Title
Directional phosphorylation and dephosphorylation of the SR protein ASF/SF2

Permalink
https://escholarship.org/uc/item/1pf654m4

Author
Ma, Chen-Ting

Publication Date
2009

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA, SAN DIEGO

Directional Phosphorylation and Dephosphorylation of the SR Protein ASF/SF2

in

Biomedical Sciences

by

Chen-Ting Ma

Committee in charge:

    Professor Joseph A. Adams, Chair

    Professor Xiang-Dong Fu

    Professor Gourisankar Ghosh

    Professor Patricia Jennings

    Professor Susan Taylor

2009

The dissertation of Chen-Ting Ma is approved, and it is acceptable in quality and form for publication on microfilm or electronically

_____

_____

_____

_____

_____

Chair

University of California, San Diego

2009

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

SRPK: Serine arginine protein kinase

CLK (Clk/Sty):   Cdk-like kinase/serine-threonine-tyrosine

ATP: Adenosine 5'-triphosphate

GST: Glutathione S-transferase

wt: wild-type

ASF/SF2:   Alternative Splicing Factor/Splicing Factor 2

PP1(γ):   Protein phosphatase 1 (gamma isoform)

Pin1:   Peptidyl-prolyl cis/trans isomerase

FPLC: Fast protein liquid chromatography

RRM: RNA-recognition motif

RS/SR: domain/protein rich in serine-arginine dipeptide repeats

mRNA: messenger ribonucleic acid

snRNP: small nuclear ribonucleoprotein

hnRNP: heteronuclear ribonucleoprotein

ESE: exonic splicing enhancer

CIP: calf intestinal alkaline phosphatase

PMSF: phenylmethylsulfonyl fluoride

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I thank my father Kuo-Tung Ma, my mother Hung-Ti Yu, and my sister Yu-Ting Ma for their unconditional love and support throughout my career.

Chapter 3 is, in part, a reprint of the material as it appears in *The Journal of Molecular Biology*, 2008, 376(1), 55-68, Ma, C.T., Velazquez-Dones, A., Hagopian, J.C., Ghosh, G., Fu, X.D., Adams, J.A. The dissertation author was the primary researcher and author of this publication. Additional text of Chapter 3 is a reprint of the material as it appears in *The Journal of Molecular Biology*, 2009, 390(4), 618-34, Ma, C.T., Hagopian, J.C., Ghosh, G., Fu, X.D., Adams, J.A. The dissertation author was the primary researcher and author of this publication.

Chapter 4 is, in part, a reprint of the material as it appears in *The Journal of Molecular Biology*, 2008, 328(4), 894-909, Hagopian, J.C., Ma, C.T., Meade, B.R., Albuquerque, C.P., Ngo, J., Ghosh, G., Jennings, P.A., Fu, X.D., Adams, J.A. The dissertation author was a collaborative researcher and co-author of this publication. Additional text of Chapter 4 is a reprint of the material as it appears in *The Journal of Molecular Biology*, 2009, 390(4), 618-34, Ma, C.T., Hagopian, J.C., Ghosh, G., Fu, X.D., Adams, J.A. The dissertation author was the primary researcher and author of this publication.

Chapter 5 is, in part, a reprint of the material as it appears in *Molecular Cell,* 2008, 29(5), 563-576, Ngo, J., Giang, K., Chakrabarti, S., Ma, C.T., Huynh, N., Hagopian, J.C., Dorrestein, P.C., Fu, X.D., Adams, J.A., Ghosh, G. The dissertation author was a collaborative researcher and co-author of this publication.

VITA

EDUCATION

1995-1999   B.S., University of California, Berkeley
1999-2000   M.Eng., University of California, San Diego
2004-2009   Ph.D., University of California, San Diego

HONORS

1995-1999   UC Berkeley Chancellor's Scholarship
2004-2007   NIGMS UCSD Pharmacology Training Grant
2004-2009   UCSD Chancellor's Fellowship

ADDITIONAL RESEARCH EXPERIENCE

2000-2004   Research Associate I-III, Triad Therapeutics, Inc., Biochemistry Department

PUBLICATIONS

Ma, C.T., Hagopian, J.C., Ghosh, G., Fu, X.D., Adams, J.A. "Regiospecific Phosphorylation Control of the SR Protein ASF/SF2 by SRPK1." *J Mol Biol.* 2009, 390(4): 618-34.

Hagopian, J.C., Ma, C.T., Meade, B.R., Albuquerque, C.P., Ngo, J., Ghosh, G., Jennings, P.A., Fu, X.D., Adams, J.A. "Adaptable Molecular Interactions Guide Phosphorylation of the SR Protein ASF/SF2 by SRPK1." *J Mol Biol.* 2008, 382(4):894-909.

Ngo, J.C., Giang, K., Chakrabarti, S., Ma, C.T., Huynh, N., Hagopian, J.C., Dorrestein, P.C., Fu X.D., Adams, J.A., Ghosh, G. "A sliding docking interaction is essential for sequential and processive phosphorylation of an SR protein by SRPK1." *Mol Cell.* 2008, 29(5):563-76.

Ma, C.T., Velazquez-Dones, A., Hagopian, J.C., Ghosh, G., Fu, X.D., Adams, J.A. "Ordered multi-site phosphorylation of the splicing factor ASF/SF2 by SRPK1." *J Mol Biol*. 2008, 376(1):55-68.

Velazquez-Dones, A., Hagopian, J.C., Ma, C.T., Zhong, X.Y., Zhou, H., Ghosh, G., Fu, X.D., Adams, J.A. "Mass spectrometric and kinetic analysis of ASF/SF2 phosphorylation by SRPK1 and Clk/Sty." *J Biol Chem*. 2005, 280(50):41761-8.

ABSTRACT OF THE DISSERTATION

Directional Phosphorylation and Dephosphorylation of the SR Protein ASF/SF2

by

Chen-Ting Ma

Doctor of Philosophy in Biomedical Science

University of California, San Diego, 2009

Professor Joseph A. Adams, Chair

Constitutive and alternative splicing are necessary steps in mRNA processing that contributes to a diverse proteome, but errors in splicing machinery lead to numerous degenerative diseases and cancers. The SR proteins are essential splicing factors that control splice-site selection as well as the transport of processed mRNA for translation. The SR proteins consist of one or two RNA-binding domains and an RS domain rich in arginine-serine repeats. Processive poly-phosphorylation of the RS domain by splicing kinases such as SRPK1, and subsequent dephosphorylation by nuclear phosphatases such as PP1 regulate cellular localization and mRNA processing functions of SR proteins.

In this thesis we focused on SRPK1-catalyzed phosphorylation and PP1-catalyzed dephosphorylation of the prototypical SR protein, ASF/SF2. Our overall goals were to establish whether these two enzymes add and remove phosphates in a directional manner and determine specific sequence elements that control initiation and sequential modification of the RS domain of ASF/SF2. A novel LysC digestion phosphomapping technique was developed to observe the order of phosphorylation. We found that SRPK1 phosphorylates ASF/SF2 in a C->N direction preferring to initiate between S221-S225 of ASF/SF2, but the kinase can adapt to other start sequences if the initiation box is made dysfunctional through mutation.

Having established that SRPK1 is a directional kinase with a strong preference for phosphorylation initiation in the center of the RS domain, we then explored structural determinants within the enzyme and substrate that might control this novel mechanism of phosphorylation. Mutagenesis within the RS domain revealed that an uninterrupted RS/SR stretch is necessary for initiation preference. Using the SRPK1:ASF/SF2 complex as a guide for mutagenesis studies, we found that RRM2 (but not RRM1) and N'-RS1 of the RS domain of ASF/SF2 form important and cooperative contacts with SRPK1 that guide directional phosphorylation. Thus, while RS/SR content in the RS domain alone governs regiospecificity, directional phosphorylation is controlled by distal structural elements offered by RRM2 and the docking groove in SRPK1.

To analyze how phosphates are removed from the RS domain of ASF/SF2, we studied the dephosphorylation reaction catalyzed by the nuclear protein phosphatase, PP1. PP1 was found to remove phosphates from the RS1 segment of ASF/SF2 in a single kinetic phase suggesting that a common rate-limiting step exists. Deletion analyses indicated that RRM1 protects against dephosphorylation. Through digestion mapping, we found that PP1 dephosphorylates ASF/SF2 in an N->C direction and that removal of RRM1 leads to random dephosphorylation. These data suggest that while RRM2 is important for the directional phosphorylation of ASF/SF2, RRM1 is important for the directional dephosphorylation. Overall, the data suggest that the opposing directional activities of SRPK1 and PP1 may foster the preferential modification of serines in the C-terminal end of RS1 in ASF/SF2, with possible biological relevance.

Chapter 1

Introduction

A. mRNA Processing and the Role of SR Proteins

The bridge between genetics and proteomics has three steps: DNA genes, mRNA transcripts, and finally translated proteins. After transcription occurs with the help of RNA polymerase, the resultant pre-mRNA undergoes extensive modification: 5' capping, 3' poly-adenylation, and splicing before the mature mRNA is ready for translation into actual proteins. Whereas the first two steps prolong mRNA stability, splicing is necessary to stitch together short protein-coding segments called exons and exclude long intervening non-coding sequences called introns [1, 2]. The definition of exon boundaries called 5' and 3' splice sites depends on splicing signals that are local mRNA sequences recognized by splicing machinery within the cell nucleus. Selection of splice sites leads to two types of splicing events: constitutive and alternative splicing. In constitutive splicing, all exons defined by the strong splice sites are connected together, whereas in alternative splicing exons defined by weaker splice sites may be chosen [3], sometimes resulting in truncated exons or missing exons (Fig. 1.1). Protein isoforms translated from these alternative splicing transcripts may have enhanced, deficient, or opposing functions. Investigation into alternative splicing has helped to resolve the mysterious disparity between human genome size and human systemic complexity. Recent studies suggested that over 90% of human genes have alternative splice forms [4, 5], and together with post-translational modifications such as phosphorylation and methylation, a complex proteome is created from a relatively small genome to respond to the diverse environmental stimuli and challenges.

The earliest splicing events in evolutionary history occurred with self-splicing of rRNA in which introns are removed through lariat formations, and this process is still observed in lower-order eukaryotes [2]. In humans, pre-mRNA splicing occurs co-transcriptionally for increased efficiency [6], in a dynamic RNA-protein complex called the spliceosome (Fig. 1.2), with more than 100 components identified [7-9]. In the initiation complex snRNP (small nuclear ribonucleoprotein) U1 interacts with the 5' splice site, U2AF (U2 snRNP auxillary factor) with the 3' splice site, and SF1 (splicing factor 1) with the branchpoint site, to recognize all consensus cis-acting elements. This complex then facilitates the binding of U2 at the branchpoint to form complex A and commit to the splicing process. Upon association of the
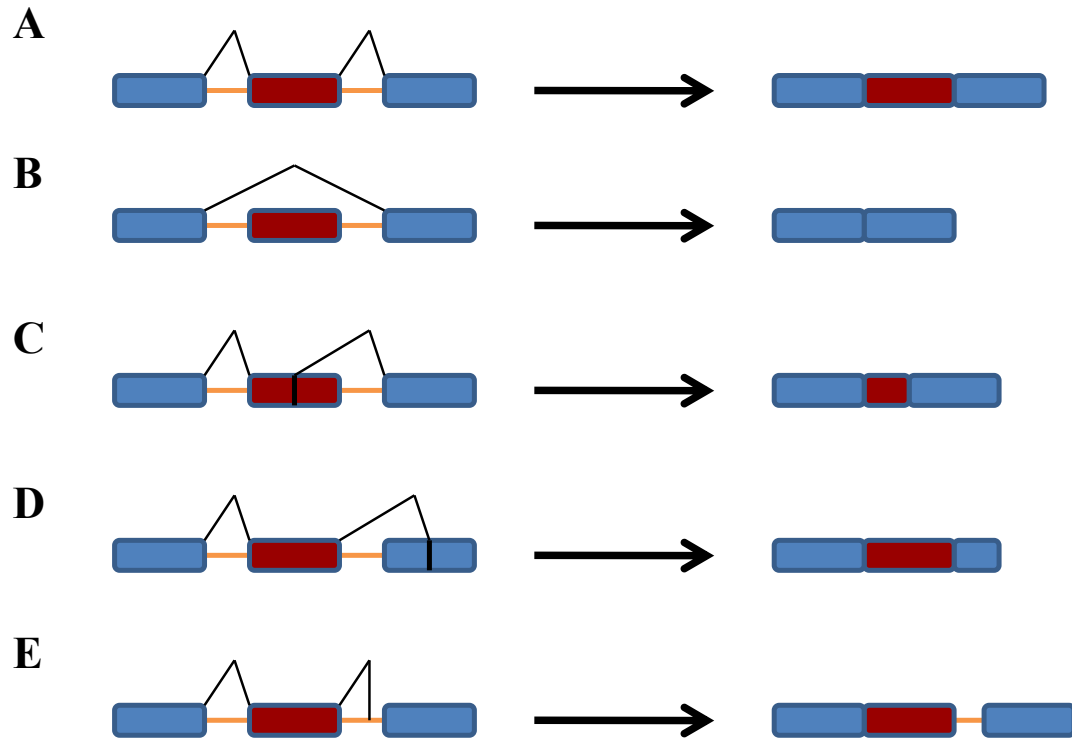
Figure 1.1 Examples of Alternative Splicing. Adapted from [3]. Weaker splice sites may be adopted to generate a large number of distinct transcripts. A) Exon inclusion. B) Exon exclusion. C) Alternative 5' splice site. D) Alternative 3' splice site. E) Intron inclusion.

U4/U6/U5 tri-snRNP in complex B, a large conformational change occurs to form the catalytically active complex C.

SR proteins are essential splicing factors that interact with the RNA and protein components of the spliceosome complex (Fig. 1.2) and participate in different stages of mRNA processing. They are characterized by one or more RRM (RNA-recognition motif) domains and usually one RS (serine-arginine dipeptide repeats) domain (Fig. 1.3A). In constitutive splicing, SR proteins help define exons and bridge across introns through association between their own RS domain and that of the 70 kDa subunit of U1 (U1-70K) and 35 kDa subunit of U2AF (U2AF[35]) [10, 11]. These protein-protein interactions help define 5'- and 3'-splice sites across great lengths [12]. Recruitment of U4/U6/U5 and splicing catalysis also requires SR protein participation [13, 14]. As for RNA-protein interactions, RS domain contacts the branchpoint in complex A to promote spliceosome assembly. In alternative splicing, RS domain also mediates binding of SR protein RRM domain to exon splicing enhancer (ESE) gene sequence (Fig. 1.3B), facilitating recruitment of U2AF to weak alternative 3' splice sites and antagonizing the splice site silencing effect of hnRNPs (heterogenous nuclear ribonucleoproteins) [12, 15-17]. At the 5' splice site, SR proteins such as ASF/SF2 (alternative splicing factor / human splicing factor 2) promote binding of spliceosome to all possible alternative splice sites, also antagonizing the hnRNPs that inhibit binding (Fig. 1.3C). In rare cases, SR proteins may have similar functions as hnRNPs and bind silencer gene sequence to suppress exon inclusion and splicing, but the evidence thus far is inconclusive [18]. After the completion of splicing, SR proteins promote mRNA quality control through nonsense-mediated decay (NMD), a process that removes at least one-third of all mRNA transcripts [19, 20]. Finally, SR proteins are known to interact with nuclear export protein TAP for mRNA export [21, 22] into cytoplasm and to participate in protein translation [23].

The key to maintaining SR protein functions as well as its subcellular localization is the phosphorylation extent of the RS domain. Addition of protein phosphatase 1 (PP1) to nuclear extracts inhibit the formation of the E complex and binding of U2 and U4/U6/U5 to pre-mRNA [24], suggesting that phosphorylated RS domain is essential for spliceosome assembly at the early stage. The interaction between ASF/SF2 and U1-70K during exon definition is also greatly enhanced by phosphorylation of the
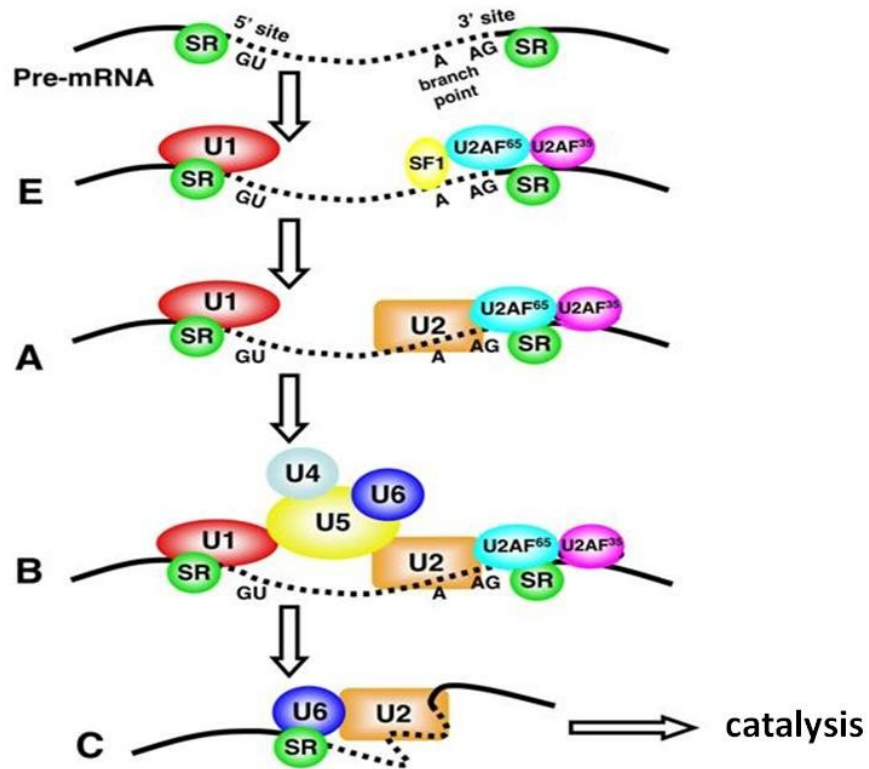
Figure 1.2 Spliceosome Assembly. Sequential recruitment and exchange of U1 through U6 snRNPs to splice sites occur with assistance from additional splicing factors. Solid lines are exons and dotted lines are introns. E, A, B, and C refer to specific stages in spliceosome assembly.

A

SRp75

SRp55

SRp40c

SRp30a
(ASF/SF2)

SRp30b

SRp20

9G8

| 316 | 160 | 88 | 50 | 105 | 79 | 118 |

RS Domain

RRM Domain

B

3' splice site:

Splicing Enhancer Model:

U2AF 65 | U2AF 35 | ASF/SF2 | ESE

weak seq.

Splicing Silencer Model:

U2AF 65 | U2AF 35 | ASF/SF2 | hnRNPA1 | ESE | ESS

C

SR Protein

ASF/SF2 | 5' Proximal Site

U1 | U1

hnRNPA1 | 5' Distal Site
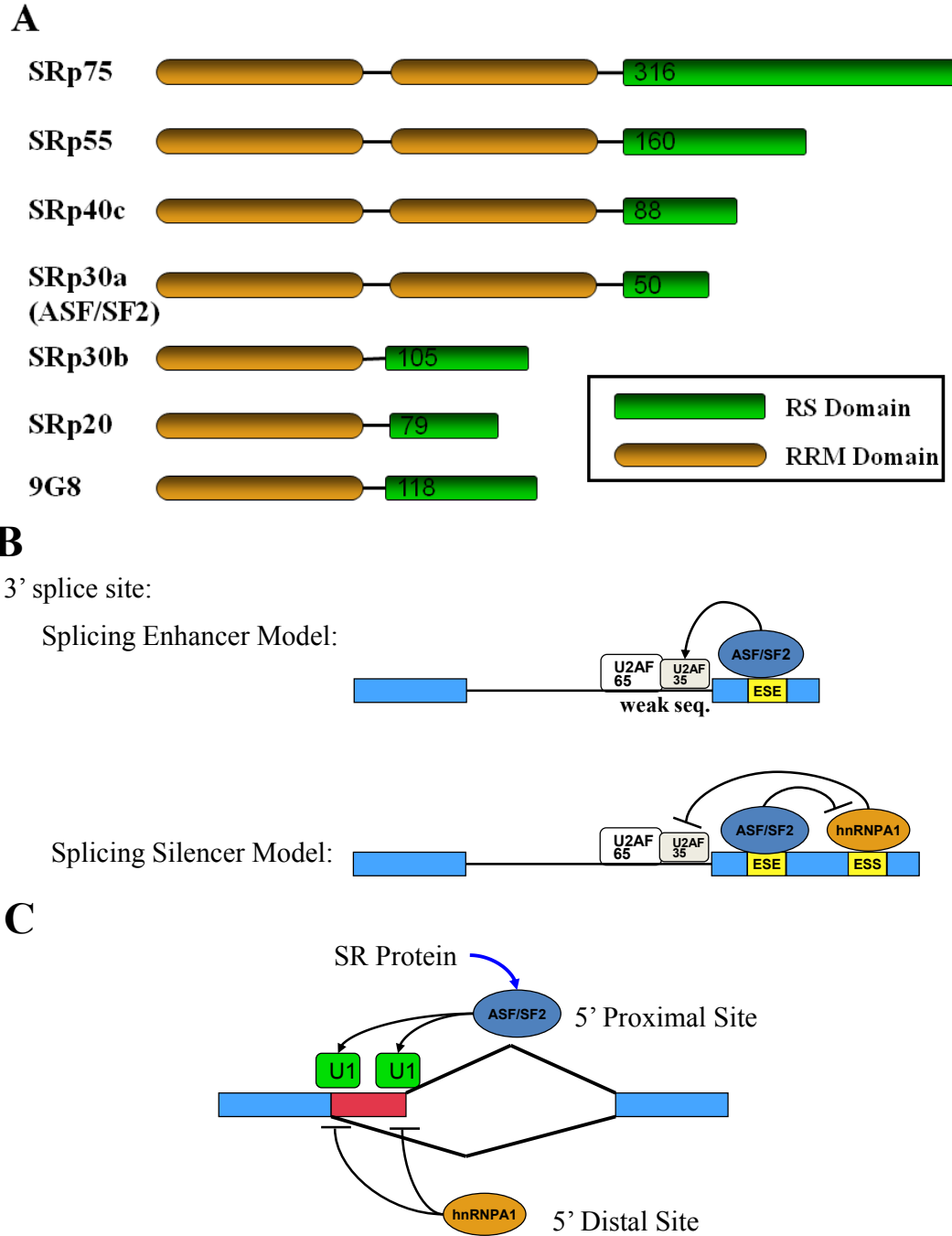
Figure 1.3 SR Proteins and Their Roles in Alternative Splicing. Adapted from [25]. SR proteins such as ASF/SF2 may counteract inhibitory effects of hnRNPs. A) Domain structure of classical SR proteins includes one or more RRM domains and one RS domain. B) ASF/SF2 enhances splicing signal of weak 3' splice sites in alternative splicing. C) ASF/SF2 enhances binding of U1 at alternative 5' splice sites.

RS domain [26].   In comparison, phosphatase inhibitors block splicing activity without affecting spliceosome assembly [27], indicating that dephosphorylation is important for splicing catalytic activity at a later stage.   As for intronless mRNA export, TAP can only interact with a subset of hypo-phosphorylated SR proteins but not the hyper-phosphorylated form [28, 29], emphasizing a need for tight regulation by splicing kinases and possibly phosphatases.

B.   Regulation of SR Protein Localization and Function

SRPK1 (SR Protein Kinase 1) and Clk/Sty (Cdk-like kinase/serine-threonine-tyrosine) represent two distinct families of splicing kinases that phosphorylate SR proteins and regulate splicing function via subcellular localization of SR protein splicing factors.   SRPK1 is mainly a cytoplasmic kinase, phosphorylating RS domain serines [30-32].   Only phosphorylated SR proteins would interact with transportin-SR for nuclear import [33-35], a necessary step for participation in intranuclear splicing reactions (Fig. 1.4A).   Once in the nucleus, SR proteins aggregate in complex structures called nuclear speckles, dynamic centers of mRNA processing and transcription [36].   Unlike SRPK1, Clk/Sty is a dual-specific kinase localized in the nucleus [37].   Clk/Sty autophosphorylates on tyrosine residues, but phosphorylates only serine and threonines on substrates [38].   Overexpression of Clk/Sty induces redistribution of SR proteins from nuclear speckles [37, 39], suggesting a possible role for Clk/Sty in recruitment of SR proteins for spliceosome assembly.

Only a few studies identified phosphatases dephosphorylating SR proteins within the nucleus for splicing catalysis and mRNA export.   Microinjection of PP1 (protein phosphatase 1) into the nucleus was shown to disrupt localization of SR proteins to the speckles, confirming that SR proteins need to be at least hypo-phosphorylated for them to remain in speckles [40].   In yeast, an eukaryote without alternative splicing, Npl3 is an SR-like shuttling protein that functions in mRNA export and possibly interacts with U1 and U2 snRNP in splicing [41].   Nuclear phosphatase Glc7 dephosphorylates Npl3 in vivo, an essential step for recruitment of mRNA export proteins and the transfer of mRNA from 3' end processing machinery to export protein complex [42].   Interestingly, Glc7 has the same catalytic domain as PP1 in humans and is considered the yeast homolog [43], lending support to the importance of human PP1.   In

addition, calf intestinal phosphatase (CIP) and protein phosphatase 1 (PP1) have been shown to dephosphorylate SR proteins in cell models [44, 45], and phosphatase inhibitors specific for PP1 and PP2A block splicing catalysis as mentioned before [27]. PP1 and PP2A also dephosphorylates U2 and U5 snRNPs as a necessary step to splicing catalysis [46], but further studies are urgently needed to reach a more complete understanding of the interactions between phosphatases and SR protein splicing factors.

The key to SR protein regulation by SRPK1, CLK/Sty, and PP1 is the phosphoryl content on the SR protein (Fig. 1.4A), and previous studies have demonstrated that both hyper- and hypo-phosphorylation inactivates splicing [47]. For each step of the mRNA processing, there is a corresponding phosphoryl content essential for splicing factor functionality, and detailed studies have begun on this subject. Previous studies on ASF/SF2, a classical SR protein, showed that the two splicing kinases recognize different parts of the RS domain [48]. SRPK1 exclusively phosphorylates the N-terminal portion of RS domain (designated as RS1) at about 12 sites (Fig. 1.4B). In contrast, Clk/Sty phosphorylates up to all 22 sites in the RS domain, with or without initial SRPK1 phosphorylation. Another study demonstrated that lack of RS2 domain phosphorylation prevents dispersion of SR proteins from nuclear speckles, confirming that a cooperative phosphorylation cascade controls SR protein subcellular localization and participation in spliceosome assembly. Finally, a study involving CIP showed that dephosphorylation of splicing factors *ex vivo* may not be complete [45], consistent with the understanding that only hypo-phosphorylated SR proteins can interact with TAP export [28, 29].

C.   Splicing Kinases Are Structurally and Catalytically Distinct

Both SRPK1 and Clk/Sty belong to the CMGC kinase group [49] and phosphorylate multiple SR proteins [50], but their structures differ significantly. SRPK1, its mammalian homolog SRPK2, and its yeast counterpart Sky1 have several unique features (Fig. 1.5A). All three have an autonomously folding domain insert bifurcating the kinase core, as shown in the representative SRPK1 domain organization figure. Insert domains are also observed in receptor tyrosine kinases, but none are so large as in SRPK1.
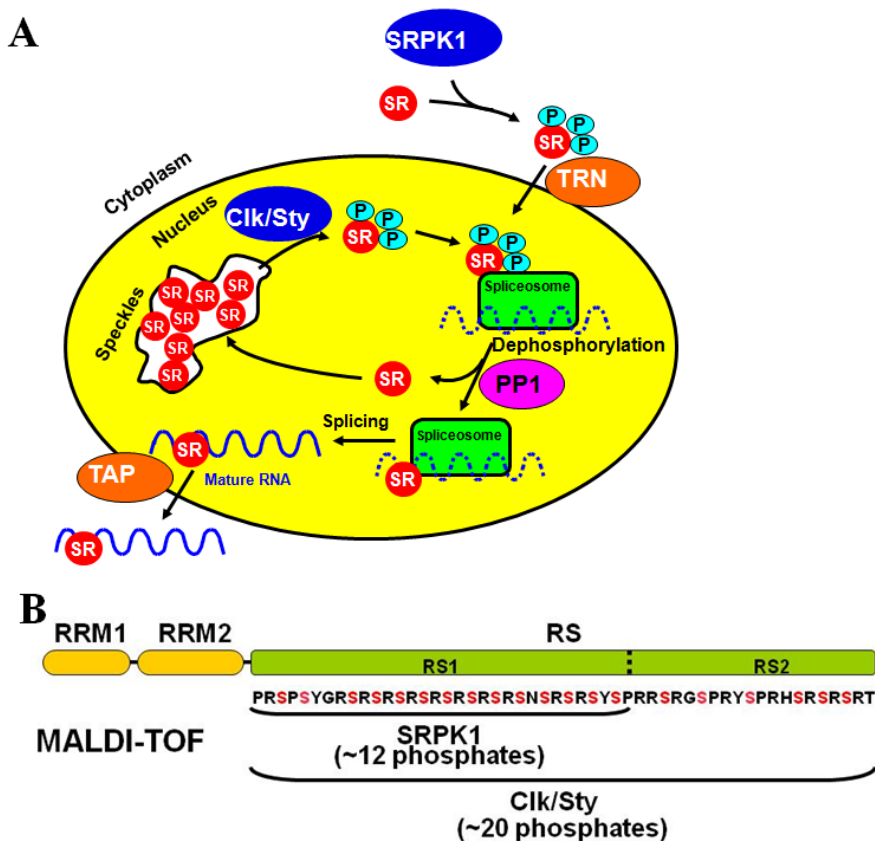
Figure 1.4 Role of Splicing Kinases and Phosphatase in mRNA Processing. A) Subcellular localization of SR proteins. Hypo-phosphorylation of SR proteins by SRPK1 drives nuclear import with transportin-SR.   Hyper-phosphorylation by Clk/Sty drives dispersion from nuclear speckles to participate in spliceosome assembly. Dephosphorylation by PP1 enables splicing catalysis and export of mature RNA. B) Detailed domain organization of ASF/SF2. ASF/SF2 is a classical SR protein, and its RS domain can be divided into RS1 and RS2 segments based on substrate preferences of SRPK1 and Clk/Sty.

Despite the proximity of the insert to the hinge region and possible role in catalysis, removal of the insert in SRPK1 has only a small effect on association or kinetic parameters but leads to complete nuclear translocation of the enzyme [30], suggesting that the insert serves as a cytoplasmic anchor, enforcing compartmentalization of SRPK1 and Clk/Sty. The X-ray structure of SRPK1 without the insert domain shows that the kinase adopts a typical kinase fold [39], with a glycine-rich loop, activation loop, catalytic loop, as well as a MAP-kinase insert typical of the CMGC kinase group (Fig. 1.5B). SRPK1 also contains N- and C-terminal extensions, but the functions are currently unknown. Previous studies demonstrated that ASF/SF2 binds SRPK1 with high affinity ($K_D \sim 50$ nM) [51], and the SRPK1 structure reveals a docking site near the MAP kinase insert in the large C-terminal lobe [39, 52]. Initial binding event involves N-terminal part of the RS domain, while the $\beta 4$ strand of RRM2 is likely to bind at a later stage of the multi-site phosphorylation.

In terms of substrate recognition, SRPK1 can phosphorylate serine and not threonine residues unlike most serine/threonine kinases, and more importantly, strongly prefers serines next to arginines, as opposed to lysines or glycines [32, 50, 53]. Finally, SRPK1 needs no priming phosphate on the activation loop for activation, which remains active without much intra-protein interactions [54].

While little is known regarding upstream signal transduction pathways, it was recently discovered that stress signals such as osmotic shock induces dissociation of SRPK1 from a chaperone complex and subsequent nuclear import, leading to differential phosphorylation of SR proteins and alternative splice site selection, confirming the importance of the spacer insert domain for localization [55].

Clk/Sty is a member of the LAMMER kinase family, with the conserved "EHLAMMERILG" motif that may participate in substrate recognition [37]. There are four human homologs, named CLK1 through CLK4 [56]. Clk/Sty has no spacer insert domain bifurcating the kinase lobes [57], but it contains an unusual N-terminal RS domain that interacts with RNA-binding proteins [37]. Deletion of this RS domain also increases enzymatic activity, suggesting an auto-regulatory role [58]. The crystal structure shows a typical kinase fold for Clk/Sty (Fig. 1.5C), with activation loop, catalytic loop, and MAP-kinase insert typical of CMGC kinase family [57]. One major difference between SRPK1 and Clk/Sty lies in the

putative docking groove (Fig. 1.5D). The structural equivalent in Clk/Sty is covered by the unique Clk/Sty helix αH, possibly explaining the different substrate specificities. The helix αH also covers the LAMMER motif and makes it inaccessible to surrounding solvents, and it is unlikely that the LAMMER motif can participate in substrate recognition. Finally, the β-hairpin insert in the C-terminal lobe may interact with substrate and replace the function of the putative docking groove.

Clk/Sty auto-phosphorylates at several serine/threonine and tyrosine residues, and tyrosine auto-phosphorylation modulates phosphorylation of ASF/SF2 by Clk/Sty, while serine/threonine auto-phosphorylation modulates phosphorylation of SC35 by Clk/Sty [58, 59], a much more complex system of regulation compared against SRPK1. A recent study suggested that Akt2 may be the upstream kinase for Clk/Sty, as a part of the insulin signal transduction pathway leading to alternative splicing of PKCβ2 [60], implying that Clk/Sty may need to be phosphorylated for activation. Finally, Clk/Sty is a dual-specific kinase that phosphorylates both tyrosines and serine/threonines, and it has broader substrate specificities than SRPK1, allowing it to phosphorylate beyond the RS domain and tolerate serines next to lysines and perhaps residues other than arginines [50]. However, a recent peptide library study suggested that CLK1, the founding member of the Clk/Sty family, shows two-fold phosphorylation preference for serine over threonines, and no detectable phosphorylation of tyrosines [57]. CLK1 also selects for proline at P+1 position and arginine at P-3 position. Further studies with various ASF/SF2 constructs are necessary to resolve the contradictions.

Clk/Sty is involved in alternative splicing of its own gene, producing the full-length isoform and a truncated, inactive isoform lacking the kinase domain [61]. It also regulates erythroid cell differentiation through alternative splicing [62], but additional studies are necessary to explore its downstream targets.
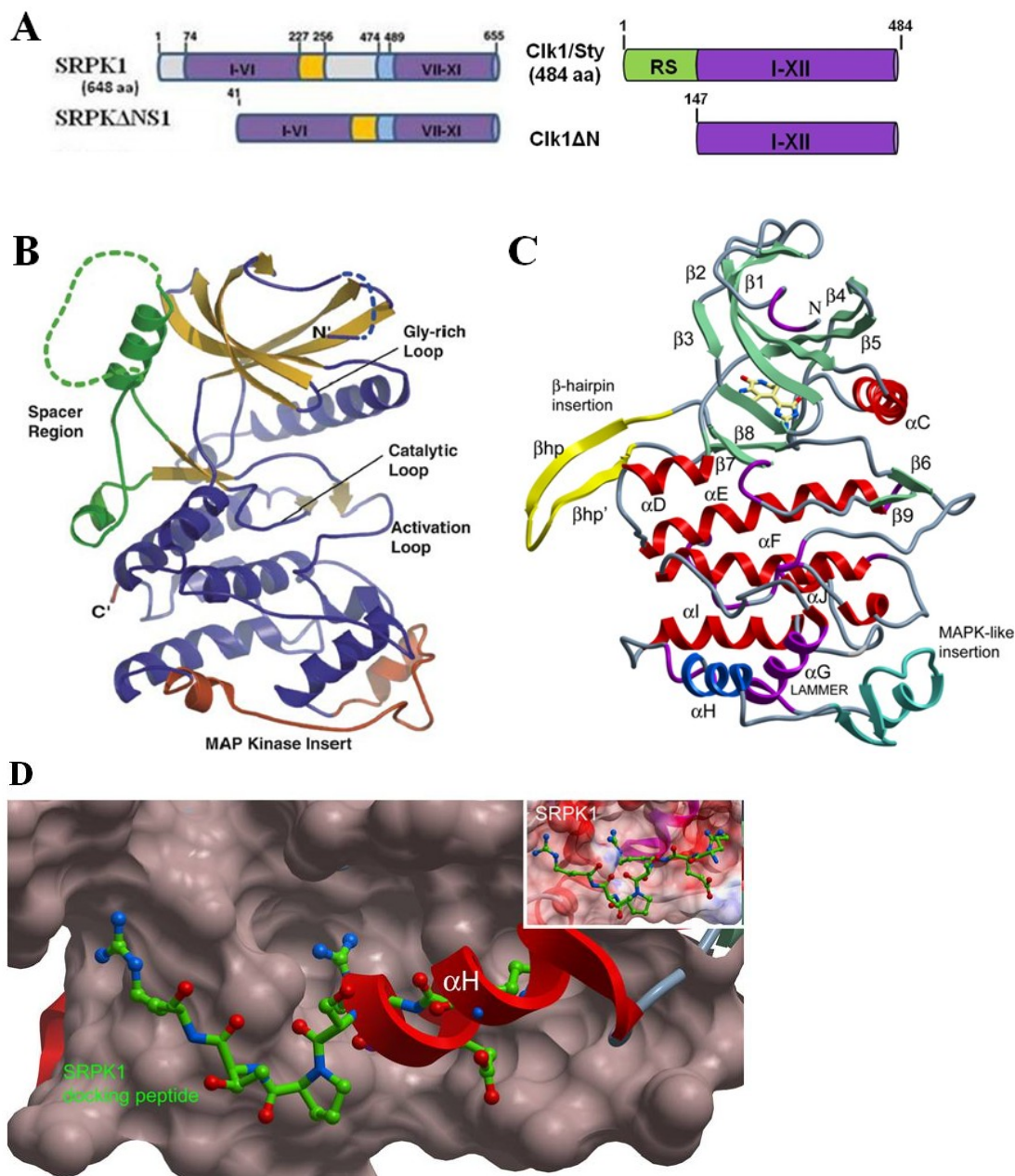
Figure 1.5 X-ray Crystal Structures of SRPK1 and Clk/Sty. Taken from [39, 57]. A) Domain organization of SRPK1 and Clk/Sty splicing kinases. B) Structure of SRPKΔNS1. To improve stability of the protein and success rate of crystal formation, the spacer insert and the N-terminus of the SRPK1 kinase have bee truncated for this structure. C) Structure of CLK1. CLK1 is a member of the Clk/Sty kinase family. D) Comparison of the substrate binding groove of SRPK1 and CLK1. Peptide derived from yeast putative splicing factor Npl3 represents SR protein substrate shared by both kinases. αH helix of CLK1 impairs binding of Npl3 and other possible substrates.

D. Phosphatase Structure and Substrate Recognition

PP1 is a member of PPP (phosphoprotein phosphatases) family, is constitutively active and dephosphorylates phosphoserines and phosphothreonines in different signal transduction pathways. Other members include PP2A, PP2B, PP2C, PP4, and PP5. The catalytic subunit of PP1 is fairly non-discriminating in terms of substrates, but there are hundreds of regulators and inhibitors that bind PP1to form huge complexes and to regulate substrate binding and enzymatic activity [63]. Random peptide library studies suggest that the PP1 binding cleft is too shallow for high-affinity binding of peptides, and regulators provide extended surface interactions for a larger, natural substrate [64]. The X-ray crystal structure of PP1γ bound to a natural toxin inhibitor okadaic acid shows three grooves on the surface of the enzyme (Fig. 1.6A), forming a large Y-shaped cleft [65]. The double ring of okadaic acid binds into the hydrophobic groove and may interact with a protruding loop separating the acidic and the C-terminal grooves. The overall fold of PP1 consists of two tightly linked domains, an N-terminal α/β domain and a C-terminal β domain.   The first eight residues in the N-terminus may modulate substrate recognition and affinity [66]. Another X-ray structure of PP1δ bound to MYPT1, a targeting regulator involved in smooth muscle relaxation, revealed contacts between the conserved RVxF motif and the N-terminus of PP1 at the hydrophobic groove (Fig. 1.6B) [67]. Acidic and basic sequences flanking the RVxF motif may interact with PP1, but in this structure only the amino acids N-terminal to the motif in MYPT1 form contacts with PP1. Myosine phosphatase is a trimeric holoenzyme, consisting of the catalytic subunit of PP1δ, the targeting subunit MYPT1, and a 20kDa subnint M20 of unknown function [68]. This example underscores the prevalence of regulators for PP1 substrate selection.

In mammalian tissues there are three PP1 genes and four individual PP1 isoforms, in contrast to the single PP1-analog gene in yeast. In mice PP1γ knockout induces impaired spermiogenesis, suggesting that these isoforms are not entirely interchangeable [69]. Many regulators show preferential binding to individual PP1 isoforms, leading to substrate and functional specificity for each isoform. Previous studies related to SR proteins have shown interaction between PP1γ and SRp38 [70], as well as between PP1γ and Tra2β1 or ASF/SF2 [71]. The contact may occur at either a conserved RVxF motif at
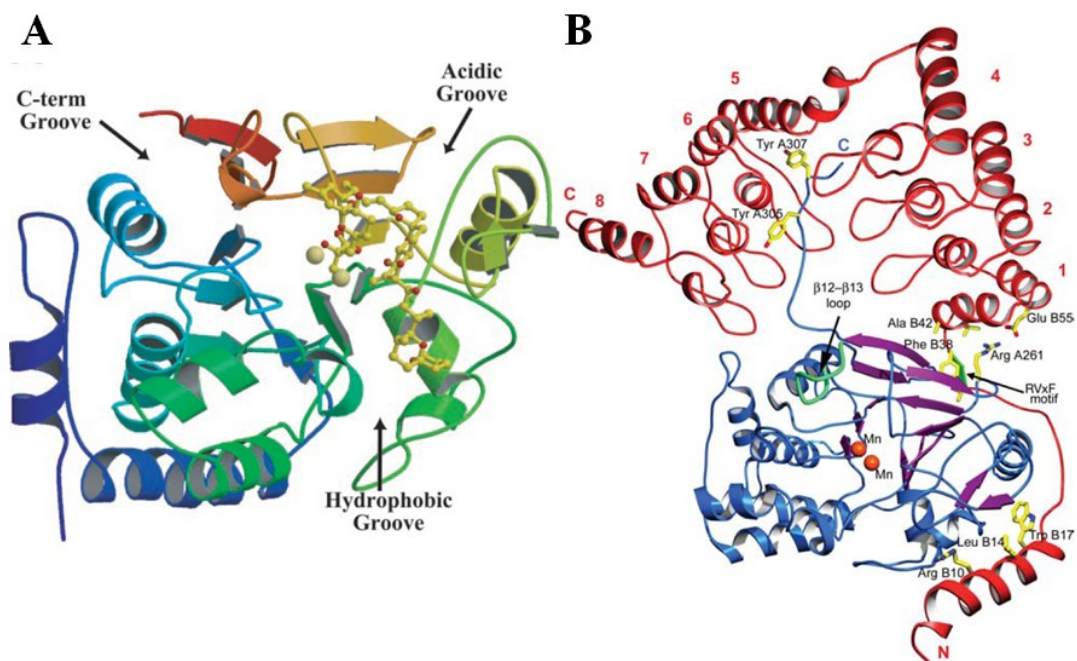
Figure 1.6 X-ray Crystal Structures of PP1γ. Taken from [65, 67]. A) PP1 in complex with inhibitor okadaic acid (ball-and-sticks model with carbon atoms in yellow and oxygen atoms in red). Blue color for the N-terminus changes to red color for the C-terminus. Two manganese atoms in the active site are shown as yellow spheres. B) PP1 in complex with regulator MYPT1 (in red). α-helices and loops (blue), β-strands (magenta) and two manganese cations (yellow) in the catalytic site are shown.

the C-terminus of the RRM domain [71] or at the phosphorylated RS domain [70]. In these two cases, SR protein domains may serve as targeting subunits. ASF/SF2, with a RVEF motif, was shown to immunoprecipitate PP1 in cell models [71], but further studies are needed to characterize the dephosphorylation mechanism. Finally, leucine-rich repeats may also serve as a substrate recognition motif for PP1, but they are not commonly found in SR proteins [72].

Phosphatase cascades regulate many key processes. The yeast homolog Glc7 is associated with over 200 genes required for survival of yeast, controlling such diverse processes as glycogen metabolism, vesicle trafficking, cell polarity, DNA damage, transcription, and cell-cycle progression [73]. In mammalian systems, PP1 is controlled by association with protein inhibitors that block all activity and other regulators that direct PP1 to specific substrates. PP1 in turn modulates many downstream pathways, including SR protein dephosphorylation and alternative splicing. Heat shock dissociates PP1 from its inhibitor protein 14-3-3 and activates dephosphorylation of SRp38, switching it from splicing activator to splicing inhibitor [46]. Tra2β1 dephosphorylation by PP1 controls alternative splicing of the motor neurons, whereas chemotherapy-induced ceramide signaling activates dephosphorylation of ASF/SF2 to regulate caspase-9 alternative splicing [44].

E. Structures Resolved and Proposed for ASF/SF2

ASF/SF2, also known as SRp30, is the prototypical member of the SR protein splicing factor family, and a well-established substrate of SRPK1 and Clk/Sty kinases years ago. Topoisomerase I and Akt are also known to phosphorylate ASF/SF2, but the biological effect is unclear [74]. Gene knockout studies in chicken DT40 cells and mouse embryo fibroblasts (MEFs) have shown that ASF/SF2 is essential for cell viability. It also regulates alternative splicing of CaMKIIδ, which is critical for development of cardiomyocytes. Despite its biological importance, its structure has never been fully resolved. ASF/SF2 has two RNA-binding RRM domains in the N-terminus and one RS domain in the C-terminus. Structures of RRM domains have been resolved by X-ray crystallography and nuclear magnetic resonance (NMR), but the RS domain structure remains elusive.
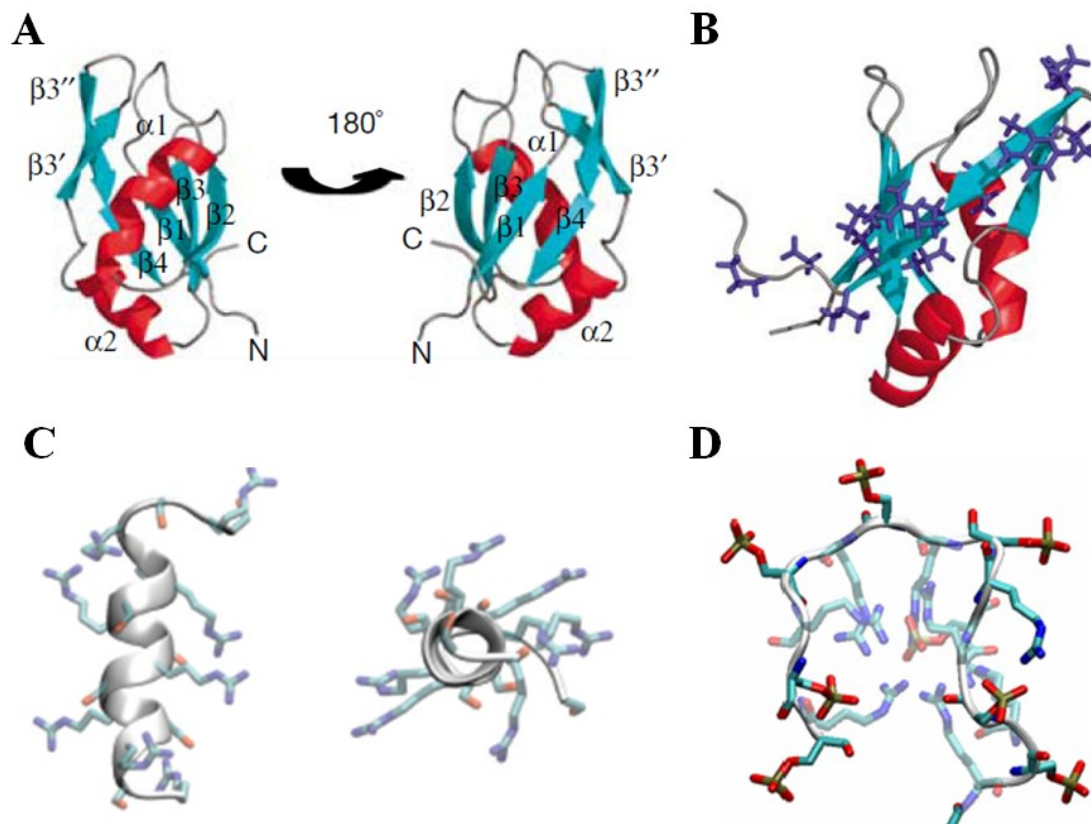
Figure 1.7 NMR Solution Structure of RRM and Proposed RS Structure. Taken from [75, 76]. A) ASF/SF2 RRM2 (aa 107-215) NMR solution structure in two orientations. Helices are in red, and β-sheets in blue. B) SRPK1 docking motif (blue side chains) highlighted on the RRM2 structure. C) Proposed helical structure of unphosphorylated RS repeats. D) Proposed "arginine claw" structure of phosphorylated RS repeats. Six arginines form interactions with centered phosphoserine.

A published NMR solution structure of RRM2 domain of ASF/SF2 showed typical RRM fold (Fig. 1.7 A) despite lacking crucial residues within the ribonucleoprotein RNP-1 and RNP-2 motifs commonly associated with RNA binding [75]. The fold contains a four-stranded β-sheet packed against 2 helices. Residues 196-215 of ASF/SF2 is the start of the RS domain and shows increased mobility as predicted by structure calculations. Residues 184 through 197 form the docking sequence for SRPK1 and extend into the RRM fold (Fig. 1.7B), deep into the hydrophobic region. The SWQLKD motif, and residue W134 in particular, is important for RNA binding and it is conserved among SR proteins containing a second RRM domain. Finally, residues 89 through 120 in the linker between RRM1 and RRM2 in wt ASF may mediate binding to TAP nuclear export protein.

Although no structure of an RS domain has been solved, numerous models have been proposed. In one sequence analysis model, SR proteins have been characterized as intrinsically disordered based on predicted randomness of their RS domains [77]. Indeed, based on a circular dichroism study of ASF/SF2, the RS domain lacks helical structure [78]. In a separate all-atom energy landscape modeling study, charge neutralization associated with phosphorylation would favor collapse of the RS domain into a more compact structure, and an increased helical content [79]. CD spectral analysis of phosphor-mimic RS domain of U1-70K snRNP clearly displays helical structure and supports this second model [80]. Finally, helix formation with lysine side-chains outward is the most favorable conformation based on a molecular dynamics model (Fig. 1.7C & D) [76]. The conformation of phosphorylated RS repeats becomes more extended, with a few of the phosphoserines stabilized by arginines in a "claw" formation and most of the remaining phosphoserines extending outward for transportin recognition. It is clear that additional studies are necessary to resolve the question of RS domain structure.

F. Processive Phosphorylation of ASF/SF2 by Splicing Kinases

Phosphorylation is an important post-translational modification of proteins for many signal transduction pathways, and in many cases more than one phosphoryl transfer is necessary for downstream effects [81], catalyzed by one or more enzymes in conjunction. Both SRPK1 and Clk/Sty catalyzes multiple phosphorylation events on ASF/SF2, and several recent studies have started to characterize these
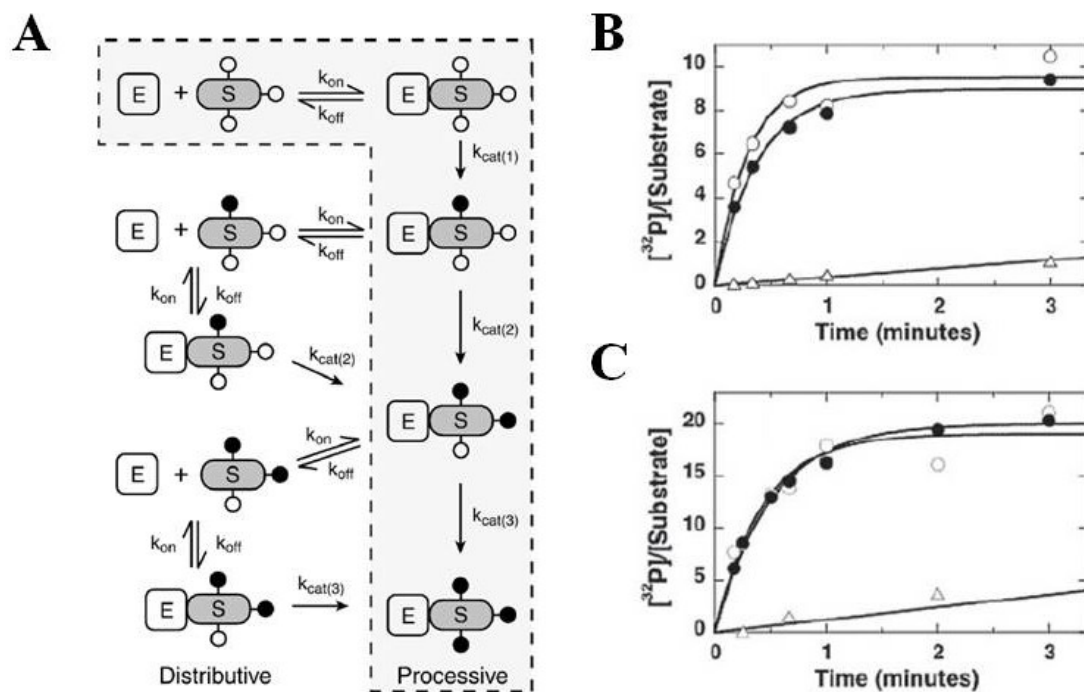
Figure 1.8 Mechanism of Phosphorylation. Taken from [48, 81]. A) Processive and distributive mechanism. No intermediate product will be release in processive mechanism. B-C) SRPK1 and Clk/Sty processively phosphorylate ASF/ASF2. Start-trap and control curves (circles) overlap, indicative of no intermediate product release. Trap-start curve (Δ) confirms efficiency of inhibition by kinase-dead SRPK1 or Clk/Sty, respectively.

phosphorylation mechanisms. In a fully processive mechanism, the substrate is not released from the enzyme until all possible sites have been modified, whereas in a distributive mechanism partially modified intermediate products may be released at any step of the reaction (Fig. 1.8A). Processive phosphorylation mechanisms are rarely found in nature, and no such mechanism has been demonstrated for SR proteins other than ASF/SF2 [81].

Processive phosphorylation was initially demonstrated with a heterodimeric SRPKΔS:ASF/SF2 complex [51]. A spacer-deleted SRPK1 enzyme was co-purified with ASF/SF2 by FPLC to create the high affinity enzyme:substrate complex. This study showed that the SRPKΔS:ASF/SF2 complex finishes the phosphorylation reaction despite the presence of an inhibitor peptide added at the start of the "start-trap" reaction, suggesting no dissociation of intermediate products in a processive mechanism. In a "trap-start" control reaction the inhibitor was incubated with the enzyme:substrate complex to demonstrate that the inhibitor is functional and abolishes >95% of kinase activity. A later study employing kinase-dead SRPK1 and Clk/Sty to trap out phosphorylated intermediate products demonstrated that both SRPK1 and Clk/Sty processively phosphorylate ASF/SF2 (Fig. 1.8B) [48]. The forward catalytic rate for each phosphate is much higher than the rate of dissociation due to the extraordinarily high affinity of the complex, allowing for the uncommon processive mechanism.

G. Diseases Associated with Defects in Splicing Machinery

Alternative splicing can generate an extremely complex human proteome necessary for normal functions, but abnormal pre-mRNA splicing has been associated with many human degenerative diseases, including frontal dementia and parkinsonism, amyotrophic lateral sclerosis, myotonic dystrophy, and spinal muscular atrophy [82-85]. Alternative splicing defects have also been linked to several types of cancers [86-90] and important stages in neoplastic onset: angiogenesis, invasion, apoptosis, and cellular differentiation [3, 44, 91-93]. New evidence suggests that errors in trans-acting splicing machinery can contribute to these disorders as much as cis-acting genetic defects specific to each gene [94]. Control of cell motility by ASF/SF2, via regulation of alternative splicing of Ron oncogene, is a fine example of SR protein participation [95]. While RNAi and related RNA-processing-targeting therapies have been

pursued in various studies to treat cis-acting defects and erroneous mRNAs [90, 91], inhibitors of splicing kinases may offer novel treatment options against trans-acting splicing defects. For example, SRPK1 downregulation in testicular germ cell tumor and in retinoblastoma is linked to cisplatin resistance [96, 97]. In addition, overexpression of SRPK1 is observed in colon, breast, and pancreatic cancer [98], suggesting that SRPK1 is a valid anti-cancer drug target. Even though inhibitors have been developed against viral replication utilizing host splicing machinery, ATP-mimics that bind at nucleotide-binding pocket of SRPK1 or Clk/Sty invariably cross-inhibit other kinases as well and lead to undesirable side effects [99-101]. Targeting the substrate docking may be the key to minimizing cross-reactivity, and an in-depth study of the phosphorylation mechanism and associated structural changes may lead the way toward a better understanding of the whole system. In contrast, PP1 may not be a good drug target due to its non-discriminating dephosphorylation mechanism. Targeting regulatory subunits may lead to drugs that target only the disease at hand, but there are hundreds to choose from and possibly hundreds more undiscovered. On the one hand, PP1 binds to BRCA1 and its expression is lowered, just like BRCA1, in breast cancer tissues [102]. On the other hand, okadaic acid and other PP1 and PP2A inhibitors are known carcinogens in animal models [103], suggesting that sustained phosphorylation is in general more tumorigenic than not. An in-depth study of the dephosphorylation mechanism and associated structural changes will start addressing the dearth of knowledge regarding biological function of phosphatases.

H. Goals of the Dissertation

SR proteins are essential for constitutive and alternative splicing, and their subcellular localization and function are modulated by phosphorylation of the RS domain by splicing kinases SRPK1 and Clk/Sty and dephosphorylation by nuclear phosphatases such as PP1. Previous studies on the mechanism of phosphorylation showed that SRPK1 and Clk/Sty processively phosphorylate ASF/SF2 for 12 and 20 sites respectively. To analyze the nature of these reactions, we wished to draw some possible parallels to DNA polymerase which catalyze highly processive reactions with distinct initiation and elongation phases. In the initiation stage, DNA polymerase requires a priming base inserted by DNA primase which recognizes a specific initiation sequence. In the elongation stage, DNA polymerase is

processive, directional, and sequential in adding additional bases. Our goal was to study the initiation and the elongation stage of ASF/SF2 phosphorylation by SRPK1, and to a much lesser extent, Clk/Sty. The greater goal was to study the whole phosphorylation cycle necessary for splicing, including the mechanism of dephosphorylation by PP1. In summary, my primary goals of research were as follows: 1) to investigate mechanism of initiation and elongation for SRPK1 phosphorylation of ASF/SF2. 2) to analyze structural components of SRPK1 and ASF/SF2 that regulate the phosphorylation mechanism. 3) to characterize the dephosphorylation mechanism of ASF/SF2 by PP1.

Chapter 2

Materials and Methods

A. Materials

Lysobacter enzymogenes endoproteinase Lys-C, ATP, ADP, AMP, 3-(N-morpholino)propanesulphonic acid (MOPS), 2-[N-morpholino]ethanesulphonic acid (MES), Tris, MgCl$_2$, NaCl, CaCl$_2$, ethylenediaminetetraaceticacid (EDTA), glycerol, acetic acid, Phenix imaging film, trifluoroacetic acid (TCA), bovine serum albumin (BSA), dry milk powder, acetonitrile, methanol, Tween 20, NP40 Igepal detergent, Hydrologix low-retention pipette tips, and liquid scintillant were obtained from Fisher Scientific. [γ-32P]ATP was obtained from NEN Products, a division of Perkin-Elmer Life Sciences. α-Cyano-4-hydroxy cinnamic acid matrix solution (α-CHC) was obtained from Aldrich Chemicals and recrystallized once from ethanol. Alternatively it could be obtained from Agilent Technologies in resuspended solution form (Santa Clara, CA). Inhibitor peptide Alapep and Npl3-derived substrate peptide (YRTRDAPRERSPTR) were obtained from the USC Norris Comprehensive Cancer Center Microchemical Core Facility. Nickel sepharose resin was purchased from Qiagen. Glutathione immobilized on 4% agarose beads, free glutathione, and AMP-PNP were purchased from Sigma-Aldrich or GE Healthcare. Isopropyl β-D-thiogalactoside was purchased from BioPioneer (San Diego, CA). Protease inhibitor cocktail tablets and NBT/BCIP reagent tablets were purchased from Roche Diagnostic. Immobilon-P PVDF membranes were purchased from Millipore. PP1α and phosphatase buffer (see section E for detailed components) were purchased from New England Biolabs. Protein and DNA Lo-Bind tubes were purchased from Eppendorf.

B. Preparation of Recombinant DNA

All single and multiple site mutations of eight amino acids or less were generated by a single or a series of PCRs using the QuikChange mutagenesis kit (Stratagene, La Jolla, CA). Manufacturer's protocol was followed in full, except that DH5α cells (gift from Jin Yang laboratory) were used for the final transformation and plasmid amplification. C-terminal deletions of protein domains in part or full were generated by inserting stop codon by mutagenesis for N-terminal Histidine tagged constructs and by subcloning into new vector for C-terminal Histidine tagged constructs. N-terminal deletions of protein domains in part or full were generated by subcloning into new vectors containing the appropriate

Histidine tag. Every mutant construct was confirmed by sequencing (Eton Bioscience, San Diego, CA). The template vector used was that of the wt unless otherwise stated.

SRPK1(wt), in pet15b vector, was a gift from Xiang-Dong Fu laboratory. It might be referred to as SRPK1 or SRPK for short.

A kinase-inactive form of SRPK1 called kdSRPK1 was generated by replacing Lys at position 109 with Met, a gift from Dr. Jonathan C. Hagopian.

SRPK1(ΔS), in pet15b vector, was a gift from Xiang-Dong Fu laboratory. Residues 255 through 474 of the wt-SRPK1construct were deleted.

SRPK1(4M) was generated by performing 4-alanine mutations: D564A, E571A, D548A, and K615A, a gift from Dr. Jonathan C. Hagopian. SRPK1(6M) was generated with SRPK1(4M) as template, adding two additional alanine mutations E558A and E557A, a gift from Dr. Jonathan C. Hagopian.

SRPK1(RA) was generated by performing single-alanine mutation at R561A. SRPK1(2RA) was generated by performing double-alanine mutation at R515A and R518A. SRPK1(3RA) was generated with SRPK1(2RA) as template, with all three alanine mutations.

Wt-ASF, in pet19b N-terminal Histidine tagged vector, was a gift from Xiang-Dong Fu laboratory. It might be referred to as ASF/SF2 or ASF for short.

ASF(4SA205), ASF(4SA211), ASF(4SA215), ASF(4SA219), and ASF(4SA221) were generated by performing 4-alanine mutations on consecutive serines mostly in the RS1 domain starting at the designated residue. The original design of the mutagenesis came from the Gourisankar Ghosh lab. ASF(4RA212), ASF(2RA216), and ASF(2RA222) were generated by performing up to 4 alanine mutations on consecutive arginines in the RS1 domain starting at the designated residue. ASF(block1) through ASF(block4) were generated by Jon Hagopian with the arginine mutant constructs above as template, employing multiple rounds of mutagenesis (see Fig. 5.6E for details).

ASF(W134A), ASF(Q135A), and ASF(R154A) were generated with single-alanine mutations using Wt-ASF as template. ASF(4M) was generated with 4-alanine substitutions: W134A, Q135A, R154A, and E183A. The original design for the mutagenesis came from the Gourisankar Ghosh lab.

ASF(RS$_8$-RS$_8$), ASF(SA$_8$-RS$_8$), ASF(RS$_8$-RA$_4$), ASF(RA$_4$-RS$_8$), ASF(RS$_8$-RS$_5$), and ASF(RS$_8$-RS$_4$) were generated with multiple rounds of mutations to create artificially long or interrupted stretches of serine-arginine dipeptide repeats in RS1 or RS2, using Wt-ASF as template. The RS/SR pairs were created to be consecutive, unless they were disrupted by alternating arginine-to-alanine mutations in RA$_4$ constructs.

ASF(P200A) and ASF(P228A) were generated by single-alanine mutations at the designated proline, using wt-ASF as template. ASF(2PA197) and ASF(2PA200) were generated by 2-alanine mutations at P197A and P200A, as well as P200A and P228A, respectively. ASF(3PA197) and ASF(3PA200) were generated by 3-alanine mutations at P197A, P200A, and P228A, as well as P200A, P228A, and P235A, respectively. ASF(4PA) was generated by 4-alanine mutations at all 4 residues mentioned above.

ASF(ΔRRM1) and ASF(ΔRRM2) were gifts from the Xiang-Dong Fu laboratory.

ASF(i4A), ASF(i8A) and ASF(i16A) were generated with multiple rounds of mutagenesis to create a long, consecutive stretch of 8-alanine or 16-alanine insertion after residue, based on the template of Wt-ASF. The original design of the mutagenesis came from the Gourisankar Ghosh lab.

c-wt-ASF, in pet28a C-terminal Histidine tagged vector, was generated by subcloning in frame the PCR product of the wt gene into the NcoI and XhoI sites.

ASF(1-219), ASF(1-226), ASF(188-248), ASF(197-248), ASF(188-235), ASF(188-228), ASF(188-220), and ASF(220-248) were generated by subcloning in frame the PCR product of part of the wt gene into the NcoI and XhoI sites of pet28a. Alternative versions with N-terminal histidine tag in pet15b vector were also generated for some of these mutant constructs. All were gifts from Dr. Jonathan C. Hagopian.

ASF(R214K) was generated with single-lysine substitution, using c-wt-ASF as template. ASF(ΔRRM2-214) was generated with single-lysine substitution, using ASF(ΔRRM2) as template. The ASF(5R1K) mutant was generated with 5-arginine and single-lysine substitutions: K138R, K165R, K174R, K179R, K193R, and R214K, with ASF(R214K) as template. It is also known as cl-ASF(214), the first successful digestion mutant. cl-ASF(193), cl-ASF(210), cl-ASF(218), cl-ASF(222), cl-

ASF(224), and cl-ASF(229) were generated by shifting the single-lysine substitution in the RS domain to the designated residue through mutagenesis. cl-ASF(210) is also known as ASF(R210K), and cl-ASF(218) is also known as ASF(R218K). cl-ASF(214)dRS2 was generated with deletion of residues 227-248, with cl-ASF(214) as template. ASF(4M-5R1K) was generated with 4-alanine mutations: W134A, Q135A, R154A, and E183A, with cl-ASF(214) as template.

Digestion mutants contained cl- in their names, and they were used in LysC proteolysis experiments described below to probe for possible sequence in multi-site phosphorylation. cl-ΔRRM1(214) was generated with subcloning of part of cl-ASF(214) gene in frame into pet28a vector. cl-RRM2-4M(214) was generated with 4-alanine mutations similar to ASF(4M-5R1K), using cl-ΔRRM1(214) as template. cl-RS(K214) was generated with subcloning of residues 188-248 of cl-ASF(214) gene in frame into pet15b vector. cl-ASFΔRS2(214) was generated with subcloning of residues 1-224 of cl-ASF(214) in frame into pet28a vector.

cl-RS-RRM2(214) and cl-RS-RRM2(224) were generated by sequential subcloning steps inserting into the NdeI and XhoI sites and then XhoI and BamHI sites of pet15b vector. RS domain from cl-ASF(214) and cl-ASF(224) were inserted respectively, followed by RRM2 domain insertion. n-cl-ASF(214) was generated by subcloning inserting into the NdeI and XhoI sites of pet15b vector. cl-ASF(214) was used as the template.

cl-RS$_8$-RS$_8$(224), cl-SA$_8$-RS$_8$(224), cl-RS$_8$-RA$_4$(224), cl-RA$_4$-RS$_8$(224), cl-RS$_8$-RS$_5$(224), and cl-RS$_8$-RS$_4$(224) were generated with multiple rounds of mutations to create artificially long or interrupted stretches of serine-arginine dipeptide repeats in RS1 or RS2, using cl-ASF(224) as template. cl-RS$_8$-RS$_4$(193) was generated using cl-RS8RS4 as template, by shifting the lysine insertion.

wt-GST-ASF and GST-RS, in pGEX4T2 N-terminal glutathione-tagged vector, were a gift from Gourisankar Ghosh laboratory.

PP1γ, in pTacTac N-terminal Histidine tagged vector, was a gift from Jie Yang and the Anna A. DePaoli-Roach lab.

Clk/Sty, in pRSETa N-terminal Histidine tagged vector, was a gift from Xiang-dong Fu lab.

C. Expression and Purification of Recombinant Proteins

SRPK1 was transformed into the BL21(DE3) *E. coli* strain. The cultures were grown at 37°C in three to ten liters of LB broth supplemented with 100 μg/ml ampicillin, and induced with 0.1-0.4 mM isopropyl β-D-thiogalactoside at room temperature for 12–16 hours. Pelleted cells were lysed by sonication in 150 ml of lysis buffer [50 mM NaCl, 20mM MES, pH 6.5, 20% (vol/vol) glycerol, 1 mM PMSF, and protease inhibitor cocktail] and spun down at 13K with ultracentrifuge, performed at 4 ºC or on ice similar to all subsequent steps. The soluble fraction was applied to a 20-ml Q-Sepharose column, the subsequent flow-through was applied to a 2-ml Ni^{2+}-Sepharose column, and the Nickel-bound protein was then washed in 40 mM imidazole and then eluted in 300 mM imidazole. The eluted protein was dialyzed against 20 mM MES (pH 6.5), 500 mM NaCl, 20% glycerol, 1 mM EDTA, and 1 mM dithiothreitol, and concentrated to 1 ml with Amicon 50-kDa MWCO spin columns and stored at -80 °C. Concentration measurements of SRPK1 were made by A_{280} absorbance and by Bradford assay. Mutant SRPK1 constructs followed the same protocol.

Clk/Sty was transformed into *E. coli* BL21-DE3, grown at 37°C in 3 to 10 liters of LB broth supplemented with 100 μg/ml ampicilin, and induced with 0.8 mM isopropyl 1-thio-β-D-galactopyranoside for 5 h at room temperature. Cells were lysed by French Press in 50mM NaH_2PO_4 (pH 7.8), 500mM NaCl, 10mM imidazole, protease inhibitor cocktail (Roche Diagnostic), and 1 mM PMSF, performed at 4 ºC or on ice similar to all subsequent steps. The lysed material was spun down at 13K with ultracentrifuge. The soluble fraction was loaded onto a Ni^{2+}-Sepharose column, washed thoroughly with the same buffer containing 40 mM imidazole, and eluted with 300 mM imidazole. The eluted protein was dialyzed against 50 mM NaH_2PO_4 (pH 7.8), 10 mM NaCl, 1 mM EDTA, and 1 mM dithiothreitol, and concentrated to 1 ml with Amicon 30-kDa MWCO spin columns and stored at -80 °C. Concentration measurements were made by A_{280} absorbance and by Bradford assay.

The plasmids for the wild-type and mutant forms of His-tagged ASF/SF2 were transformed into the BL21 (DE3) *E. coli* strain, and the cells were then grown at 37 °C in 2 to 10 liters LB broth supplemented with 50 μg/mL kanamycin for pET28a constructs or with 100 μg/mL ampicillin for pET15b and pet19b constructs. Protein expression was induced with 0.4 mM isopropyl 1-thio-β-D-

galactopyranoside at room temperature for 5 hours. Cells were then pelleted and lysed by sonication using 30 ml of lysis buffer (0.1 M MOPS, 20 mM Tris-HCl, 0.3 M NaCl, 10% glycerol, 1 mM PMSF, and 8 M urea), performed at 4 ℃ or on ice similar to all subsequent steps. There was no insoluble fraction when spun at 14000 rpm for 30 minutes. The soluble fraction was applied to a 1.5-ml Ni^{2+}-Sepharose column and washed with lysis buffer containing 20 mM imidazole. The protein was re-folded by passing decreasing concentrations of urea through the column (8, 6, 4, 2, 1, and 0 M urea, in lysis buffer). The re-folded protein was eluted with 10 to 40 ml of elution buffer (0.1 M MOPS, 20 mM Tris-HCl, 0.3 M NaCl, and 10% glycerol, pH 2.2), concentrated to 1 to 10 ml with the appropriate MWCO Amicon spin columns and stored at -80 °C. Concentration measurements were made by A_{280} absorbance and by Bradford assay.

PP1γ was transformed into the BL21(DE3) *E. coli* strain. The cultures were grown at 37°C in 3 to 10 liters of LB broth supplemented with 100 μg/ml ampicillin, and induced with 0.1-0.4 mM isopropyl β-D-thiogalactoside at room temperature for 12–16 hours. Pelleted cells were lysed by sonication in 150 ml of lysis buffer [50 mM NaCl, 20mM MES, pH 6.5, 20% (vol/vol) glycerol, 1 mM PMSF, and protease inhibitor cocktail] and then spun down at 13K with ultracentrifuge, performed at 4 ℃ or on ice similar to all subsequent steps. The soluble fraction was applied to a 20-ml Q-Sepharose column, the subsequent flow-through was applied to a 2-ml Ni^{2+}-Sepharose column, and the Nickel-bound protein was then washed in 40 mM imidazole and then eluted in 300 mM imidazole. The eluted protein was dialyzed against 20 mM MES (pH 6.5), 500 mM NaCl, 20% glycerol, 1 mM EDTA, and 1 mM dithiothreitol, and concentrated to 1 ml with Amicon 50-kDa MWCO spin columns and stored at -80 °C. Concentration measurements of SRPK1 were made by A_{280} absorbance and by Bradford assay. Mutant SRPK1 constructs followed the same protocol.

Pin1 cell aliquots were gifts from Patricia Jennings lab. The cultures were grown at 37°C in three to ten liters of LB broth supplemented with 50 μg/ml kanamycin, and induced at OD of 1.2 with 0.5 mM isopropyl β-D-thiogalactoside at room temperature for 12–16 hours. Pelleted cells were lysed by sonication in 50 ml of lysis buffer [500 mM NaCl, 25 mM Tris, pH 8, 10 mM imidazole, 2 mM β-mercaptoethanol, 1% Tween 20, 1 mM PMSF, and protease inhibitor cocktail], performed at 4 ℃ or on

ice similar to all subsequent steps. The soluble fraction was applied to a 4-ml Ni^{2+}-Sepharose column for 45 minutes at 4 ºC, and the Nickel-bound protein was then washed in 10X volume lysis buffer, and then eluted in lysis buffer with 250 mM imidazole but no Tween 20. The eluted protein was dialyzed against 25 mM ammonium sulfate, 1 mM β-mercaptoethanol, 5 mM magnesium chloride, 2.5 mM calcium chloride, 50 mM Tris pH 8, 150 mM NaCl, and 10% glycerol. Dialyzed enzyme stock was concentrated to 1 ml with Amicon 50-kDa MWCO spin columns and stored at -80 °C. Concentration measurements were made by A_{280} absorbance and by Bradford assay.

GST-ASF plasmid was transformed into BL21(DE3) *E. coli* cells. The cultures were grown in 3 liters of LB media supplemented with 50 µg/ml kanamycin, and induced at OD of 0.4 to 10 with 0.4 mM isopropyl β-D-thiogalactoside at room temperature for 4 to 5 hours. Cell pellet was resuspended in lysis buffer [20 mM Tris pH 7.5, 100 mM NaCl, 1 mM DTT, and 10% glycerol] supplemented with PMSF and protease inhibitor cocktail tablet, performed at 4 ºC or on ice similar to all subsequent steps. Lysozyme might be added for improved lysis. The suspension was sonicated and then spun down at 13K with ultracentrifuge. The supernatant was passed through Q resin column and the elution was applied to 2-4 ml glutathione resin with salt added to reach 500 mM NaCl concentration. The glutathione resin was washed with 30-60 ml lysis buffer and 500 mM NaCl, and eluted with wash buffer plus 20 mM reduced form glutathione (pH readjusted for glutathione). The enzyme stock was dialyzed against wash buffer plus 1 mM EDTA, and then concentrated by Amicon 30-kDa MWCO spin columns and stored at -80 °C. Concentration measurements were made by Bradford assay due to contaminating DNA/RNA. Mutant GST-ASF constructs followed the same protocol.

D. Protein Phosphorylation Assay

The phosphorylation of ASF/SF2 (0.2 µM) by SRPK1 or Clk/Sty (1 µM) was carried out in the presence of 50mM MOPS (pH 7.4), 10mM free $MgCl_2$, 5 mg/ml bovine serum albumin, and [^{32}P]ATP (2000-20000 cpm/pmol) at 23 °C unless otherwise stated. Reactions were initiated with the addition of [^{32}P]ATP (100 µM) in a total reaction volume of 10 µl and then were quenched with 10 µl of SDS-PAGE loading buffer. The whole quenched reaction (20 µl) was loaded onto a 10-18% SDS-PAGE gel,

depending on the size of the substrate. The gels were run at 170-200 V for 45 to 90 minutes and then dried in cellophane overnight. Dried gels were then exposed with Kodak or Phenix imaging film, and protein bands corresponding to phosphorylated substrate were excised and counted on the ^{32}P channel in liquid scintillant. Specific activity determination was made with 1 µl aliquots from reaction mixture. The control group and the amount of enzyme and substrate may vary depending on the specific goal of the experiment, and details would be provided in figure legends for each specific experiment.

For equilibrium ATP-dependent phosphorylation mapping, also known as the ATP limitation assay, various amounts of [^{32}P]ATP were added to initiate the 10-µl reactions with enzyme (1 µM) and digestion mutant substrate (0.25 µM). After 20 min at 23 °C, proteolysis with 100 ng of LysC endoproteinase was carried out in a digestion buffer of 25 mM Tris–HCl (pH 8.5) and 1 mM EDTA for 1 hour (less than 1 µM ATP), 2 hours (between 1 and 5 µM ATP), or 4 hours (more than 5 µM ATP) at 37 °C. For kinetic time-dependent phosphorylation mapping, reactions were initiated with [^{32}P]ATP (1 or 100 µM) in a total reaction volume of 10 µL and then were cold-chased at 8 s or 10 min with 2.5 µL of 5 to 50mM ATP. Proteolysis with 100 ng of LysC was then carried out in the digestion buffer for 4 hours to overnight at 37 °C. Each 20-µL proteolysis reaction was quenched with 10 µL of SDS-PAGE loading buffer. The N-terminal and C-terminal fragments were excised and counted on the 32P channel in liquid scintillant.

Steady-state kinetic parameters for SRPK1(wt) and additional mutant constructs with a substrate peptide (YRTRDAPRERSPTR) were determined at 23°C in 50 mM MOPS, pH 7.4, 10 mM MgCl$_2$, [^{32}P]ATP (600–1,000 cpm/pmol). Enzyme, MgCl$_2$, and substrate were pre-equilibrated for 1 min before initiating the reaction with the addition of ATP. Reactions (20 µl) were quenched with 180 µl of 30% acetic acid, the products were applied to pre-equilibrated DE 52 columns (3 ml), which were washed with 30% acetic acid (5 ml), and the phosphorylated substrate was collected and its radioactivity was measured. Control experiments were performed to determine the background phosphorylation without enzyme.

E. Protein Dephosphorylation Assay

0.1-1 μM ASF/SF2 substrate was pre-phosphorylated with catalytic or stoichiometric amount of SRPK1 or Clk/Sty and limiting amount of [^{32}P]ATP. Dephosphorylation of phosphorylated ASF/SF2 was carried out with various concentrations of PP1α or PP1γ at 37 ℃ before the reaction was quenched with SDS-PAGE loading buffer or stopping buffer (vanadate and/or β-glycerolphosphate) at different time points.   In most cases phosphatase buffer from New England Biolab was used for pre-phosphorylation and dephosphorylation reactions [50 mM HEPES pH 7.5, 100 mM NaCl, 2 mM DTT, 0.025% Tween-20, and 0.1 mM EGTA].   Quenched samples were loaded onto SDS-PAGE gel and subsequent steps were described in section D.

F. Pull-down Assay

For the pull-down assay determining identity of the LysC-digested fragments, each 20-μL proteolysis reaction was incubated with Ni^{2+}-Sepharose resin at 23 °C for 1 hour. The resin was washed three to five times with 450 μL of wash buffer [25 mM Tris (pH 7.5), 1 M NaCl, 0.5% Triton X-100, and 15% glycerol], until there was minimal counts from the wash buffer. The bound proteins were eluted with 50 mM EDTA and SDS-PAGE loading buffer, boiled for at least 10 minutes. Each reaction was loaded onto 18% SDS-PAGE gel and the rest of the protocol for drying and counting were described above.

G. Mass Spectrometry Assay

MALDI-TOF analyses were carried out using a PerSeptive Biosystems Voyager DE PRO spectrometer. Preparation of phosphorylated protein samples was typically carried out using SRPK1 (5 to 10 μM) and ASF/SF2 (5 to 10 μM) in the presence of 50 mM Tris–HCl (pH 7.4), 10 mM free Mg^{2+}, and 1 mM DTT at room temperature (Chapter 3). Reactions were initiated with the addition of 3 mM ATP into a total volume of 40 μL. Reactions then were quenched after 1 hour with 3 vol of 8 M urea, desalted with ZipTip C18, and eluted with 80% acetonitrile and 0.08% TFA for MALDI-TOF analysis.

Unphosphorylated sample controls were prepared in the same manner, except that ATP was omitted. The matrix solution consisted of 5 mg/mL α-cyano-4-hydroxy cinnamic acid in 1:1:1 acetonitrile, ethanol, and 0.52% TFA. The final pH of the matrix solution was 2.0.

Alternatively SRPK1 (0.3 μM) and ASF/SF2 (1 μM) was used in the presence of 50 mM MOPS (pH 7.4) and 10 mM free Mg^{2+} at room temperature (Chapter 3 & 4). Data were then collected on a different Voyager machine, and the alteration in protocol was necessary due to machine-specific differences. Reactions were initiated with the addition of 0.3 mM ATP into a total volume of 200 μL. Reactions then were quenched after 3 hours with 5% acetic acid, desalted with ZipTip C18, and eluted with 80% acetonitrile and 2% acetic acid for MALDI-TOF analysis. The α-cyano-4-hydroxy cinnamic acid matrix solution was purchased from Agilent Technologies.

H. Circular Dichroism Assay

ASF/SF2 (2 μM) and ASF(188–248) (4.8 μM) were mixed with SRPK1 (90 nM) and a 12-fold molar excess of ATP (24 and 60 μM) in 5 mM acetate assay buffer (pH 5.5). After a 60 min reaction time, the samples were diluted in assay buffer and concentrated three times, using an Amicon Ultra 10K MWCO centrifugal filter. The circular dichroism (CD) spectra were recorded on an Aviv Circular Dichroism Spectrophotometer Model 202 purged with nitrogen using 2 mm thick quartz cells and 0.8 ml sample volume. Spectra were recorded from 205–260 nm, and appropriate baseline buffer spectra was collected and subtracted. Data points were converted from millidegrees to units of molar ellipticity (deg*cm2/dmole). Unphosphorylated proteins were mock phosphorylated with AMP-PNP and processed in the same manner as the phosphorylated samples.

I. Data Analysis

The progress-curve data for K_d determination, in which enzyme concentration varied, were plotted as a ratio of phosphate incorporated to total ASF/SF2 concentration ($[^{32}P]/[ASF/SF2]_o$) as a function of time and were fit to either a single or double-exponential function. In the latter case, the

amplitude of the first phase (α_1) represents the fraction of sites phosphorylated in the enzyme–substrate complex. The fraction bound of ASF/SF2 is then calculated from the ratio of α_1 and the total amplitude (α_{tot}) (Fraction bound = α_1/α_{tot}). The dissociation constant for the SRPK1:ASF/SF2 complex (K_d), measured by using dilution studies, was determined by plotting the fraction bound as a function of $[E]_o$ and fitting the curve to Eq. 1.

$$BF = \frac{K_d + [S]_o + [E]_o - \sqrt{(K_d + [S]_o + [E]_o)^2 - 4[S]_o[E]_o}}{2[S]_o}$$

(Eq. 1)

Equation 1 represents a quadratic solution for a high-affinity, enzyme-substrate interaction in which BF is the fraction bound, and $[S]_o$ and $[E]_o$ are the total substrate and enzyme concentration. For rough estimation of K_d, Michaelis-Menton equation may be used instead.

Progress curve data for ASF/SF2 phosphorylation, in which biphasic behavior were observed, were plotted as ratios of incorporated phosphate to the total substrate concentration as a function of time and were fit to a double-exponential function with rate constants k_1 and k_2 and amplitudes α_1 and α_2. See Table 3.1 for examples in which this mode of curve-fitting was used. Most parameters could be found in figure legends for each specific kinetic experiment.

In competition experiments, catalytic amount of enzyme was used to turn over less than 10% of 50 nM of substrate ASF/SF2 construct in a steady-state experiment. The putative competitor should bind at the same site as substrate and compete for the limited enzyme, and by varying the concentration of the competitor, the K_i inhibition constant may be determined. The K_i is the same as K_d for a purely competitive system in which substrate and competitor binds at the same site. The relative initial velocities for the competition data were fit to Eq. (2):

$$\frac{v_i}{v_o} = \frac{[ASF(\Delta RRM2)] + K_m}{[ASF(\Delta RRM2)] + K_m\left(1 + \frac{[I]}{K_I}\right)}$$

(Eq. 2)

where v_i/v_o is the relative initial velocity (the ratio of v in the presence of an inhibitor to v in the absence of an inhibitor), K_m is the Michaelis constant for ASF(ΔRRM2), [I] is the concentration of the inhibitor, and K_I is the dissociation constant for the inhibitor. This formula could be used for other substrates if the K_m is determined beforehand. See Chapter 3 for examples.

Chapter 3

Ordered Multi-site Phosphorylation of ASF/SF2 by SRPK1

A. Introduction

Previous studies showed an unusually stable SRPK:ASF/SF2 complex ($K_D \sim 50$ nM), in which multi-site phosphorylation has a higher forward catalytic rate compared to dissociation and proceeds processively [51]. However, the product release step is still not fully resolved, as a stable enzyme:substrate complex may have a slow product release step. While other studies showed that phosphorylated ASF/SF2 has low affinity for SRPK1 [104] and quick release of phosphorylated product is possible, this model of decreasing affinity does not support a processive mechanism for the whole RS1 segment. Additional evidence supporting the weakening affinity came from MALDI-TOF studies, in which the final phosphoryl content shifts from assay to assay, suggesting premature termination of the reaction due to poor binding [48, 51]. However, if the affinity weakens, the rate of reaction may actually slow down. The final piece of information prompting fresh inquiries came from the original K_d determination assay. At saturating SRPK1 concentration the data was fitted to a single-exponential curve, but sometimes the curve fit was not perfect (Fig. 3.1A) [48]. This raises the possibility that a slower phase of the phosphorylation reaction exists but was ignored in the past. An important goal of the present study was to fully characterize the full extent of ASF/SF2 phosphorylation by SRPK1 and to assess its processive nature, in both cases focusing on the latter stage of phosphorylation.

Processive enzymes are rare in nature, and DNA polymerase, a well-known example, follows a highly ordered catalytic mechanism. DNA polymerase requires a priming base placed by DNA primase and a specific initiation sequence to initiate its reaction. During the elongation stage it is directional, sequential, and processive. An interesting question to be addressed is whether a processive enzyme is always characterized by additional ordered mechanisms of catalysis. SRPK1 does not require any priming phosphate catalyzed by any other enzyme, but it may require a specific initiation sequence on ASF/SF2, and it may catalyze a directional phosphorylation mechanism starting at either terminus of the RS1 domain. An important goal of the present study was to characterize any initiation sequence requirement and the order of phosphorylation.

A. Results

a. Characterize the Slow Phosphorylation Phase by SRPK1

Single-turnover kinetics were applied, in which a saturating amount of SRPK1 enzyme was used to bind all available ASF/SF2 molecules. The phosphorylation assay was carried out for up to one hour with SRPK1 (1 μM) and ASF/SF2 (0.2 μM), and the progress curve displayed a biphasic behavior previously ignored in 10-minute time courses. This was confirmed in many repeat assays, and the representative progress curve is shown in Fig. 3.1B (solid triangle). The first phase lasted for about one minute, and it can be called the "fast phase" due to the 3.7 min^{-1} high exponential reaction rate. During the fast phase about 12 phosphoserine sites were catalyzed by SRPK1, the majority of the reaction complete at the end of this phase. The second phase spanned from 1 minute to 15 minutes, and it can be called the "slow phase" due to the 0.13 min^{-1} low exponential reaction rate. During the slow phase the remaining 2 phosphoserine sites were catalyzed by SRPK1, until no noticeable progress was made beyond 15 minutes. The magnitude of these phases were larger than errors commonly associated with this assay (<= 20%) and biphasic behavior was also observed in most mutant SRPK:ASF/SF2 complex reactions (Fig. 3.1B), suggesting that it is truly part of SRPK1 phosphorylation mechanism. Overall, the data suggested that SRPK1 phosphorylates ASF/SF2 in a biphasic manner, and a single-turnover saturating enzyme time course that lasts for 15-20 minutes can better characterize that slow phase than a common 3-minute time course in previous studies.

b. Characterize the Maximum ASF/SF2 Phosphorylation Possible by SRPK1

Given the discovery of the slow phase of phosphorylation of ASF/SF2 by SRPK1, the maximum ASF/SF2 phosphoryl content must be re-assessed under longer time courses. The [^{32}P] phosphorylation assay cannot provide this data due to possible decay of phosphorylated ASF/SF2 and the presence of inactive substrate in the preparation. MALDI-TOF mass spectrometry identifies the most common unphosphorylated and phosphorylated ASF/SF2 species, where the difference in molecular weight indicative of phosphoryl content. MALDI-TOF analyses were performed by Adolfo Velazquez-Dones using the urea quench protocol and by Jonathan Hagopian using the acetic acid quench protocol (see Chapter 2), and the results suggested that SRPK1 phosphorylates ASF/SF2 on 12-15 sites (Fig. 3.1C &
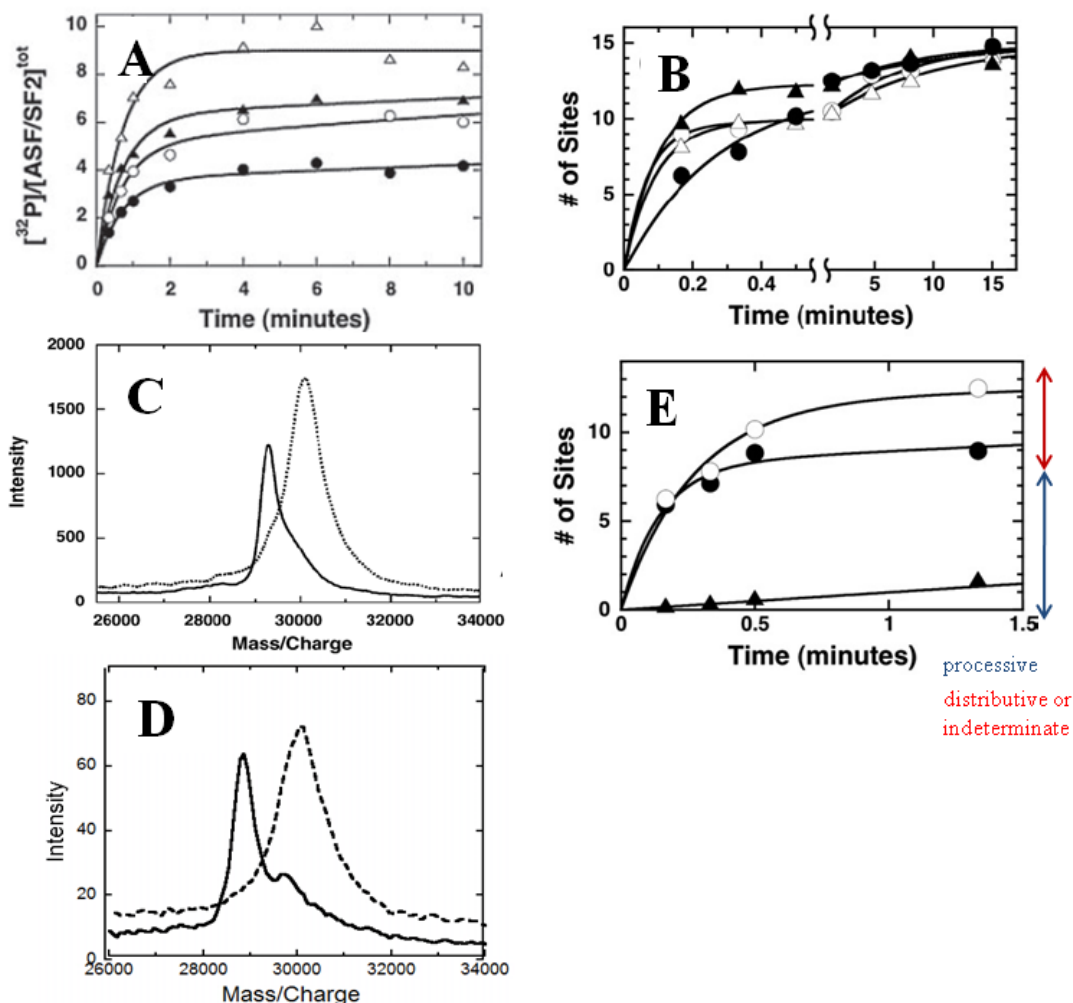
Figure 3.1 Full Extent of Processive Phosphorylation. A) Single-turnover kinetics is not always a single-phase exponential. Progress curve at saturating SRPK1 concentration shows deviation from single-exponential fit ($\triangle$). B) SRPK1 phosphorylate ASF/ASF2 in a bi-phase. The data was fitted to a double-exponential equation. The wild-type ($\blacktriangle$) fast and slow phase magnitude are 12 and 2 sites respectively. The fast and slow phase rate constant are 3.7 and 0.13 min^{-1} respectively. Rest of the data were obtained with mutant ASF constructs. C) Older MALDI-TOF spectra of c-wt-ASF. Data was prepared by Adolfo Velazquez-Dones. Solid line represents unphosphorylated ASF/SF2, and dotted line represents phosphorylated ASF/SF2. Increase in molecular weight is consistent with 12 sites being phosphorylated. D) Newer MALDI-TOF spectra of c-wt-ASF. Solid line represents unphosphorylated ASF/SF2, and dotted line represents phosphorylated ASF/SF2. Increase in molecular weight is consistent with 15 sites being phosphorylated. E) Start-trap experiment. Start-trap progress curve ($\bullet$) fitted to single-exponential plus linear function shows 8 sites processively phosphorylated, compared against 12 possible sites shown in the control curve ($\circ$). 4 sites are most likely distributively phosphorylated. Trap-start curve ($\blacktriangle$) is fitted to linear function with rate constant of 1 site/min.

D), an expected increase from 8-10 sites reported previously. The inherent error range associated with MALDI-TOF mass spectrometry of large proteins makes it difficult to fully characterize the slow phase of the reaction, and throughout this thesis 12 and 15 sites may be used interchangeably as the wt maximum phosphorylation. Overall, the data suggested that SRPK1 phosphorylates ASF/SF2 on 12 to 15 sites in a 15-minute single-turnover time course.

c.   Re-assess Processive Phosphorylation of ASF/SF2 by SRPK1

In order to investigate the processive nature of the other phase of SRPK1 phosphorylation, the processivity assay was improved in two ways: the time course was extended to 15 minutes and doping points were added to ensure common endpoints of the start-trap, trap-start, and control curves (Fig. 3.1B). For the doping points, part of the reaction mixture was added to 5 μM extra SRPK1 for 25 to 35 minutes to ensure the same amount of ASF/SF2 was added to all three progress curves and remained phosphorylatable throughout the reaction. Phosphorylated ASF/SF2 is unstable at physiological pH, and a new control experimental set is necessary to monitor the overall experiment for irregularities.

The start-trap progress curve was generated with pre-mixed SRPK1 (1 μM) and ASF (0.2 μM), the reaction initiated by the mixture of ATP and kinase-dead SRPK1 trap (40 μM). The purpose of the trap was to prevent additional phosphorylation of released phosphorylated ASF/SF2 intermediate, and its efficiency was tested with trap-start progress curve, in which kinase-dead SRPK1 was pre-mixed with enzyme:substrate mixture.

The first 2 minutes of the trap-start progress curve was fitted to a linear slow rate, and the observed slow phase of the control curve (2 phosphorylation sites) was roughly as slow as the trap-start progress curve, so that part of the reaction was not inhibited by kinase-dead SRPK1, and no conclusion could be drawn as to its mechanism of action. However, given the 100-fold reduction in catalytic rate, the slow phase is most likely characterized by distributive phosphorylation. The remaining 12 sites of the fast phase of the control curve represent the maximum possible extent of processivity observable. The first 2 minutes of the start-trap progress curve was fitted to a fast phase plus a linear slow phase with the same rate as trap-start progress curve (Fig. 3.1E). The fast phase in the start-trap curve accounted for 8

sites out of 12, suggesting that SRPK1 dissociates after phosphorylating 8 sites processively. SRPK1 may rebind to finish the other 4 sites processively, but that cannot be determined with the current assay.

Finally, when pre-phosphorylated ASF/SF2 (8-10 sites) was used as substrate in the start-trap experiment, the trap did not show efficient inhibition, consistent with previous results showing that the second phase cannot be effectively inhibited or observed in terms of processivity and the current assay is limited to characterizing only the initial fast phase. Overall, the data suggested that roughly 8 phosphorylation events occurs processively, and the rest are too slow to be characterized but most likely distributively catalyzed.

d. Probing for an Initiation Site

SRPK1 phosphorylates at least 12 serines in N-terminal part of the RS domain of ASF/SF2, designated as the RS1 segment. If a strictly required initiation site exists, mutation of that serine would block all subsequent phosphorylation of the remaining serines, leading to significantly reduced phosphoryl content at completion of reaction. To scan for this potential initiation site, 5 serine-to-alanine block mutant ASF constructs were created to cover the whole RS1 segment (Fig. 3.2A), and phosphorylation parameters were compared against wt control in different assays. All constructs contained N-terminal Histidine tags, to control for possible phosphorylation artifacts in the C-terminus of the substrate. The MALDI-TOF mass spectrometry spectra showed that the maximum phosphoryl content catalyzed by SRPK1 decreased by roughly 4 in each case (Fig. 3.2B), suggesting that no single serine within RS1 is essential and required for subsequent phosphorylation. Single-turnover time courses with saturating enzyme led to the observation that the mutant constructs exhibits similar biphasic behavior with dominating fast phase and the whole reaction is complete by 15 minutes (Fig. 3.2C). Finally, steady-state competition studies revealed that the mutant constructs compete as well as wt-ASF and have similar affinities for SRPK1, measured by the dissociation constant K_d (Fig. 3.2D). Only the ASF(4SA205) construct showed a three-fold increase in K_d, close to the error ranges associated with this type of assay. Overall, the data suggested that SRPK1 phosphorylates ASF/SF2 without a strict initiation
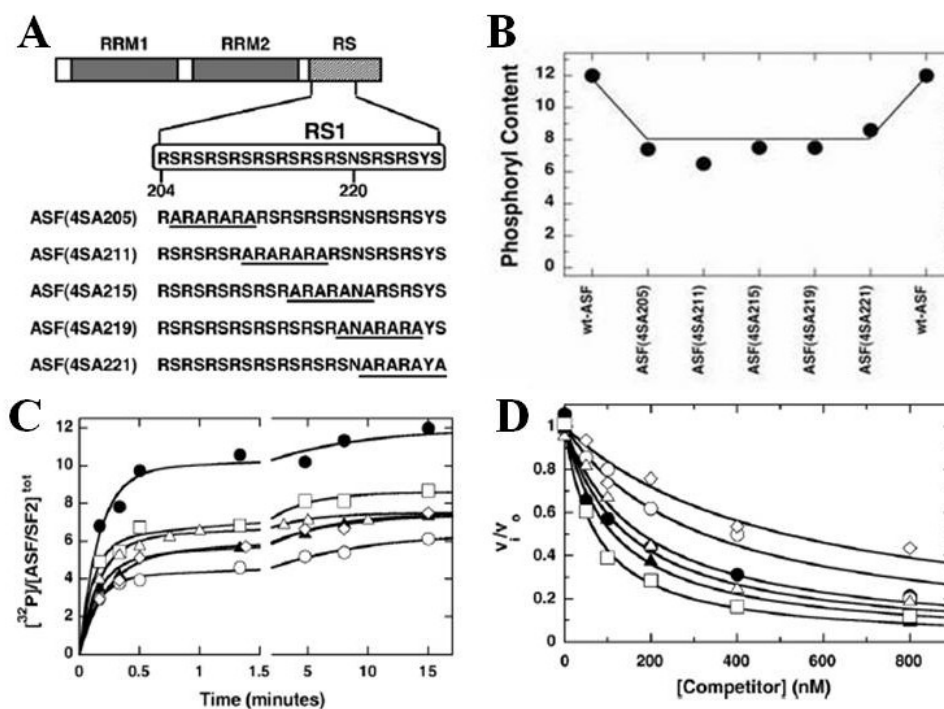
Figure 3.2 Phosphorylation Analysis of Alanine Block Mutants. A) Serine-to-alanine block mutations in ASF/SF2. B) Phosphoryl content. Maximum number of phosphoserines catalyzed by SRPK1 were calculated from MALDI-TOF spectra of various ASF constructs. C) Progress curves. SRPK1 (1 μM) was mixed with 100 nM wt-ASF (●), ASF(4SA205) (◊), ASF(4SA211) (○), ASF(4SA215) (□), ASF(4SA219) (△), and ASF(4SA221) (▲), and the reactions were initiated with 100 μM [^{32}P]ATP. All data are fitted to double-exponential functions, and the parameters displayed in Table 3.1. D) Binding affinities.   Relative velocities (v_i/v_o) for the phosphorylation of ASF(ΔRRM2) were monitored as a function of wt-ASF and mutant substrates (competitors). ASF(ΔRRM2) were mixed with SRPK1 (2 nM) and varying competitor concentrations (0-800 nM) in the presence of 100 μM [^{32}P]ATP. Inhibition constants were measured using Eq. 2 and displayed in Table 3.1.

Table 3.1 Kinetic and Mechanistic Analyses of wt and Mutant ASF/SF2. a) All data were obtained through MALDI-TOF mass spectra. b) Inhibition constants were obtained through steady-state competition studies against 50 nM ASF(ΔRRM2). c) Single-turnover rate constants were measured with 1 μM SRPK1 and 100 nM substrate.

Substrate	Phosphoryl content[a]	K_I (nM)[b]	k_1 (min^{-1})[c]	α_1[c]	k_2 (min^{-1})[c]	α_2[c]
ASF/SF2	12	100$\pm$10	6.0$\pm$1.3	9.8$\pm$0.70	0.13$\pm$0.09	2.2$\pm$1.0
ASF(4SA205)	7.4	340$\pm$50	4.4$\pm$1.0	5.3$\pm$0.41	0.18$\pm$0.06	2.1$\pm$0.5
ASF(4SA211)	6.5	220$\pm$20	7.0$\pm$0.91	4.1$\pm$0.15	0.12$\pm$0.02	2.4$\pm$0.1
ASF(4SA215)	7.5	80$\pm$6	6.4$\pm$1.0	5.1$\pm$0.27	0.19$\pm$0.04	2.4$\pm$0.3
ASF(4SA219)	7.5	120$\pm$10	6.0$\pm$0.53	6.1$\pm$1.3	0.28$\pm$0.08	1.4$\pm$0.2
ASF(4SA221)	8.6	50$\pm$3	7.7$\pm$2.3	6.0$\pm$0.47	0.30$\pm$0.12	2.6$\pm$0.5
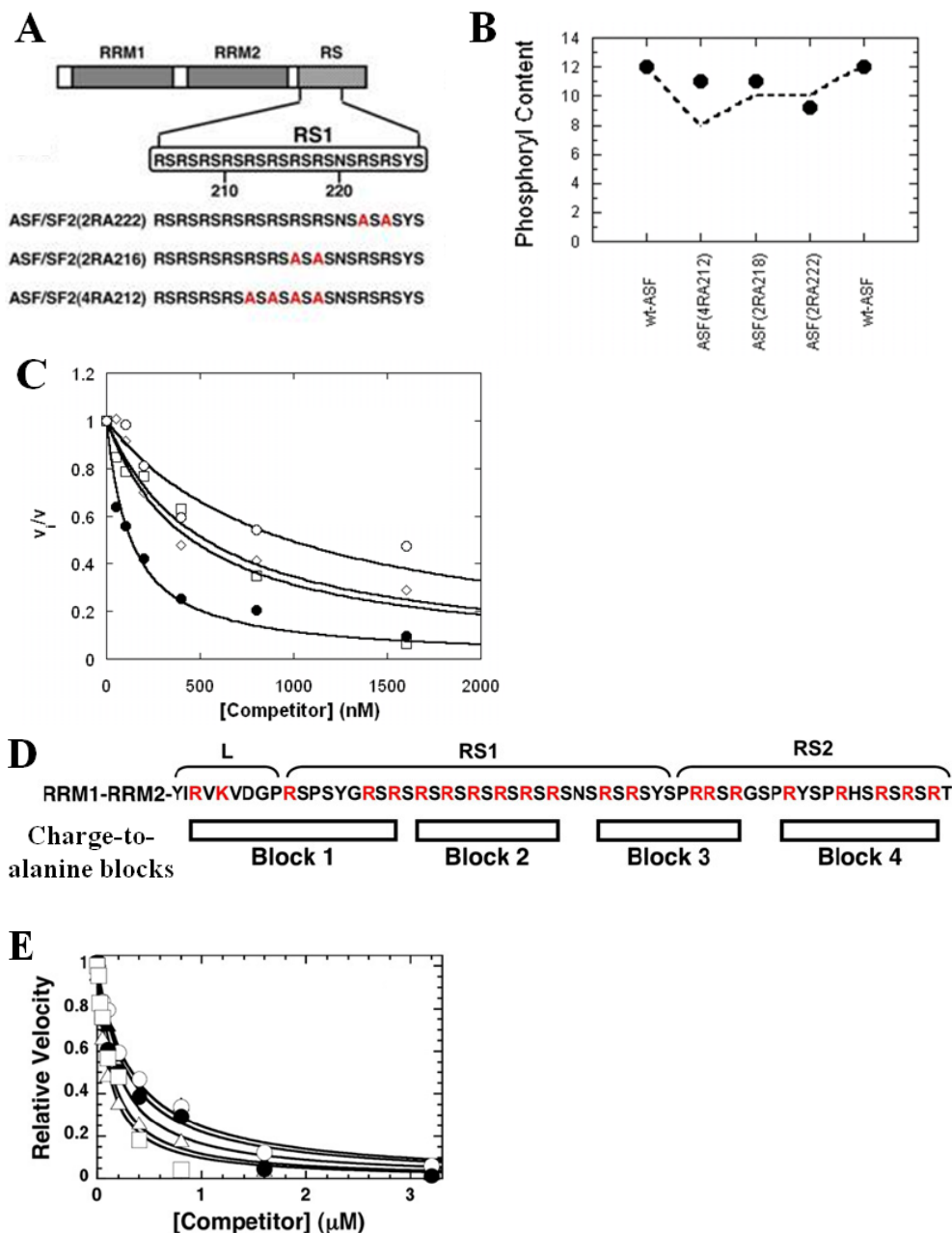
Figure 3.3 Phosphorylation Analysis of Arginine Block Mutants.   A)   Arginine-to-alanine block mutations in ASF/SF2 RS1 domain. B) Phosphoryl content. Maximum number of phosphoserines catalyzed by SRPK1 were calculated from MALDI-TOF spectra of various ASF constructs. C) Binding affinities.   Relative velocities (v_i/v_o) for the phosphorylation of ASF(ΔRRM2) were monitored as a function of wt-ASF and mutant substrates (competitors). ASF(ΔRRM2) were mixed with SRPK1 (2 nM) and varying competitor concentrations (0-1600 nM) in the presence of 100 μM [^{32}P]ATP. Competitors include wt-ASF (●), ASF(2RA216) (□), ASF(2RA222) (◊), and ASF(4RA212) (○). D) Arginine-to-alanine block mutations spanning the whole ASF/SF2 RS domain. E) Binding affinities. Relative velocities for the phosphorylation of ASF(ΔRRM1) were monitored as a function of wt-ASF and mutant substrates (competitors).   Competitors include wt-ASF (□), ASF(block1) (●), ASF(block2) (○), ASF(block3) (▲), and ASF(block4) (Δ). Data from E) were collected by Jonathan Hagopian.

center, and SRPK1 can bypass any serine mutation and finish phosphorylation of the remaining residues without significantly changed kinetic or binding profiles.

In addition to the serines, there are also multiple arginines in the RS1 domain, and it is possible that SRPK1 recognizes and requires certain arginines for initiation. Double and quadruple mutations of arginines to alanines in the RS1 domain (Fig. 3.3A), and the maximum phosphorylation catalyzed by SRPK1 did not change significantly beyond error of the MALDI-TOF mass spectrometry assay (Fig. 3.3B). Binding studies revealed small changes in K_d values for the double mutants and significant looser binding affinity for the 4-arginine mutant (Fig. 3.3C), leading to additional affinity studies conducted by Jonathan Hagopian. Larger blocks of the whole RS domain were mutated to probe for SRPK1 binding site (Fig. 3.3D), but none of the block mutants displayed the change in binding affinity previously observed in smaller block mutants. Phosphoryl content by MALDI-TOF spectra were also normal compared against wt-ASF. Kinetic studies were also conducted but none of the arginine constructs showed significant change in biphasic reaction progress or rate. Overall, the data suggests that SRPK1 can bypass any arginine mutation and phosphorylate a fixed number of serines.

Having established that SRPK1 phosphorylates ASF/SF2 in a very flexible manner, an additional series of alanine insertion mutant constructs were tested to determine whether SRPK1 can bypass 4, 8 or even 16 alanines inserted into the middle of the RS1 segment (after S211). While binding affinity and phosphoryl content did not change significantly, the number of sites phosphorylated processively on ASF(i8A) and ASF(i16A) mutants decreased significantly, suggesting that SRPK1 may need to scan for distant RS/SR repeats to maintain the processive mechanism (data not shown).

e.   **Developing a New Gel-based Phosphomapping Assay**

Lack of a strict initiation site on ASF/SF2 precludes the possibility of a simple determination of phosphorylation direction, so new assays must be developed to localize the initial phosphorylation event and to track subsequent progress. Peptide sequencing through trypsin digestion, HPLC separation, and mass spectrometry identification is extremely difficult due to the prevalent arginine repeats in the RS domain. An endoproteinase that cut at much fewer sites is needed, and Lys-C was chosen because

ASF/SF2 has only 8 natural lysines in its sequence, none of them in the RS domain. An arginine was substituted by lysine in the middle of the RS1 domain (R214K) to divide the serines phosphorylated by SRPK1 into roughly two equal parts, and this simple assay limited the model of phosphorylation to three possibilities: N->C directional, C->N directional, or non-directional. MALDI-TOF mass spectrometry is not an option for subsequent analysis because fragments with different phosphoryl content may fly at different efficiencies, and it is very easy to overlook peaks representing certain phosphorylated fragments. This is unacceptable when accurate quantification of the fragments is necessary to investigate possible regiospecific phosphorylation preference. Instead, a C-terminal Histidine tagged vector was used to separate the two phosphorylated RS fragments by Nickel resin pull-down of the Histidine tagged C-terminal fragment. Unfortunately, Nickel resin only bound half or less of the Histidine tagged C-terminal fragment, and thorough washing to remove nonspecific binding of the no-tag N-terminal fragment led further loss of the bound C-terminal fragment, making it unfeasible as a quantitative assay (data not shown). Gel separation of the two fragments was attempted, but the two fragments were similar in size (3 & 5 kDa for N- & C-fragments) and co-migrated as one band in 18% or 20% SDS-PAGE gels necessary to trap the small fragments (data not shown).

In order to obtain two fragments of unequal size separable on SDS-PAGE gel for accurate quantitative analysis of phosphate placement, two additional strategies were tested. R214K substitution was performed on ASF(ΔRRM2) to generate a digestion mutant ASF(ΔRRM2-214) with all 5 lysines in the RRM2 and the linker region removed, but the resultant N- and C-fragments were still fairly close in molecular weight (8 & 5 kDa), leading to overlapping of bands inseparable by SDS-PAGE gel (data not shown). Finally, using ASF(R214K) as a template, 5 lysines were mutated into arginines in the RRM2 domain and in the docking linker region connecting RRM2 and RS domains (Fig. 3.4A). Upon Lys-C digestion of the new ASF(5R1K) construct, the phosphorylated RS domain was split into two fragments of 5 and 19 kDa, easily separable by SDS-PAGE gel. An autoradiogram was obtained to detect only fragments marked with [^{32}P]phosphates (Fig. 3.4B & C).

However, as in the case of most proteolysis, complete digestion could not be achieved without
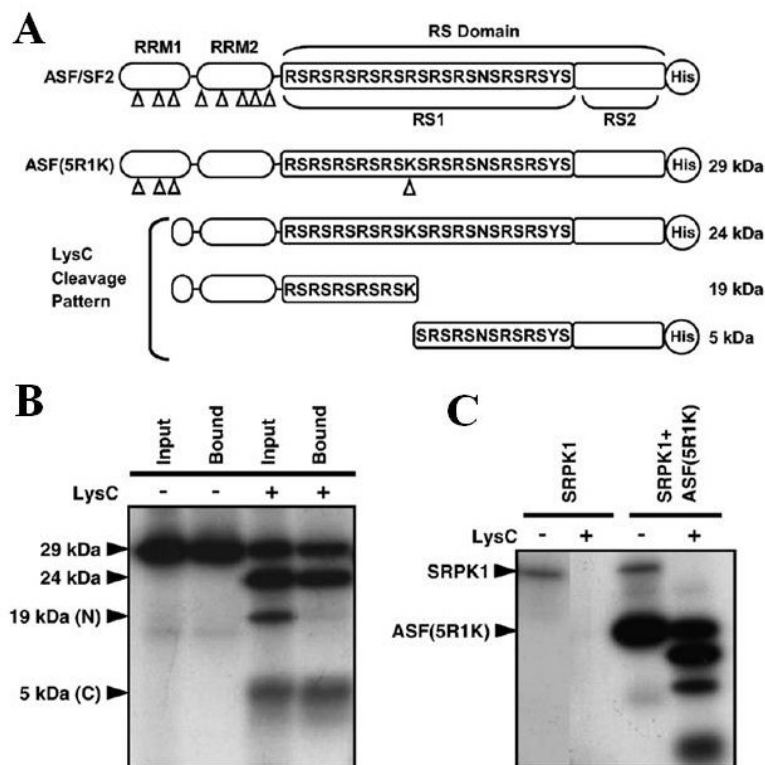
Figure 3.4 wt-ASF/SF2 Structure and ASF(5R1K) Digestion Pattern. A) Domain organization of wt-ASF/SF2 and expected LysC digestion fragments of ASF(5R1K). LysC protease digestion sites are shown using open arrows. Complete sequence for RS1 domain is shown. Predicted peptide fragments carrying phosphorylation sites for ASF(5R1K) are shown. B) Identification of phosphorylated peptide fragments by autoradiography. ASF(5R1K) (250 nM) was phosphorylated using SRPK1 (1 μM) and [^{32}P]ATP (100 μM) for 20 minutes and then treated with LysC (5 ng/μL) for 4 hours at 37 ºC before a 1-hour incubation with Ni^{2+}-Sepharose resin. After the peptides had been washed, they were eluted from the resin with 50 mM EDTA and boiled in SDS-PAGE loading buffer. C) LysC fragments from SRPK1 or ASF(5R1K). The LysC cleavge of SRPK1 (1 μM) was carried out in the absence and in the presence of ASF(5R1K), and the products were displayed by autoradiogram.

over-digestion of fragments into unrecognizable small chunks of amino acids, when LysC concentration or duration of digestioned increased 4-10 fold. The optimized LysC digestion pattern was slightly more complicated than planned (Fig. 3.4A). Each fragment observed must be analyzed through its migration relative to prestained molecular weight standards and the presence or the lack of Histidine tag. The top band in the input lane of Fig. 3.4B was the undigested ASF(5R1K), confirmed by molecular weight of 28.5 kDa and presence of the Histidine tag detected by Nickel resin pull-down assay. The next band below it was an unexpected intermediate band. This fragment was roughly 24 kDa, the combined size of the two phosphorylated fragment, and the presence of the C-terminal Histidine tag confirmed its identity. Possibly due to conformational changes of the RS domain upon phosphorylation, LysC digestion was inefficient in the crucial RS1 lysine (K214), and this intermediate band resisted digestion. The 19 kDa band was the N-terminal fragment containing the first half of RS domain due to the correct molecular weight and the lack of Histidine tag. The final 5 kDa band was the C-terminal fragment due to the correct size and the presence of Histidine tag. Guanidine hydrochloride, acetonitrile, Tween80, boiling of substrate, urea and even SDS were tested to solubilize the fragments for improved digestion, but in every condition the intermediate band persisted and resisted complete digestion (data not shown). However, assay optimization efforts led to the use of Protein Lo-Bind tubes from Eppendorf and Hydrologix pipet tips from Fisher Scientific. Loss of protein fragments to tube and tip surfaces were reduced (data not shown), and the yield of the digested N-terminal and C-terminal fragments improved.

As a new assay, this phosphomapping technique and the ASF(5R1K) mutant involved must be fully tested and characterized to ensure its validity. SRPK1 is known to auto-phosphorylate very slowly, and at much fewer residues than Clk/Sty. Proteolysis of auto-phosphorylated SRPK1 showed no noticeable contribution to the ASF(5R1K) digestion pattern (Fig. 3.4C). MALDI-TOF spectra of the C-terminal Histidine tagged wt construct c-wt-ASF and ASF(5R1K) both showed 12 sites in phosphoryl content catalyzed by SRPK1 (Fig. 3.5A & B). Binding studies revealed that ASF(5R1K) has similar dissociation constant as wt (Fig. 3.5C), and both were shown to be processively phosphorylated by SRPK1 for at least the first 5 sites out of 12 (Fig. 3.5D & E). Overall, the data suggested that ASF(5R1K) behaves the same as wt-ASF, arginine-lysine mutations do not affect the phosphorylation of
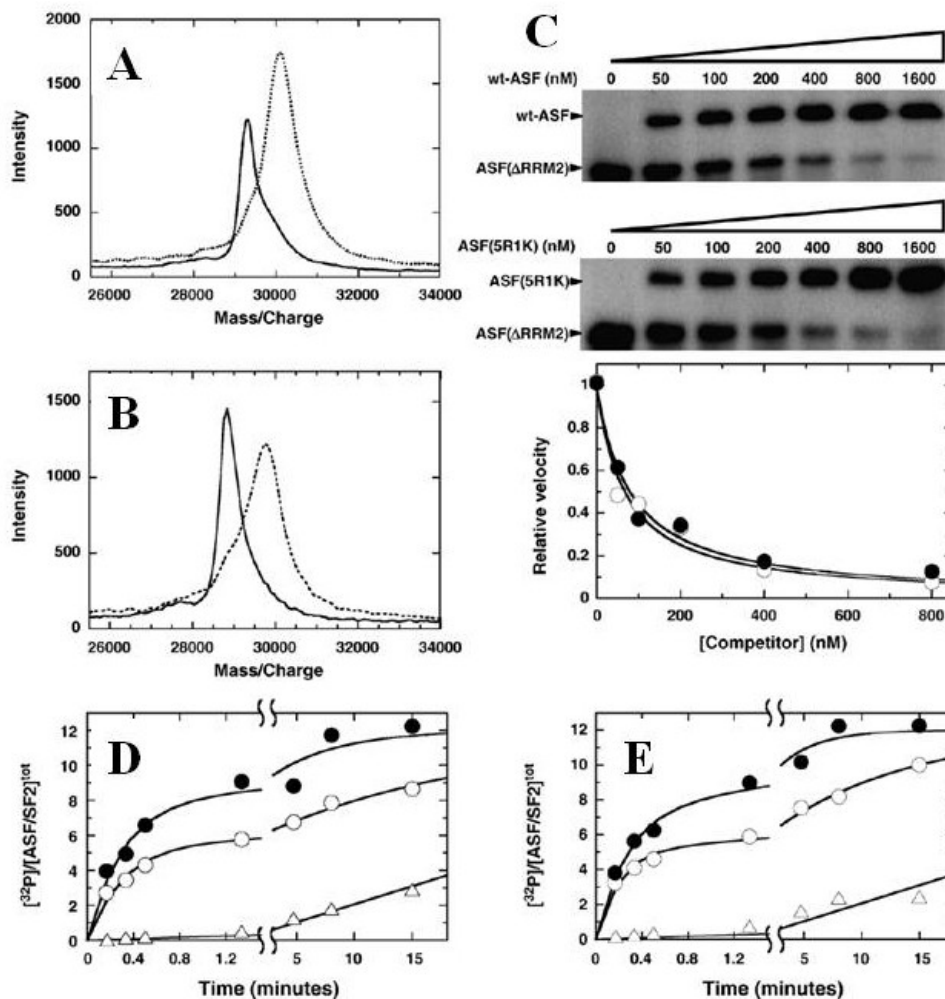
Figure 3.5 Arginine-lysine Mutations Do Not Affect the Phosphorylation of the RS Domain. A) and B) MALDI-TOF spectra of wt-ASF/SF2 and ASF(5R1K). Data for wt-ASF/SF2 are taken from Fig. 3.1C. The substrate (10 μM) was mixed with SRPK1 (10 μM) for 30 minutes in the absence and in the presence of ATP (3 mM). The increase in mass was consistent with the phosphorylation of 12 serines. C) Binding affinity of ASF(5R1K). The phosphorylation of ASF(ΔRRM2) was monitored as a function of increasing wt-ASF/SF2 or ASF(5R1K) using SDS-PAGE autoradiography. Decreases in the relative initial velocity for ASF(ΔRRM2) are plotted as a function of competitor concentration and fitted to Eq. 2 to obtain K_i values of 50 ± 5 and 45 ± 6 nM for wt-ASF and ASF(5R1K), respectively. D) and E) Single-turnover kinetic analyes of wt-ASF/SF2 and ASF(5R1K). SRPK1 (1 μM) was mixed with 200 nM substrate, and the reactions were initiated with 100 μM [^{32}P]ATP. All data are fitted to double-exponential functions with similar rate constants and amplitudes of 8 sites and 3-4 min^{-1} for the fast phase, and 4 sites and 0.2-0.3 min^{-1} for the slower phase, respectively (●). Both reactions were inhibited by preequilibration with 60 μM kinase-dead SRPK1 prior to the start of the reaction with [^{32}P]ATP in trap-start experiments (Δ). In comparison, the addition of kdSRPK1 to [^{32}P]ATP in start-trap experimens led to a rapid incorporation of about 5 phosphates into both cases before inhibition (○).

the RS domain, and ASF(5R1K) can be used interchangeably to study the order of phosphorylation in wt constructs.

f.   Equilibrium Analysis of Regiospecific Phosphorylation

Having designed a digestion mutant ASF/SF2 construct that yields two distinct RS domain fragments and is phosphorylated by SRPK1 indistinguishably from wt-ASF, the next step is to track phosphorylation initiation and progress. In order to study the initial phosphorylation event where the first phosphate is placed on unphosphorylated ASF, the ATP limitation assay was developed to regulate phoshoryl content of ASF(5R1K). Under single-turnover mode with saturating SRPK1 enzyme, all ASF(5R1K) molecules would be bound in the enzyme:substrate complex, and ATP would be added in small amounts to slowly titrate each phosphorylation site, shared equally by all ASF(5R1K) molecules. By limiting ATP concentration the reaction proceeds until a specific number of sites are phosphorylated, but the reaction terminates before completion due to a lack of ATP. As shown in Fig. 3.6A, 0.2 μM ATP titrated one phosphoserine on ASF(5R1K) and 5 μM ATP minimum was required for almost full titration of all phosphorylation sites.

Digestion of these phosphorylation intermediate products yielded the expected digestion profile with four phosphorylated bands, three of which shown in Fig. 3.6B, the last one the undigested phosphorylated ASF. Reaction progress was tracked through ^{32}P incorporation into digested fragments. Specific radioactivity was adjusted for each ATP concentration used to normalize the C-terminal fragment intensity in autoradiography, but high ATP concentration negatively impacted LysC digestion efficiency. Thus, the C-terminal fragment was lighter than expected and the intermediate fragment heavier than expected for 100 μM ATP. Despite imperfect normalization, the N-terminal fragment gradually increased in intensity with increasing phosphoryl content, and the C-terminal fragment suffered a slight decrease in intensity. At lowest phosphoryl content the dominant majority of ^{32}P incorporation occurred at the C-terminal fragment. The ratio of ^{32}P in the N-terminal fragment to ^{32}P in the C-terminal fragment, or the N/C ratio, increased gradually with ATP concentration and phosphorylation content plotted in Fig. 3.6A. This trend suggested that SRPK1 catalyzes C->N directional phosphorylation. This
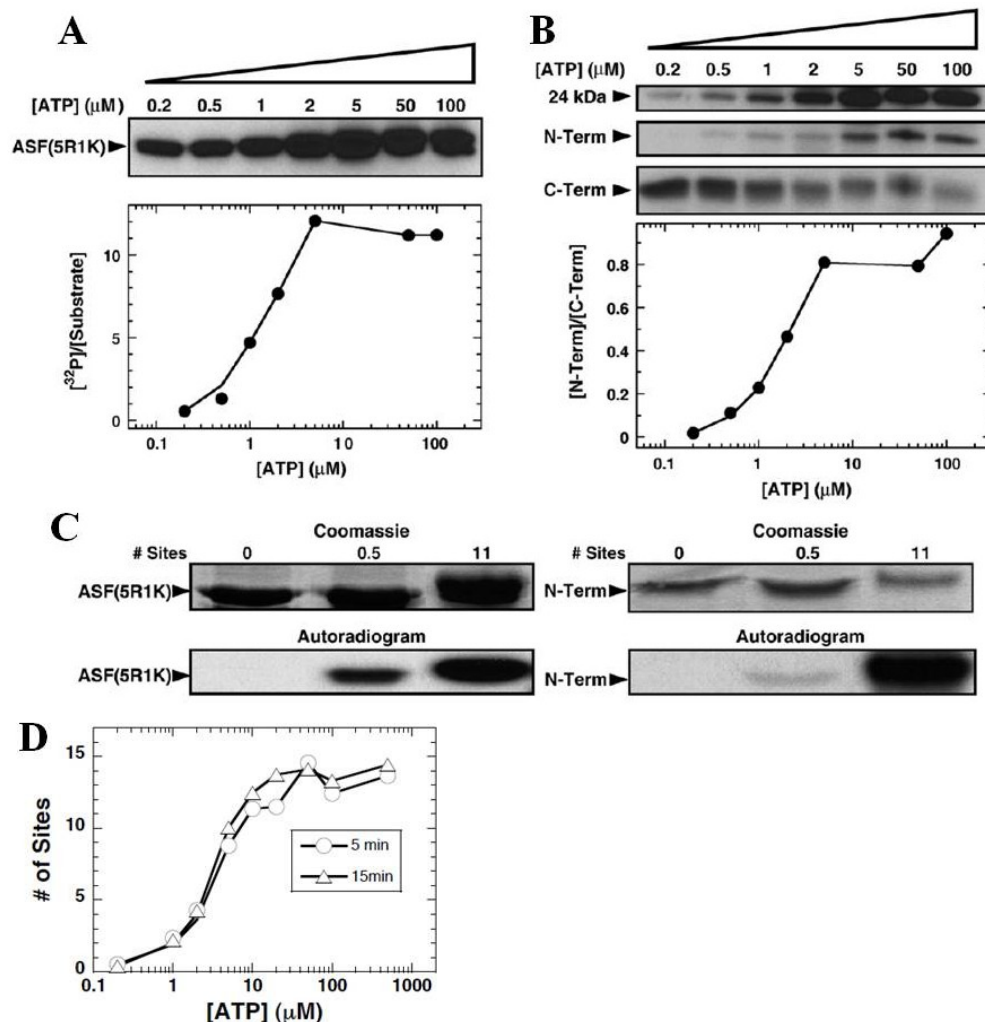
Figure 3.6 Equilibrium Analysis of Regiospecific Phosphorylation. A) Controlling ASF(5R1K) phosphoryl content through ATP limitation. ASF(5R1K) (250 nM) and SRPK1 (1 μM) were pre-equilibrated before a series of ATP concentrations (0.2, 0.5, 1, 2, 5, 50, and 100 μM ATP) were used to initiate a 20-minute reaction and displayed by SDS-PAGE autoradiography. Phosphoryl content is plotted as a function of increasing ATP concentration. B) N- and C-terminal fragments after LysC digestion. After the phosphorylation reaction described in A), half of each reaction volume was treated with LysC (5 ng/μL) at 37 ºC for either 1 hour (0.2, 0.5, 1, and 2 μM) or 4 hour (5, 50, and 100 μM). N- and C-terminal and 24-kDa LysC fragments from the ASF(5R1K) RS domain are shown, and the ratio of ^{32}P in the N-terminal fragment to ^{32}P in the C-terminal fragment is plotted against ATP concentrations. C) Coomassie staining of the N-terminal fragment. The N-terminal fragment upon the LysC treatment of ASF(5R1K) at various phosphoryl contents (0, 0.5, and 11 sites) is visualized by autoradiography and Coomassie staining. D) ATP limitation experiment for wt-ASF/SF2 at varying quench times. A complex of SRPK1 (1 μM) and wt-ASF (250 nM) was mixed with varying amounts of [^{32}P]ATP (0.2-500 μM) for 5 and 15 minutes before the reaction was quenched with SDS loading buffer and the total phosphoryl contents are plotted as a function of ATP.

N/C ratio reached ~0.8 at full phosphorylation, close to the equal splitting of phosphates predicted by the original design of the digestion site at K214. The N/C ratio at lowest phosphoryl content suggested roughly 20-fold preference for initiation of phosphorylation at C-terminal end of RS1 domain.

In order to discount the possibility that the digestion was overdone and the N-terminal fragment had degraded for the low ATP concentrations, Coomassie staining was performed alongside autoradiography for ASF(5R1K) at different phosphoryl content (Fig. 3.6C). The phosphoryl content was confirmed by autoradiography, and the N-terminal fragment was present in similar quantity for 0.5 and 11 site phosphorylated ASF, suggesting that no nonspecific digestion or degradation has occurred.

g.   Kinetic Analysis of Regiospecific Phosphorylation

ATP limitation is an endpoint assay, in which the SRPK1 enzyme has reached equilibrium in the phosphorylation of ASF/SF2. Since many enzymatic reactions are reversible, it is possible that SRPK1 may scramble its placement of phosphates toward the C-terminus of RS1 domain over time because it is thermodynamically favored. An alternative way to regulate phosphoryl content would be to stop the reaction at different time points in the progress curve, and this kinetic analysis does not allow any scrambling to occur. Using a pulse-chase experiment, the reaction was initiated with [^{32}P]ATP, and cold ATP (chase) was added at early or late time points so that only early phosphorylation with radioactive ATP would be recorded on autoradiography. The low ATP concentration was used to capture the extremely fast initial phosphorylation event. An impurity band was noticeably phosphorylated and appeared beneath the N-terminal fragment, possibly a digestion artifact caused by presence of large amounts of ATP (Fig. 3.7). The endpoint control group had N/C ratio of 0.7, roughly the same as full-phosphorylation sample in the ATP limitation experiment. At lowest phosphoryl content the dominant majority of ^{32}P incorporation occurred at the C-terminal fragment. The N/C ratio increased gradually with phosphoryl content, indicating that C->N directional phosphorylation is also preferred kinetically. The N/C ratio at lowest phosphoryl content suggested roughly 11-fold preference for initiation of phosphorylation at C-terminal end of RS1 domain. Overall, the data suggested that C->N phosphorylation is favored kinetically and thermodynamically.
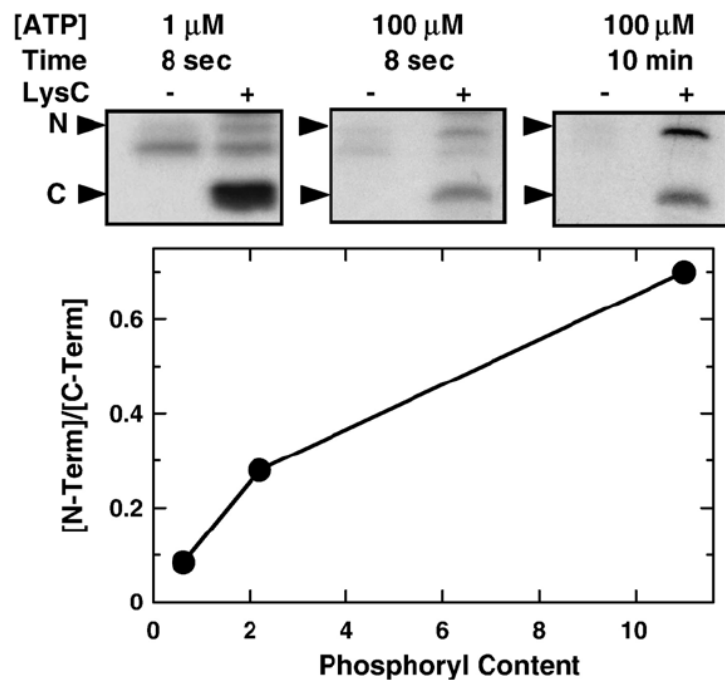
Figure. 3.7 Kinetic Analysis of Regiospecific Phosphorylation. ASF(5R1K) (250 nM) and SRPK1 (1 µM) were pre-equilibrated before the reaction was initiated with 1 or 100 µM [32P]ATP (pulse). Further 32P incorporation was stopped at 8 seconds or 10 minutes by the addition of 10 mM cold ATP (chase). The reaction was then treated immediately with LysC (5 ng/µL) for four hours at 37 ℃, and the products were analyzed by autoradiography. The ratio of ^{32}P in the N-terminal fragment to ^{32}P in the C-terminal fragment is plotted against phosphoryl content of undigested ASF(5R1K).

h.   Digestion Site Location Has No Impact on Regiospecific Preference

Given the extensive modification to the ASF/SF2 protein necessary for creating a digestion mutant, one last control experiment was implemented to confirm that the digestion site location has no impact on C->N regiospecific preference. ASF(R210K) and ASF(R218K) were generated to confirm C->N directionality. As in the case of ASF(5R1K), both mutants were phosphorylated to the same extent as wt-ASF (12 sites) based on MALDI-TOF measurements and bound to SRPK1 with high affinity based on competition experiments. Using the ATP limitation protocol coupled with LysC digestion, the incorporation of ^{32}P into the full-length proteins and the digested fragments of ASF(R210K) and ASF(R218K) were monitored by autoradiography. Significant gel shift was observed for the N- and C-terminal fragments and larger gel band zones were excised for quantification purpose. Based on the position of the digestion site, N/C ratio of 0.33 and 1.4 were expected, and experimental value of 0.45 and 1.2 (Fig. 3.8) for ASF(R210K) and ASF(R218K), respectively, were totally consistent. For both substrates, the dominating majority of the 32P incorporation occurred at the C-terminal fragment for the lowest phosphoryl content. The N/C ratio gradually increased, consistent with C->N directional phosphorylation. The overall data suggested that digestion site has no impact on C->N regiospecific preference exhibited by SRPK1. At the lowest phosphoryl content, the C-terminal fragment is favored over the N-terminal fragment by factors of 50 and 12, respectively. Finally, a stoichiometric ATP limitation experiment was performed with SRPK1 and ASF(R210K) (500 nM) and the results were similar to Fig.3.8A, suggesting that C-terminal preference is not due to multiple SRPK1 binding to the same ASF molecule.

i.   Digestion Mutants Establish an Intiation Box for SRPK1 Phosphorylation

Although SRPK1 does not require a strict initiation site on ASF/SF2, it exhibits strong preference for C->N directional phosphorylation, and in the wt construct there may exist a preferred start residue or region at the C-terminus of RS1 domain, to be investigated in depth with the recently developed phosphomapping technique. The digestion site was moved sequentially from position 229 outside RS1 to more N-terminal residues within RS1 (Fig. 3.9A), using the ASF(5R1K) [renamed c-ASF(214)] digestion construct as the template. ATP limitation and LysC digestion experiment on mutant
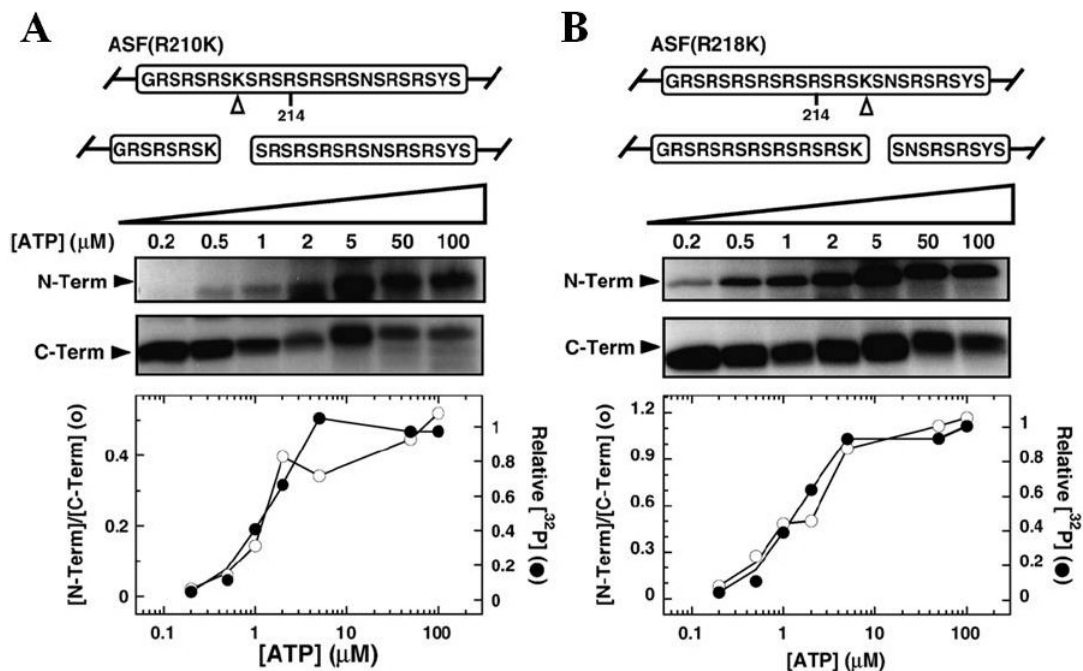
Figure 3.8 Phosphorylation mapping of ASF(R210K) A) and ASF(R218K) B). Substrates (250 nM) and SRPK1 (1 µM) were pre-equilibrated before a series of ATP concentrations (0.2-100 µM ATP) were used to initiate a 20-minute reaction. After phosphorylation, half of each reaction was treated with LysC (5 ng/µL) at 37 ºC for either 1 hour (0.2, 0.5, 1, and 2 µM ATP) or 4 hours (5, 50, and 100 µM ATP) and counted using autoradiography. [32]P incorporation N/C ratio (○) and normalized phosphoryl content of undigested substrate (●) are plotted against ATP concentrations.
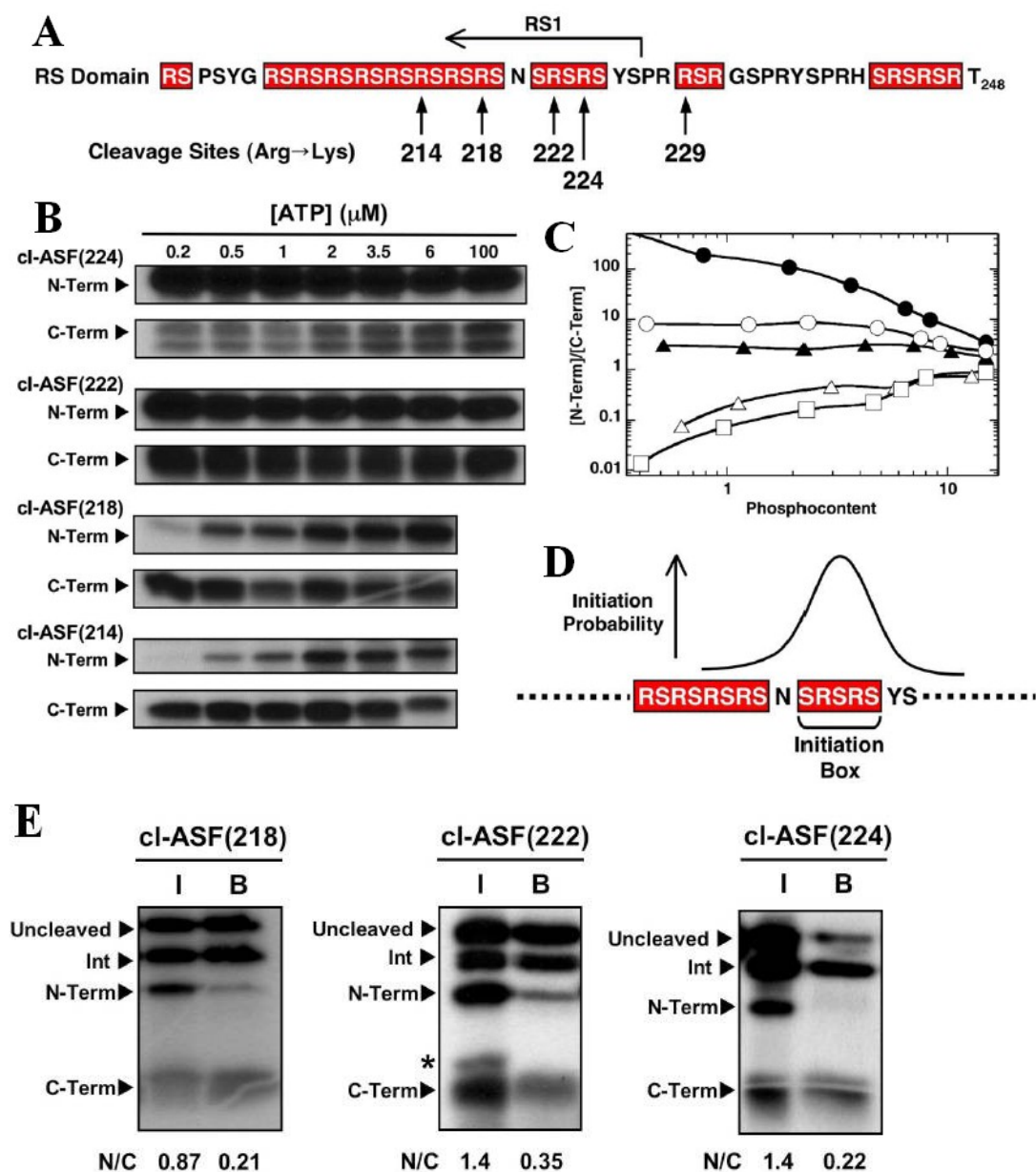
Figure 3.9 Digestion Site Position and Phosphorylation Initiation. A) Digestion sites. Alternate digestion sites in the RS domain are made within the ASF/SF2 digestion construct. B) ATP limitation studies. Complexes of SRPK1 (1 µM) and digestion mutants (250 nM) were reacted with varying amounts of [32P]ATP before the N- and C-terminal fragments were resolved by SDS-PAGE. Data for cl-ASF(229) can be found in Fig. 3.11C. C) N/C plot. The ratio of ^{32}P in the N-terminal fragment to ^{32}P in the C-terminal fragment is plotted against phosphoryl content of cl-ASF(229) (●), cl-ASF(224) (○), cl-ASF(222) (▲), cl-ASF(218) (Δ), and cl-ASF(214) (□). D) Initiation box. N-terminal and C-terminal phosphorylation preferences at low ATP are used to establish the probably start region for SRPK1. E) Pull-down assays for cl-ASF(218), cl-ASF(222), and cl-ASF(224). SRPK1, [^{32}P]-ATP, and substrates reacted for 20 minutes before LysC treatment. The reaction mixtures were then bound to Nickel resin and washed. The samples before (I) and after (B) resin binding are displayed in the SDS-PAGE autoradiograms. Asterisks mark minor impurity bands.

constructs were performed to establish their phosphorylation preferences. The N/C ratio at lowest phosphoryl content was especially important in identifying preferred site of phosphorylation initiation. While cl-ASF(229) showed strong preference for N-terminal initiation by more than 100-fold (Fig. 3.11C), cl-ASF(224) showed a much lower preference of about 8-fold at low ATP concentration (Fig. 3.9 B & C). In contrast, cl-ASF(214) and cl-ASF(218) showed strong preferences for initial C-terminal phosphorylation. The transition occurred at residue R222, where the cl-ASF(222) mutant demonstrated only a modest 2-fold N-terminal preference at low ATP (Fig. 3.9 B & C). The almost even distribution suggested that there is not a specific initiation serine but an initiation box where each serine is assigned a certain probability of being the intiation serine. Overall, the data suggested that S221, S223, and S225 belong in this initiation box. In terms of digestion mutant validation, in all cases, the N/C ratios converged on expected values at high ATP concentrations based on the position of the digestion sites in the RS domain. The final experimental N/C ratios for cl-ASF(229), cl-ASF(224), cl-ASF(222), cl-ASF(218), and cl-ASF(214) are 3.4, 2.4, 1.7, 0.9, and 0.7, respectively, comparing favorably with theoretical ratios of 3, 1.8, 1.4, 1, and 0.6 based on RS/SR content in the RS domain. All mutants were characterized kinetically and exhibited biphasic kinetics with rate constants similar to wt ASF construct (data not shown).

j.   Characterizing Potential RS2 Phosphorylation by SRPK1

While SRPK1 phosphorylates ASF/SF2 on 12-15 sites based on MALDI-TOF spectra and there are insufficient RS/SR pairs within RS1 (ending at S227) to account for all 15 sites, this raises the possibility that SRPK1 could phosphorylate ASF/SF2 outside RS1 segment. The investigation into potential phosphorylation of the RS2 domain started with deletion ASF/SF2 constructs. The full-length wt-ASF/SF2 and a deletion construct lacking RS2 [ASF(1-226)] were incubated with SRPK1 and ATP for 3 hours and analyzed by MALDI-TOF (Fig. 3.10A), and a difference of 3 sites suggested that SRPK1 could possibly phosphorylate serines beyond S227. However, the RS2 deletion could result in instability of RS1 domain, so further deletion constructs were created to explore the relationship between the RS/SR
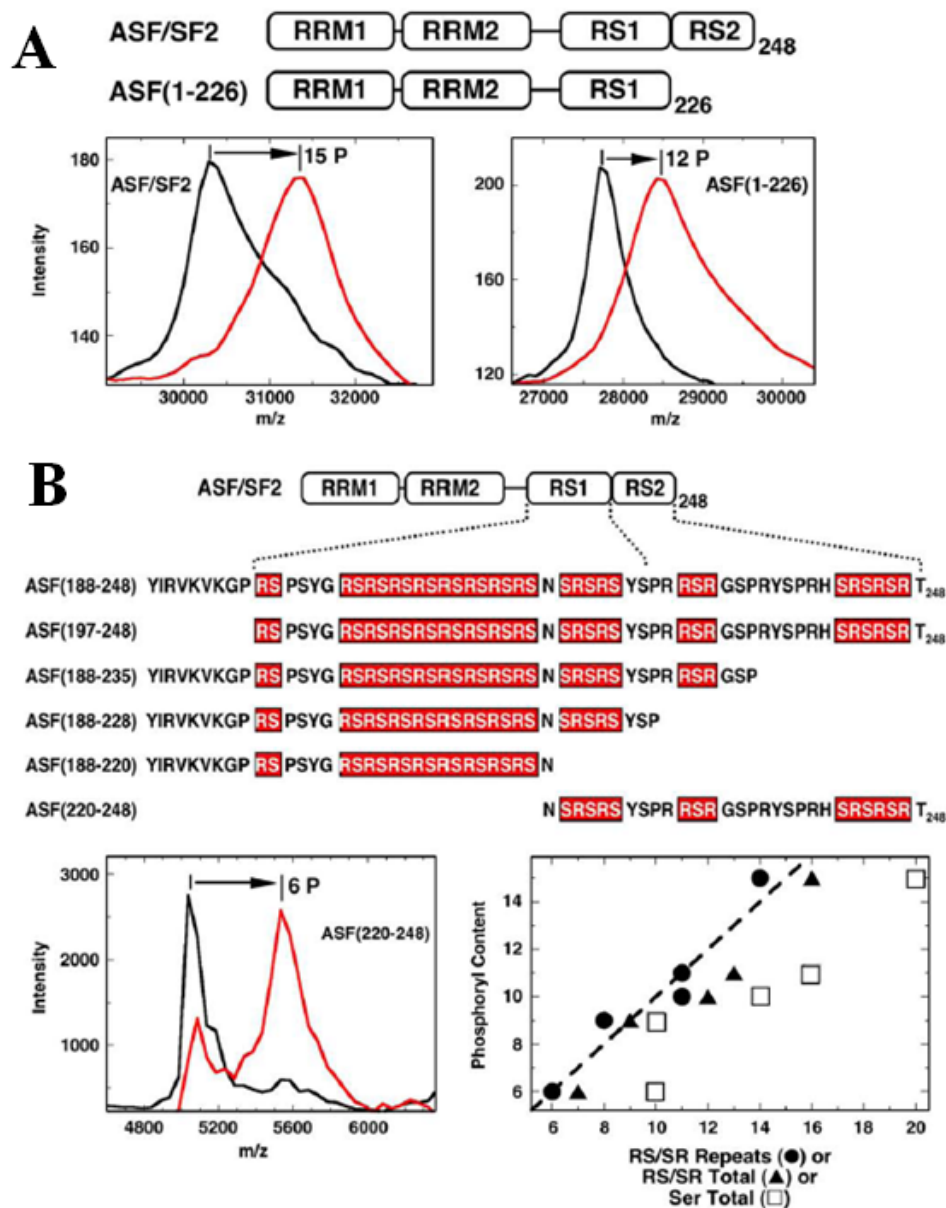
Figure 3.10 Phosphorylation of the RS Domain Dependent on Arg-Ser Content. A) RS2 deletion in ASF/SF2. The MALDI-TOF spectra of wt-ASF/SF2 and an RS2-deleted substrate [ASF(1-226)] were collected in the presence of SRPK1 and in the absence (black) and in the presence of ATP (red). B) Phosphoryl contents of RS domain constructs. MALDI-TOF spectra for all constructs were collected in the presence of SRPK1 and in the absence and in the presence of ATP. The total phosphoryl contents for each construct are plotted as a function of RS/SR repeats (●), total RS/SR content (▲), and total number of serines (□). RS/SR repeats are a series of contiguous Arg-Ser or Ser-Arg, whereas total RS/SR content reflects all serines directly flanked on the N- or C-terminal side by an arginine. Dotted line represents the expected correlation (slope =1; y-intercept = 0) between observed ordinate and abscissa.
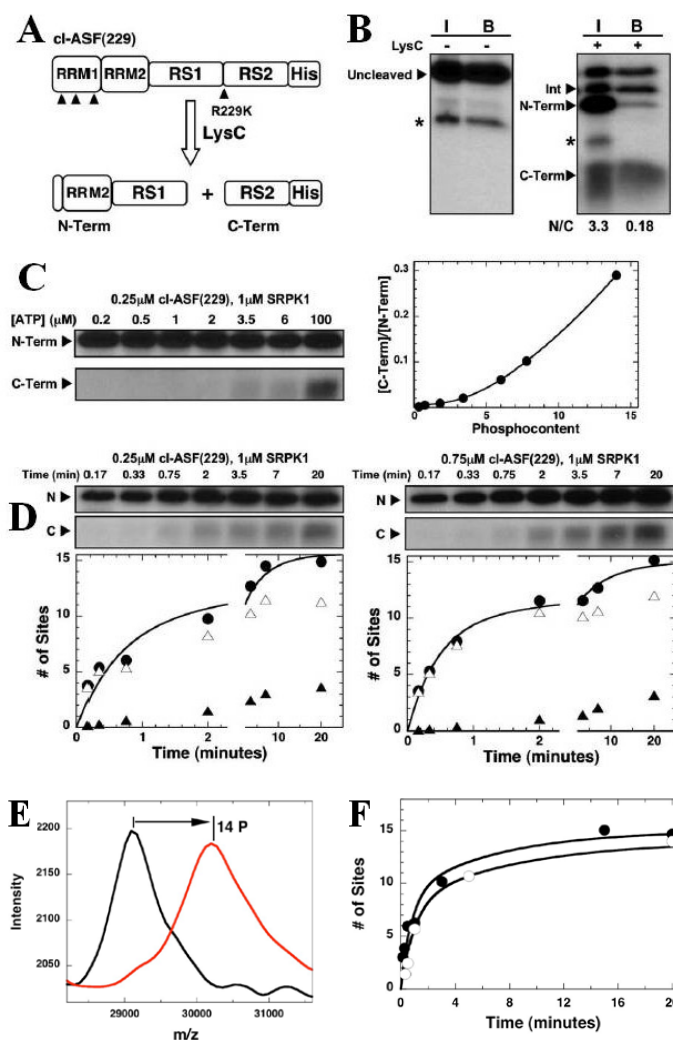
Figure 3.11 Serines in RS2 Phosphorylated by SRPK1. A) Digestion construct. All lysines in cl-ASF(229) are designated with continuous arrows. LysC digestion produced two major N-terminal and C-terminal fragments. B) Pull-down assays. SRPK1, [^{32}P]ATP, and cl-ASF(229) reacted for 20 minutes before LysC treatment, resin binding and wash. The samples before (I) and after (B) resin binding and washing are displayed in the SDS-PAGE autoradiogram. N/C ratio is the ratio of ^{32}P counts in the N-terminal and C-terminal bands. Asterisk marks minor impurity bands. C) ATP limitation experiment. A complex of SRPK1 and cl-ASF(229) was mixed with varying amounts of [^{32}P]ATP for 20 minutes before LysC treatment. The C/N ratio (ratio of C-terminal to N-terminal bands) is plotted as a function of total phosphoryl content. D) Pulse-chase experiments. Complexes of SRPK1 and cl-ASF(229) in ratios of 4:1 or 1.3:1 were reacted with 100 μM [^{32}P]ATP (pulse) for varying times before 4 mM cold ATP (chase) was added to stop the reaction. The samples are then treated with LysC to generate N-terminal (Δ) and C-terminal fragments (▲). The phosphorylation of cl-ASF(229) prior to LysC digestion is plotted as a function of time (●) for both enzyme:substrate stoichiometries. The time dependent data are fitted to a double-exponential function with rate constants of 1.6 and 0.18 min^{-1} and amplitudes of 9 and 6 sites for 250 nM cl-ASF(229), and rate constants of 2.1 and 0.13 min^{-1} and amplitudes of 10 and 5 sites for 750 nM cl-ASF(229), respectively. E) MALDI-TOF spectra of cl-ASF(229). cl-ASF(229) was mixed with SRPK1 in the absence (black) and in the presence of ATP. F) Complexes of SRPK1 and wt-ASF (●) or cl-ASF(229) were reacted with 100 μM [^{32}P]ATP, and the total phosphoryl contents are plotted as a function of time.

content and phosphoryl content. All constructs were made with N-terminal Histidine tags, except for ASF(197-248) and ASF(220-248), which contained C-terminal Histidine tags. Placement of the histidine tag on C-terminal or N-terminal end of the wt-ASF/SF2 construct had no impact on total phosphoryl content (Fig. 3.1D and Fig. 3.10A, respectively). The data showed a direct linear correlation between the observed phosphoryl and Arg-Ser contents of the peptides (Fig. 3.10B). Also, a construct with only three serines from RS1 [ASF(220-248)] was phosphorylated at a total of 6 sites, suggesting that SRPK1 may have the capacity to modify serines in RS2. Overall, the data suggested that while SRPK1 largely phosphorylates serines in N-terminus of RS domain, the C-terminus may be targeted as well.

To gather additional evidence, cl-ASF(229) was created with the lysine digestion site at RS1/RS2 border (Fig. 3.11A). Upon LysC digestion the RS1 and RS2 domains could be easily separated by SDS-PAGE gel. cl-ASF(229) was phosphorylated by SRPK1 at the same phosphoryl content and reaction rate as wt-ASF. LysC digestion yielded three major species: an intermediate band (24 kDa), an N-terminal band (19 kDa), and a diffuse C-terminal band (5 kDa), consistent with site of digestion (Fig. 3.11B). Nickel resin pull-down experiments confirmed the identity of the fragments, where the N-terminal fragment was the only one lacking Histidine tag and lost in the wash. An impurity was present in both undigested and digested sample, but it was <5%, clearly separate and distinguishable from the rest of the bands, and ignored in the analysis. The ratio of N-terminal fragment to C-terminal fragment was 3.3, consistent with phosphorylation of about 3-4 sites in RS2 and 11-12 sites in RS1.

The next question to address would be the order of phosphorylation. Phosphorylation of RS2 was monitored as a function of reaction progress. In ATP limitation experiments, it was discovered that the C-terminal fragment was phosphorylated only at higher ATP concentrations (Fig. 3.11C). Monitoring the corresponding phosphoryl content, it was revealed that no significant RS2 modification occurs until after eight serines in RS1 had been modified. At the highest concentration the C/N ratio of 0.28 was consistent with the modification of about 12 serines in RS1 and 3 serines in RS2, validating the digestion profile.

Although SRPK1 phosphorylates RS1 prior to RS2 in ATP limitation experiments, the preference could be the result of phosphorylation-dependent changes in K_m for the nucleotide, with

serines in RS1 having lower K_m values. To address this possibility, the reaction was initiated with a constantly high amount of $[^{32}P]$ATP (100 μM) and then stopped with 4 mM cold ATP (chase) in a pulse-chase experiment prior to LysC treatment. This experiment was performed at 4:1 and 1.3:1 ratio of SRPK1 to cl-ASF(229) to ensure that the observed RS2 phosphorylation is not the result of reaction priming at more than one site in higher-order complexes. The late appearance of phosphorylated C-terminal fragment suggested that SRPK1 does not bind simultaneously at two positions in the RS domain (Fig. 3.11D). Kinetic analysis of the undigested cl-ASF(229) revealed a fast and slow phase similar to wt-ASF, and within the fast phase of the reaction(<2 minutes) the dominating majority of the 32P incorporation occurred in the N-terminal fragment. Overall, both thermodynamic and kinetic data showed that while RS1 phosphorylation is strongly favored, SRPK1 can also modify serines in RS2 in a longer time scale.

k.   Clk/Sty Phosphorylates ASF/SF2 Non-directionally

Whereas SRPK1 phosphorylates up to 15 sites in the RS domain of ASF/SF2, Clk/Sty phosphorylates the whole RS domain for a total of 20 sites. The phosphomapping technique developed in the previous sections was applied to the question of directionality for Clk/Sty. cl-ASF(224) was designed to roughly divide the RS domain into RS1 and RS2. Extensive control experiments described above for SRPK1 and cl-ASF(214) were performed to ensure the validity of this new digestion construct. MALDI-TOF spectra confirmed 20 site phosphorylation by Clk/Sty (data not shown), and competition studies showed high binding affinity of the enzyme:substrate complex similar to wt-ASF (Fig. 3.12C).

Given the evidence that cl-ASF(224) behaves the same as wt-ASF, LysC digestion was performed and the digestion profile showed a strange migration pattern of the N- and C-terminal fragments. Extensive gel-shifts were observed when comparing low and high amounts of ATP (0.2 and 100 μM). In order to align the N- and C-terminal fragments for all possible phosphoryl content, the pulse-chase experiment was applied in which 10 mM cold ATP was added to quench the sample and to ensure a uniform digestion profile. Owing to inefficient proteolysis, it took two days instead of four hours for sufficient LysC digestion. This new digestion profile showed complete digestion of the original
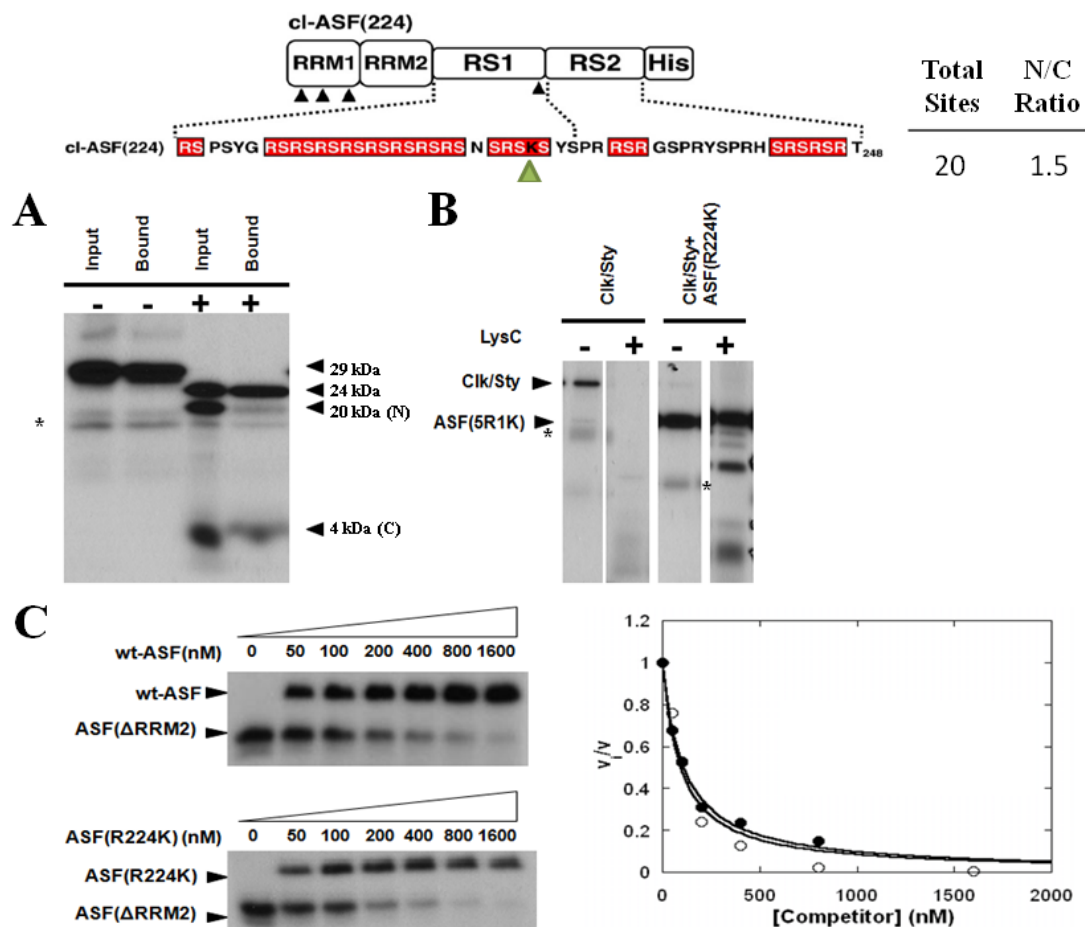
Figure 3.12 Clk/Sty Phosphorylation of cl-ASF(224) and Digestion Profile. Green arrow indicates digestion site on cl-ASF(224) RS domain. A) and B) Digestion profile and Nickel-resin pull-down assay. Clk/Sty and [^{32}P]-ATP react in the presence or absence of cl-ASF(224) for 20 minutes before LysC treatment. The reaction mixtures are then bound to Nickel resin and washed. The samples before (input) and after (bound) resin binding are displayed in the SDS-PAGE autoradiograms. Asterisks mark minor impurity bands. C) Binding affinity of cl-ASF(224), also known as ASF(R224K). The phosphorylation of ASF(ΔRRM2) was monitored as a function of increasing wt-ASF/SF2 (●) or cl-ASF(224) (○) using SDS-PAGE autoradiography. Decreases in the relative initial velocity for ASF(ΔRRM2) are plotted as a function of competitor concentration and fitted to Eq. 2.

cl-ASF(224) at 29 kDa but extensive build-up of the 24 kDa intermediate band, consistent with previous observation that at high ATP concentration the mid-RS1 domain digestion site is somehow inaccessible (Fig. 3.12A). The identities of the N- and C-terminal fragments were confirmed through a Nickel-resin pull-down assay, in which the N-terminal un-tagged fragment was not pulled down after washes. In a different control experiment, auto-phosphorylated Clk/Sty, much higher in intensity than auto-phosphorylated SRPK1, was digested alone and its profile compared against Clk/Sty:cl-ASF(224) mixture. Some low-molecular weight smear was observed in the general position of the C-terminal fragment, but the intensity was weak and should not significantly impact the N/C ratio calculated from the relative intensity of the N- and C-terminal fragments. In addition, the intermediate fragment and the N-terminal fragment migrated very close to each other, and extreme care was taken to cut out the two bands as cleanly as possible. Fully phosphorylated cl-ASF(224) showed an N/C ratio of 1.5, in line with the theoretical phosphate distribution of 12 phosphates in the N-terminal side of K224 and 8 phosphates in the C-terminal direction.   The minor impurity bands phosphorylated by Clk/Sty were minimal (<5%) in intensity compared against the main bands.

Having validated the cl-ASF(224) digestion and phosphorylation profile, an ATP limitation experiment was conducted to investigate possible regiospecific preferences (Fig. 3.13A). Some of the phosphorylation sites on Clk/Sty showed higher than expected K_m value, since at least 15 μM ATP was necessary to fully phosphorylate 250 nM ASF, while 5 to 6 μM ATP would be the minimal required amount if Clk/Sty could catalyze ATP efficiently. Alternatively, the auto-phosphorylation of Clk/Sty might contribute to the greater requirement for ATP for full phosphorylation of ASF. The N/C ratio started at roughly 0.8 at the lowest ATP concentration, increased to as high as 2.4 with increasing ATP concentration, and then fell again at full-phosphorylation to roughly 1.4. If a standard line was drawn across 1.5 as uniform phosphorylation of the N- and C-terminal fragments based on RS/SR content, then all the deviations did not exceed two-fold from standard, suggesting that Clk/Sty phosphorylates cl-ASF(224) without preference for RS1 vs. RS2. The ATP limitation assay with cl-ASF(218) showed similar nondiscriminating results, suggesting that the digestion site has no impact on Clk/Sty phosphorylation mechanism. Kinetic analysis was conducted with pulse-chase experiment, in which cold
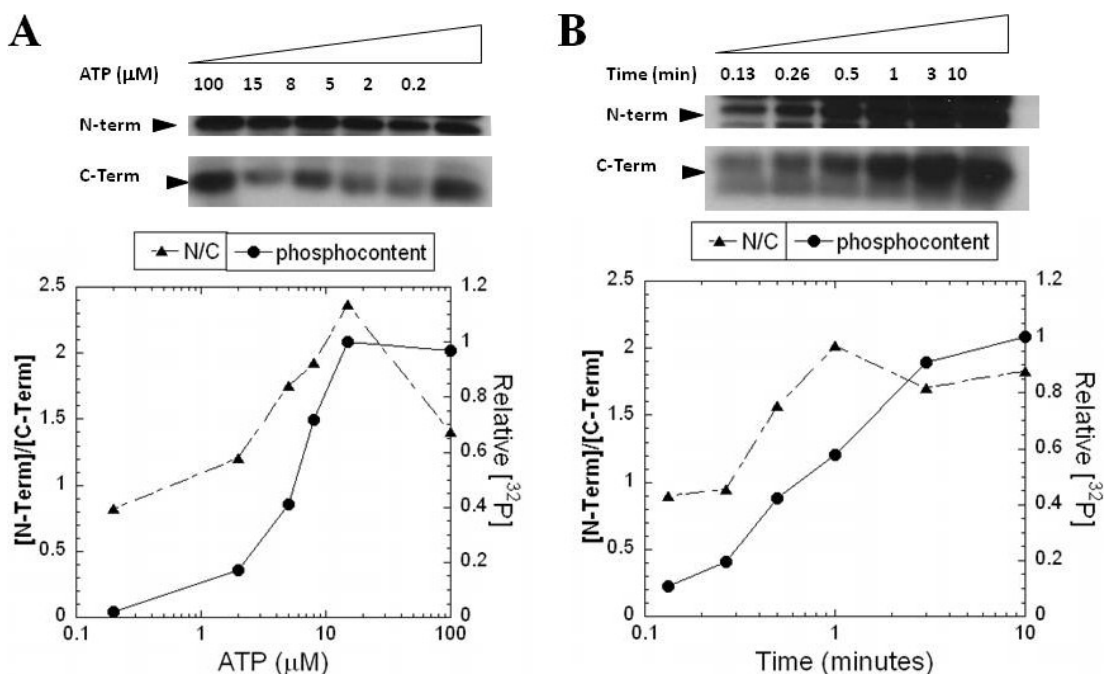
Figure 3.13 Equilibrium and Kinetic Analysis of Clk/Sty Non-Directional Phosphorylation. A) Equilibrium analysis of Clk/Sty Phosphorylation. Cl-ASF(224) (250 nM) and Clk/Sty (1 μM) were pre-equilibrated before a series of ATP concentrations (0.2, 2, 5, 8, 15, and 100 μM ATP) were used to initiate a 20-minute reaction that was quenched by 10 mM cold ATP. Results are displayed by SDS-PAGE autoradiography. Phosphoryl content is plotted as a function of increasing ATP concentration. After the phosphorylation reaction, half of each reaction volume was treated with LysC (5 ng/μL) at 37 ºC for two days. N- and C-terminal fragments from the cl-ASF(224) RS domain are shown, and the ratio of ^{32}P in the N-terminal fragment to ^{32}P in the C-terminal fragment is plotted against ATP concentrations. B) Kinetic analysis of Clk/Sty phosphorylation. cl-ASF(224) (250 nM) and Clk/Sty (1 μM) were pre-equilibrated before the reaction was initiated with 100 μM [32P]ATP. Further ^{32}P incorporation was stopped at various times with 10 mM cold ATP. The reaction was then treated immediately with LysC (5 ng/μL) for two days at 37 ºC, and the products are analyzed by autoradiography. The ratio of ^{32}P in the N-terminal fragment to ^{32}P in the C-terminal fragment and the phosphoryl content are plotted against various time points.

ATP was added to quench the reaction at different time points (Fig. 3.13B). The phosphoryl content reached maximum at around 10 minutes, and the N/C ratio increased from 0.9 at the lowest ATP concentration to 2 at the 1-minute mark. The N/C ratio then dropped slightly and reached 1.8 at the end of the reaction. The general shape was reproduced kinetically and thermodynamically, and Clk/Sty phosphorylates ASF/SF2 without significant preference for either N-terminal or C-terminal initiation. However, the large size of the RS domain involved and the deviation from flat, constant N/C ratio suggested that there may be some order to pockets of phosphorylation by Clk/Sty, defined not by RS1:RS2 but some other local RS domain feature.

B. Discussion

The phosphorylation state of the RS domain in ASF/SF2 is important for several processes, including localization in nuclear speckles, protein-protein and protein-RNA interactions in the spliceosome, and export of mature mRNA [22, 26, 39, 104]. Such findings underscore the importance of investigating ASF/SF2 phosphorylation as a means of understanding splicing regulation. While peptide studies demonstrate that SRPK1, a principal regulator of SR protein function, will phosphorylate serine (but not threonine) when flanked by arginine [53], the enzyme specifically phosphorylates only the N-terminal portion of this domain (designated as RS1 segment) [48], despite additional RS/SR repeats inside and outside the RS domain. SRPK1 phosphorylates ASF/SF2 processively without dissociation of intermediate phosphorylated ASF/SF2, possible through initial high-affinity SRPK1:ASF/SF2 complex [48, 51]. In contrast, Clk/Sty, another splicing kinase, phosphorylates the whole RS domain [48] and, at most, one or two residues outside the RS domain (Velazquez-Dones, pers. comm.). Although the restrictive phosphorylation of the RS1segment is sufficient for the localization of ASF/SF2 in nuclear speckles, phosphorylation by Clk/Sty at other sites in the RS domain (designated as RS2) disperses ASF/SF2 from speckles [39]. Thus, specific segments of the RS domains in SR proteins, despite their charge periodicity, elicit unique cellular functions when phosphorlated.

a. **Refinement of the Phosphorylation Kinetics for ASF/SF2**

In single-turnover experiments, SRPK1 phosphorylates the RS1 segment in ASF/SF2 rapidly, within the first minute of the reaction. While previous studies showed that excess SRPK1 saturates all ASF/SF2 molecules and processes them in one fast single turnover, when SRPK1 concentration is low and not all ASF/SF2 is bound by SRPK1, a biphasic progress curve is observed in which the SRPK1 processes the pre-bound part of the ASF/SF2 population in a fast phase and then dissociates and processes the remaining ASF/SF2 in a slower phase. This kinetic observation which underpins the original assay is used to measure the dissociation constant of the complex (K_d). However, I have observed that, in some cases, the single-exponential fit is poor even for saturating amounts of enzyme (Fig. 3.1A), suggesting that there is a slower phosphorylation phase that was ignored in previous 3-minute or 10-minute studies. By extending the time course to 15 or 20 minutes and repeating the single-turnover experiment many times, the data suggested that 12 sites are phosphorylated in the fast phase with a rate constant of 3.7 min^{-1}, and 2-3 sites are phosphorylated in the slow phase with rate constant of 0.13 min^{-1} (Fig. 3.1B). Mutant ASF/SF2 constructs also demonstrate the same biphasic progress curves, further confirming the existence of a secondary phosphorylation phase. Finally, more recent MALDI-TOF spectrometric studies revealed that SRPK1 phosphorylates ASF/SF2 for 12-15 sites (Fig. 3.1C & D), consistent with a slower phase that is not easily detected without extending the duration of phosphorylation. The Gourisankar Ghosh lab reported recently that if left overnight SRPK1 would phosphorylate as many sites as Clk/Sty (Ghosh, pers. comm.). While it is difficult to say what time frame is physiologically relevant, we feel that 1 to 15 minutes may be a reasonable time frame for cellular trafficking events, whereas overnight is too long to be relevant. Also, phosphorylated ASF/SF2 is unstable and decays over time, and *in vitro* kinetic studies operate under a time frame best characterized by 15-site phosphorylation.

While previous studies suggested that SRPK1 is fully processive in phosphorylating ASF/SF2, discovery of the slow phase and better analysis of the kinetic data raise the need to reassess the extent of processivity. Data from an improved start-trap experiment revealed that SRPK1 processively and rapidly phosphorylates about 8 sites on ASF/SF2 before dissociating. Whether SRPK1 can rebind and then

initiate an additional processive reaction after this initial one is not clear, as pre-phosphorylated ASF/SF2 (8-10 sites) is not effectively inhibited by kinase-dead SRPK1 and the extent of processivity cannot be assessed (data not shown). The second, slower phase at the end of reaction is also not easily distinguished from the trap-start reaction, and no conclusion can be drawn for this phase. These results suggest that as phosphates are added to ASF/SF2, progressively weakened interactions coupled with a reduced rate for catalysis may promote dissociation over forward catalysis. The slower rate of the second phase could reflect the difficulty of rebinding a partially phosphorylated molecule to complete the reaction. Overall, the kinetic data support a model where SRPK1 is strongly driven by processive attachment of phosphates in the initial fast phase of the reaction and then favors dissociation over forward catalysis in later stages.

b.   SRPK1 Strongly Recognizes C-terminal Sequences in RS1

SR protein phosphorylation has been very difficult to study owing to the repetitive nature of the RS domain and inability to identify unique proteolytic fragments for analyses. This problem was addressed using an engineered protease footprinting technique that introduces unique LysC digestion sites into the RS domain [48]. When coupled to MALDI-TOF spectrometry, a determination of the minimal number of phosphates on a given peptide can be made. However, quantifying phosphopeptide concentrations is still difficult and remains a serious limitation for mechanistic studies. In this new investigation, the LysC fragment lengths covering RS1 have been manipulated so that the relative concentrations of various regions of RS1 can be easily measured with sensitive SDS-PAGE autoradiography (Fig. 3.4). Using both equilibrium (ATP limitation) and kinetic (pulse-chase) methods, it was shown that SRPK1 prefers to phosphorylate C-terminal serines prior to N-terminal serines in RS1 (Fig. 3.6 & 3.7). This trend was found to be independent of the position of the lysine digestion site (Fig. 3.8), suggesting that this is a general phenomenon within RS1. To achieve this preference, it is likely that the active site of SRPK1 strongly binds the C-terminal sequences of RS1. The ATP dependence on the fragmentation ratios for the digestion mutants can be used to assess the inherent specificity of SRPK1 for either the N-terminus or the C-terminus of RS1. The N/C ratio at low ATP concentration informs us on

the intrinsic preference for N-terminal over C-terminal modifications. For all three mutants, the C-terminal fragment is preferred by at least a factor of 12 at the lowest ATP concentration tested (0.2 μM), where approximately one site is modified (Fig. 3.6 & 3.8). These findings suggest that the active site of SRPK1 initially grabs onto the C-terminus of RS1 to start the reaction.

c.   SRPK1 Catalyzes Directional Phosphorylation of ASF/SF2

Using a combination of limited phosphorylation, LysC proteolysis, and start-trap experiments, it was found that SRPK1 docks initially to the C-terminus of the RS1 segment of ASF/SF2 and then processively phosphorylates up to 8 serines before dissociating. These experimental observations support an unusually ordered mechanism where SRPK1 catalyzes directional C-terminal to N-terminal phosphorylation within RS1. The best known example of ordered enzymatic catalysis is DNA polymerase, which catalyzes base pair addition directionally and sequentially. The current digestion phosphomapping technique is insufficient to track the placement of every single phosphate, and it is impossible at this point to determine whether SRPK1 phosphorylates sequential residues in the RS1 domain of ASF/SF2. Ordered polyphosphorylation also has some precedence within the protein kinase family. Glycogen synthase kinase-3 (GSK-3) phosphorylates glycogen synthase at several serines in sequence [105, 106]. In this case, however, a priming phosphorylation step, catalyzed by casein kinase II, is required for subsequent N-terminal phosphorylation following the consensus sequence, -S-X-X-X-S(P)- [107]. For GSK-3, replacement of one of the serines with alanine in the substrate abolishes the phosphorylatio nof all subsequent N-terminal serines. Thus, GSK-3 follows a compulsory ordered mechanism where a P+4 phosphoserine serves as a determinant for upstream serine phosphorylatino. Interestingly, it was recently found that GSK-3 also polyphosphorylates the RS domain of ASF/SF2, but, as in the case of its natural substrate glycogen synthase, the reaction requires a priming step [108]. SRPK1 differs from GSK-3 in that it lacks a mandatory priming step catalyzed by a heterologous kinase and possesses a flexible initiation site for phosphorylation. Removing blocks of serines or blocks of arginines did not abolish the phosphorylation of the remaining serines in RS1, suggesting that no single serine or arginine in RS1 is essential for full phosphorylation (Fig. 3.2 & 3.3).

d.   Multiple C-terminal Initiation Modes

SRPK1 adapts well to sequence changes in and near RS1 (Fig. 3.2 & 3.3). This could be due to a malleable enzyme that adjusts to mutations by simply translocating to further N-terminal sites. However, the data also suggest that SRPK1 may naturally incorporate multiple initiation modes near the C-terminus of RS1 into the wild-type substrate. While SRPK1 supports a directional phosphorylation pathway, the enzyme may not be constrained to initiation at the farthest C-terminal serine(s). In fact, if SRPK1 were to initiate at a single site at the C-terminal fringe, N/C ratios should be near zero for the digestion construct ASF(R210K) even at intermediate phosphoryl contents when the enzyme should not have reached the N-terminal fragment. Instead, there are noticeable levels of N-terminal phosphorylation at these stages of the reaction. Under conditions where ASF(R210K) is phosphorylated at a total of 4.7 sites at 1 μM ATP (Fig. 3.8A), the N/C ratio suggests that 4.1 serines are modified in the C-terminal fragment while 0.6 site is modified in the N-terminal fragment. This early incorporation supports the model that the enzyme-substrate complex may adopt several possible modes of initiation near the C-terminus so that some SRPK1 molecules are more advanced (further N-terminal) along the RS domain than others. Since this phenomenon is also observed under stoichiometric SRPK1:ASF(R210K) levels, advanced phosphorylation in the N-terminal fragment is not due to the simultaneous binding of multiple SRPK1 molecules to a single RS domain. Finally, this flexible initiation mode could serve a valuable function, since other SR proteins phosphorylated by SRPK1 have larger more complex RS domains. By avoiding a strict initiation serine and by incorporating multiple start modes, SRPK1 may broaden its substrate specificity to modify more diverse SR proteins. It is worth remarking that the tyrosine kinase Src catalyzes the processive phosphorylation of its substrate Cas without a compulsory initiation site [109], so there are other examples of relaxed initiation site selection among enzymes catalyzing polyphosphorylation.

e.   SRPK1 Starts Phosphorylation in the Initation Box

While SRPK1 has a flexible initiation site on ASF/SF2, it is highly likely that some serines should receive higher preference over others. The X-ray crystal structure does not define this initiation region clearly (Fig. 4.1), but with the new phosphomapping technique and an itinerant digestion site, the previous hypothesis was confirmed, and it was shown that SRPK1 does not initiate phosphorylation at a specific serine but rather selects a small group of serines at the C-terminus of RS1, a small region referred to as the initiation box (Fig. 3.9). Interestingly, mutation of all three serines in the initiation box to alanine does not halt or negatively affect the kinetics of phosphorylation at other series in the RS domain (Fig. 3.2). The data imply that while SRPK1 has a preference for initiation here, it readily adapts to changes in the initiation box and modifies other residues in RS1. Although the possibility that Arg-to-Lys mutation causes broadening of initiation preference cannot be dismissed, these findings further support the model that SRPK1 has flexibility in binding the RS domain and is not likely to form a unique enzyme-substrate complex. Instead, the initial unphosphorylated SRPK1-ASF/SF2 complex may exist in three major forms that differ by the position of the first serine in the active site (Ser221, Ser223, or Ser225). Additional studies are necessary to characterize local topological features within the ASF/SF2 RS domain that attracts SRPK1 to initiate in this region.

f.   SRPK1 Is a Residue-specific Protein Kinase with Regiotemporal Preferences

Earlier MALDI-TOF analyses of ASF/SF2 suggested that SRPK1 not only is specific for the RS domain but also limit itself to only serines in the RS1 segment [48]. However, the original data are based on the detection of phosphoserines in the N-terminal peptide fragments on a mass spectrometer, the absence of observable phosphoserines in the C-terminal region does not exclude a broader specificity for SRPK1. By monitoring the incorporation of ^{32}P into N-terminal and C-terminal fragments of the RS domain upon LysC proteolysis, it was found that SRPK1 not only is capable of phosphorylating peptides based on the C-terminal region of the RS domain (Fig. 3.10B) but also can modify RS2 in the full-length substrate in a time-dependent manner (Fig. 3.11). While SRPK1 rapidly phosphorylates 10-12 serines in RS1 in about 2 minutes, RS2 can be phosphorylated at another 3-4 sites within 15-20 minutes.

g.   Clk/Sty Shows No Preference for RS1 vs. RS2 Phosphorylation

Clk/Sty is another splicing kinase found in the cell nucleus, processively phosphorylating all serines within the RS domain of ASF/SF2 and regulating ASF/SF2 nuclear localization [39, 48]. The same phosphomapping technique involving limited phosphorylation and LysC proteolysis was applied to investigate the order of phosphorylation by Clk/Sty (Fig. 3.12), and it was shown that Clk/Sty phosphorylates the RS domain without noticeable thermodynamic or kinetic initiation preference for either RS1 or RS2 segment (Fig. 3.13).   The lack of an initiation region means a uni-directional phosphorylation mechanism is impossible. However, the N/C ratio is not constant throughout the assay, suggesting that there may be ordered phosphorylation within small segments of the RS domain. Given the large size of the RS domain and the limit of the assays, it is impossible to fully characterize the order of phosphorylation and designate the mechanism as totally random.   The recent X-ray structure of Clk/Sty revealed that helix αH blocks the putative docking groove also found in SRPK1 [57]. This could relax the substrate specificity of Clk/Sty and lead to apparent non-directional phosphorylation of the RS domain of ASF/SF2.   Further studies are needed to determine whether this helix abolishes potential ordered phosphorylation of ASF/SF2.

h.   Conclusion

Experimental observations made above help to establish a model for the phosphorylation mechanism by SRPK1 (Fig. 3.14). SRPK1prefers to initiate phosphorylation reaction in an initiation box at the C-terminus of RS1 (Ser221/223/225), but these residues are not strictly required for subsequent phosphorylation events. SRPK1 rapidly and processively phosphorylates about 8 serines in RS1 before dissociation.   This fast-phase phosphorylation is also C->N directional, such that RS2 is not phosphorylated until the dissociation event occurs and SRPK1 rebinds at either RS1 or RS2. The total phosphoryl content at the end of reaction is 15 sites based on recent MALDI-TOF data that accounts for the slow phase of the phosphorylation reaction, and only 3 sites are slowly phosphorylated in the RS2 segment. We think that this overall reaction is likely to be highly dynamic where the kinase may initiate at different serines within the initiation box and then 'back-track' to phosphorylate remaining serines that

are not modified in the first pass. Molecular contacts driving the kinetic preference for RS1 over RS2 (~20-fold) and directional phosphorylation would be elucidated in Chapter 4.

Chapter 3 is, in part, a reprint of the material as it appears in *The Journal of Molecular Biology*, 2008, 376(1), 55-68, Ma, C.T., Velazquez-Dones, A., Hagopian, J.C., Ghosh, G., Fu, X.D., Adams, J.A. The dissertation author was the primary researcher and author of this publication. Additional text of Chapter 3 is a reprint of the material as it appears in *The Journal of Molecular Biology*, 2009, 390(4), 618-34, Ma, C.T., Hagopian, J.C., Ghosh, G., Fu, X.D., Adams, J.A. The dissertation author was the primary researcher and author of this publication.
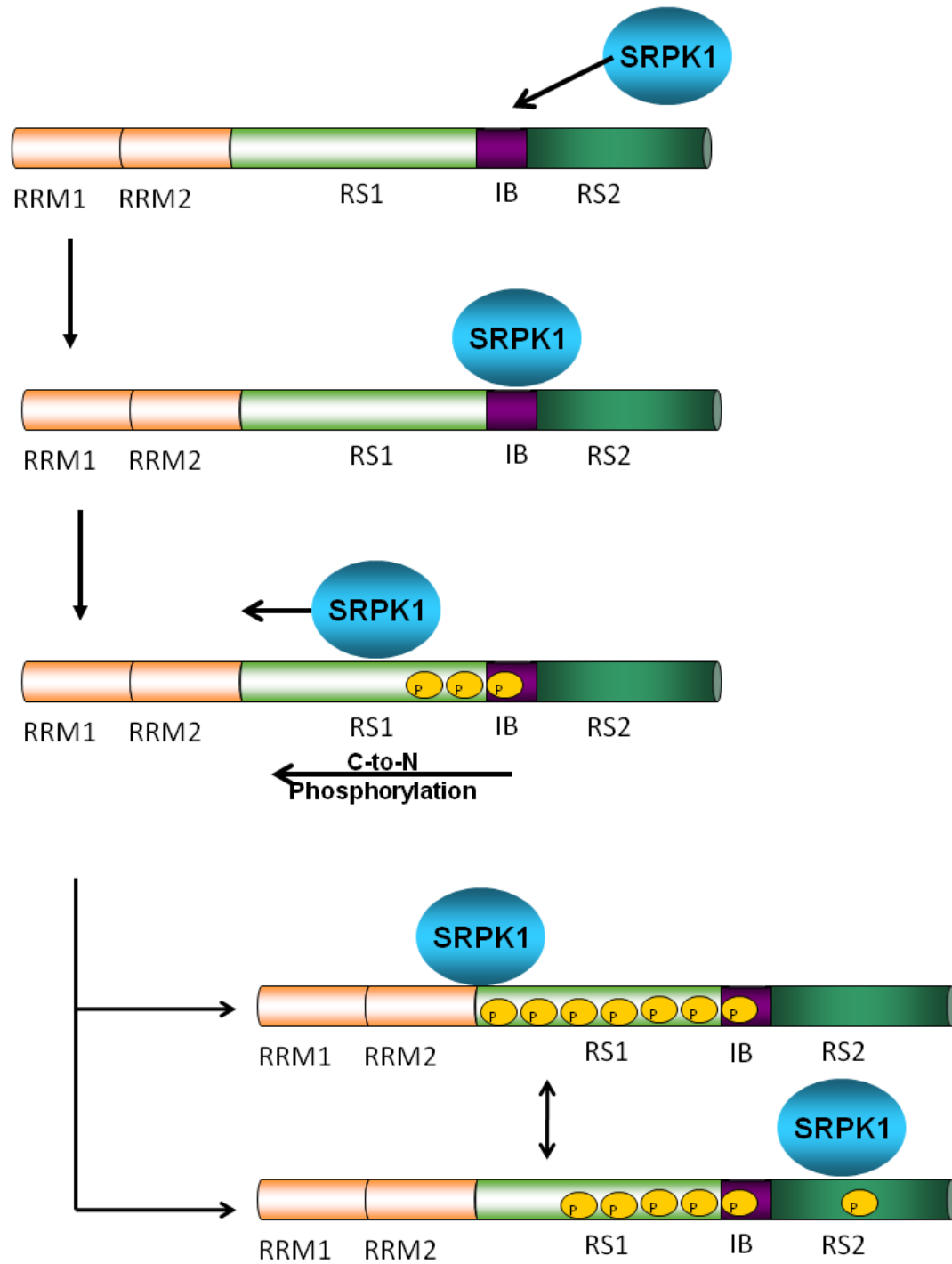
Figure 3.14 Model for ASF/SF2 Phosphorylation by SRPK1. SRPK1 initiates phosphorylation in initiation box at C-terminus of RS1 (IB, purple), catalytzes C->N phosphorylation processively for 8 serines, and then dissociates. SRPK1 will re-bind at either RS2 or RS1 to phosphorylate 3-4 remaining sites each slowly.

Chapter 4

Molecular Interactions Guide Regiospecific Phosphorylation of

ASF/SF2 by SRPK1

A. Introduction

Extensive characterization of the kinetic mechanism of SRPK1 has contributed to deeper understanding of ASF/SF2 phosphorylation by SRPK1, but much is unknown regarding molecular interactions. A recently published X-ray crystal structure of the SRPK1:ASF/SF2 complex, together with cross-linking experimental data, provided fresh evidence for C->N directional phosphorylation [52], as well as delineating potential molecular contacts that may modulate the initiation and elongation phases observed in biochemical assays. The crystal structure made use of a truncated form of SRPK1 lacking the spacer insert and the N-terminal insert (roughly first fifty residues), as well as a truncated ASF/SF2 lacking RS2 and RRM1 domains. This structure shows that the N-terminus of the RS domain docks in the docking groove of SRPK1, near the MAP-kinase insert. The ASF molecule is mono-phosphorylated in this structure, and the phosphoserine is located closer to the active site and far away from the docking groove of SRPK1, suggesting that phosphorylation initiation starts at C-terminus of RS1. These findings now confirm our own directional mechanism. The P+2 pocket near the phosphoserine might act as an electropositive sink, guide the phosphoserine in a certain direction and prevent backtracking of the phosphorylated ASF molecule. Three additional contacts between RRM2 of ASF/SF2 and both kinase lobes might modulate SRPK1 conformation and movement during catalysis. Finally, recent data suggested that the RS2 domain of ASF/SF2 might interact in an intramolecular fashion, and its phosphorylation state might determine RNA binding affinity and specificity (Cho, unpublished data), in addition to RS1 phosphorylation regulation. The goal of this study was to investigate molecular interactions regulating the ordered phosphorylation mechanism observed and characterized in the Chapter 3, and the focus will be on: 1. SRPK1 docking groove and ASF/SF2 RS domain. 2. SRPK1 P+2 pocket and ASF/SF2 phosphoserines. 3. SRPK1 kinase lobes and ASF/SF2 RRM2 domain. 4. RS2 region sequence and function.

B. Results

a. SRPK1 Docking Groove Regulates Initiation and Directional Phosphorylation

Two X-ray crystal structures of SRPK1 were published, and both the RS domain-derived peptide and the truncated ASF/SF2 molecule docked at an acidic groove far away from the catalytic site, near the MAP-kinase insert. The electrostatic interaction between multiple arginines in the ASF/SF2 RS domain and the negatively charged residues in the acidic docking groove is sufficient for strong affinity between the enzyme:substrate complex and necessary for processive phosphorylation (data not shown), as demonstrated by RS domain constructs and an SRPK1 docking groove mutant [SRPK1(6M)]. Such an important docking site might impact directional phosphorylation as well. The SRPK1(6M) mutant construct has four mutant residues that originally interact with N-terminus of RS1 (D564A, E571A, D548A, and K615A) and two additional mutations on residues anticipated to interact with RS1 domain (E558A and E557A) (Fig. 4.2A). SRPK(4M) construct was unstable and did not express well in *E. coli*, so the only viable docking groove mutant is SRPK(6M). As for the P+2 pocket, three positively charged arginines from the P+1 loop and helix αG forms an electropositive sink that might interact with negatively charged phoshoserines and guide them in one single direction. The SRPK1(3RA) mutant construct has three arginines mutated at αG helix (R561A) and at the P+1 loop of SRPK1 (R515A and R518A) (Fig. 4.2B). Both groups of contacts were investigated in ATP limitation assays to discover their potential roles in regulating directional phosphorylation of ASF/SF2 by SRPK1. cl-ASF(214), also known as ASF(5R1K), was the digestion mutant construct utilized due to the digestion site located in the central of the RS1 domain (Fig. 4.2C). wt-SRPK1 served as the control group for comparison. Both SRPK1(3RA) and SRPK1(6M) phosphorylates 15 sites on ASF/SF2 according to MALDI-TOF spectra, similar to wt-SRPK1 phosphorylation.

ASF(5R1K) fragments phosphorylated by all three constructs showed increasing N/C ratio with increasing phosphoryl content and the N/C ratio converged at 0.6-0.8, consistent with normal digestion profile with no significant nonspecific digestion. At lowest phosphoryl content generated by lowest ATP concentration applied, LysC digestion of ASF(5R1K) phosphorylated by wt-SRPK1 and SRPK1(3RA) generated a heavily phosphorylated C-terminal fragment and a clean N-terminal fragment, suggesting that
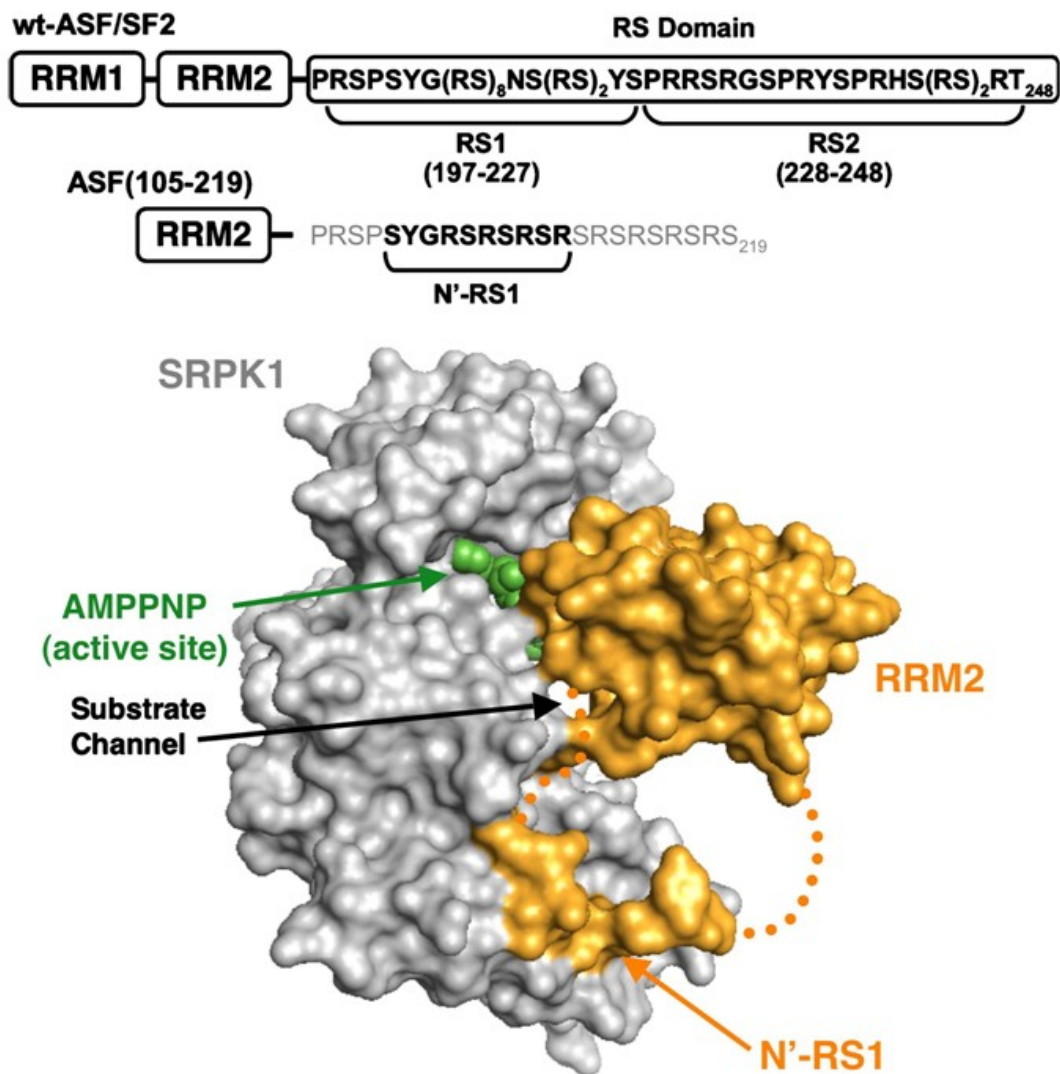
Figure 4.1 X-ray Crystal Structure of the SRPK1:ASF/SF2 Complex. SRPK1 lacking the insert domain (gray) is shown bound with a truncated form of ASF/SF2 [ASF(105-219)] (gold). The densities for RRM2 and a short piece of the RS domain (N'-RS1) are visible. Residues 211-219 and the PRSP sequence linking RRM2 and N'RS1 are disordered (dotted gold line). The active site is designated by AMPPNP.
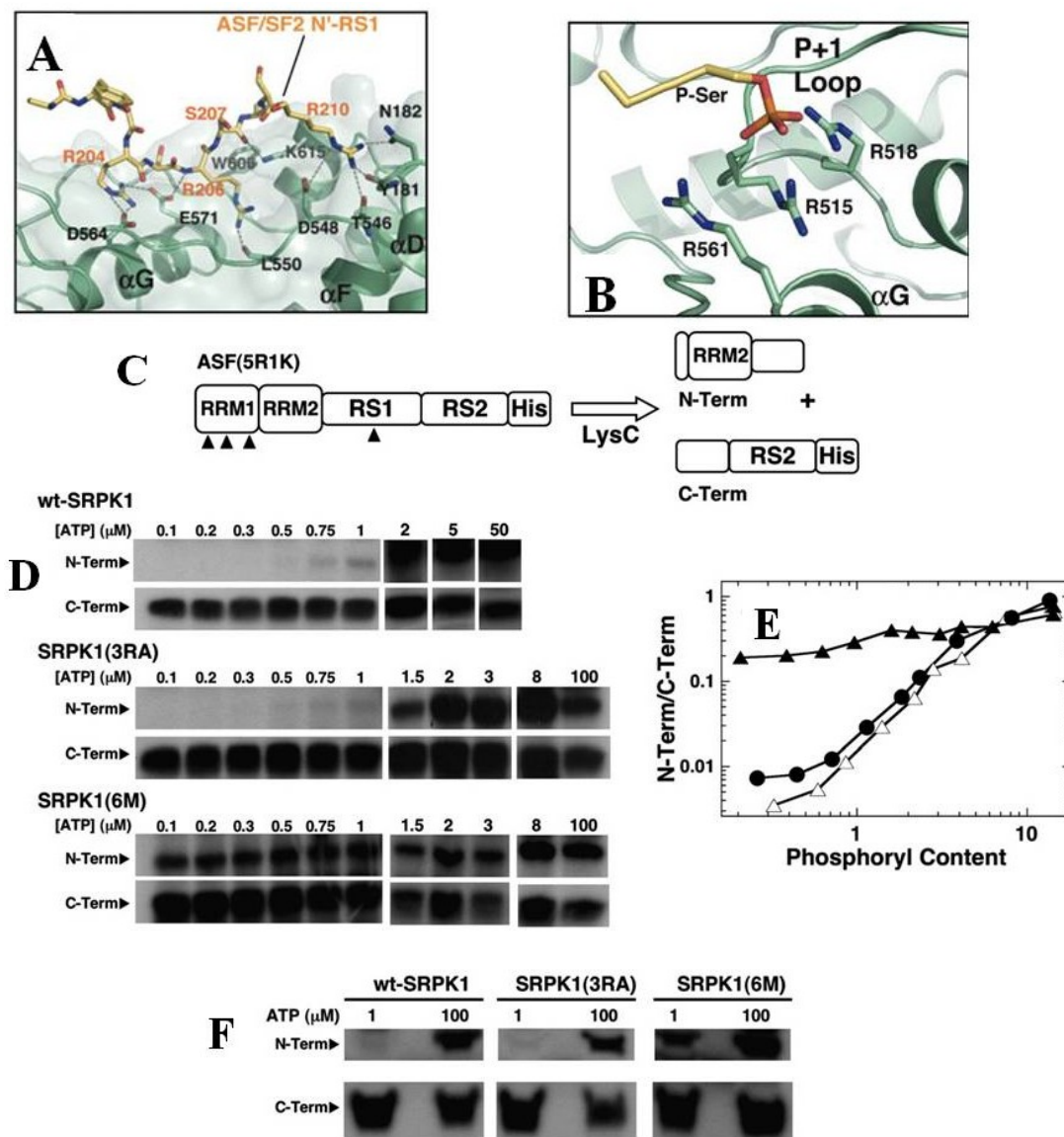
Figure 4.2 SRPK1 Docking Groove Mutants Affect Directional Phosphorylation. A) Docking groove in SRPK1. Residues from N'-RS1 are shown in gold. B) Electropositive pocket. Phosphoserine (P-Ser) interacts with three positively charged arginine residues. C) Digestion fragmentation of ASF/SF2. The mutant form of ASF/SF2 [ASF(5R1K)] with a single lysine mutation at center of RS1 (R214) and several Lys-to-Arg mutations in RRM2 can be digested by LysC into two phosphorylated fragments. Solid arrowheads indicate the locations of LysC digestion sites. D) ATP limitation assays. ASF(5R1K) (250 nM) was pre-equilibrated with 1 μM wt-SRPK, SRPK1(3RA), and 3 μM SRPK1(6M) before the addition of varying amounts of [32P]ATP. N-terminal and C-terminal fragments upon LysC treatment were then resolved by SDS-PAGE autoradiography. E) Fragmentation pattern as a function of phosphoryl content. The ratios of N-terminal proteolytic fragments to C-terminal fragments are plotted as a function of the total phosphoryl content of ASF(5R1K). The total phosphoryl content was measured from undigested reactions. F) Pulse-chase experiments. SRPK1:ASF(5R1K) complexes were mixed with 1 or 100 μM [32P]ATP (pulse) and allowed to react for 8 seconds before the addition of 10 mM cold ATP (chase). N-terminal and C-terminal fragments were then generated with LysC.

C-terminal initiation preference is strong (Fig. 4.2D). For SRPK1(6M) the distribution of phosphates was much more even between the N- and C-terminal fragments, suggesting that the C-terminal initiation preference was partially abolished. The C-terminal preference changed from 50~100 fold for wt-SRPK1 and SRPK1(3RA) to ~5 fold for SRPK1(6M). In order to establish whether this phenomenon, measured under equilibrium conditions, was also observed under kinetic conditions, the order of phosphorylation was observed under pulse-chase experiment as well. SRPK1:ASF(5R1K) complexes were mixed with 1 or 100 μM [^{32}P]ATP (pulse) and allowed to react for 8 seconds before the addition of 10 mM cold ATP (chase). The low ATP concentration was used to capture the extremely fast initial phosphorylation event. For wt- SRPK1 or SRPK1(3RA), very little N-terminal phosphorylation was observed after the pulse chase using 1 μM [^{32}P]ATP, consistent with ordered C->N phosphorylation of RS1domain (Fig. 4.2E). However, extensive phosphorylation of the N-terminal fragment was observed at 1 μM [^{32}P]ATP for SRPK1(6M).   At low ATP concentration, the N/C ratio was 0.06-0.07 for wt-SRPK1 and SRPK1(3RA) and 0.40 for SRPK1(6M). These findings confirmed the ATP limitation assay results and indicated that the docking groove plays an essential role in regulating phosphorylation initiation at the RS1/RS2 border.

b.   Electropositive Pocket in SRPK1 Assists Late Phosphorylation Steps

Despite its lack of participation in directional phosphorylation regulation, the electropositive P+2 pocket was investigated for alternative roles in modulating the mechanism of ASF/SF2 phosphorylation. In addition to the SRPK1(3RA) mutant construct, a single-alanine mutation at R561 and a double-alanine mutation at R515 and R518 were studied. None of these mutations had significant impact on overall binding affinity or on the phosphoryl content of ASF/SF2 (Fig. 4.3A & B). Instead, the mutations had a significant effect on the kinetics of ASF/SF2 phosphorylation.   In single-turnover experiments, the mutations did not affect the initial phosphorylation rate, but significantly lowered the amplitude of the fast phase (Fig. 4.3A). Whereas wt-SRPK1 rapidly phosphorylated the first 12 sites, all three mutants lowered this initial phase amplitude to about 5-7 sites. Start-trap experiments were also performed to determine the mechanism of phosphate addition. The SRPK1(3RA)-catalyzed reaction was
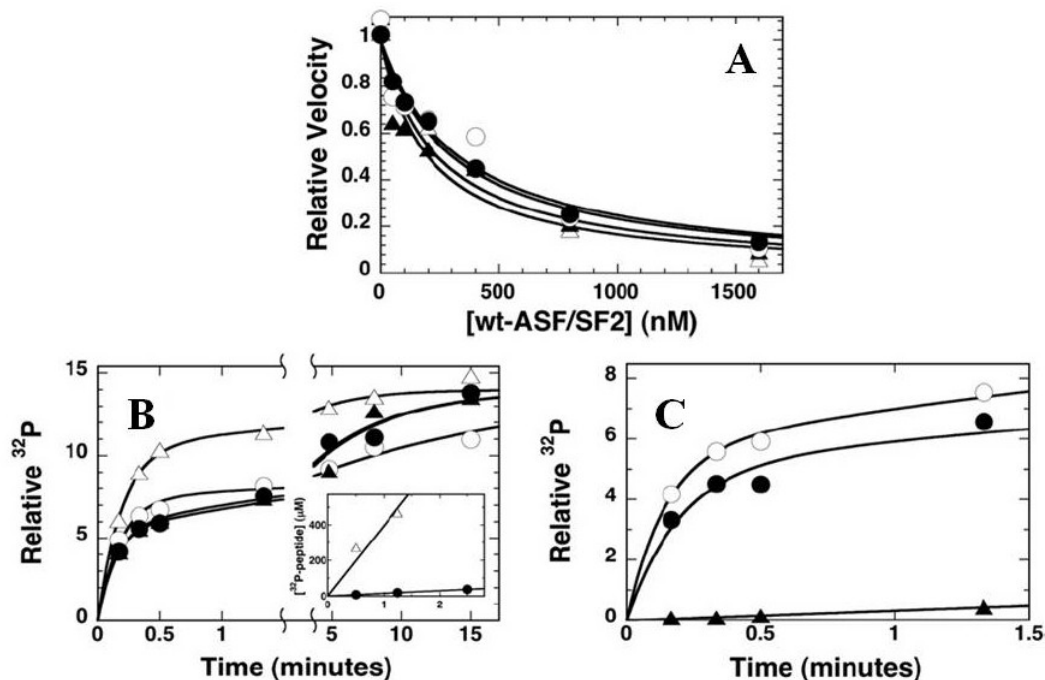
Figure 4.3 Role of the Electropositive Pocket in Late Phosphorylation Steps. A) Binding affinities. Competition experiments were performed using 50 nM ASF(ΔRRM2), varying amounts of ASF/SF2, and up to 15 nM of each enzyme. True K_i values of 150±23, 110±21, 240±36, and 220±15 were obtained for wt-SRPK1 (Δ), SRPK1(RA) (▲), SRPK1(2RA) (○), and SRPK1(3RA) (●), respectively. K_m values of 95, 110, 139, and 134 nM were obtained for wt-SRPK, SRPK1(RA), SRPK1(2RA), and SRPK1(3RA), respectively, and used for determining the true K_i values. B) Single-turnover kinetics. Enzymes (2 μM) were pre-equilibrated with ASF/SF2 (250 nM) and reacted with 100 μM [^{32}P]ATP. The data are fitted to double-exponential functions with initial amplitudes of 10.5, 5.5, 7.1, and 5.6 sites for wt-SRPK, SRPK1(RA), SRPK1(2RA), and SRPK1(3RA), respectively. C) Start-trap kinetics. SRPK1(3RA) (2 μM) was pre-equilibrated with ASF/SF2 (250 nM) before 100 μM [^{32}P]ATP was added in the absence (○) and in the presence (●) of 60 μM kinase-dead SRPK1 at start time. In the trap-start experiments, kdSRPK1 was added to the enzyme-substrate complex before the addition of 100 μM [^{32}P]ATP. The slope of the trap-start data is 0.2 site min^{-1}.

largely processive within the fast initial phase (5 sites) of the reaction (Fig. 4.3C). Due to the slower rate of the secondary phase and the intrinsic background rate of the trap-start control, no conclusion could be drawn as to the mechanism in the secondary phase. Overall, the data suggested that this P+2 electropositive pocket is very important in facilitating the rate of later phosphorylation steps.

c.   Specific RRM2 Contacts Affect Processive, but Not Directional Phosphorylation

Deletion analyses suggest that ASF/SF2 requires an RRM directly N-terminal to the RS domain for processive phosphorylation of later serines , and the X-ray crystal structure described 4 possible residues on RRM2 interacting with SRPK1 [52]. All 4 residues were mutated into alanines in the ASF(4M) construct and it was phosphorylated to the same extent as wt-ASF based on MALDI-TOF spectrometric analysis. In single-turnover experiment, ASF(4M) displayed similar kinetic parameters as the RRM-deleted substrates in that the initial phase was faster than wt-ASF (4 *versus* > 10 min^{-1}) (Fig. 4.4B). Mutation of these 4 residues also led to a reduction in processive phosphorylation relative to wt-ASF (eight *versus* five sites) (Fig. 4.4C) by an amount larger than removal of either RRM, but smaller than removal of both RRM domains [110]. Consistent with deletion results, unphosphorylated ASF(4M) showed similar affinity as wt-ASF. Overall the data showed that disruption of specific RRM2 contacts predicted by the X-ray crystal structure impacts the phosphorylation mode of later serines in a manner intermediate between single RRM removal and complete RRM1/2 removal.

To address whether these specific RRM2 contacts serve a role in directional phosphorylation, these four mutations were introduced into the digestion mutant ASF(5R1K). The new mutant ASF(4M-5R1K) showed similar affinity to SRPK1 and similar phosphoryl content as wt-ASF. The mutations did not change the strong initiation preference for C-terminus of RS1 domain, and at the lowest ATP concentration the phosphates were predominantly (by ~100 fold) placed in the C-terminal fragment, similar to wt-SRPK1 (Fig. 4.4D). Furthermore, the final N/C ratio at full phosphorylation was 0.8, the same as it was for wt-SRPK. Phosphorylation preference for RS1 domain over RS2 is not altered. Overall, the data suggested that specific RRM2 contacts affect the extent of processivity, but not the directionality or the region-specificity of RS domain phosphorylation.
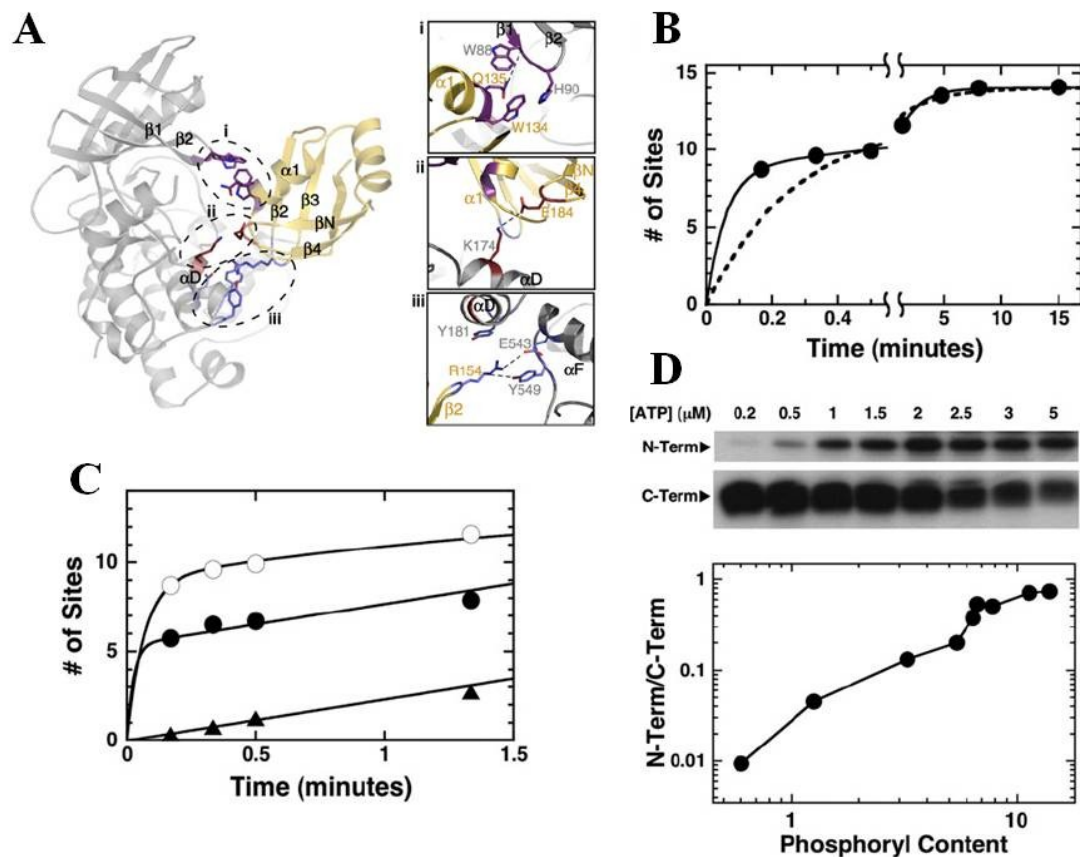
Figure 4.4. Effects of Specific RRM2 Mutations on Phosphorylation Mechanism. A) Specific RRM2 contacts. SRPK1 (gray) is contacted in both kinase lobes by the RRM2 domain (gold) of ASF/SF2. 4 residues on RRM2 are involved (W134, Q135, R154, and E184). B) Single-turnover kinetics. SRPK1 (1 μM) is incubated with 250 nM ASF(4M), and the reaction is initiated with the addition of 100 μM [^{32}P]ATP. The data are fitted to a double-exponential equation with rate constants of > 18 and 0.57 min^{-1} and amplitudes of 8.7 and 5.3 phosphates. The dotted line is from wt-ASF. C) Start-trap experiment. SRPK1 (1 μM) was incubated with 250 nM ASF(4M), and the reaction was initiated with the addition of 100 μM [^{32}P]ATP in the absence (○) and in the presence (●) of 60 μM kinase-dead SRPK1 added at start of reaction. In trap-start experiments, kdSRPK1 was added prior to reaction start (▲) and fitted to a linear function with a slope of 2.3 sites min^{-1}. The start-trap data are fitted with a single-exponential function with a rate constant of > 20 min^{-1} and an amplitude of 5.4 sites, followed by a linear function of 2.3 sites min^{-1}. D) ATP limitation experiments. ASF(4M-5R1K) (250 nM) was pre-equilibrated with 3 μM wt-SRPK1 before the addition of varying amounts of [^{32}P]ATP. N-terminal and C-terminal fragments upon LysC treatment were then resolved by SDS-PAGE autoradiography, and the ratios of N- and C-terminal proteolytic fragments are plotted as a function of the total phosphorylation content. Total phosphorylation content is measured from undigested reactions.

d. **A Flanking RRM Is Important for Phosphorylation Initiation Preference**

Even though specific contacts predicted by the X-ray crystal structure do not impact directional phosphorylation, the structure lacks an RRM1 domain within ASF/SF2 and is, therefore, incomplete. Additional studies are necessary to show categorically that RRMs are not important for phosphorylation initiation regulation. A series of deletion and domain swap mutations were made to reduce ASF/SF2 into individual modules that so that the order of phosphorylation can be assessed (Fig. 4.6A). RRM1 was deleted in the digestion construct cl-ASF(214) [cl-ΔRRM1(214)] to address whether two RRMs are necessary for C-terminal phosphorylation initiation. The positions of the RS domain and RRM2 domain were swapped in [cl-RS-RRM2(214)] to address the role of RRM2 in the absence of RRM1. The swap should remove all existing contacts defined by the X-ray structure (Fig. 4.4A). To address the specificity of these interactions in a construct similar to the one used in the X-ray structure, 4 alanines were mutated in the RRM2 of cl-ΔRRM1(214) (W134A, Q135A, E184A, and R154A) to generate the construct cl-RRM2-4M(214). To address the phosphorylation region-specificity in these new constructs, the digestion site in the RS domain was shifted to cl-RS-RRM2(214) position 224 [cl-RS-RRM2(224)] near the RS1/RS2 boundary. All these constructs were phosphorylated to the same approximate level as wt-ASF (13-15 sites) and possessed high catalytic activity (Fig. 4.5). To identify the fragments in the domain swap mutants, Nickel-resin pull-down assays were performed. The N- and C-terminal fragments migrated in reverse order owing to their different molecular weights, and the C-terminal fragment, lacking the His-tag, was not pulled down by Nickel resin (Fig. 4.6B). No intermediate fragment was observed for both domain swap constructs since they only contain one cleavable lysine.

Having validated the digestion mutants, ATP limitation assays were conducted to establish phosphorylation initiation preferences. While cl-ΔRRM1(214) showed strong preference for C-terminal initiation, cl-RS-RRM2(214) showed a significantly reduced preference (Fig. 4.6C). The N/C ratio curve for cl-ΔRRM1(214) was similar to that for cl-ASF(214), suggesting that RRM1 is not essential for C->N directionality when RRM2 is still present (Fig. 4.6D). In contrast to the roughly 80-fold preference exhibited by cl-ASF(214) and cl-ΔRRM1(214), cl-RS-RRM2(214) showed only 3-fold preference for C-terminal initiation. Thus, swapping RRM2 from its N-terminal position to a C-terminal position relative
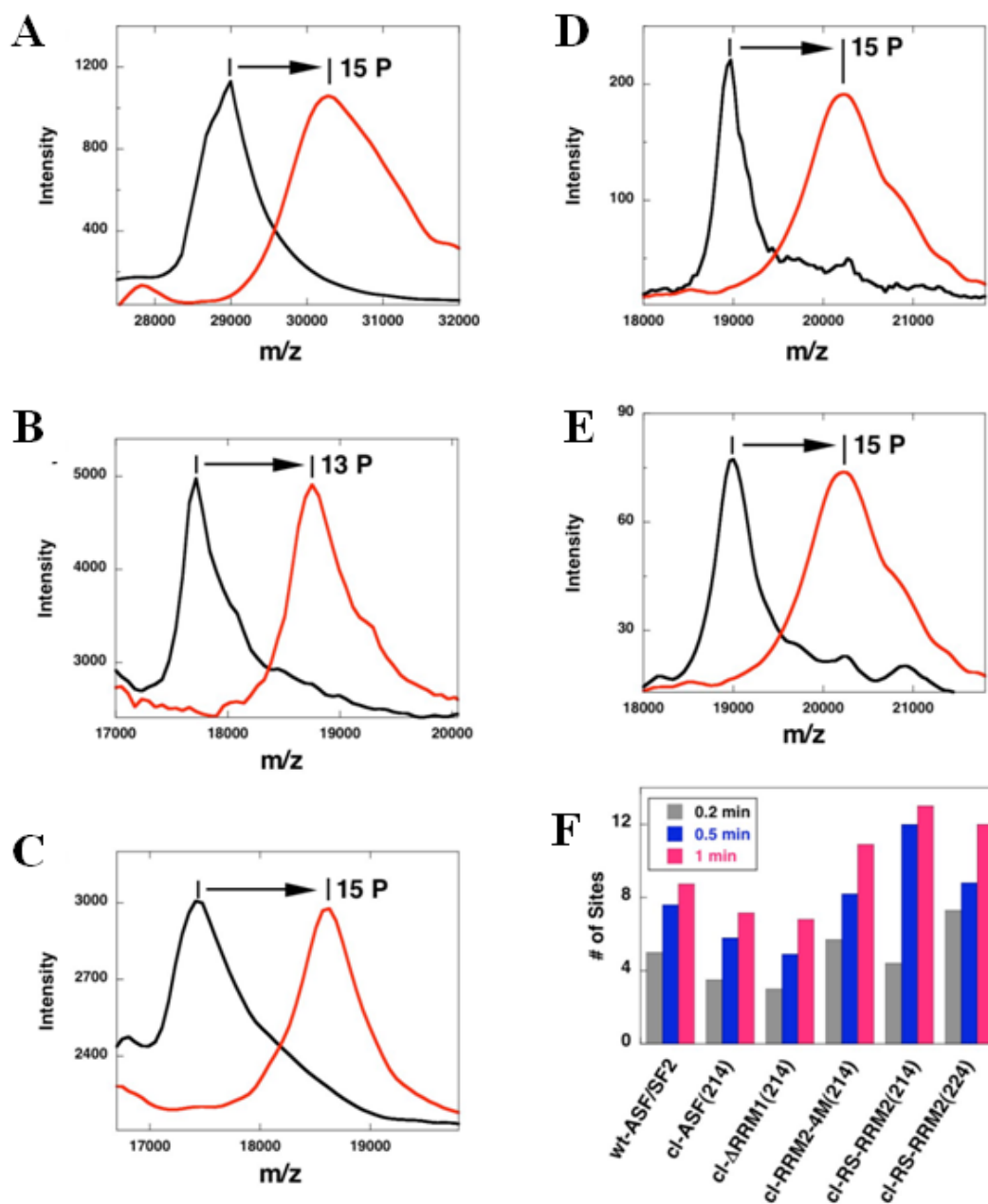
Figure 4.5 Phosphorylation of Several Digestion Substrates. A-E) Phosphoryl contents. MALDI-TOF spectra of SRPK1 and cl-ASF(214) (A), cl-ΔRRM1(214) (B), cl-RRM2-4M(214) (C), cl-RS-RRM2(214) (D), and cl-RS-RRM2(224) (E) in the absence (black) and presence (red) of ATP are recorded. F) Kinetic analyses. Complexes of 1 μM SRPK1 and 250 nM of wt-ASF/SF2, cl-ASF(214), cl-ΔRRM1(214), cl-RRM2-4M(214), cl-RS-RRM2(214), and cl-RS-RRM2(224) were reacted with 100 μM [^{32}P]ATP, quenched with SDS loading buffer at 0.2, 0.5, and 1 minute and the phosphoryl contents are plotted in the bar graph. For these studies, wt-ASF/SF2 contains a C-terminal His-tag for comparison with the digestion proteins.
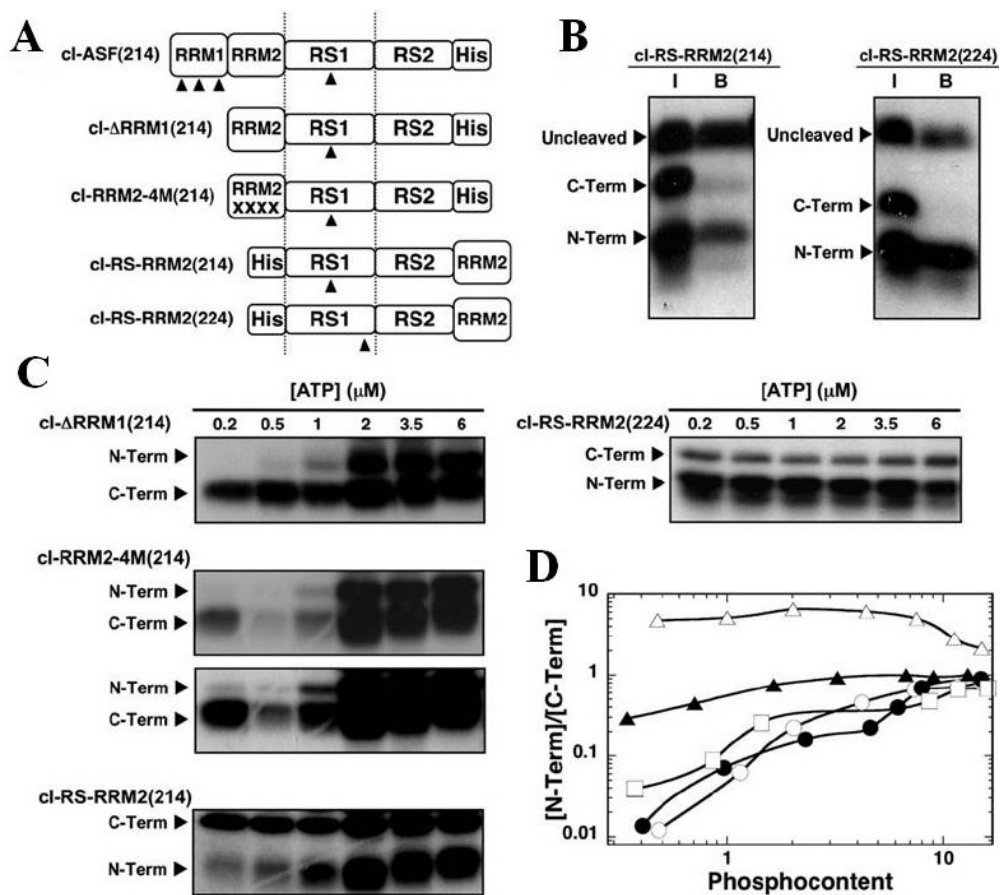
Figure 4.6 Role of RRMs for Phosphorylation Initiation. A) Deletion and swap mutants. Numbers in parentheses define the Arg-to-Lys digestion site in the RS domain. B) Nickel-resin pull-down assays. SRPK1 (1 μM) was used to phosphorylate 250 M cl-RS-RRM2(214) and cl-RS-RRM2(224) before digestion by LysC. Samples before binding to the Ni-resin (I) and after binding and washing (B) are displayed in the autoradiogram. (C) ATP limitation experiments. Complexes of SRPK1 (1 μM) and mutants (250 nM) were phosphorylated for 20 minutes using varying amounts of $[^{32}P]ATP$, digested with LysC, and separated on SDS-PAGE. Two time exposures (20 and 70 minutes) for the cl-RRM2-4M(214) gel were performed for better visualization of low and high ATP concentrations. D) N/C plot. The ratios of ^{32}P in the N-terminal and C-terminal fragments of cl-ΔRRM1(214) (●), cl-ASF(214) (○), cl-RRM2-4M(214) (□), cl-RS-RRM2(214) (▲), and cl-RS-RRM2(224) (Δ) are plotted against the total phosphoryl content.

to the RS domain reduced regiospecific preference by about 25-fold. Follow-up ATP limitation experiment on a swap construct where the digestion site was moved closer to the RS1/RS2 boundary [cl-RS-RRM2(224)] demonstrated that RS1 is still preferred over RS2 for phosphorylation. The N/C ratio was 5 at low phosphoryl content for cl-RS-RRM2(224) and 8 for cl-ASF(224) (Fig. 3.9B). Also, the N/C ratio was 2 at full phosphorylation for cl-RS-RRM2(224), the same for cl-ASF(224), suggesting that SRPK1 still phosphorylates serines in RS2 even if RRM2 is attached to RS2 region instead. A follow-up mutant construct lacking both RRMs but with single Arg-to-Lys mutation at R214 [cl-RS(K214)] displayed a poor digestion profile at high phosphoryl content, but at low phosphoryl content it displayed N/C ratio of roughly 0.3, consistent with the observation that RRM2 is important in regulating phosphorylation initiation. Finally, SRPK1 still initiated the phosphorylation of cl-RRM2-4M(214) in the C-terminal fragment (30-fold preference at low phosphoryl content), confirming the hypothesis that RRM2 contacts with SRPK1 are flexible and not reliant on discrete contacts found in X-ray structure. Overall, the data suggested that an RRM directly flanking the N-terminus of the RS domain regulates phosphorylation initiation but not RS1/RS2 region-specificity.

e.   **Interdependent Docking Surfaces Control Initiation**

Since the above deletion and swap mutants suggested that an N-terminal RRM flanking the RS domain assists in phosphorylation initiation, the question whether the RRM interaction and the SRPK1 docking groove interaction (Fig. 4.2) act independently would be addressed. ATP limitation studies were conducted on cl-ΔRRM1(214) and cl-RS-RRM2(214) using a docking-defective mutant of SRPK1 [SRPK1(6M)]. Using SRPK1(6M), cl-ΔRRM1(214) displayed fully random phosphorylation with an N/C ratio of 1 at low ATP concentrations (Fig. 4.7A & B), suggesting that the removal of both the docking groove residues and RRM1 leads to further reduction in initiation preference. Interestingly, SRPK1(6M) showed a weak preference for N-terminal phosphorylation initiation with cl-RS-RRM2(214) (N/C=2). To determine whether mutations in the docking groove affects region-specificity of the enzyme, LysC digestion of cl-ASF(229) phosphorylated at low (0.2 μM) and high ATP concentrations (100 μM) were analyzed (Fig. 4.7A). While SRPK1(6M) phosphorylated exclusively in the N-terminal
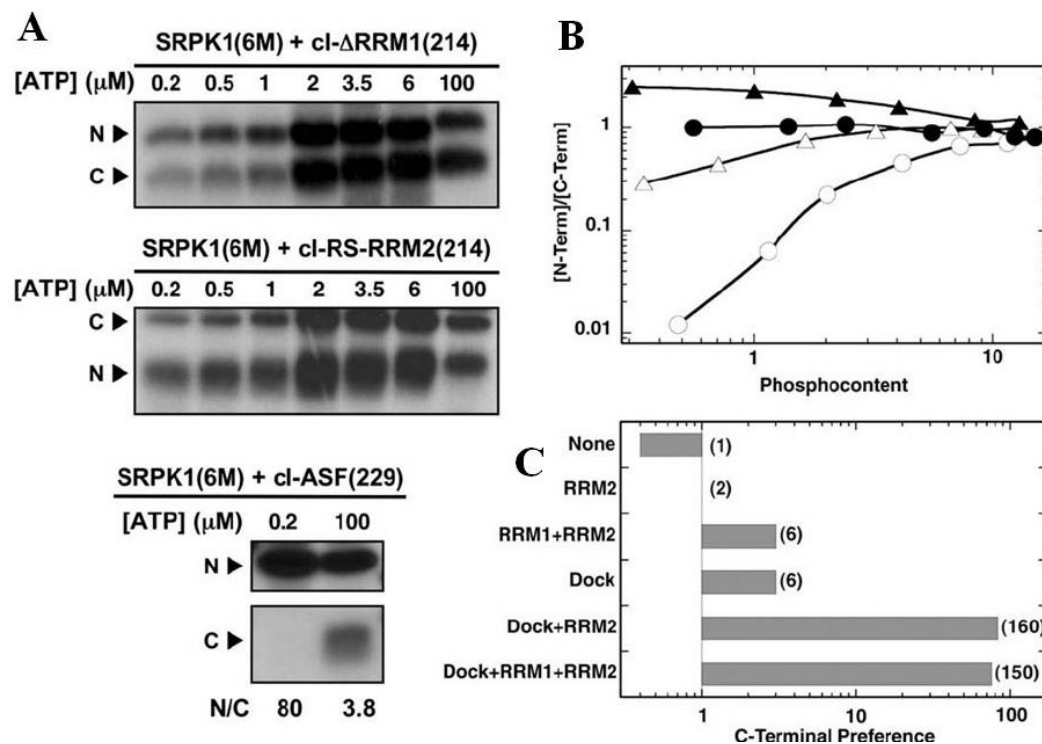
Figure 4.7 SRPK1 Docking Mutations Impact Phosphorylation Initiation. A) ATP limitation experiments. Complexes of SRPK1(6M) (3 μM) and 250 nM cl-ΔRRM1(214), cl-RS-RRM2(214), cl-ASF(229) were reacted with varying amounts of [^{32}P]ATP for 20 minutes and digested with LysC. N-terminal and C-terminal fragments were separated on SDS-PAGE. (B) N/C plot. The ratios of ^{32}P in the N-terminal and C-terminal fragments for SRPK1(6M) with cl-ΔRRM1(214) (●) and cl-RS-RRM2(214) (▲), and for wt SRPK1 with with cl-ΔRRM1(214) (○) and cl-RS-RRM2(214) (Δ), are plotted against the total phosphoryl content. The data for wt SRPK1 are taken from Fig. 4.6D. C) Effects of contacts on C-terminal preference. N/C ratios at low ATP concentrations are plotted in the bar graph as a function of added contacts. Numbers in parentheses represent the fold increase in C-terminal initiation preference relative to SRPK1(6M) and cl-RS-RRM2(214).

fragment (RS1) at low phosphoryl content, the mutant displayed the same overall region-specificity as wt-SRPK1 at high phosphoryl content (N/C=3.8). Based on total phosphoryl content determined by MALDI-TOF and the N/C ratio at full phosphorylation under high ATP concentration, SRPK1(6M) modified about three serines in RS2, similar to wt-SRPK1 (Fig. 3.11). Overall, the data suggested that no single contact surface between SRPK1 and ASF/SF2 can adequately account for the strong preference for C-terminal phosphorylation initiation in RS1 domain, and contacts between SRPK1 and the N-terminal portion of RS1 and its flanking RRMs function in an interdependent manner to control phosphorylation initiation (Fig. 4.7C).

f.   Investigating SRPK1 Spacer Insert and ASF/SF2 RS2 Segment

Even though interactions explored in previous sections have fully accounted for the 80-100-fold C-terminal phosphorylation initiation preference shown in the wt SRPK:ASF/SF2 complex, additional structural determinants within SRPK1 and ASF/SF2 were also investigated for potential regulatory role secondary to the N-terminal RS1 domain and RRM contacts. The spacer insert regulates subcellular localization of SRPK1 and has minor impact on catalytic rate [30, 55], while the RS2 domain may form contacts with RRM domains intramolecularly (Cho, unpublished data). ATP limitation assays were conducted on the SRPK1 spacer deletion mutant construct [SRPK1(ΔS)] and ASF/SF2 digestion construct lacking RS2 domain [cl-ASFΔRS2(214)]. At the lowest ATP concentration (0.2 μM), phosphates were placed predominantly in the C-terminal fragment (N/C = 0.02, 50-fold preference) for SRPK1(ΔS) (Fig. 4.8C), and at highest ATP concentration (100 μM) the N/C ratio was 0.9, similar to wt SRPK1. The data suggested that the C-terminal initiation preference is unchanged by deletion of the spacer domain. As for cl-ASFΔRS2(214), the ATP limitation digestion profile suggested incomplete phosphorylation due to the late phosphorylation of the N-terminal fragment (Fig. 4.8B). Either the K_m value is very high for later sites, or the phosphorylation has slowed down for low ATP concentrations and the endpoint has not been reached. However, clearly at low phosphoryl content the C-terminal fragment was preferentially phosphorylated, implying that RS2 domain deletion does not change C->N directional phosphorylation. Finally, single-turnover kinetic analysis demonstrated that ASF(1-219) and ASF(1-226)
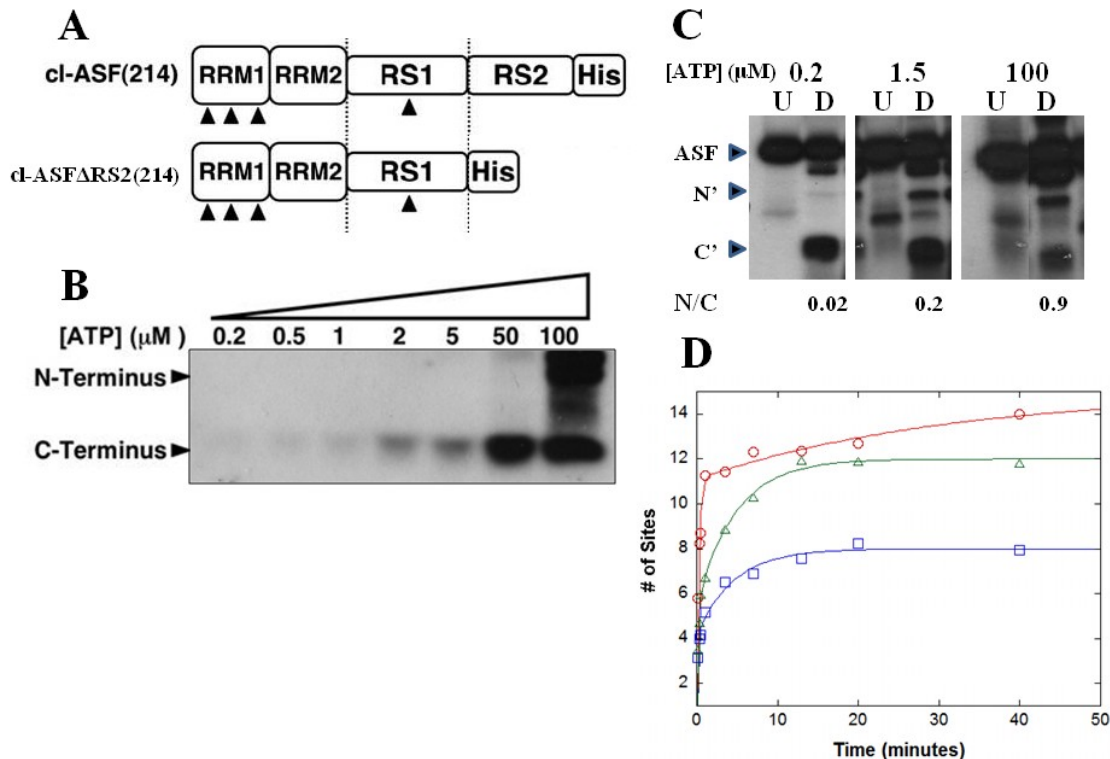
Figure 4.8 Investigating SRPK1 Spacer Domain and ASF/SF2 RS2 Segment. A) Domain structure of digestion mutants. B) ATP limitation assay. Complex of wt SRPK1 (1 μM) and 250 nM cl-ASFΔRS2(214) was reacted with varying amounts of [^{32}P]ATP for 20 minutes and digested with LysC. N-terminal and C-terminal fragments were separated on SDS-PAGE. C) ATP limitation assay and N/C calculations.   Complex of SRPK1(ΔS) (1 μM) and 250 nM cl-ASF(214) was reacted with varying amounts of [^{32}P]ATP for 20 minutes and digested with LysC. N-terminal and C-terminal fragments were separated on SDS-PAGE. The ratios of ^{32}P in the N-terminal and C-terminal fragments are calculated and displayed. D) Single-turnover kinetics. SRPK1 (1 μM) was incubated with 250 nM wt-ASF (○), ASF(1-219) (□), and ASF(1-226) (△),   and the reaction was initiated with the addition of 100 μM [^{32}P]ATP. The data are fitted to a double-exponential equation with rate constants of 3.9 and 0.03 min^{-1} and amplitudes of 11.1 and 3.9 phosphates for wt-ASF, with rate constants of 7.0 and 0.22 min^{-1} and amplitudes of 4.2 and 3.8 phosphates for ASF(1-219), and with rate constants of 5.0 and 0.21 min^{-1} and amplitudes of 5.6 and 6.4 phosphates for ASF(1-226).

display biphasic phosphorylation kinetics even in the absence of RS2 region, suggesting that the RS2 segment has no significant impact on phosphoryl content of RS1 region but enhances fast-phase phosphorylation (Fig. 3.10A & 4.8D). This observation also suggested that a longer time course would be necessary for full phosphorylation of cl-ASFΔRS2(214) so that N/C ratio upon LysC digestion could reach 0.7-0.8. Overall, the data suggested that the spacer domain or RS2 segment do not regulate C-terminal regiospecific phosphorylation of ASF/SF2 by SRPK1.

g.   Sequence Requirements for RS1/RS2 Preference

Having resolved molecular interactions that regulate directional phosphorylation of the RS1 region of ASF/SF2 by SRPK1, the investigation turned toward the RS1/RS2 preference exhibited by SRPK1, a phenomenon unaffected by the docking groove or RRMs. As shown in Fig. 3.11, SRPK1 strongly prefers to phosphorylate the N-terminus of the RS domain first (RS1), followed by late phosporylation of the C-terminus (RS2). Several possible explanations were explored in mutagenesis constructs.

The RS1 segment is flanked by two prolines on either the N- or the C-terminus (P197 and P200 *versus* P228 and P235, respectively), and previous studies suggested that prolines can define a protein segment with flanking turn motifs and present it for protein-protein interactions [111]. Also, within the 4 prolines, 3 have vicinal serines, and phosphorylation of these serines not only favors the formation of a cis-pSer-Pro peptide bond, but also decelerates the interconversion of the cis-trans isomeric states even further [112], another possible molecular switch for SRPK1 phosphorylation. Pin1, a peptidyl-prolyl isomerase involved in mitosis and residing in nuclear speckles, may interact with partially or fully phosphorylated SR proteins and catalyze cis-trans isomerization [112, 113]. It was purified in *E. coli* and tested with an AAPF-peptide substrate in the chymotrypsin assay to ensure enzymatic activity (data not shown) [114]. Pin1(30 nM) in the presence of stoichiometric SRPK1 and ASF/SF2 (500 nM), or pre-incubated with ASF/SF2 for 10 minutes, exhibited no significant increase in phosphoryl content or change in catalytic rate (Fig. 4.9A). However, the slow-phase phosphorylation was removed and the data were best fitted to a single-exponential equation, suggesting that RS2 phosphorylation may be impacted.

90
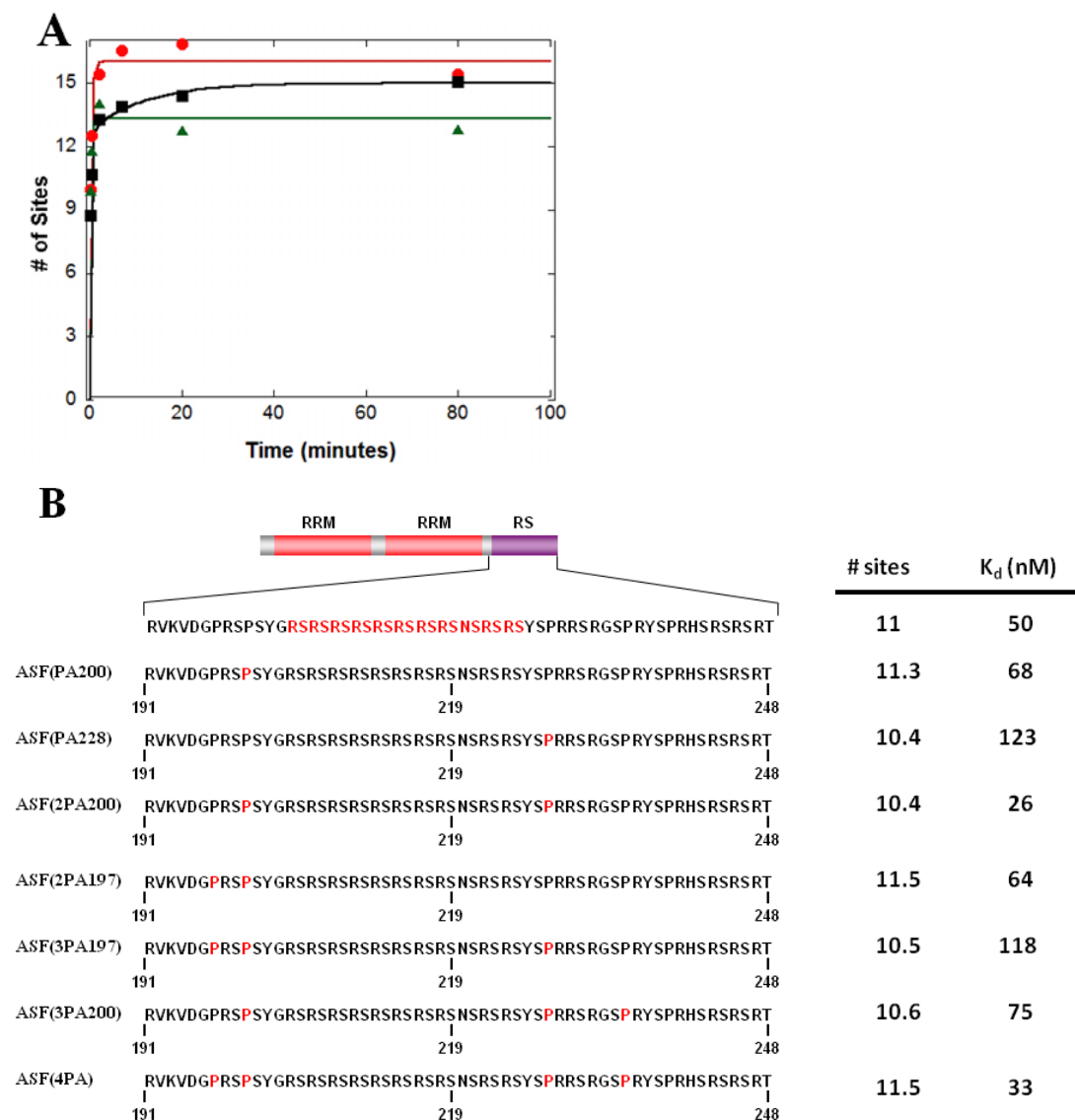
Figure 4.9 Investigating Flanking Prolines. A) Pin1 effects in single-turnover kinetics. SRPK1 (1 μM) was incubated with 250 nM wt-ASF (■), wt-ASF and 30 nM Pin1 (●), or wt-ASF and 30 nM Pin1 pre-incubated for 10 minutes first (▲), and the reaction is initiated with the addition of 100 μM [^{32}P]ATP. The data are fitted to a double-exponential equation with rate constants of 3.5 and 0.085 min^{-1} and amplitudes of 12.7 and 2.3 phosphates for wt-ASF. The data are best fitted to a single-exponential equation with rate constant of 4.3 min^{-1} and amplitude of 13.3 for wt-ASF and 30 nM Pin1. The data are best fitted to a single-exponential equation with rate constant of 2.9 min^{-1} and amplitude of 16.0 for wt-ASF and 30 nM Pin1.

Prolines were also mutated into alanines in various patterns on both ends of RS1, but no significant change in steady-state or single-turnover kinetics was observed. The binding affinity, as measured in K_d values, and phosphoryl content were also similar to wt-ASF/SF2 (Fig. 4.9B), and start-trap experiments demonstrated processive phosphorylation (data not shown). Overall, the data suggested that prolines do not play an important role in SRPK1 phosphorylation of ASF/SF2.

The second possibility explored involved another unique feature of RS1. The RS1 segment has an uninterrupted stretch of eight pairs of RS repeats, while the RS2 region has at most three RS pairs next to one another. The RS2 segment was mutated to be more "RS1-like" by the substitution of residues to create lengthy, uninterrupted RS repeat region into RS2, and this new mutant [cl-RS$_8$-RS$_8$(224)] was phosphorylated to a higher extent (20 *versus* 15 sites) than wt construct cl-ASF(224) (Fig. 4.10A & B). Removing eight consecutive serines in RS1 from this construct to generate cl-SA$_8$-RS$_8$(224) lowered the phosphoryl content by about nine sites, consistent with a direct correlation between RS count and phosphoryl content observed in Fig. 3.10. However, another mutant that contains the same number of serines but fewer arginines [cl-RS$_8$-RA$_4$(224)] was phosphorylated to the same extent as cl-ASF(224) (Fig. 4.10A), clearly less than cl-RS$_8$-RS$_8$(224), suggesting that phosphoryl content is dependent on a consecutive array of RS repeats and not on total serine content.

In order to further investigate RS1/RS2 specificity, the mutant substrates were phosphorylated at low and high ATP concentrations and the N-terminal and C-terminal fragments were separated by SDS-PAGE after LysC treatment (Fig. 4.10C). At high ATP concentration, both the N-terminal and C-terminal fragments of cl-RS$_8$-RS$_8$(224) were phosphorylated to the same extent (N/C=1), suggesting that 10 serines in the C-terminal fragment are modified in this construct. The phosphorylation in mutant RS2 segment is not dependent on phosphorylation of RS1 region or distance from RRM domains, as was shown by 10 site phosphoryl content of C-terminal fragment of cl-SA$_8$-RS$_8$(224). The N-terminal fragment of cl-SA$_8$-RS$_8$(224) was hardly phosphorylated, producing a very low N/C ratio (Fig. 4.10C). Interestingly, disruption of the consecutive RS repeats in the RS2 region in cl-RS$_8$-RA$_4$(224) led to very high N/C ratio at both low and high ATP concentrations, consistent with only two sites phosphorylated in RS2 (Fig. 4.10A & B). The cl-RA$_4$-RS$_8$ construct, in contrast, was hyper-phosphorylated on 20 sites and
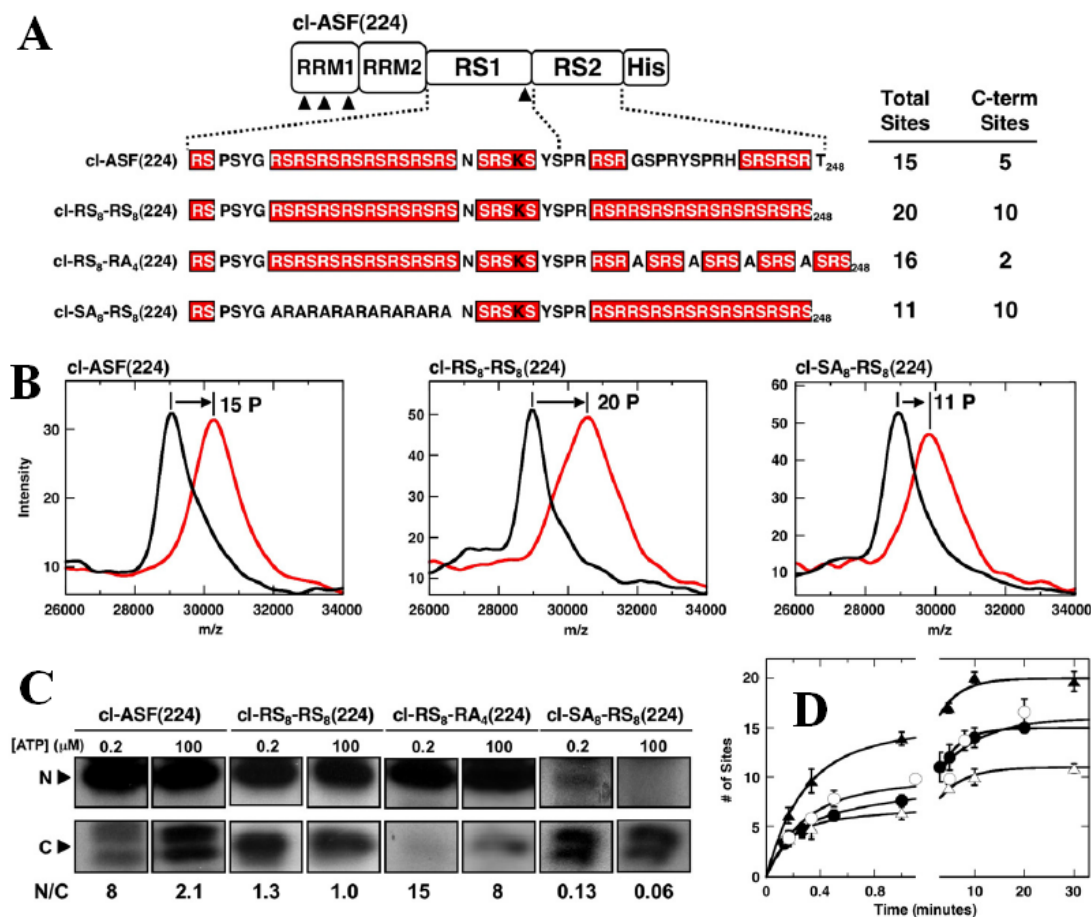
Figure 4.10 Effects of Altered RS/SR Content on Phosphoryl Content. A) Mutations in RS2. Mutations in the RS2 segment are made in the context of the parent construct cl-ASF(224). MALDI-TOF analyses of the total phosphoryl contents of each mutant are displayed along with estimates of the phospohryl contents of the C-terminal fragments based on LysC treatment. [see (C) for details]. B) MALDI-TOF spectra. Mass spectrometric data were recorded for cl-ASF(224), cl-RS$_8$-RS$_8$(224), and cl-SA$_8$-RS$_8$(224) in the presence of SRPK1 and in the absence (black) and in the presence (red) of ATP. C) ATP limitation experiments. Complexes of SRPK1 (1 μM) and mutants (250 nM) were phosphorylated for 20 minutes using 0.2 and 100 μM [^{32}P]ATP, digested with LysC, and separated on SDS-PAGE. N/C ratios at each ATP are shown. D) Single-turnover kinetic analyses. Complexes of SRPK1 (1 μM) and 250 nM cl-ASF(224) (●),cl-RS$_8$-RS$_8$(224) (▲),cl-RS$_8$-RA$_4$(224) (○) and cl-SA$_8$-RS$_8$(224) (Δ) were reacted with 100 μM [^{32}P]ATP and quenched with SDS loading buffer, and the total phosphoryl contents are plotted as a function of time. The amplitudes and rate constants for the initial phases are 5.5±0.48 and 5.4±1.0 min^{-1}for cl-ASF(224); 12±1.2 and 3.7±0.7 min^{-1}for cl-RS$_8$-RS$_8$(224); 8.2±1.9 and 3.8±1.9 min^{-1} for cl-RS$_8$-RA$_4$(224); and 5.3±0.31 and 6.3±0.93 min^{-1}for cl-SA$_8$-RS$_8$(224), respectively. All progress curves are performed in triplicate, and error bars for each time point are displayed.

showed C-terminal initiation preference. Together, the data suggested a strong link between initiation preference and consecutive stretch, as well as a tendency to phosphorylate most of the residues N-terminal of iniation region.

Additional studies were conducted with digestion mutants cl-RS$_8$-RS$_5$(224) and cl-RS$_8$-RS$_4$(224) to explore the minimal stretch of RS repeats preferred by SRPK1, and phosphoryl content determined by MALDI-TOF increased to 20 sites for both mutants as well (data not shown). However, cl-RS$_8$-RS$_4$(193) showed clear evidence for 5 sites phosphorylated outside the RS domain, indicating that the extra phosphoserines are not placed in RS2 as was expected, a phenomenon that needs to be further studied (data not shown). The investigation into the minimal stretch of RS/SR recognizable by SRPK1 ended with single mutations. Since cl-RS$_8$-RS$_4$ was generated with T248S and H243R double mutations, two individual mutations were carried out. However, these single mutation did not change the phosphoryl content up to 20 sites (18 and 16 sites, respectively, by MALDI-TOF), suggesting that the minimal RS/SR pair necessary for significant change in phosphorylation reaction termination is 4 (data not shown). It is important to note that with phosphorylation extending outside the RS domain, extensive mutagenesis may have created an artificial phenomenon so different from wt ASF that nothing more can be learned from it.

Finally, single-turnover kinetic analyses were performed to determine whether the RS domain mutations affect the rate of phosphorylation (Fig. 4.10D). No significant change in catalytic rate was observed, although the reaction end points varied according to total phosphorylatable series. Overall, the data implied that the preference for RS1 phosphorylation is dependent on a long uninterrupted RS/SR repeat, and SRPK1 may phosphorylate most residues N-terminal of the initiation site.

h.   **Phosphorylation Mechanism in RS2 Segment**

While C-terminal residues in the RS domain of wt-ASF/SF2 are slowly phosphorylated by SRPK1, the introduction of a long RS repeat in this region led to rapid and high-level phosphorylation (Fig. 4.11). In order to address whether changes in region-specificity can lead to changes in processive phosphorylation, start-trap experiments were conducted with N-terminal His-ASF constructs with similar
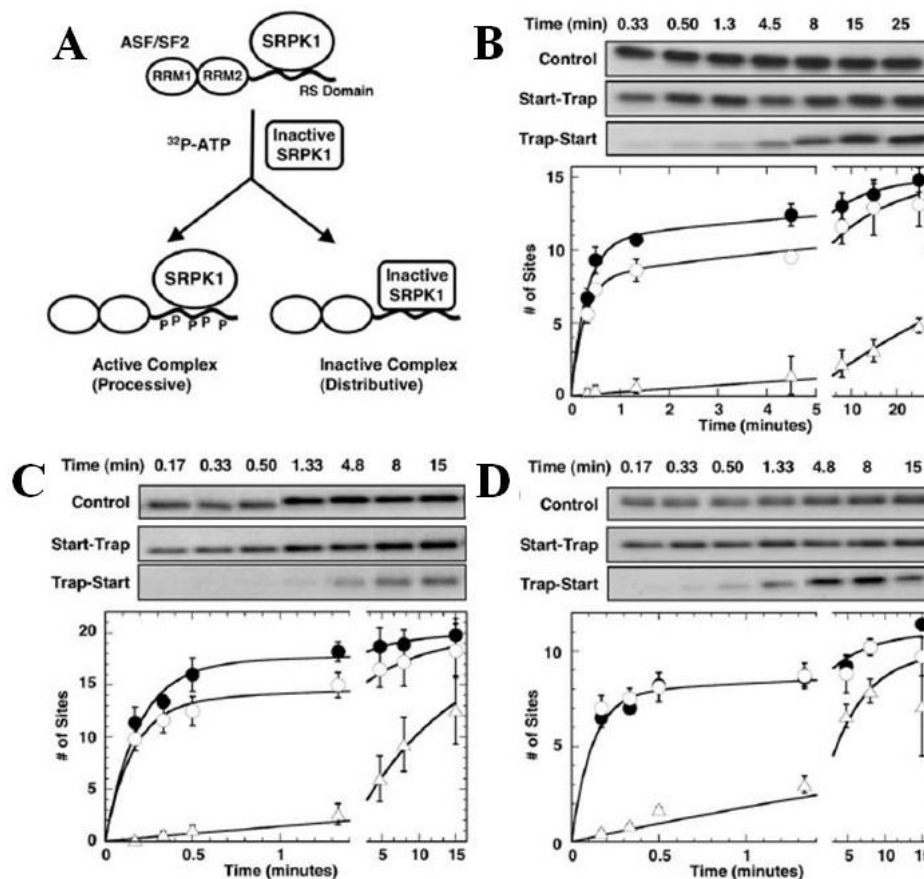
Figure 4.11 Processivity Maintained in ASF/SF2 with Modified RS Domains. A) Start-trap protocol. B-D) Start-trap experiments. SRPK1 (1 μM) was pre-incubated with 200 nM wt-ASF/SF2 (B), RS_8-RS_8 (C), and SA_8-RS_8 (D), and then the reaction was started with 100 μM [^{32}P]ATP in the absence (control; ●), and in the presence (start-trap; ○) of 60 μM kdSRPK1. In trap-start experiments, the complexes first mixed with kdSRPK1 before [^{32}P]ATP was added (Δ). The experiments were performed in triplicate, and error bars for each time point are displayed. One representative kinetic experiment for each substrate is shown in the gel panels. The amplitudes of the fast phases are 11 ± 0.6 and 8.0 ± 0.36 sites in the absence and in the presence of kdSRPK1 for wt-ASF/SF2; 17 ± 1.0 and 14 ± 0.80 sites in the absence and in the presence of kdSRPK1 for RS_8-RS_8; 7.8 ± 0.56 and 7.8 ± 0.48 sites in the absence and in the presence of kdSRPK1 for SA_8-RS_8.

altered RS content described above but no Lys-to-Arg mutations. For wt-ASF/SF2, SRPK1 rapidly phosphorylates about 11 serines in the first minute in the absence of kdSRPK1 (control) and 8 sites in the presence of kdSRPK1 added with ATP (Fig. 4.11B), suggesting that SRPK1 adds about eight phosphates processively to ASF/SF2 before dissociating from the RS domain. For the substrate containing an additional RS_8 stretch (RS_8-RS_8), a higher level of processivity is achieved, suggesting that added arginines have a beneficial effect on processive phosphorylation (Fig. 4.11C). Interestingly, removal of RS1 serines did not negatively impact processivity compared to wt-ASF/SF2 (Fig. 4.11D). Overall, the data implied that SRPK1 can support the processive phosphorylation of RS repeats anywhere in the RS domain, regardless of its distance relative to the RRMs.

C. Discussion

SRPK1 catalyzes a highly specialized multi-site phosphorylation reaction that is essential for the localization of SR proteins in the nucleus and subsequent splicing function [39]. Prior investigation reveals that SRPK1 places these phosphates in the RS domain of ASF/SF2 in a regiotemporal, directional, and processive manner. Together with the X-ray crystal structure of the SRPK1:ASF/SF2 complex, these studies showed that while the N-terminus of the RS1 segment (N'-RS1) is initially bound in the SRPK1 docking groove at the start of phosphorylation, this segment moves out of the groove as a function of C-terminal to N-terminal phosphorylation. This process not only requires movement of the RS domain but also induces changes in secondary structure within ASF/SF2. In place of N'-RS1, β4 of ASF/SF2 RRM2 domain unfolds and moves into the docking groove for the phosphorylation of residues at the N-terminal end of RS1 [52]. In this present study, a broad range of both local and non-local structural elements within SRPK1 and ASF/SF2 would be investigated in their roles in the observed residue preferences for SRPK1 and the sliding RS domain movement throughout phosphorylation progress.

a. Proper Alignment Promotes Efficient Phosphorylation of the RS1 Segment

SRPK1 phosphorylates RS1 directionally, and the X-ray crystal structure (Fig. 4.1) has suggested possible residue contacts between SRPK1 and RRM2 or the RS1 segment in ASF/SF2 that

could regulate this ordered mechanism. While mutations in RRM2 that are expected to disrupt interactions with SRPK1 have effects on reaction processivity, they do not impact the directional mechanism (Fig. 4.4). In contrast, the docking site defective mutant SRPK1(6M) not only is a fully distributive kinase but also displays a greatly relaxed preference for C->N phosphorylation (Fig. 4.2). This phenomenon is a direct result of a misaligned RS1 segment that can no longer bind dominantly in the manner presented in the X-ray structure but instead binds in different initiation modes where N'-RS1 is not consistently engaged in the docking groove at reaction start. Interestingly, while the docking groove controls directional phosphorylation, it plays no role in modulating RS1/RS2 regiospecificity (Fig. 4.2). Such findings support that other protein-protein interactions regulate this phenomenon.

b.   SRPK1 Uses a Distinct Pathway for Multi-site Phosphorylation

Directional phosphorylation has been observed in another protein kinase, GSK-3, where C-terminal to N-terminal movement is conducted through an electropositive pocket that binds a P+4 priming phosphoserine [105, 115]. Removal of one of the three positively charged residues in the P+4 pocket renders the mutant incapable of phosphorylating a primed substrate [116]. Although SRPK1 catalyzes its own priming step, there is an analogous P+2 eletropositive pocket in the crystal structure that may direct processive and directional phosphorylation. In contrast to GSK-3, removal of any or all of the electropositive residues in this P+2 pocket does not halt phosphorylation of the RS1 segment of ASF/SF2, suggesting that SRPK1 differs greatly from GSK-3 in phosphorylation mechanism (Fig. 4.3). Although the P+2 pocket does not impact directional phosphorylation either (Fig. 4.2), these residues serve a kinetic role in later steps of the reaction. While the wild-type enzyme readily completes the phosphorylation of the first 10 residues in RS1 in about 90 seconds, the P+2 mutants phosphorylate only half of these reisdues in this time (Fig. 4.3). Thus, the electropositive pocket in SRPK1 may balance negative charge in the RS domain late in the phosphorylation process.

c. Multiple Protein-protein Interactions Drive Orderly Phosphorylation

The SRPK1 docking groove has been shown to regulate directional phosphorylation of ASF/SF2, but the role of RRMs is still unclear since the X-ray structure lacks RRM1. To disrupt interactions between SRPK1 and the RRMs, RRM1 was deleted and RRM2 was moved relative to the RS domain to remove any specific contacts. Mutations within SRPK1 were considered and rejected because residues contacting RRMs are also catalytically important. Alanine mutations in two consecutive residues (Trp-His) in the glycine-rich loop, designed to disrupt one of the binding surfaces between RRM2 and SRPK1, results in a mutant kinase that is about 100-fold slower than wild-type SRPK1 at phosphorylating a short peptide substrate (Npl). This result is consistent with previously published results on glycine-rich loop mutants of protein kinase A and v-Fps that display very low phosphoryl transfer rates and bidning affinities for ATP [117, 118]. Such findings suggest that probing the role of docking residues involving catalytically important residues is likely to be problematic owing to secondary effects on phosphoryl transfer and/or nucleotide binding steps.

By considering RS domain interactions in the docking groove and those made with RRM2, the contributions of these distal contacts were analyzed in detail (Fig. 4.7). Although SRPK1 strongly prefers to start at the initiation box, mutation of the docking groove and removal of both RRMs completely eliminates this preference and imparts a slight preference for N-terminal initiation. Adding back either the docking groove or both RRMs increases the preference for C-terminal initiation by about 6-fold (Fig. 4.6), suggesting that either group of contacts participates equally in the initiation process. Although the role of RRM1 in the enzyme-substrate complex is not defined, this domain has a small impact on phosphorylation initiation (2-fold), suggesting that it could make some direct interactions with SRPK1 or possibily stabilize/strengthen RRM2 contacts with the enzyme. Nonetheless, no single contact accounts for the whole enhancement effect, and it is necessary to have both a functional docking groove and a neighboring RRM N-terminal to the RS domain to elicit strong C-terminal initiation preference. In fact, these binding elements are highly cooperative, as the addition of the docking groove and the RRM contacts increases C-terminal preference by about 160-fold. A very interesting outcome of this analysis is that whatever the means by which RRM2 imposes a directional bias towards the C-terminal region of

RS1, the molecular mechanism appears highly flexible. The removal of all four contact residues in RRM2 does not significantly impact C-terminal initiation. SRPK1 may recognize general electrostatic residues on the RRM surface rather than discrete residues, but the X-ray structure of the mutated RRM2 is necessary to search for alternative binding sites.

d. Additional Structural Elements Investigated

In addition to RRM1, the X-ray structure of the SRPK1:ASF/SF2 complex lacks the RS2 segment of ASF/SF2 and the spacer domain of SRPK1, and potential contacts may exist in those absent parts of the complex. Deletion mutants for both were created and investigated with the ATP limitation assay, and neither one impacts C-terminal to N-terminal directional phosphorylation (Fig. 4.8). Interestingly, the SRPK1 phosphorylation progress curve of RS2 domain deletion mutant constructs clearly fit to a fast-slow bi-phase, even if RS1 domain is phosphorylated predominantly in the fast phase (10-12 sites) for wild-type ASF/SF2. Deletion of RS2 segment may not impact phosphoryl content of RS1, but RS2 segment may be important in maintaining a certain conformation that assists in rapid phosphorylation of the RS1 segment. Additional studies are necessary to elucidate the structure of the RS domain with and without RS2 segment before any conclusion could be drawn.

e. SRPK1 Substrate Recognition Motif Investigated

SRPK1 phosphorylates the RS1 segment up to 12 sites very rapidly and the RS2 segment for 3-4 sites very slowly, raising the question of how certain RS/SR pairs are selected. Two different approaches were taken to investigate this phenomenon. First, prolines flanking the RS1 segment have been mutated to alanine in different combinations, because prolines may help present a protein segment for protein-protein interactions through turn motifs. However, the phosphoryl content, binding affinity, and phosphorylation kinetics do not exhibit any changes upon mutation (Fig. 4.9). The addition of peptidyl prolyl isomerase Pin1 to unphosphorylated ASF/SF2 does not change phosphoryl content either, but the progress curve is best fitted to a single-exponential, suggesting that RS2 phosphorylation may be affected

by Pin1. Given that most of the phosphorylatable serines are in RS1, prolines play, at most, a minor role in the SRPK1 phosphorylation mechanism.

While the RS domain is rich in Arg-Ser repeats, they are not distributed evenly. The RS1 segment has 8 consecutive RS/SR pairs, while RS2 segment has only 3. It was discovered that the kinetic preference of RS1 over RS2 (~20-fold) is linked directly to the overall length of the RS/SR repeat. Inserting long RS repeats into the C-terminal end of the RS domain leads to rapid high-level phosphorylation of RS2 (~20 sites total *versus* 15 sites for wt) that no longer displays preference for RS1 (Fig. 4.10), suggesting that the initiation box has expanded drastically. Furthermore, these modified RS/SR repeats in the RS2 can be processively phosphorylated to the same extent as RS1 (Fig. 4.11), implying that there is no requirement for a directly collinear RRM and RS/SR repeat.

The next question to address is what is the exact sequence requirement for RS2 domain phosphorylation? Interestingly, phosphorylation of RS2 is indeed dependent on the continuity of the RS/SR repeat, as four alanine mutations at every other arginine disable most modifications in RS2. It is also worth remarking that the same alternating arginine mutation in RS1 while leaving RS2 a continuous stretch of RS pairs leads to RS2 initiation and high-level phosphorylation, implying that SRPK1 will phosphorylate any serine residues N-terminal of the initiation region. Another layer of complexity is added here, as cl-RS_8-RS_4(193) shows that $(RS)_4$ is the minimal requirement for high-level phosphorylation, but up to 5 sites are phosphorylated outside RS domain. Previous investigations have shown that SRPK1 does not phosphorylate outside the RS domain, where an RS pair can be found only once. At some point the mutations in RS2 segment may have led to conformational changes that elicit distal effects too far removed from wild-type phosphorylation mechanism, and care must be taken to limit the conclusion that can be drawn from the tangle of data. Finally, SRPK1 does not require an N-terminal flanking RRM for regiospecific preference for RS1 over RS2 segment or rapid overall phosphorylation, as the swap substrate cl-RS-RRM2(224) is readily modified despite possessing a C-terminal RRM. Overall, the data suggest that SRPK1 kinetically favors certain regions owing to a high consecutive RS/SR content.

f.   Conclusion

Together with observations from Chapter 3, an updated phosphorylation model can be assembled in Fig. 4.12.   RRM2 and the N-terminus of the RS1 segment directs processive, C->N directional phosphorylation of the RS1 segment by SRPK1, followed by dissociation of the enzyme/substrate complex and slow phosphorylation of the RS2 segment upon rebinding.   RRM2 forms nonspecific contacts with SRPK1, while the N-terminus of the RS1 segment forms contacts with the SRPK1 docking groove.   These two groups act interdependently to enforce high regiospecific preference for initiation in the initiation box.   We also discovered the additional requirement that SRPK1 prefers long, uninterrupted RS/SR stretch for regiospecific phosphorylation initiation, thus explaining the RS1/RS2 substrate specificity.   Having made extensive progress in characterizing the mechanism of phosphorylation and relevant molecular interactions for the SRPK1:ASF/SF2 complex, in the next chapter we will focus on the next stage of the phosphorylation cycle.

Chapter 4 is, in part, a reprint of the material as it appears in *The Journal of Molecular Biology*, 2008, 328(4), 894-909, Hagopian, J.C., Ma, C.T., Meade, B.R., Albuquerque, C.P., Ngo, J., Ghosh, G., Jennings, P.A., Fu, X.D., Adams, J.A.   The dissertation author was a collaborative researcher and co-author of this publication.   Additional text of Chapter 4 is a reprint of the material as it appears in *The Journal of Molecular Biology*, 2009, 390(4), 618-34, Ma, C.T., Hagopian, J.C., Ghosh, G., Fu, X.D., Adams, J.A. The dissertation author was the primary researcher and author of this publication.
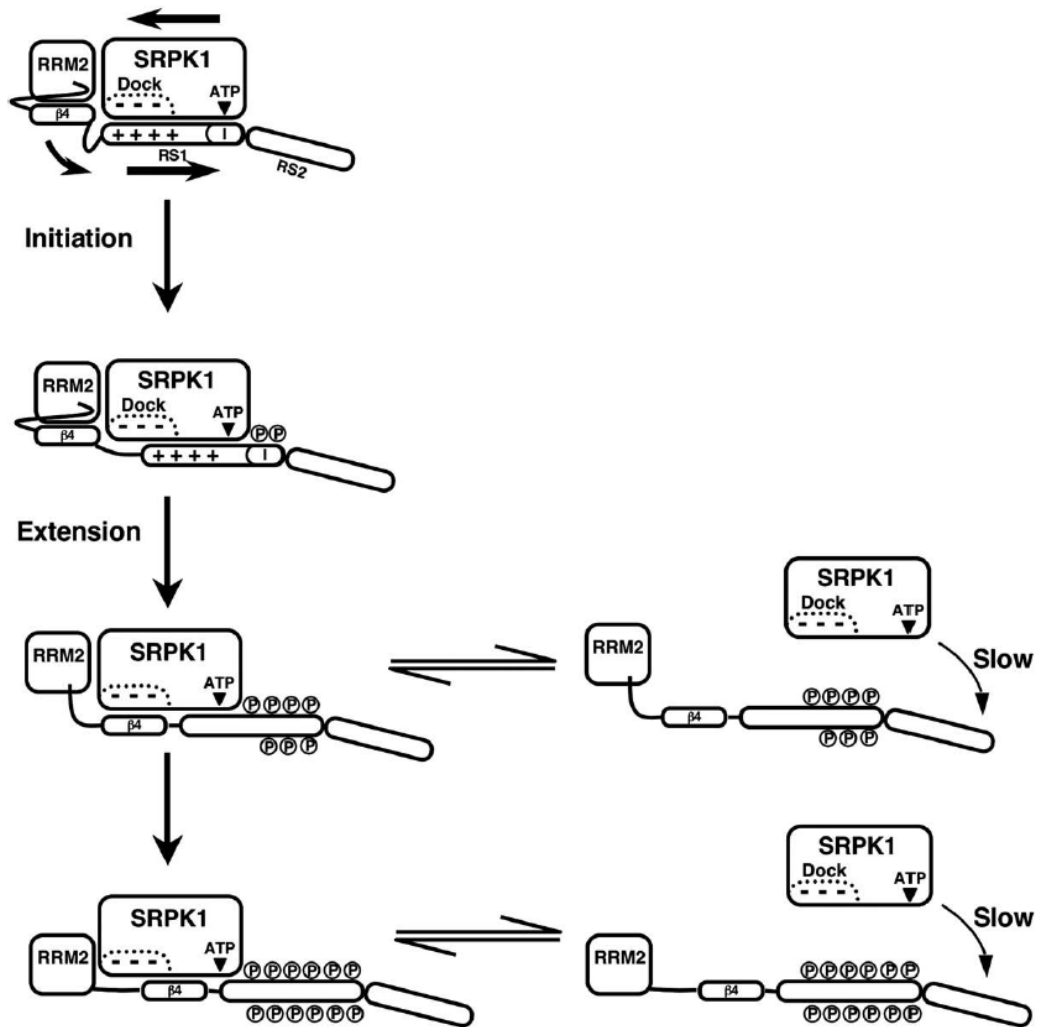
Figure. 4.12 Proposed Mechanism for Regiospecific Phosphorylation Control. SRPK1 forms a high affinity complex with ASF/SF2 and starts the phosphorylation cycle in the "initiation box" at the C-terminal end of RS1(I). Seven to eight serines are modified in the extension phase before the stability of the enzyme-substrate complex diminishes and before SRPK1 can dissociate from and rebind to modify the remaining serines in RS1 or begin to modify serines in RS2. RRM2 and N'-RS1 interact with SRPK1 to control initiation of phosphorylation.

Chapter 5

Dephosphorylation Mechanism of ASF/SF2

A. Introduction

While phosphorylation of the SR proteins is necessary for spliceosome assembly, dephosphorylation is a key step in splicing catalysis and mature mRNA nuclear export. In conjunction with the splicing kinases SRPK1 and Clk/Sty, nuclear phosphatases such as PP1 play an important role modulating the phosphoryl content of the SR proteins. Despite the importance of SR protein dephosphorylation, very few mechanistic studies have been performed on the PP1 reaction in context of the phosphatases associated with splicing [70, 71]. In this chapter we will investigate the dephosphorylation of the prototypical SR protein ASF/SF2 by PP1 providing insights into the nature of the enzyme:substrate interaction, the role of the RRMs and the pathway by which phosphates are removed from the RS1 segment of the RS domain. We will also investigate the effects of phosphorylation/dephosphorylation on ASF/SF2 structure using circular dichroism spectrometry.

In mammalian tissues there are three serine-threonine phosphatase PP1 genes and four individual PP1 isoforms. For this study we will focus on PP1γ, since this isoform is commonly used in cell-based studies of PP1 and SR protein interactions and can be readily expressed and purified for mechanistic studies. The goals of the present study was to characterize binding affinity and mechanism of dephosphorylation of phosphorylated ASF/SF2 by PP1, including the order of dephosphorylation. As we have shown previously using SRPK1 that the phosphorylation reaction occurs in a specifc C-terminal to N-terminal direction, we would now like to address whether these phosphates are removed in a similar order. Using the different isoform PP1α in parallel will allow us to test potential difference in the mechanism of dephosphorylation across isoforms.

B. Results

a. Change in ASF/SF2 Secondary Structure upon Phosphorylation

The recent X-ray structure of the SRPK1:ASF/SF2 complex and additional cross-linking studies suggest that phosphorylation leads to movement of ASF/SF2 within the complex, most likely characterized by the unpacking of the β4 strand of ASF/SF2 RRM2 domain and its subsequent binding to the SRPK1 docking groove [52]. In order to observe any potential changes in secondary structure

associated with movement of ASF/SF2 relative to SRPK1, circular dichroism far UV spectra were recorded for unphosphorylated and phosphorylated ASF/SF2, both with SRPK1 and Clk/Sty. The MOPS (pH 7.4) buffer used for most phosphorylation reactions was deemed unsuitable for circular dichroism studies due to a significant background absorbance of the buffer. To avoid this problem, wt-ASF/SF2 was dialyzed into acetate buffer (pH 4.5) which showed much less interference (Fig. 5.1A). Given this lower background, acetate buffer was used for all circular dichroism studies at pH 5.5, the highest pH in which SRPK-phosphorylated ASF/SF2 is soluble. Single-turnover kinetics were performed to ensure that SRPK1 can phosphorylate ASF/SF2 and the RS domain construct [ASF(188-248)] at roughly the same catalytic rate in this new assay buffer (Fig. 5.1B). To minimize spectral interference from SRPK1, a ratio of 50:1 ASF/SF2:SRPK1was used to phosphorylate the RS domain of ASF/SF2 under steady-state kinetic conditions (Fig. 5.1C). Similar single-turnover and steady-state kinetic time courses were also conducted to confirm that Clk/Sty can phosphorylate ASF/SF2 and ASF(188-248) efficiently in acetate buffer pH 5 (Fig. 5.1D), the highest pH in which Clk-phosphorylated ASF/SF2 is soluble.

Having optimized the experimental conditions, CD spectra of mock phosphorylated (by AMP-PNP, a substrate-analog inhibitor) and phosphorylated RS domain were recorded. The absence of any apparent changes in the two spectra suggested that no significant secondary structure change occurs (Fig. 5.1E). In comparison, mock phosphorylated and phosphorylated ASF/SF2 showed significant differences in circular dichroism spectra, suggesting that phosphorylation-dependent conformational change involving the RRM domains had occurred. The phosphorylated ASF/SF2 spectra showed a lower ellipticity value, suggesting that the secondary structure has changed. For Clk/Sty phosphorylation, similar patterns were observed in which RS domain showed no change in secondary structure, while ASF/SF2 shows noticeable reductions in ellipticity value upon phosphorylation (Fig. 5.1F). Finally, circular dichroism spectra were taken on the same samples of unphosphorylated and SRPK1-phosphorylated ASF/SF2 over three consecutive days. Over this time frame the samples were normalized by concentrations determined by A_{280} absorbance and Bradford assay. The curve representing phosphorylated ASF/SF2 (Fig. 5.1G) approached the curve representing unphosphorylated ASF/SF2 over
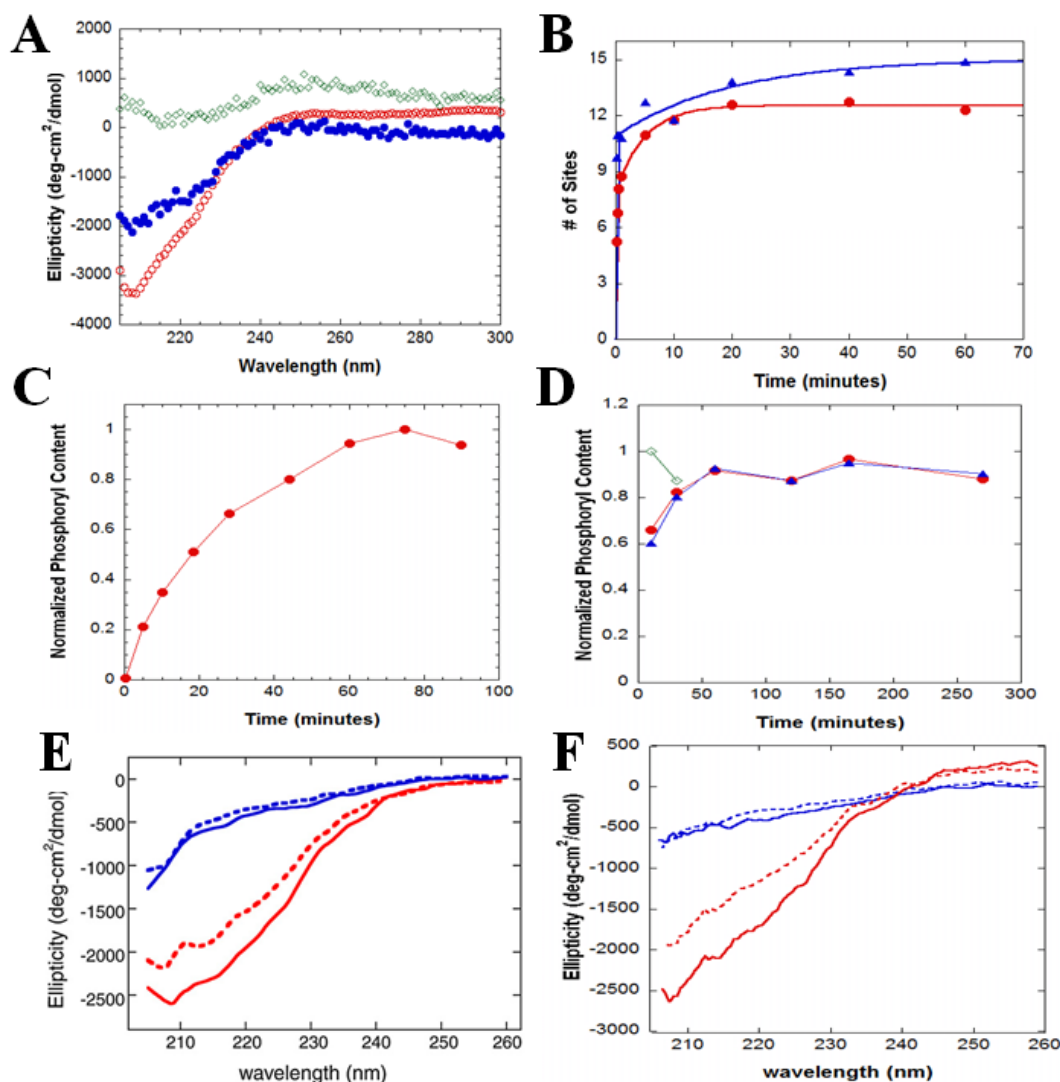
Figure 5.1 CD Spectra of ASF/SF2 with and without Phosphorylation. A) Buffer selection. 1 μM wt-ASF was mixed with 30 nM SRPK1 and then the reaction was initiated with 13 μM ATP for one hour to generate phosphorylated ASF. A portion of unphosphorylated wt-ASF was dialyzed into acetate buffer at pH 4.5 for buffer exchange. Circular dichroism spectra are displayed for wt-ASF (●), wt-ASF phosphorylated (pASF) (◊), and wt-ASF dialyzed (○). B) Single-turnover kinetics in different buffers. SRPK1 (1 μM) was pre-incubated with 200 nM ASF(188-248) in MOPS pH 7.4 buffer (▲) or acetate pH 5.5 buffer (●), and then the reaction was started with 100 μM [^{32}P]ATP. The data are fitted to a double-exponential function. The reaction rate constants are 5.8 and 0.2 and amplitudes are 7.9 and 4.6 sites for fast and slow phase in acetate buffer. The reaction rate constants are 13.3 and 0.05 and amplitudes are 10.9 and 4.1 sites for MOPS buffer. C) Steady-state kinetics. SRPK1 (90 nM) was pre-incubated with 4.8 μM ASF(188-248) in acetate buffer, and then the reaction was started with 100 μM [^{32}P]ATP. D) Kinetic analyses for Clk/Sty. 4.8 μM ASF(188-248) was mixed with 160 nM Clk/Sty (catalytic 1:30,), 96 nM Clk/Sty (catalytic 1:50,), or 4.8 μM Clk/Sty in acetate buffer pH 5 and then the reaction was started with 100 μM [^{32}P]ATP. E) CD spectra with SRPK1. ASF/SF2 (red) and the RS domain [ASF(188-248)] were incubated with SRPK1 in the presence of AMP-PNP (solid lines) and ATP (dotted lines), and CD spectra are displayed between 205 and 260 nm. F) CD spectra with Clk/Sty. ASF/SF2 (red) and the RS domain [ASF(188-248)] were incubated with SRPK1 in the presence of AMP-PNP (solid lines) and ATP (dotted lines), and CD spectra are displayed between 205 and 260 nm.

Figure 5.1 (Continued)
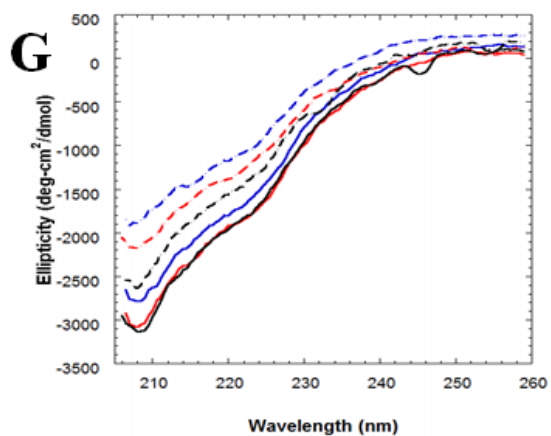
G) CD spectra with SRPK1. ASF/SF2 was incubated with SRPK1 in the presence of AMP-PNP (solid lines) and ATP (dotted lines) for 1 hour, and CD spectra are displayed between 205 and 260 nm on the same day of the experiment (blue), the next day (red), or the third day (black).

time, suggesting that the substrate may undergo a conformational change on a time scale of days. An alternative explanation would be that phosphorylated ASF/SF2 is naturally dephosphorylated, although we have not been able to substantiate this claim using autoradiographic techniques.

b. Assay Optimization for PP1α

SRPK1-phosphorylated wt-ASF/SF2 was dephosphorylated about 4-fold faster in MOPS SRPK1 buffer than the NEB phosphatase buffer (HEPES pH 7.5, see Chapter 2 section E) from New England Biolab (Fig. 5.2A), but the phosphorylated wt-ASF was more stable over time in NEB buffer, and subsequent reactions were, therefore, carried out in this buffer. PP1 dialysis buffer (see Chapter 2 section C) was used in place of PP1 enzyme to ensure volume and salt uniformity across samples, and ASF/SF2 concentration was monitored throughout the reaction to confirm the stability of the substrate over time. The no-enzyme control was always performed for substrates tested the first time, although it was not always shown beyond Figure 5.3B. In general, 8-10 sites were phosphorylated in ASF/SF2 at 7 μM ATP, and on later figures the Y-axis will be represented by normalized phosphoryl content or number of sites for simplicity. For reaction quenching, 20 mM vanadate was tested as a stopping agent and the results showed effective inhibition (Fig. 5.2B) using 125 to 250 nM PP1α. Small differences in phosphoryl contents on ASF/SF2 using this procedure did not impact initial reaction rates, as evidenced in roughly the same slope values for all three curves in their initial linear region (Fig. 5.2C).

c. RRMs Protect ASF/SF2 from Dephosphorylation by PP1α

Having established optimal conditions for the dephosphorylation assay, the rates of dephosphorylation of various mutant ASF/SF2 constructs were assessed to explore potential important molecular interactions. To test for possible proline-associated conformational changes upon phosphorylation of vicinal serines, the dephosphorylation rates of a proline-mutant [ASF(4PA)] and wt-ASF in the presence of peptidyl-prolyl isomerase Pin1 were studied. We found that neither the mutant ASF nor the administration of Pin1 to wt-ASF significantly change the dephosphorylation rates of the RS domain (Fig. 5.2D). To test the role of N-terminal sequences, several RRM deletion mutants of ASF/SF2
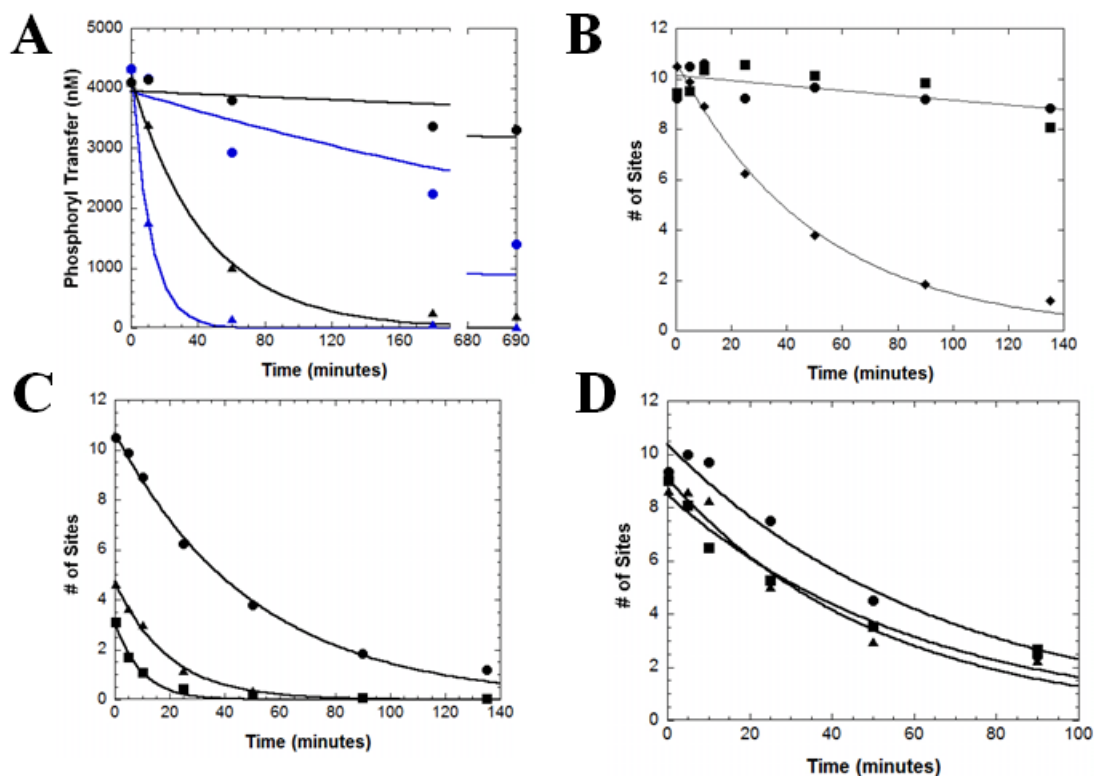
Figure 5.2 Dephosphorylation of ASF/SF2 by PP1α. A) Buffer comparison for dephosphorylation experiment. 0.5 μM wt-ASF/SF2 was pre-phosphorylated with stoichiometric amount of SRPK1 and 7 μM [^{32}P]ATP for 20 minutes, in MOPS (pH 7.4) buffer (blue) or in NEB (Tris pH 7.5) buffer (black). Near-stoichiometric amount (250 nM) of PP1α (▲) or equal amounts of PP1dialysis buffer control (●) was added to initiate phosphatase reaction at 37 ºC, and the reaction was quenched with SDS-PAGE loading buffer at different time points. Data are fitted to single-exponential decay equation. For 250 nM PP1α, rate constants are 0.090 and 0.022 for MOPS and NEB buffer, respectively. For no-enzyme buffer control, rate constants are 0.00030 and 0.0022 for MOPS and NEB buffer, respectively. B) Phosphatase inhibitor tested. 0.5 μM wt-ASF/SF2 was pre-phosphorylated with stoichiometric amount of SRPK1 and 7 μM [^{32}P]ATP for 20 minutes, in NEB buffer. PP1α (125 nM) (♦), PP1 dialysis buffer (●), or PP1α plus 20 mM vanadate (■) were added to the pre-phosphorylated reaction mixture and the reaction was quenched with SDS-PAGE loading buffer at different time points. Data are fitted to single-exponential decay equation. Rate constants are 0.020 and 0.0010 for PP1α and vanadate-stopped PP1α, respectively. C) Phosphoryl content does not impact dephosphorylation rate. 0.5 μM wt-ASF/SF2 was pre-phosphorylated with stoichiometric amount of SRPK1 and 7 μM (●), 3.5 μM (▲), or 2 μM (■) [^{32}P]ATP for 20 minutes, in NEB buffer. PP1α (125 nM) was added to the pre-phosphorylated reaction mixture and the reaction was quenched with SDS-PAGE loading buffer at different time points. Data are fitted to single-exponential decay equation. Amplitudes are 10.7, 4.71 and 3.10 and rate constants are 0.020 and 0.051, and 0.10 for 7 μM, 3.5 μM, or 2 μM [^{32}P]ATP, respectively. D) Proline effects. 0.5 μM wt-ASF/SF2 (●) or ASF(4PA) (■) was pre-phosphorylated with stoichiometric amount of SRPK1 and 7 μM [^{32}P]ATP for 20 minutes, in NEB buffer. Part of wt-ASF/SF2 reaction mixture was pre-incubated with 60 nM Pin1 (▲) for 20 minutes. All samples were treated with PP1α (125 nM) and the reaction was quenched with SDS-PAGE loading buffer at different time points. Data are fitted to single-exponential decay equation. Rate constants are 0.015, 0.020, and 0.016 for wt-ASF/SF2, wt-ASF/SF2 plus Pin1, and ASF(4PA), respectively.

Figure 5.2 (Continued)
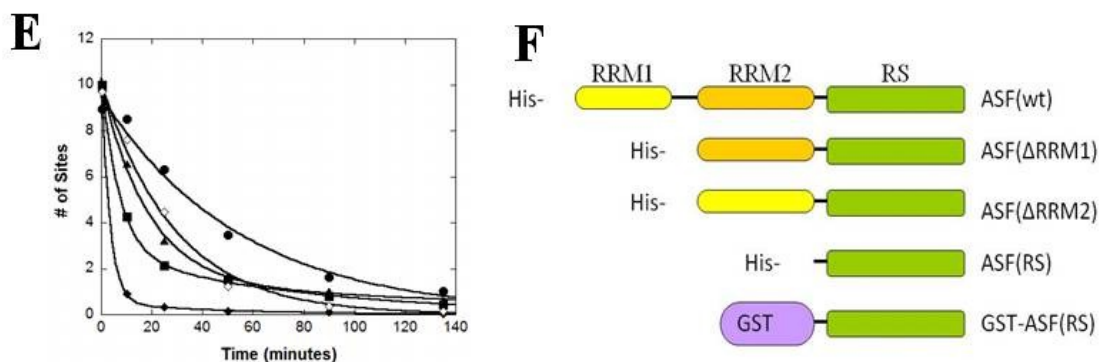
E) RRM effects. 0.5 μM wt-ASF/SF2 (●), ASF(ΔRRM1) (■), ASF(188-248) [also known as ASF(RS), ♦], ASF(ΔRRM2) (▲), or GST-RS (◊) was pre-phosphorylated with stoichiometric amount of SRPK1 and 7 μM [^{32}P]ATP for 20 minutes, in NEB buffer. All samples were treated with PP1α (125 nM) and the reaction was quenched with SDS-PAGE loading buffer at different time points. Data are fitted to single-exponential decay equation for wt-ASF/SF2 and GST-RS. Rate constants are 0.018 and 0.034, respectively. Data are fitted to double-exponential decay equation for ASF(ΔRRM1), ASF(188-248), and ASF(ΔRRM2). Amplitudes of the fast and slow phases are 7.47 and 2.71, 9.77 and 0.52, 8.62 and 1.42, respectively. Rate constants are 0.14 and 0.013, 0.30 and 0.022, 0.054 and 0.0055, respectively. F) Domain organization of RRM deletion mutants.

were also tested. We found that removal of RRM1, RRM2 and both RRM1 and RRM2 increased the dephosphorylation rate by 8-, 2-, and 17-fold, respectively (Fig. 5.2E). Interestingly, the rate of a GST-tagged RS domain was increased by only 2-fold, suggesting that the dephosphorylation rate enhancement may be partly due to steric hindrance of domains N-terminal to the RS domain, as opposed to specific contacts formed by PP1α and the RRM domains (Fig. 5.2E). It is also important to remark that, unlike wt-ASF/SF2 and GST-RS, deletion of RRM domains in Histidine-tagged constructs led to biphasic dephosphorylation, implying that this rapid dephosphorylation may be followed by dissociation of the enzyme:substrate complex similar to SRPK1. Overall, the data suggested that RRM1, and RRM2 to a minor extent, exert a protective effect against PP1α dephosphorylation.

d.   RRM1 Regulates Directional Dephosphorylation of ASF/SF2

To determine whether PP1 removes phosphates from the RS domain of ASF/SF2 in a directional manner, we employed the digestion construct approach developed in the preceding chapters. cl-ASF(214) was used in dephosphorylation assay with vanadate as stopping buffer to test for possible directional dephosphorylation (Fig. 5.3A). Vanadate did not change the LysC digestion rate at 20 mM. Interestingly, the rate of dephosphorylation of cl-ASF(214) was about 3-fold faster than wt-ASF/SF2, suggesting that Lys-to-Arg mutations may have a slight influence on RRM and RS structures (Fig. 5.3B). Nevertheless, we did not consider this a significant deviation from wt-ASF/SF2. The autoradiogram clearly showed that the phosphoryl content of the N-terminal fragment upon LysC digestion decreased faster than that of the C-terminal fragments, with the N/C ratio calculated from relative intensities decreased with time (Fig. 5.3C).   At 3 remaining phosphoserines the preference for N-terminus dephosphorylation was about 8 to 9 fold, weaker than SRPK1 preference for C-terminus initiation in the RS1 segment. Overall, the data implied that PP1α prefers to initiate dephosphorylation in the N-terminus of the RS1 segment.

PP1α is a commercially available isoform for protein phosphatase 1, and a series of simple studies were conducted with this enzyme. However, the high price, batch-dependent catalytic activity, and loss of mutagenesis capability prevented further in-depth studies into dephosphorylation mechanisms.
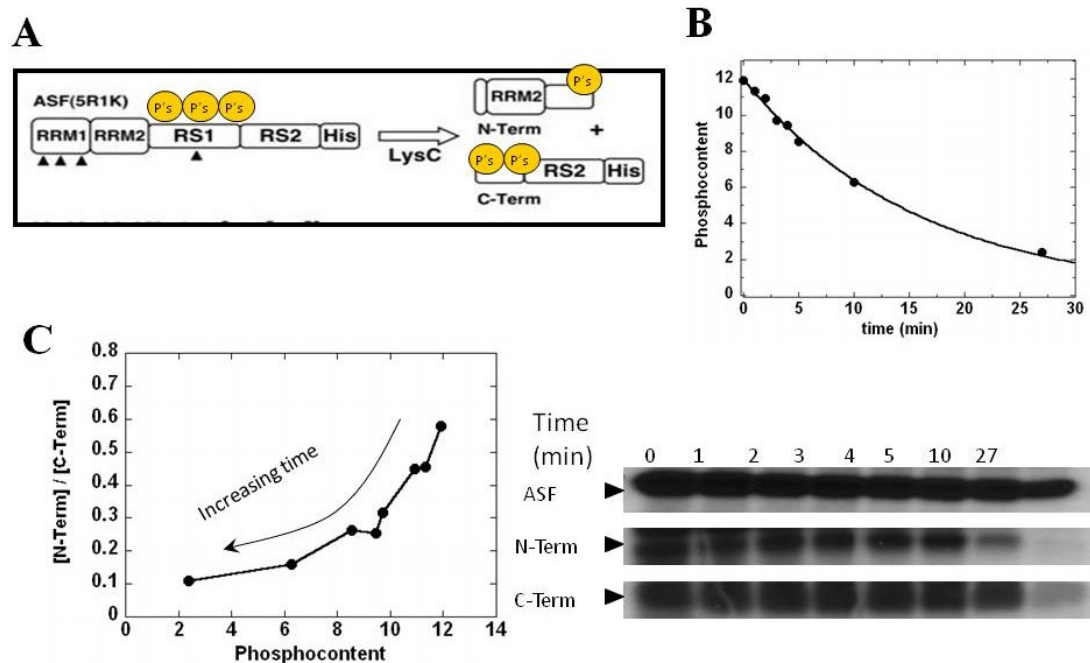
Figure 5.3 Directional Dephosphorylation of ASF/SF2 by PP1α. A) Expected LysC digestion profile. B) Time course. 0.5 µM cl-ASF(214) was pre-phosphorylated with stoichiometric amount of SRPK1 and 7 µM [^{32}P]ATP for 20 minutes in NEB (Tris pH 7.5) buffer. PP1α (125 nM) was added to initiate phosphatase reaction at 37 ºC, and the reaction was quenched with 20 mM vanadate at different time points. All samples were quenched in SDS-PAGE buffer right before gel loading. Data are fitted to single-exponential decay equation, with rate constant of 0.063. C) N/C ratio and digestion profile. Samples from (B) were treated with LysC (5 ng/µL) for 2 hours (majority of samples) or 1 hour (10-, 27- minute) at 37 ºC and then separated by SDS-PAGE and developed into autoradiograms. N/C ratio is calculated from phosphoryl content of N-terminal and C-terminal fragments and plotted against total phosphoryl content determined from undigested samples.

Nevertheless, the data suggested that RRM1 protects ASF/SF2 from dephosphorylation by PP1, and PP1 dephosphorylation proceeds in N->C direction.

e. Assay Optimization of PP1γ (buffer)

Using a bacterial plasmid construct we were able to express and purify PP1γ, opening up the possibility of greater mechanistic investigations. For our studies, the concentration of the purified enzyme was determined by Bradford assays and by A_{280} absorbance. To begin the study of PP1γ, a set of control experiments were performed to optimize and validate the dephosphorylation assay. First the no-enzyme control was performed to ensure stability of the substrate for the time frame under consideration. Phosphorylated ASF/SF2 was stable in Tris NEB buffer for up to 2 hours with minimal degradation, while significant dephosphorylation occurred even at a non-stoichiometric PP1 concentration (Fig. 5.4A). The raw phosphoryl transfer counts indicated that 9-10 sites of 500 nM wt-ASF/SF2 were phosphorylated at the end of pre-phosphorylation reaction and that these phosphates were readily removed by the recombinant PP1, similar to the commercial enzyme isoform (Fig. 5.4A).

The effects of $MnCl_2$ concentration were also investigated in terms of dephosphorylation rate. Previous studies showed that PP1γ expressed in *E. coli* requires $MnCl_2$ for enzymatic activation, and commonly 1-2 mM are used for maximal activity [119, 120]. Using NEB buffer (Tris pH 7.5), various amounts of $MnCl_2$ were added during the pre-phosphorylation stage, and, interestingly, at the lowest $MnCl_2$ concentration (0.4 mM, the lowest concentration possible because PP1γ stock solution contains $MnCl_2$) the rate of dephosphorylation was highest and the initial phosphoryl content was a little higher (Fig. 5.4F). For the rest of the $MnCl_2$ concentrations tested, PP1γ activity and initial phosphoryl content catalyzed by SRPK1 remained roughly the same, and 1.4 mM was chosen as the standard phosphatase buffer based on the stability of the dephosphorylation rate and protocols from previous literatures.
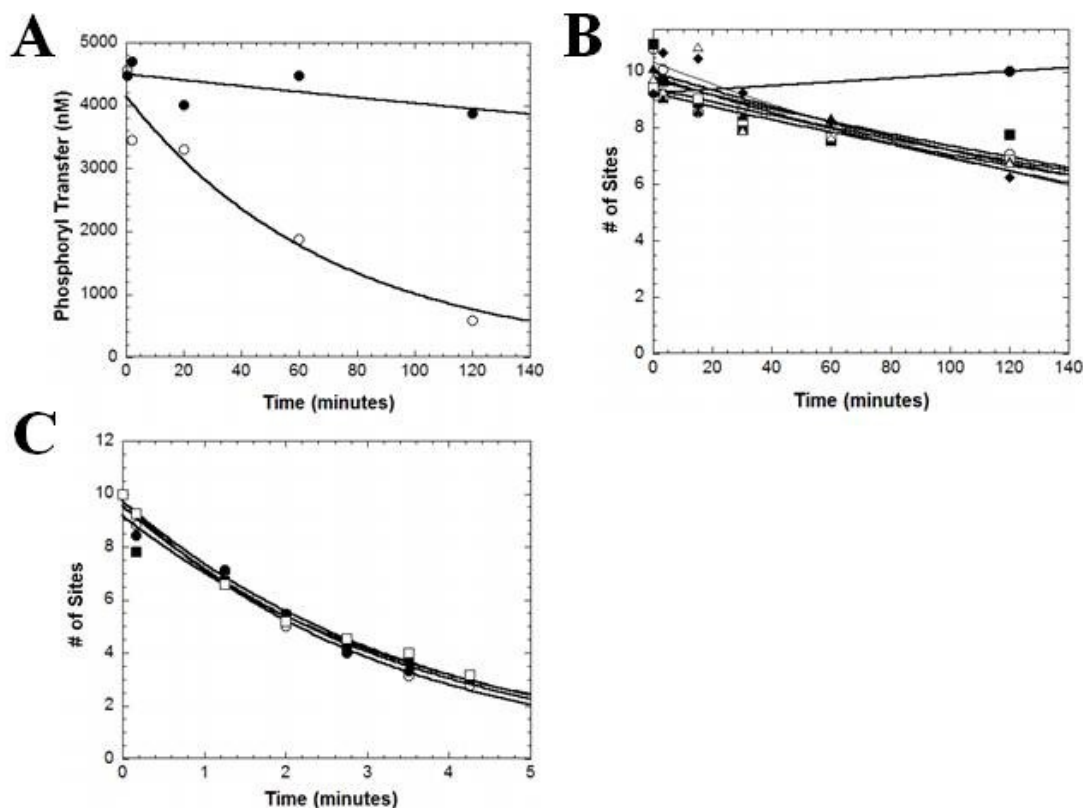
Figure 5.4 Dephosphorylation of Phosphorylated ASF/SF2 by PP1γ. A) No-enzyme control. 0.5 μM wt-ASF/SF2 was pre-phosphorylated with stoichiometric amount of SRPK1 and up to 8 μM [^{32}P]ATP for 25 minutes in NEB (Tris pH 7.5) buffer. PP1α (125 nM) (○) or equal amounts of PP1dialysis buffer control (●) was added to initiate phosphatase reaction at 37 ºC, and the reaction was quenched with SDS-PAGE loading buffer at different time points. Data are fitted to single-exponential decay equation. Rate constants are 0.014 and 0.0011 for 125 nM PP1γ and no-enzyme control, respectively. B) Temperature dependence. 0.1 μM wt-ASF/SF2 was pre-phosphorylated with catalytic amount of SRPK1 (20 nM) and up to 2 μM [^{32}P]ATP for 120 minutes in NEB (Tris pH 7.5) buffer. PP1γ at 3 μM (○), 2 μM (■), 1 μM (□), 500 nM (▲), 250 nM (Δ), 125 nM (♦), or equal amounts of PP1dialysis buffer control (●) were added to initiate phosphatase reaction at 37 ºC, and the reaction was quenched with SDS-PAGE loading buffer at different time points. Data are fitted to single-exponential decay equation. Rate constants are 0.00322, 0.0027, 0.0027, 0.0026, 0.0035, and 0.0039 for PP1γ (3 μM to 125 nM), respectively. For no-enzyme control, the data are fitted linearly with a slope of 0.00063. C) Phosphatase inhibitor test. 0.1 μM cl-ASF(214) was pre-phosphorylated with catalytic amount of SRPK1 (20 nM) and up to 2 μM [^{32}P]ATP for 120 minutes in NEB (Tris pH 7.5) buffer. 3 μM PP1γ was added to initiate phosphatase reaction at 37 ºC, and the reaction was quenched with SDS-PAGE loading buffer (●), 33 mM vanadate (○), 16.5 mM vanadate and 10 mM β-glycerolphosphate (■), or 33 mM vanadate and 20 mM β-glycerolphosphate (□), at different time points. Data are fitted to single-exponential decay equation. Rate constants are 0.29, 0.31, 0.26, and 0.28 for the various stopping agents.

Figure 5.4 (Continued)
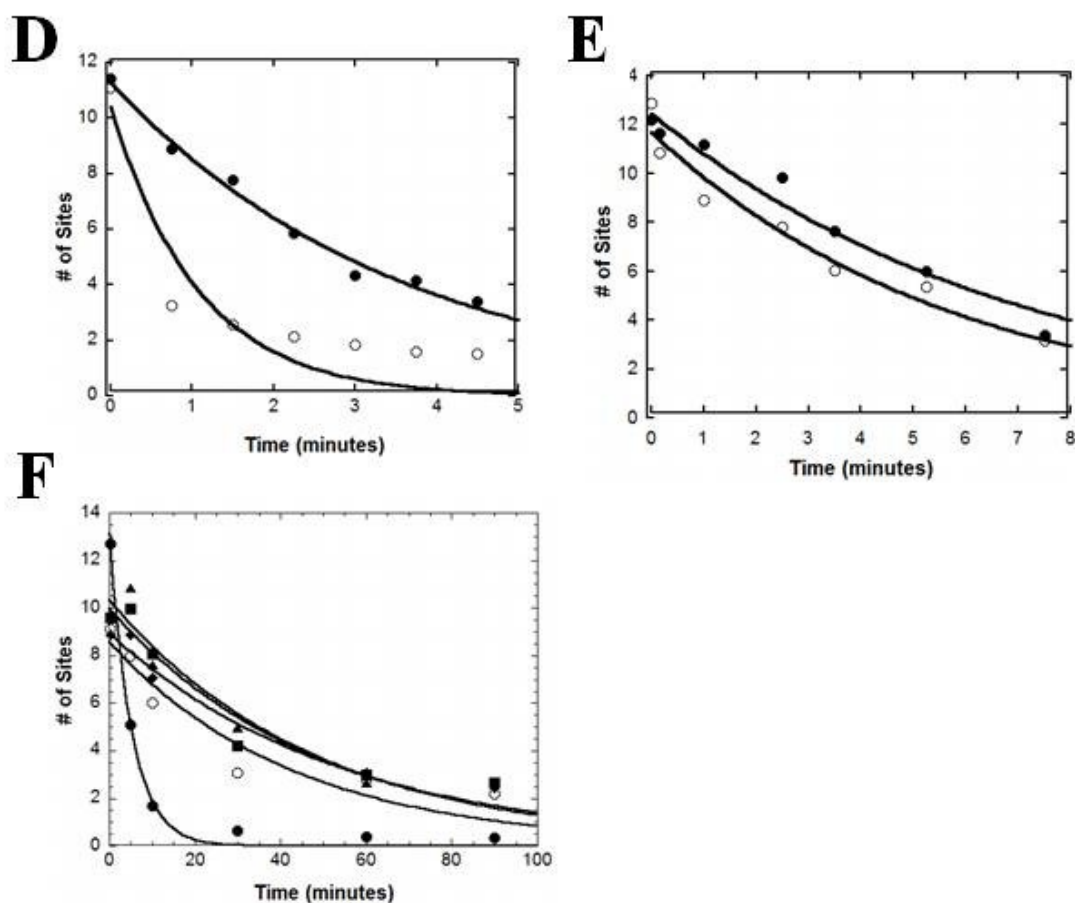
D & E) LysC digestion accelerates dephosphorylation. 0.1 μM cl-ASF(214) was pre-phosphorylated with catalytic amount of SRPK1 (20 nM) and up to 2 μM [^{32}P]ATP for 120 minutes in NEB (Tris pH 7.5) buffer. 3 μM (D) or 250 nM (E) PP1γ was added to initiate phosphatase reaction at 37 ℃, and the reaction was quenched with SDS-PAGE loading buffer (●) or 33 mM vanadate and 20 mM β-glycerolphosphate (○) at different time points. Samples stopped with vanadate were treated with LysC for up to two hours and all fragments identified on autoradiogram were collected for total phosphoryl content. Data are fitted to single-exponential decay equation. Rate constants are 0.28 and 0.93 (D) or 0.14 and 0.17 (E) for undigested and LysC-digested samples. F) MnCl$_2$ effects. 0.1 μM wt-ASF/SF2 was pre-phosphorylated with catalytic amount of SRPK1 (20 nM) and up to 2 μM [^{32}P]ATP for 120 minutes in Tris pH 7.5 buffer plus 0.4 mM MnCl$_2$ (●), 1.4 mM (▲), 2.4 mM (■), 4.4 mM (◆), or 10 mM (○). 250 nM PP1γ was added to initiate phosphatase reaction at 37 ℃, and the reaction was quenched with SDS-PAGE loading buffer at different time points. Data are fitted to single-exponential decay equation. Rate constants are 0.19, 0.021, 0.020, 0.018, 0.023 min^{-1}, respectively. Initial phosphoryl content are 13 sites, 10 sites, 10 sites, 8.9 sites, and 8.6 sites, respectively.

f. Assay Optimization of PP1γ (temperature)

To optimize dephosphorylation we also adjusted reaction temperature. Although protocols provided by New England Biolab recommends 37 ºC for PP1α reactions, room temperature studies were first conducted for PP1γ purified within our lab to check the temperature dependence of dephosphorylation. Since studies conducted at 37 ºC are more complicated and error-prone, requiring the use of hot water bath for quick temperature adjustment, we first attempted the reaction at room temperature. At 23 ºC the rate of the reaction was reduced by 35-fold, much more than the drop in temperature could account for (Fig. 5.4B). Interestingly, PP1γ enzyme titration did not yield significant changes in dephosphorylation rate, suggesting that additional complex processes may be operational. Given these findings, subsequent reactions were performed at 37 ºC, the same condition for PP1α for easy comparison and optimized reaction rate.

g. Assay Optimization of PP1γ (inhibitors)

Finally, the non-specific phosphatase inhibitor vanadate and the false substrate β-glycerolphosphate were tested for inhibition of PP1γ, and they performed as well as SDS-PAGE loading buffer in quenching the dephosphorylation reaction (Fig. 5.4C). However, LysC treatment accelerated the dephosphorylation despite the presence of the inhibitors, such that the combined phosphoryl content of all digested fragments was much less than undigested substrate for every single time point of the dephosphorylation reaction (Fig. 5.4D). Since 3 μM PP1 was too high to be inhibited effectively (Fig. 5.4D), 250 nM PP1 would be used in all subsequent directionality assay. At 250 nM PP1, dephosphorylation activity can be stopped with stopping agents that do not interfere with subsequent LysC digestion (Fig. 5.4E).

h. Binding Affinity of PP1γ:ASF/SF2 Complex

Having optimized all assay parameters, an in-depth characterization of the PP1γ enzyme is now possible. Previous cell-based studies showed that PP1γ can be used to pull down ASF/SF2 or Tra2β [71]. Since these studies suggest that PP1 interacts well with these proteins, we were interested in establishing

the exact binding affinity of the complex using single-turnover analyses. In these studies, the substrate is held at a low concentration (100 nM phosphorylated ASF/SF2) and various concentrations of enzyme (up to 3 μM PP1γ) are added. The dephosphorylation rate was found to change in a hyperbolic manner as a function of PP1. Analysis of the curve provided a K_d for the PP1γ:ASF/SF2 complex of 250 nM (Fig. 5.5A). Overall, the data suggest that a fairly tight complex forms between PP1γ and ASF/SF2.

In a separate experiment, various concentrations of PP1γ were added to the SRPK1 phosphorylation assay to see how well PP1γ binds ASF/SF2 compared to SRPK1. By observing the first 80 seconds of the reaction, 3 μM PP1γ was necessary to inhibit 0.5 μM SRPK1 for 30-35% of the activity (data not shown), suggesting that PP1γ affinity for unphosphorylated ASF/SF2 is less than that for SRPK1 and ASF/SF2.

i. Dephosphorylation Rate Is Independent of Initial RS Domain Phosphoryl Content

Since very little is known about the molecular interactions between PP1γ and ASF/SF2, we studied the dephosphorylation rates of several mutant ASF constructs at various phosphoryl contents to better understand this interaction. To explore the effects of varying varying phosphoryl content, ATP was added to the SRPK1 pre-phosphorylation reaction at different amounts to generate ASF/SF2 phosphorylated at10- and 15-sites. ADP was added to ensure no re-phosphorylation by SRPK1 may occur, even at catalytic amount of SRPK1 enzyme. We found that the initial maximal dephosphorylation rates, calculated by the product of initial phosphoryl content and exponential rate constant, were similar for both phosphoryl contents, suggesting that PP1γ can recognize partially phosphorylated ASF/SF2 as well as fully phosphorylated ASF/SF2 (Fig. 5.6A).

j. Dephosphorylation Rates Are Not Dependent on Sequences in RS2, Prolines throughout the RS Domain, or Engagement of the RRMs with RNA

The role of the RS2 segment in RS1 dephosphorylation was investigated using mutant forms of

117
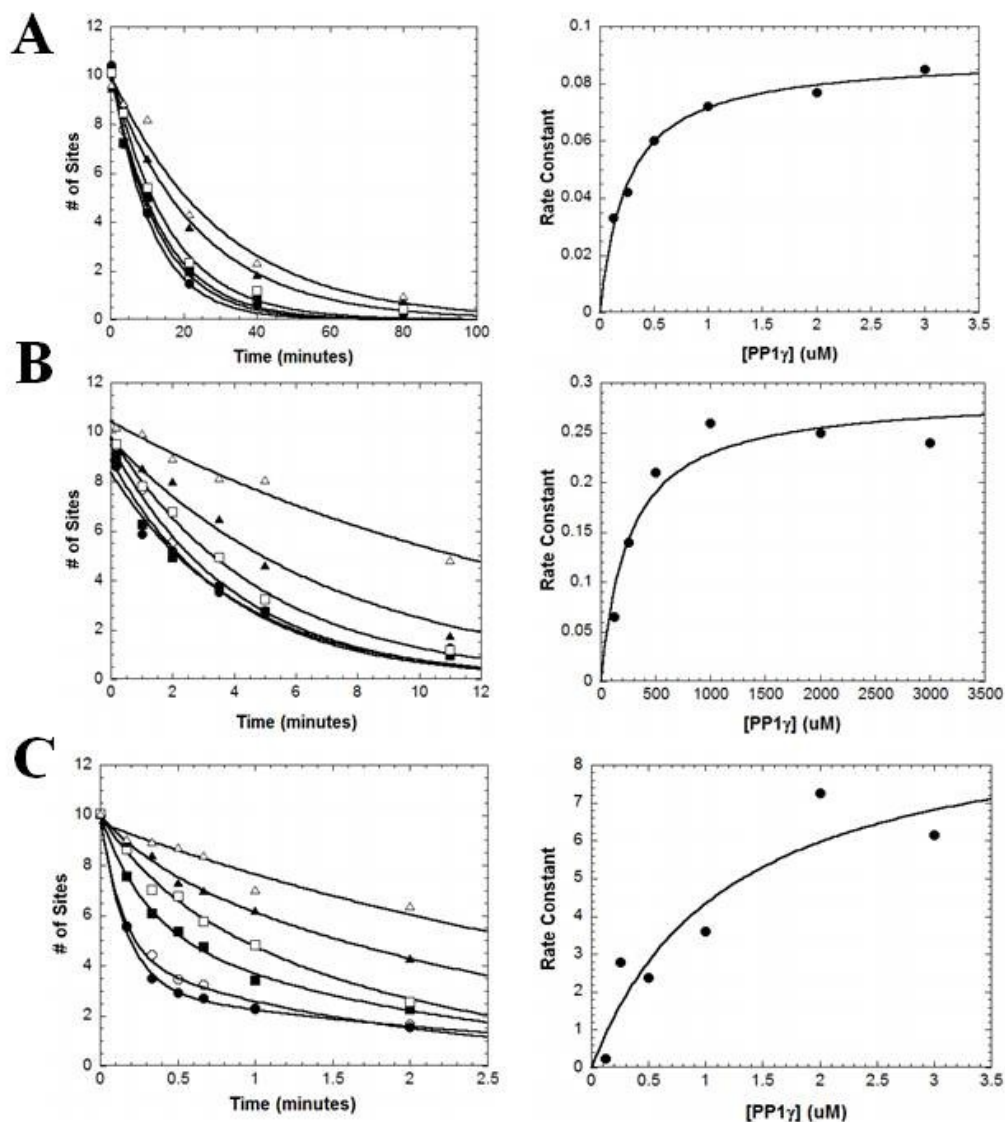
Figure 5.5 PP1γ Titration Curves. 100 nM wt-ASF (A), cl-ASF(214) (B), ASF(ΔRRM1) (C), ASF(ΔRRM2) (D), ASF(188-248) (E), and ASF(RATA) (F) were pre-phosphorylated with catalytic SRPK1 (20 nM) and 2 μM [^{32}P]ATP for 2 hours in NEB (Tris pH 7.5) buffer. 3 μM (●), 2 μM (○), 1 μM (■), 0.5 μM (□), 250 nM (▲), or 125 nM (Δ) PP1γ was added to initiate phosphatase reaction at 37 ºC, and the reaction was quenched with SDS-PAGE loading buffer at different time points. Data are fitted to single-exponential decay equation unless otherwise stated. For wt-ASF, rate constants are 0.085, 0.076, 0.071, 0.06, 0.042, and 0.033for PP1γ concentrations described above. For cl-ASF(124), rate constants are 0.24, 0.25, 0.26, 0.21, 0.14, and 0.065 for PP1γ concentrations described above. For ASF(ΔRRM1), data are fitted to double-exponential decay equation except for 125 nM PP1γ curve, fixing amplitude sum to be 10 sites. Rate constants are 6.1 and 0.34, 7.2 and 0.53, 3.6 and 0.48, 2.4 and 0.49, 2.8 and 0.34, and 0.24 for concentrations described above. Fast-phase amplitudes are 6.9, 5.6, 4.2, 2.9, and 1.5 for the double-exponential fits. Graphs on the right are for binding affinity determination. Rate constants of ASF constructs described above (fast phase only, if biphasic) are fitted to Michaelis-Menton equation to approximate the K_d values. When the dephosphorylation progress curve is biphasic, the rate constant of the fast phase is applied. The K_d values are 250 nM, 250 nM, and 1200 nM for constructs described above.

Figure 5.5 (Continued)
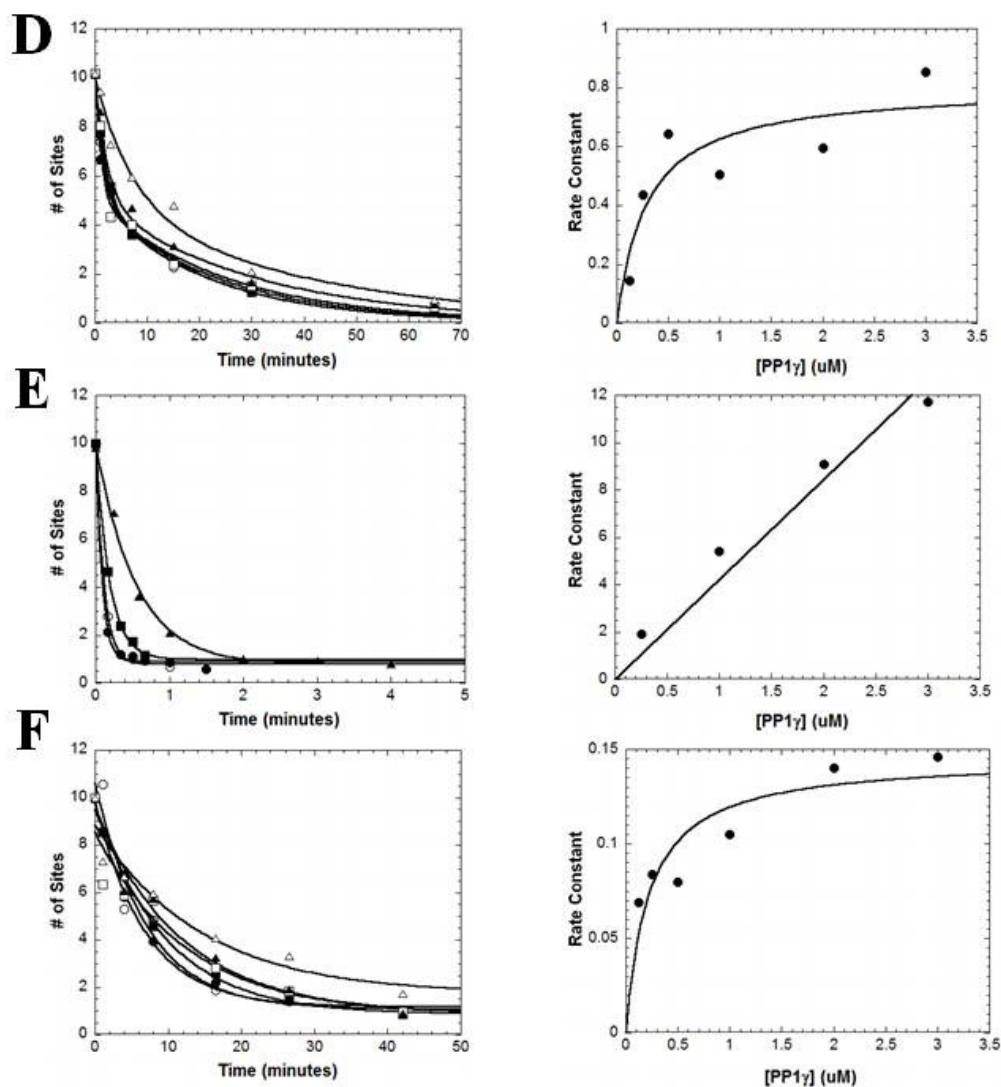
100 nM ASF(ΔRRM2) (D), ASF(188-248) (E), and ASF(RATA) (F) were pre-phosphorylated with catalytic SRPK1 (20 nM) and 2 μM [^{32}P]ATP for 2 hours in NEB (Tris pH 7.5) buffer. 3 μM (●), 2 μM (○), 1 μM (■), 0.5 μM (□), 250 nM (▲), or 125 nM (Δ) PP1γ was added to initiate phosphatase reaction at 37 °C, and the reaction was quenched with SDS-PAGE loading buffer at different time points. Data are fitted to single-exponential decay equation unless otherwise stated. For ASF(ΔRRM2), data are fitted to double-exponential decay equation, fixing amplitudes to 5 sites for fast and slow phases. Rate constants are 0.85 and 0.040, 0.59 and 0.046, 0.50 and 0.046, 0.64 and 0.042, 0.43 and 0.032, and 0.14 and 0.025 for concentrations described above. For ASF(RATA), data are fitted to double-exponential decay equation plus constant background rate. Rate constants are 0.15, 0.14, 0.10, 0.08, 0.084, and 0.069 for concentrations described above. Amplitudes of the fast phase and the background are 8.7 and 1.2 sites, 9.5 and 1.1, 8.9 and 0.85, 7.7 and 0.84, 8.6 and 0.88, and 7.1 and 1.7 for concentrations described above. For ASF(188-248), data are fitted to single-exponential decay equation with constant background. Rate constants are 11.7, 9.1, 5.4, and 1.9 for 3 μM, 2 μM, 1 μM, and 250 nM PP1γ. Amplitudes are 9.1 and 0.89, 9.2 and 0.79, 9.0 and 0.99, and 9.2 and 0.79 for the fast and slow phase of the concentrations described above. The K_d values are 280 nM, >= 3 μM, and 220 nM for constructs described above.
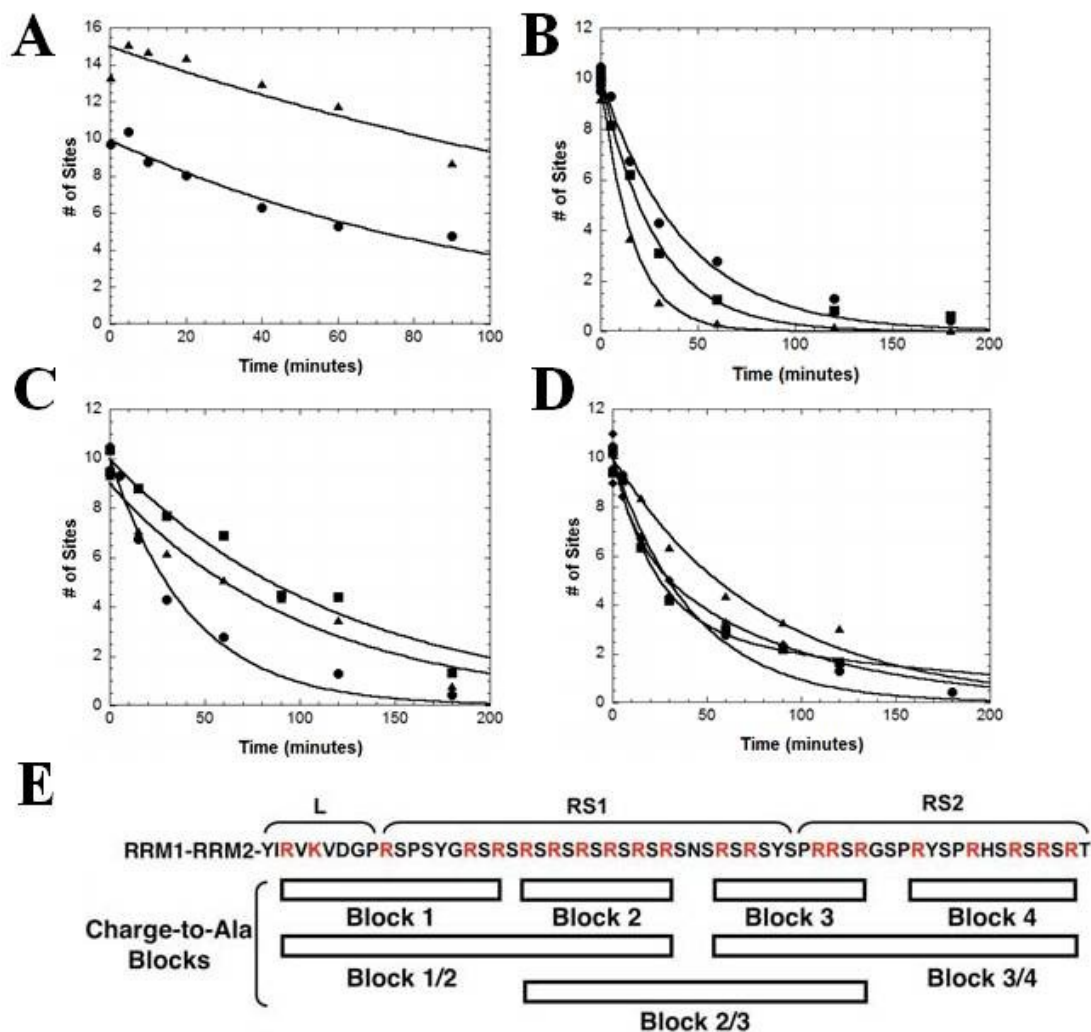
Figure 5.6 ASF/SF2 Structure Affects Dephosphorylation Rate. A) Phosphoryl content does not impact dephosphorylation rate. 0.1 µM wt-ASF/SF2 was pre-phosphorylated with catalytic amount of SRPK1 (20 nM) and 1.6 (●) or 100 µM (▲) [^{32}P]ATP for 120 minutes in NEB (Tris pH 7.5) buffer, with 1 mM or 10 mM ADP added at end of phosphorylation reaction. 250 nM PP1γ was added to initiate phosphatase reaction at 37 ºC, and the reaction was quenched with SDS-PAGE loading buffer at different time points. Data are fitted to single-exponential decay, with rate constants of 0.0097 and 0.0048, and initial phosphoryl content of 10 sites and 15 sites for 1.6 or 100 µM [^{32}P]ATP, respectively. B) Location of phosphoserines. 0.1 µM wt-ASF/SF2 (●), ASF(RS$_8$-RS$_8$) (▲), and ASF(SA$_8$-RS$_8$) (■) were dephosphorylated as described above. Data are fitted to single-exponential decay, with rate constants of 0.023, 0.062 and 0.035. C) Deletion of RS2 domain. 0.1 µM wt-ASF/SF2 (●), ASF(1-219) (▲), and ASF(1-226) (■) were dephosphorylated as described above. Data are fitted to single-exponential decay, with rate constants of 0.023, 0.010 and 0.0082. D) Mutation of RS2 domain. 0.1 µM wt-ASF/SF2 (●), ASF(block3) (▲), ASF(block4) (■), and ASF(block3/4) (♦) were dephosphorylated as described above. Data are fitted to single-exponential decay for wt-ASF/SF2 and ASF(block3), with rate constants of 0.023 and 0.012. For ASF(block4), data are fitted to double-exponential decay equation with rate constants of 0.045 and 0.005 and amplitudes of 6.8 and 3.2 sites for fast and slow phase. For ASF(block3/4), data are fitted to double-exponential decay equation with rate constants of 0.080 and 0.012 and amplitudes of 3.3 and 6.7 sites for fast and slow phase.

the RS2 segment. We studied two mutants in which the RS content in RS1 and RS2 were altered in the ASF(RS_8-RS_8) and ASF(SA_8-RS_8) constructs (Fig. 4.10). ASF(SA_8-RS_8) dephosphorylation was similar in rate to wt-ASF/SF2, suggesting that location of the phophoserines relative to the RRM domains did not change PP1 activity. ASF(RS_8-RS_8) dephosphorylation was about 2 to 3-fold faster than wt-ASF/SF2, suggesting that PP1γ may show a slight preference for hyper-phosphorylated (20 sites) ASF/SF2 (Fig. 5.6B).

Deletion in RS2 domain in ASF(1-219) and ASF(1-226) mutant constructs did not affect dephosphorylation rate significantly (Fig. 5.6C), with a 2-3-fold rate decrease for ASF(1-226). However, mutation of multiple arginines into alanines in the RS2 domain changed the dephosphorylation mechanism from monophasic to biphasic, with the fast phase being 2- to 4- fold faster than wt-ASF/SF2 dephosphorylation rate (Fig. 5.6D & E). Overall, the data suggested only a small role for the RS2 domain in regulating PP1γ activity.

Flanking proline mutations were first developed to explore regiospecificity of the SRPK1 kinase (Fig. 4.9). The effects of flanking proline mutations were also tested in the dephosphorylation assay. We observed a 3-fold decrease in rate for a four-proline-mutation ASF(4PA) and single-proline mutation ASF(PA228), suggesting that C-terminal prolines may play a minor role in dephosphorylation mechanism (Fig. 5.7A). ASF(PA200) showed no change in rate. However, Pin1 added during and at the end of the pre-phosphorylation reaction did not decrease dephosphorylation rate significantly (Fig.5.7B), failing to support the hypothesis that prolines or their isomerization direct dephosphorylation.

Finally, the dephosphorylation rate in the presence and in the absence of Ron ESE RNA was tested. ASF/SF2 must be phosphorylated for the assembly of the spliceosome with pre-mRNA, so the presence of RNA known to bind ASF/SF2 may induce certain conformational changes, possibly detectable via dephosphorylation rate. Rate changes from 0.01 to 0.0024 were observed, a 4-fold decrease with (Fig. 5.7C) 200 nM RNA and RNAse inhibitor cocktail added after pre-phosphorylation by SRPK1. Given that K_d value for the RNA-pASF/SF2 interaction should be around 20-40 nM (Cho, pers.
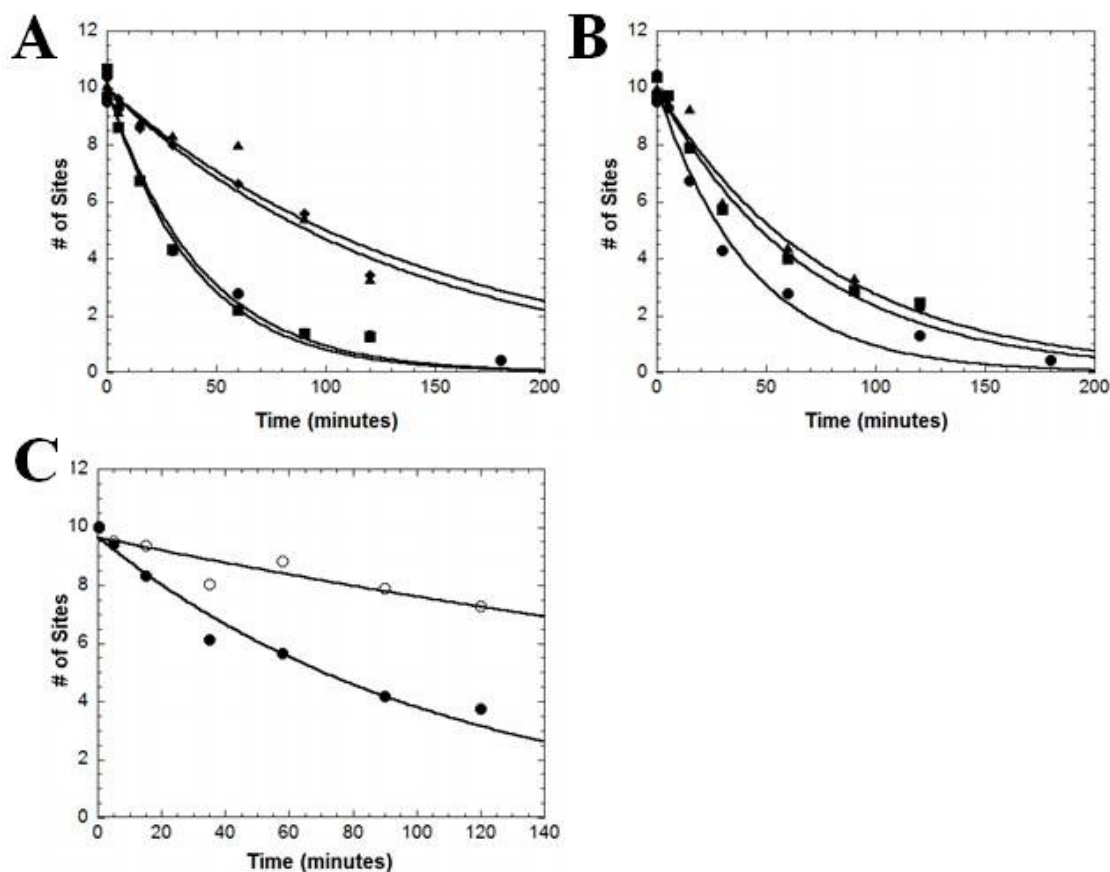
Figure 5.7 Proline and RNA Effects on PP1γ Activity. A) Proline mutations. 0.1 μM wt-ASF/SF2 (●), ASF(4PA) (▲), ASF(PA200) (■), and ASF(PA228) (♦) were pre-phosphorylated with catalytic amount of SRPK1 (20 nM) and 2 μM [^{32}P]ATP for 120 minutes in NEB (Tris pH 7.5) buffer. 250 nM PP1γ was added to initiate phosphatase reaction at 37 ºC, and the reaction was quenched with SDS-PAGE loading buffer at different time points. Data are fitted to single-exponential decay, with rate constants of 0.023, 0.007, 0.0025, and 0.008. B) Pin1 effects. 0.1 μM wt-ASF/SF2 was pre-phosphorylated with catalytic amount of SRPK1 (20 nM) and 2 μM [^{32}P]ATP for 120 minutes in NEB (Tris pH 7.5) buffer. Pin1 dialysis buffer (●) was added during phosphorylation, or 60 nM Pin1 was added during phosphorylation (▲) or during the last 20 minute of phosphorylation (■). 250 nM PP1γ was added to initiate phosphatase reaction at 37 ºC, and the reaction was quenched with SDS-PAGE loading buffer at different time points. Data are fitted to single-exponential decay, with rate constants of 0.023, 0.013, and 0.015. C) RNA effects. 0.1 μM wt-ASF/SF2 was pre-phosphorylated with catalytic amount of SRPK1 (20 nM) and 2 μM [^{32}P]ATP for 120 minutes in NEB (Tris pH 7.5) buffer. 250 nM PP1γ was added in the absence (●) and presence (○) of 200 nM Ron ESE RNA to initiate phosphatase reaction at 37 ºC. Data are fitted to single-exponential decay, with rate constants of 0.010 and 0.0024.

comm..), the data suggested that at a level of RNA that should saturate the RRMs dephosphorylation rate decreases slightly.

k. RRM1 Protects ASF/SF2 from Dephosphorylation by PP1γ

RRM1 contains the canonical RVEF binding motif that could be a binding site for PP1. To test this possibility we mutated RVEF to RATA to remove these potential contacts but did not detect any changes in dephosphorylation rate (Fig. 5.8A). Instead, similar to the PP1α results (Fig. 5.2), deletion of RRM1 accelerated the dephosphorylation rate by almost 300-fold, deletion of RRM2 accelerated by 7-fold, and deletion of both RRM domains accelerated by 83-fold (Fig. 5.8B), the magnitude of increase much higher than that for PP1α. In addition, the deletion of the RRMs reproduced the biphasic effect on the dephosphorylation progress curve, similar to that of PP1α. The slow phase rate constant was close to the wt-ASF/SF2 rate. With both RRMs deleted, the amplitude of the fast phase reached 90% of total amplitude, and the slow phase was approximated by a constant background due to the relatively slow rate (Fig. 5.8B). Parallel studies with GST-ASF and GST-RS showed similar fast-phase kinetic rates but increased magnitude of the fast phase from 3 to more than 8 sites upon deletion of the RRMs, suggesting that important RRM interactions are independent of tag type (Fig. 5.8C). Overall, the data suggested that RRM1, and RRM2 to a minor extent, protects ASF/SF2 from dephosphorylation by PP1 without specific contacts through RVEF motif.

While RRM deletions accelerate dephosphorylation rates, the affinity of the PP1γ:ASF/SF2 complex, measured by PP1γ titration, was affected in different ways. Due to the biphasic progress curves the number of variables increased and certain limits (amplitude fixed at 10 sites, or slow phase approximated as constant background) must be set up to avoid large errors in fitting. Removal of RRM1 increased the binding constant K_d from 250 nM to 1.2 μM (Fig. 5.5C), and removal of both RRMs increased the binding constant even higher such that only the linear phase of the curve could be observed (Fig. 5.5E). The K_d for RS domain and PP1γ probably exceeds 3 μM. Interestingly, removal of RRM2 did not impact binding affinity (Fig. 5.5D), suggesting that RRM1, not RRM2, is mediating interaction in the PP1γ:ASF complex. Mutation of the RVEF binding motif did not change binding affinity (Fig. 5.5F),
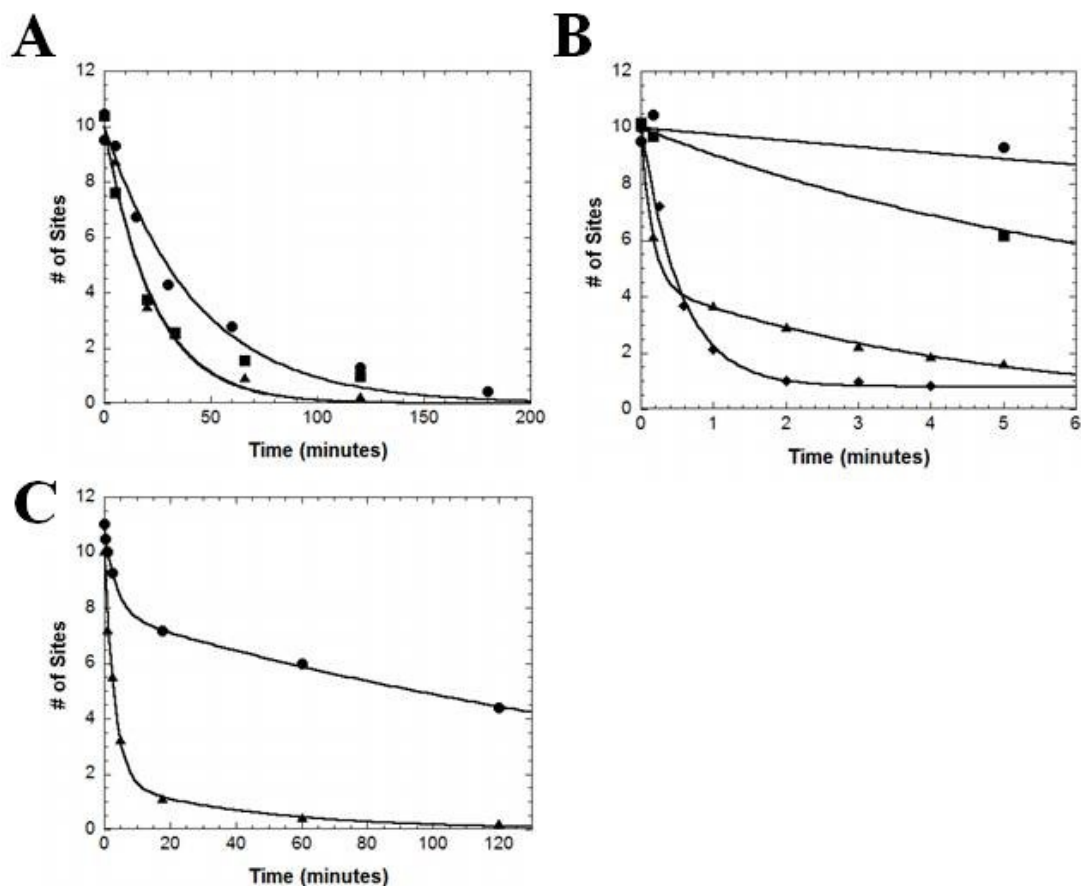
Figure 5.8 RRM Effects on PP1γ Activity. A) RVEF motif mutation. 0.1 μM wt-ASF/SF2 (●), ASF(RATA) (▲), and cl-ASF(RATA) (■) were pre-phosphorylated with catalytic amount of SRPK1 (20 nM) and 2 μM [^{32}P]ATP for 120 minutes in NEB (Tris pH 7.5) buffer. 250 nM PP1γ was added to initiate phosphatase reaction at 37 ºC, and the reaction was quenched with SDS-PAGE loading buffer at different time points. Data are fitted to single-exponential decay, with rate constants of 0.023, 0.044, and 0.0043. B) RRM deletions. 0.1 μM wt-ASF/SF2 (●), ASF(ΔRRM1) (▲), ASF(ΔRRM2) (■), and ASF(188-248) (♦) were pre-phosphorylated with catalytic amount of SRPK1 (20 nM) and 2 μM [^{32}P]ATP for 120 minutes in NEB (Tris pH 7.5) buffer and dephosphorylated by 250 nM PP1γ. Data are fitted to single-exponential decay for wt-ASF/SF2, with rate constants of 0.023. For ASF(ΔRRM1), data are fitted to double-exponential decay equation with rate constants of 6.8 and 0.22 and amplitudes of 5.5 and 4.5 sites for fast and slow phase. For ASF(ΔRRM2), data are fitted to double-exponential decay equation with rate constants of 0.15 and 0.019 and amplitudes of 6.2 and 3.8 sites for fast and slow phase. For ASF(188-248), data are fitted to a single-exponential decay equation plus constant background, with rate constant of 1.9 and amplitudes of 9.4 and 0.8 sites for fast and slow phase. C) GST-tagged RRM deletions. 0.1 μM wt GST-ASF (●) and GST-RS (▲) were pre-phosphorylated with catalytic amount of SRPK1 (20 nM) and 2 μM [^{32}P]ATP for 120 minutes in NEB (Tris pH 7.5) buffer and dephosphorylated by 250 nM PP1γ. For GST-ASF, data are fitted to double-exponential decay equation with rate constants of 0.28 and 0.0047 and amplitudes of 3 and 7.8 sites for fast and slow phase. For GST-RS, data are fitted to double-exponential decay equation with rate constants of 0.33 and 0.022 and amplitudes of 8.4 and 1.7 sites for fast and slow phase.

further supporting the model that those residues in RRM1 are not the sole contacts with PP1.

l.   Directional Dephosphorylation by PP1γ and Relevant Molecular Interactions

To study the order of dephosphorylation, we used the digestion mutant ASF/SF2 constructs from Chapter 4 to localize the removal of phosphates. PP1γ enzyme titration K_d determination studies were also carried out for cl-ASF(214), which exhibits low binding affinity similar to that of wt-ASF/SF2 (Fig. 5.5B), but higher catalytic efficiency compared to wt-ASF/SF2 (3-to-10 fold higher). Similar results from PP1α studies suggested that the mutations may affect RRM structure and PP1 dephosphorylation (Fig. 5.3B). Directionality studies were performed under kinetic conditions only, since ATP limitation experiments could only regulate initial phosphoryl content, not dephosphorylation rate. Pulse-chase was not an option due to inefficient lysis upon addition of high amounts of chase ATP. To maximize the signal of the samples with low phosphoryl content, the time course was stopped at roughly 70% dephosphorylation. To accentuate the progress of dephosphorylation, most of the time course points would be between 25% dephosphorylation and 70% dephosphorylation in a duplicate experiment (Fig. 5.9B).

In two separate experiments, the N-terminal fragment showed a visible decrease in signal intensity in the autoradiogram with the progression of time, while the C-terminal fragment was more stable. The ratio of $[^{32}P]$ incorporation in the C-terminal fragment to N-terminal fragment increased from 1.2 to 3.4 or 7.8, suggesting that dephosphorylation occurs preferentially in the N-terminal end of the RS1 domain. This decrease was also clearly noticeable on the autoradiogram. At 3 remaining phosphoserines, up to 8-fold preference for the N-terminus was demonstrated, not as strong as SRPK1 phosphorylation initiation preference of 50-100 fold but still significant, and similar to PP1α results (Fig. 5.3C).

Next, additional digestion mutants were used to explore the molecular interactions necessary to maintain this N->C directional dephosphorylation mechanism. Removal of RRM1 abolished the N-terminal preference, and with cl-ΔRRM1(214) the C/N ratio remained relatively constant between 1.1 and
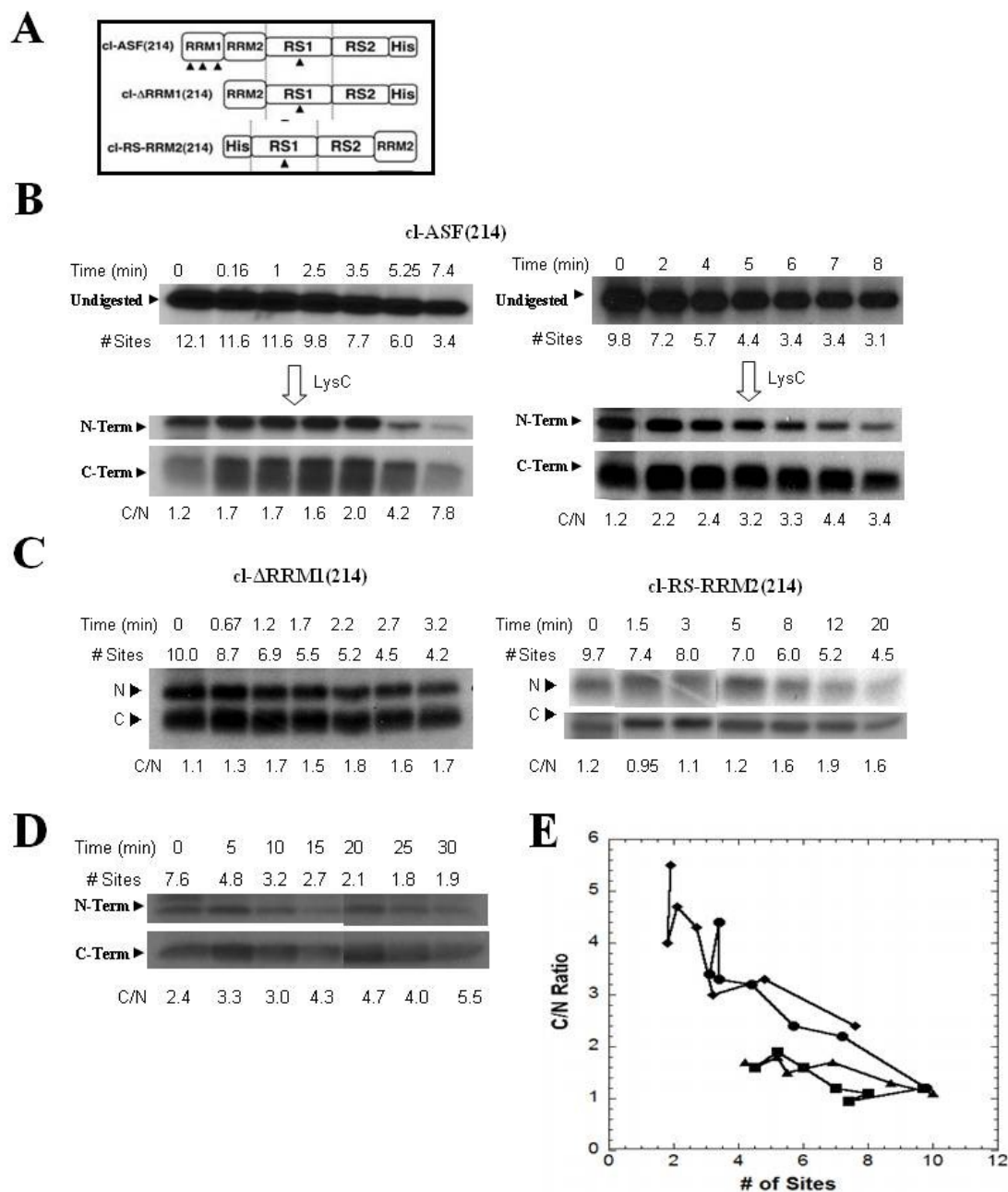
Figure 5.9 Directional Dephosphorylation by PP1γ. A) Domain organization of digestion mutants. B) ASF/SF2 is preferentially dephosphorylated at the N-terminus.   0.1 µM cl-ASF(214) was pre-phosphorylated with catalytic amount of SRPK1 (20 nM) and up to 2 µM [^{32}P]ATP for 120 minutes in NEB (Tris pH 7.5) buffer. 250 nM PP1γ was added to initiate phosphatase reaction at 37 ºC, and the reaction was quenched with 33 mM vanadate at different time points. C/N ratio is calculated from LysC-digested C-terminal and N-terminal fragments. Similar protocols for subsequent constructs. C) RRM effects. 0.1 µM cl-ΔRRM1(214) and cl-RS-RRM2(214) are used for digestion studies. D) RVEF motif effects. 0.1 µM cl-ASF(RATA) was used for digestion studies. E) C/N ratio plot. C/N ratio is plotted against phosphoryl content for cl-ASF(214) (●),cl-ΔRRM1(214) (▲), cl-RS-RRM2(214) (■), and cl-ASF(RATA) (♦).

1.8. Further swapping of RRM2 to the C-terminus of the RS domain did not restore the N-terminal preference, and the C/N ratio remained relatively constant between 0.95 and 1.9 (Fig. 5.9C). The RRM1 deletion accelerated dephosphorylation, as was observed in Fig. 5.10, but interestingly the swapping of the RRM2 domain slowed down the dephosphorylation rate. Finally, digestion data from cl-ASF(RATA) mutant dephosphorylation time course showed that mutation of the RVEF PP1 binding motif does not impact directional phosphorylation. Despite the incomplete phosphorylation (< 10 sites), the C/N ratio increased to 5.5 at 2 sites phosphorylated, suggesting that dephosphorylation occurs at the N-terminus preferentially (Fig. 5.9D). The cl-ASF(RATA) mutant was dephosphorylated 3-fold faster than wt-ASF, and slightly faster than ASF(RATA) (Fig. 5.8A). Overall, the data shown in the C/N ratio plot (Fig. 5.9E) suggested that RRM1 is an important regulator of directional N->C dephosphorylation by PP1γ, but the RVEF binding motif is not the sole contact.

C. Discussion

Whereas phosphorylation of SR proteins is necessary for spliceosome assembly, dephosphorylation is necessary for splicing catalysis and mRNA export [27-29], as part of the phosphorylation cycle for the SR proteins. Protein phosphatase 1 (PP1) is an important nuclear phosphatase involved in SR protein dephosphorylation [27, 42, 43], and recent studies found that ASF/SF2 dephosphorylation by PP1 regulates apoptosis through alternative splicing of the caspase-9 gene [44], underscoring the biological importance of investigating dephosphorylation mechanisms. The focus of our investigation was on the human isoform PP1γ, which has been used in two previous studies on SR protein dephosphorylation [70, 71]. These studies show that PP1 uses a dephosphorylation mechanism for ASF/SF2 that is highly complementary to that for the SRPK1 phosphorylation reaction. Namely, while SRPK1 places phosphates on RS1 in a preferential C- to N-terminal direction, PP1 takes these phosphates off in the opposite direction. Interestingly, it appears that one of the RRMs (RRM1) plays a role in this preferential mechanism.

a. Phosphorylated ASF/SF2 Structure

The spectra for wt-ASF shows noticeable decrease in ellipticity upon phosphorylation by either SRPK1 or Clk/Sty, while the spectra for the RS domain alone [ASF(188-248)] with and without phosphorylation show no change in secondary structure. The decrease in ellipticity indicates a clear shift from a beta sheet to a random coil, consistent with unpacking of the β4 strand of ASF/SF2 RRM2 domain upon phosphorylation. However, the spectra minimum is in the wrong place to suggest extensive alpha helical content in the unphosphorylated ASF/SF2 or in the RS domain. The far UV spectra cannot further resolve the structure of the RS domain, and it is possible that the RS domain adopts a random conformation phosphorylated or unphosphorylated, based on the overall low ellipticity values.

Finally, phosphorylated ASF/SF2 may exhibit conformational changes on a time scale of days that bring it closer to the unphosphorylated structure. These findings are supported by the diminishing spectral difference between unphosphorylated and phosphorylated ASF/SF2 by the third day at room temperature. Further studies, perhaps with hydrogen-deuterium exchange mass spectrometry, are necessary to eliminate the possibility that ASF/SF2 is simply dephosphorylated naturally over time.

b. PP1 Can Efficiently Process ASF/SF2 with Different Phosphoryl Content

Phosphorylation-dependent conformational changes in ASF/SF2 have been demonstrated using circular dichroism far UV spectra and cross-linking studies, in the presence and in the absence of SRPK1. However, since changes in phosphoryl content of the RS domain do not impact dephosphorylation rates (Fig. 5.2C & 5.6A), PP1 may interact with parts of ASF/SF2 that are unchanged by phosphorylation, possibly in RRM1. We will discuss this possibility in latter sections. Regardless of the mechanism, PP1 does not appear to discriminate between the various phosphates within RS1 since the dephosphorylation kinetics are cleanly monophasic and unchanged by the number of phosphates initially on the segment.

c. RRM1 Is Important for ASF/SF2 Binding

Previous studies have shown that PP1γ can pull down ASF/SF2 or Tra2β in cell-based studies [71], but the exact binding affinity has yet to be characterized. Since the competition assay is not suitable

for high starting signal of phosphorylated substrate, single-turnover assay were applied instead, where enzymatic rate depends on extent of enzyme-substrate complex formation. The K_d for PP1γ:ASF/SF2 complex was shown to be 250 nM (Fig. 5.7A), higher than K_d for the SRPK1:ASF/SF2 complex (50-100 nM) [48] but still fairly tight. Since there are no previous studies reporting aggregation or cooperative behavior of PP1, it is unlikely that the dephosphorylation rate increase is instead due to aggregation of PP1 at high concentrations. However, we cannot rule out the possibility that the hyperbolic response in the rate versus PP1 plot (Fig. 5.5A) may be due to a slow change in rate-limiting step from enzyme-substrate association to a conformational change in p-ASF/SF2 at high PP1 concentrations.

Having established that PP1 interacts well with ASF/SF2, molecular interactions between the enzyme and substrate were characterized, and dephosphorylation rate is a good kinetic measurement of stable interaction. PP1 dephosphorylation of wt-ASF/SF2 is monophasic and can be characterized by a single-exponential decay equation (Fig. 5.4A). These findings suggest that either PP1 does not discriminate between the individual phosphates in RS1, or that the reaction is sequential but is controlled by an early, rate-limiting step. To address the roles of residues and phosphates in the RS domain for binding we made a series of mutations. In general, we found that many individual residues within the RS domain do not appear to be very important for dephosphorylation. Hyper phosphorylation (20 sites) (Fig. 5.6B) and Arg-to-Ala mutations increase dephosphorylation rate only by 2-to-4 fold, suggesting minor interactions with RS2 domain. In the case of Arg-to-Lys mutations in RS2 (Fig. 5.6D) the progress curves are clearly biphasic, implying a significant shift in mechanism. As for prolines flanking RS1 domain, some mutations decelerates dephosphorylation by only 3-fold, and peptidyl-prolyl isomerase Pin1 addition shows no effect (Fig. 5.7A & B). Without a clear trend, the data implies that prolines do not have any global effects on the dephosphorylation of the whole RS1 segment. Parallel studies with PP1α showed similar results (Fig. 5.2D). Finally, addition of Ron ESE RNA decelerates dephosphorylation by about 4-fold (Fig. 5.7C), suggesting that dephosphorylation of the SR proteins in spliceosome complex may be slower than elsewhere in the nucleus.

The only important interaction observed occurs between the RRM domains and PP1. Deletion of RRM1 accelerates dephosphorylation rate by almost 300-fold, deletion of RRM2 accelerates it by 7-

fold, and deletion of both RRM domains accelerates it by 83-fold (Fig. 5.8B). Interestingly, removal of the RRMs induce a shift in the progress curve from monophasic to biphasic dephosphorylation, and removal of both RRMs drives the fast-phase further to 80-90% dephosphorylation, compared against 50-55% dephosphorylation in the fast phase of ΔRRM1 construct. Thus, removal of both RRMs still results in faster overall dephosphorylation than sole removal of RRM1. Interestingly, parallel studies in GST-tagged constructs showed that removal of RRMs drives the fast-phase further to cover 90% of dephosphorylation progress, compared against 20% shown for wt-ASF (Fig. 5.8C). Parallel studies with PP1α revealed similar effects with RRM deletion (Fig. 5.2E). Corroborated by LysC digestion acceleration data from section b, these observations suggest that the RRMs, and RRM1 in particular, protects ASF/SF2 from dephosphorylation by PP1. Going one step further, the PP1 binding motif RVEF in RRM1 was mutated to RATA to confirm specific contacts between PP1γ and ASF/SF2, but PP1 dephosphorylates both wild-type and mutant ASF/SF2 equally well (Fig. 5.8A), suggesting that, similar to SRPK1 interaction with RRM2, general surface contacts are more relevant than individual residues.

Our studies show that RRM1 plays an opposing role in dephosphorylation rate and ASF/SF2 binding affinity. While the removal of RRM1 accelerates the rate of dephosphorylation, this observation is accompanied by a decrease in binding affinity, as judged in the single-turnover assays (Fig. 5.5C & E). This phenomenon appears to be specific for RRM1, as the removal of RRM2 accelerates dephosphorylation much less, and correspondingly, does not impact binding affinity (Fig. 5.5D). In general, the change in binding affinity always accompanies the change in dephosphorylation rate, lending support to the model that RRM1 mediates binding of ASF/SF2 to PP1 and negatively regulates catalysis. Without structural data, it is difficult to explain this observation. It is possible that PP1 binds through RRM1, placing it in a position that is not easily accessible to the RS domain. Removal of RRM1 then frees PP1 to rapidly dephosphorylate the RS domain, although the affinity is negatively impacted by this disruption.

d. Regiotemporal Dephosphorylation of ASF/SF2 by PP1

ASF/SF2 is phosphorylated at roughly 10 sites in RS1 by SRPK1 using a highly ordered, C-to-N-terminal direction mechanism. In this chapter we now describe how PP1 removes these phosphates using an opposing N-to-C-terminal mechanism. In the case of SRPK1, the directional mechanism has been shown to be linked to specific contacts between the RS domain and RRM2 with the enzyme. Although we do not have an x-ray structure for PP1 with ASF/SF2 bound, we have shown through deletion analyses that RRM1 plays an important role in controlling this directionality (Fig. 5.9). Thus, while RRM2 is important for directional phosphorylation of RS1 by SRPK1, RRM1 is important for directional dephosphorylation by PP1. Such findings suggest that the kinase and phosphatase involved in controlling the phosphoryl content of ASF/SF2 may interact with the SR protein through very different binding determinants.

Dephosphorylation specificity has been reported in previous PP1 studies. Troponin in muscle fibers are phosphorylated and then dephosphorylated in phosphorylation cycles similar to SR proteins. While rat cardiac myofibril troponin I is uniformly dephosphorylated by rat cardiac PP1, troponin T is preferentially dephosphorylated at specific sites (Ser43/45 and Thr 144 resistant to PP1) [121]. Ser 23/24 are most vulnerable to PP1 dephosphorylation, with Ser 78 roughly in the middle in terms of dephosphorylation accessibility. It is important to note that ordered dephosphorylation does not have to be directional, but PP1 dephosphorylates ASF/SF2 in a unique, directional mechanism.

e. Conclusion

Overall, the data presented in this chapter suggest that dephosphorylation of the wild-type ASF/SF2 is a slow but directional process (Fig. 5.10). While SRPK1 places phosphates onto the RS domain in an N-terminal direction, PP1 removes these phosphates in the opposite, C-terminal direction. Interestingly, both enzymes use different RRMs to achieve this directionality. While SRPK1 requires RRM2 for C-to-N-terminal phosphorylation, PP1 requires RRM1 for N-to-C-terminal dephosphorylation (Fig. 5.10). These two directional activities may be vital for the biological function of the SR protein. Indeed, both enzymes may work cooperatively in the cell to maintain phosphorylation levels towards the

C-terminal end of the RS1 segment compared to the N-terminal end. In this scenario, SRPK1 would add phosphates to the C-terminal end of RS1 while PP1 would prune phosphates away from the N-terminal end (Fig. 5.11).

Chapter 5 is, in part, a reprint of the material as it appears in *Molecular Cell,* 2008, 29(5), 563-576, Ngo, J., Giang, K., Chakrabarti, S., Ma, C.T., Huynh, N., Hagopian, J.C., Dorrestein, P.C., Fu, X.D., Adams, J.A., Ghosh, G. The dissertation author was a collaborative researcher and co-author of this publication.
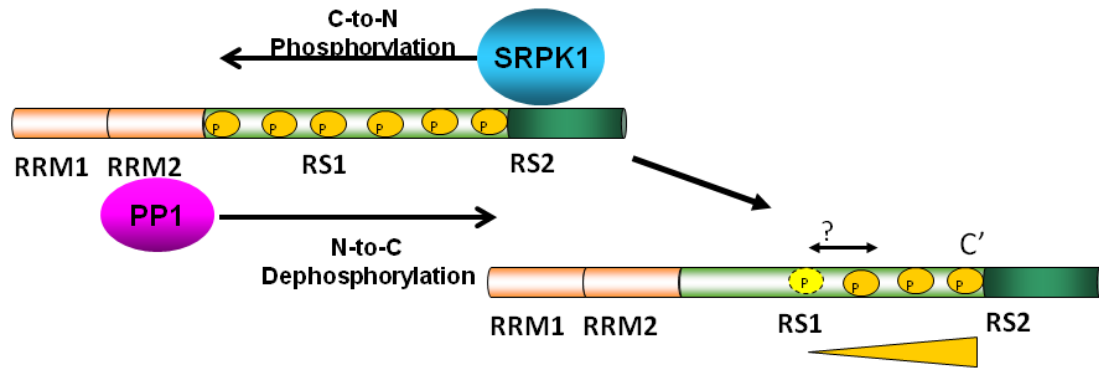
\

Fig. 5.10 Potential Interactions between SRPK1 and PP1. SRPK1 phosphorylates ASF/SF2 in C->N direction, while PP1 dephosphorylates in N->C direction. Interaction of the two enzymes may fine-tune phosphoryl content of ASF/SF2 in response to cell signaling pathways.

Chapter 6

Conclusion

SR protein splicing factors are essential for the maturation of pre-mRNA in mammalian cells, from exon definition and spliceosome assembly to mRNA export for protein translation. SR proteins consist of one or two RNA-recognition motif (RRM) domains and an RS domain rich in arginine-serine dipeptide repeats. The phosphorylation and dephosphorylation of the RS domain regulates protein-protein and protein-RNA interactions and controls subcellular localization of the SR proteins. Our lab has shown previously that phosphorylation of the RS1 segment by SRPK1 leads to entry of the SR protein into the nucleus and localization in nuclear speckles whereas phosphorylation of the RS2 segment by Clk/Sty leads to dispersion of the speckles. Although phosphorylation by SRPK1 and Clk/Sty is processive where multi-site phosphorylation occurs without dissociation of the enzyme:substrate complex, very little is known about the dephosphorylation mechanism by nuclear phosphatases such as PP1. Analysis of both parts of the phosphorylation cycle through mechanistic investigations should help us to better understand the function of SR proteins in splicing and to develop regulators of kinase function for chemotherapeutic purposes.

Since SRPK1 uses a processive mechanism to add phosphates onto the RS domain of ASF/SF2, we wished to establish whether this addition occurs in an orderly fashion. Here, we relied on the analogy of SRPK1 to DNA polymerases which add nucleic bases to a template primer in a 5' to 3' direction. To assess whether SRPK1 may also be a directional enzyme we engineered a mutant form of the substrate SR protein ASF/SF2 that could be separated into two identifiable LysC fragments on PAGE representing the N- and C-terminal ends of the RS1 segment of the RS domain. Using both equilibrium (ATP limitation) and kinetic (pulse-chase) experiments we showed that SRPK1 phosphorylates RS1 in a preferred C-to-N-terminal direction. Although mutagenesis studies showed that SRPK1 does not require a strict serine initiation site, engineered phosphomapping studies showed that the enzyme prefers to initiate in a small segment toward the C-terminal end of RS1, a region that includes Ser 221, 223, and 225 (initiation box). This flexible initiation site is in sharp contrast to the priming P+4 phosphoserine which is strictly necessary for subsequent phosphorylation by glycogen synthase kinase-3, also a directional kinase. In this regard, SRPK1 shares more similarity to the tyrosine kinase Src which processively phosphorylates numerous tyrosines in the protein substrate Cas but does so without a strict initiation site

[109]. Using extensive mutagenesis studies we showed that while SRPK1 strongly prefers to phosphorylate RS1 it can also phosphorylate RS2 but a a much slower rate. The preference for RS1 is dependent entirely on a long RS/SR repeat in RS1 and not on any distal contacts with the RRMs. Overall, we showed that SRPK1 binds with high affinity to ASF/SF2 and initiates a directional and processive phosphorylation reaction.

Having established a directional phosphorylation model for SRPK1, we next explored the molecular interactions that underpin this unique mechanism. For this study, we relied upon the recent X-ray structure of the SRPK1:ASF/SF2 complex. We found that there are two main contact groups that act cooperatively to induce high C-terminus regiospecificity. The electronegative docking groove in SRPK1 interacts with arginines in the N-terminus of RS1 to align the C-terminal portion of this segment for phosphorylation initiation. Mutations in the docking groove lead to substantial decreases in C-terminal initiation and a striking shift toward more random phosphorylation. Phosphorylation of RS1 generates phosphoserines that are likely to be repulsed by the docking groove, thereby, maintaining unidirectional phosphorylation. Despite the necessity for these charged contacts, other structural features of ASF/SF2 are required for directional phosphorylation. We showed that surface contacts between RRM2 and both SRPK1 kinase lobes are also essential regulators of directional phosphorylation. Removal of RRM2, but not RRM1, leads to a substrate that is phosphorylated with only a very weak preference for C-terminal serines in RS1. Interestingly, the contacts within RRM2 are not discrete as they may be mutated without any changes in directional phosphorylation. These findings suggest that RRM2 makes diffuse contacts with SRPK1 and that the binding surface is highly malleable. This is consistent with data from the X-ray structure of the complex. While the interaction surface of RRM2 and SRPK1 is large, the contacts do not appear to be very specific with few 'hot-spots' for direct contact. Overall, our mutagenesis studies show that SRPK1 achieves a directional mechanism through both local contacts in the RS domain, most notably those made with the N-terminal portion of RS1 (N'-RS1), and a larger, topological surface offered by RRM2. Surprisingly, when these two groups of contacts are removed, SRPK1 has a preference for N-terminal rather than C-terminal phosphorylation.

In the final stage of the thesis we investigated the mechanism of ASF/SF2 dephosphorylation by PP1, an important serine-threonine phosphatase in the nucleus. We found that PP1 can dephosphorylate RS1 in a monophasic reaction that is much slower than the phosphorylation reaction catalyzed by SRPK1. Using deletion analyses, we found that RRM1 plays a considerable role in protecting the RS domain for dephosphorylation. Deletion of RRM1 accelerates the dephosphorylation rate by almost 300-fold and leads to biphasic kinetics. Although RRM1 has a potential docking site for PP1 (RVEF), mutation of this site (RATA) had no substantial effect on the kinetics of dephosphorylation suggesting that the interactions with RRM1 may occur through a unique binding mechanism. For example, it is possible that RRM1 may limit RS domain dephosphorylation by directly interacting with the phosphorylated domain, thereby, itnererfering with PP1 contacts at RS1.

The final goal of the thesis was to analyze the order of dephosphorylation. For this study we utilized the digestion mutants of ASF/SF2 developed for determining the order of the SRPK1 reaction. Controlled phosphorylation reactions using SRPK1 were implemented so that only the dephosphorylation of serines in RS1 (about 10 sites) were investigated. With kinetic analysis, we showed that PP1 removes phosphates from RS1 in an N-to-C-terminal manner. Thus, while SRPK1 add phosphates in the C-to-N-terminal direction, PP1 operates in the opposite direction. To address potential structural aspects of the substrate that might drive this mechanism, we studied the directional dephosphorylation of several deletion mutants of ASF/SF2. We found that RRM1 and not RRM2 is essential for controlling the directional reaction. While wild-type ASF/SF2 is dephosphorylated in a directional manner, a substrate mutant lacking RRM1 or both RRMs is dephosphorylated in a random manner.

In this thesis, we have shown that the phosphorylation and dephosphorylation cycles of the SR protein ASF/SF2 occur through unique mechanisms. While SRPK1 phosphorylates from the initiation box to the N-terminus of RS1, PP1 functions in the opposite direction. We have described relevant molecular contacts that control these two complementary reactions showing that RRM2 is important for directional phosphorylation while RRM1 is important for directional dephosphorylation. We think that these activities may have a biological function for SR protein phosphorylation particularly in the nucleus where rounds of phosphorylation and dephosphorylation are needed for pre-mRNA splicing. As shown

in Fig. 5.10, SRPK1 will rapidly modify all serines in RS1, an event that is essential for entry of the SR protein into the nucleus and into speckles. In the nucleus, PP1 will then begin to dephosphorylate RS1 starting from the N-terminal end. This unique contrast in activities between kinase and phosphatase would ensure that a partially phosphporylated RS domain would be enriched in phosphates at the C-terminal rather than the N-terminal end. Such regiospecificity might be important for splicing regulation in the nucleus. One could imagine other effector molecules that might alter RS domain conformation and convert PP1 into a phosphatase with C-terminal initiation preferences leading to a different regiospecific modification of the RS domain. While these possibilities await other experimental protocols, the studies presented in this thesis now establish a framework for analyzing the regiospecific phosphorylation control of ASF/SF2 and other SR proteins in a biological context.

References

1.   Padgett, R.A., et al., *Splicing of messenger RNA precursors.* Annu Rev Biochem, 1986. **55**: p. 1119-50.

2.   Sharp, P.A., *On the origin of RNA splicing and introns.* Cell, 1985. **42**(2): p. 397-400.

3.   Pajares, M.J., et al., *Alternative splicing: an emerging topic in molecular and clinical oncology.* Lancet Oncol, 2007. **8**(4): p. 349-57.

4.   Lee, C. and Q. Wang, *Bioinformatics analysis of alternative splicing.* Brief Bioinform, 2005. **6**(1): p. 23-33.

5.   Wang, E.T., et al., *Alternative isoform regulation in human tissue transcriptomes.* Nature, 2008. **456**(7221): p. 470-6.

6.   Ares, M., Jr. and N.J. Proudfoot, *The spanish connection: transcription and mRNA processing get even closer.* Cell, 2005. **120**(2): p. 163-6.

7.   Matlin, A.J., F. Clark, and C.W. Smith, *Understanding alternative splicing: towards a cellular code.* Nat Rev Mol Cell Biol, 2005. **6**(5): p. 386-98.

8.   Jurica, M.S. and M.J. Moore, *Pre-mRNA splicing: awash in a sea of proteins.* Mol Cell, 2003. **12**(1): p. 5-14.

9.   Wahl, M.C., C.L. Will, and R. Luhrmann, *The spliceosome: design principles of a dynamic RNP machine.* Cell, 2009. **136**(4): p. 701-18.

10.   Kohtz, J.D., et al., *Protein-protein interactions and 5'-splice-site recognition in mammalian mRNA precursors.* Nature, 1994. **368**(6467): p. 119-24.

11.   Wu, J.Y. and T. Maniatis, *Specific interactions between proteins implicated in splice site selection and regulated alternative splicing.* Cell, 1993. **75**(6): p. 1061-70.

12.   Reed, R., *Initial splice-site recognition and pairing during pre-mRNA splicing.* Curr Opin Genet Dev, 1996. **6**(2): p. 215-20.

13.   Roscigno, R.F. and M.A. Garcia-Blanco, *SR proteins escort the U4/U6.U5 tri-snRNP to the spliceosome.* RNA, 1995. **1**(7): p. 692-706.

14.   Chew, S.L., et al., *Evidence for the function of an exonic splicing enhancer after the first catalytic step of pre-mRNA splicing.* Proc Natl Acad Sci U S A, 1999. **96**(19): p. 10655-60.

15.   Graveley, B.R., *Sorting out the complexity of SR protein functions.* RNA, 2000. **6**(9): p. 1197-211.

16. Shen, H., J.L. Kan, and M.R. Green, *Arginine-serine-rich domains bound at splicing enhancers contact the branchpoint to promote prespliceosome assembly.* Mol Cell, 2004. **13**(3): p. 367-76.

17. Shen, H. and M.R. Green, *A pathway of sequential arginine-serine-rich domain-splicing signal interactions during mammalian spliceosome assembly.* Mol Cell, 2004. **16**(3): p. 363-73.

18. Pozzoli, U. and M. Sironi, *Silencers regulate both constitutive and alternative splicing events in mammals.* Cell Mol Life Sci, 2005. **62**(14): p. 1579-604.

19. Zhang, Z. and A.R. Krainer, *Involvement of SR proteins in mRNA surveillance.* Mol Cell, 2004. **16**(4): p. 597-607.

20. Lejeune, F. and L.E. Maquat, *Mechanistic links between nonsense-mediated mRNA decay and pre-mRNA splicing in mammalian cells.* Curr Opin Cell Biol, 2005. **17**(3): p. 309-15.

21. Caceres, J.F., G.R. Screaton, and A.R. Krainer, *A specific subset of SR proteins shuttles continuously between the nucleus and the cytoplasm.* Genes Dev, 1998. **12**(1): p. 55-66.

22. Huang, Y., et al., *SR splicing factors serve as adapter proteins for TAP-dependent mRNA export.* Mol Cell, 2003. **11**(3): p. 837-43.

23. Sanford, J.R., et al., *A novel role for shuttling SR proteins in mRNA translation.* Genes Dev, 2004. **18**(7): p. 755-68.

24. Mermoud, J.E., P.T. Cohen, and A.I. Lamond, *Regulation of mammalian spliceosome assembly by a protein phosphorylation mechanism.* EMBO J, 1994. **13**(23): p. 5679-88.

25. Caceres, J.F. and A.R. Kornblihtt, *Alternative splicing: multiple control mechanisms and involvement in human disease.* Trends Genet, 2002. **18**(4): p. 186-93.

26. Xiao, S.H. and J.L. Manley, *Phosphorylation of the ASF/SF2 RS domain affects both protein-protein and protein-RNA interactions and is necessary for splicing.* Genes Dev, 1997. **11**(3): p. 334-44.

27. Mermoud, J.E., P. Cohen, and A.I. Lamond, *Ser/Thr-specific protein phosphatases are required for both catalytic steps of pre-mRNA splicing.* Nucleic Acids Res, 1992. **20**(20): p. 5263-9.

28. Huang, Y., T.A. Yario, and J.A. Steitz, *A molecular link between SR protein dephosphorylation and mRNA export.* Proc Natl Acad Sci U S A, 2004. **101**(26): p. 9666-70.

29. Lai, M.C. and W.Y. Tarn, *Hypophosphorylated ASF/SF2 binds TAP and is present in messenger ribonucleoproteins.* J Biol Chem, 2004. **279**(30): p. 31745-9.

30.   Ding, J.H., et al., *Regulated cellular partitioning of SR protein-specific kinases in mammalian cells.* Mol Biol Cell, 2006. **17**(2): p. 876-85.

31.   Wang, H.Y., et al., *Localization of serine kinases, SRPK1 (SFRSK1) and SRPK2 (SFRSK2), specific for the SR family of splicing factors in mouse and human chromosomes.* Genomics, 1999. **57**(2): p. 310-5.

32.   Gui, J.F., et al., *Purification and characterization of a kinase specific for the serine- and arginine-rich pre-mRNA splicing factors.* Proc Natl Acad Sci U S A, 1994. **91**(23): p. 10824-8.

33.   Yun, C.Y. and X.D. Fu, *Conserved SR protein kinase functions in nuclear import and its action is counteracted by arginine methylation in Saccharomyces cerevisiae.* J Cell Biol, 2000. **150**(4): p. 707-18.

34.   Kataoka, N., J.L. Bachorik, and G. Dreyfuss, *Transportin-SR, a nuclear import receptor for SR proteins.* J Cell Biol, 1999. **145**(6): p. 1145-52.

35.   Lai, M.C., R.I. Lin, and W.Y. Tarn, *Transportin-SR2 mediates nuclear import of phosphorylated SR proteins.* Proc Natl Acad Sci U S A, 2001. **98**(18): p. 10154-9.

36.   Hall, L.L., et al., *Molecular anatomy of a speckle.* Anat Rec A Discov Mol Cell Evol Biol, 2006. **288**(7): p. 664-75.

37.   Colwill, K., et al., *The Clk/Sty protein kinase phosphorylates SR splicing factors and regulates their intranuclear distribution.* EMBO J, 1996. **15**(2): p. 265-75.

38.   Nayler, O., S. Stamm, and A. Ullrich, *Characterization and comparison of four serine- and arginine-rich (SR) protein kinases.* Biochem J, 1997. **326 (Pt 3)**: p. 693-700.

39.   Ngo, J.C., et al., *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, 2005. **20**(1): p. 77-89.

40.   Misteli, T. and D.L. Spector, *Serine/threonine phosphatase 1 modulates the subnuclear distribution of pre-mRNA splicing factors.* Mol Biol Cell, 1996. **7**(10): p. 1559-72.

41.   Kress, T.L., N.J. Krogan, and C. Guthrie, *A single SR-like protein, Npl3, promotes pre-mRNA splicing in budding yeast.* Mol Cell, 2008. **32**(5): p. 727-34.

42.   Gilbert, W. and C. Guthrie, *The Glc7p nuclear phosphatase promotes mRNA export by facilitating association of Mex67p with mRNA.* Mol Cell, 2004. **13**(2): p. 201-12.

43.   Trinkle-Mulcahy, L. and A.I. Lamond, *Mitotic phosphatases: no longer silent partners.* Curr Opin Cell Biol, 2006. **18**(6): p. 623-31.

44.   Massiello, A. and C.E. Chalfant, *SRp30a (ASF/SF2) regulates the alternative splicing of caspase-9 pre-mRNA and is required for ceramide-responsiveness.* J Lipid Res, 2006. **47**(5): p. 892-7.

45. Lin, S., et al., *Dephosphorylation-dependent sorting of SR splicing factors during mRNP maturation.* Mol Cell, 2005. **20**(3): p. 413-25.

46. Shi, Y., B. Reddy, and J.L. Manley, *PP1/PP2A phosphatases are required for the second step of Pre-mRNA splicing and target specific snRNP proteins.* Mol Cell, 2006. **23**(6): p. 819-29.

47. Prasad, J., et al., *The protein kinase Clk/Sty directly modulates SR protein activity: both hyper- and hypophosphorylation inhibit splicing.* Mol Cell Biol, 1999. **19**(10): p. 6991-7000.

48. Velazquez-Dones, A., et al., *Mass spectrometric and kinetic analysis of ASF/SF2 phosphorylation by SRPK1 and Clk/Sty.* J Biol Chem, 2005. **280**(50): p. 41761-8.

49. Manning, G., et al., *The protein kinase complement of the human genome.* Science, 2002. **298**(5600): p. 1912-34.

50. Colwill, K., et al., *SRPK1 and Clk/Sty protein kinases show distinct substrate specificities for serine/arginine-rich splicing factors.* J Biol Chem, 1996. **271**(40): p. 24569-75.

51. Aubol, B.E., et al., *Processive phosphorylation of alternative splicing factor/splicing factor 2.* Proc Natl Acad Sci U S A, 2003. **100**(22): p. 12601-6.

52. Ngo, J.C., et al., *A sliding docking interaction is essential for sequential and processive phosphorylation of an SR protein by SRPK1.* Mol Cell, 2008. **29**(5): p. 563-76.

53. Wang, H.Y., et al., *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, 1998. **140**(4): p. 737-50.

54. Ngo, J.C., et al., *SR protein kinase 1 is resilient to inactivation.* Structure, 2007. **15**(1): p. 123-33.

55. Zhong, X.Y., et al., *Regulation of SR protein phosphorylation and alternative splicing by modulating kinetic interactions of SRPK1 with molecular chaperones.* Genes Dev, 2009. **23**(4): p. 482-95.

56. Duncan, P.I., et al., *The Clk2 and Clk3 dual-specificity protein kinases regulate the intranuclear distribution of SR proteins and influence pre-mRNA splicing.* Exp Cell Res, 1998. **241**(2): p. 300-8.

57. Bullock, A.N., et al., *Kinase domain insertions define distinct roles of CLK kinases in SR protein phosphorylation.* Structure, 2009. **17**(3): p. 352-62.

58. Menegay, H.J., et al., *Biochemical characterization and localization of the dual specificity kinase CLK1.* J Cell Sci, 2000. **113 (Pt 18)**: p. 3241-53.

59. Prasad, J. and J.L. Manley, *Regulation and substrate specificity of the SR protein kinase Clk/Sty.* Mol Cell Biol, 2003. **23**(12): p. 4139-49.

60. Jiang, K., et al., *Akt2 regulation of Cdc2-like kinases (Clk/Sty), serine/arginine-rich (SR) protein phosphorylation, and insulin-induced alternative splicing of PKCbetaII messenger ribonucleic acid.* Endocrinology, 2009. **150**(5): p. 2087-97.

61. Duncan, P.I., et al., *In vivo regulation of alternative pre-mRNA splicing by the Clk1 protein kinase.* Mol Cell Biol, 1997. **17**(10): p. 5996-6001.

62. Garcia-Sacristan, A., et al., *Protein kinase clk/STY is differentially regulated during erythroleukemia cell differentiation: a bias toward the skipped splice variant characterizes postcommitment stages.* Cell Res, 2005. **15**(7): p. 495-503.

63. Virshup, D.M. and S. Shenolikar, *From promiscuity to precision: protein phosphatases get a makeover.* Mol Cell, 2009. **33**(5): p. 537-45.

64. Zhao, S. and E.Y. Lee, *A protein phosphatase-1-binding motif identified by the panning of a random peptide display library.* J Biol Chem, 1997. **272**(45): p. 28368-72.

65. Maynes, J.T., et al., *Crystal structure of the tumor-promoter okadaic acid bound to protein phosphatase-1.* J Biol Chem, 2001. **276**(47): p. 44078-82.

66. Xie, X.J., et al., *The N-terminal domain influences the structure and property of protein phosphatase 1.* Mol Cell Biochem, 2009. **327**(1-2): p. 241-6.

67. Terrak, M., et al., *Structural basis of protein phosphatase 1 regulation.* Nature, 2004. **429**(6993): p. 780-4.

68. Hartshorne, D.J., M. Ito, and F. Erdodi, *Myosin light chain phosphatase: subunit composition, interactions and regulation.* J Muscle Res Cell Motil, 1998. **19**(4): p. 325-41.

69. Varmuza, S., et al., *Spermiogenesis is impaired in mice bearing a targeted mutation in the protein phosphatase 1cgamma gene.* Dev Biol, 1999. **205**(1): p. 98-110.

70. Shi, Y. and J.L. Manley, *A complex signaling pathway regulates SRp38 phosphorylation and pre-mRNA splicing in response to heat shock.* Mol Cell, 2007. **28**(1): p. 79-90.

71. Novoyatleva, T., et al., *Protein phosphatase 1 binds to the RNA recognition motif of several splicing factors and regulates alternative pre-mRNA processing.* Hum Mol Genet, 2008. **17**(1): p. 52-70.

72. Wang, R. and A.O. Sperry, *Identification of a novel Leucine-rich repeat protein and candidate PP1 regulatory subunit expressed in developing spermatids.* BMC Cell Biol, 2008. **9**: p. 9.

73.   Logan, M.R., et al., *Genetic interaction network of the Saccharomyces cerevisiae type 1 phosphatase Glc7.* BMC Genomics, 2008. **9**: p. 336.

74.   Rossi, F., et al., *Specific phosphorylation of SR proteins by mammalian DNA topoisomerase I.* Nature, 1996. **381**(6577): p. 80-2.

75.   Tintaru, A.M., et al., *Structural and functional analysis of RNA and TAP binding to SF2/ASF.* EMBO Rep, 2007. **8**(8): p. 756-62.

76.   Hamelberg, D., T. Shen, and J.A. McCammon, *A proposed signaling motif for nuclear import in mRNA processing via the formation of arginine claw.* Proc Natl Acad Sci U S A, 2007. **104**(38): p. 14947-51.

77.   Haynes, C. and L.M. Iakoucheva, *Serine/arginine-rich splicing factors belong to a class of intrinsically disordered proteins.* Nucleic Acids Res, 2006. **34**(1): p. 305-12.

78.   Labourier, E., et al., *Interaction between the N-terminal domain of human DNA topoisomerase I and the arginine-serine domain of its substrate determines phosphorylation of SF2/ASF splicing factor.* Nucleic Acids Res, 1998. **26**(12): p. 2955-62.

79.   Shen, T., et al., *The folding energy landscape and phosphorylation: modeling the conformational switch of the NFAT regulatory domain.* FASEB J, 2005. **19**(11): p. 1389-95.

80.   Pelsue, S. and P.F. Agris, *Immunoreactivity between a monoclonal lupus autoantibody and the arginine/aspartic acid repeats within the U1-snRNP 70K autoantigen is conformationally restricted.* J Protein Chem, 1994. **13**(4): p. 401-8.

81.   Patwardhan, P. and W.T. Miller, *Processive phosphorylation: mechanism and biological importance.* Cell Signal, 2007. **19**(11): p. 2218-26.

82.   D'Souza, I., et al., *Missense and silent tau gene mutations cause frontotemporal dementia with parkinsonism-chromosome 17 type, by affecting multiple alternative RNA splicing regulatory elements.* Proc Natl Acad Sci U S A, 1999. **96**(10): p. 5598-603.

83.   Lin, C.L., et al., *Aberrant RNA processing in a neurodegenerative disease: the cause for absent EAAT2, a glutamate transporter, in amyotrophic lateral sclerosis.* Neuron, 1998. **20**(3): p. 589-602.

84.   Buxton, J., et al., *Detection of an unstable fragment of DNA specific to individuals with myotonic dystrophy.* Nature, 1992. **355**(6360): p. 547-8.

85.   Coovert, D.D., et al., *The survival motor neuron protein in spinal muscular atrophy.* Hum Mol Genet, 1997. **6**(8): p. 1205-14.

86.   Fischer, D.C., et al., *Expression of splicing factors in human ovarian cancer.* Oncol Rep, 2004. **11**(5): p. 1085-90.

87.   Ghigna, C., et al., *Altered expression of heterogenous nuclear ribonucleoproteins and SR factors in human colon adenocarcinomas.* Cancer Res, 1998. **58**(24): p. 5818-24.

88.   Stickeler, E., et al., *Stage-specific changes in SR splicing factors and alternative splicing in mammary tumorigenesis.* Oncogene, 1999. **18**(24): p. 3574-82.

89.   Zerbe, L.K., et al., *Relative amounts of antagonistic splicing factors, hnRNP A1 and ASF/SF2, change during neoplastic lung growth: implications for pre-mRNA processing.* Mol Carcinog, 2004. **41**(4): p. 187-96.

90.   Hayes, G.M., et al., *Targeting the RNA splicing machinery as a novel treatment strategy for pancreatic carcinoma.* Cancer Res, 2006. **66**(7): p. 3819-27.

91.   Kalnina, Z., et al., *Alterations of pre-mRNA splicing in cancer.* Genes Chromosomes Cancer, 2005. **42**(4): p. 342-57.

92.   Venables, J.P., *Unbalanced alternative splicing and its significance in cancer.* Bioessays, 2006. **28**(4): p. 378-86.

93.   Schwerk, C. and K. Schulze-Osthoff, *Regulation of apoptosis by alternative pre-mRNA splicing.* Mol Cell, 2005. **19**(1): p. 1-13.

94.   Hagiwara, M., *Alternative splicing: a new drug target of the post-genome era.* Biochim Biophys Acta, 2005. **1754**(1-2): p. 324-31.

95.   Ghigna, C., et al., *Cell motility is controlled by SF2/ASF through alternative splicing of the Ron protooncogene.* Mol Cell, 2005. **20**(6): p. 881-90.

96.   Schenk, P.W., et al., *Resistance to platinum-containing chemotherapy in testicular germ cell tumors is associated with downregulation of the protein kinase SRPK1.* Neoplasia, 2004. **6**(4): p. 297-301.

97.   Krishnakumar, S., et al., *SRPK1: a cisplatin sensitive protein expressed in retinoblastoma.* Pediatr Blood Cancer, 2008. **50**(2): p. 402-6.

98.   Hayes, G.M., P.E. Carrigan, and L.J. Miller, *Serine-arginine protein kinase 1 overexpression is associated with tumorigenic imbalance in mitogen-activated protein kinase pathways in breast, colonic, and pancreatic carcinomas.* Cancer Res, 2007. **67**(5): p. 2072-80.

99.   Szekelyhidi, Z., et al., *Synthesis of selective SRPK-1 inhibitors: novel tricyclic quinoxaline derivatives.* Bioorg Med Chem Lett, 2005. **15**(13): p. 3241-6.

100.   Muraki, M., et al., *Manipulation of alternative splicing by a newly developed inhibitor of Clks.* J Biol Chem, 2004. **279**(23): p. 24246-54.

101.   Fukuhara, T., et al., *Utilization of host SR protein kinases and RNA-splicing machinery during viral replication.* Proc Natl Acad Sci U S A, 2006. **103**(30): p. 11329-33.

102.   Winter, S.L., et al., *The interaction of PP1 with BRCA1 and analysis of their expression in breast tumors.* BMC Cancer, 2007. **7**: p. 85.

103.   Fujiki, H. and M. Suganuma, *Carcinogenic aspects of protein phosphatase 1 and 2A inhibitors.* Prog Mol Subcell Biol, 2009. **46**: p. 221-54.

104.   Koizumi, J., et al., *The subcellular localization of SF2/ASF is regulated by direct interaction with SR protein kinases (SRPKs).* J Biol Chem, 1999. **274**(16): p. 11125-31.

105.   Fiol, C.J., et al., *Ordered multisite protein phosphorylation. Analysis of glycogen synthase kinase 3 action using model peptide substrates.* J Biol Chem, 1990. **265**(11): p. 6061-5.

106.   Fiol, C.J., et al., *Phosphoserine as a recognition determinant for glycogen synthase kinase-3: phosphorylation of a synthetic peptide based on the G-component of protein phosphatase-1.* Arch Biochem Biophys, 1988. **267**(2): p. 797-802.

107.   Fiol, C.J., et al., *Formation of protein kinase recognition sites by covalent modification of the substrate. Molecular mechanism for the synergistic action of casein kinase II and glycogen synthase kinase 3.* J Biol Chem, 1987. **262**(29): p. 14042-8.

108.   Hernandez, F., et al., *Glycogen synthase kinase-3 plays a crucial role in tau exon 10 splicing and intranuclear distribution of SC35. Implications for Alzheimer's disease.* J Biol Chem, 2004. **279**(5): p. 3801-6.

109.   Patwardhan, P., et al., *Individual Cas phosphorylation sites are dispensable for processive phosphorylation by Src and anchorage-independent cell growth.* J Biol Chem, 2006. **281**(30): p. 20689-97.

110.   Hagopian, J.C., et al., *Adaptable molecular interactions guide phosphorylation of the SR protein ASF/SF2 by SRPK1.* J Mol Biol, 2008. **382**(4): p. 894-909.

111.   Kini, R.M. and H.J. Evans, *A hypothetical structural role for proline residues in the flanking segments of protein-protein interaction sites.* Biochem Biophys Res Commun, 1995. **212**(3): p. 1115-24.

112.   Schutkowski, M., et al., *Role of phosphorylation in determining the backbone dynamics of the serine/threonine-proline motif and Pin1 substrate recognition.* Biochemistry, 1998. **37**(16): p. 5566-75.

113.   Albert, A., S. Lavoie, and M. Vincent, *A hyperphosphorylated form of RNA polymerase II is the major interphase antigen of the phosphoprotein antibody MPM-2 and interacts with the peptidyl-prolyl isomerase Pin1.* J Cell Sci, 1999. **112 (Pt 15)**: p. 2493-500.

114.   Kofron, J.L., et al., *Determination of kinetic constants for peptidyl prolyl cis-trans isomerases by an improved spectrophotometric assay.* Biochemistry, 1991. **30**(25): p. 6127-34.

115. Dajani, R., et al., *Crystal structure of glycogen synthase kinase 3 beta: structural basis for phosphate-primed substrate specificity and autoinhibition.* Cell, 2001. **105**(6): p. 721-32.

116. Frame, S., P. Cohen, and R.M. Biondi, *A common phosphate binding site explains the unique substrate specificity of GSK3 and its inactivation by phosphorylation.* Mol Cell, 2001. **7**(6): p. 1321-7.

117. Grant, B.D., et al., *Kinetic analyses of mutations in the glycine-rich loop of cAMP-dependent protein kinase.* Biochemistry, 1998. **37**(21): p. 7708-15.

118. Hirai, T.J., I. Tsigelny, and J.A. Adams, *Catalytic assessment of the glycine-rich loop of the v-Fps oncoprotein using site-directed mutagenesis.* Biochemistry, 2000. **39**(43): p. 13276-84.

119. Endo, S., et al., *Conversion of protein phosphatase 1 catalytic subunit to a Mn(2+)-dependent enzyme impairs its regulation by inhibitor 1.* Biochemistry, 1997. **36**(23): p. 6986-92.

120. Barik, S., *Expression and biochemical properties of a protein serine/threonine phosphatase encoded by bacteriophage lambda.* Proc Natl Acad Sci U S A, 1993. **90**(22): p. 10633-7.

121. Jideama, N.M., et al., *Dephosphorylation specificities of protein phosphatase for cardiac troponin I, troponin T, and sites within troponin T.* Int J Biol Sci, 2006. **2**(1): p. 1-9.