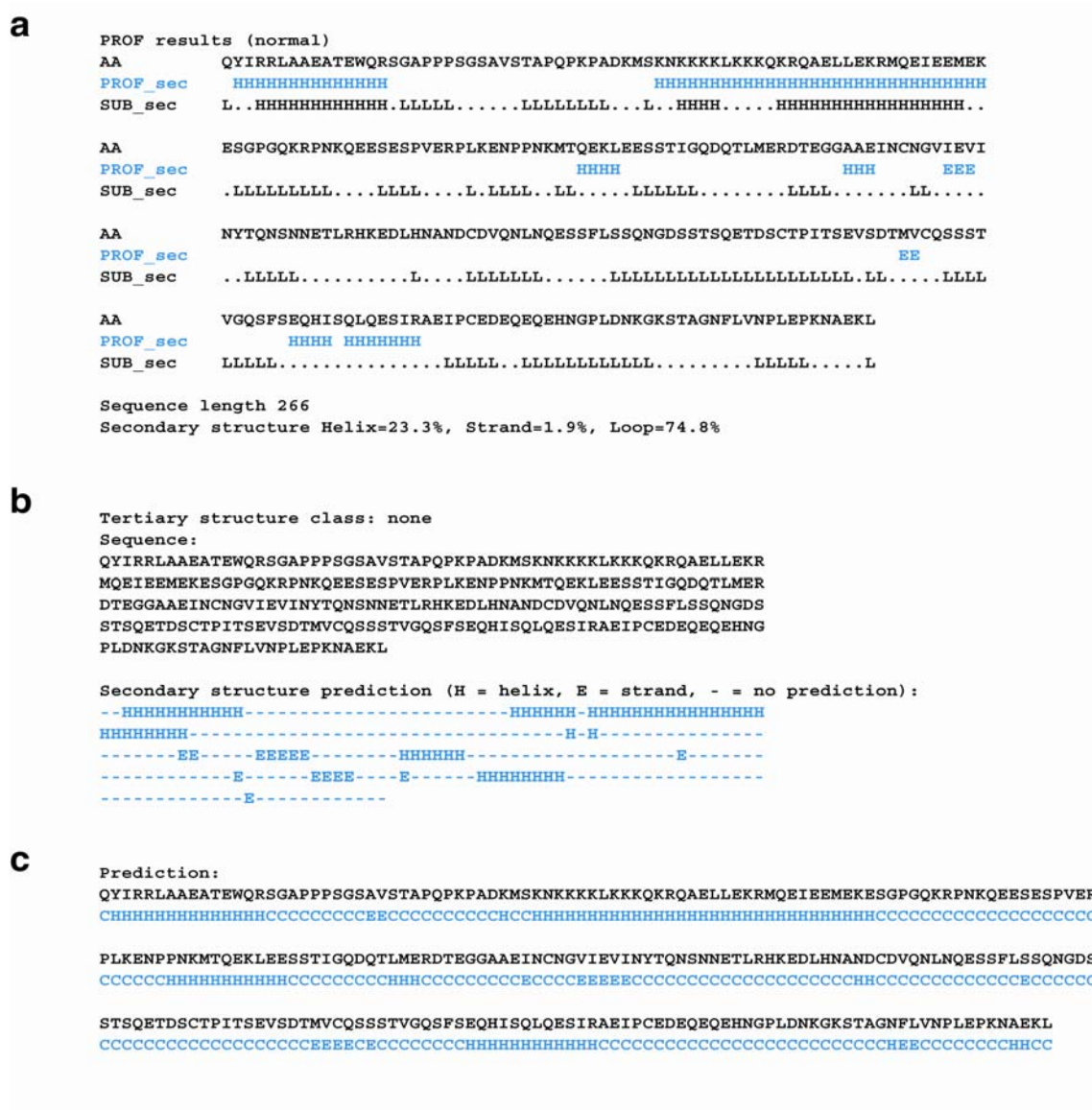


Figure 3.3 Regions of the spacer domain adopt helical structures. Secondary structures prediction programs were used to analyze the spacer domain of SRPK1 and show that the domain is partly helical. a) Prediction by program PredictProtein. b) Prediction by program nnPredict c) Prediction by program SSpro. In all cases, H represents  $\alpha$ -helix, E represents  $\beta$ -strand and C represents coil or unstructured region.

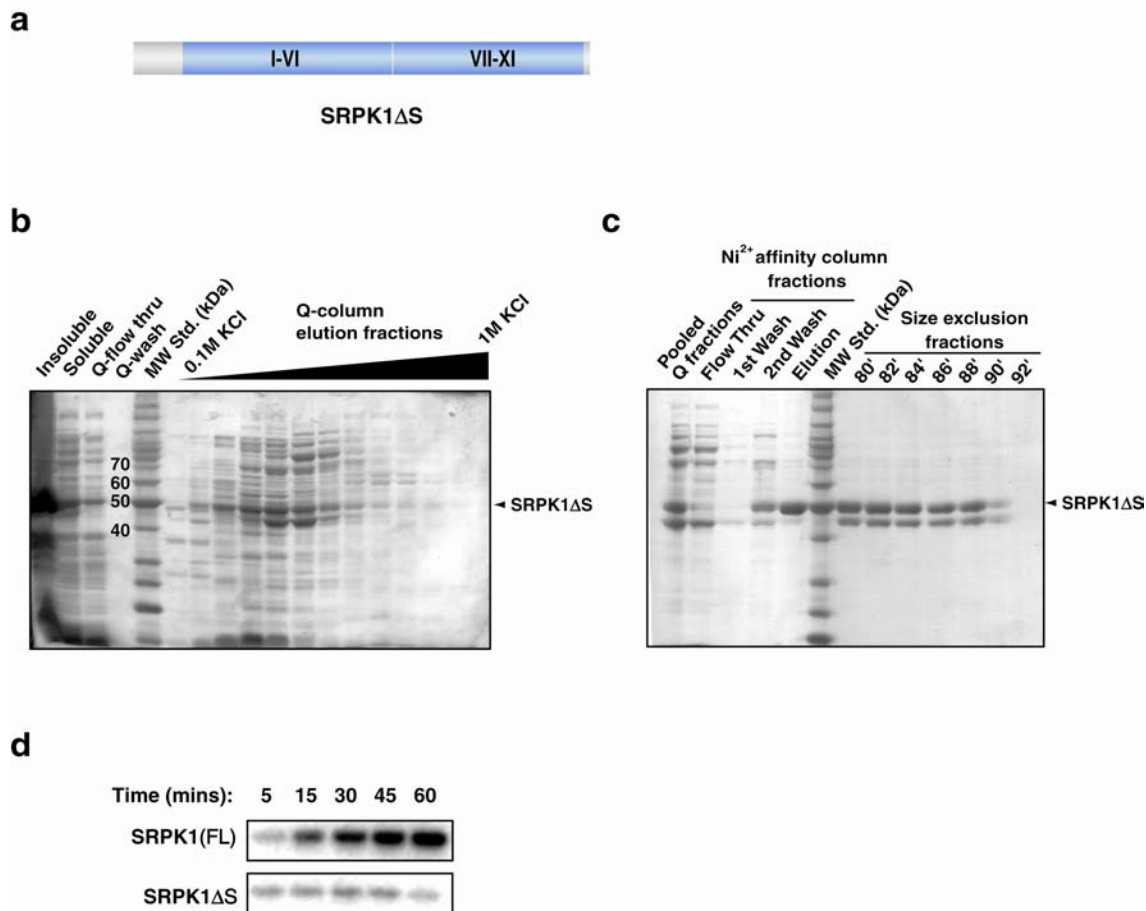


Figure 3.4 Purification of spacer-deleted SRPK1. a) Schematic representation of SRPK1 $\Delta$ S. The complete spacer domain is removed in this construct. b) 12.5% SDS PAGE analysis of crude SRPK1 $\Delta$ S lysate purified by anion exchange chromatography (Q sepharose). c) SDS PAGE analyses of fractions from Ni<sup>2+</sup> affinity column and peak fractions from size exclusion chromatography. A major degraded product of size ~45 kDa was generated after size exclusion chromatography. d) SRPK1 $\Delta$ S shows lower catalytic activity than the full-length kinase. SR-1 $\kappa$ B $\alpha$  was used as a substrate in the assays.

In order to identify a more stable fragment of SRPK1, we aligned and compared the sequences of different SRPK1 sequences from 9 species (Figure 3.1). The sequence alignment revealed that two short regions outside the kinase core, one at the N-terminus and the other at the C-terminus of the spacer region, are conserved among different species of SRPK1 except in yeast (Sky1p). More importantly, the N-terminus of this spacer region was previously predicted to adopt an  $\alpha$ -helical structure (Figure 3.3). We therefore constructed a new spacer-deleted form of SRPK1 where only the central 217 residues (256-473) of the spacer domain were removed, but the two flanking segments were retained (SRPK1 $\Delta$ S1). This protein was purified and shown to have higher solubility than the previous construct (soluble in 500mM KCl vs 1M NaCl) and its activity is similar to that of the wild type kinase (Figure 3.5).

Concurrent with SRPK1 $\Delta$ S1 crystallization trials, proteolysis studies of SRPK1 $\Delta$ S with trypsin and chymotrypsin and SRPK1 $\Delta$ S1 with chymotrypsin were performed respectively (Figure 3.6). In all three experiments, stable fragments that were ~5 kD smaller than the starting proteins were obtained, suggesting a non-homologous region beyond the kinase domain core of human SRPK1 might be solvent exposed and prone to degradation. Since the two constructs are almost identical except the retention of two short regions of spacer in SRPK1 $\Delta$ S1, we speculated that the solvent exposed and flexible region is the N-terminus of the kinase. In fact, while most SRPK1s contain N-extensions of different lengths and these regions are mostly divergent among species, there is significant sequence homology among SRPK1s in higher

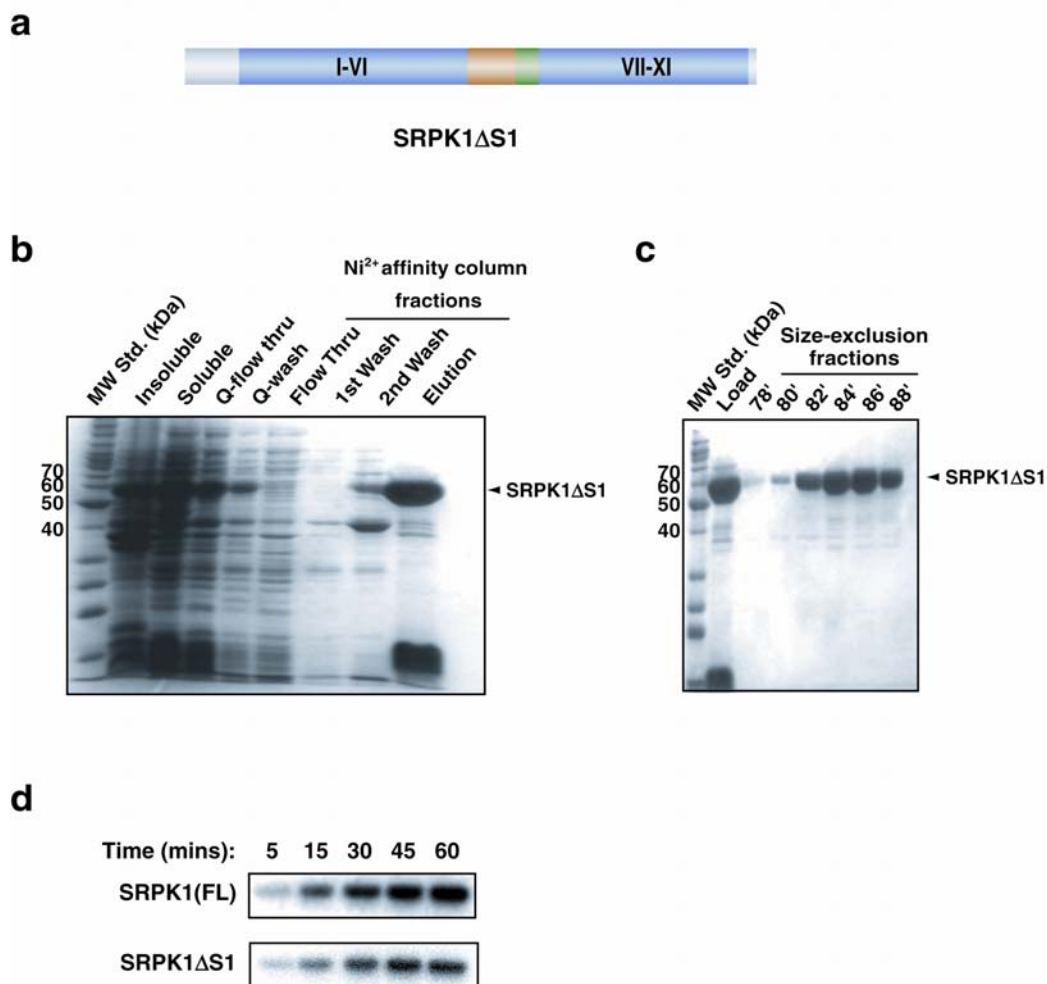


Figure 3.5 Purification of SRPK1 $\Delta$ S1. a) Schematic representation of SRPK1 $\Delta$ S1. In this construct, two short regions at the N- (colored yellow) and C- (colored green) termini of the spacer region were retained when compare to SRPK1 $\Delta$ S. b) 12.5% SDS PAGE analysis of SRPK1 $\Delta$ S1 purified by Q-sepharose column and Ni<sup>2+</sup> affinity column. The flow-through fraction instead of the elution fractions from Q-sepharose column was loaded onto the Ni<sup>2+</sup> affinity column. Elution from the Ni<sup>2+</sup> affinity column shows one major degradation product. c) 12.5% SDS PAGE analysis of peak fractions from size exclusion shows resolution of SRPK1 $\Delta$ S1 and the degradation product. d) SRPK1 $\Delta$ S1 shows similar catalytic activity to the full-length kinase when SR-I $\kappa$ B $\alpha$  was used as a substrate.

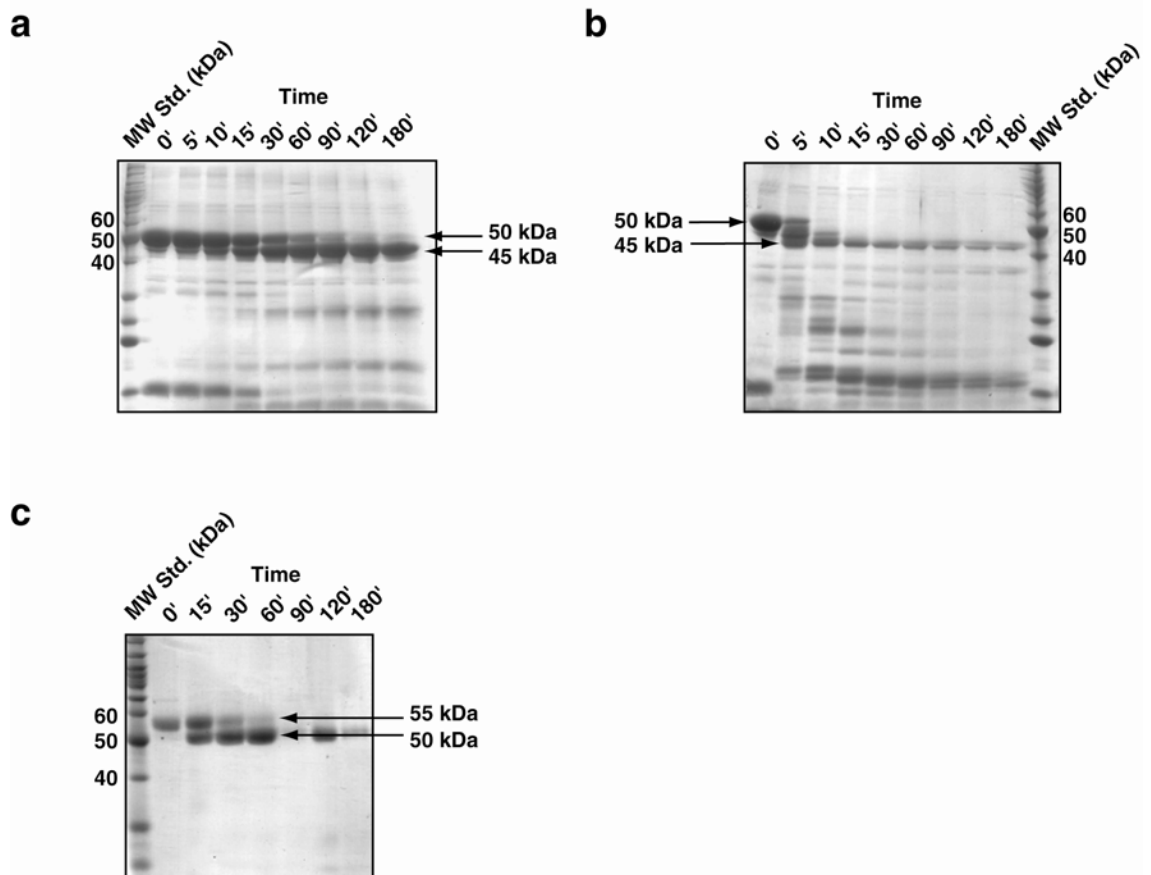


Figure 3.6 Stable fragment of SRPK1. SRPK1 $\Delta$ S was digested with a) chymotrypsin and b) trypsin for the indicated period of time. Digestion with either proteases generated stable products of size  $\sim$ 45 kDa, which are 5 kDa lower than the uncut SRPK1 $\Delta$ S. c) Proteolytic cleavage of SRPK1 $\Delta$ S1 by chymotrypsin. A stable product of size  $\sim$ 50 kDa was obtained and is 5 kDa lower than the uncut protein.

organisms within the 20 residues N-terminal to the start of kinase domain core. We therefore retained this 20 residues region and deleted the first 41 N-terminal residues of SRPK1 $\Delta$ S1 and referred to this truncated construct as SRPK1 $\Delta$ NS1. This protein was expressed in *E.coli* and first purified by anion-exchange and Ni-NTA affinity chromatography. The hexa-histidine tag was then removed by thrombin and the untagged protein was loaded onto a size-exclusion column and purified to ~95% homogeneity (Figure 3.7). In order to acquire a monodispersed protein, a dynamic light scattering (DLS) experiment was performed to analyze the protein dialysed against different buffers. The results from DLS experiment suggested MES pH 6.5 as the optimal buffer to be used during the purification process (Figure 3.8). The protein purified in the presence of this buffer remained soluble in 250mM KCl and could be concentrated to 12 mg/ml. The new protein construct also showed similar in vitro phosphorylation activity as the wild-type kinase (Figure 3.7f).

## **2. Crystallization of an Active Fragment of SRPK1**

Crystallization screening of SRPK1 $\Delta$ NS1 using commercially available sparse matrix screen successfully provide a condition that yields small but well ordered crystals. An optimal crystallization condition was achieved by finer screening using hanging drop vapor diffusion method at 18 °C and consistently grew large size crystals (400  $\mu$ M X 120  $\mu$ m X 100  $\mu$ m) (Figure 3.9a). This condition contains 200 mM ammonium acetate, 100 mM sodium citrate pH5.6 and 15% PEG 3350 as the precipitant. Isolated crystals were dissolved in buffer

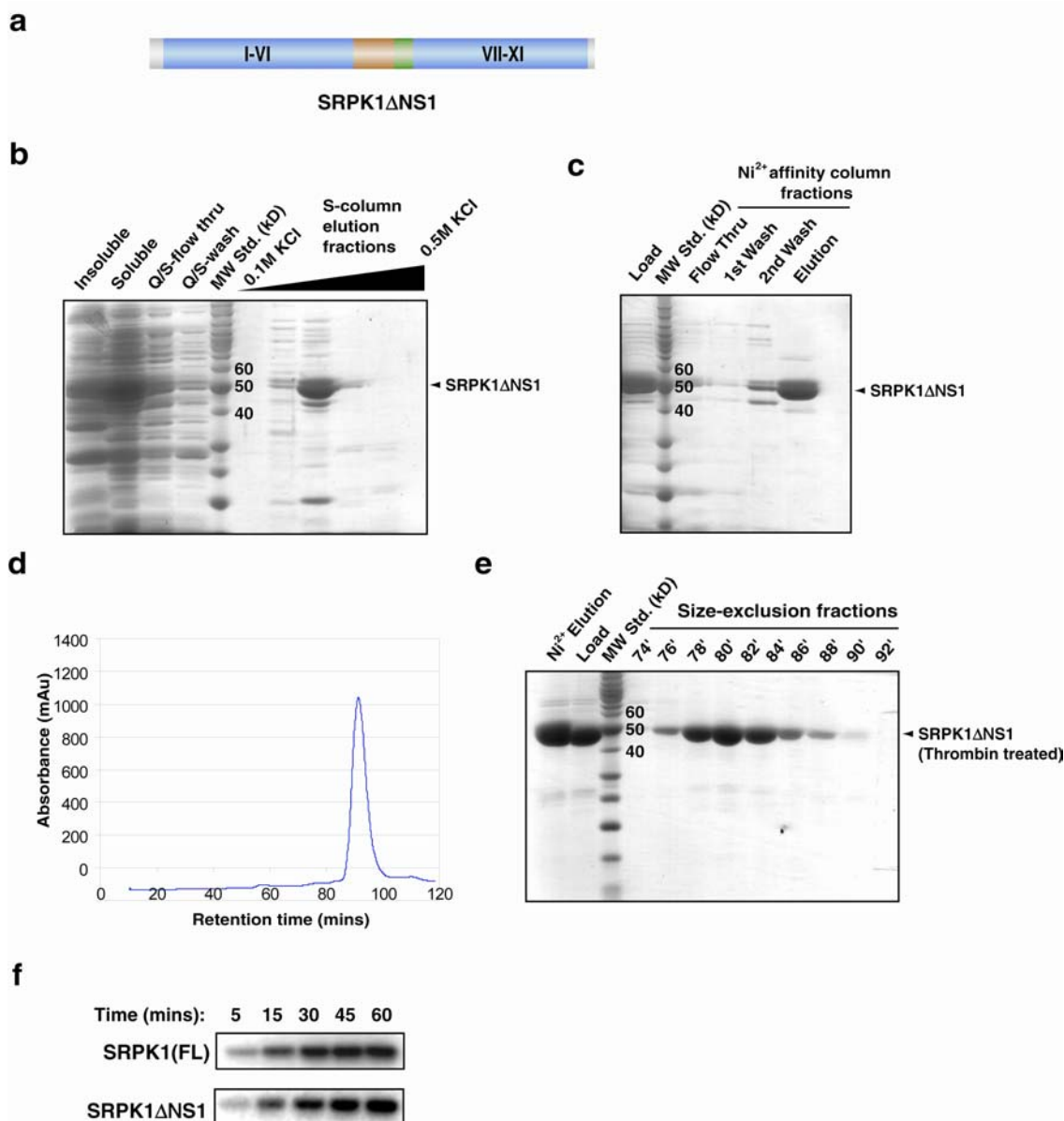


Figure 3.7 Purification of SRPK1 $\Delta$ NS1. a) Schematic representation of SRPK1 $\Delta$ NS1. 41 residues N-terminal to SRPK1 $\Delta$ S1 were truncated in this construct. b) 12.5% SDS PAGE analyses of SRPK1 $\Delta$ NS1 purified by Q- and S-sepharose column. The flow-through fraction from Q-sepharose column was loaded directly onto the S-sepharose column (cation exchange). c) 12.5% SDS PAGE analysis shows that SRPK1 $\Delta$ NS1 is about 90 % pure after Ni<sup>2+</sup> affinity chromatography. d) Elution profile of size exclusion chromatography shows that the SRPK1 $\Delta$ NS1 is homogeneous. The retention time of the protein corresponds to a monomeric form. e) SDS-PAGE analysis of size-exclusion fractions of SRPK1 $\Delta$ NS1. f) Catalytic activity of SRPK1 $\Delta$ NS1 is similar to the full-length kinase when SR-I $\kappa$ B $\alpha$  was used as substrates.

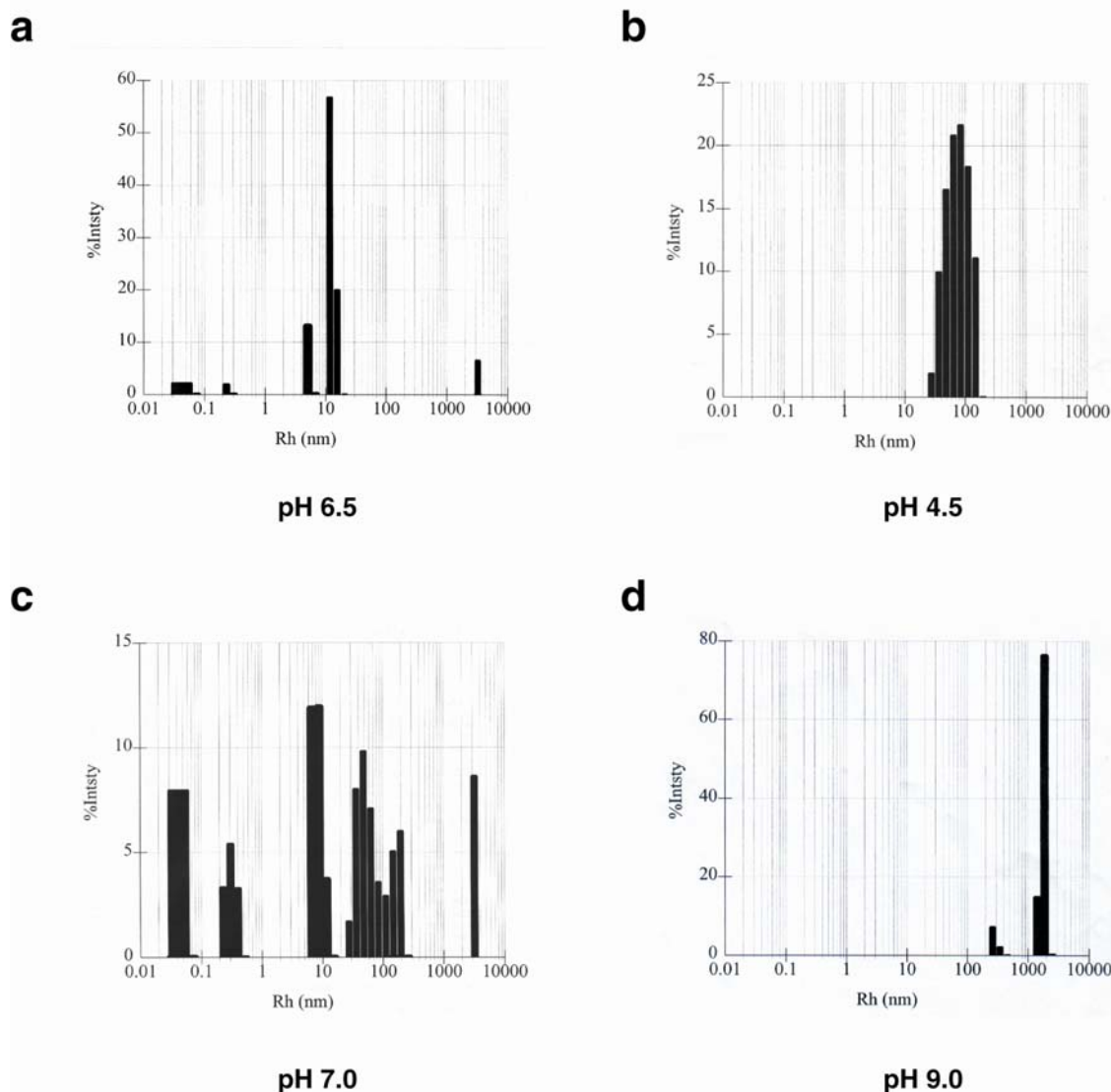


Figure 3.8 SRPK1 is monodisperse at pH 6.5. Dynamic light scattering profiles of SRPK1 $\Delta$ NS1 dialysed against different buffers: a) MES pH 6.5, b) Sodium Acetate pH 4.5, c) Tris pH 7.0 and d) CAPSO pH 9.0. While SRPK1 $\Delta$ NS1 is either polydisperse or aggregated at pH 4.5, 7.0 or 9.0, the protein remains monodisperse in buffer MES pH 6.5. A calibration curve was not generated, therefore the molecular mass or the size of the protein cannot be estimated from the measured hydrodynamic diameter (Rh).



and analyzed by SDS-PAGE to confirm that the crystals consisted of SRPK1 $\Delta$ NS1 protein (Figure 3.9b). The crystals were cryo-protected by step-wise addition of ethylene glycol to a final concentration of 20% (v/v) and flash frozen in liquid nitrogen prior to data collection. The crystals were tested using our home source X-ray and diffracted to  $\sim 2.1$  Å.

### **3. Data Collection of Apo- and Seleno-Methionine Derivative Crystals of SRPK1 $\Delta$ NS1**

Higher resolution X-ray diffraction data of the apo-enzyme crystals were collected using a SBC2 CCD detector at SBC-CAT synchrotron beamline ID-19 of the Advanced Photon Source (APS) at Argonne National Laboratory. The diffraction quality of crystals improved significantly using the synchrotron radiation and the best diffracted crystal (1.73 Å) was characterized to the hexagonal space group P6<sub>5</sub>22, with unit cell dimensions  $a = b = 75.11$  Å,  $c = 313.33$  Å,  $a = b = 90^\circ$  and  $c = 120^\circ$ . Due to the long dimension of the c-vector of the unit cell, data set of complete  $120^\circ$  was collected at  $0.25^\circ$  oscillations in order to prevent overlapping of reflections (Figure 3.8c). Data collection statistics are summarized in Table 3.1.

In order to obtain the initial phase solution of apo-SRPK1 $\Delta$ NS1, we decided to perform a multi-wavelength anomalous dispersion (MAD) experiment. Five methionines within the protein construct were substituted for seleno-methionines and the selenium used as a source of anomalous scattering. Seleno-methionine derivative proteins were expressed in *E. Coli* by

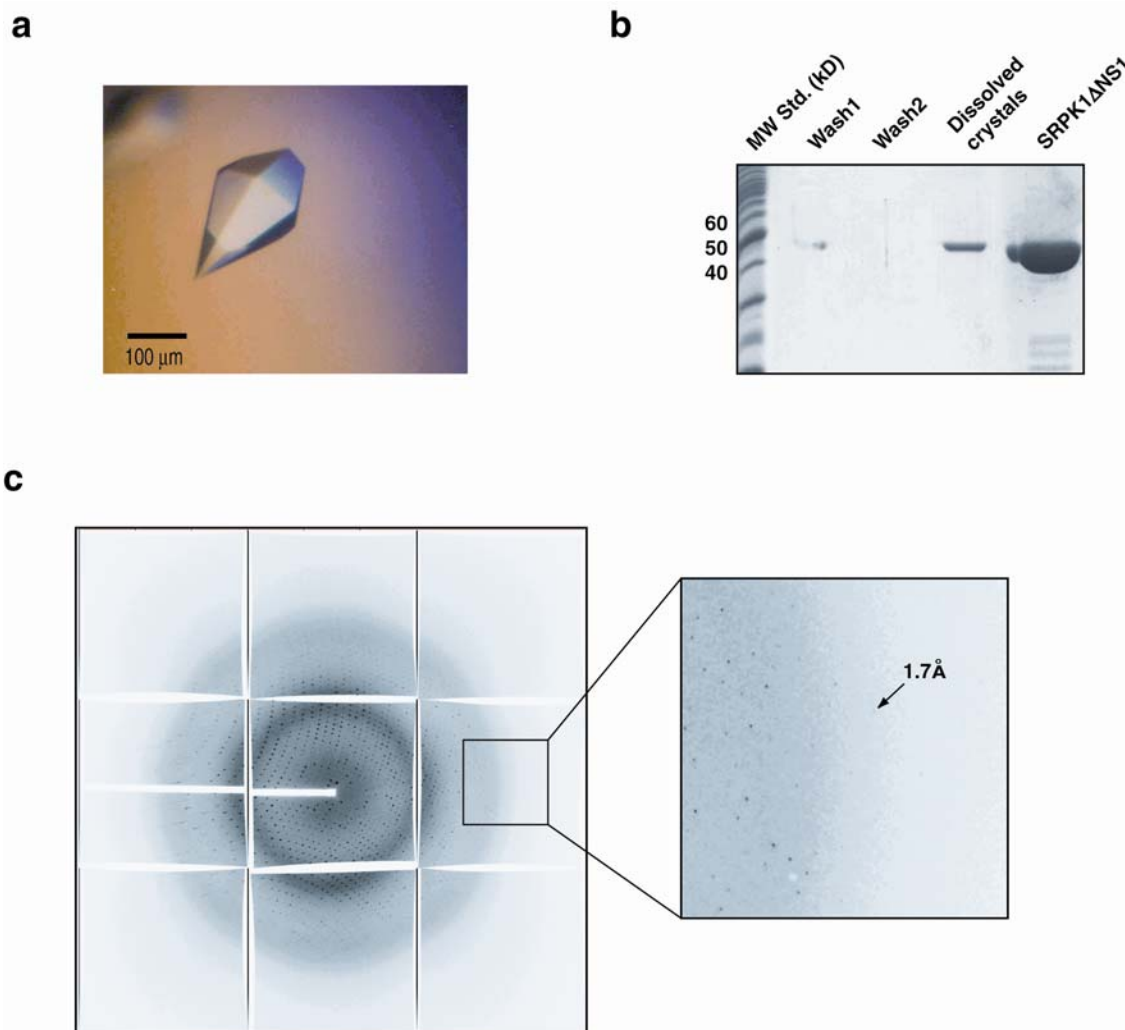


Figure 3.9 Crystal of SRPK1 $\Delta$ NS1 (a) SRPK1 $\Delta$ NS1 crystal grown in 15% PEG 3350, 200mM Ammonium acetate and 100mM Sodium citrate pH5.6 (b) SDS-PAGE analysis of dissolved crystals show the crystals are composed of SRPK1 $\Delta$ NS1 protein. (c) Diffraction image of an SRPK1 $\Delta$ NS1 crystal collected at 0.5 $^\circ$  oscillation at synchrotron beamline ID-19 of the Advanced Photo Source. Spots are visible to 1.7 $\text{\AA}$  (inset).

Table 3.1: Data Collection of SRPK1ΔNS1 crystal

Crystal form	Native	Peak (SeMet)	Edge (SeMet)	Remote (SeMet)
<b>Data Collection</b>				
Data Collection Source	APS ID-19		ALS 5.0.2	
Wavelength (Å)	1.0332	0.97976	0.97996	0.9649
Resolution (Å)	30 - 1.73	50-2.0	50 - 2.0	50-2.14
No. of measured reflections	562787	496903	556508	540921
No. of unique reflections	55769	35757	35797	29780
Completeness (outer shell) (%)	99.5 (99.9)	98.0 (87.1)	98.6(90.3)	95.3(61.9)
$I/\sigma$ (overall / outer shell)	11.5 (5.87)	16.5(1.55)	13.5(1.61)	11.0(0.91)
$R_{\text{sym}}^a$ (overall / outer shell) (%)	7.9 (58.4)	5.0 (37.0)	5.8 (55.2)	8.1 (78.8)

$^a R_{\text{sym}} = \sum |I - \langle I \rangle| / \sum I$

culturing the bacteria in a methionine-deprived minimal growth media and in the presence of seleno-methionine. The derivative crystals were crystallized under the same conditions as the apo-enzyme crystals. MAD data of the seleno-methionine derivative crystal was collected using an ADSC Quantum-4 CCD area detector at ALS synchrotron beamline 5.0.2. of Advanced Light Source (ALS) at Lawrence Berkeley National Laboratory. Measurement of the absorption edge of the seleno-methionine was first performed by measuring the fluorescence from the crystal. The crystal diffracted to 2.0 Å and three data sets were collected for 120° at 0.5° oscillations at three different wavelengths (Table 3.1). The tested crystal suffered minor radiation-induced decay throughout the course of data collection, resulting in a slightly lower resolution data set at the remote wavelength. The Friedel pairs were recorded by first collecting a small wedge of diffraction data and then rotated the crystal by 180° to collect the same wedge of data. The space group of the tested derivative crystal is the same as the apo-crystal with similar unit cell dimensions ( $a = b = 74.84$  Å,  $c = 312.94$  Å).

### **C. Discussion**

The first requirement for protein X-ray structure determination is to obtain well-ordered crystals. However, obtaining suitable crystals is the least understood process in protein X-ray crystallography. Indeed, the crystallization process is mostly a trial-and-error procedure due to the presence of many uncertainties. For instance, the choices of precipitant, buffer and even

temperature for crystallization can be critical to promote crystal growth, but these parameters are mostly unknown until the first crystal is acquired. More importantly, the impurities present during crystallization, heterogeneity of proteins and the presence of flexible disordered regions within the proteins can also hinder the growth of protein crystals. Fortunately, advances in protein purification and molecular cloning technologies allow structural biologists to overcome some of these obstacles. Target proteins can be engineered to remove local disordered regions or independent domains can be isolated to provide better crystallization targets. Indeed, a significant fraction of the 27800 X-ray protein structures deposited in the protein data bank (PDB) to date were obtained using these techniques.

The catalytic domains of all protein kinases adopt the same scaffold despite the diversity in their functions, substrate specificities or regulation mechanisms. However, kinases frequently contain non-conserved regions beyond the kinase domain, making it difficult to develop a general scheme for purification or crystallization of these proteins. This is proven in the case of SRPK1. Although Sky1p (SRPK1 ortholog in yeast) and human SRPK1 share high sequence homology within the kinase domain, the truncation of the non-conserved spacer region induces different effects on the two kinases. Sky1p remains fully stable and crystallizable in the absence of this region but SRPK1 $\Delta$ S is rendered unstable and cannot be crystallized. Moreover, difficulties encountered during purification, like salt sensitivity and protein aggregation, were unseen before in the case of Sky1p. Such dramatic

differences suggest a special role of the spacer region in the context of SRPK1. Further sequence analysis of the spacer regions of different SRPK1s led us to retain short N- and C-terminal flanking regions of the spacer. This modification significantly improved the kinase behavior in solution and expedited the purification process. However, crystallization of this SRPK1 construct (SRPK1 $\Delta$ S1) did not succeed until the removal of the non-conserved N-terminus (SRPK1 $\Delta$ NS1). The pH of buffer used during purification was also shown to affect significantly the monodispersity of the target protein, which is a long-known essential factor during the crystallization process. In conclusion, we report here the crystallization of SRPK1, which was critically dependent on the rational truncation of the non-kinase domain regions and the monodispersity of the protein.

**Chapter IV: Insights into the Constitutive Activity from  
the X-ray Crystal Structure of SRPK1 Apoenzyme**

## A. Introduction

Pre-mRNA splicing is an essential step for the expression of most eukaryotic genes. This process is regulated by a complex series of protein-protein, protein-RNA and RNA-RNA interactions and requires small ribonucleoprotein particles (snRNPs) and non-snRNP proteins. One family of these non-snRNP proteins is the SR protein family. SR proteins are characterized by the presence of one or two N-terminal RNA-recognition motifs (RRMs) and a C-terminal RS domain, which contains repeats of consecutive arginine-serine dipeptides. The serines within the RS domain are extensively phosphorylated and their phosphorylation states govern multiple facets of SR protein activity, including their localization, splice site selection, mRNA transport and translation (Caceres et al., 1998; Huang et al., 2003; Huang and Steitz, 2001; Misteli et al., 1997; Sanford et al., 2004).

Members from two protein kinase families, SRPKs and Clk/Stys, have been shown to specifically phosphorylate SR proteins (Colwill et al., 1996; Gui et al., 1994a). Both SRPKs and Clk/Stys belong to the CMGC group of the kinase superfamily. Although SRPKs are classified as serine/threonine kinases, they are highly RS-specific and only phosphorylate serine, not threonine, in the presence of arginine (not lysine) in their substrates (Gui et al., 1994b). SRPKs are highly conserved from budding yeast to humans and share high sequence homology within the kinase domains (Figure 4.1) (Siebel et al., 1999; Takeuchi and Yanagida, 1993; Tang et al., 1998; Wang et al., 1998a). They are constitutively active kinases as they are active when isolated from their native



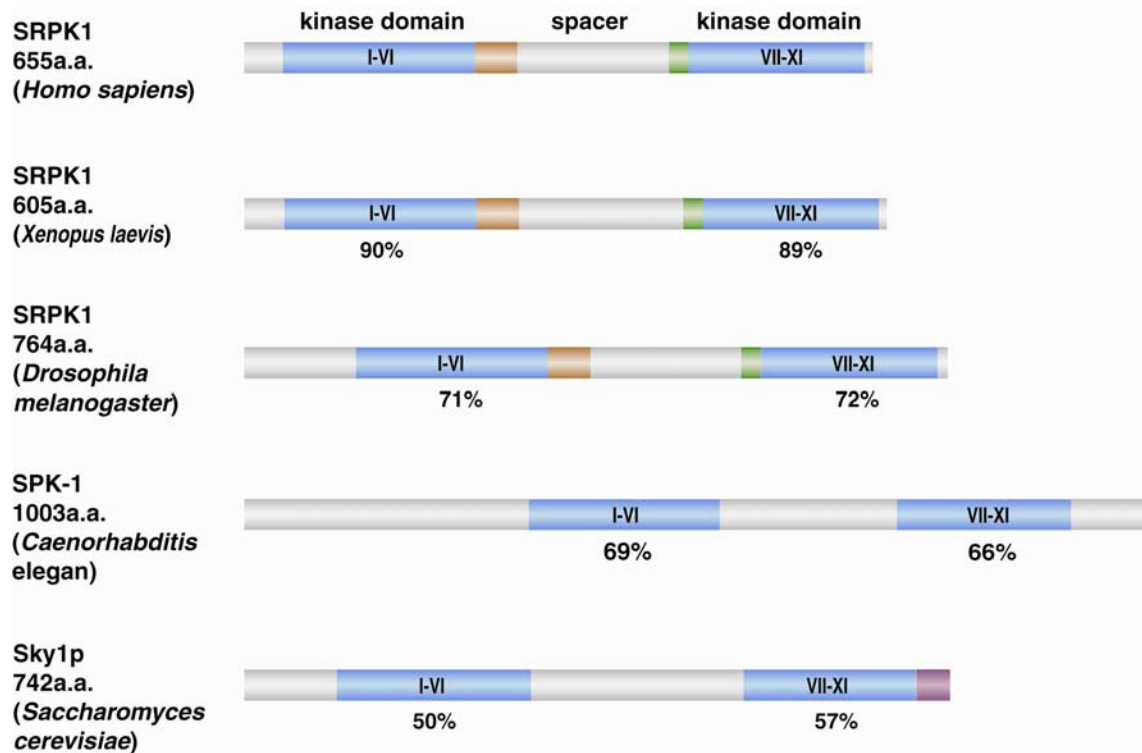


Figure 4.1 Kinase domains of SRPKs are conserved. The kinase domains of SRPK1 from different species are highly homologous and are colored in blue. Their sequence identities compared to human SRPK1 are indicated. Yellow and green segments represent the homologous spacer inserts in human and *Drosophila* SRPKs. The red segment present in Sky1p is essential for its kinase activity.

sources or expressed in bacteria. All members of the SRPK family are characterized by the presence of a spacer region of ~250 – 300 residues that bifurcates the kinase domains (Figure 4.1) (Gui et al., 1994a; Wang et al., 1998a). This spacer domain of SRPKs has been shown to regulate the sub-cellular localization of the kinases but is not required for the in vitro kinase activity (Siebel et al., 1999). In human, SRPK1 is predominantly cytoplasmic in interphase cells, but translocates to the nucleus at the G2/M boundary and is dependent on the spacer domain (Ding et al., 2006). SRPKs also contain non-homologous N- and C-terminal extensions that vary in length. In the case of Sky1p (SRPK in yeast), the C-terminal extension is important to maintain its catalytic activity. However, the roles of these extensions remain largely unknown for other SRPK family members.

Protein kinase activity can be regulated by many ways. While many kinases are regulated by phosphorylation at their activation segments, some are controlled by inter- or intra-molecular interactions with positive or negative regulatory elements. In more complex cases, these two mechanisms are used together to regulate the kinase activity. For instances, full activation of CDK2 requires both binding of cyclin and phosphorylation at the activation loop, while c-Src is inhibited by intramolecular interactions between the phosphorylated C-terminal tail and the SH2 domain, and between the kinase-SH2 linker and SH3 domain (Gonfloni et al., 2000; Jeffrey et al., 1995; Moarefi et al., 1997; Russo et al., 1996). As mentioned above, SRPKs are constitutive kinases and do not require phosphorylation or regulator binding to become activated. Indeed, X-ray

structure of Sky1p has revealed that the kinase contains multiple intrinsic regulatory elements to maintain its constitutive activity (Nolen et al., 2001). However, the 36 residues C-terminal extension of Sky1p, which stabilizes the active conformation of the activation segment and is essential for the kinase activity, is not conserved among species and is missing in human SRPK1. This suggests the members of the SRPK family utilize drastic different mechanisms to establish their constitutive activities.

In this chapter, we present the 1.73 Å resolution X-ray crystal structure of an active fragment of SRPK1. We describe multiple distinct secondary structure elements present in the model and address how these non-kinase core regions maintain SRPK1 in its active conformation in a manner distinct from Sky1p and other kinases.

## **B. Results**

### **1. Structure Solution of SRPK1 $\Delta$ NS1**

The active fragment of SRPK1 (SRPK1 $\Delta$ NS1) used in the crystallization contains residues 42 to 255 and 474 to 655 ligated by the addition of a single alanine. The protein crystallized in a hexagonal space group P6<sub>5</sub>22 with solvent content of 56.6%. The phasing problem was solved by multi-wavelength anomalous diffraction (MAD) methods. The structure was refined to free R factor of 21.76% and experimental R factor of 19.72% with good stereochemistry. Table 4.1 summarizes the phase determination and the model

Table 4.1: Phasing and Refinement Statistics of SRPK1ΔNS1 crystal

Crystal form	Native	Peak (SeMet)	Edge (SeMet)	Remote (SeMet)
<b>Phasing</b>				
Resolution (Å)	30 - 1.73	50-2.0	50 - 2.0	50-2.14
Phasing power <sup>a</sup> (centric/acentric)		2.086/2.030	2.480/2.324	1.488/1.256
Figure of merit (centric/acentric)		0.843/0.648		
<b>Refinement</b>				
Resolution (Å)	30 - 1.73			
R <sub>crys</sub> <sup>b</sup> (%)	19.72			
R <sub>free</sub> <sup>c</sup> (%)	21.76			
R.m.s. deviations				
Bond lengths (Å)	0.0106			
Bond angles	1.431			
<sup>a</sup> Phasing power = $\langle  F_H  \rangle / \langle   F_{PH}  -  F_P + F_H   \rangle$ .				
<sup>b</sup> R <sub>crys</sub> = $\sum   F_{obs}  -  F_{calc}   / \sum  F_{obs} $ , where $F_{obs}$ and $F_{calc}$ are the observe and calculated structure factors, respectively.				
<sup>c</sup> R <sub>free</sub> was calculated with 5% of the data excluded from the refinement calculation.				

and 298 water molecules. Residues 42-62, 65 and 66 of the N-terminus and residues 237-255 and 474 of the spacer region are poorly ordered and were not modeled. The coordinates for this structure have been deposited in the Protein Data Bank under ID code 1WAK.

Crystals of SRPK1 $\Delta$ NS1 contain one molecule per asymmetric unit. Two SRPK1 $\Delta$ NS1 molecules form a dimer organized in a head-to-tail arrangement across a crystallographic 2-fold axis, reminiscent of those observed for GSK3 $\beta$  and phosphorylase kinase (PhK) (Figure 4.2a) (Dajani et al., 2001; Lowe et al., 1997). The dimer interface buries 2582  $\text{\AA}^2$  of accessible surface area, which is more than those observed for both GSK3 $\beta$  and PhK dimers. However, size exclusion chromatography shows that SRPK1 $\Delta$ NS1 is a monomer and different pH conditions do not alter its oligomeric state. This suggests that the observed dimerization might not be a physiological event (Figure 4.2b).

## **2. Overall Architecture of SRPK1 $\Delta$ NS1**

The overall structure of SRPK1 $\Delta$ NS1 is similar to other protein kinases where the kinase core is folded into two lobes - the N-terminal small lobe composed primarily of  $\beta$ -strands and the C-terminal large lobe composed mostly of  $\alpha$ -helices (Figure 4.3). The SRPK1 $\Delta$ NS1 structure shares high structural homology with its yeast counterpart Sky1p with two notable differences. First, a C-terminal 40 residue segment of Sky1p that surrounds the back of its kinase core domain is absent in SRPK1 $\Delta$ NS1 (Figure 4.4).

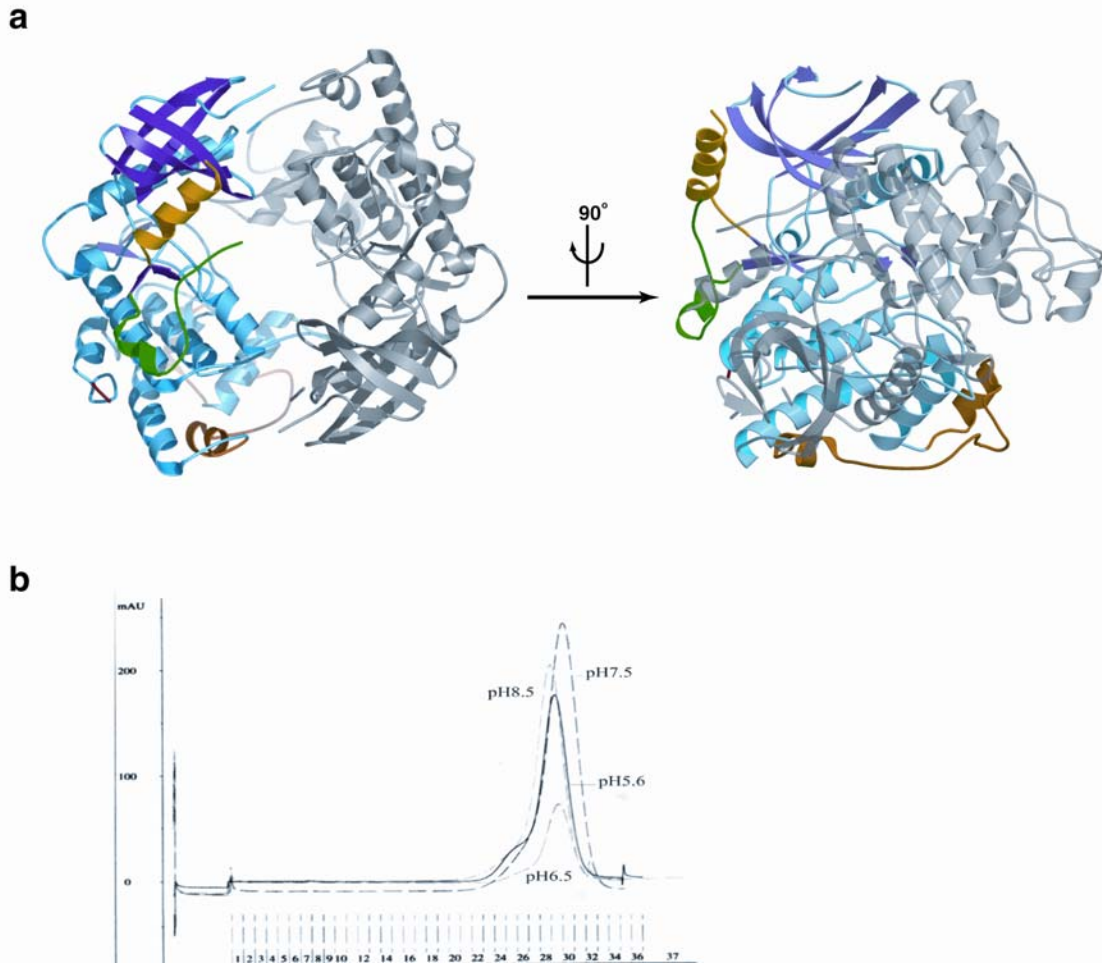


Figure 4.2 SRPK1 $\Delta$ NS1 is a monomer in solution. a) Ribbon diagram of SRPK1 crystal dimer. Crystals of SRPK1 $\Delta$ NS1 contain one molecule per asymmetric unit. Two SRPK1 $\Delta$ NS1 molecules form a dimer organized in a head-to-tail arrangement across a crystallographic 2-fold axis. b) Size exclusion profile of SRPK1. Size exclusion chromatography shows that SRPK1 $\Delta$ NS1 is a monomer and different pH conditions do not alter its oligomeric state. This suggests that the observed dimerization might not be a physiological event.

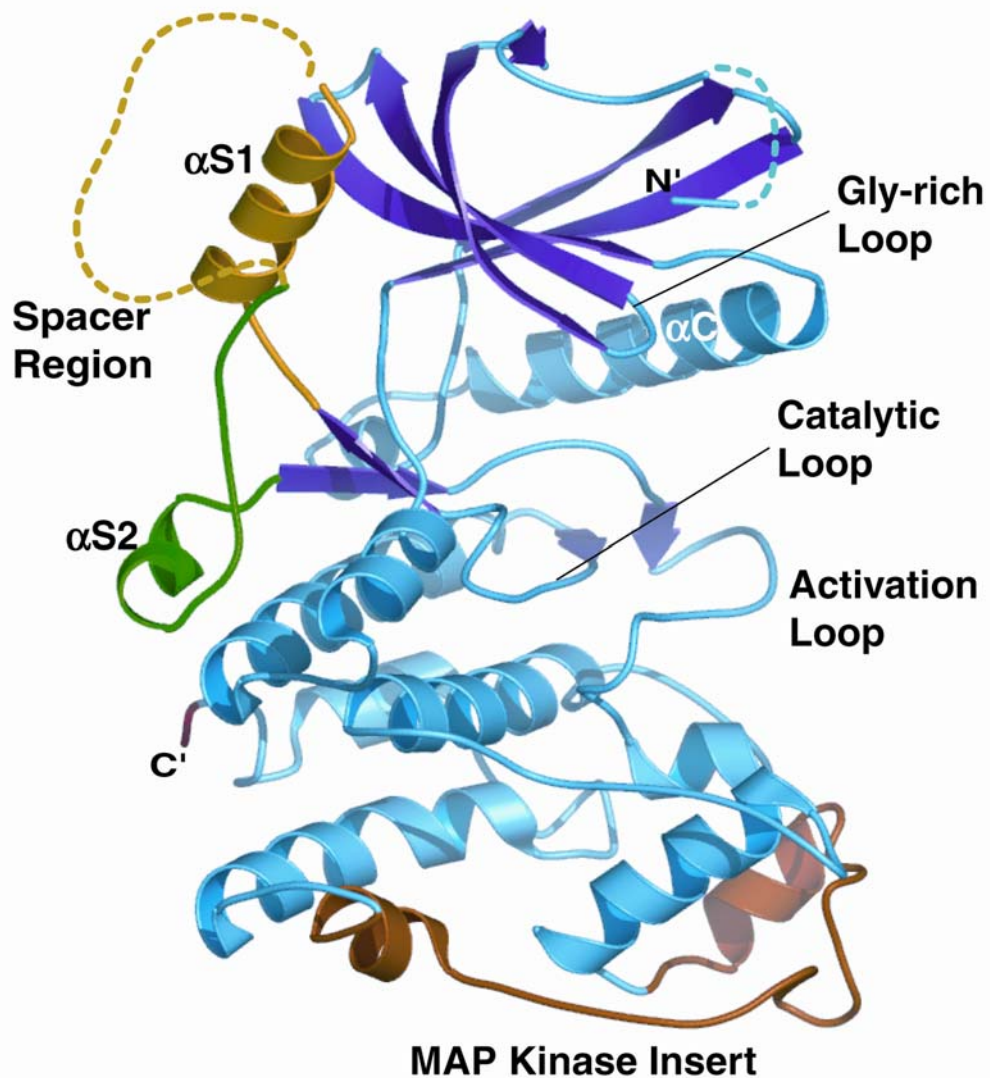


Figure 4.3 Overall architecture of SRPK1 $\Delta$ NS1. The overall structure of SRPK1 $\Delta$ NS1 is similar to other protein kinases where the kinase core is folded into two lobes - the N-terminal small lobe composed primarily of  $\beta$ -strands and the C-terminal large lobe composed mostly of  $\alpha$ -helices. Disordered regions in the structure are represented by dotted lines. The N-terminal spacer insert including helix  $\alpha$ S1 is colored yellow, the C-terminal spacer insert including helix  $\alpha$ S2 is colored green and the MAP kinase insert is colored orange. Helices and strands of the kinase core are colored blue and purple respectively.

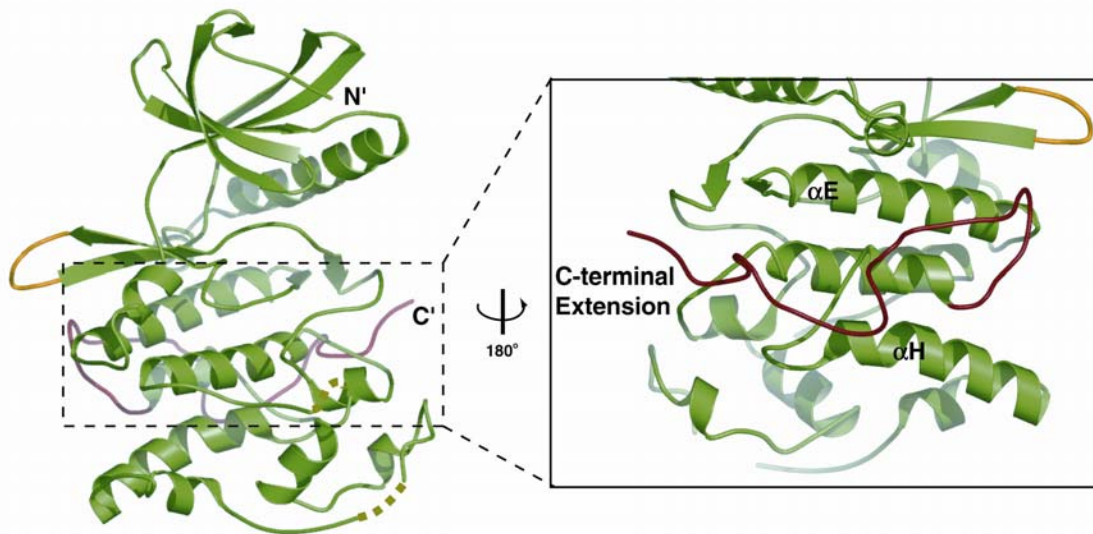


Figure 4.4 Crystal structure of Sky1p (Nolen et al., 2001). The structures of SRPK1 and Sky1p share high structural homology except the presence of a C-terminal extension in Sky1p. The C-terminal extension shown in red wraps the back of the kinase and interacts with the activation loop. The loop in yellow represents an artificial connection resulting from the removal of its spacer.



Secondly, the two spacer inserts in SRPK1 $\Delta$ NS1 adopt helical structures and form a discontinuous wall on the side of the kinase. Despite these differences, SRPK1 $\Delta$ NS1 contains multiple non-kinase core segments in addition to the conserved kinase core like Sky1p. Firstly, SRPK1 $\Delta$ NS1 contains an extra  $\beta$  strand ( $\beta_0$ ) at its N-terminus spanning from residues 74-76. Secondly, residues 140-142 in SRPK1 $\Delta$ NS1 form a  $3_{10}$  helix immediately after helix  $\alpha_C$ . This segment, termed helix  $\alpha_C'$ , is first seen as a seven amino acids helix in Sky1p and was suggested to stabilize the correct position of the helix  $\alpha_C$ . The role of this segment in SRPK will be discussed later. One structural characteristic of the CMGC group of kinases is the presence of an insert, termed MAP kinase insert, between helices  $\alpha_G$  and  $\alpha_H$ . Although the structures of these inserts are not conserved among CMGC kinases, they pack onto the same surface of the large lobe and mask one of the solvent-exposed sides of helix  $\alpha_G$  (Canagarajah et al., 1997; Dajani et al., 2001; De Bondt et al., 1993). The MAP kinase insert of SRPK1 $\Delta$ NS1 forms an extended loop flanked by two helices and plays an important role in its substrate recognition, which will be discussed in Chapter V. The positions of the N- and C- terminal inserts of the spacer region suggest that the rest of the region, which was removed from the SRPK1 $\Delta$ NS1 construct, can be folded independently from the kinase domain and would not affect the structural integrity of the kinase. The critical role of these two spacer inserts in stabilizing an active conformation of SRPK1 $\Delta$ NS1 will be discussed later. Figure 4.5 shows a sequence alignment of SRPK1, Sky1p, PKA and other CMGC

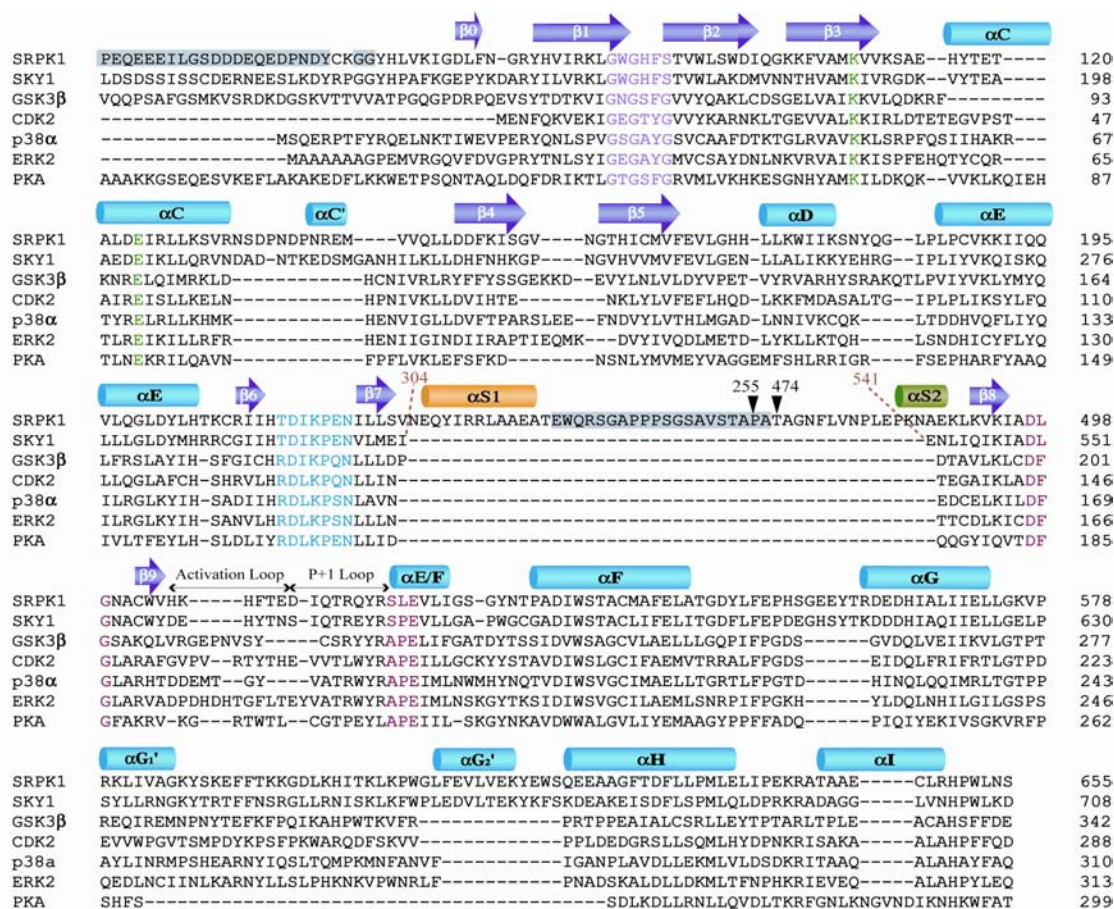


Figure 4.5 The primary sequence and secondary structures of SRPK1 $\Delta$ NS1. The sequence of SRPK1 $\Delta$ NS1 is compared with six other kinases including its close homologue Sky1p. SRPK1, Sky1p, GSK3 $\beta$ , CDK2, p38 and ERK2 belong to the CMGC group of kinases while PKA belongs to the AGC group of kinases. Residues that are disordered and not included in the model are highlighted in gray. Residues 255 to 474 of the spacer of SRPK1 were deleted in the construct crystallized. The spacer of Sky1p spans residues 304 to 541. Red tubes denote  $\alpha$ -helices and purple arrows  $\beta$ -strands. Helices  $\alpha$ S1 and  $\alpha$ S2 are denoted by yellow and green tubes respectively. Homologous residues in the glycine rich loop and the catalytic loop of these kinases are shown in pink and blue. Green denotes the invariant lysine and glutamic acid. The anchor points of the activation segment are denoted by red.

kinases, illustrating the positions of the non-kinase core segments of SRPK1.

### 3. Orientation of Helix $\alpha$ C and Active Conformation

A signature feature, observed in all active kinases, is the proper positioning of helix  $\alpha$ C in the small lobe such that an invariant glutamate of helix  $\alpha$ C forms an ion-pair with an invariant lysine of  $\beta$ 3 (Huse and Kuriyan, 2002; Johnson et al., 1996). This ion-pair is important as it positions the lysine residue to interact with the  $\alpha$  and  $\beta$  phosphates of ATP. The structure of Sky1p suggests a non-kinase core insert, the  $\alpha$ C' insert, helps to position the helix  $\alpha$ C properly. Comparison of the  $\alpha$ C' inserts in Sky1p and SRPK1 $\Delta$ NS1 reveals that SRPK1 contains unique interactions within this region and orients the helix  $\alpha$ C through a different mechanism. The  $\alpha$ C' insert in Sky1p is stabilized by hydrogen bonds between Y283 of helix  $\alpha$ E and the backbone of the insert. However, the observed temperature factors of the insert's backbone are comparatively high except at M220, suggesting this residue also plays an important role in stabilizing the  $\alpha$ C' insert. In fact, M220 of the short helix  $\alpha$ C' is positioned in a hydrophobic pocket formed by L279 from helix  $\alpha$ E, M712 and I715 from the C-terminal extension and is stabilized by van der Waals interactions (Figure 4.6a). These interactions together help to pack the  $\alpha$ C' insert against the kinase core and in turn hold the helix  $\alpha$ C in place. In SRPK1 $\Delta$ NS1, the conserved Y202 from helix  $\alpha$ E forms hydrogen bonds with the backbone carbonyl group of V131 and backbone NH group of D135 of this

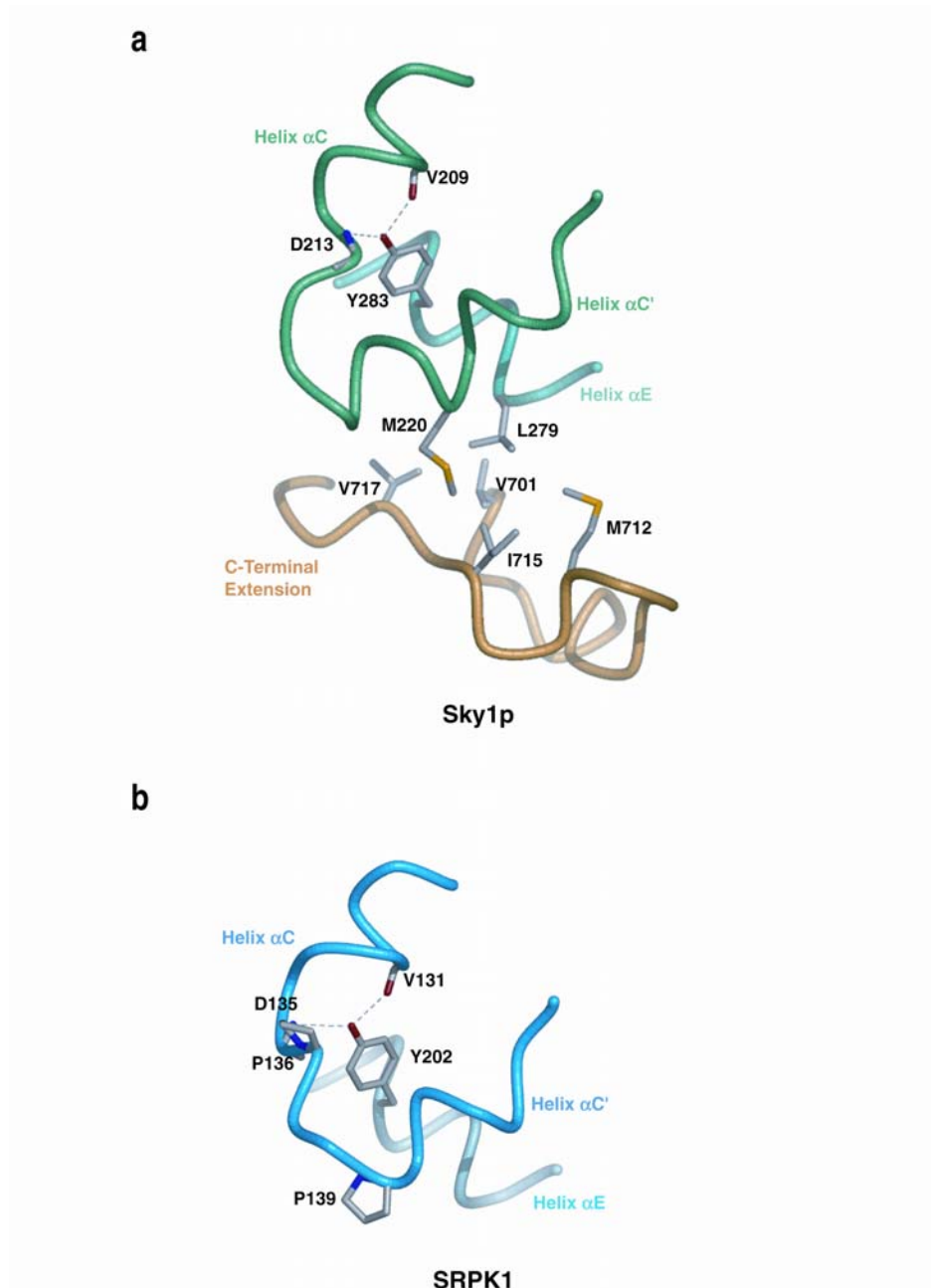


Figure 4.6 Conformation of  $\alpha$ C' insert. a) The  $\alpha$ C' insert of Sky1p interacts with helix  $\alpha$ E and the C-terminal extension and anchors the helix  $\alpha$ C into its correct position. M220 from the helix  $\alpha$ C' is oriented into a hydrophobic pocket formed by residues from helix  $\alpha$ E and the C-terminal tail. b) The correct position of helix  $\alpha$ C in SRPK1 is stabilized by two prolines within the  $\alpha$ C' insert.

segment to steady its position. However, in contrast to Sky1p, SRPK1 does not utilize a hydrophobic pocket to stabilize the insert due to the absence of a C-terminal extension. Instead, two prolines at positions 136 and 139 appears to stabilize the insert by restraining the motion of the backbone and help to lock the helix  $\alpha C$  into its correct position (Figure 4.6b).

Despite the correct positioning of the helix  $\alpha C$  in the structure of SRPK1 $\Delta$ NS1, protein-protein interactions at the dimer interface and local interactions between one of the spacer inserts (helix  $\alpha S1$ ) and the small lobe results in a distortion of the  $\beta$ -sheet formed by  $\beta 1$ , 2 and 3, rotating it slightly toward helix  $\alpha S1$  (Figure 4.7a). Due to this structural movement, K109 and E124, the two invariant residues mentioned above, are pulled apart by  $\sim 4.6$  Å (Figure 4.7b). However, these two residues are ordered and oriented correctly, requiring only a small adjustment to form the important ion-pair. Barring this local distortion within the small lobe, SRPK1 $\Delta$ NS1 reveals molecular features consistent with its constitutive activity.

#### **4. Spacer Domain of SRPK1**

A unique feature of the structure of SRPK1 $\Delta$ NS1 is the intimate interaction between its kinase core and the inserts of the spacer domain (Figure 4.8a). The N-terminal spacer insert, which folds into a long helix ( $\alpha S1$ ), interacts with the small lobe whereas the C-terminal spacer insert, which folds into a  $3_{10}$  helix ( $\alpha S2$ ), interacts with the large lobe. The helix  $\alpha S1$  anchors the

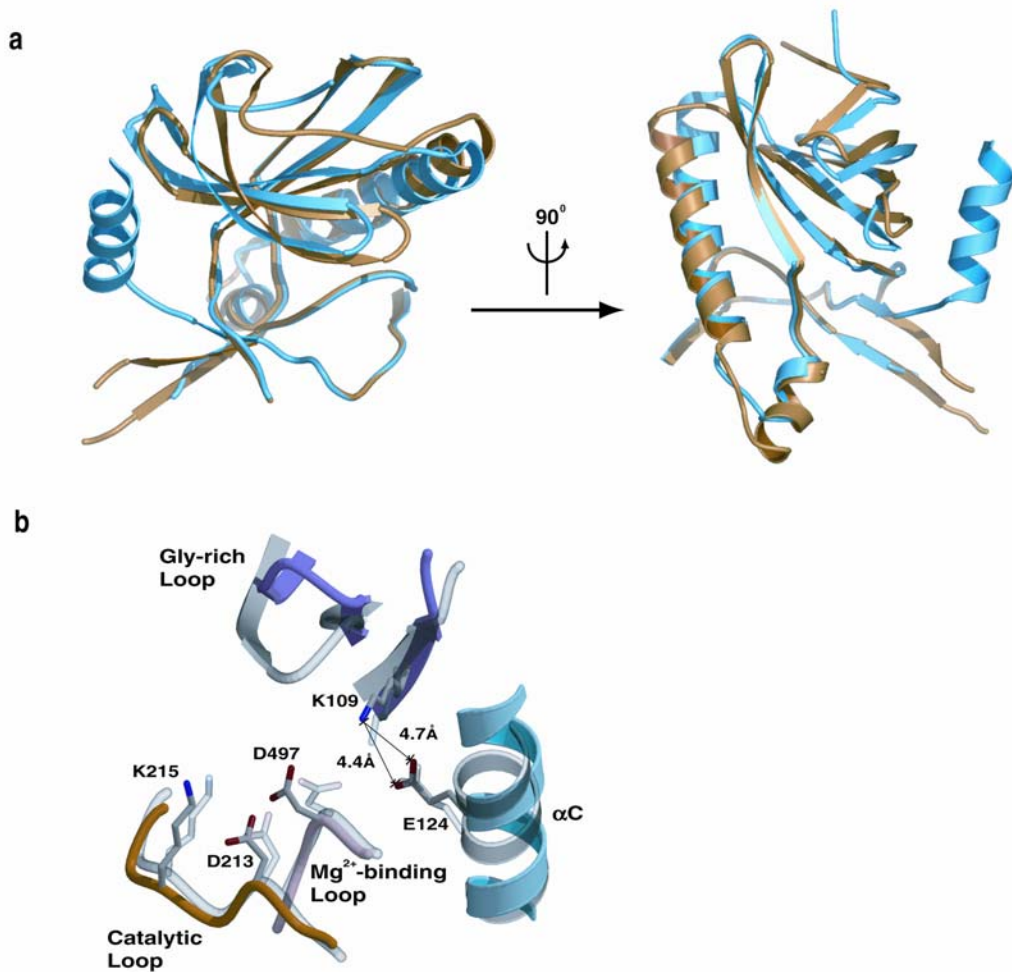


Figure 4.7 Small distortion of SRPK1 small lobe. a) Ribbon diagram of small lobe of SRPK1 (blue) overlays with that of Sky1p (gold). Crystal packing and local interactions cause the small lobe to rotate slightly toward helix  $\alpha$ S1, resulting in a distortion of the small lobe when compared to other active kinase structures. b) Catalytic residues of SRPK1 structure are ordered and positioned correctly for catalysis except K109 and E124 which although ordered, are separated by  $\sim 4.6$  Å and require slight conformational adjustment to form the invariant ion pair. The corresponding residues of PKA (gray, transparent) are overlaid for comparison.

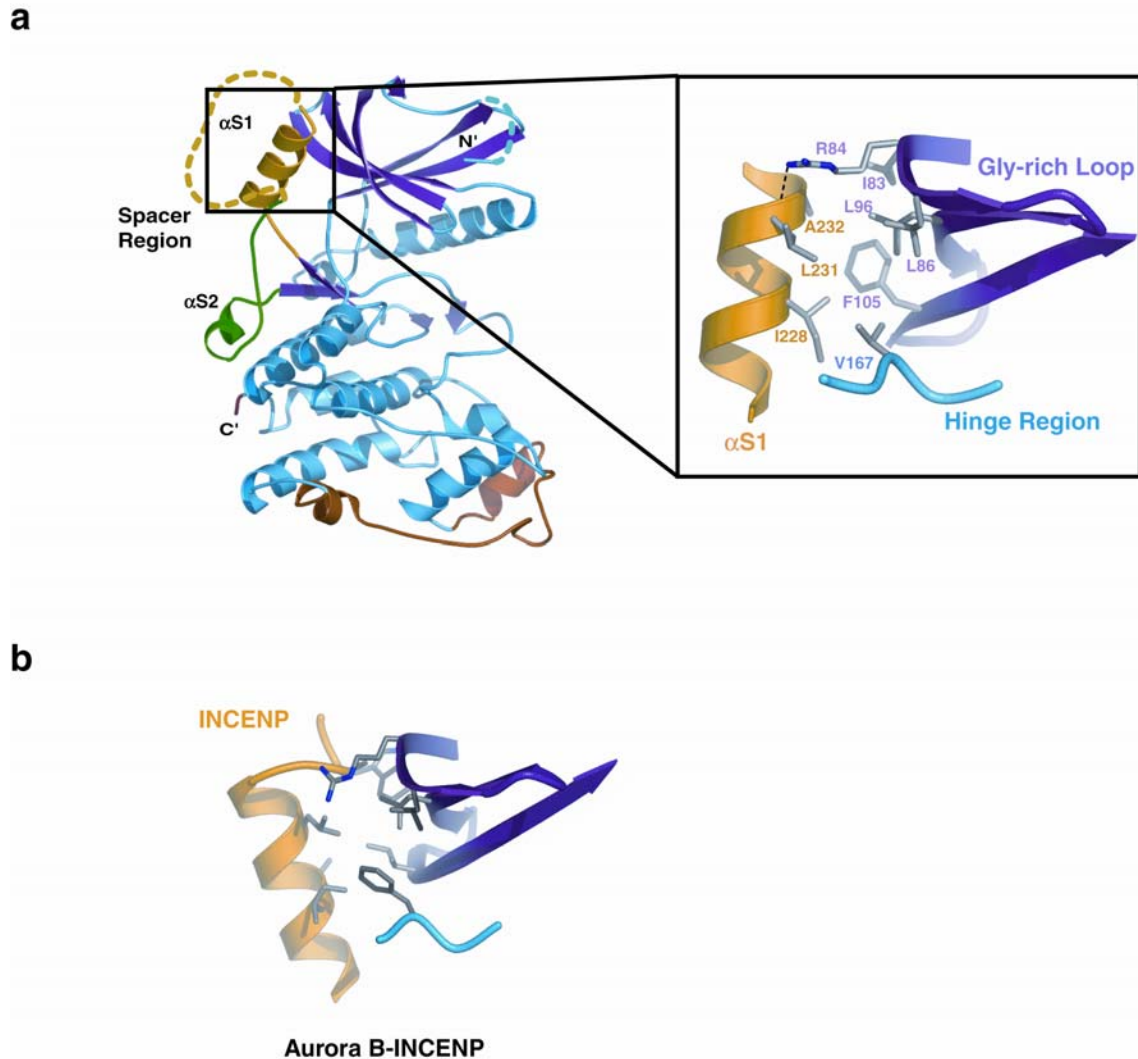


Figure 4.8 Spacer helix  $\alpha$ S1 resembles activator of Aurora B kinase. a) Ribbon diagram showing extensive hydrophobic interactions mediated by helices  $\alpha$ S1. Residues from the spacer helix, small lobe and the hinge region of the kinase are involved in the interactions. b) The hydrophobic interactions seen in INCEP/Aurora B kinase complex is structurally similar to the groove formed by the small lobe and large spacer helix in SRPK1 $\Delta$ NS1.

small lobe and the hinge region that links the two lobes by forming a deep and extended hydrophobic core (Figure 4.8a). Residues I228 of helix  $\alpha$ S1 and V167 of the hinge region are important for the formation of this core. It is apparent that such interactions restrict the relative mobility of the two lobes around the hinge and this hydrophobic core would also affect binding of the adenine moiety of the nucleotide. This intra-molecular interaction between  $\alpha$ S1 and the SRPK1 kinase core domain is remarkably similar to the intermolecular interaction observed in the case of aurora B kinase bound to INCENP, an activator of the kinase (Sessa et al., 2005). INCENP interacts with the kinase through its three helices allowing the inactive apo-enzyme to achieve its active conformation. Figure 4.8b show the detail intermolecular interactions between the small lobe of aurora B kinase and helix  $\alpha$ 1 of ICENP. Although the INCENP and  $\alpha$ S1 helices have opposite orientation, the hydrophobic core and the detailed interactions are very similar.

The helix  $\alpha$ S2 interacts with a hydrophobic groove in the large lobe formed by helices  $\alpha$ D and  $\alpha$ E, and the loop between  $\beta$ 7 and  $\alpha$ S1. This groove is stabilized by van der Waals interactions mediated by residues L479, V480, P482 and L483 from the loop that precedes helix  $\alpha$ S2 together with P485, A488 and L491 within the helix (Figure 4.9a). MAP kinases utilize an identical hydrophobic groove (the DEJL motif binding site) to dock their substrates, activators, inhibitors or scaffold proteins (Bhattacharyya et al., 2006; Chang et al., 2002; Lee et al., 2004) (Figure 4.9b). We refer to this kinase docking



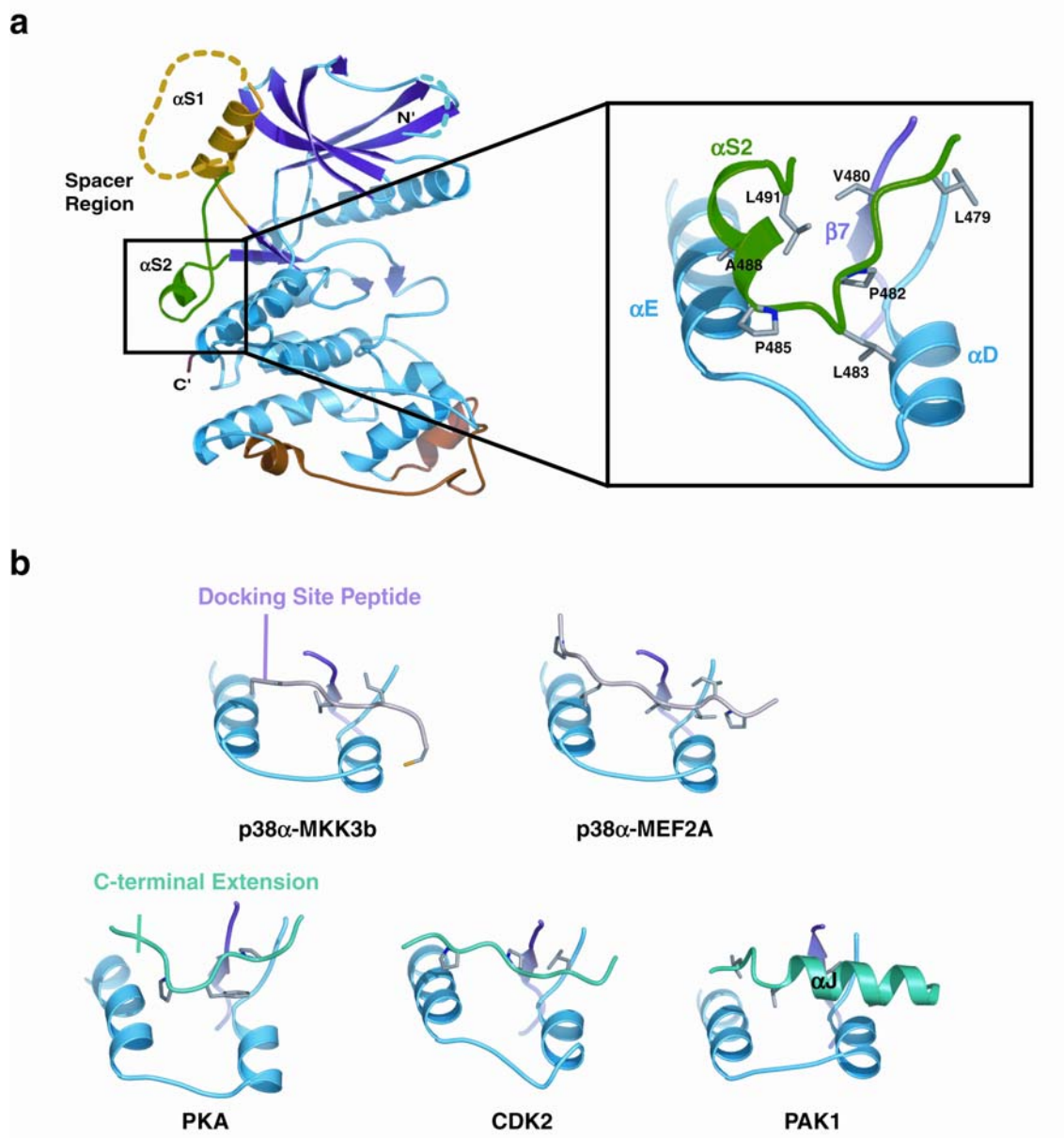


Figure 4.9 Spacer helix  $\alpha$ S2 is structural mimic of binding proteins of other kinases. a) Helix  $\alpha$ S2 interacts with a hydrophobic groove in the large lobe formed by helices  $\alpha$ D and  $\alpha$ E, and the loop between  $\beta$ 7 and  $\alpha$ S1. Hydrophobic residues involved in the interaction are indicated. b) The identical hydrophobic groove in the large lobe in other kinases is also involved in either inter- or intramolecular interactions. MAP kinase p38 $\alpha$  interacts with its substrate (MEF2A) or upstream activating enzyme (MKK3b) using this hydrophobic groove whereas PKA, CDK2 and PAK1 mediate intramolecular interactions.

groove as  $\alpha$ D/E $\beta$ 7/8 groove. Docking motifs of MKK3b and MEF2A, the activator and substrate of p38 $\alpha$  respectively, recognize and bind this hydrophobic pocket of p38 $\alpha$  with high specificity (Chang et al., 2002). The same groove is recognized by JIP1 (JNK-interacting protein-1) to dock onto JNK and inhibit its activity (Heo et al., 2004). Fus3, a MAP kinase in yeast also utilizes this groove to bind to the scaffold protein Ste5 (Bhattacharyya et al., 2006). Intriguingly, although the docking grooves are distal to the active sites of these kinases, the docking interactions significantly affect their catalytic activities. Comparison of the apo- and peptide-bound forms of these kinases revealed that there is significant hinge motion between the small and large lobes and also rearrangement of the activation loops (Chang et al., 2002; Heo et al., 2004).

The same docking groove is also involved in intra-molecular interactions with segments outside of the kinase core in several other kinases including PKA, CDK2 and PAK1 (Figure 4.9b) (De Bondt et al., 1993; Knighton et al., 1991a; Lei et al., 2000). In PKA, the C-terminal extension is thought to be important to maintain the kinase in an active conformation (Batkin et al., 2000). It thus appears that this site in protein kinases is a critical determinant of kinase regulation. In all cases, either intra- or inter-molecular, docking interactions induce allosteric conformational changes and affect inter-lobe motion and activation segment integrity. Our results thus suggest that SRPK1 has evolved to integrate its own activator within its own coding sequence, resembling the in-trans interactions observed in aurora B and MAP kinases.

## 5. Activation Segment of SRPK1

In most protein kinases, intra-loop side chains of the activation segment interact with each other in a phosphorylation dependent manner to maintain the activation loop, and in turn the whole activation segment, in an active conformation (Johnson et al., 1996; Nolen et al., 2004). For instance, in PKA, T197 within the activation loop is phosphorylated and is involved in electrostatic interactions with an arginine preceding the conserved aspartate of the catalytic loop and promotes the active conformation of the activation segment (Figure 4.10a). In contrast, the short activation loop of SRPK1 forms a tight  $\beta$ -turn stabilized by backbone-backbone interactions between residues within the loop, allowing it to adopt an active conformation without phosphorylation (Figure 4.10b). A glutamine residue, Q513, conserved within the SRPK family, further stabilizes the activation loop conformation. This residue assumes a strained backbone conformation ( $\phi = 70.7^\circ$ ,  $\psi = 141.7^\circ$ ) stabilized by an extensive hydrogen bonding network that links the catalytic loop and helix  $\alpha F$  with the P+1 loop to maintain a conformation compatible with substrate binding and catalysis. This unusual geometry is also seen in other known structures of active serine/threonine kinases that have a non-glycine residue at this position (ERK2, p38 $\gamma$ , CDK2, CK2, GSK3 $\beta$  and Sky1p). Furthermore, mutation of the glutamine residues at the corresponding positions in Sky1p and DYPK1A significantly reduce their catalytic activities, suggesting this residue plays an important role in maintaining kinase activity (Nolen et al., 2001; Wiechmann et al., 2003).

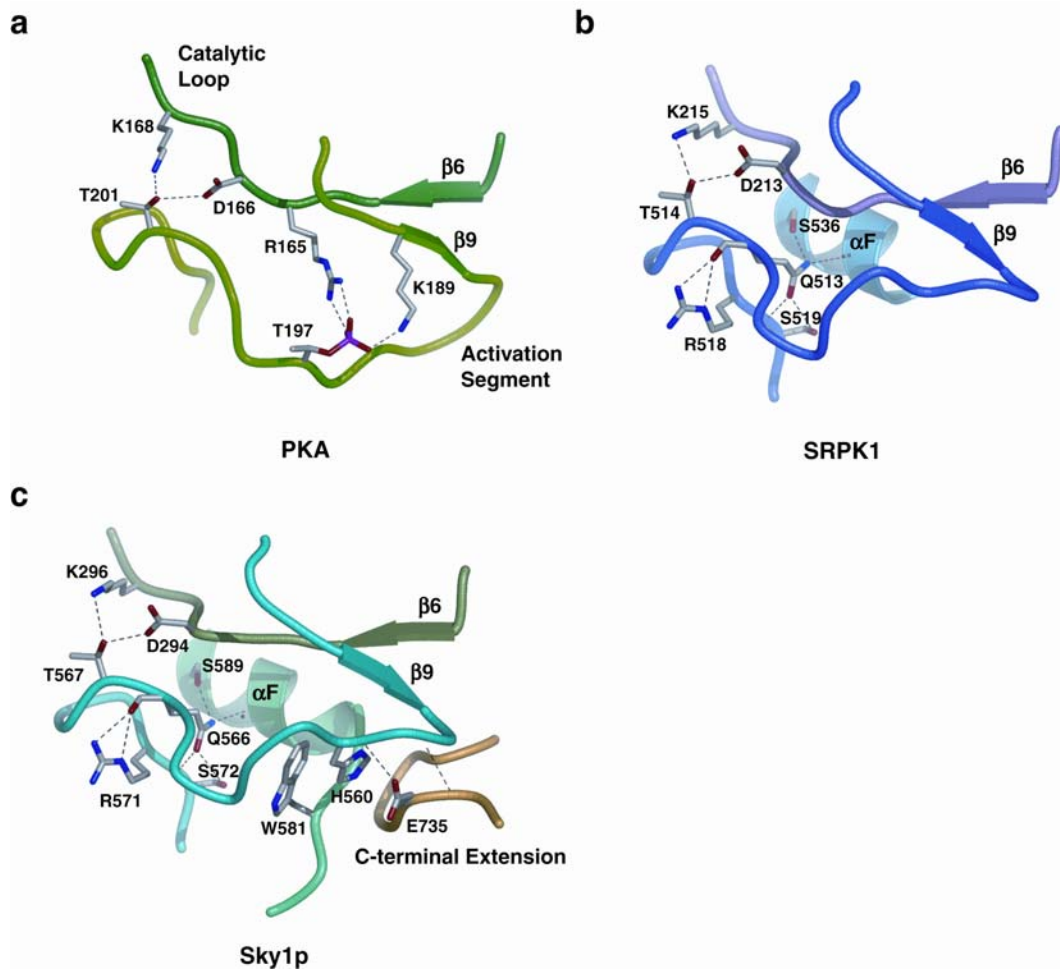


Figure 4.10 Activation loop conformation of SRPK1 and comparison with those of PKA and Sky1p. a) Active conformation of the PKA activation loop is achieved by phosphorylation at T197, which interacts with a cluster of basic residues at the active site, and forms a broad turn. Negatively and positively charged residues are denoted by red and blue, respectively. Hydrogen bonds are denoted by dotted lines. b) The activation loop of SRPK1 adopts a tight  $\beta$ -turn conformation and does not require phosphorylation. The presence of the  $\beta$ -sheet formed by  $\beta 6$  and  $\beta 9$  along with the critical interactions between T514, D213 and K215 suggest that the activation segment is correctly anchored in an active conformation. The side chain of Q513 is also involved in a network of interactions that further stabilizes the activation loop by assuming a strained backbone conformation. c) The activation loop of Sky1p highly resembles that of SRPK1 except that it makes additional contacts with the C-terminal extension denoted in yellow.

Comparison of the human and yeast SRPKs show that overall conformations of their activation segments are highly similar except one significant difference: a non-kinase core segment that interacts with Sky1p's activation loop is absent in SRPK1 (compare Figures 4.10b and 4.10c). In the model of Sky1p, the tip of its C-terminal extension lies underneath the activation loop and stabilizes it in the active conformation through hydrogen-bonding interactions (Figure 4.10c). These interactions were shown to be critical for the catalytic activity of Sky1p (Nolen et al., 2001). However, SRPK1 does not contain any C-terminal extension beyond helix  $\alpha$ I and its activation loop does not shown to interact with any other segment within the kinase. Surprisingly, the activation loop of SRPK1 appears to be as stable as that of Sky1p based on the overall fold and observed temperature factor of the loop residues. This unexpected stability of the activation loop prompted us to investigate the local side-chain interactions at the activation segment of SRPK1 in further details.

## **6. SRPK1-specific Interactions at the Activation Loop**

In the model of SRPK1 $\Delta$ NS1, the tip of the activation loop is mostly solvent exposed and multiple well ordered water molecules mediate a network of side chain-side chain and side chain-backbone interactions (Figure 4.11a). In particular, R208 from helix  $\alpha$ E interacts with N529 from helix  $\alpha$ F through water-mediated interactions. These interactions do not exist in Sky1p due to the placement of the C-terminal extension at this region. In fact, Sky1p contains glycines at the corresponding positions of R208 and N529 of SRPK1 in order to

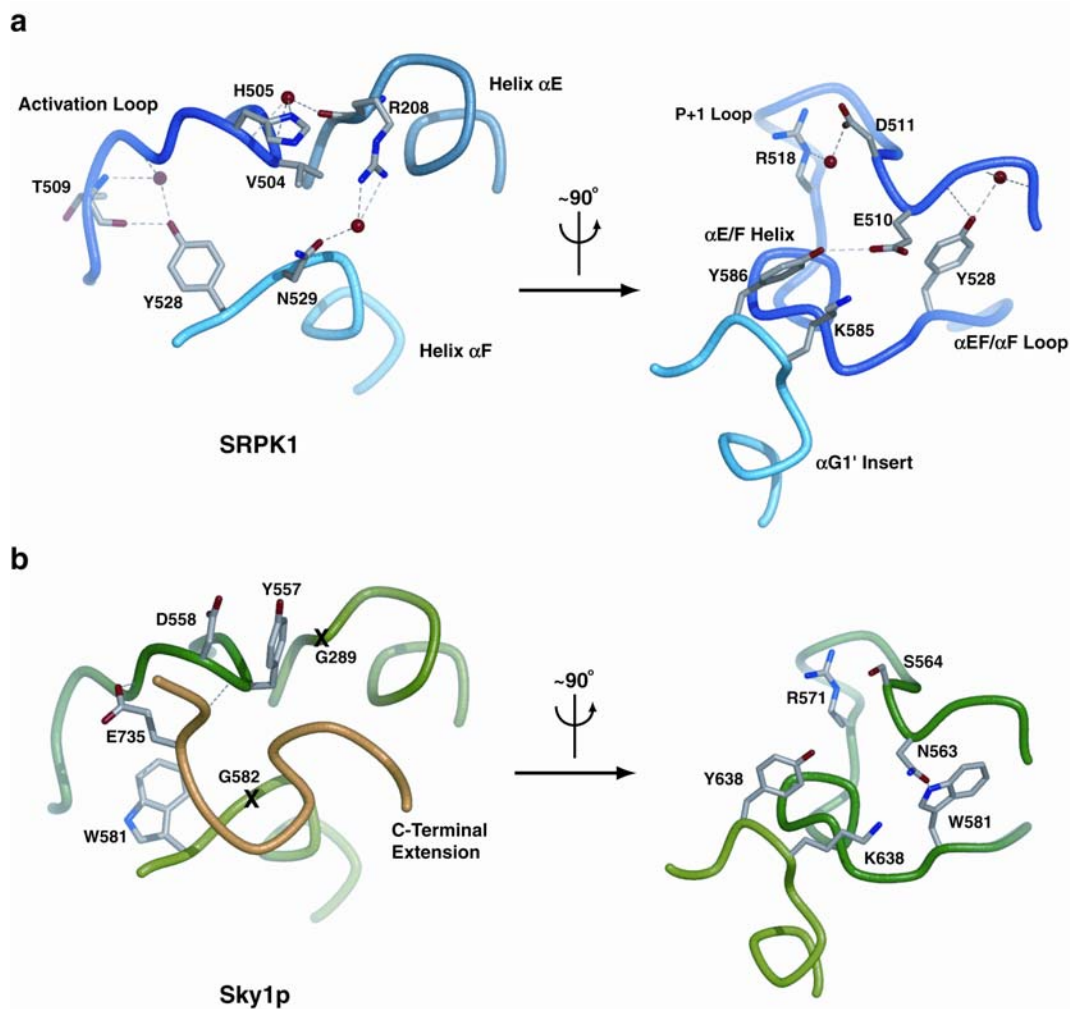


Figure 4.11 Unique SRPK1 interactions at the activation segment. a) SRPK1 contains multiple side-chains near or at the activation loop that mediate contacts not seen in Sky1p. Residues that were mutated in the M7 mutant are labeled in blue. b) Sky1p structure shows that the C-terminal extension directly contacts the activation loop. Due to the placement of the C-terminal extension, the corresponding residues of R208 and N529 in SRPK1 are glycines in Sky1p and are marked with crosses on the diagram.

prevent steric clashes (Figure 4.11b). Another water molecule in the model of SRPK1 $\Delta$ NS1 bridges H505 to the backbone of helix  $\alpha$ E and this interaction is prevented by Y557 in Sky1p. The hydroxyl group of Y528 from the  $\alpha$ EF/ $\alpha$ F loop of SRPK1 directly hydrogen bonds to the backbone carbonyl group of T509 and a water molecule further bridges this side chain to the backbone of the activation loop. In the active structures of PKA and ERK2, the corresponding tyrosines at this position also hydrogen bond to the backbone carbonyls of their phosphothreonines respectively, suggesting this residue might play an important role in stabilizing the active conformation of the activation loop. In contrast, Sky1p contains a tryptophan at this position and does not make any contact with the activation loop. Furthermore, E510 (N563 in Sky1p) of SRPK1 forms an ionic pair with K585 and hydrogen bonds to Y586 from the MAP kinase insert, and D511 (S564 in Sky1p) is bridged to the guanidinium group of R518 by another water molecule. These extensive SRPK1-specific intramolecular interactions might pose a solution to the constitutive activity of SRPK1. We speculated that the stability of the activation loop is either indirectly regulated from a distance by the docking interactions as mentioned before, or achieved by these unique intra-loop side chains interactions.

## **7. SRPK1 shows Unusual Resistance to Inactivation**

Since both spacer helices of SRPK1 appear to interact with the kinase core in ways that mimic the activators of other kinases, we wanted to investigate the importance of this region in kinase activity. To test this, we

mutated two residues (V167A and I228A) that appear to stabilize the interaction between spacer helices and the core of the kinase. However, the mutant is not defective in substrate phosphorylation (Figure 4.12a). Next we completely deleted both helices (SRPK1 $\Delta$ S) to test if they play any role and compared it to that of SRPK1 $\Delta$ S1 (Figure 4.12b). While the catalytic activity of SRPK1 $\Delta$ S reaches maxima within 15 minutes in our assay, SRPK1 $\Delta$ S1 continues to phosphorylate ASF/SF2 for a longer time (Figure 4.12c). The loss of activity of SRPK1 $\Delta$ S is not a result of the kinase precipitated out of solution. This observation suggests that SRPK1 $\Delta$ S becomes inactive overtime under the assay condition and fails to complete substrate phosphorylation. To further test the activities of these kinases under different assay conditions, we incubated both enzymes at 25 °C in conditions containing high or low content of salt respectively for different times before initiating the phosphorylation reaction. While the activity of SRPK1 $\Delta$ S1 remained the same even after 60 mins of incubation, SRPK1 $\Delta$ S lost most of its activity between 15 to 30 mins of incubation in the presence of low ionic strength buffer solution (Figure 4.12d). In contrast, at higher ionic strength condition, SRPK1 $\Delta$ S1 and SRPK1 $\Delta$ S showed similar activity. This suggests the inactivation of the kinase is due to the absence of stabilization from the spacer helices.

When we removed the N-terminal 41 residues from SRPK1 $\Delta$ S (Figure 4.12b), the mutant kinase lost most of its activity, whereas the same deletion in SRPK1 $\Delta$ S1 (the crystallized fragment SRPK1 $\Delta$ NS1) showed no effect (Figure 4.12e). Furthermore, the complete removal of the N-terminus beyond the



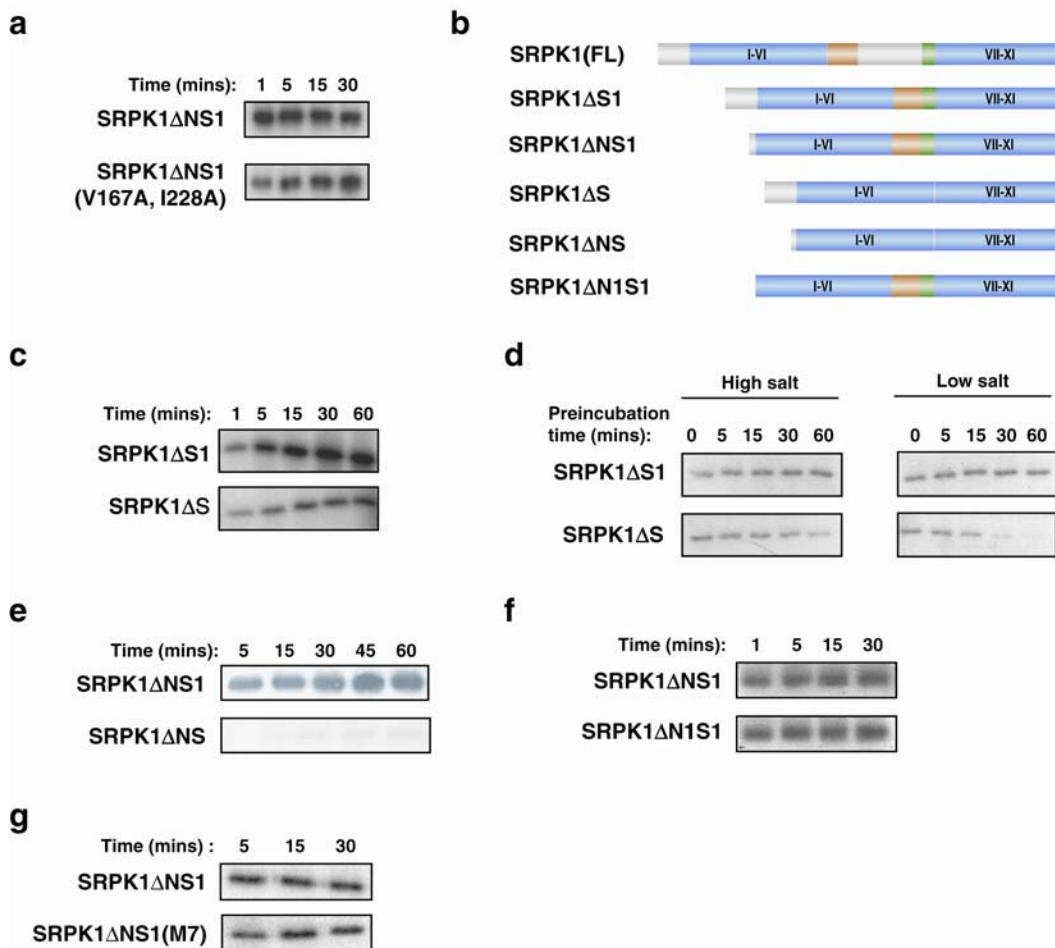


Figure 4.12 Spacer inserts mediated interactions are important for maintaining kinase activity. A) Mutant proteins containing two mutated residues (V167A and I228A) show similar catalytic activity as the wild type construct. b) Domain organization of SRPK1 and different constructs generated. N- and C-terminal spacer regions are colored yellow and green respectively. c) The catalytic activity of SRPK1 $\Delta$ S reaches maxima within 15 mins while SRPK1 $\Delta$ S1 continues to phosphorylate for longer time. d) Kinases are preincubated at room temperature for the indicated time period before performing activity assay and all reactions are carried out for 10 mins. Both SRPK1 $\Delta$ S1 and SRPK1 $\Delta$ S deletion mutants remain stable after 60 mins when the kinases are suspended in buffer that contains high content of salt. In the presence of low content of salt, SRPK1 $\Delta$ S1 retains most of its activity while SRPK1 $\Delta$ S starts to lose its catalytic activity after 15 mins. e) The spacer regions are essential for the catalytic activity of SRPK1 in the absence of first 41 residues of the N-terminus. f) The entire N-terminal extension beyond the kinase domain core is dispensable when the spacer regions are intact. g) Mutant SRPK1 (M7) contains 7 mutations around the activation segment does not show significant defect in catalytic activity suggesting that SRPK1 specific interactions at the activation segment may not play an important role in maintaining the constitutive activity.

kinase core in the presence of the spacer helices (SRPK1 $\Delta$ N1S1) did not reduce phosphorylation activity either (Figure 4.12f). These results suggest that the spacer helices show mild effects on kinase stability under normal assay conditions, however, in the absence of the N-terminus or at low ionic strength condition, the importance of the spacer helices in maintaining kinase activity become apparent. These observations are surprising; first, the spacer helices do not appear to stabilize only the activation loop but rather globally affect the kinase stability. Secondly, the N-terminal segment can compensate the stabilization function of the spacer helices.

Next we examined whether interactions mediated by segments flanking the activation loop have any effect on the catalytic activity. We reasoned that these interactions must be unique to SRPK1 and might play a role in preserving the kinase activity and function equivalent to the C-terminal extension of Sky1p. We mutated seven side chains that mediate unique interactions in SRPK1 to the corresponding residues in Sky1p to test their effects on the catalytic activity (R208G, V504Y, H505D, E510N, D511S, Y528W, N529G) (Figure 4.11). We created seven cumulative mutants (M1, M2... M7) where the first mutant (M1) has only one residue changed and each mutant thereafter contains one additional mutation. However, even the most drastic M7 mutant does not show significant defect in catalytic activity, suggesting that interactions mediated by these residues may not play an important role in maintaining the constitutive activity of SRPK1 (Figure 4.12g).

Together these results clearly indicated SRPK1 is a unique kinase that possess alternate networks of interaction to maintain the global stability of the kinase domain core as the N-terminus or spacer helices can compensate each other's stabilization function. Furthermore, the catalytically competent conformation of the activation loop of SRPK1 does not rely upon any direct stabilization mechanism observed for other active Ser/Thr kinases.

### **C. Discussion**

Protein kinases are the key regulatory components of most signaling pathway in eukaryotes. By phosphorylating their substrates, protein kinases can propagate and amplify extra- and intracellular signals and regulate diverse cellular processes such as apoptosis, transcription, and mRNA processing in the case of SRPK1. Given the important roles of kinases, the molecular mechanism of the regulation of these enzymes are of central importance in the study of signaling processes. The most well studied regulatory mechanisms of kinases involve phosphorylation of their activation loops or binding of regulatory elements. However, these regulatory mechanisms cannot be applied to SRPK1 as the kinase is constitutively active and does not require any post-translational modification. Furthermore, while some other kinases like CK2 and Sky1p utilize non-kinase core segments to stabilize their activation loops, the crystal structure of SRPK1 shows that these interactions are absent yet the enzyme maintains its catalytic activity (Niefind et al., 1998; Nolen et al., 2004). The molecular architecture of SRPK1 has also revealed that two helices from the

spacer region of SRPK1 interact with the back of the kinase core, and indeed, biochemical studies demonstrate that these spacer helices are important for the stabilization of SRPK1. However, this stabilization effect is not sufficient to explain the activity of SRPK1, since under specific conditions the spacer-deleted SRPK1 functions almost identical to the wild type kinase. We further show that the flexible N-terminus region can substitute for the function of the spacer helices. Surprisingly, residues involved in mediating a network of interactions surrounding the activation segment can also be substituted without having any effect on the catalytic activity.

### **1. Alternative Interactions at Activation Segment of SRPK1**

The lack of effect of the M7 mutant on catalysis was surprising because of the location and non-conservational nature of the side chain substitutions. In order to understand the molecular basis for these observations, molecular dynamic simulations on the M7 mutant was carried out by our collaborator Dr. Justin Gullingsrud. The M7 mutant simulations were consistent with the observation from our biochemical study, as the mutant activation loop adopted a stable structure close to the native form based on the computed RMSD. The mutant loop retained this high degree of stability in the simulations in part because contacts found in the native structure were replaced by contacts with nearby residues in the mutant (Table 4.2). Although the new contacts formed in the model of the mutant structure are less stable than in the wild type, they appear to be sufficient at maintaining a conformation similar to wild type. This

Table 4.2. Interactions at the activation segment of SRPK1

Original interactions in wild-type construct		Type of interaction
E510 side-chain	K585 side-chain	Ionic
Y528 side-chain	T509 backbone	H-bond
H505 side-chain	V504 side-chain	van der Waals
New interactions in mutant construct		Type of interaction
N510 side-chain	G527 backbone	H-bond
W528 side-chain	H507, Y504 side-chain	Stacking
D505 side-chain	K506 side-chain	H-bond

preponderance of compensatory interactions to preserve the active state of SRPK1 has not been observed in other kinases. Our studies thus describe a new mode of active state preservation of a protein kinase.

## **2. Spacer Helices Stabilize Hinge Motion between the Two Lobes**

To further understand the role of the spacer helices in the stability of SRPK1, molecular dynamic simulations on the spacer-deleted mutant were also performed. Simulations were initiated from the crystal structure of SRPK1 $\Delta$ NS1 in the presence or absence of an ATP molecule and two Mg<sup>2+</sup>. While the wild-type construct retains high stability regardless of the presence of ATP, the stability of the spacer deletion mutant appeared to depend on the presence of the nucleotide in the active site. The overall backbone RMSD of the spacer deletion mutant stabilized at the same level as the native system in the presence of ATP; however, without ATP the small lobe exhibited large hinge movements relative to the large lobe in the spacer-deleted mutant (Figure 4.13). This suggests that the spacer residues do not bias the relative position of the small lobe, but rather simply serve to stabilize the motion and conformation of the lobe.

## **3. Biological Role of the Spacer Domain**

Since the SRPK1 kinase domain is highly stable and remains constitutively active as described in this chapter, we wondered whether its regulation is achieved by its sub-cellular localization and the spacer domain

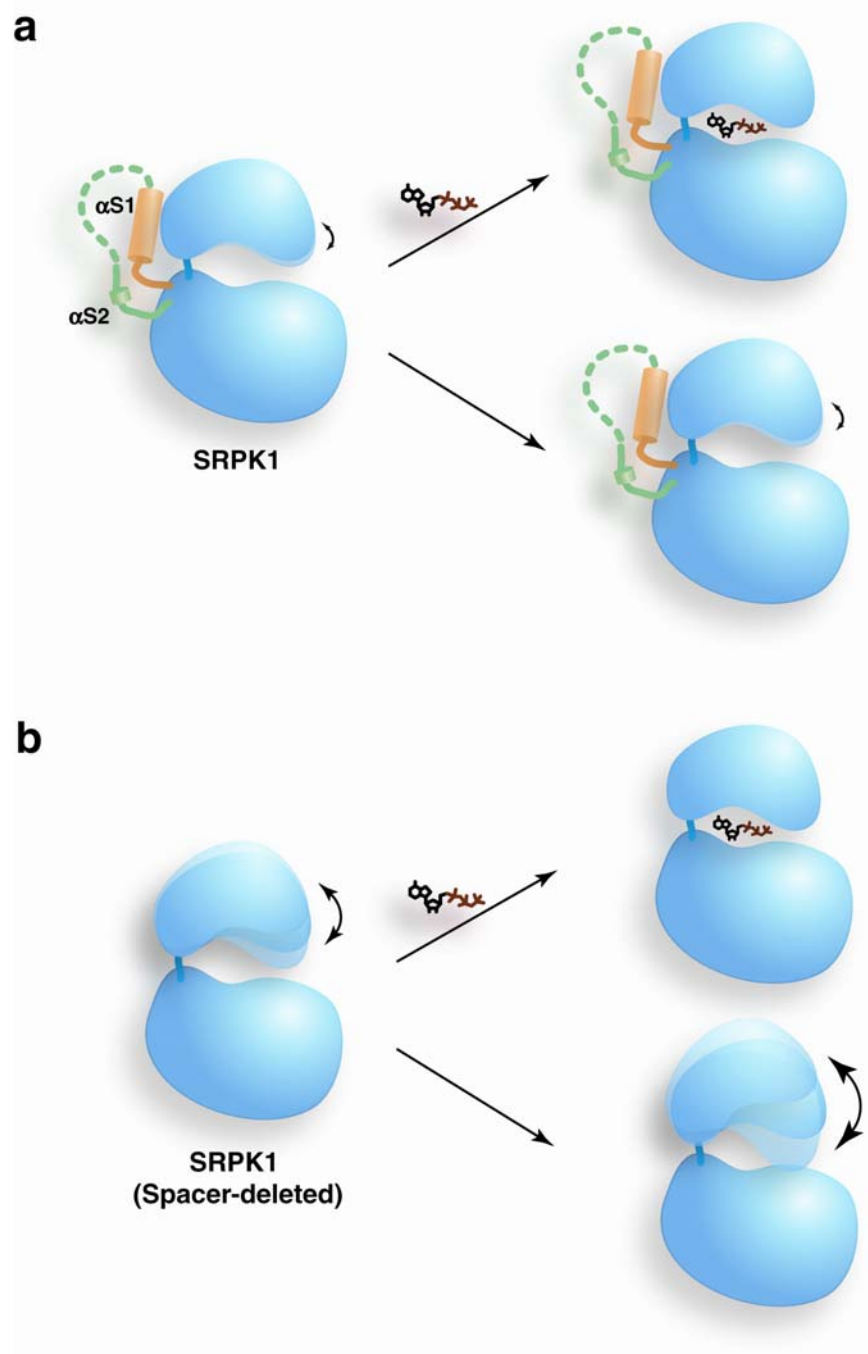


Figure 4.13 Spacer helices are important to stabilize the conformation of small lobe. a) In the presence of the spacer helices, SRPK1 remains stable in the presence or absence of ATP. b) When the spacer helices are truncated, the small lobe of the mutant protein exhibits large hinge motion, resulting in an unstable kinase.

might play an important role in this regard. Although the crystal structure of SRPK1 shows that the two helices from the spacer domain make intimate contacts with the kinase domain core, the rest of the spacer domain retained in the construct is disordered, suggesting that the full spacer domain can be folded into an independent module and would not affect the structural integrity of the kinase domain. Furthermore, in agreement with our speculation, a recent study has shown that the spacer domain of SRPK1 contains anchoring motifs that are responsible for its retention in the cytoplasm in interphase cell (Ding et al., 2006). Despite the fact that the protein(s) that is/are responsible for anchoring SRPK1 in the cytoplasm is unknown, the observation that SRPK1 is released into the nucleus during G2/M phases boundary suggests that its localization is cell cycle regulated. However, the signaling pathway that triggers the dissociation of the kinase from the anchoring protein is yet to be identified.



**Chapter V: The SRPK1/Peptide/ADP Ternary Complex  
Structure Reveals a Docking Groove and its  
Interactions with ASF/SF2**

## A. Introduction

Several human SR proteins are known to date, ASF/SF2 being one of the most well studied members of this family (Figure 5.1). ASF/SF2 is a modular protein containing two N-terminal RRM's followed by a short linker which connects it to the C-terminal RS domain. The RS domain of ASF/SF2 contains a tract of eight RS dipeptides (RS1 motif) followed by short interrupted stretches of RS dipeptides (RS2 motif) (Figure 5.1a). Previous studies have shown a precise requirement for distinct phosphorylated states of ASF/SF2 for its different biological functions, including spliceosome assembly, splicing activity, mRNA export and translation (Cao and Garcia-Blanco, 1998; Huang et al., 2003; Huang and Steitz, 2001; Huang et al., 2004; Lai and Tarn, 2004; Shen and Green, 2004; Shen et al., 2004; Tange and Kjems, 2001). Phosphorylation of ASF/SF2 is also essential for its own nuclear import (Duncan et al., 1998; Kataoka et al., 1999; Lai et al., 2001). In the nucleus, ASF/SF2 and most other splicing factors are concentrated in distinct nuclear bodies called nuclear speckles, the formation of which is dependent on the degree of phosphorylation of these proteins (Cazalla et al., 2002; Colwill et al., 1996; Lamond and Spector, 2003; Misteli et al., 1998).

The RS domain of ASF/SF2 is phosphorylated by SR protein kinases (SRPKs) and Clk/Sty kinases. SRPKs are serine-specific kinases where they only accept serine at the phosphorylation site (P0), while most serine/threonine kinases will accept threonine as well. Furthermore, mutations of the arginines that flanked the P0 residue to lysines abolish the phosphorylation of the

substrate, suggesting that substrate recognition depends on more than general charge complementarity (Gui et al., 1994b). However, the X-ray crystal structure of SRPK1 apoenzyme alone does not clarify how SRPK1 acquires such remarkable specificity/selectivity.

Recently, SRPK1 has been shown to interact with ASF/SF2 with high affinity ( $K_d = 50$  nM) phosphorylating approximately half the sites in its RS domain in a processive manner in vitro, where the kinase locks onto the substrate and does not dissociate until phosphorylation is completed (Aubol et al., 2003). It is unknown as to why ASF/SF2 binds SRPK1 with high affinity and is phosphorylated at only half of its available serines. The biological significance of such restricted phosphorylation is also unknown.

Unlike SRPK1, Clk/Sty belongs to a family of dual specificity kinases which phosphorylate serine, threonine as well as tyrosine residues (Duncan et al., 1995; Prasad and Manley, 2003). Clk/Sty is a nuclear kinase and thus appears to regulate SR proteins in the nucleus through phosphorylation (Duncan et al., 1998; Hartmann et al., 2001; Misteli et al., 1998; Prasad et al., 1999). It is thought to phosphorylate significantly more residues in the RS domain of ASF/ SF2 than SRPK1 (Colwill et al., 1996; Prasad et al., 1999). Since transient overexpression of Clk/Sty causes release of ASF/SF2 from nuclear speckles, phosphorylation of ASF/SF2 by Clk/Sty was thought to serve as a means of recruiting splicing factors when required. However, how SRPK1 and Clk/Sty coordinate their activities to regulate SR proteins was not clear.

We reasoned that a high resolution X-ray structure of SRPK1 in complex to a substrate peptide will likely illustrate how SRPK1 selects against a single methyl group on the P0 residue and selects arginines against lysines at the flanking positions. Moreover, such complex might elucidate the molecular basis of the restricted and processive phosphorylation of ASF/SF2 by SRPK1. In this chapter, we describe the X-ray crystal structure of human SRPK1 bound to a peptide derived from an SR protein. Serendipitously, instead of binding to the substrate binding site of SRPK1, this peptide mimics a docking motif of ASF/SF2 and interacts with a docking groove of SRPK1. We show that the docking groove-docking motif interaction is critical for the restricted phosphorylation of ASF/SF2 by SRPK1. Our results also suggest that the docking motif of ASF/SF2 is a key regulatory element for sequential phosphorylation by SRPK1 and Clk/Sty and thus essential for its subcellular localization.

## **B. Results**

### **1. Rational Design of SRPK1 Substrate Peptide**

SRPK1 processively phosphorylates its substrate ASF/SF2 at a stretch of eight RS dipeptides repeats (RS1) within the RS domain (Figure 5.1a)(Aubol et al., 2003). However, a substrate peptide of multiple RS dipeptides poses significant difficulties for crystallization. As every serine present within the peptide could potentially be the P0 residue, formation of the kinase/peptide complex could result in a mixed pool of kinases bound to the peptides at

different positions and will hinder the crystallization process or provide poorly diffracting crystals. Therefore, we reasoned that an optimal substrate peptide for crystallization purpose should contain a single phosphorylation site. SRPK1 can phosphorylate the yeast SR-like protein, Npl3p, in vitro and functionally complement Sky1p in vivo (Siebel et al., 1999; Yeakley et al., 1999). Unlike the SR proteins, Npl3p is phosphorylated at one RS dipeptide near the C-terminus (Figure 5.1b). Therefore, a peptide derived from Npl3p could serve as a potential substrate peptide for crystallization. Based on the proposed model of Sky1p complex to an Npl3p peptide and the electrostatic surface potential of SRPK1 $\Delta$ NS1 model, we proposed that the two most critical residues for substrate recognition are the arginines at P-3 and P+3 positions, and additional basic residues at positions -4, -5 and -6 would enhance substrate binding (Figure 5.2). We designed a 9mer substrate peptide (RRRERSPTR) which was derived from the single phosphorylation site on Npl3p (residues 408 – 414 with S411 being the P site serine) and modified by adding two arginines at the N-terminus (Figure 5.1b).

## **2. Formation and Data Collection of SRPK1 $\Delta$ NS1/ Peptide/ADP Complex**

### **Crystals**

We attempted to obtain crystals of SRPK1/nucleotide/ peptide complex by soaking the 9mer peptide into the crystals of apo-SRPK1 $\Delta$ NS1 in the presence of ADP or AMP-PNP (a non-hydrolyzable analog of ATP). Crystals of apo-SRPK1 $\Delta$ NS1 were obtained as described in the previous chapters. The

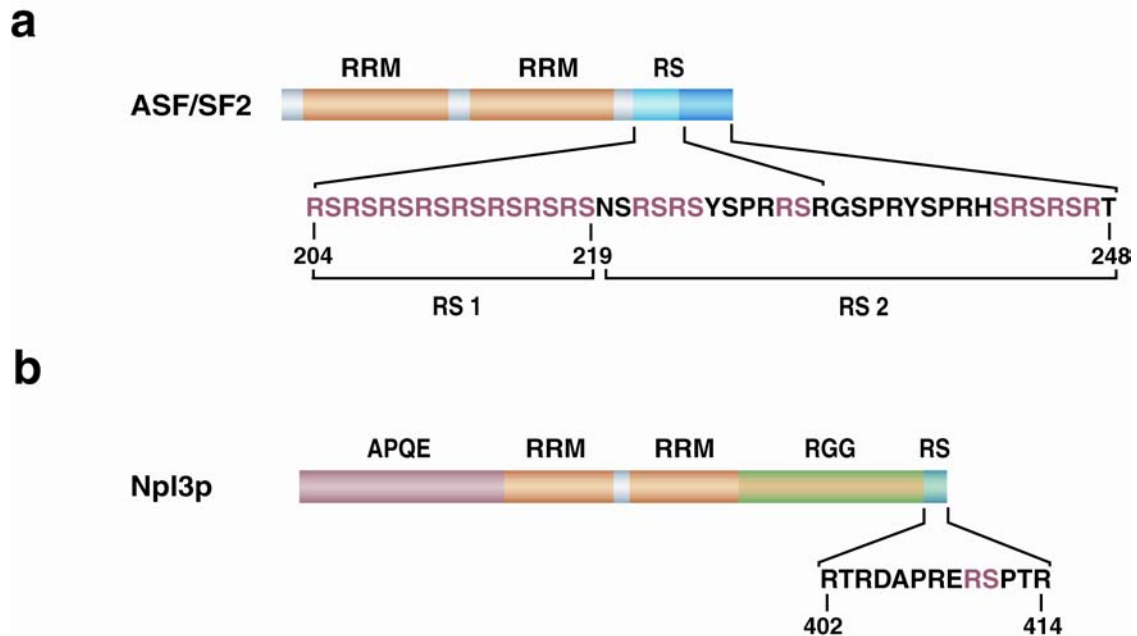


Figure 5.1 Domain organization of ASF/SF2. a) ASF/SF2 contains two RRM domains followed by an RS domain at the C-terminus. The RS domain of ASF/SF2 can be further divided into two subdomains: RS1 contains eight RS dipeptides in tandem and RS2 contains multiple discontinuous RS dipeptides. b) The RS domain of Npl3p, an SR-like protein in yeast, contains only one RS dipeptide at the C-terminus and is phosphorylated by Sky1p.

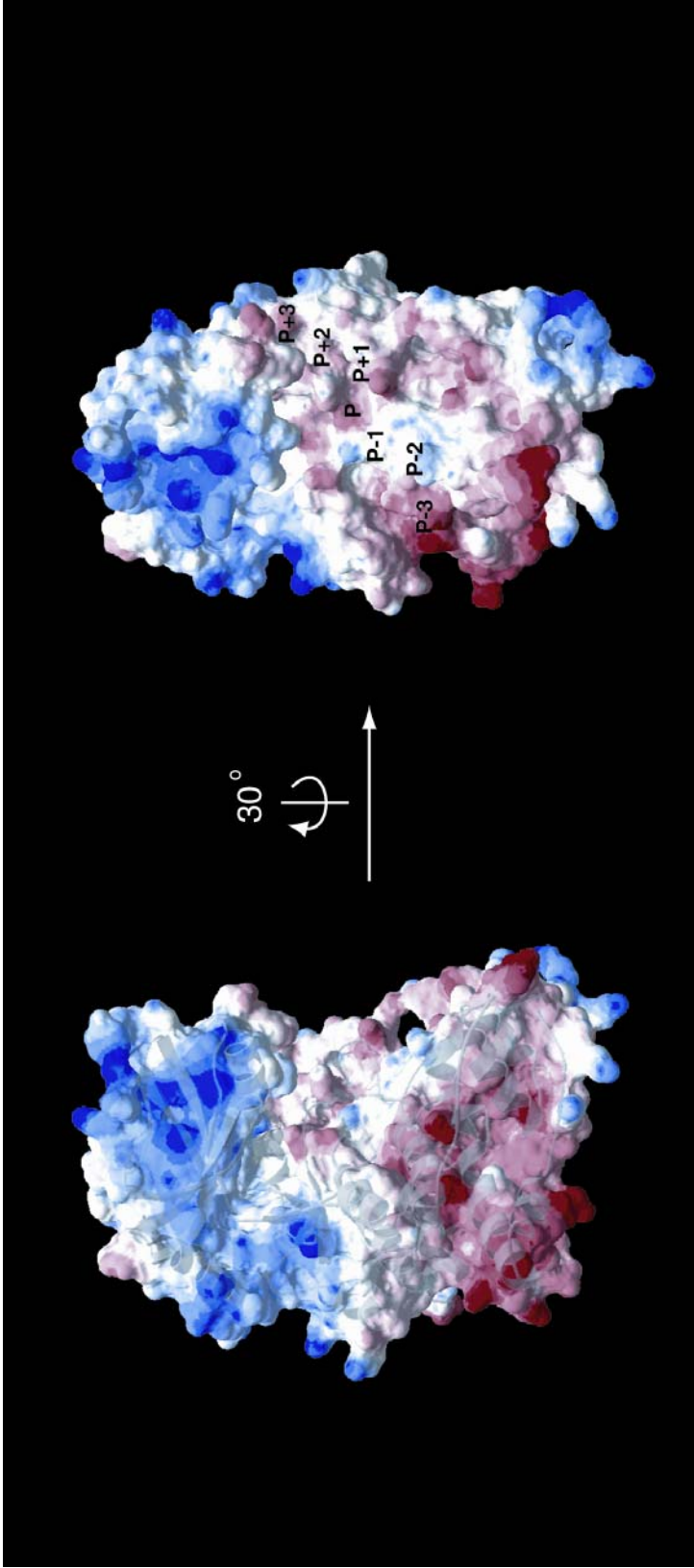


Figure 5.2 Electrostatic surface potential of SRPK1 $\Delta$ NS1. View on the left is from the front of the kinase. After rotated by  $\sim 30^\circ$ , the substrate binding groove is exposed and shown on the right. The corresponding P-3 to P+3 positions of an Npl3p peptide modeled into Sky1p previously (ref) are indicated. Red denotes negative potential and blue is positive. The negative potential beyond the P-3 position suggest that basic residues at P-4 to P-6 positions of the substrate peptide might enhance binding.

apoenzyme crystals were then transferred into a soaking buffer contained 17% PEG 3350, 100 mM sodium citrate (pH 5.6), 200 mM sodium acetate, 20% ethylene glycol, 10 mM MgCl<sub>2</sub>, 5 mM ADP or AMP-PNP and 5 mM peptide. The soaking process was performed for 24 hrs at 18 °C, and the crystals were directly flash frozen in liquid nitrogen after the soaking period. X-ray diffraction data of these crystals were collected at SBC-CAT synchrotron beamline ID-19 of the Advanced Photo Source (APS) at Argonne National Laboratory. The peptide-soaked crystals diffracted to lower resolutions than the apoenzyme crystals (Table 5.1). All crystals were identified to the same space group (P6<sub>5</sub>22) as the apoenzyme crystal. While most peptide-soaked crystals shared similar unit cell dimensions as the apo-crystal, two crystals showed deviations (highest resolution at 2.4 Å and 3.2 Å respectively) (Table 5.1). An electron density map was generated for each collected data set to confirm the presence of the peptide and/or nucleotide. Only the two data sets with unit cell dimensions different than those of the apo-enzyme crystals showed good electron density for both nucleotide and peptide. The crystals for both data sets were prepared similarly by soaking the 9mer peptide along with ADP, therefore we only used the 2.4 Å resolution data set to determine the structure of SRPK1ΔNS1/ADP/9mer ternary complex (Table 5.2).

### **3. Structure Solution of SRPK1ΔNS1/Peptide/ADP Ternary Complex**

The structure of SRPK1ΔNS1/peptide/ADP ternary complex was initially solved by molecular replacement using the coordinates of apo-SRPK1ΔNS1 as



Table 5.1: Summary of Data Collection of 9mer soaked crystals

Crystals	Highest Resolution (Å)	Space Group	Unit cell dimension (Å)
apo	1.73	P6 <sub>5</sub> 22	a = b = 75.11, c = 313.33
9mer1	2.4	P6 <sub>5</sub> 22	a = b = 78.68, c = 310.54
9mer2	3.2	P6 <sub>5</sub> 22	a = b = 78.68, c = 310.58
9mer3	2.76	P6 <sub>5</sub> 22	a = b = 76.11, c = 312.64
9mer4	2.45	P6 <sub>5</sub> 22	a = b = 75.92, c = 314.70
9mer5	2.95	P6 <sub>5</sub> 22	a = b = 75.94, c = 312.62
9mer6	2.85	P6 <sub>5</sub> 22	a = b = 76.21, c = 314.88
9mer7	2.3	P6 <sub>5</sub> 22	a = b = 75.01, c = 313.96

Table 5.2: Data Collection and Refinement of SRPK1/peptide/ADP complex crystal

Crystal	SRPK1 $\Delta$ NS1/peptide/ADP
Data Collection	
Data Collection Source	APS ID-19
Wavelength (Å)	1.0332
Resolution (Å)	30-2.4
No. of measured reflections	255743
No. of unique reflections	22765
Completeness (outer shell) (%)	97.8 (84.0)
I/ $\sigma$ (overall / outer shell)	8.1 (2.86)
R <sub>sym</sub> <sup>a</sup> (overall / outer shell) (%)	7.2 (56.5)
Refinement	
Resolution (Å)	30-2.4
R <sub>crys</sub> <sup>b</sup> (%)	22.84
R <sub>free</sub> <sup>c</sup> (%)	24.71
R.m.s. deviations	
Bond lengths (Å)	0.0085
Bond angles	1.383

$$^a R_{\text{sym}} = \frac{\sum |I - \langle I \rangle|}{\sum I}$$

$$^b R_{\text{crys}} = \frac{\sum ||F_{\text{obs}}| - |F_{\text{calc}}||}{\sum |F_{\text{obs}}|}$$
, where  $F_{\text{obs}}$  and  $F_{\text{calc}}$  are the observed and calculated structure factors, respectively.

<sup>c</sup>R<sub>free</sub> was calculated with 5% of the data excluded from the refinement calculation.

the search model, yielding clear rotation function and translation peaks (Figure 5.3). After rigid body refinement, the  $F_o-F_c$  electron density map calculated with the model revealed large positive peaks and readily interpretable density for the peptide and ADP (Figure 5.4). After model building of the peptide, the structure was refined with several cycles of manual refitting and refinements ( $R_{\text{crys}} = 22.84\%$ ,  $R_{\text{free}} = 24.71\%$ ). The final model of the ternary complex includes residues 67-237 and 477-655 of the kinase, residues 3-9 of the peptide (RERSPTR), 1 molecule of ADP and 1 molecule of acetate. Although  $\text{MgCl}_2$  is present in the soaking buffer, no  $\text{Mg}^{2+}$  ion was modeled in the structure due to poor electron density.

#### **4. Overall Structure of SRPK1 $\Delta$ NS1/Peptide/ADP Complex**

The overall structure of the kinase in the complex is almost identical to apo-SRPK $\Delta$ NS1 structure (Figure 5.5). The ternary structure reveals that the nucleotide binds to a conserved nucleotide binding cleft between the two lobes and the peptide binds in an extended conformation to a groove within the large lobe of kinase. Due to crystal contacts, the width of the nucleotide binding cleft does not change upon ADP binding and the ADP interacts only with the small lobe of the kinase. The adenine ring of ADP makes van der Waals contacts with three residues (L85, V93 and A106) present in the small lobe of SRPK1, which form the hydrophobic lining at the upper cleft of the nucleotide-binding pocket. In contrast to other nucleotide-bound kinase structures, the adenine ring in SRPK1 does not interact with the conserved hydrophobic residue (L219) from

```

Data line--- FITFUN NMOL 1 RESOLUTION 9 4
Data line--- CRYSTAL ORTH 1
Comment line--- # fixme: caveat about y or z refinement for polar SGs
Data line--- REFSOL AL BE GA X Y Z BF
Comment line--- # Here are the solutions to fit:
Data line--- SOLUT_1 1 16.17 78.76 310.54 0.6191 0.9064 0.2696 56.9 44.2 57.4 1 28.8
Data line--- SOLUT_1 1 16.17 78.76 310.54 0.6186 0.9067 0.2249 34.0 52.6 36.2 7 30.3
Data line--- SOLUT_1 1 16.17 78.76 310.54 0.6190 0.9059 0.3026 33.9 52.5 35.9 9 41.8
Data line--- SOLUT_1 1 16.17 78.76 310.54 0.6185 0.9056 0.4671 33.8 53.1 38.2 2 24.1
Data line--- SOLUT_1 1 16.17 78.76 310.54 0.6191 0.9063 0.4165 33.7 52.8 36.0 6 15.2
Data line--- SOLUT_1 1 16.17 78.76 310.54 0.6191 0.9056 0.1614 32.9 53.2 36.8 4 42.7
Data line--- SOLUT_1 1 16.17 78.76 310.54 0.6183 0.9066 0.2525 32.8 53.3 36.1 10 26.0
Data line--- SOLUT_1 1 16.17 78.76 310.54 0.6200 0.9069 0.4845 32.7 53.4 36.3 3 15.9
Data line--- SOLUT_1 1 16.17 78.76 310.54 0.6194 0.9063 0.4904 32.3 53.7 35.8 5 14.1
Data line--- SOLUT_1 1 16.17 78.76 310.54 0.6207 0.9071 0.2012 31.9 53.6 35.6 8 40.1

```

Figure 5.3 Phasing of SRPK1 $\Delta$ NS1//peptide/ADP ternary complex structure by molecular replacement. Molecular replacement using apo-SRPK1 $\Delta$ NS1 structure as a searching model yielded a clear solution as indicated by the significant improvement of R-factor (red dotted box) and correlation factor (blue dotted box).

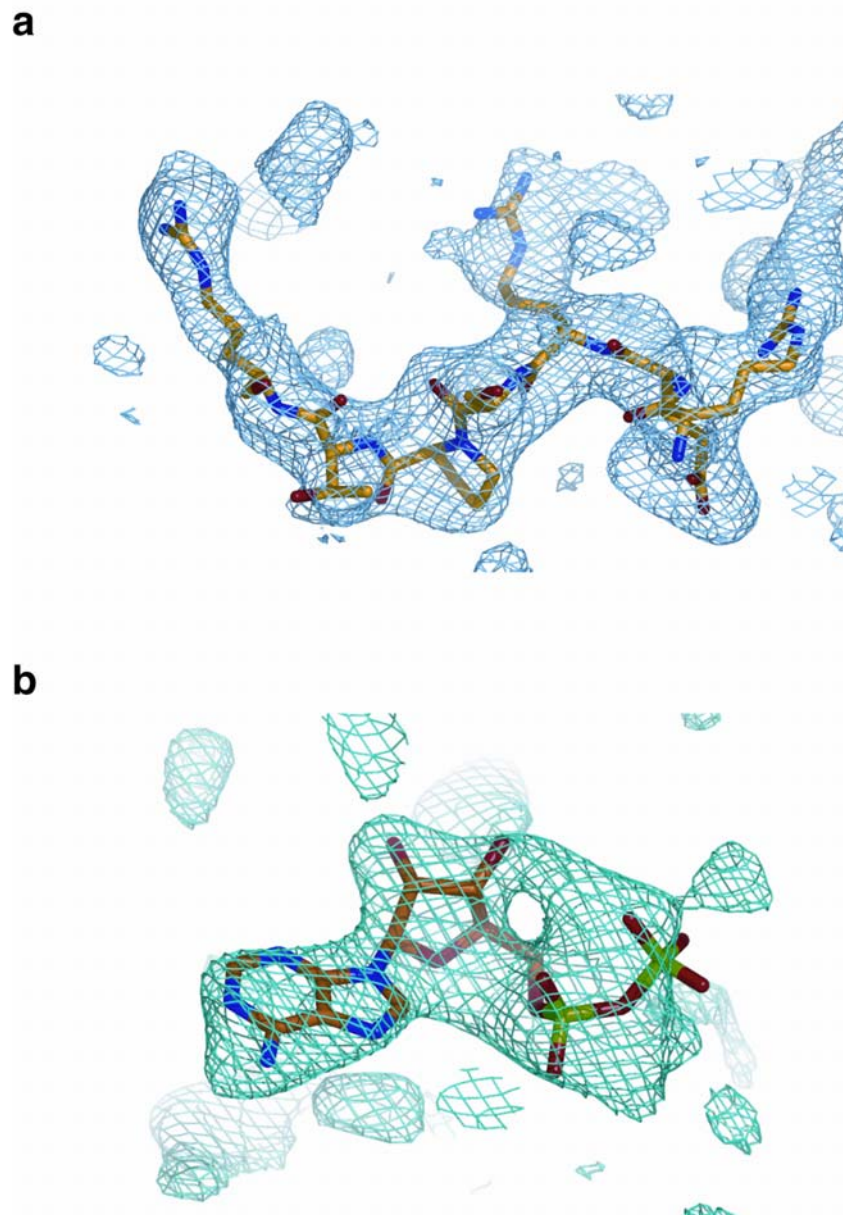


Figure 5.4 Presence of a peptide and ADP in the complex crystal. a)  $F_o-F_c$  electron density map of the peptide contoured at  $3\sigma$ . b)  $F_o-F_c$  electron density map of the ADP contoured at  $5\sigma$ .

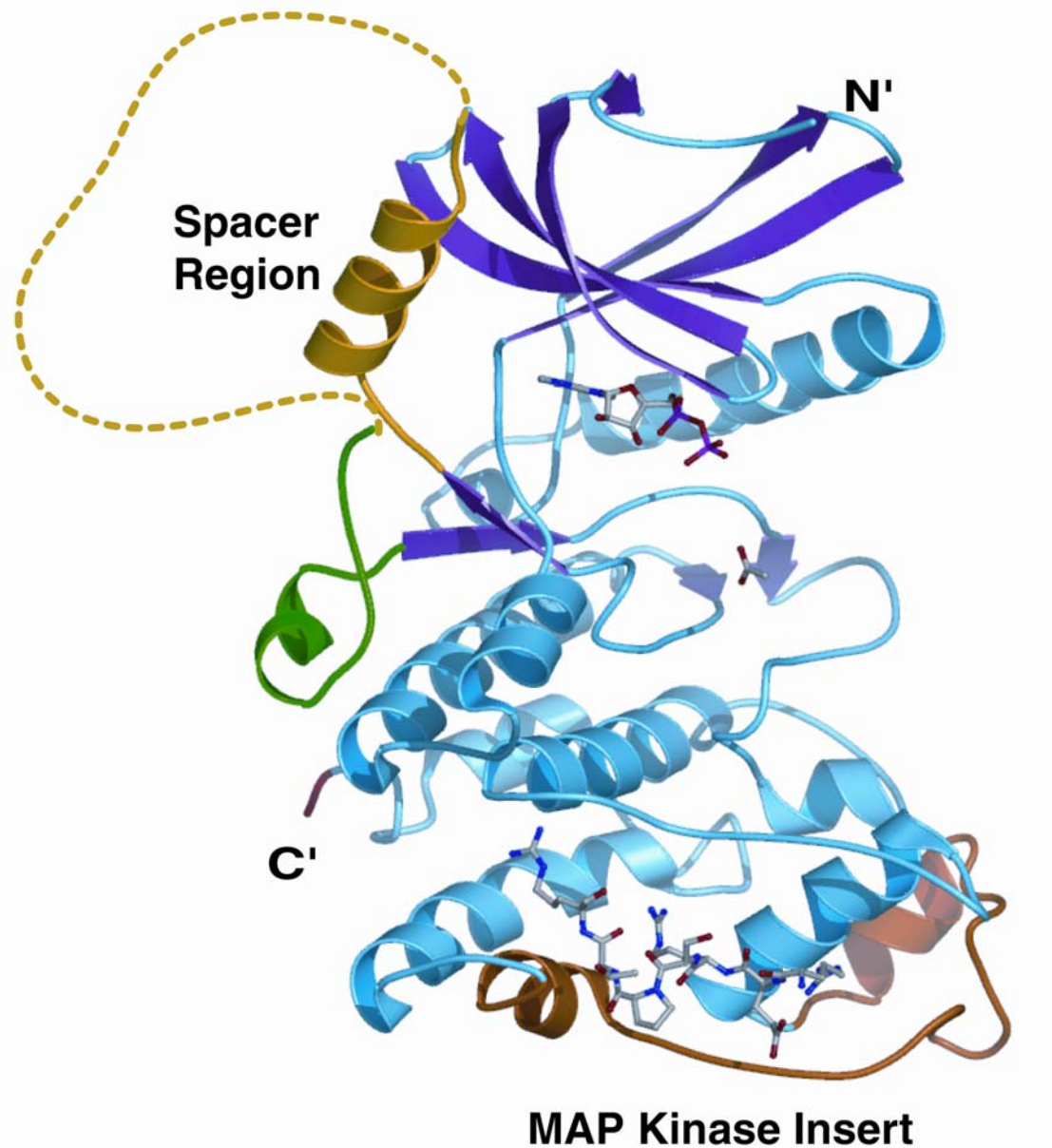


Figure 5.5 Ribbon diagram of SRPK1 $\Delta$ NS1/peptide/ADP tertiary complex X-ray structure. The overall structure of the kinase in the complex is similar to the apo-SRPK $\Delta$ NS1 structure. The ADP binds to the nucleotide binding pocket as in other nucleotide-bound kinase structures. The peptide is bound to a groove formed by the MAP kinase insert, the loop connecting helices  $\alpha$ F and  $\alpha$ G and the helix  $\alpha$ G

the bottom of the cleft (Figure 5.6).

Although the peptide was designed to bind to the active site groove, the electron density for the last seven residues (R<sup>1</sup>E<sup>2</sup>R<sup>3</sup>S<sup>4</sup>P<sup>5</sup>T<sup>6</sup>R<sup>7</sup>) of the peptide was not observed in this area. Instead, the strong electron density of the peptide was located at a distal site adjacent to the DEF motif binding site in ERK2 (Jacobs et al., 1999; Lee et al., 2004) (Figures 5.5 and 5.7a). The peptide binds into a deep groove in the kinase, which we refer to as the docking groove. The docking groove is formed by the MAP kinase insert, the loop connecting helices  $\alpha$ F and  $\alpha$ G and the helix  $\alpha$ G (Figure 5.7b).

## 5. Interactions at the Docking Groove of SRPK1

The kinase-peptide complex buries 1257 Å<sup>2</sup> of accessible surface area, a value typical of other peptide- protein complex structures. The structure reveals that the peptide docks in the docking groove through near-perfect charge complementarity, where the positively charged peptide neutralizes the negatively charged docking groove formed by residues D564 and E571 from helix  $\alpha$ G and D548 from the  $\alpha$ F/ $\alpha$ G loop of the kinase (Figure 5.7b). In addition to the carboxylates of these acidic residues, backbone carbonyls of L550 and T546 form hydrogen bonds with the guanidinium groups of the arginines (R1, R3 and R7) from the peptide. K615 from  $\alpha$ G2' of SRPK1 hydrogen bonds to the backbone carbonyl of S4 of the peptide and stabilizes the peptide backbone. Finally, the side chain of the peptide arginine R3 is buried in a deep hydrophobic pocket in SRPK1 and stabilized by van der Waals contact with

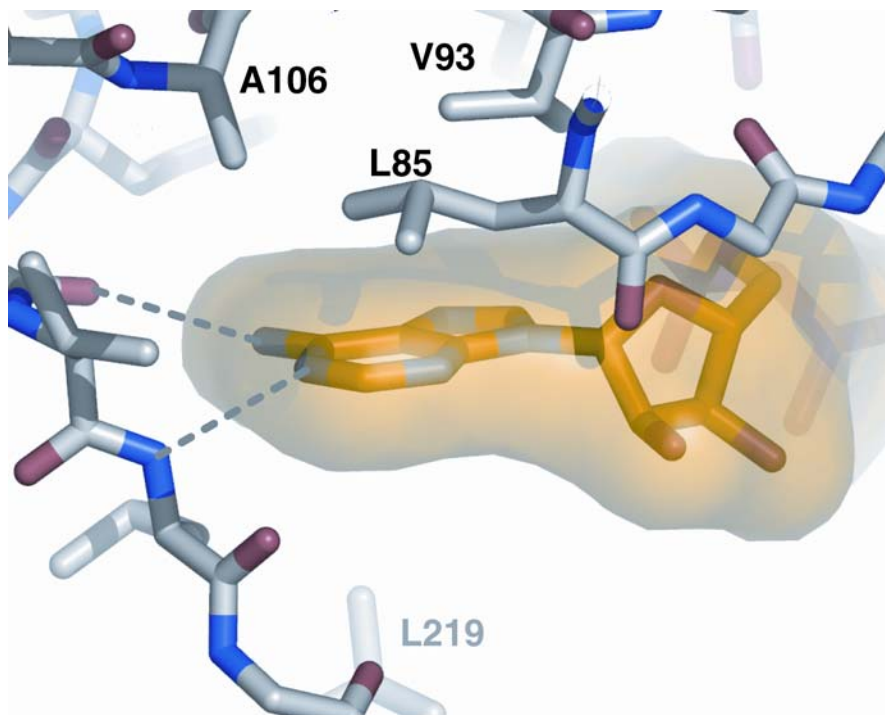


Figure 5.6 Nucleotide binding pocket of SRPK1 $\Delta$ NS1. In the complex structure, adenine ring of ADP hydrogen bonds to the backbone of the hinge region (indicated by the dotted lines) and makes van der Waals contacts with three hydrophobic residues (L85, V93 and A106) from the small lobe of SRPK1. However, another conserved hydrophobic residue, L219 from the large lobe, is too distal to make any direct contact with the adenine ring.



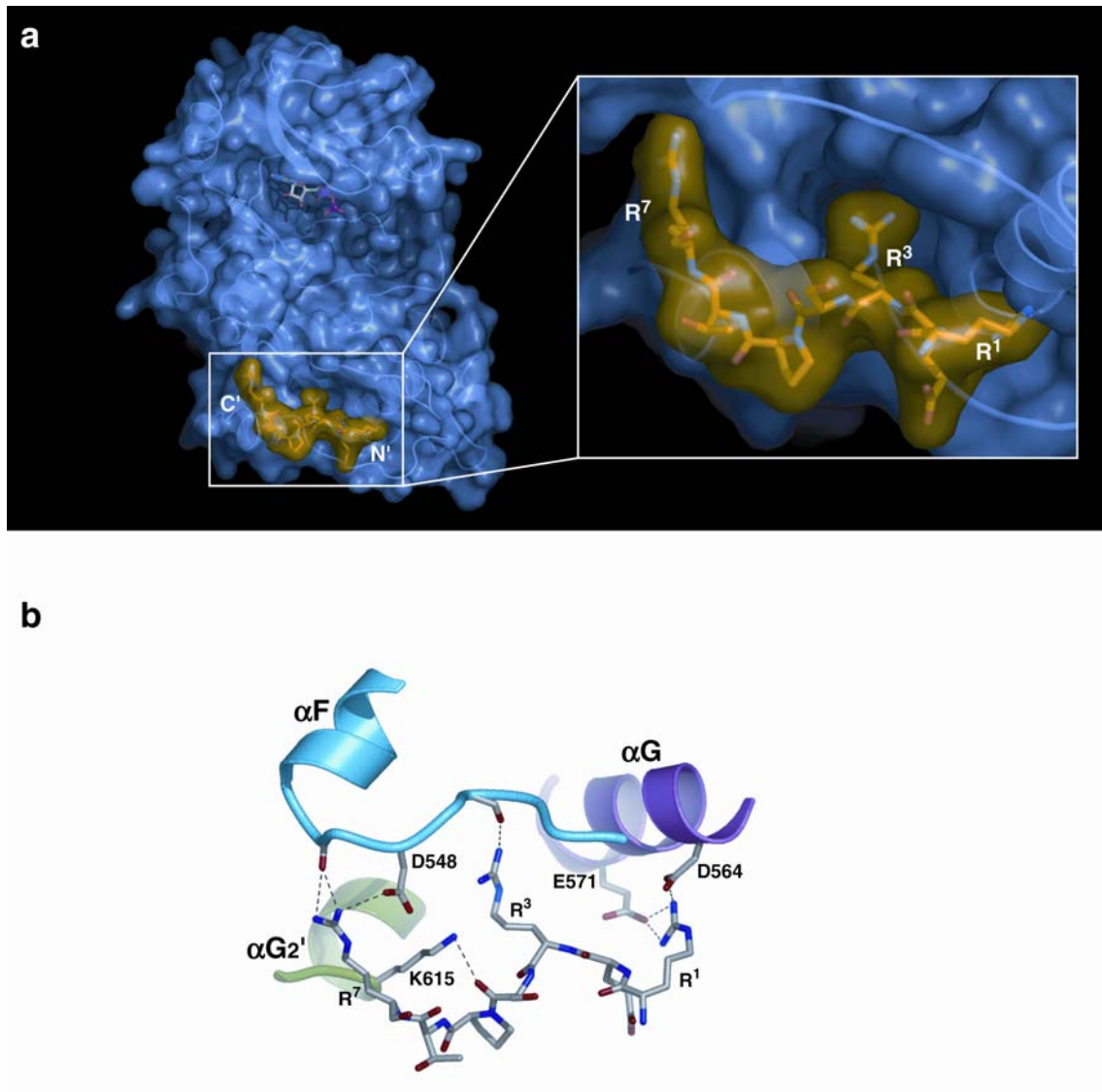


Figure 5.7 Docking groove of SRPK1. a) Surface rendition of SRPK1 $\Delta$ NS1 bound to ADP and docking peptide. The peptide docks in the docking groove through high shape and charge complementarities. b) Interactions mediated by arginines of the docking motif peptide with residues from the docking groove of SRPK1.

W606 of SRPK1. The side chain interactions of this residue suggest a “lock and key” recognition mechanism that can only be mediated by long basic side chain such as arginine and lysine. The complex structure therefore suggests that the docking groove on SRPK1 might recognize a consensus motif of R-X-R/K-X-X-X-R on its substrate.

There are two explanations as to why the substrate peptide binds the docking groove instead of the active site groove. First, the dimeric arrangement of SRPK1 $\Delta$ NS1 in the crystal closes the accessible solvent channel of the substrate binding groove in both molecules, preventing entry and binding of the peptide to the expected site. Thus the peptide docks to a second available site in the kinase due to occlusion of the substrate binding groove. The second and more likely explanation of the docking groove binding by this peptide is that this peptide mimics a docking motif of physiological SR protein substrates and thus fits well in the docking groove of SRPK1 as observed in the crystal structure.

## **6. Substrate Docking Regulates the Mode of Phosphorylation**

To test if the peptide-kinase interactions in the crystal structure truly mimic docking motif/docking groove interactions between SRPK1 and its substrate ASF/SF2, we have mutated four peptide-contacting charged residues in the docking groove of SRPK1 $\Delta$ NS1 (D548, D564, D571 and K615) to alanines. While the mutant SRPK1 $\Delta$ NS1 does not abrogate binding of ASF/SF2, there appears to be a decrease in affinity based on a co-precipitation assay (Figure 5.8a, compare lanes 3 and 4). Also, the overall rate of

phosphorylation by this mutant kinase is much lower than that of wild type SRPK1 $\Delta$ NS1. SRPK1 $\Delta$ NS1 rapidly phosphorylates 8-10 serines in ASF/SF2 with a half-life of approximately 1 minute (Aubol et al., 2003). However, the mutant SRPK1 $\Delta$ NS1 appears to phosphorylate ASF/SF2 at a lower rate than its wild type counterpart, suggesting that the docking groove of SRPK1 may influence its catalytic efficiency (Figure 5.8b). To gain further insight into any mechanistic differences in phosphorylation by the wild type and mutant kinases, we performed start-delay- trap assays, a modification of the start-trap assays (Aubol et al., 2003). The assay involves the use of a kinase active site-directed inhibitor peptide (trap) that has been previously used to demonstrate processive phosphorylation of ASF/SF2 by SRPK1. Our results show that phosphorylation of ASF/SF2 by the mutant SRPK1 is inhibited in presence of the trap, whereas phosphorylation by wild type SRPK1 is unaffected (Figure 5.8c). For the inhibitor peptide to bind to the kinase, the phosphorylated RS domain must be released from its active site after each phosphorylation event, implying that phosphorylation now occurs in a distributive manner. However, to stringently differentiate between processive and distributive phosphorylation, one needs to determine accurately the dissociation constants ( $K_d$ ) of complexes between SRPK1 and ASF/SF2. We suggest that the interaction between the docking motif of ASF/SF2 and the docking groove of SRPK1 is important for regulating the mechanism of phosphorylation of ASF/SF2.

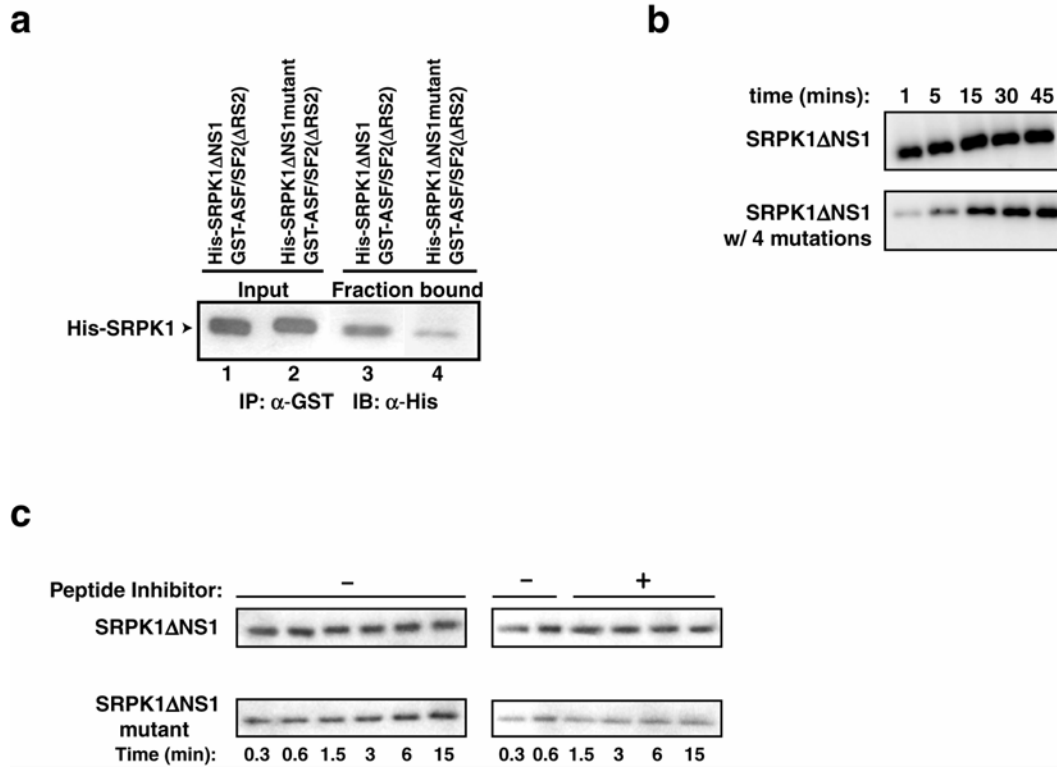


Figure 5.8 The docking groove of SRPK1 plays an important role in its interaction with ASF/SF2. a) Co-precipitation of wild type and mutant SRPK1 with ASF/SF2( $\Delta$ RS2) shows that mutant kinase no longer interacts efficiently with ASF/SF2. b) The mutations at the docking groove significantly affect the phosphorylation of ASF/SF2 by SRPK1. c) The docking groove of SRPK1 is important for processive phosphorylation. Trap-delay-start assays with wild type and mutant SRPK1 show that mutant SRPK1 is inhibited by the trap, whereas wild type SRPK1 is unaffected.

## 7. Identification of a Docking Motif in ASF/SF2

The docking site interactions observed in the peptide bound structure of the kinase led us to identify a docking motif in ASF/SF2. Sequence analysis of ASF/SF2 reveals a small stretch of residues between the RRM and the RS domain which is homologous to the suggested consensus sequence for the docking motif of the substrate (Figure 5.9a). In particular, two of the three basic residues involved in contacting the kinase are conserved (R<sup>191</sup>VKVDGPR<sup>198</sup>) and the third one shifted one position towards the C-terminus. We synthesized this peptide and tested its ability to function as an inhibitor for phosphorylation of ASF/SF2. We observed that this peptide is capable of inhibiting SRPK1 activity at a concentration of 10 mM, suggesting that this eight residue long peptide is the docking motif on the SR protein ASF/SF2 (Figure 5.9b).

## 8. Conservation of Docking Groove in SRPKs

The MAP kinase insert is a canonical structural feature of the CMGC group of kinases. It has been implicated to play an important role in protein interactions between MAP kinases and the substrates, activators, and scaffolding proteins. For instances, Erk2 possesses a groove between helix  $\alpha$ G and the MAP kinase insert that binds its upstream activator Elk (Lee et al., 2004). GSK3 $\beta$  utilizes the same groove to dock its binding proteins axin and FRAT (Bax et al., 2001; Dajani et al., 2001). Despite the fact that the primary sequences and structures of this insert are not conserved among members of

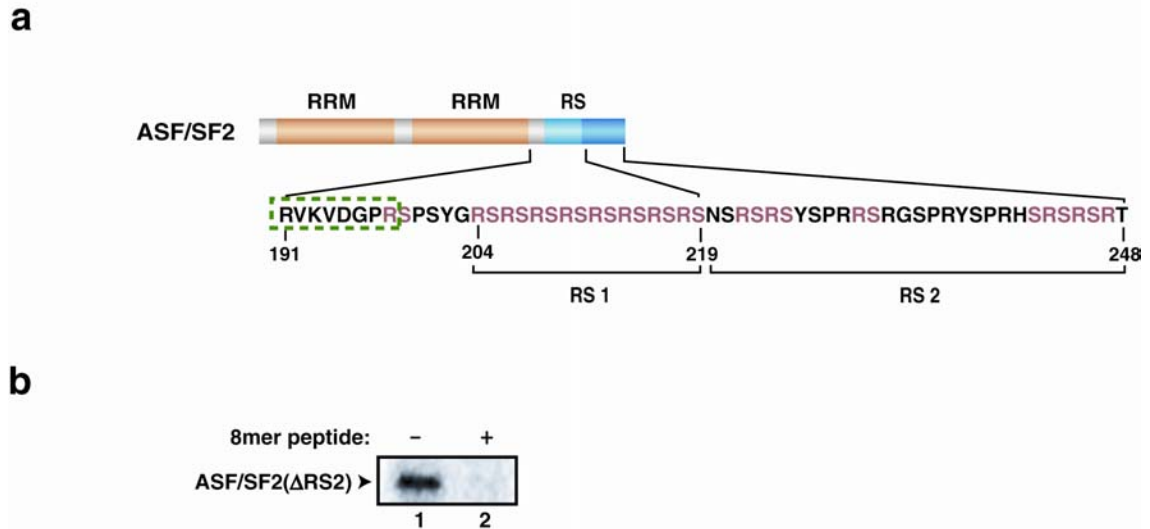


Figure 5.9 Identification of a docking motif in ASF/SF2. a) Domain organization of ASF/SF2. A small stretch of residues (191-198) between the RRMs and the RS domain is identified as the docking motif of the substrate and is indicated by the green box. b) An 8mer peptide derived from the docking motif of ASF/SF2 is a potent inhibitor of SRPK1.

the CMGC group, it is highly conserved in the SRPK family (Figure 5.10). Sequence analysis reveals that all SRPKs contain an extended  $\alpha F/\alpha G$  loop, which shapes the top cleft of the docking groove in SRPK1 (Figures 5.10 and 5.11). As shown earlier, this extended loop and the MAP kinase insert form a deep groove and play a significant role in ASF/SF2 phosphorylation by SRPK1. Furthermore, the four residues that are shown to be important to interact with the docking motif side chains are highly conserved within the family (Figure 5.10). The conservations of  $\alpha F/\alpha G$  loop, MAP kinase insert and the key docking groove residues in SRPKs lead us to speculate that other members of the family might employ the same groove to bind their substrates, or to greater extend, to regulate the mode of phosphorylation of their substrates. Figure 5.11 shows a structural comparison of the  $\alpha F/\alpha G$  loop and the MAP kinase insert in SRPK1, Sky1p and several CMGC kinases.

## 9. Conservation of Docking Motif in SR Proteins

Members of the SR protein family can be divided into two sub-classes based on their structural features: 1) SR proteins that contain only one RRM and 2) members that contain two RRM. While the first RRM of all SR proteins can be defined as canonical RRM, the second RRM cannot. All SR proteins containing two RRM contain a short stretch of highly conserved amino acids, referred as SWQDLKD motif, within the second RRM. The role of this motif will be discussed in detail in Chapter VI. Besides this conservation, sequence comparison between ASF/SF2 and other SR proteins reveals that other SR

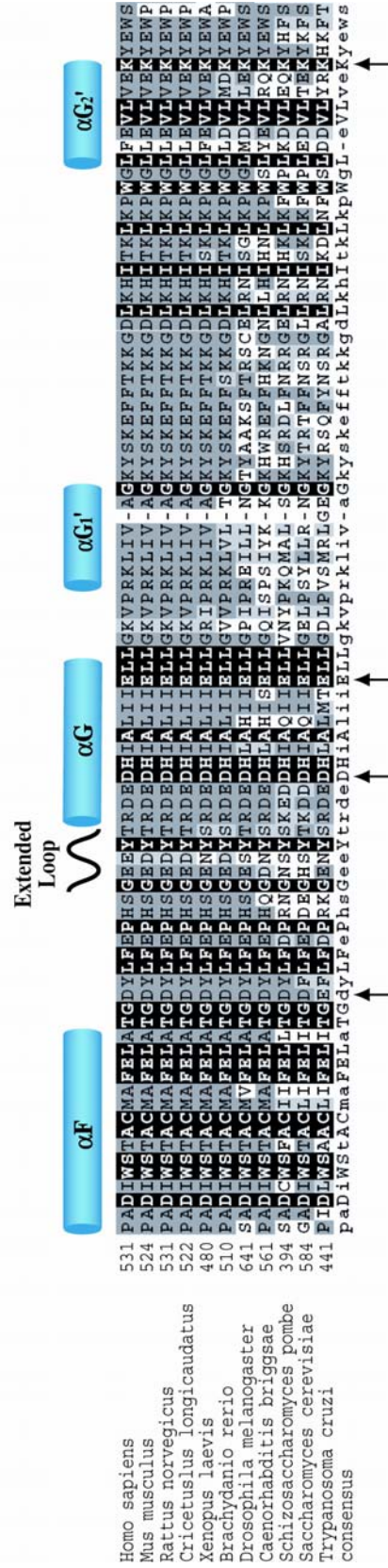


Figure 5.10 Conservation of docking groove in SRPK1. Sequence alignment of SRPK1 from different species shows that the MAP kinase insert is highly conserved. All SRPK1s contain extended  $\alpha F/\alpha G$  loops and the critical residues that interact with the docking peptide are also conserved. The positions of the critical residues are indicated by the arrows. Residues that are identical are boxed in black, similar residues are boxed in grey and different residues in light grey.



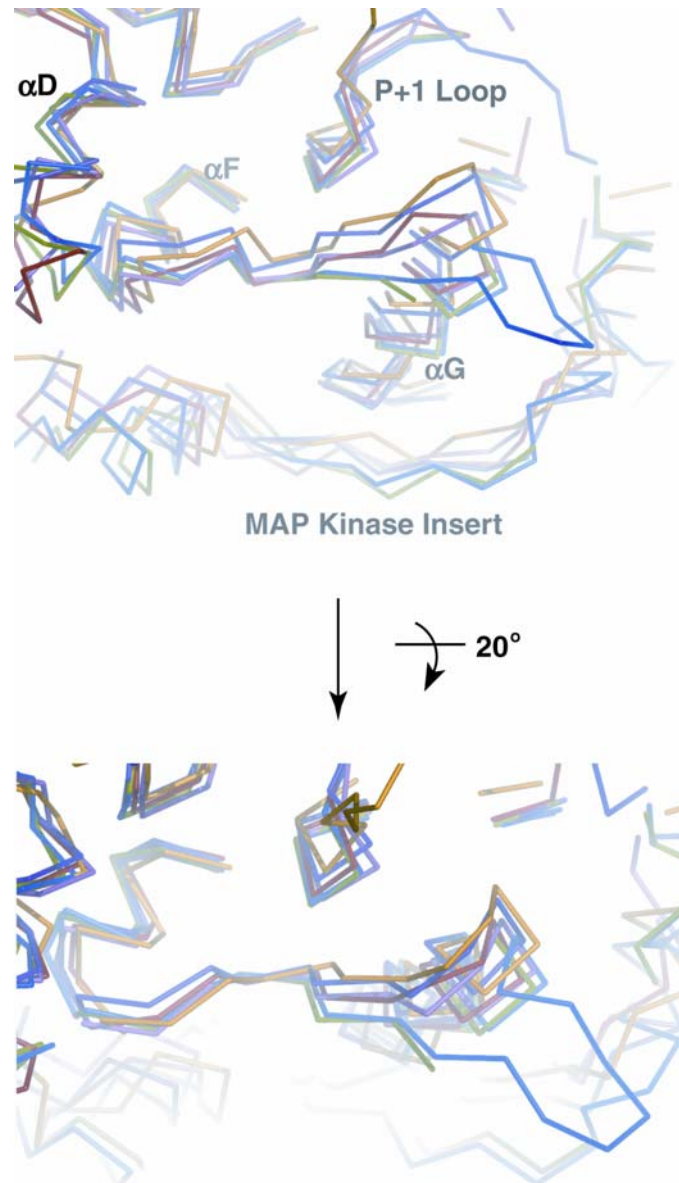


Figure 5.11 SRPK1 contains extended  $\alpha F/\alpha G$  loop. Overlay of large lobes of SRPK1 and other CMGC kinases. While most structural elements of the large lobe including the helices  $\alpha D$  and  $\alpha F$ , and P+1 loop can be overlaid with good r.m.s.d., the  $\alpha F/\alpha G$  loop, helix  $\alpha G$  and the MAP Kinase insert show significant conformational changes in different CMGC kinases. Furthermore, SRPK1 contains an extended  $\alpha F/\alpha G$  loop which shapes the upper cleft of the docking groove and provides an extra surface platform for binding of the docking peptide. SRPK1 is colored blue, Sky1p green, JNK3 yellow, GSK3 $\beta$  orange, CDK2 pink and Erk2 purple.

proteins containing two RRM domains also contain sequences homologous to the first four residues of the optimal docking motif, R/K $\phi$ R/K $\phi$ x(2-3)R (x and  $\phi$  denote any amino acid and hydrophobic residues, respectively) (Figure 5.12). Furthermore, comparison of the sequences of these SR proteins reveals that the number of residues between the docking motif and the first RS dipeptides are similar. We suggest that SR proteins containing two RRM domains possess a common docking motif which is critical in substrate recognition and phosphorylation by SRPK1.

## 10. Functional Role of Helix $\alpha$ G

As mentioned above, helix  $\alpha$ G together with the MAP kinase insert serves as a docking site in SRPK1 and other CMGC kinases, implying that helix  $\alpha$ G plays an important role in substrate binding in many kinases. Indeed, several existing kinase complex structures have confirmed the protein-binding function of helix  $\alpha$ G: complex of CDK2 with a cell cycle regulatory protein CksHs1, CDK2 bound to a phosphatase KAP, PAK1 bound to its autoregulatory domain, PKA bound to the R1 $\alpha$  inhibitory subunit and PKR complexed to its substrate eIF2 $\alpha$  (Bourne et al., 1996; Dar et al., 2005; Kim et al., 2005; Lei et al., 2000; Song et al., 2001). In all cases, residues from the helices  $\alpha$ G make direct contact with the binding proteins. For instance, a surface formed by helix  $\alpha$ G and the L14 (MAP kinase insert) in CDK2 serves as the recognition site for both KAP and CksHs1 (Bourne et al., 1996; Song et al., 2001). In

	RRM2	RS Domain
ASF/SF2	SWQDLKDHMREAGDVCYADVYRDG--TGVVEFVRKEDMTYAVRKLDTKFRSHEGETAYI	RVKVDGFRSPSYGRSRSRSRRS
SRp40	SWQDLKDFMRQAGEVTFADAHKPKLNEGVEFFASYGDLKNAIEKLSGKEING-----	RKIKLEGSKRHSRSRSRSRSRRT
SRp75	SWQDLKDYMRQAGEVTYADAHKGRKNEGVIIEFVSYSDMKRALDKLDGTEVNG-----	RKIRLVEDKPGSRRRSYSRSRRS
SRp55	SWQDLKDFMRQAGEVTYADAHKERTNEGVIIEFRSISDMKRALDKLDGTEING-----	RNIRLIIEKPRTSRHSYSGSRS

Figure 5.12 SR proteins with two RRM2s contain docking motifs. Sequence analysis of SR proteins containing two RRM2s reveals that these proteins, in addition to the conserved SWQDLKD motif, also possess sequences homologous to the proposed consensus docking motif. The basic residues in the consensus motif is colored blue, hydrophobic residues green and the serine of the first RS dipeptide red.

SRPK1, residues at the N-terminus of the helix  $\alpha$ G also make significant contributions to the binding of the peptide by forming ionic pairs with the arginine within the consensus docking motif. Intriguingly, all existing X-ray structures of kinases from the CMGC family contain a conserved glycine residue between the C-terminal end of helix  $\alpha$ G and the MAP kinase insert. The importance of this residue is unknown but one can speculate that a glycine residue at this position can serve as a hinge for independent motion of the helix  $\alpha$ G and MAP kinase insert upon substrate binding in SRPK1 and other CMGC kinases.

## **C. Discussion**

### **1. The docking motif on ASF/SF2 is responsible for high affinity interaction with SRPK1 and its mode of phosphorylation by SRPK1**

The X-ray crystal structure of a complex of SRPK1 and a 9mer peptide derived from Npl3p led us to identify a docking groove in SRPK1 and a docking motif in ASF/SF2 (R<sup>191</sup>VKVDGPR<sup>198</sup>). Mutations of residues within SRPK1 have shown that direct interactions between the docking groove and the docking motif govern the binding affinity of the two proteins, which result in processive phosphorylation of ASF/SF2. We also observed that deletion (ASF/SF2( $\Delta$ Dock)) or mutation (ASF/SF2( $\Delta$ RS2 R191A/K193A)) of the docking motif of ASF/SF2 significantly reduces binding to SRPK1 $\Delta$ NS1. Furthermore, the start-delay-trap assay showed that phosphorylation of ASF/SF2( $\Delta$ Dock) and ASF/SF2( $\Delta$ RS2 R191A/K193A) is inhibited in the presence of the inhibitor peptide, suggesting

that the docking motif in ASF/SF2 plays an important role in governing the mechanism of phosphorylation by SRPK1.

## **2. The docking motif of ASF/SF2 limits its phosphorylation by SRPK1 and is important for its localization to nuclear speckles**

Subcellular localization of ASF/SF2 depends on its phosphorylation state. ASF/SF2 is imported into the nucleus and localizes to the nuclear speckles upon hypophosphorylation (Lai et al., 2001; Yun et al., 2003). ASF/SF2 is then released from the nuclear speckles and recruited to the site of splicing upon hyperphosphorylation (Colwill et al., 1996; Misteli et al., 1998). In fact, SRPK1 phosphorylates only ~10 serines within the first stretch of RS dipeptides (RS1) (Figure 5.1a) (Aubol et al.). We also observed that such restriction of phosphorylation is dependent on the docking motif of ASF/SF2. Furthermore, ASF/SF2 localizes to the nuclear speckles in a docking motif-dependent manner *in vivo*. While the release of wild-type ASF/SF2 from the nuclear speckles requires the presence of Clk/Sty and the RS2 domain of ASF/SF2, suggesting the release is due to hyperphosphorylation of the RS domain, ASF/SF2 mutants with the docking motif either deleted or mutated are mostly diffused throughout the nucleus (Ngo et al., 2005). The fact that these molecules can escape restricted phosphorylation by SRPK1 (due to the defects of the docking motif) suggests that hyperphosphorylation might be the reason why the ASF/SF2 mutants failed to localize to the nuclear speckles.

### 3. Regulation of ASF/SF2 Phosphorylation by SRPK1 and Clk/Sty

In summary, we propose that SRPK1 phosphorylates the RS1 motif of ASF/SF2 in a docking groove-docking motif interaction-dependent manner to generate a hypophosphorylated species in the cytoplasm. This modification mediates nuclear import and recruitment of ASF/SF2 to nuclear speckles. ASF/SF2 is then hyperphosphorylated by Clk/Sty, leading to its release from the nuclear speckles (Figure 5.13). Nuclear speckles are thought to be sites of storage for splicing factors, where they maybe recruited to the sites of active transcription (and consequently splicing) by the action of Clk/Sty (Misteli et al., 1998; Misteli et al., 1997). We can now begin to understand the sequence of phosphorylation events of ASF/SF2 and its relation to biological function. Since splicing factors have to be released from nuclear speckles in order to participate in splicing, it is likely that the hyperphosphorylation and release event is stringently regulated in cells by specific signals to activate Clk/Sty.

It has been shown earlier that dephosphorylation of ASF/SF2 is required for splicing activity (Mermoud et al., 1992; Tazi et al., 1993). Moreover, nuclear export of mRNA mediated by ASF/SF2 requires hypophosphorylated ASF/SF2. Therefore, it is likely that at some point during spliceosome assembly or catalysis, ASF/SF2 undergoes dephosphorylation, which converts a hyperphosphorylated state back to a hypophosphorylated state. However no specific phosphatase responsible for this event has been identified as yet.

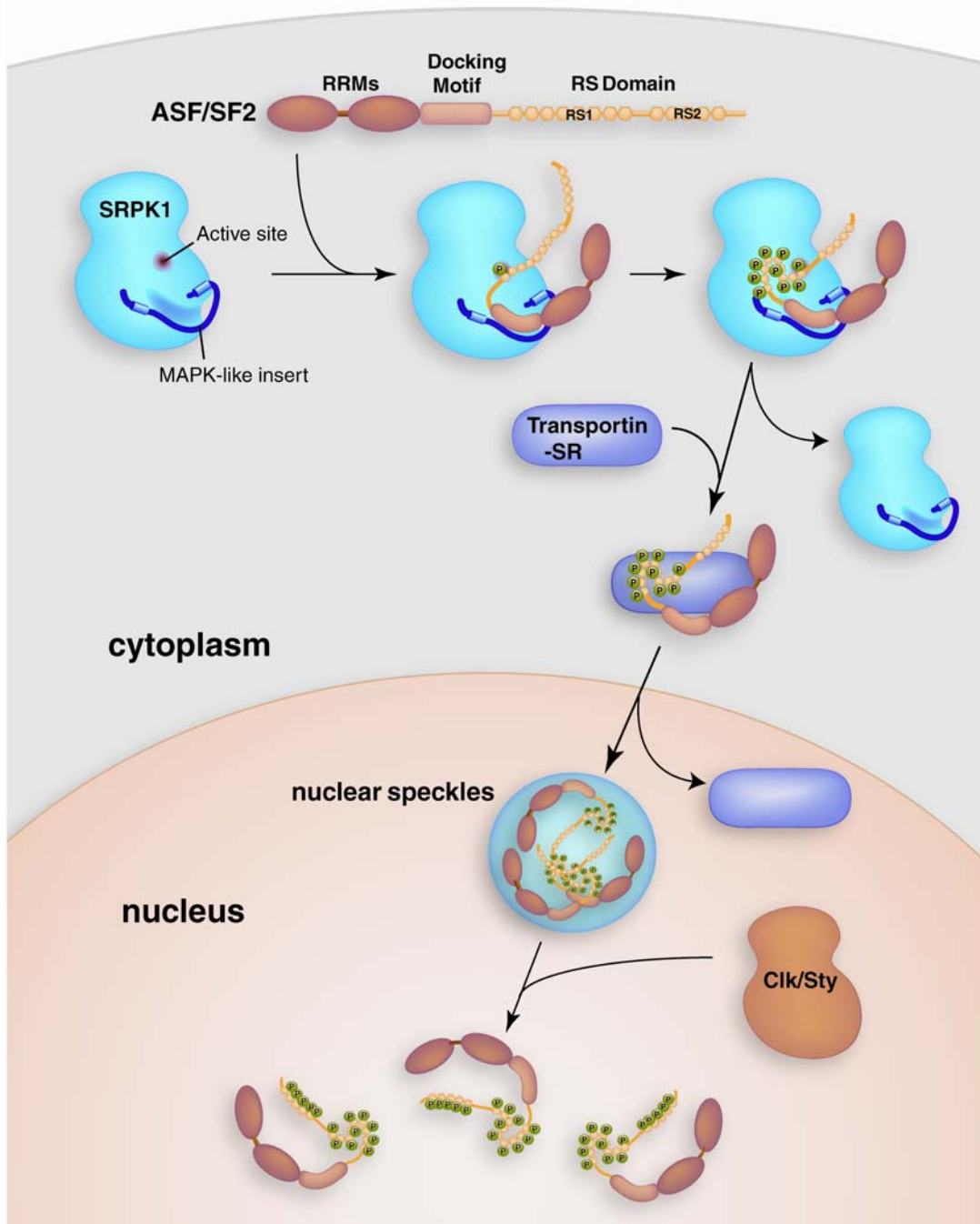


Figure 5.13 Proposed model for a role of both SRPK1 and Clk/Sty kinases in phosphorylation and subcellular localization of ASF/SF2. SRPK1 present in the cytoplasm hypophosphorylates ASF/SF2 in a processive manner. The hypophosphorylated form of ASF/SF2 is imported into the nucleus through interaction with transportin-SR and is deposited into the nuclear speckles. Further phosphorylation by Clk/Sty results in a hyperphosphorylated form of ASF/SF2, which is released from the speckles and recruited to the site of splicing.

#### 4. Docking Interactions between Kinases and their Binding Proteins

Several kinases rely on docking interactions with substrates, using sites distinct from the phosphor-acceptor sequences to achieve substrate specificity (Bhattacharyya et al., 2006; Biondi et al., 2002; Biondi and Nebreda, 2003; Chang et al., 2002; Holland and Cooper, 1999; Lee et al., 2004; Tanoue et al., 2000; Tanoue et al., 2001). These docking sites on the surface of the kinase are distinct from the active site, and the interactions mediated by these grooves are critical for both highly efficient as well as specific phosphorylation. For instance, p38, a member of the MAP kinase family, phosphorylates a full-length substrate with an efficiency nearly two orders of magnitude higher than that for a minimal 14-residue substrate derived from the same protein (Hawkins et al., 2000). The docking motifs of substrates, upstream activators and regulators, and scaffold proteins of MAP kinases bind to these docking grooves. One such docking groove identified in Erk2 and GSK3 $\beta$  is formed by the MAP kinase insert and helix  $\alpha$ G (Bax et al., 2001; Dajani et al., 2003; Lee et al., 2004). The MAP kinase insert is present in all members of the CMGC group of kinases, including SRPKs, suggesting that this docking groove might be a common feature in CMGC kinases. Indeed, we have shown here that the docking groove in SRPK1 not only facilitates substrate binding, but also regulates the mode of phosphorylation.



## **5. Biological Implications of High Affinity Binding and Processive Phosphorylation**

SRPK1 also phosphorylates non-SR proteins including protamine 1, lamin B-receptor and viral HBV core protein, which do not contain any RRM or docking motifs (Daub et al., 2002; Papoutsopoulou et al., 1999a; Papoutsopoulou et al., 1999b). However, these proteins contain RS dipeptide stretches resembling those of SR proteins. Unlike SR proteins, these RS dipeptides do not exist as continuous repeats. One would speculate that these proteins might not use the docking motif and thus bind with lower affinity than the SR proteins. Such changes in affinity may lead to a distributive phosphorylation mechanism. Why do ASF/SF2, and possibly other SR proteins containing two RRMs, bind SRPK1 with such high affinity and undergo processive phosphorylation? First of all, by exploiting stable binding interactions in a processive mechanism, SRPK1 ensures high substrate selectivity and phosphorylation content, thus assuring that a consistent number of serines on ASF/SF2 are phosphorylated before its release. In contrast, a distributive mechanism may result in a mixed pool of differentially phosphorylated substrates will result. Secondly, as mentioned in Chapter IV, the localization of SRPK1 is regulated by its spacer domain. However, the import mechanism of SRPK1 into the nucleus is unknown. Phosphorylated ASF/SF2 has been shown to interact with transportin-SR, a member of the importin- $\beta$  protein family, and can be imported into the nucleus. We suggest that through high affinity binding, the kinase remains associated with its phosphorylated substrate, interacts with

transportin-SR as a ternary complex and is colocalized into the nucleus through a “piggy-back” mechanism. In agreement with this, a recent study has shown that the nuclear import of SRPK1 is dependent on its catalytic activity, suggesting the phosphorylation event and its association with its substrate could be essential for its nuclear import (Ding et al., 2006). Finally, by forming a stable complex, SRPK1 may also protect phosphorylated ASF/SF2 from being dephosphorylated by cellular phosphatases before its recruitment to the speckles.

The text of this chapter is, in part, reprints of material as it appears in *Molecular Cell*, 2005, **20**, 77-89, Ngo, J. C., Chakrabarti, S., Ding, J. H., Velazquez-Dones, A., Nolen, B., Aubol, B. E., Adams, J. A., Fu, X. D. and Ghosh, G.. The dissertation author was the primary researcher and author of this publication.

**Chapter VI: X-Ray Crystal Structure of SRPK1 bound to  
an SR Protein ASF/SF2**

## A. Introduction

The RS domains of SR proteins contain many serines, each in a slightly different primary sequence context. In the RS domain of ASF/SF2, for example, 20 out of 50 residues are serines. Eight of these serines are within an extended RS dipeptide repeat (RS1 domain); the others are within shorter repeats or in single dipeptides surrounded by residues other than arginines (RS2 domain). It has been shown that ~10 moles of phosphate are transferred per mole of ASF/SF2 by SRPK1 and that the phosphorylation occurs in a processive manner (Aubol et al., 2003). The identifications of a docking groove in SRPK1 and a docking motif in ASF/SF2, as described in Chapter V, has provided invaluable insight into the regulation of ASF/SF2 phosphorylation. We showed that the docking interaction between SRPK1 and ASF/SF2 restricts the phosphorylation of the dipeptides to largely RS1 domain and promotes nuclear import of the SR protein, whereas hyperphosphorylation of the rest of the RS domain by Clk/Sty results in its dissociation from the speckles. However, the molecular basis for the substrate recognition and mechanism to achieve the restricted and processive phosphorylation remains to be delineated.

Through small adaptations to the common kinase fold, protein kinases can recognize and phosphorylate a unique subset of the cellular proteins. Generally, the recognition of residues flanking the phosphorylatable residue, referred to the consensus sequence motifs, through specific intermolecular interactions in the active site, is important for recognition. However, this rule cannot be applied to SRPKs as the substrates contain only RS repeats

surrounding the P0 site, whereas the P0 site is also a RS dipeptide. Furthermore, while most Ser/Thr kinases can recognize both serine and threonine as the phosphate acceptor, SRPKs are highly stringent toward serine and basicity of the surrounding residues is not the only determinant for substrate specificity (Gui et al., 1994b; Wang et al., 1998a).

In this study, we attempted to clarify the structural basis for both substrate recognition and restricted phosphorylation of ASF/SF2 by SRPK1. We have identified an experimental protocol for the formation and purification of various SRPK1:ASF/SF2 complexes. We have also obtained crystals of a complex between functional fragments of SRPK1 and ASF/SF2. While structure determination of this complex is underway, the current model has revealed important information on substrate recognition of ASF/SF2 by SRPK1.

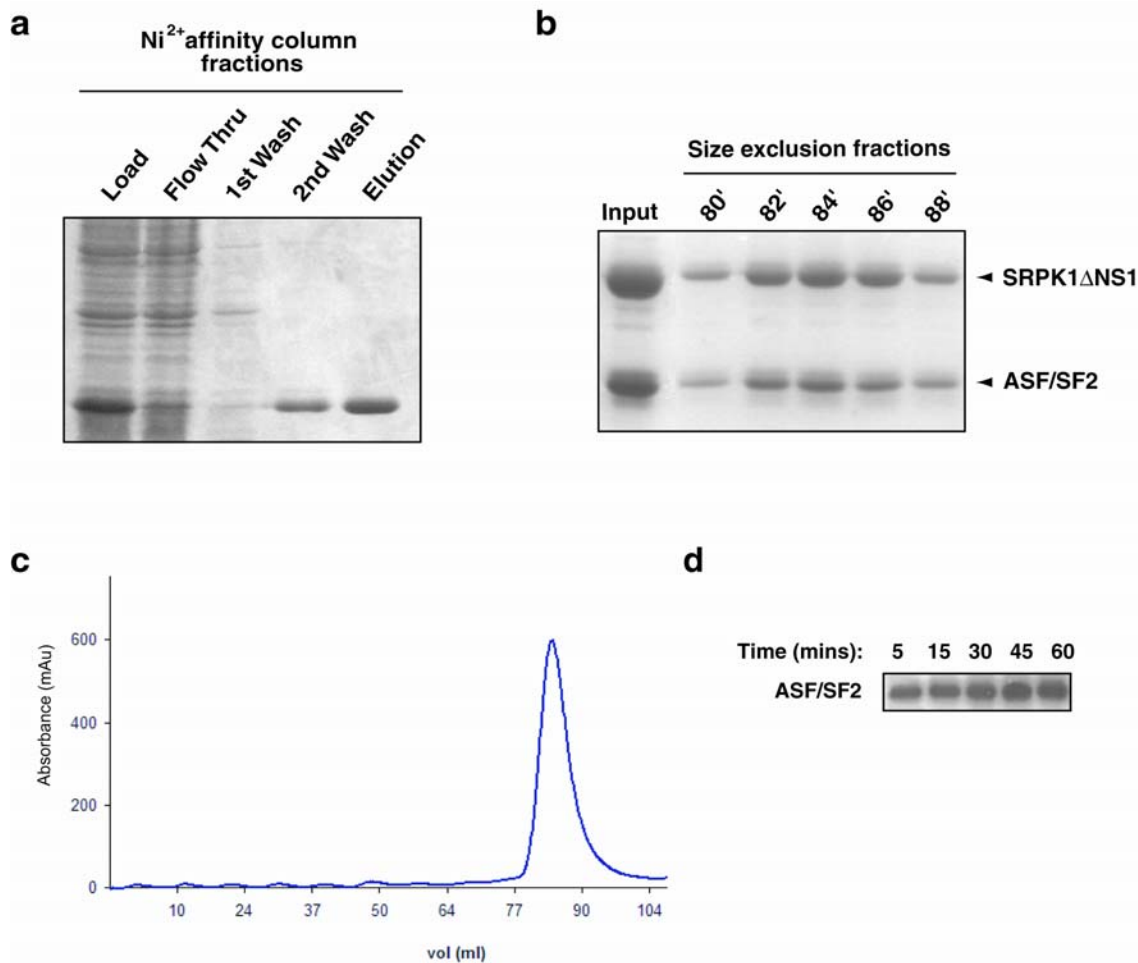
## **B. Results**

### **1. Purification and Complex Formation of SRPK1:ASF/SF2**

Biochemical and kinetic studies of SRPK1 and its substrate ASF/SF2 reveal that the kinase binds to the substrate with unusually high affinity ( $K_d = 20 \pm 6\text{nM}$ ) ((Aubol et al., 2003; Ding et al., 2006). This observation strongly suggests that SRPK1 could be complexed to ASF/SF2 and could serve as potential crystallization target. However, efforts to purify ASF/SF2 as an N-terminal hexa-histidine fusion protein have proven to be problematic. Although a high level of protein expression of His-tagged ASF/SF2 in E.coli can be easily achieved, the protein remains mostly insoluble and forms inclusion bodies. By

solubilizing ASF/SF2 in the presence of denaturant, we were able to purify the protein to ~90% homogeneity using Ni<sup>2+</sup> affinity chromatography (Figure 6.1a). However, attempts to refold the denatured protein by step-wise dialysis proved unsuccessful. The recovery yield of soluble ASF/SF2 by this method was minimal as less than 0.5 mg of refolded protein could be recovered from 100 mg of denatured protein. This low yield also discouraged the idea to form complex of SRPK1 and ASF/SF2 by simply mixing the two proteins.

Since ASF/SF2 is a natural substrate of SRPK1, we reasoned that the introduction of SRPK1 might assist the refolding of ASF/SF2 from inclusion bodies and thus form a stable complex during the refolding process. We first purified both proteins independently to near homogeneity and then mixed the two proteins together in the presence of denaturant. The denatured proteins were then refolded together by dialysis. Surprisingly, through this method, ASF/SF2 remained soluble. SDS-PAGE analysis of the dialysate showed the presence of both SRPK1 and ASF/SF2 suggesting the SRPK1:ASF/SF2 complex was formed in stoichiometric ratio (Figure 6.1b). The dialysate was concentrated and the His-tag of the protein was removed by thrombin cleavage overnight at room temperature. The untagged protein was further purified by size exclusion chromatography. The size exclusion profile further confirmed the formation of stoichiometric SRPK1:ASF/SF2 complex (Figure 6.1c). Although a significant amount of the protein complex could be recovered after the refolding process and size exclusion chromatography, the purified protein has strong tendency to aggregate during concentration and could only be concentrated



purified to near homogeneity by Ni<sup>2+</sup> affinity column before refolded with SRPK1. b) Both SRPK1ΔNS1 and ASF/SF2 were eluted together from size exclusion column. SDS-PAGE analysis suggests the proteins were in stoichiometric ratio. c) Elution profile from size exclusion chromatography shows a single peak, confirming that SRPK1ΔNS1 and ASF/SF2 were eluted together as a complex. d) Kinase activity assay using the refolded complex showed the complex was fully functional.

to ~3 mg/mL. Such a low concentration usually discourages the formation of crystals. We therefore attempted to test the effects of different buffers on the solubility of the complex. Complexes obtained after refolding were subjected to another dialysis using buffers with different pH. Remarkably, the complexes dialysed in the presence of 20 mM CAPSO pH 9.5 showed significant improved solubility and could be concentrated to 12 mg/mL. Finally, kinase assays performed with the refolded complexes show the complex is fully functional (Figure 6.1d). We concluded that we have developed an efficient protocol for the formation of SRPK1:ASF/SF2 complexes for crystallization.

## **2. Design of SRPK1:ASF/SF2 Crystallization Targets**

Initially, the crystallizable active fragment of SRPK1, SRPK1 $\Delta$ NS1, was used to form a complex with full length ASF/SF2. However, the crystallization trials did not produce any lead. We reasoned that RS domains of all SR proteins, including ASF/SF2, are highly disordered and could hinder the crystallization process. Therefore, we removed most of the RS domain of ASF/SF2, referred as ASF/SF2(1RS), retaining only the first RS dipeptide repeat at the C-terminus (Figure 6.2). Surprisingly, unlike the full-length ASF/SF2, this truncated protein did not form a tight complex with SRPK1 $\Delta$ NS1. SRPK1 $\Delta$ NS1 and ASF/SF2(1RS) were resolved by size exclusion chromatography and the two proteins eluted at different retention times (Figure 6.3a). The apparently lower affinity of this complex suggested that the RS



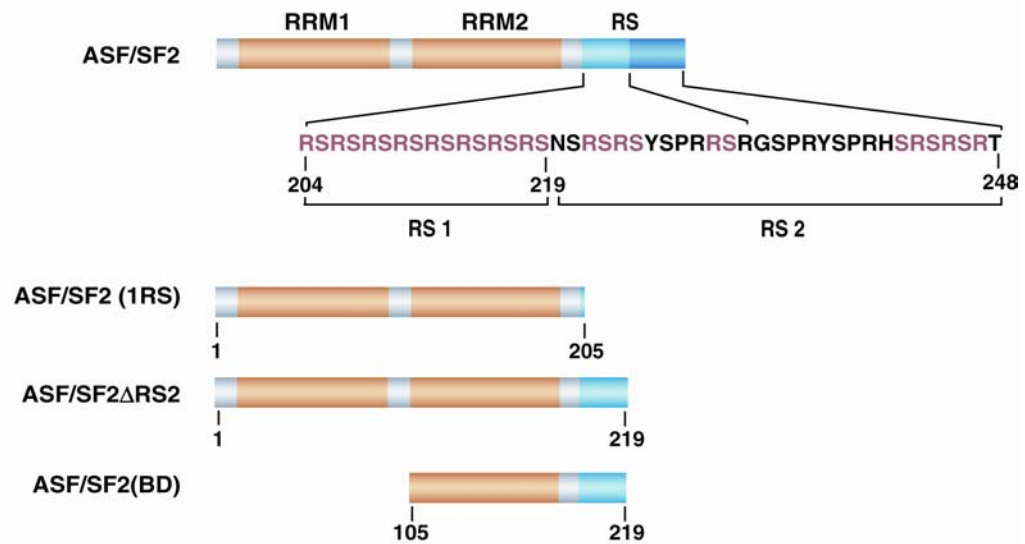
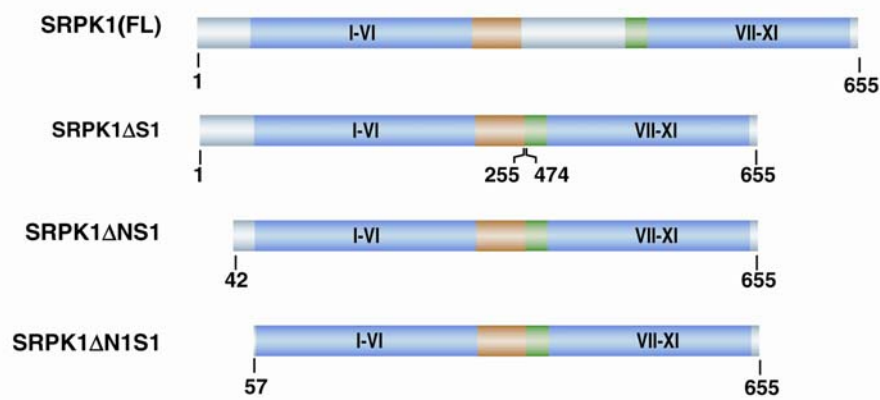
**a****b**

Figure 6.2 Constructs of ASF/SF2 and SRPK1 used for crystallization. a) Different constructs of ASF/SF2 designed for refolding and crystallization experiments with SRPK1. b) SRPK1 constructs used for complex formation with ASF/SF2. Yellow and green segments denote the important spacer region as that described in Chapter IV.

domain might contribute to the high affinity between the kinase-substrate pair.

Next, we retained the first eight RS dipeptide repeats and only removed the RS2 domain of ASF/SF2, referred as ASF/SF2 $\Delta$ RS2 (Figure 6.2). This protein formed tight complexes with SRPK1 $\Delta$ NS1 as shown by the size exclusion profile and SDS-PAGE analysis (Figures 6.3b and c). Crystallization trials of SRPK1 $\Delta$ NS1:ASF/SF2 $\Delta$ RS2 complexes were carried out in the presence and absence of different nucleotides (ADP and AMP-PNP), but no complex crystal was obtained.

The crystal structure of apo-SRPK1 $\Delta$ NS1 reveals that although portion of the N-terminal region was retained, no electron density was observed, suggesting that this region is flexible and disordered. We reasoned that the flexibility of this region might have impeded the formation of any complex crystal. Therefore, we further removed 15 residues of the N-terminus in the context of SRPK1 $\Delta$ NS1. This construct contains residues 57-255 linked to residues 474-655 of SRPK1 and is referred as SRPK1 $\Delta$ N1S1 (Figure 6.2). This SRPK1 construct readily formed a complex with ASF/SF2 $\Delta$ RS2 (Figures 6.4a and b). Crystallization screening of apo-complex successfully produced small clusters of needle-shaped crystals in 100 mM Tris-HCl pH 8.5, 200 mM NH<sub>4</sub>H<sub>2</sub>PO<sub>4</sub> and 50% MPD (Figure 6.4c). Unfortunately, optimization of these crystals was unsuccessful and did not produce suitable crystals for diffraction studies. On the other hand, when the crystallization trials were performed in the presence of AMP-PNP, similar needle shaped crystals were formed in

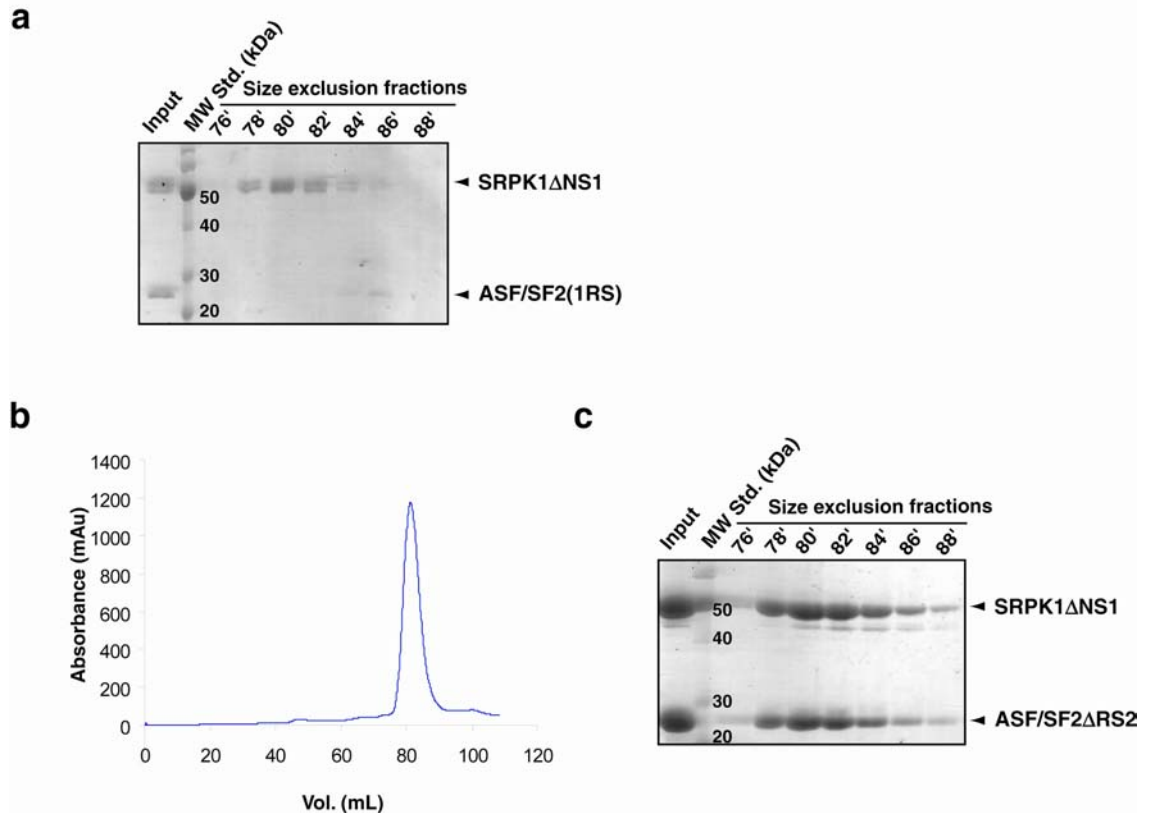


Figure 6.3 Purification of different SRPK1:ASF/SF2 complexes. a) SDS-PAGE analysis of SRPK1 $\Delta$ NS1:ASF/SF2(1RS) complex reveals that although the two proteins could be refolded together, binding affinity between the two proteins was weak and they were resolved by size exclusion chromatography and eluted at different retention time. b) Size exclusion profile of SRPK1 $\Delta$ NS1:ASF/SF2 $\Delta$ RS2 complex confirms the formation of tight complex. c) 12.5 % SDS-PAGE analysis of elution from size exclusion chromatography shows stoichiometric ratio of SRPK1 $\Delta$ NS1 and ASF/SF2 $\Delta$ RS2.

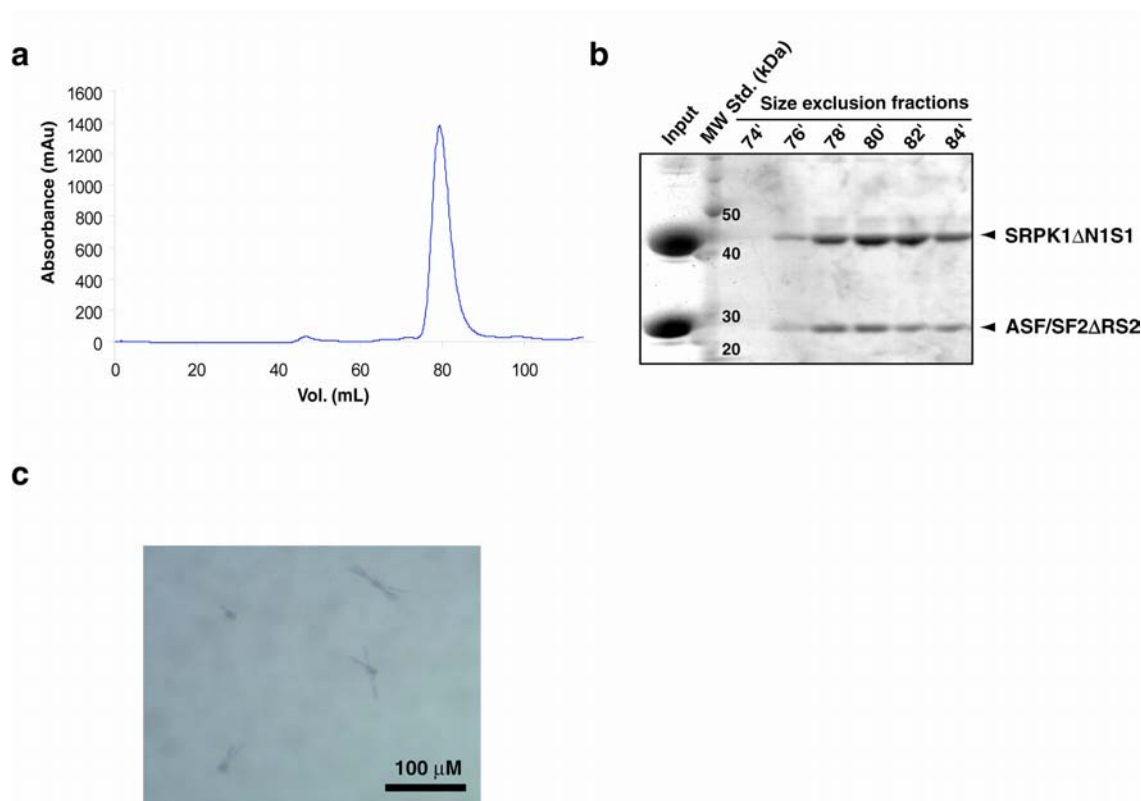


Figure 6.4 Purification of crystallizable SRPK1:ASF/SF2 complex. a) Size exclusion profile of SRPK1 $\Delta$ N1S1:ASF/SF2 $\Delta$ RS2 complex shows a single peak. b) SDS-PAGE analysis of elution from size exclusion chromatography confirms the formation of complex. c) Small clusters of needle-shaped crystals of SRPK1 $\Delta$ N1S1:ASF/SF2 $\Delta$ RS2 complex were obtained from 100 mM Tris-HCl pH 8.5, 200 mM NH<sub>4</sub>H<sub>2</sub>PO<sub>4</sub> and 50 % MPD.

4 different conditions from a commercial sparse matrix screen after 10 days (Figure 6.5a).

### **3. Identification of Diffraction Quality SRPK1:ASF/SF2 Complex Crystals**

All four conditions that produced crystals contain polyethylene glycol (PEG) as precipitant. Optimization of the crystals by screening different concentrations of PEG, different buffers and ions eventually yielded small rod shape crystals of sizes around 30  $\mu\text{M}$  X 30  $\mu\text{M}$  X 150  $\mu\text{M}$  (Figure 6.5b). The optimized condition contained 200 mM sodium citrate pH 5.6, 200 mM sodium acetate and 8% PEG 5000MME. The crystals diffracted to 4.5 Å at the home source X-ray and 3.2 Å at synchrotron beamline 8.2.2 at ALS at Lawrence Berkeley National Laboratory respectively. Surprisingly, dissolution of isolated crystals and analysis by SDS-PAGE revealed that the complex crystals contained SRPK1 $\Delta$ NS1 and a fragment of protein that has lower molecular weight than ASF/SF2 $\Delta$ RS2 (Figure 6.5c). Detail sequence analysis of ASF/SF2 unveiled that the protein contains a plausible thrombin cleavage site between the two RRM. We reasoned that trace amount of thrombin was retained after the purification and eventually cleaved ASF/SF2 $\Delta$ RS2 into two fragments, leaving only one fragment bound to SRPK1 $\Delta$ N1S1 and crystallized. As described in Chapter IV, the docking motif of ASF/SF2 lies between the RRM2 and RS1 domains and contributes to stable interactions between the kinase and substrate. Therefore, it was persuasive that the retained fragment was RRM2 and the RS1 domains. We truncated both RRM1 and RS2 domains of ASF/SF2

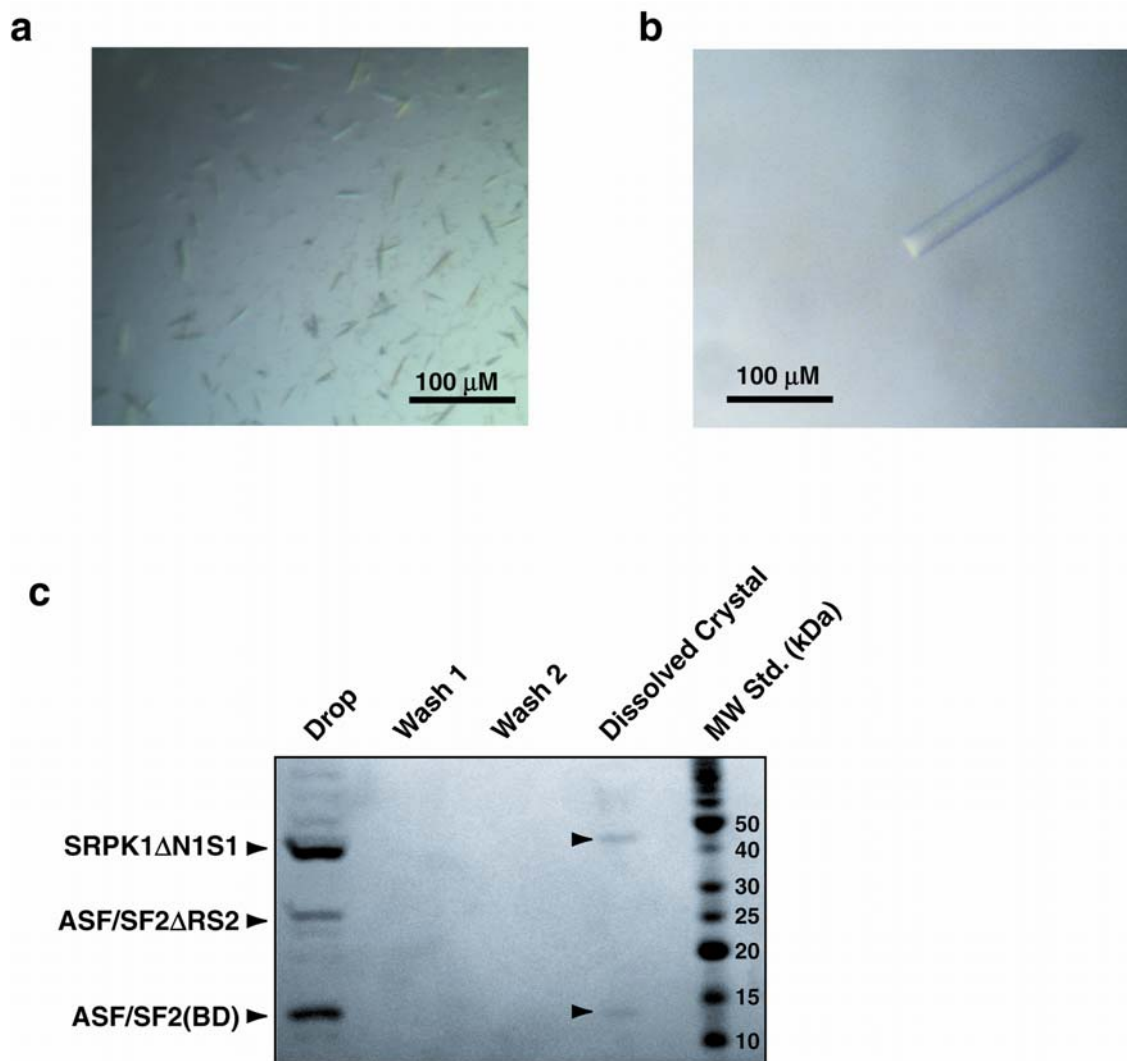


Figure 6.5 SRPK1:ASF/SF2 complex crystals. a) Needle-shaped crystals of SRPK1 $\Delta$ N1S1:ASF/SF2 $\Delta$ RS2 complex were obtained in the presence of AMP-PNP. b) Crystal of SRPK1 $\Delta$ N1S1:ASF/SF2 $\Delta$ RS2 complex obtained from optimized condition containing 200 mM sodium citrate pH 5.6, 200 mM ammonium acetate and 8% PEG5000MME. c) 12.5 % SDS-PAGE analysis of dissolved crystals revealed that the complex crystals contained, in addition to SRPK1 $\Delta$ N1S1, a truncated fragment of ASF/SF2 $\Delta$ RS2, which was later identified as ASF/SF2(BD).

and referred this construct as ASF/SF2(BD) (BD for binding domain) (Figure 6.2). This new complex of SRPK1 $\Delta$ N1S1:ASF/SF2(BD) was purified from size exclusion chromatography (Figure 6.6). Crystallization of the new complex was set up using the optimized crystallization condition, the complex crystals with the same crystal form and diffraction quality were reproduced in 1 day, suggesting the original complex crystals obtained after thrombin cleavage was indeed SRPK1 $\Delta$ N1S1:ASF/SF2(BD). Table 6.1 summarizes all of the protein complexes that were screened for crystallization conditions.

#### **4. Data Collection of SRPK1:ASF/SF2:AMP-PNP Complex**

In order to improve the crystal size for better diffraction quality, we screened different additives and identified that either 5% MeOH or EtOH significantly improved the size of crystals to around 80  $\mu$ M X 80  $\mu$ M X 350  $\mu$ M. The crystals were then dehydrated and cryo-protected by dialysis overnight against 200 mM sodium citrate pH 5.6, 100 mM sodium acetate, 5% MeOH of EtOH, 15% PEG 5000MME and 25% (V/V) ethylene glycol. All dialysed crystals were flash frozen in liquid nitrogen. X-ray diffraction data of the ternary complex crystals were collected using a MAR CCD detector at GM/CA-CAT synchrotron beamline ID-23 of the Advanced Photon Source (APS) at Argonne National Laboratory. The diffraction limit of crystals improved to  $\sim$ 2.9 Å using the synchrotron radiation (Table 6.2). The complex crystals belong to orthorhombic space group I222 with unit cell dimensions  $a = 57.406$  Å,  $b =$

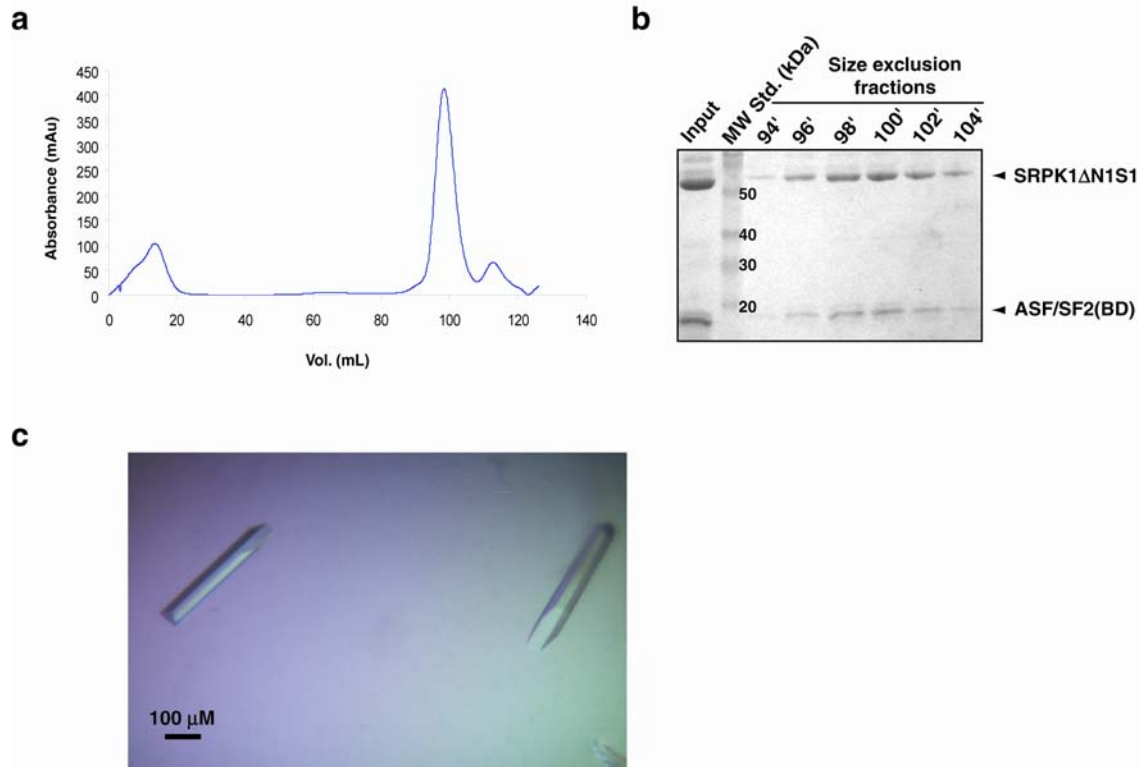


Figure 6.6 Formation of SRPK1 $\Delta$ N1S1:ASF/SF2(BD):AMP-PNP complex crystals. a) Elution profile of size exclusion chromatography shows that SRPK1 $\Delta$ N1S1:ASF/SF2(BD) complex was homogeneous. b) 12.5 % SDS-PAGE analysis of the size exclusion fractions of SRPK1 $\Delta$ N1S1:ASF/SF2(BD) complex. c) Big complex crystals were obtained after one day.



Table 6.1: Summary of SRPK1:ASF/SF2 Complex Formation and Crystallization

SRPK1	ASF/SF2 construct	Complex Formation	Crystals	Diffraction
SRPK1 $\Delta$ S1	ASF/SF2	Yes	None	
SRPK1 $\Delta$ NS1	ASF/SF2	Yes	None	
SRPK1 $\Delta$ N1S1	ASF/SF2	Yes	None	
SRPK1 $\Delta$ NS1	ASF/SF2(1RS)	No	None	
SRPK1 $\Delta$ NS1	ASF/SF2 $\Delta$ RS2	Yes	None	
SRPK1 $\Delta$ N1S1	ASF/SF2 $\Delta$ RS2	Yes	Yes	4.5 Å at home 3.2 Å at synchrotron
SRPK1 $\Delta$ N1S1	ASF/SF2(BD)	Yes	Yes	2.9 Å

Table 6.2: Data Collection and Refinement of SRPK1:ASF/SF2:AMP-PNP ternary complex crystal

Crystal	SRPK1 $\Delta$ N1S1:ASF/SF2(BD):AMP-PNP
Data Collection	
Data Collection Source	APS ID-23
Wavelength (Å)	1.0124
Resolution (Å)	50-2.9
No. of measured reflections	79579
Completeness (outer shell) (%)	93.0 (62.2)
$I/\sigma$ (overall / outer shell)	19.9 (2.54)
$R_{\text{sym}}^a$ (overall / outer shell) (%)	5.8 (44.3)
Refinement	
Resolution (Å)	50-2.9
$R_{\text{crys}}^b$ (%) (current)	29.23
$R_{\text{free}}^c$ (%) (current)	38.88

<sup>a</sup> $R_{\text{sym}} = \sum |I - \langle I \rangle| / \sum I$

<sup>b</sup> $R_{\text{crys}} = \sum ||F_{\text{obs}}| - |F_{\text{calc}}|| / \sum |F_{\text{obs}}|$ , where  $F_{\text{obs}}$  and  $F_{\text{calc}}$  are the observe and calculated structure factors, respectively.

<sup>c</sup> $R_{\text{free}}$  was calculated with 5% of the data excluded from the refinement calculation.

117.525 Å,  $c = 193.554$  Å,  $a = b = c = 90^\circ$ .

## 5. Structure Solution and Overall Architecture of the Complex

The phase solution of the SRPK $\Delta$ N1S1:ASF/SF2(BD):AMP-PNP ternary complex was obtained by molecular replacement. The coordinates of truncated apo-SRPK1 $\Delta$ NS1 were used as the search model, which yielded clear rotation function and translation peaks (Figure 6.7). The calculated  $2F_o - F_c$  electron density map revealed clear electron density for AMP-PNP, the secondary structures of the RRM2 of ASF/SF2, and a peptide at the docking groove of SRPK1 (Figures 6.8a and b). The contour of electron density of the peptide is almost identical to that observed in the peptide-bound structure, confirming the presence of an ASF/SF2 motif at the docking groove of SRPK1 (Figure 6.8c). The model of SRPK1 $\Delta$ N1S1 was first refined with several cycles of manual refitting and refinements. Then, the AMP-PNP and a poly-alanine backbone of ASF/SF2 were built to fit the observed electron density. The complex structure was first refined by rigid body refinement, allowing the small lobe and large lobe of the kinase, and the polypeptide of ASF/SF2 to move independently, followed by minimization and annealing. The poly-alanine backbone was replaced by residues of ASF/SF2 subsequently. The R-factor and R-free of the current model of the ternary complex are 29.23% and 38.88% respectively. This model includes residues 68-237 and 477-655 of the kinase, residues 122-194 of ASF/SF2, a 7 residues peptide at the docking groove of SRPK1 and 1 molecule

```

Data line--- FITFUN NMOL 1 RESOLUTION 9 4
Data line--- CRYSTAL ORTH 1
Comment line--- # fixme: caveat about y or z refinement for polar SGs
Data line--- REFSOL AL BE GA X Y Z BF
Comment line--- # Here are the solutions to fit:
Data line--- SOLUT_1 1 5.97 77.79 186.51 0.3233 0.1663 0.3487 37.0 52.9 40.8 1 44.0
Data line--- SOLUT_1 1 5.97 77.79 186.51 0.3218 0.2495 0.3488 26.1 56.3 31.4 3 47.8
Data line--- SOLUT_1 1 5.97 77.79 186.51 0.3232 0.1038 0.3481 25.3 56.8 31.6 6 31.7
Data line--- SOLUT_1 1 5.97 77.79 186.51 0.3234 0.0248 0.3482 25.1 56.5 30.1 5 21.1
Data line--- SOLUT_1 1 5.97 77.79 186.51 0.3237 0.4707 0.3482 24.9 56.8 29.0 7 21.4
Data line--- SOLUT_1 1 5.97 77.79 186.51 0.3216 0.4036 0.3485 24.8 56.7 30.7 4 30.6
Data line--- SOLUT_1 1 5.97 77.79 186.51 0.0176 0.4384 0.0301 22.8 57.7 28.1 2 11.8
Data line--- SOLUT_1 1 5.97 77.79 186.51 0.2627 0.4384 0.0286 22.2 57.9 27.5 9 18.2

```

Figure 6.7 Phasing of SRPK1ΔN1S1:ASF/SF2(BD):AMP-PNP ternary complex structure. Molecular replacement using apo-SRPK1ΔN1S1 structure as a searching model yielded a clear solution as indicated by the significant improvement of R-factor (red dotted box) and correlation factor (blue dotted box).

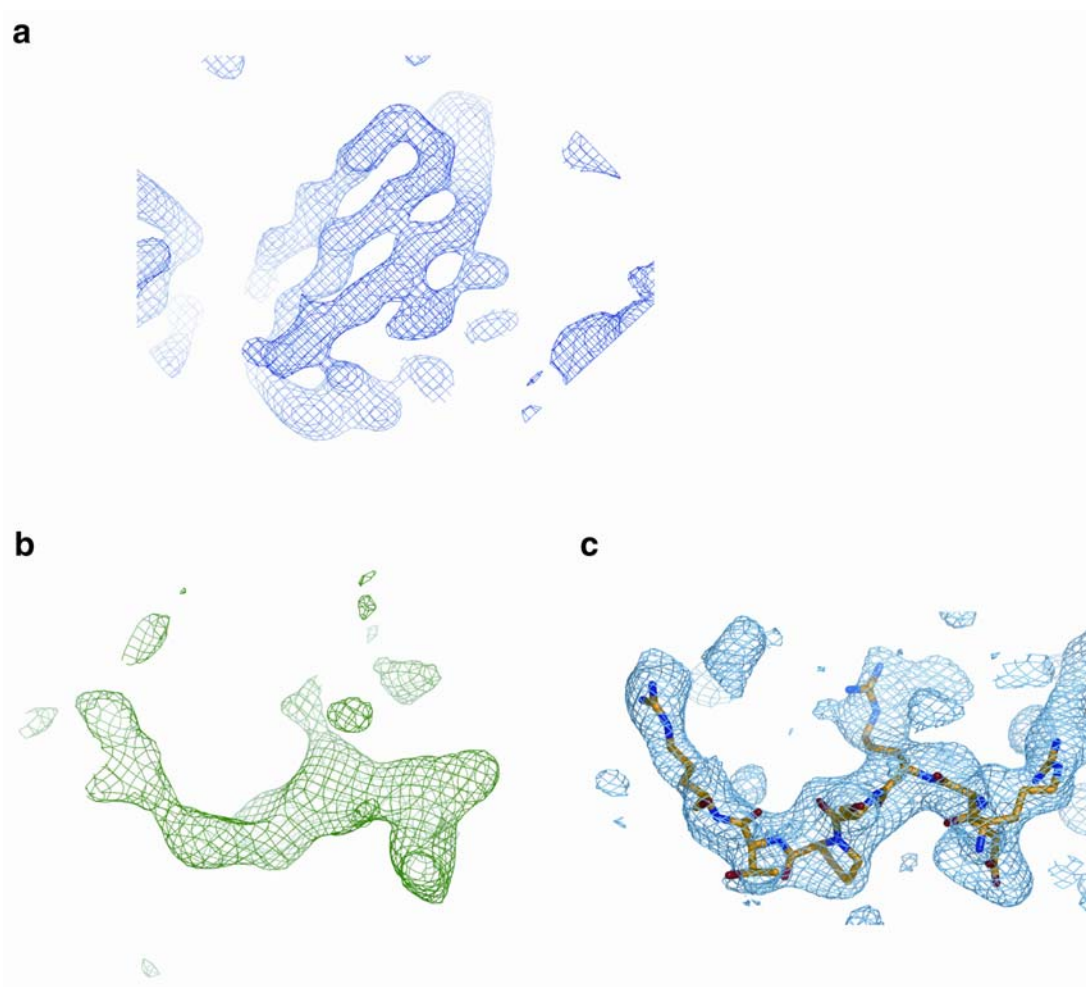


Figure 6.8 Electron density of ASF/SF2 and docking groove peptide. a)  $2F_o - F_c$  electron density map of the RRM2 of ASF/SF2 contoured at  $1\sigma$  clearly shows the presence of secondary structures. b) Clear electron density for a peptide is observed at the SRPK1 docking groove, the contour of the density is similar to that seen for c) the docking peptide that is described in Chapter V.

of AMP-PNP (Figure 6.9).

In the current model of the ternary complex, the SRPK1 molecule adopts similar conformation to that seen in the apo-structure. However, the invariant lysine-glutamate ion pair (K109 and E124 in SRPK1), which is too far apart to interact in the apo-structure, is now in close proximity to form the ionic interaction. This suggests that the small lobe of the kinase might have acquired slight conformational change upon binding to AMP-PNP and ASF/SF2. More structural analysis will be performed to address this observation.

At the current stage of refinement, the observed electron density of the RRM2 of ASF/SF2 starts at residues 122 and ends at 194. Although the RRM2 contains four anti-parallel  $\beta$  strands and two  $\alpha$  helices that adopts the  $\beta 1-\alpha 1-\beta 2-\beta 3-\alpha 2-\beta 4$  topology of the canonical RRM fold, residues 177-182 of ASF/SF2 incorporate into an extra  $\beta$  strand ( $\beta N$ ) preceding  $\beta 4$  and sandwiched between  $\beta 1$  and  $\beta 4$ , forming a five-strand antiparallel  $\beta$  sheet (Figure 6.10). ASF/SF2 packs on the front of SRPK1, interacting with the kinase at both the small and large lobes. The binding interface covers the surfaces of the Gly-rich loop, helices  $\alpha D$  and  $\alpha F$ , and the front of the nucleotide-binding pocket and buries 1302 Å<sup>2</sup> of accessible surface area (Figure 6.11).

## **6. SRPK1:ASF/SF2 Binding Interface**

Since the current model is not completely refined, we will only discuss about the obvious structural features for substrate recognition observed at the binding interface. The conserved SWQDLKD motif of all SR proteins that

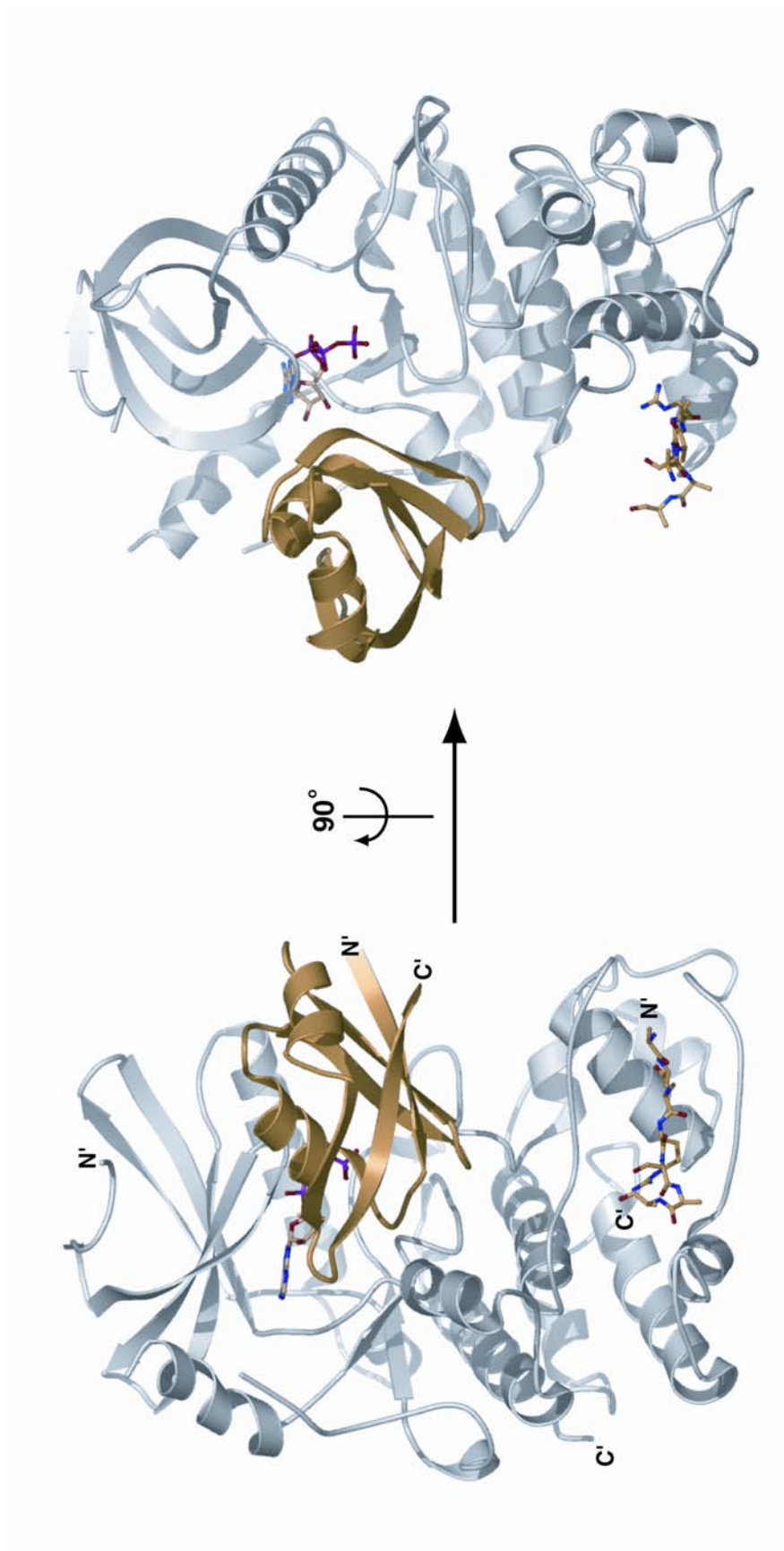


Figure 6.9 Overall structure of SRPK1:ASF/SF2 complex. SRPK1 is colored in silver and ASF/SF2 in gold. The model also contains a 7 residues peptide at the docking groove of SRPK1 and a molecule of AMP-PNP at the nucleotide-binding cleft.

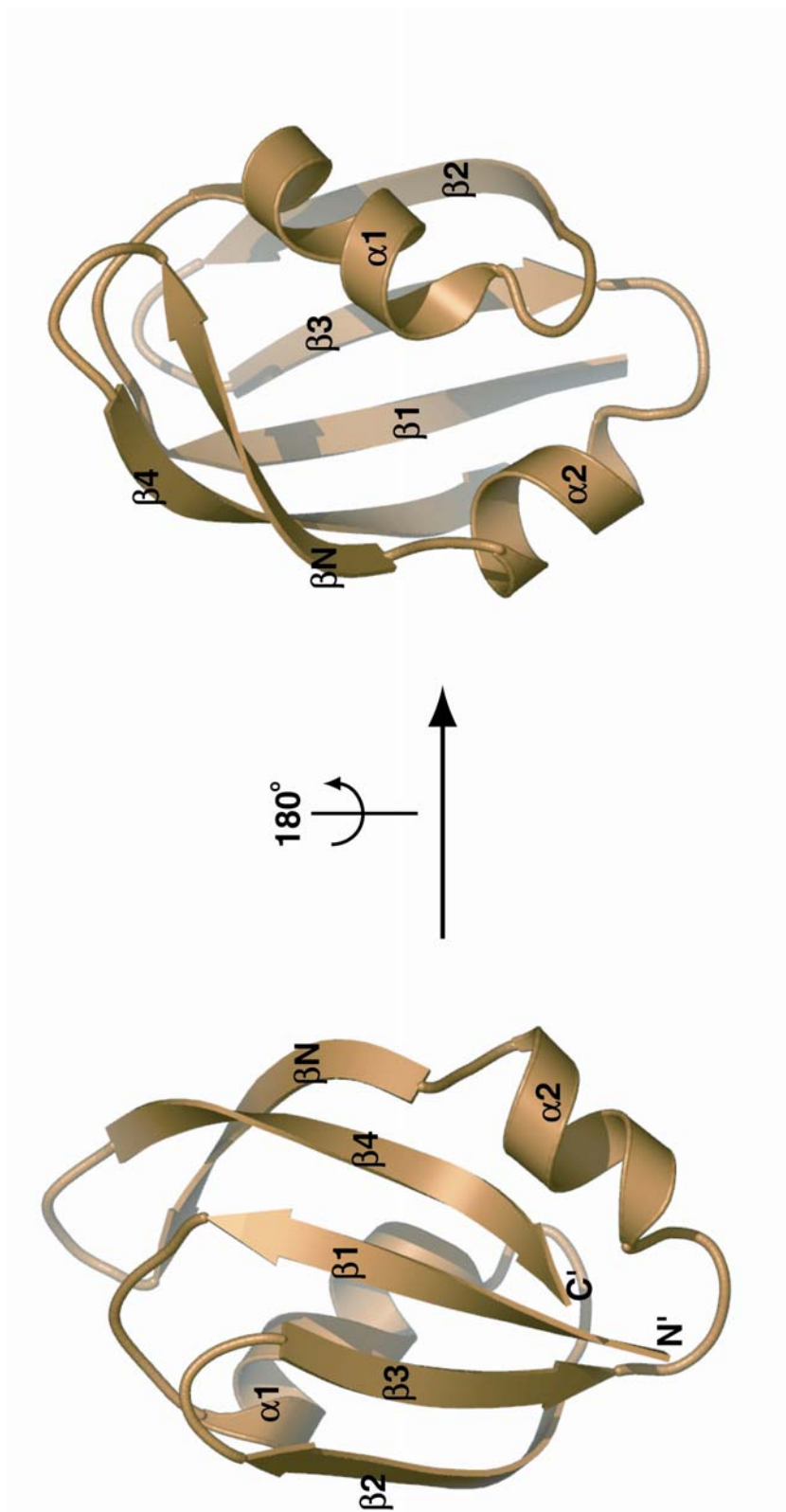


Figure 6.10 Structure of RRM2 of ASF/SF2. RRM2 of ASF/SF2 contains five anti-parallel  $\beta$  strands and two  $\alpha$  helices that are arranged in a  $\beta 1$ - $\alpha 1$ - $\beta 2$ - $\beta 3$ - $\alpha 2$ - $\beta 4$  order with a non-canonical  $\beta$  strand ( $\beta N$ ) sandwiched between  $\beta 1$  and  $\beta 4$ , forming a five-strand antiparallel  $\beta$  sheet



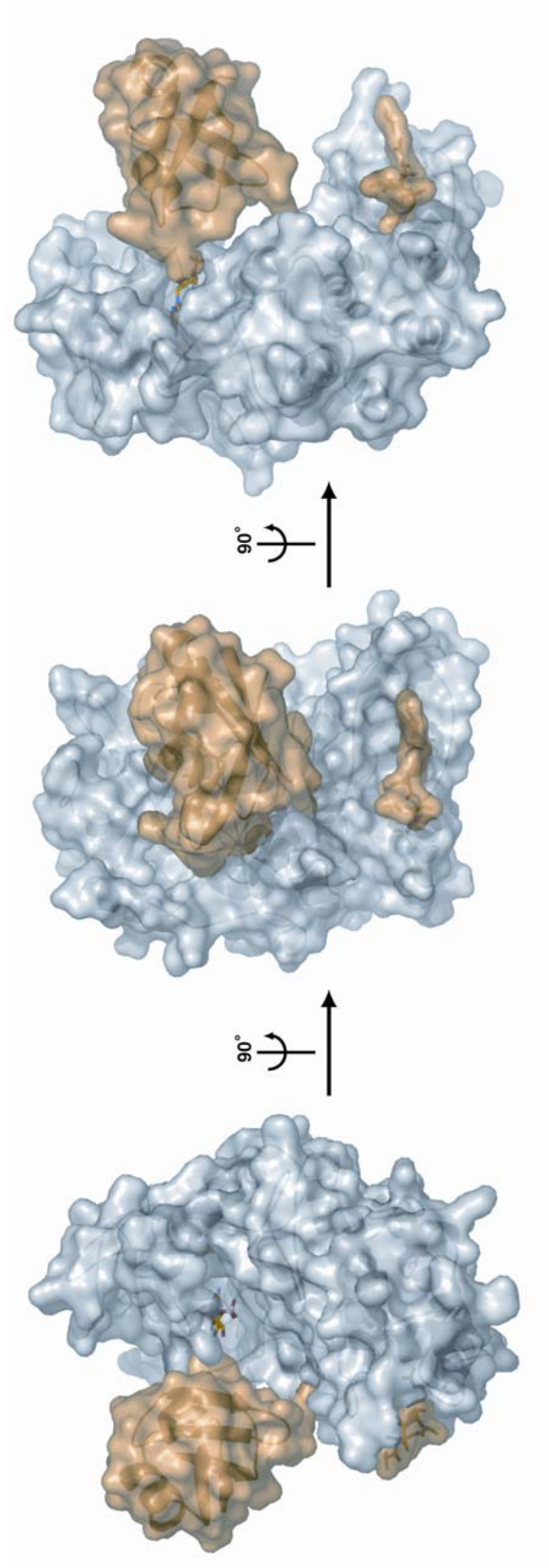


Figure 6.11 Buried surface area in SRPK1:ASF/SF2 complex. Surface renderings of the complex structure from three different sides. The complex binding interface buries 1302 Å of accessible surface area

contain two RRM is located at the C-terminus of the helix  $\alpha 1$  of ASF/SF2. This motif caps the helix and directly contacts the small lobe of SRPK1 at the tip of the Gly-rich loop. In particular, W134 forms a stacking interaction with H90 of SRPK1 and appears to stabilize the tip of the glycine rich loop. Q135 is stacked against W88 and the bottom of  $\beta 1$  of the glycine rich loop and forms hydrogen bond with the backbone of the loop (Figure 6.12a). Besides the interaction between the SWQDLKD region and SRPK1, ASF/SF2 also engages the large lobe of the kinase at helix  $\alpha F$ . R154 from the  $\beta$  turn between  $\beta 2$  and  $\beta 3$  is projected into a pocket formed by residues from helices  $\alpha D$  and  $\alpha F$  of SRPK1 and is stabilized by hydrophobic and ionic interactions, together with hydrogen bonds contributed by the side chains of L173, E543, Y549 and backbone carbonyl of G547 from the kinase (Figure 6.12b).

The additional  $\beta$  strand in ASF/SF2 also allows the formation of an extra hairpin loop, creating three finger-like projections on the same surface of the molecule that serves as the binding surface for SRPK1. In summary, ASF/SF2 projects three “fingers”: the  $\alpha 1/\beta 1$  loop, the two hairpin loops between  $\beta 2$  and  $\beta 3$ , and between  $\beta N$  and  $\beta 4$  respectively, onto the surface of SRPK1, gripping the kinase at both small and large lobes (Figure 6.12c). However, due to the incompleteness of model building, the complete detail interactions between the kinase and the substrates have yet to be elucidated.

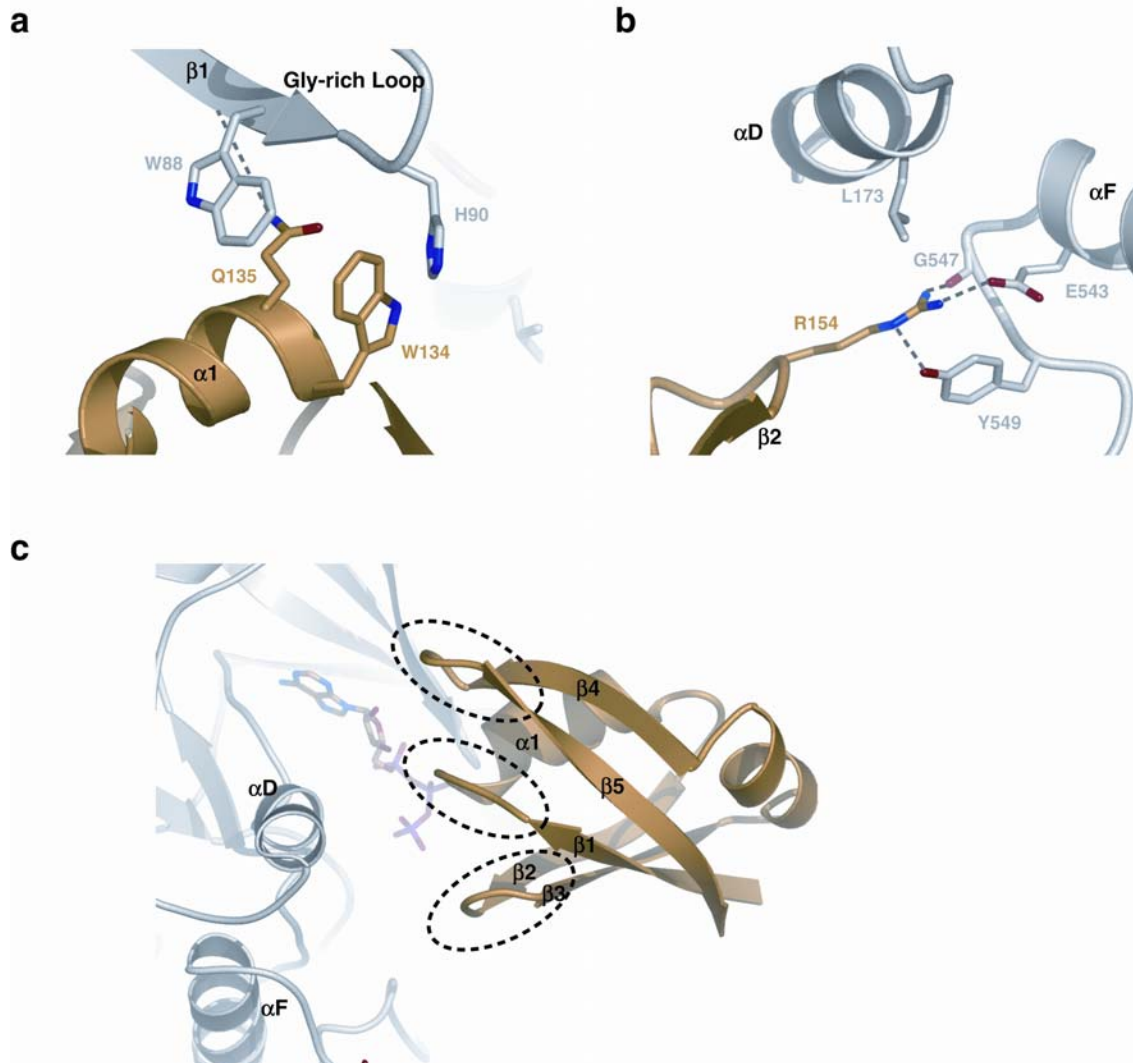


Figure 6.12 Binding interface of SRPK1:ASF/SF2 complex. a) The SWQDLKD motif of ASF/SF2 interacts with the Gly-rich loop of SRPK1. W134 from ASF/SF2 is stacked against H90 of SRPK1; Q135 from ASF/SF2 is stacked against the Gly-rich loop and W88 of SRPK1, and hydrogen bonded to the backbone of Gly-rich loop. b) R154 from ASF/SF2 is projected into a pocket formed by helices  $\alpha D$  and  $\alpha F$ . It forms ionic pair with E543 and hydrogen bonds to Y549 and the backbone carbonyl of G547 from SRPK1. The side chain of R154 is further stabilized by van der Waals contact of L173 from the kinase. c) The RRM2 of ASF/SF2 forms three finger-like projections to contact SRPK1.

## 7. Peptide at the Docking Groove of SRPK1

The present model of the complex contains a 7 residue peptide at the docking groove of SRPK1. Although not all specific side chains for the peptide are modeled, electron density was observed for an arginine at residue 4 (R<sub>4</sub>) and the side chain is modeled at this position (Figure 6.13). The preliminary analysis of this side chain reveals that it mediates interactions identical to that observed in the peptide-bound structure as described in Chapter V. Weak continuous electron density was observed between the C-terminus of RRM2 and the N-terminus of the peptide, suggesting they are indeed connected (Figure 6.13). However, further model building and refinement is required before any conclusive comment about this peptide can be made.

### C. Discussion

While further steps are needed to establish a model of the SRPK1:ASF/SF2 complex, the present model has already provide invaluable information on substrate recognition. Firstly, the RRM2 of ASF/SF2 contains a non-canonical  $\beta$  strand that allows it to adopt an extra hairpin loop that forms part of the binding surface. Secondly, the presence of a docking groove peptide confirms the importance of this groove in substrate binding. Further model building will allow us to identify the sequence of this peptide. Moreover, weak continuous electron density is observed near helix  $\alpha$ G and P+1 loop, suggesting that the completion of the structure will very likely illustrate the substrate binding near the active site and lead us to understand the

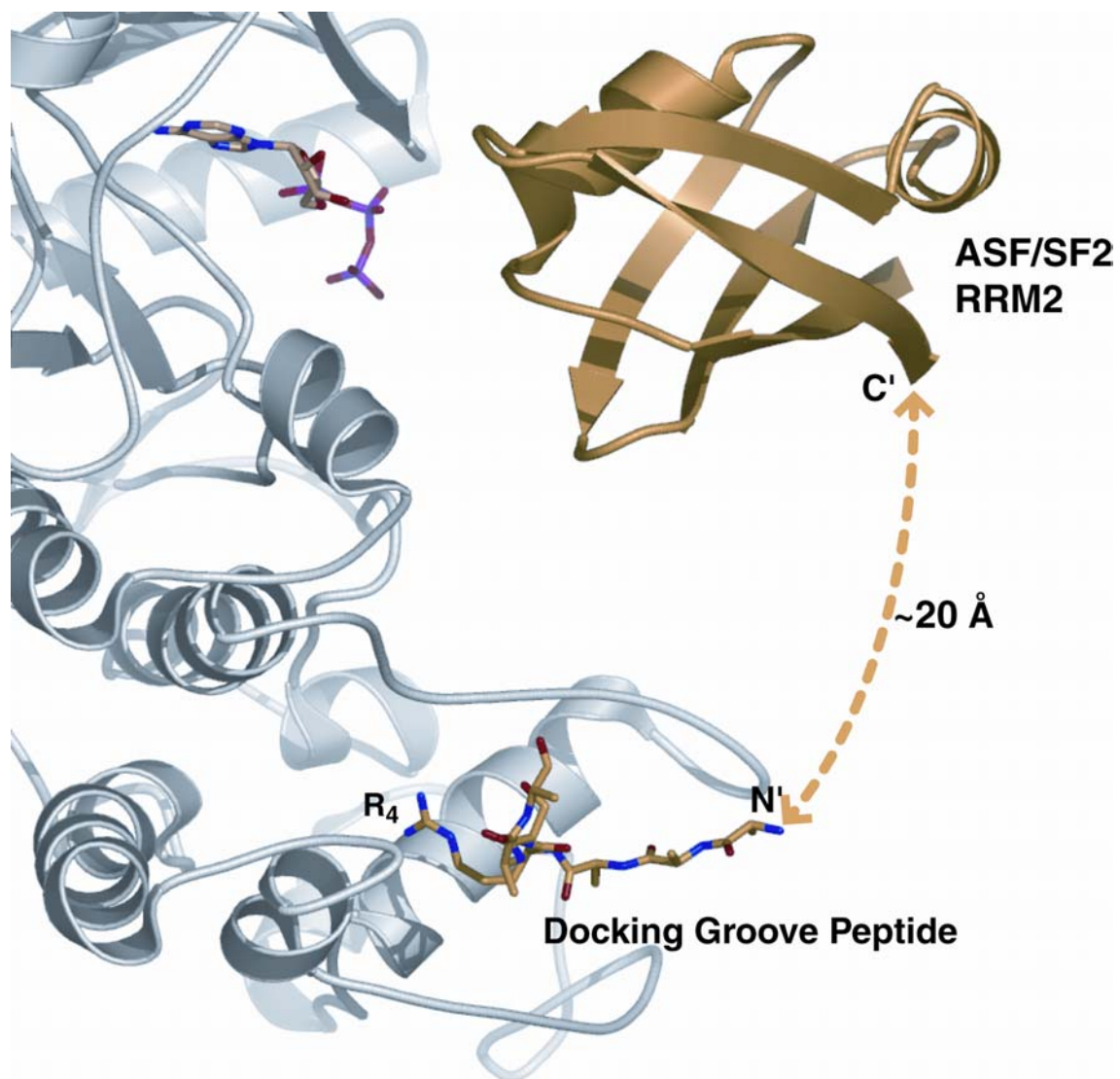


Figure 6.13 Docking groove peptide. In the current model, a 7mer peptide is located at the docking groove of SRPK1. An arginine from the peptide interacts with the kinase through interactions similar to that described in the SRPK1/peptide/ADP ternary complex structure (Chapter V). Weak electron density is observed between the C terminus of RRM2 of ASF/SF2 and the N terminus of the peptide, which are separated by about 20 Å.

directionality of the processive phosphorylation event and the specificity towards RS dipeptides.

Recently, the first kinase-substrate complex structure, PKR:eIF2 $\alpha$ , was solved and the structure has revealed important information on kinase-substrate interaction (Dar et al., 2005). In the case of PKR:eIF2 $\alpha$ , the major binding interface happens between the protein substrate eIF2 $\alpha$  and the helix  $\alpha$ G of PKR. The interactions at the interface are mostly ionic and hydrophobic, involving F489 from PKR deeply buried in a hydrophobic pocket formed on eIF $\alpha$ . Secondly, the helix  $\alpha$ G of PKR is longer than those in other kinases and adopts a noncanonical conformation to form the platform for substrate binding. Whether such an atypical position of the helix is intrinsic or induced upon binding of the substrate is unclear, nevertheless it illustrates that helix  $\alpha$ G plays a central role in substrate binding. Finally, the region of eIF $\alpha$  that contains the phosphorylatable residue is transitioned from a well-ordered state into a disordered state. This structural rearrangement is essential to position the phosphorylatable residue to the catalytic site, implying that the motif containing the phosphorylation site, unlike that seen in other kinase-peptide structures, might be in a transient state when in context of the full length substrate.

The observations made from the PKR:eIF2 $\alpha$  structure are similar to SRPK1:ASF/SF2 complex in many respects. First, in both cases, the helix  $\alpha$ G plays an important role in substrate binding. While the helix  $\alpha$ G adopts a non-canonical conformation in PKR, SRPK1 contains a family-conserved extension

at the N-terminus of helix  $\alpha$ G (Chapter V). Secondly, the region flanking the phosphorylation site in eIF2 $\alpha$  is disordered. In SRPK1:ASF/SF2 structure, only weak electron density is observed for the RS1 domain, suggesting that the RS dipeptides are highly flexible and might be in a transient state due to the redundancy of the RS sequence. Finally, in either case, the structures showed the importance of distal determinants in kinase-substrate recognition in addition to the primary structure surrounding the phosphorylation site.

However, unlike the PKR:eIF2 $\alpha$  structure and other peptide bound kinase structures, substrate binding in SRPK1 is not limited to the substrate binding groove in the large lobe, but spans both small and large lobes of the kinase instead. Although the completed model will very likely reveal more interactions between SRPK1 and ASF/SF2, the current model suggests that the Gly-rich loop and helices  $\alpha$ D and  $\alpha$ F of SRPK1 are involved in substrate binding. Intriguingly, one of the reciprocal binding site in ASF/SF2 is the SWQDLQD motif. This motif is highly conserved in all SR proteins that contain two RRM. Together with the observation that docking motifs are also present in these SR proteins (Chapter V), we speculate that SR proteins containing one or two RRM are two different classes of substrate of SRPK1. Therefore, SRPK1 might possess multiple mechanisms of substrate recognition for 1) SR proteins that only contain one RRM, 2) SR proteins containing two RRM and 3) non-SR proteins.

## **Chapter VII: Conclusion**



### **A. Constitutive Active SRPK1**

The molecular architecture of SRPK1 has revealed that the catalytically competent conformation of the activation loop does not rely upon any direct stabilization mechanism as observed for other active Ser/Thr kinase. SRPK1 integrates several in-trans activation mechanisms observed in different kinases into its structure allowing it to be constitutively active. Surprisingly, our biochemical and molecular dynamics studies showed that the global network of interactions within the three-dimensional structure of SRPK1 integrates into a precise chemical environment required for substrate binding and catalysis, allowing the kinase to create this catalysis-competent chemical environment in multiple ways. We suggest that the functional role of SRPK1 in splicing regulation requires it to efficiently phosphorylate a large number of serines in each substrate molecule. Therefore the kinase has evolved to adopt a unique mechanism to maintain a highly stable and active conformation. Furthermore, the absence of any direct regulatory mechanism strongly suggests that the regulation of SRPK1 maybe achieved through subcellular localization, where the spacer domain plays a central role.

### **B. Docking Interactions between SRPK1 and ASF/SF2**

The X-ray crystal structure of a complex of SRPK1 and a 9mer peptide derived from Npl3p led us to identify a docking groove in SRPK1 and a docking motif in ASF/SF2. Direct interactions between the docking groove and the docking motif contribute to the high binding affinity of the two proteins, and thus

may participate in the processive phosphorylation of ASF/SF2. Even more intriguing is the fact that the docking groove-docking motif interactions between the kinase and the substrate restrict the number of serines in ASF/SF2 that can be phosphorylated by SRPK1. We have shown that in while only serines within the RS1 motif of ASF/SF2 are phosphorylated by SRPK1 and upon deletion of the docking motif, significantly more serines in the RS domain are phosphorylated. In vivo, SRPK1 phosphorylates ASF/SF2 to generate a hypophosphorylated species which is deposited into nuclear speckles. Release of ASF/SF2 from these speckles requires its hyperphosphorylation by Clk/Sty. We showed conclusively that distinct phosphorylation states of ASF/SF2 determine its subcellular and subnuclear localization.

Sequence analyses suggest that the docking groove in SRPKs and docking motif in SR proteins containing two RRM domains are conserved features of the two classes of proteins. We hypothesize that restrictive phosphorylation by SRPK1 might be a common feature of the RS domains of SR proteins containing two RRM domains. However, how SRPKs phosphorylate SR proteins containing only one RRM and other non-SR protein remains unclear.

### **C. SRPK1 Substrate Recognition**

SRPK1 binds to its substrate ASF/SF2 with unusually high affinity and phosphorylates the RS domain with high specificity and efficiency. The ternary complex structure of SRPK1 bound to AMP-PNP and substrate ASF/SF2 has revealed the molecular basis of substrate recognition and the completion of this

structure will help us to understand the structural basis for the high binding affinity and activity. ASF/SF2 contains a non-canonical  $\beta$  strand which allows it to create a unique binding surface for SRPK1. Surprisingly, in contrast to the general belief that the large lobes of kinases serve as the major site of substrate binding, ASF/SF2 contacts both small and large lobes of SRPK1. One of the important recognition determinants in ASF/SF2 is the SWQDLKD motif. This motif is highly conserved in all SR proteins containing two RRMS. The involvement of this conserved motif of ASF/SF2 in substrate recognition suggests that the SR proteins containing one or two RRMs are indeed two unique classes of substrate of SRPK1 and require different modes of substrate recognition.

## References

Aubol, B. E., Chakrabarti, S., Ngo, J., Shaffer, J., Nolen, B., Fu, X. D., Ghosh, G., and Adams, J. A. (2003). Processive phosphorylation of alternative splicing factor/splicing factor 2. *Proc Natl Acad Sci U S A* *100*, 12601-12606.

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

Bax, B., Carter, P. S., Lewis, C., Guy, A. R., Bridges, A., Tanner, R., Pettman, G., Mannix, C., Culbert, A. A., Brown, M. J., *et al.* (2001). The structure of phosphorylated GSK-3 $\beta$  complexed with a peptide, FRATtide, that inhibits beta-catenin phosphorylation. *Structure* *9*, 1143-1152.

Bhattacharyya, R. P., Remenyi, A., Good, M. C., Bashor, C. J., Falick, A. M., and Lim, W. A. (2006). The Ste5 scaffold allosterically modulates signaling output of the yeast mating pathway. *Science* *311*, 822-826.

Biondi, R. M., Komander, D., Thomas, C. C., Lizcano, J. M., Deak, M., Alessi, D. R., and van Aalten, D. M. (2002). High resolution crystal structure of the human PDK1 catalytic domain defines the regulatory phosphopeptide docking site. *Embo J* *21*, 4219-4228.

Biondi, R. M., and Nebreda, A. R. (2003). Signalling specificity of Ser/Thr protein kinases through docking-site-mediated interactions. *Biochem J* *372*, 1-13.

Birney, E., Kumar, S., and Krainer, A. R. (1993). Analysis of the RNA-recognition motif and RS and RGG domains: Conservation in metazoan pre-mRNA splicing factors. *Nucleic Acids Research* *21*, 5803-5816.

Blaustein, M., Pelisch, F., Tanos, T., Munoz, M. J., Wengier, D., Quadrana, L., Sanford, J. R., Muschietti, J. P., Kornblihtt, A. R., Caceres, J. F., *et al.* (2005). Concerted regulation of nuclear and cytoplasmic activities of SR proteins by AKT. *Nat Struct Mol Biol* *12*, 1037-1044.

Bourne, Y., Watson, M. H., Hickey, M. J., Holmes, W., Rocque, W., Reed, S. I., and Tainer, J. A. (1996). Crystal structure and mutational analysis of the human CDK2 kinase complex with cell cycle-regulatory protein CksHs1. *Cell* *84*, 863-874.

Brown, N. R., Noble, M. E., Endicott, J. A., and Johnson, L. N. (1999). The structural basis for specificity of substrate and recruitment peptides for cyclin-dependent kinases. *Nat Cell Biol* 1, 438-443.

Brunger, A. T., Adams, P. D., Clore, G. M., Delano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., *et al.* (1998). Crystallography and NMR system: A new software suite for macromolecular structure determination. *Acta Crystallographica Section D Biological Crystallography* 54, 905-921.

Caceres, J. F., and Krainer, A. R. (1993). Functional analysis of pre-mRNA splicing factor SF2/ASF structural domains. *Embo J* 12, 4715-4726.

Caceres, J. F., Sreaton, G. R., and Krainer, A. R. (1998). A specific subset of SR proteins shuttles continuously between the nucleus and the cytoplasm. *Genes Dev* 12, 55-66.

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

Cao, W., and Garcia-Blanco, M. A. (1998). A serine/arginine-rich domain in the human U1 70k protein is necessary and sufficient for ASF/SF2 binding. *J Biol Chem* 273, 20629-20635.

Cao, W., Jamison, S. F., and Garcia-Blanco, M. A. (1997). Both phosphorylation and dephosphorylation of ASF/SF2 are required for pre-mRNA splicing in vitro. *Rna* 3, 1456-1467.

Cavaloc, Y., Popielarz, M., Fuchs, J. P., Gattoni, R., and Stevenin, J. (1994). Characterization and cloning of the human splicing factor 9G8: a novel 35 kDa factor of the serine/arginine protein family. *Embo J* 13, 2639-2649.

Cazalla, D., Zhu, J., Manche, L., Huber, E., Krainer, A. R., and Caceres, J. F. (2002). Nuclear export and retention signals in the RS domain of SR proteins. *Mol Cell Biol* 22, 6871-6882.

Chandler, S. D., Mayeda, A., Yeakley, J. M., Krainer, A. R., and Fu, X. D. (1997). RNA splicing specificity determined by the coordinated action of RNA recognition motifs in SR proteins. *Proc Natl Acad Sci U S A* 94, 3596-3601.

- Chang, C. I., Xu, B. E., Akella, R., Cobb, M. H., and Goldsmith, E. J. (2002). Crystal structures of MAP kinase p38 complexed to the docking sites on its nuclear substrate MEF2A and activator MKK3b. *Mol Cell* 9, 1241-1249.
- Colwill, K., Pawson, T., Andrews, B., Prasad, J., Manley, J. L., Bell, J. C., and Duncan, P. I. (1996). The Clk/Sty protein kinase phosphorylates SR splicing factors and regulates their intranuclear distribution. *Embo J* 15, 265-275.
- Dajani, R., Fraser, E., Roe, S. M., Yeo, M., Good, V. M., Thompson, V., Dale, T. C., and Pearl, L. H. (2003). Structural basis for recruitment of glycogen synthase kinase 3beta to the axin-APC scaffold complex. *Embo J* 22, 494-501.
- Dajani, R., Fraser, E., Roe, S. M., Young, N., Good, V., Dale, T. C., and Pearl, L. H. (2001). Crystal structure of glycogen synthase kinase 3 beta: structural basis for phosphate-primed substrate specificity and autoinhibition. *Cell* 105, 721-732.
- Dar, A. C., Dever, T. E., and Sicheri, F. (2005). Higher-order substrate recognition of eIF2alpha by the RNA-dependent protein kinase PKR. *Cell* 122, 887-900.
- Daub, H., Blencke, S., Habenberger, P., Kurtenbach, A., Dennenmoser, J., Wissing, J., Ullrich, A., and Cotten, M. (2002). Identification of SRPK1 and SRPK2 as the major cellular protein kinases phosphorylating hepatitis B virus core protein. *J Virol* 76, 8124-8137.
- De Bondt, H. L., Rosenblatt, J., Jancarik, J., Jones, H. D., Morgan, D. O., and Kim, S. H. (1993). Crystal structure of cyclin-dependent kinase 2. *Nature* 363, 595-602.
- Ding, J. H., Zhong, X. Y., Hagopian, J. C., Cruz, M. M., Ghosh, G., Feramisco, J., Adams, J. A., and Fu, X. D. (2006). Regulated cellular partitioning of SR protein-specific kinases in mammalian cells. *Mol Biol Cell* 17, 876-885.
- Duncan, P. I., Howell, B. W., Marius, R. M., Drmanic, S., Douville, E. M., and Bell, J. C. (1995). Alternative splicing of STY, a nuclear dual specificity kinase. *J Biol Chem* 270, 21524-21531.
- Duncan, P. I., Stojdl, D. F., Marius, R. M., Scheit, K. H., and Bell, J. C. (1998). The Clk2 and Clk3 dual-specificity protein kinases regulate the intranuclear distribution of SR proteins and influence pre-mRNA splicing. *Exp Cell Res* 241, 300-308.

Fu, X.-D. (1995). The superfamily of arginine/serine-rich splicing factors. *RNA (New York)* *1*, 663-680.

Fu, X. D., and Maniatis, T. (1992). Isolation of a complementary DNA that encodes the mammalian splicing factor SC35. *Science* *256*, 535-538.

Gonfloni, S., Weijland, A., Kretzschmar, J., and Superti-Furga, G. (2000). Crosstalk between the catalytic and regulatory domains allows bidirectional regulation of Src. *Nat Struct Biol* *7*, 281-286.

Graveley, B. R. (2000). Sorting out the complexity of SR protein functions. *RNA (New York)* *6*, 1197-1211.

Graveley, B. R., Hertel, K. J., and Maniatis, T. (1998). A systematic analysis of the factors that determine the strength of pre-mRNA splicing enhancers. *Embo J* *17*, 6747-6756.

Graveley, B. R., and Maniatis, T. (1998). Arginine/serine-rich domains of SR proteins can function as activators of pre-mRNA splicing. *Mol Cell* *1*, 765-771.

Gui, J.-F., Lane, W. S., and Fu, X.-D. (1994a). A serine kinase regulates intracellular localization of splicing factors in the cell cycle. *Nature (London)* *369*, 678-683.

Gui, J. F., Tronchere, H., Chandler, S. D., and Fu, X. D. (1994b). Purification and characterization of a kinase specific for the serine- and arginine-rich pre-mRNA splicing factors. *Proc Natl Acad Sci U S A* *91*, 10824-10828.

Hanks, S. K., and Quinn, A. M. (1991). Protein kinase catalytic domain sequence database: identification of conserved features of primary structure and classification of family members. *Methods Enzymol* *200*, 38-62.

Hartmann, A. M., Rujescu, D., Giannakouros, T., Nikolakaki, E., Goedert, M., Mandelkow, E. M., Gao, Q. S., Andreadis, A., and Stamm, S. (2001). Regulation of alternative splicing of human tau exon 10 by phosphorylation of splicing factors. *Mol Cell Neurosci* *18*, 80-90.

Hawkins, J., Zheng, S., Frantz, B., and LoGrasso, P. (2000). p38 map kinase substrate specificity differs greatly for protein and peptide substrates. *Arch Biochem Biophys* *382*, 310-313.



Hayes, G. M., Carrigan, P. E., Beck, A. M., and Miller, L. J. (2006). Targeting the RNA splicing machinery as a novel treatment strategy for pancreatic carcinoma. *Cancer Res* 66, 3819-3827.

Haynes, C., and Iakoucheva, L. M. (2006). Serine/arginine-rich splicing factors belong to a class of intrinsically disordered proteins. *Nucleic Acids Res* 34, 305-312.

Heo, Y. S., Kim, S. K., Seo, C. I., Kim, Y. K., Sung, B. J., Lee, H. S., Lee, J. I., Park, S. Y., Kim, J. H., Hwang, K. Y., *et al.* (2004). Structural basis for the selective inhibition of JNK1 by the scaffolding protein JIP1 and SP600125. *Embo J* 23, 2185-2195.

Hishizawa, M., Imada, K., Sakai, T., Ueda, M., Hori, T., and Uchiyama, T. (2005). Serological identification of adult T-cell leukaemia-associated antigens. *Br J Haematol* 130, 382-390.

Holland, P. M., and Cooper, J. A. (1999). Protein modification: docking sites for kinases. *Curr Biol* 9, R329-331.

Huang, Y., Gattoni, R., Stevenin, J., and Steitz, J. A. (2003). SR splicing factors serve as adapter proteins for TAP-dependent mRNA export. *Mol Cell* 11, 837-843.

Huang, Y., and Steitz, J. A. (2001). Splicing factors SRp20 and 9G8 promote the nucleocytoplasmic export of mRNA. *Mol Cell* 7, 899-905.

Huang, Y., Yario, T. A., and Steitz, J. A. (2004). A molecular link between SR protein dephosphorylation and mRNA export. *Proc Natl Acad Sci U S A* 101, 9666-9670.

Hubbard, S. R. (1997). Crystal structure of the activated insulin receptor tyrosine kinase in complex with peptide substrate and ATP analog. *Embo J* 16, 5572-5581.

Hubbard, S. R., Wei, L., Ellis, L., and Hendrickson, W. A. (1994). Crystal structure of the tyrosine kinase domain of the human insulin receptor. *Nature* 372, 746-754.

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

Jacobs, D., Glossip, D., Xing, H., Muslin, A. J., and Kornfeld, K. (1999). Multiple docking sites on substrate proteins form a modular system that mediates recognition by ERK MAP kinase. *Genes Dev* 13, 163-175.

Jeffrey, P. D., Russo, A. A., Polyak, K., Gibbs, E., Hurwitz, J., Massague, J., and Pavletich, N. P. (1995). Mechanism of CDK activation revealed by the structure of a cyclinA-CDK2 complex. *Nature* 376, 313-320.

Johnson, J. M., Castle, J., Garrett-Engele, P., Kan, Z., Loerch, P. M., Armour, C. D., Santos, R., Schadt, E. E., Stoughton, R., and Shoemaker, D. D. (2003). Genome-wide survey of human alternative pre-mRNA splicing with exon junction microarrays. *Science* 302, 2141-2144.

Johnson, L. N., Noble, M. E., and Owen, D. J. (1996). Active and inactive protein kinases: structural basis for regulation. *Cell* 85, 149-158.

Kamachi, M., Le, T. M., Kim, S. J., Geiger, M. E., Anderson, P., and Utz, P. J. (2002). Human autoimmune sera as molecular probes for the identification of an autoantigen kinase signaling pathway. *J Exp Med* 196, 1213-1225.

Kanopka, A., Muhlemann, O., Petersen-Mahrt, S., Estmer, C., Ohrmalm, C., and Akusjarvi, G. (1998). Regulation of adenovirus alternative RNA splicing by dephosphorylation of SR proteins. *Nature* 393, 185-187.

Kataoka, N., Bachorik, J. L., and Dreyfuss, G. (1999). Transportin-SR, a nuclear import receptor for SR proteins. *J Cell Biol* 145, 1145-1152.

Kim, C., Xuong, N. H., and Taylor, S. S. (2005). Crystal structure of a complex between the catalytic and regulatory (RI $\alpha$ ) subunits of PKA. *Science* 307, 690-696.

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

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

- Koizumi, J., Okamoto, Y., Onogi, H., Mayeda, A., Krainer, A. R., and Hagiwara, M. (1999). The subcellular localization of SF2/ASF is regulated by direct interaction with SR protein kinases (SRPKs). *J Biol Chem* 274, 11125-11131.
- Lai, M. C., Lin, R. I., Huang, S. Y., Tsai, C. W., and Tarn, W. Y. (2000). A human importin-beta family protein, transportin-SR2, interacts with the phosphorylated RS domain of SR proteins. *J Biol Chem* 275, 7950-7957.
- Lai, M. C., Lin, R. I., and Tarn, W. Y. (2001). Transportin-SR2 mediates nuclear import of phosphorylated SR proteins. *Proc Natl Acad Sci U S A* 98, 10154-10159.
- Lai, M. C., and Tarn, W. Y. (2004). Hypophosphorylated ASF/SF2 binds TAP and is present in messenger ribonucleoproteins. *J Biol Chem* 279, 31745-31749.
- Lamond, A. I., and Spector, D. L. (2003). Nuclear speckles: a model for nuclear organelles. *Nat Rev Mol Cell Biol* 4, 605-612.
- Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., *et al.* (2001). Initial sequencing and analysis of the human genome. *Nature* 409, 860-921.
- Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *Journal of Appl Crystallogr* 26, 283-291.
- Lazar, G., Schaal, T., Maniatis, T., and Goodman, H. M. (1995). Identification of a plant serine-arginine-rich protein similar to the mammalian splicing factor SF2/ASF. *Proc Natl Acad Sci U S A* 92, 7672-7676.
- Lee, T., Hoofnagle, A. N., Kabuyama, Y., Stroud, J., Min, X., Goldsmith, E. J., Chen, L., Resing, K. A., and Ahn, N. G. (2004). Docking motif interactions in MAP kinases revealed by hydrogen exchange mass spectrometry. *Mol Cell* 14, 43-55.
- Lei, M., Lu, W., Meng, W., Parrini, M. C., Eck, M. J., Mayer, B. J., and Harrison, S. C. (2000). Structure of PAK1 in an autoinhibited conformation reveals a multistage activation switch. *Cell* 102, 387-397.

Lemaire, R., Prasad, J., Kashima, T., Gustafson, J., Manley, J. L., and Lafyatis, R. (2002). Stability of a PKCI-1-related mRNA is controlled by the splicing factor ASF/SF2: a novel function for SR proteins. *Genes Dev* 16, 594-607.

Li, X., and Manley, J. L. (2005). Inactivation of the SR protein splicing factor ASF/SF2 results in genomic instability. *Cell* 122, 365-378.

Lin, S., Xiao, R., Sun, P., Xu, X., and Fu, X. D. (2005). Dephosphorylation-dependent sorting of SR splicing factors during mRNP maturation. *Mol Cell* 20, 413-425.

Liu, H. X., Zhang, M., and Krainer, A. R. (1998). Identification of functional exonic splicing enhancer motifs recognized by individual SR proteins. *Genes Dev* 12, 1998-2012.

Lopato, S., Mayeda, A., Krainer, A. R., and Barta, A. (1996). Pre-mRNA splicing in plants: characterization of Ser/Arg splicing factors. *Proc Natl Acad Sci U S A* 93, 3074-3079.

Lowe, E. D., Noble, M. E., Skamnaki, V. T., Oikonomakos, N. G., Owen, D. J., and Johnson, L. N. (1997). The crystal structure of a phosphorylase kinase peptide substrate complex: kinase substrate recognition. *Embo J* 16, 6646-6658.

Maniatis, T., and Reed, R. (2002). An extensive network of coupling among gene expression machines. *Nature* 416, 499-506.

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

McRee, D. E. (1999). XtalView/Xfit--A versatile program for manipulating atomic coordinates and electron density. *J Struct Biol* 125, 156-165.

Mermoud, J. E., Cohen, P., and Lamond, A. I. (1992). Ser-Thr-specific Protein Phosphatases Are Required for Both Catalytic Steps of Pre-Mrna Splicing. *Nucleic Acids Research* 20, 5263-5269.

Mermoud, J. E., Cohen, P. T. W., and Lamond, A. I. (1994). Regulation of mammalian spliceosome assembly by a protein phosphorylation mechanism. *EMBO (European Molecular Biology Organization) Journal* 13, 5679-5688.

Misteli, T., Caceres, J. F., Clement, J. Q., Krainer, A. R., Wilkinson, M. F., and Spector, D. L. (1998). Serine phosphorylation of SR proteins is required for their recruitment to sites of transcription in vivo. *J Cell Biol* 143, 297-307.

Misteli, T., Caceres, J. F., and Spector, D. L. (1997). The dynamics of a pre-mRNA splicing factor in living cells. *Nature* 387, 523-527.

Moarefi, I., LaFevre-Bernt, M., Sicheri, F., Huse, M., Lee, C. H., Kuriyan, J., and Miller, W. T. (1997). Activation of the Src-family tyrosine kinase Hck by SH3 domain displacement. *Nature* 385, 650-653.

Nagai, K., Oubridge, C., Ito, N., Avis, J., and Evans, P. (1995a). The RNP domain: a sequence-specific RNA-binding domain involved in processing and transport of RNA. *Trends Biochem Sci* 20, 235-240.

Nagai, K., Oubridge, C., Ito, N., Jessen, T. H., Avis, J., and Evans, P. (1995b). Crystal structure of the U1A spliceosomal protein complexed with its cognate RNA hairpin. *Nucleic Acids Symp Ser*, 1-2.

Navaza, J. (1994). AMoRe: an automated package for molecular replacement. *Acta Cryst A* 50, 157-163.

Ngo, J. C., Chakrabarti, S., Ding, J. H., Velazquez-Dones, A., Nolen, B., Aubol, B. E., Adams, J. A., Fu, X. D., and Ghosh, G. (2005). Interplay between SRPK and Clk/Sty kinases in phosphorylation of the splicing factor ASF/SF2 is regulated by a docking motif in ASF/SF2. *Mol Cell* 20, 77-89.

Niefind, K., Guerra, B., Pinna, L. A., Issinger, O. G., and Schomburg, D. (1998). Crystal structure of the catalytic subunit of protein kinase CK2 from *Zea mays* at 2.1 Å resolution. *Embo J* 17, 2451-2462.

Nikolakaki, E., Kohen, R., Hartmann, A. M., Stamm, S., Georgatsou, E., and Giannakouros, T. (2001). Cloning and characterization of an alternatively spliced form of SR protein kinase 1 that interacts specifically with scaffold attachment factor-B. *J Biol Chem* 276, 40175-40182.

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

Nolen, B., Yun, C. Y., Wong, C. F., McCammon, J. A., Fu, X. D., and Ghosh, G. (2001). The structure of Sky1p reveals a novel mechanism for constitutive activity. *Nat Struct Biol* 8, 176-183.

Otwinowski, Z., and Minor, W. (1997). Processing of x-ray diffraction data collected in oscillation mode., In *Methods in Enzymology*, R. M. Sweet, and C. W. Carter, eds. (New York: Academic Press), pp. 307-326.

Papoutsopoulou, S., Nikolakaki, E., Chalepakis, G., Kruff, V., Chevallier, P., and Giannakouros, T. (1999a). SR protein-specific kinase 1 is highly expressed in testis and phosphorylates protamine 1. *Nucleic Acids Res* 27, 2972-2980.

Papoutsopoulou, S., Nikolakaki, E., and Giannakouros, T. (1999b). SRPK1 and LBR protein kinases show identical substrate specificities. *Biochem Biophys Res Commun* 255, 602-607.

Prasad, J., Colwill, K., Pawson, T., and Manley, J. L. (1999). The protein kinase Clk/Sty directly modulates SR protein activity: both hyper- and hypophosphorylation inhibit splicing. *Mol Cell Biol* 19, 6991-7000.

Prasad, J., and Manley, J. L. (2003). Regulation and substrate specificity of the SR protein kinase Clk/Sty. *Mol Cell Biol* 23, 4139-4149.

Reed, R. (1996). Initial splice-site recognition and pairing during pre-mRNA splicing. *Curr Opin Genet Dev* 6, 215-220.

Reed, R., and Hurt, E. (2002). A conserved mRNA export machinery coupled to pre-mRNA splicing. *Cell* 108, 523-531.

Roscigno, R. F., and Garcia-Blanco, M. A. (1995). SR proteins escort the U4/U6.U5 tri-snRNP to the spliceosome. *Rna* 1, 692-706.

Rossi, F., Labourier, E., Forne, T., Divita, G., Derancourt, J., Riou, J. F., Antoine, E., Cathala, G., Brunel, C., and Tazi, J. (1996). Specific phosphorylation of SR proteins by mammalian DNA topoisomerase I. *Nature (London)* 381, 80-82.

Roth, M. B., Murphy, C., and Gall, J. G. (1990). A monoclonal antibody that recognizes a phosphorylated epitope stains lampbrush chromosome loops and small granules in the amphibian germinal vesicle. *J Cell Biol* 111, 2217-2223.

Russo, A. A., Jeffrey, P. D., and Pavletich, N. P. (1996). Structural basis of cyclin-dependent kinase activation by phosphorylation. *Nat Struct Biol* 3, 696-700.

Salesse, S., Dylla, S. J., and Verfaillie, C. M. (2004). p210BCR/ABL-induced alteration of pre-mRNA splicing in primary human CD34+ hematopoietic progenitor cells. *Leukemia* 18, 727-733.

Sanford, J. R., Ellis, J. D., Cazalla, D., and Caceres, J. F. (2005). Reversible phosphorylation differentially affects nuclear and cytoplasmic functions of splicing factor 2/alternative splicing factor. *Proc Natl Acad Sci U S A* 102, 15042-15047.

Sanford, J. R., Gray, N. K., Beckmann, K., and Caceres, J. F. (2004). A novel role for shuttling SR proteins in mRNA translation. *Genes Dev* 18, 755-768.

Sanz, G., Mir, L., and Jacquemin-Sablon, A. (2002). Bleomycin resistance in mammalian cells expressing a genetic suppressor element derived from the SRPK1 gene. *Cancer Res* 62, 4453-4458.

Schaal, T. D., and Maniatis, T. (1999). Selection and characterization of pre-mRNA splicing enhancers: identification of novel SR protein-specific enhancer sequences. *Mol Cell Biol* 19, 1705-1719.

Schenk, P. W., Boersma, A. W., Brandsma, J. A., den Dulk, H., Burger, H., Stoter, G., Brouwer, J., and Nooter, K. (2001). SKY1 is involved in cisplatin-induced cell kill in *Saccharomyces cerevisiae*, and inactivation of its human homologue, SRPK1, induces cisplatin resistance in a human ovarian carcinoma cell line. *Cancer Res* 61, 6982-6986.

Schenk, P. W., Stoop, H., Bokemeyer, C., Mayer, F., Stoter, G., Oosterhuis, J. W., Wiemer, E., Looijenga, L. H., and Nooter, K. (2004). Resistance to platinum-containing chemotherapy in testicular germ cell tumors is associated with downregulation of the protein kinase SRPK1. *Neoplasia* 6, 297-301.

Sciabica, K. S., Dai, Q. J., and Sandri-Goldin, R. M. (2003). ICP27 interacts with SRPK1 to mediate HSV splicing inhibition by altering SR protein phosphorylation. *Embo J* 22, 1608-1619.

Screaton, G. R., Caceres, J. F., Mayeda, A., Bell, M. V., Plebanski, M., Jackson, D. G., Bell, J. I., and Krainer, A. R. (1995). Identification and

characterization of three members of the human SR family of pre-mRNA splicing factors. *Embo J* 14, 4336-4349.

Sessa, F., Mapelli, M., Ciferri, C., Tarricone, C., Areces, L. B., Schneider, T. R., Stukenberg, P. T., and Musacchio, A. (2005). Mechanism of Aurora B activation by INCENP and inhibition by hesperadin. *Mol Cell* 18, 379-391.

Shen, H., and Green, M. R. (2004). A pathway of sequential arginine-serine-rich domain-splicing signal interactions during mammalian spliceosome assembly. *Mol Cell* 16, 363-373.

Shen, H., Kan, J. L., and Green, M. R. (2004). Arginine-serine-rich domains bound at splicing enhancers contact the branchpoint to promote prespliceosome assembly. *Mol Cell* 13, 367-376.

Shin, C., and Manley, J. L. (2002). The SR protein SRp38 represses splicing in M phase cells. *Cell* 111, 407-417.

Siebel, C. W., Feng, L., Guthrie, C., and Fu, X. D. (1999). Conservation in budding yeast of a kinase specific for SR splicing factors. *Proc Natl Acad Sci U S A* 96, 5440-5445.

Smith, C. W., and Valcarcel, J. (2000). Alternative pre-mRNA splicing: the logic of combinatorial control. *Trends Biochem Sci* 25, 381-388.

Song, H., Hanlon, N., Brown, N. R., Noble, M. E., Johnson, L. N., and Barford, D. (2001). Phosphoprotein-protein interactions revealed by the crystal structure of kinase-associated phosphatase in complex with phosphoCDK2. *Mol Cell* 7, 615-626.

Soret, J., Gattoni, R., Guyon, C., Sureau, A., Popielarz, M., Le Rouzic, E., Dumon, S., Apiou, F., Dutrillaux, B., Voss, H., *et al.* (1998). Characterization of SRp46, a novel human SR splicing factor encoded by a PR264/SC35 retroseudogene. *Mol Cell Biol* 18, 4924-4934.

Staknis, D., and Reed, R. (1994). SR proteins promote the first specific recognition of Pre-mRNA and are present together with the U1 small nuclear ribonucleoprotein particle in a general splicing enhancer complex. *Mol Cell Biol* 14, 7670-7682.



Tacke, R., and Manley, J. L. (1995). The human splicing factors ASF/SF2 and SC35 possess distinct, functionally significant RNA binding specificities. *Embo J* 14, 3540-3551.

Takano, M., Koyama, Y., Ito, H., Hoshino, S., Onogi, H., Hagiwara, M., Furukawa, K., and Horigome, T. (2004). Regulation of binding of lamin B receptor to chromatin by SR protein kinase and cdc2 kinase in *Xenopus* egg extracts. *J Biol Chem* 279, 13265-13271.

Takeuchi, M., and Yanagida, M. (1993). A mitotic role for a novel fission yeast protein kinase dsk1 with cell cycle stage dependent phosphorylation and localization. *Mol Biol Cell* 4, 247-260.

Tang, Z., Yanagida, M., and Lin, R. J. (1998). Fission yeast mitotic regulator Dsk1 is an SR protein-specific kinase. *J Biol Chem* 273, 5963-5969.

Tange, T. O., and Kjems, J. (2001). SF2/ASF binds to a splicing enhancer in the third HIV-1 tat exon and stimulates U2AF binding independently of the RS domain. *J Mol Biol* 312, 649-662.

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

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

Tarn, W. Y., and Steitz, J. A. (1995). Modulation of 5' splice site choice in pre-messenger RNA by two distinct steps. *Proc Natl Acad Sci U S A* 92, 2504-2508.

Tazi, J., Kornstaedt, U., Rossi, F., Jeanteur, P., Cathala, G., Brunel, C., and Luehrmann, R. (1993). Thiophosphorylation of U1-70K protein inhibits pre-mRNA splicing. *Nature (London)* 363, 283-286.

Wang, H. Y., Lin, W., Dyck, J. A., Yeakley, J. M., Songyang, Z., Cantley, L. C., and Fu, X. D. (1998a). SRPK2: a differentially expressed SR protein-specific kinase involved in mediating the interaction and localization of pre-mRNA splicing factors in mammalian cells. *J Cell Biol* 140, 737-750.

- Wang, H. Y., Xu, X., Ding, J. H., Bermingham, J. R., Jr., and Fu, X. D. (2001). SC35 plays a role in T cell development and alternative splicing of CD45. *Mol Cell* 7, 331-342.
- Wang, J., Xiao, S. H., and Manley, J. L. (1998b). Genetic analysis of the SR protein ASF/SF2: interchangeability of RS domains and negative control of splicing. *Genes Dev* 12, 2222-2233.
- Wiechmann, S., Czajkowska, H., de Graaf, K., Grotzinger, J., Joost, H. G., and Becker, W. (2003). Unusual function of the activation loop in the protein kinase DYRK1A. *Biochem Biophys Res Commun* 302, 403-408.
- Xiao, S. H., and Manley, J. L. (1997). Phosphorylation of the ASF/SF2 RS domain affects both protein-protein and protein-RNA interactions and is necessary for splicing. *Genes Dev* 11, 334-344.
- Xu, X., Yang, D., Ding, J. H., Wang, W., Chu, P. H., Dalton, N. D., Wang, H. Y., Bermingham, J. R., Jr., Ye, Z., Liu, F., *et al.* (2005). ASF/SF2-regulated CaMKII $\delta$  alternative splicing temporally reprograms excitation-contraction coupling in cardiac muscle. *Cell* 120, 59-72.
- Yang, J., Cron, P., Good, V. M., Thompson, V., Hemmings, B. A., and Barford, D. (2002). Crystal structure of an activated Akt/protein kinase B ternary complex with GSK3-peptide and AMP-PNP. *Nat Struct Biol* 9, 940-944.
- Yeakley, J. M., Tronchere, H., Olesen, J., Dyck, J. A., Wang, H. Y., and Fu, X. D. (1999). Phosphorylation regulates in vivo interaction and molecular targeting of serine/arginine-rich pre-mRNA splicing factors. *J Cell Biol* 145, 447-455.
- Yun, C. Y., Velazquez-Dones, A. L., Lyman, S. K., and Fu, X. D. (2003). Phosphorylation-dependent and -independent nuclear import of RS domain-containing splicing factors and regulators. *J Biol Chem* 278, 18050-18055.
- Zahler, A. M., Lane, W. S., Stolk, J. A., and Roth, M. B. (1992). SR proteins: a conserved family of pre-mRNA splicing factors. *Genes Dev* 6, 837-847.
- Zhang, W. J., and Wu, J. Y. (1996). Functional properties of p54, a novel SR protein active in constitutive and alternative splicing. *Mol Cell Biol* 16, 5400-5408.
- Zhang, Z., and Krainer, A. R. (2004). Involvement of SR proteins in mRNA surveillance. *Mol Cell* 16, 597-607.

Zheng, Y., Fu, X. D., and Ou, J. H. (2005). Suppression of hepatitis B virus replication by SRPK1 and SRPK2 via a pathway independent of the phosphorylation of the viral core protein. *Virology*.

Zuo, P., and Manley, J. L. (1993). Functional domains of the human splicing factor ASF/SF2. *Embo J* 12, 4727-4737.