Search of Inhibitors That Target HIV pre-mRNA Splicing to Overcome Drug Resistance

YU, Xiyao

A Thesis Submitted in Partial Fulfillment

of the Requirements for the Degree of

Master of Philosophy

in

Biochemistry

The Chinese University of Hong Kong

August 2012

Thesis/Assessment Committee

Professor Shaw Pang Chui (Chair)

Professor Ngo Chi Ki, Jacky (Thesis Supervisor)

Professor Au Wing Ngor, Shannon (Committee Member)

Professor Zheng Yongtang (External Examiner)

Abstract

Human immunodeficient virus (HIV) is a retrovirus that cause acquired immunodeficiency syndrome (AIDS). Highly active antiretroviral therapy (HAART) is a treatment of HIV infection that uses combinations of antiretroviral drugs and has achieved great success in the past two decades. However, since the reverse transcription process of viral RNA is notoriously prone to error, HIV-1 can acquire resistance to nearly all known inhibitors and has started to develop resistance to HAART. Therefore, there is an ongoing search for new drugs with novel inhibitory mechanism such as targeting cellular proteins essential for HIV-1 replication to overcome drug resistance of the virus.

HIV-1 mRNA undergoes complex splicing and the expression of the integrated HIV-1 provirus is largely dependent on the host's splicing machinery which assembly requires splicing factors such as serine-arginine rich proteins (SR proteins). Alternative splicing factor/splicing factor 2 (ASF/SF2), a prototypic SR protein that is essential for pre-mRNA splicing, has been shown to play critical roles during HIV-1 pre-mRNA splicing and replication. ASF/SF2, like other SR proteins, is phosphorylated by SR protein-specific kinases (SRPKs) at its C-terminal arginine/serine (RS) domain, which governs its localization and metabolism. Structural and functional studies of SRPK1 in complex with ASF/SF2 has revealed that a docking groove on SRPK1 that is distal to the active site interacts strongly with a docking motif and the RS domain of ASF/SF2, leading to high affinity binding as well as regulating the mechanism of phosphorylation.

In this study, we propose that by blocking this interaction, we might interfere the phosphorylation of ASF/SF2 and inhibit its activity during splicing of HIV-1 pre-mRNA.

Structure-based *in silico* screening method is adopted to identify potential inhibitors that bind to the docking groove of SRPK1 to block the binding and phosphorylation of ASF/SF2. The compound libraries being used include the Natual Products Database and SPECS database from ZINC (UCSF). 351,473 compounds have been screened using the program Autodock Vina as well as Autodock 4.0. Until now 50 potential candidates of inhibitor have been selected for biochemical analyses. *In vitro* kinase assays showed that six compounds exhibit inhibitory activity against the phosphorylation of ASF/SF2.

To test the effect of the selected inhibitors on the splicing of HIV-1 mRNA, *ex vivo* splicing assay has been performed. Current results showed that the synthesis of splicing products extracted from drug-treated cells was less efficient when compared to untreated cells. Biological assays testing the inhibitory effects of the compounds on viral infection are currently underway. Our preliminary result suggested that one of the compounds could indeed inhibit HIV-1 viral infection.

Other biochemical and biological analyses including structural study of kinase-inhibitor complexes to understand the mode of inhibition; measurement of binding kinetics using surface plasmon resonance spectroscopy (SPR); and biological assays testing the inhibitory effects of the compounds on replication are underway.

摘要

引發獲得性免疫缺陷綜合癥（AIDS）的人類免疫缺陷病毒（HIV）是一種逆轉錄病毒。過去的十餘年間，高效抗逆轉錄病毒治療療法（HARRT），在抗病毒感染方面取得了很大的成功。高效抗逆轉錄病毒治療療法是一種將多種抗逆轉錄病毒藥物複合的藥物聯用療法。然而，因為病毒的逆轉錄過程極易突變，導致HIV已經可以對大多數使用的抑製藥物產生抗藥性。因此，有越來越多的需要去尋找新型的抗病毒複製機理，例如將人體細胞蛋白作為載體，來達到克服病毒抗藥性的目的。

HIV-1的複製離不開宿主細胞的剪接因子，例如SR蛋白。選擇性剪接因子ASF/SF2，一個典型的調控pre-RNA剪接的SR蛋白，在HIV-1的pre-mRNA剪接和複製中起到了很重要的調控作用。ASF/SF2和其他SR蛋白一樣，都被丝氨酸/苏氨酸蛋白激酶（SRPK）磷酸化，磷酸化位點位於C端的丝氨酸/苏氨酸結構域（RS domain）。SRPK通過磷酸化來調節ASF/SF2在細胞中的分佈。對於SRPK 和ASF/SF2複合物的結構學和功能學研究指出，ASF/SF2的docking motif和SRPK1的遠離活性位點的docking groove存在很強的相互作用。而這種相互作用是調節磷酸化過程關鍵。所以，在我們的研究過程中，我們希望通過阻斷2個蛋白的相互作用來干擾ASF/SF2的磷酸化，進而抑制其在HIV-1 pre-mRNA剪接過程中的活性。

我們採用以結構為基礎的藥物模擬篩選，來選擇潛在的抑制物，達到通過抑制物與docking groove的相互作用來阻斷ASF/SF2和SRPK1的相互作用，以達到抑制磷酸化的目的。我們使用的數據庫來自于ZINC數據庫（UCSF），包括天然產物數據庫和SPECS。我們採用AutoDock Vina 和AutoDock 4.2 二個模擬軟件來栓選數

據庫中351473个化合物。并從中選出50個潛在的化合物用作之後的化學生物學測試。體外的激酶活性試驗顯示，6個化合物對ASF/SF2的磷酸化有抑製作用。

體外的HIV-1 pre-mRNA剪接實驗顯示，5個化合物在逆轉錄PCR（RT-PCR）中有一定得抑制效果。和DMSO對照組相比，在抑製劑作用下剪接產物的生成被抑制。HIV-1病毒合胞體感染實驗顯示，有一個化合物對病毒的感染起到了一定的抑制作用。

其他的測試實驗還在進行中，包括對 SRPK1 和抑制物複合物的結構研究，從而更好的研究抑制物的作用機理。以及，採用表面等離子共振波譜來進行動力學研究和其他關於化合物在病毒複製過程中的實驗測試。

Acknowledgements

Firstly, I would like to express my sincere thank to my supervisor, Professor Ngo Chi Ki, Jacky for his patience guidance and providing help and suggestion throughout my M.Phil. study. In these two years, he has not only taught me the way of thinking, but also provided me with best technical supports.

Besides, I want to thank my labmates, including Dr. Mars Tao, Mr. Liang Ning, Mr. Yong Shang, Ms. Wendy Suo and Ms. Jenny Lau for their help and encouragement when I encounter difficulties in my research work. I am also thankful for Ms. Winnie Poon, for giving me technical supports of surface plasmon resonance. Moreover, I want to thank Ms. Xuemei Liao, Ms. Minghui Gao and Mr. Lin Wang for their help and suggestion during my study and life. Allow me to express my gratitude to my best friend Ms. Qidi Sun for sharing laughter and tears in the past 6 years.

Last but not least, I wish to thank my family, especially my parents for their unlimited supports and love.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter Ⅰ: Introduction

1.1 HIV, HAART and HIV Drug Resistance

Human immunodeficiency virus (HIV) is a lentivirus (a member of the retrovirus family) that causes acquired immunodeficiency syndrome (AIDS) (Weiss, 1993). It is an enveloped virus (Figure 1.1) which envelope is derived from the host cell membrane and contains viral glycoproteins such as gp120 and gp41 that are inserted into the membrane as the virus leaves the cell. Inside the core are two identical copies of single-stranded RNA viral genome and three enzymes: reverse transcriptase, protease and integrase. To establish infection, HIV must first attach to its host cell. Attachment occurs when gp120 interacts with CD4 antigen receptor on the surface of the host cell with the help of different co-receptors (Chan and Kim, 1998). Then the viral envelop and the host cell membrane fuses, resulting in the entry of the virus into the cell (Wyatt and Sodroski, 1998). The viral RNA genome of HIV is then released into the cytoplasm of the host cell. The viral reverse transcriptase makes a DNA copy of the viral RNA genome (Zheng et al., 2005) and the DNA is then transported into the nucleus and inserted into the host cell chromosome by the viral integrase enzyme, forming the proviral DNA. Transcription of the proviral DNA is carried out by the host cell's transcription machinery and the newly synthesized viral mRNA is then translated, yielding viral enzymes and structural proteins that are required for the packaging of new virus particles. Some of these functional proteins are formed by cleavage of long poly-proteins by the viral protease. Finally, the virion is released from the host cell by budding off (Gelderblom, et al., 1997) (Figure 1.2).

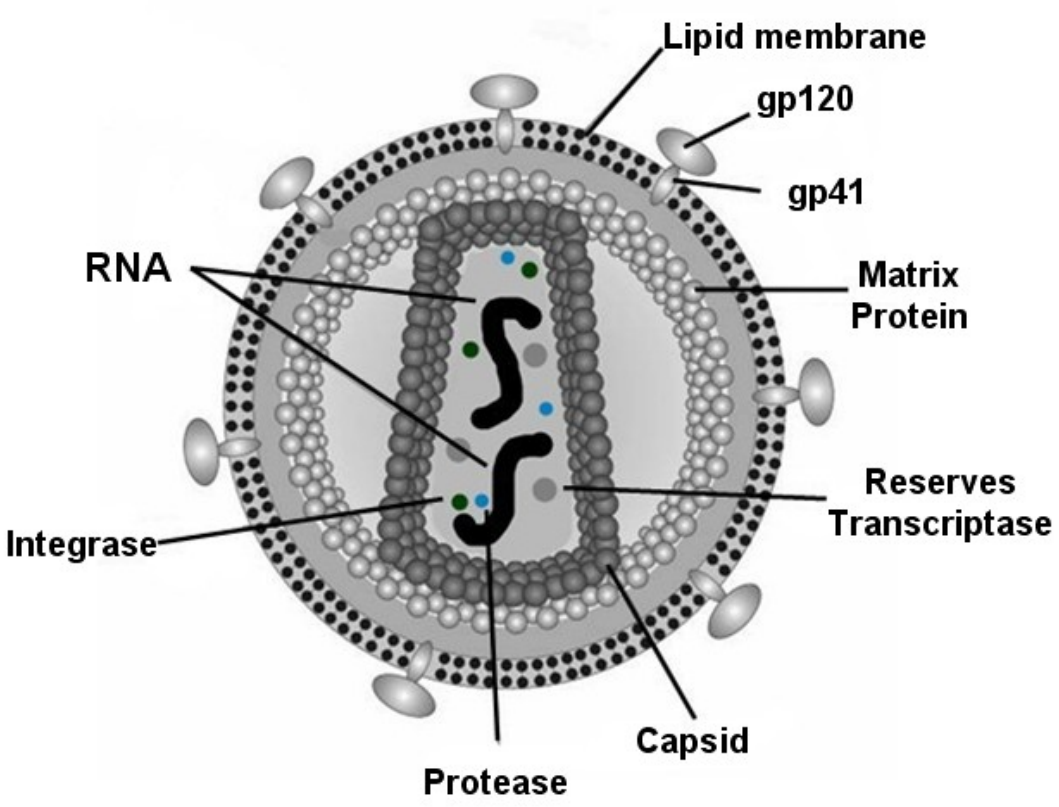


Figure 1. 1 Structure of HIV-1 virus (http://en.wikipedia.org/wiki/File:HIV_gross_ cycle_only.png). The virol envelope is derived from the host cell membrane with viral glycoproteins such as gp120 and gp41 inserted into the lipid bilayer. Inside the core are two identical copies of single-stranded RNA viral genome and three enzymes, reverse transcriptase, protease and integrase.

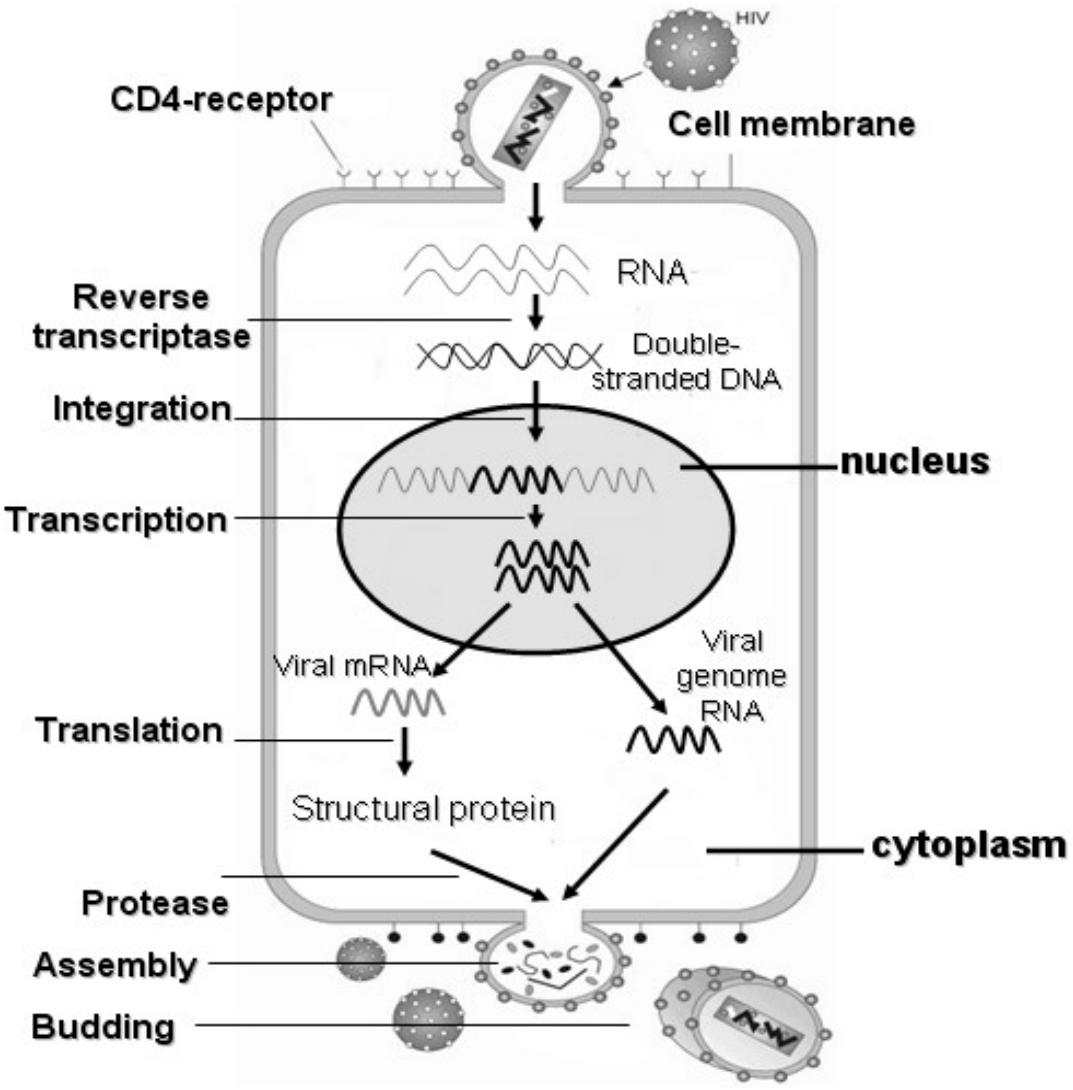


Figure 1. 2 The HIV replication cycle (http://en.wikipedia.org/wiki/File: HIV_ gross _cycle_only.png). Firstly, to establish infection, HIV attaches to its host cell through interaction with CD4-receptor. Then the viral envelope and the host cell membranes fuse, resulting in the entry of the virus into the cell. After the RNA is released into the cytoplasm of the host cell, reverse transcriptase makes a DNA copy of the viral RNA genome, which is later inserted into the host cell chromosome by integrase. After translation, some of the functional protein is formed by cleavage of a long poly-protein by the viral protease. Finally, the fully assembled virion is released by budding.

According to the estimations made by the World Health Organization (WHO) in 2011, about 33.3 million people are living with HIV. The growth rate of the total number of HIV infections has slowed down since the beginning of 21st century; meanwhile, the number of the AIDS-related deaths also falls rapidly. The reduction in AIDS-related deaths is mostly contributed by the significant scale up of antiretroviral therapy over the past few years (figure 1.3) (WHO Progress report, 2011).

In particular, highly active antiretroviral therapy (HAART), which is introduced in 1997 uses combinations of antiretroviral drugs for the treatment of HIV infection (Gelderblom, 1997). Currently, drugs being used in HAART to treat HIV-1 infection could be classified into five distinct classes: entry inhibitors, nucleoside and nucleotide analogues, nonnucleoside reverse transcriptase inhibitors, integrase inhibitors and protease inhibitors. Entry inhibitors aim to block HIV viron from penetrating into the target cells by interfering with the interactions between gp120 and host cells' receptors (Dybul et al., 2002). Nucleoside and nucleotide analogues act as DNA-chain terminators and inhibit the reverse transcription of the viral RNA genome into DNA; while nonnucleoside reverse-transcriptase inhibitors bind and inhibit the viral reverse transcriptase directly. Integrase inhibitors block the integration. Finally, protease inhibitors are designed to inhibit the proteolytic activity of the viral protease that is essential for viral particles assembly.

The usual HAART regimen combines three or more different drugs such as two nucleoside reverse transcriptase inhibitors (NRTIs) and a protease inhibitor (PI), two NRTIs and a non-nucleoside reverse transcriptase inhibitor (NNRTI) or other similar

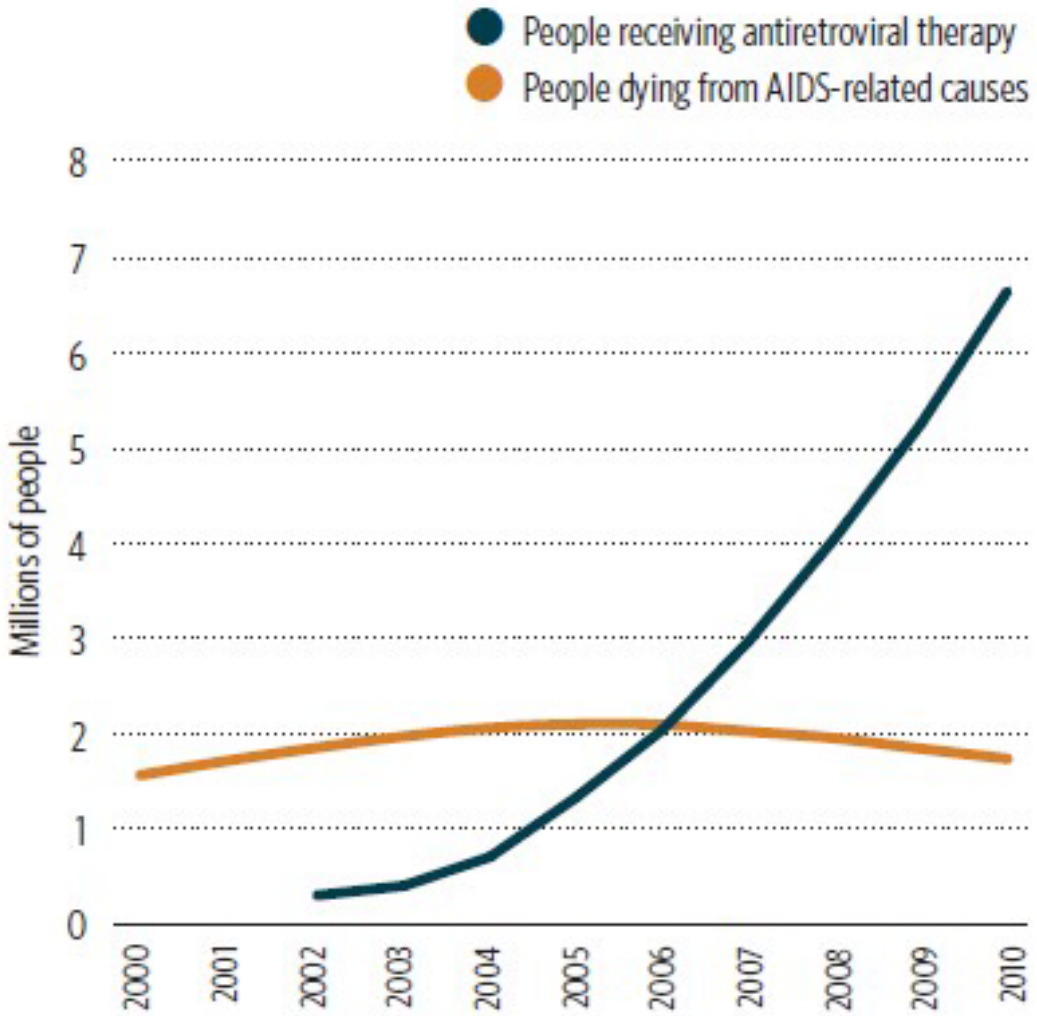


Figure 1. 3 Number of people with access to antiretroviral therapy and the number of people dying from AIDS-related causes (WHO Progress report, 2011).

combinations. These HAART regimens, commonly known as cocktail therapy, have proven to reduce the amount of active viruses in patients and in some cases to an extent where it is undetectable by current blood testing techniques (Henkel, 1999).

However, HAART is not bullet-proof and sometimes HIV drug resistance is acquired by viral mutation that promotes virus replication even in the presence of antiretroviral drugs, leading to antiretroviral therapy (ART) failure, consequently the spread of resistant strains of HIV and the need to develop new anti-HIV drugs (Geretti, 2006). The increasing prevalence of drug resistance compromised the benefits of the use of HAART and challenges the remarkable treatment of HIV-1-infected patients (Barbaro et al., 2005; del Rio, 2006).

As mentioned, HIV is a retrovirus which genome consists of RNA. By using reverse transcriptase, HIV virus synthesizes DNA from its single-strand RNA genome through a reverse transcription (RT) process that lacks a mechanism for correcting errors made while copying its genome (Freeman, 2007). The reverse transcription of viral RNA into DNA is notoriously prone to error, introducing on average one mutation for each viral genome transcribed, resulting that HIV replicates its genome with the highest known mutation rate of any pathological agents (Freeman, 2007). These mutations accumulate and lead to great genetic variation within populations of virus, while creating an ideal situation and raw material for natural selection over generations (Freeman, 2007). Selected virions with higher fitness would survive (Kozal, 2009), resulting in drug resistance virions that escape the harmful effects of anti-HIV drugs, which then reproduce and create an entirely new, drug resistant population. An

investigation report by WHO in 2004 estimated 76.3% of HIV-positive population in USA is suffering with some form of drug resistance (Richman et al., 2004).

Due to drug resistance caused by the error-prone reverse transcription process, different mechanisms have also been characterized for individual class of drug or even for drugs from the same class.

For instance, non-nucleoside reverse-transcriptase inhibitors work as potent and highly selective inhibitors of HIV-1 reverse transcriptase. These small molecules bind to the less conserved hydrophobic pocket located near the active site in the p66 subdomain of the reverse transcriptase with strong affinity. By interfering the flexibility of the enzyme, the drugs subsequently distort its ability to bind RNA and block DNA synthesis (Esnouf et al., 1997). However, mutations could occur through selection in the targeted pocket after prolonged treatment with the nonnucleoside reverse transcriptase inhibitors. These drugs then lose their specific interactions with their targeted docking pocket (Boyer et al., 1993). In particular, resistance to the treatment with a combination of efavirenz (EFV) and nevirapine (NVP), both belong to the non-nucleoside reverse transcriptase inhibitors, is typically associated with the mutations L100I, Y181C/I, K103N, V106A/M, V108I, Y188C/H/L and G190A/S on the reverse transcriptase.

The treatment regimens of HIV/AIDS currently used in HARRT are mainly targeting the viral specific enzymes and proteins, such as reverse transcriptase, protease, and gp120. Because of multidrug-resistant, toxic side effects, poor patient compliance due to complicated regimens，there is an immediate need for the search of new drugs with novel inhibitory mechanism to overcome drug resistance of the virus. For instance,

apart from the targeted viral-specific proteins, various cellular proteins also play critical roles during HIV-1 replication. Research on these essential host cell factors might provide the basis for the development of novel drugs that target cellular proteins that are not affected by the HIV-1 reverse transcription process, thus block the replication of the HIV-1 virus on the premise to reduce the burden of spread of multidrug-resistant HIV-1.

1.2 HIV-1 alternative splicing mechanism

After the HIV-1 genome is reverse transcribed and integrated into the host cell genome, a primary transcript of 9 kb pre-mRNA is transcribed from the integrated viral DNA by RNA polymerase II. This primary transcript serves as genomic RNA for progeny virus and the mRNA that encodes the viral Gag and Pol proteins. However, more importantly, the transcripts contains multiple splicing sites and generates over 40 different spliced mRNAs by alternative splicing (McLaren et al., 2008). HIV-1 pre-mRNA splicing is a highly orchestrated process by which the virus regulates proper cytoplasmic levels of both spliced mRNAs for viral protein synthesis and unspliced viral RNA to serve as both genome RNA as well as mRNA for translation.

Over the past two decades, there is significant progress on the study of mechanisms on how HIV-1 regulates its pre-mRNA splicing (Lutzelberger et al., 2006b; Pollard and Malim, 1998). Most introns start with the sequence GU and end with the sequence AG

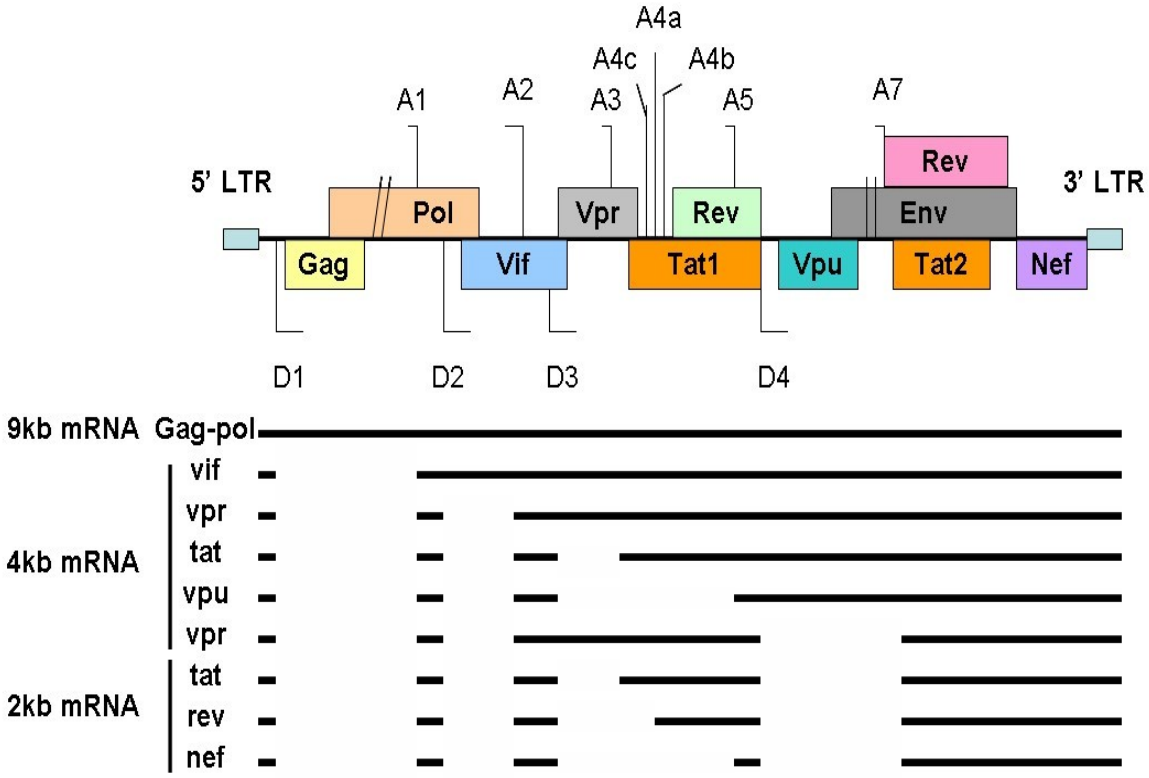


Figure 1. 4 Schematic organisation of the HIV-1 genome and different mRNA splicing products. HIV-1 5'splice sites (D1, D2, D3 and D4) and 3' splice sites (A1, A2, A3, A4c,a,b, A5 and A7) are indicated. The various open reading frames of the HIV-1 genome are denoted by different coloured boxes.

(in the 5' to 3' direction), which denominate the 5’ splice donor (Figure 1.4 D1–D4) and 3’ splice acceptor site (Figure. 1.4, A1–A7) respectively. These consensus splice site sequences are usually required for the recognition by the cellular spliceosome during constitutive splicing (Lutzelberger et al., 2006a).

To express all key viral proteins, HIV-1 uses a combination of several alternative 5′ and 3′ splice sites to generate two sorts of spliced mRNA species that can be classified as multiply spliced (~2 kb) and singly spliced (~ 4 kb) RNAs (Fig. 1.4) (Purcell and Martin, 1993). Five 3’ss (A3, A4c, A4a, A4b and A5) serve as the splice acceptor sites for the multiply spliced mRNA to produce mRNAs for Tat, Rev, or Nef proteins respectively during the early stage of HIV-1 gene expression. Sort of Tat (trans-activator of transcription) mRNA uses the A3 splice site, while the Nef (negative regulatory factor) mRNAs are spliced at site A5 (Stoltzfus and Madsen, 2006). On the other hand, the Rev (regulator of virion expression) mRNAs are spliced respectively at sites A4a, A4b or A4c. Only when these splice sites are utilized accurately during alternative splicing that the encoded functional proteins can be expressed and support the viral life cycle of HIV-1. The main function of Tat proteins is to extensively increase the level of transcription of the HIV dsRNA and plays a more direct role in the HIV disease process. The expression of Nef modulates T cell activation and ensures a persistent state of infection. Nef also facilitates the survival of host cells by down-modulating the expression of several surface molecules to satisfy the requirements of the HIV-1 viral life cycle. The Rev protein, a product of the doubly spliced mRNA, facilitates the expression of HIV proteins by controlling the export rate

of singly and unspliced mRNAs from nuclear to cytoplasm (Kim et al., 1989; Klotman et al., 1992). Completely spliced mRNAs like Rev in general will be translated in the cytoplasm first because they are smaller and exported faster. Therefore, Rev is particularly important in the regulation the expression of, the Gag and Gag-Pol polyproteins, as well as Vif, Vpr, Vpu and Env proteins by up-regulating the nuclear export of the incompletely spliced mRNAs that encode them.

The process of alternative splicing is tightly regulated by a system of cis- and trans-regulatory elements. In HIV virus, the choice of alternative splice sites is dependent on the presence of cis-regulatory elements termed exonic splicing enhancers (ESEs) and/or exonic⁄intronic splicing silencers (ESS／ISS) on the pre-mRNA due to the weak 3′ ss (Caputi and Zahler, 2002; Staffa and Cochrane, 1994; Tange et al., 2001). ESEs and ESSs bind with trans-acting cellular splicing regulatory factors, which either promote or inhibit the splicing at neighbouring splice sites, respectively (Zahler A.M. et al., 2004). Some of these trans-acting factors include members of the serine-arginine rich protein (SR protein) and heterogenous nuclear ribonucleoproteins (hnRNPs) families. During alternative splicing of HIV-1 pre-mRNA, members of the hnRNP family play a role in the downregulation of splicing at the 3'ss A1 (Kammler et al., 2006), A2 (Bilodeau et al., 2001), and A3 (Caputi et al., 1999; Jacquenet et al., 2001). On the other hand, binding of the SR proteins to exonic splicing enhancers (ESEs) can activate splicing at the 3'ss A2, A3, A5 and A7 (Zhu et al., 2001)

Since, both SR proteins and hnRNP families play crucial roles in the HIV-1 pre-mRNA splicing and consequently important for viral propagation, these proteins thus represent potential drug targets for the development of antiretroviral therapies.

1.3 SR Protein Family

Serine/arginine-rich proteins (SR proteins) constitute a family of splicing factor that is involved in the regulation of constitutive splicing and also influences the selection of the alternative splice sites during alternative splicing in eukaryotic mRNA. The first SR protein was identified by Mark Roth in 1990 and this family of proteins was given the name SR proteins because they contain long repeats of serine and arginine dipeptide at their C-terminal regions (Madhani and Guthrie, 1994; Sharp, 1994).

The primary structures of all SR proteins follow the consensus that they either contain one or two N-terminal RNA-recognition motif (RRMs; also known as an RNA recognition domain [RBD]), and a C-terminal region enriched in different lengths of Arg–Ser dipeptides (Fu and Maniatis, 1992; Zahler et al., 1992). To date, 12 human SR proteins have been identified and classified as classical SR proteins. They are listed in Table 1.1.

Based on the number of RRM(s), these 12 SR proteins could be further classified into two different groups. Group 1 proteins, consisting of SFRS1, SFRS4, SFRS5, SFRS6, and SFRS9, have two RRMs and contain a conserved SWQDLKD motif in

Table 1. 1: SR protein / gene symbol

Previous gene symbol	Aliases	New protein/gene symbol
SFRS1	ASF/SF2, SRP30a	SRSF 1
SFRS2	SC35, PR264, SRp30b	SRSF 2
SFRS3	SRp20	SRSF 3
SFRS4	SRp75	SRSF 4
SFRS5	SRp40, HRS	SRSF 5
SFRS6	SRp55, B52	SRSF 6
SFRS7	9G8	SRSF 7
SFRS2B	SRp46 (human only)	SRSF 8
SFRS9	SRp30c	SRSF 9
SFRS13A	TASR1, SRp38, SRrp40	SRSF 10
SFRS11	P54, SRp54	SRSF 11
SFRS13B	SRrp35	SRSF12

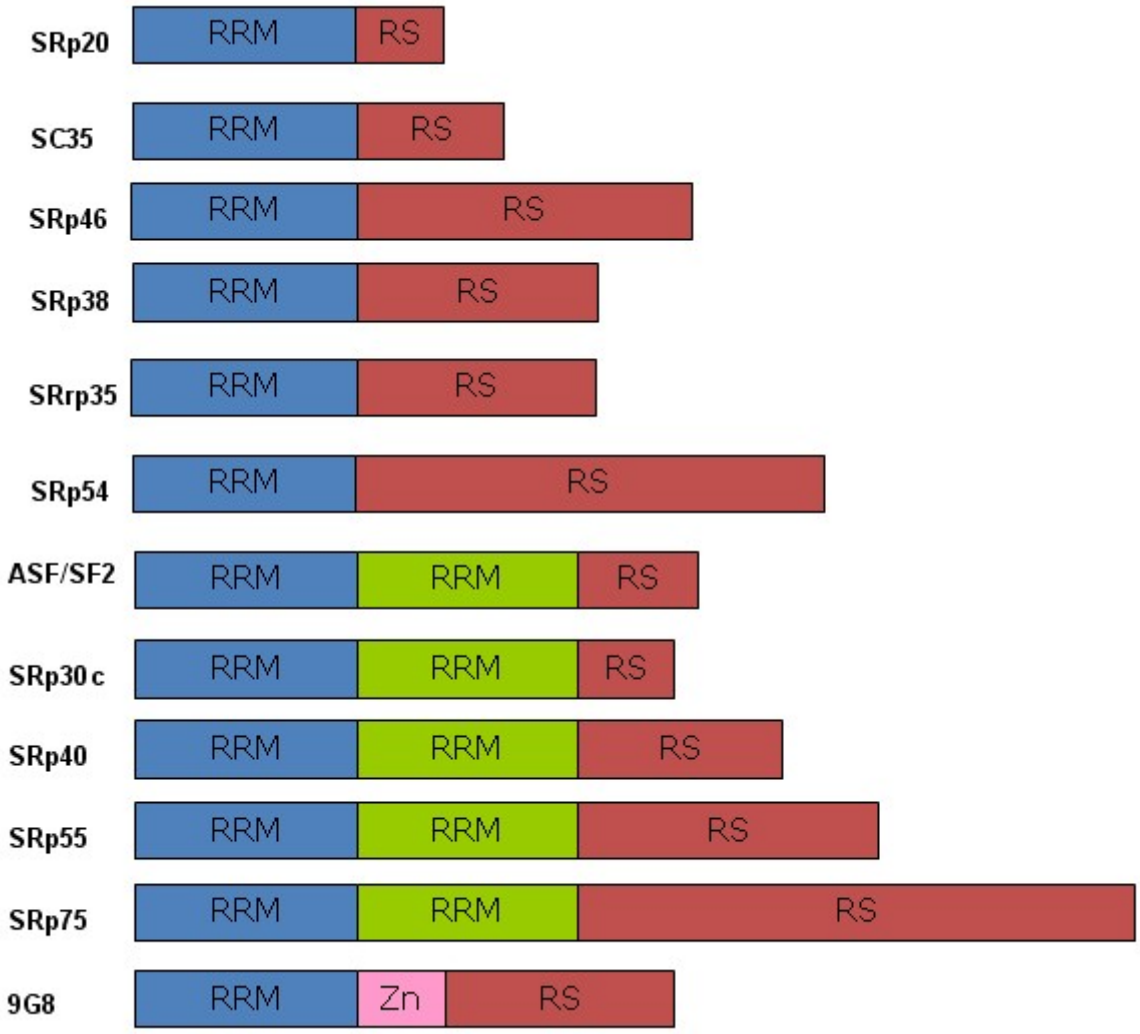


Figure 1. 5 Human SR protein family. The schematic representations of the structural organization of nine human SR proteins are shown. RRM abbreviated for RNA recognition motif; RS, arginine/serine-rich domain; Zn, zinc finger.

their second non-canonical RRM. In comparison, group 2 SR proteins, which include SFRS13A, SFRS11, SFRS2B, SFRS2, SFRS3, SFRS7, and SFRS13B, contain only a single RRM and lack the SWQDLKD motif (Figure 1.5) (Manley and Krainer, 2010).

While the N-terminal RRMs are responsible for sequence-specific RNA binding, the RS domains of SR proteins are believed to function as a protein-RNA and protein-protein interaction domain. The RS domain likely mediates interaction with other splicing proteins including other SR proteins, SR-related proteins(Blencowe et al., 1999), and components of the general splicing machinery (Kohtz et al., 1994; Wu and Maniatis, 1993), to enhance the assembly of the spliceosome (Fetzer et al., 1997; Makarova et al., 2001; Teigelkamp et al., 1997). Furthermore, the RS domain also participates in a nuclear localization pathway by mediating the interaction between the SR proteins and the β-karyopherin family member transportin-SR (Caceres et al., 1997; Kataoka et al., 1999; Lai et al., 2000).

SR proteins predominantly locate in the nucleus (Cowper et al., 2001; Sakashita and Endo, 2010). A subset of SR proteins, like ASF/SF2, 9G8 and SRP20, shuttle continuously between the cytoplasm and the nucleus at a high dynamic rate with the help of transportin-SR (Caceres et al., 1997; Caceres et al., 1998b), which contacts with the RS domains to convey the SR proteins through the nuclear pore in a phosphorylation-dependent manner (Koizumi et al., 1999; Ngo et al., 2005) .

1.4 Functional roles of SR protein in HIV pre-mRNA splicing

In general cases, SR proteins are essential for both the 5' and 3' splice sites selection in alternative splicing as well as stabilizing interactions between components of the splicing machinery during constitutive splicing. During alternative splicing, SR proteins usually recognize and bind to the exonic splicing enhancers (ESEs) and act as general activators of exon definition (Hertel et al., 1997; Lam and Hertel, 2002), which then regulate the recruitment of the splicing machinery to the adjacent splice sites. This process is mediated by at least one member of the SR protein family (Black, 2003; Blencowe, 2000; Graveley, 2000).

Based on the conservation of the splicing pattern in different HIV expressing cells (Purcell and Martin, 1993), splicing of viral pre-mRNA must be tightly regulated and is crucial to the HIV virus life cycle (Gorry et al., 1999; Sonza et al., 2002). In fact, SR proteins ASF/SF2, SC35, 9G8 and SRp40 have been proven to play active roles in mediating the balance in the HIV-1 splicing pattern *in vitro* and *ex vivo* (Mayeda et al., 1999; Zahler et al., 2004). For instance, the expression levels of Env mRNA, Nef protein synthesis and full-length viral RNA decrease obviously when the SR proteins ASF/SF2, SC35 and SRP40 are overexpressed. Such disturbed viral RNA splicing pattern would lead to weaker virus infectivity and SR proteins overexpression can induce the inhibition of virion production (Jacquenet et al., 2005). Among all the splice sites in the HIV-1 pre-mRNA transcript, splice sites A1 and A2 are specifically activated by ASF/SF2 although site A1 is less sensitive to the regulation. Previous studies have shown that binding between the SR protein ASF/SF2 and ESE1 activates the HIV splicing at the A2, while hnRNP A/B represses this process (Calarco et al., 2009;

Nakagawa et al., 2005; Papoutsopoulou et al., 1999). As a result, overexpression of ASF/SF2 increases the level of Vpr mRNA specifically at the cost of the reduction of other mRNA species. On the other hand, SC35 and SRp40 strongly activate site A3 and their overexpression results in an accumulation of tat mRNA and disturbance of the splicing of other viral mRNAs (Ropers et al., 2004). These results suggest that individual SR proteins might have specific roles during the splicing of HIV-1 pre-mRNA and their levels must be tightly regulated in order for HIV-1 to complete its life cycle.

Beside Site A2, Splice site A5, which is utilized for the production of singly spliced Env mRNA is also regulated by SF/ASF2. SF2/ASF functions as a trans-acting factor for the ESE sequence upstream of D4 and the ESE3 sequence as well (Dagher and Fu, 2001; Lim and Burge, 2001) (Table 1.2). Furthermore, HIV-1 3'ss A7 is negatively regulated by the ISS, ESS3 and ESE3. ESE3 actually functions as both splicing silencer and enhancer activities, for it can interact with both hnRNP A1 and SF2/ASF (Colwill et al., 1996b; Nieratschker et al., 2009; Wang et al., 1999).

1.5 Phosphorylation States of SR Proteins

SR proteins undergo dynamic phosphorylation and dephosphorylation during the pre-mRNA splicing process (Xiao and Manley, 1998). Phosphorylation of SR proteins has been well documented to occur in the C-terminal RS domain. Different phosphorylation states of SR protein regulate how the SR protein functions during the

Table 1. 2: Summary of concerning regulatory sites and the regulatory factors involved

3' splicing site	Regulatory elements	regulatory factors involved	position
A2	ESSV	hnRNP A1	4995-5017
	ESE1	ASF/SF2	
A3	ESSp	hnRNP H	5362-5366
	ESS2	hnRNP A1	5428-5437
	ESE2	SC35, SRp40	5558-5582
A5	ESE GAR	ASF/SF2, SRp40	8047-8062
A7	ESS3	hnRNP A1	7928-7950
		hnRNP E1/E2	
	ISS	ASF/SF2	8016-8025
	ESE3	hnRNP A1	8018-8025

assembly of spliceosome, splicing, and export of mRNA (Caceres et al., 1997; Hedley et al., 1995). RS domain phosphorylation also regulates the subcellular localization of SR proteins (Figure 1.6). SR proteins predominantly reside in nuclear speckles that serve as the sites for storage and/or reassembly of splicing factors. Hyperphosphorylation of the RS domain causes redistribution of SR proteins from nuclear speckles to the nucleoplasm and recruit them to the splicing sites (Lai et al., 2001). Furthermore, as aforementioned, some members of the SR protein family like ASF/SF2 continuously shuttle between the nucleus and the cytoplasm in a phosphorylation dependent manner (Caceres et al., 1998a).

In fact, the nuclear import of SR proteins mediated by transportin-SR has been shown to depend on the phosphorylation state of the SR protein. Hypophosphorylation of the RS domain by the SR protein kinase SRPK1 drastically enhances the binding affinity between the SR proteins and transportin-SR (Lai et al., 2000). Thus, phosphorylation plays a key role in the regulation of the biological functions of SR proteins and subsequently splicing.

1.6 SR protein Kinase

Several protein kinases have been reported to phosphorylate multiple serines in the RS domain of SR proteins, including the SR protein kinase (SRPK) family, Cdc2-like kinase/Ser-Thr-Tyr (Clk/Sty) kinase family, DNA topoisomerase I and Akt (Blaustein et al., 2005; Colwill et al., 1996a; Gui et al., 1994; Rossi et al., 1996). Serine-arginine

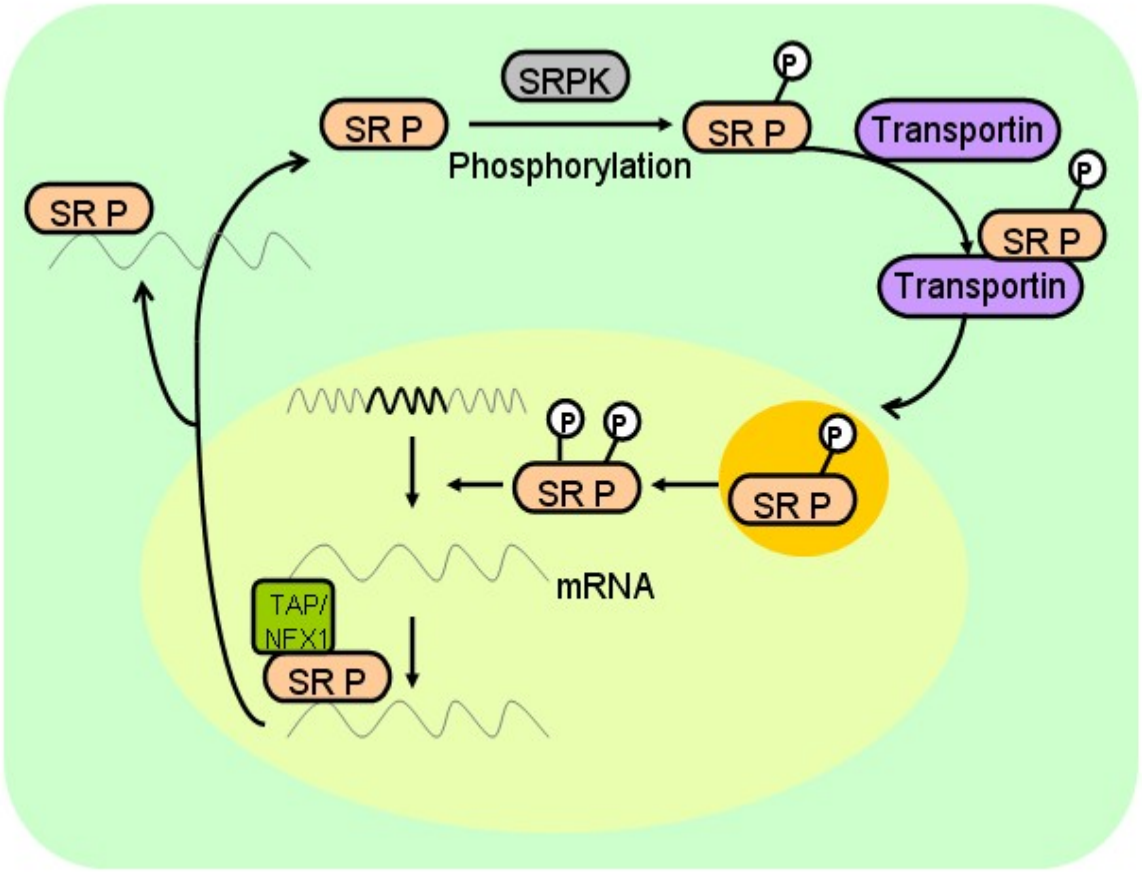


Figure 1. 6 Regulation of SR protein localization by SRPK.

protein kinases (SPRKs) predominantly phosphorylate SR proteins *in vivo* and alter their cellular distribution and activities. Members of SRPKs family are highly RS-specific kinases and have been shown to phosphorylate serine, but not threonine residues, that are immediately adjacent to arginine, but not lysine residues *in vitro* (Gui et al., 1994). Furthermore, biochemical and structural research on SRPK1 and Sky1p have provided evidence that SRPKs are constitutively active kinases (Ngo et al., 2007a; Nolen et al., 2001). These proteins, when purified from *E.coli*, are fully catalytically active *in vitro* and do not need any post-translational modification (Cao and Garcia-Blanco, 1998;Papoutsopoulou et al., 1999).

The first serine-arginine (SR) protein kinase, SRPK1, was characterized and purified in 1994 (Lai et al., 2000). In 1998, the cloning and the characterization of SRPK2, which is highly related to SRPK1 in kinase core sequence, were performed (Kuroyanagi et al., 1998). Another member of SRPKs, SRPK1a, which is encoded by an alternative spliced SRPK1 transcript that retains 171 amino acids at its N-terminal domain, was reported in 2001 (Nikolakaki et al., 2001). Finally, mouse SRPK3 was identified as a downstream target gene of the transcription factor myocyte enhancer factor 2 (MEF2) in 2005 and was shown to specifically expressed in the heart and skeletal muscle (Nakagawa et al., 2005). Another Porcine SRPK3 gene encoding a 566 amino acid protein was isolated in 2011 (Xu et al., 2011).

SRPKs have been proven to phosphorylate more than ten members from the SR protein family both *in vitro* and *in vivo*. By influencing the phosphorylation states and

the localization of these SR proteins, it is suggested that SRPKs may indicatively regulate multiple facets of cell cycle, including mRNA maturation, cell cycle progression, chromatin reorganization and metabolic signaling.

1.7 Interaction between SRPK1 and ASF/SF2

ASF/SF2 (human alternative splicing factor) is the most well-studied member of the SR protein family. It comprises two RRM domains (RRM1 and RRM2) and a C-terminal RS domain. The RS domain of ASF/SF2 contains 20 serines where 15 of which are directly adjacent to arginine residues. It has been shown that SR protein kinase 1 (SRPK1) phosphorylates approximately 12 of serines in the RS domain using a directional (C-terminal-to-N-terminal) and processive mechanism.

The crystal structure of SRPK1ΔNS1, an active truncation of human SRPK1 with 41 amino acids from the N terminus and the spacer residues (aa 256–473) deleted, has been solved in complex with a 9 mer peptide ($RRR^{\mathbf{1}}E^{2}R^{\mathbf{3}}S^{4}P^{5}T^{6}R^{\mathbf{7}}$) derived from an SR-like protein Npl3p (Figure 1.7) (Ngo et al., 2005). The structure reveals that SRPK1ΔNS1 shares the bilobal fold similar to other protein kinases. The small lobe is composed mostly of β-strands while the larger lobe is mostly α-helical. SRPK1, belonging to the CMGC group of kinases, contains a mitogen-activated protein kinase (MAPK) insert that connects helices αG and αH. This MAP kinase insert, together with the loop connecting helices αF and αG, and the helix αG of the kinase region form a

docking groove that is distal to the active site of the kinase and interacts with the soaked RS peptide. Further biochemical studies on the function of this docking groove revealed that it is essential to the interaction between SRPK1 and ASF/SF2, and is responsible for mediating the processive phosphorylation of ASF/SF2.

X-ray structure of another truncation construct of SRPK1, which includes residues 69–235 and 475–655, in complex with the core binding domain of ASF/SF2, comprises the RRM2 and part of the RS domain (residues 121–196 and 201–210 respectively), reveals more structural features that are important for the interaction between the kinase and substrate and the mechanism of the phosphoralation (Ngo et al., 2008). Most importantly, this structure confirms that the docking groove of SRPK1 binds to arginine-rich peptide (derived from the RS domain) and is critical for the interaction and processive phosphorylation of ASF/SF2 (Ngo et al., 2008). This process of processive phosphorylation finally directs ASF/SF2 to be transported into the nucleus; whereas interference of the kinase-substrate interaction results in aberrant phosphorylation, thereby disturbing the subcellular localization of ASF/SF2 (Ngo et al., 2005).

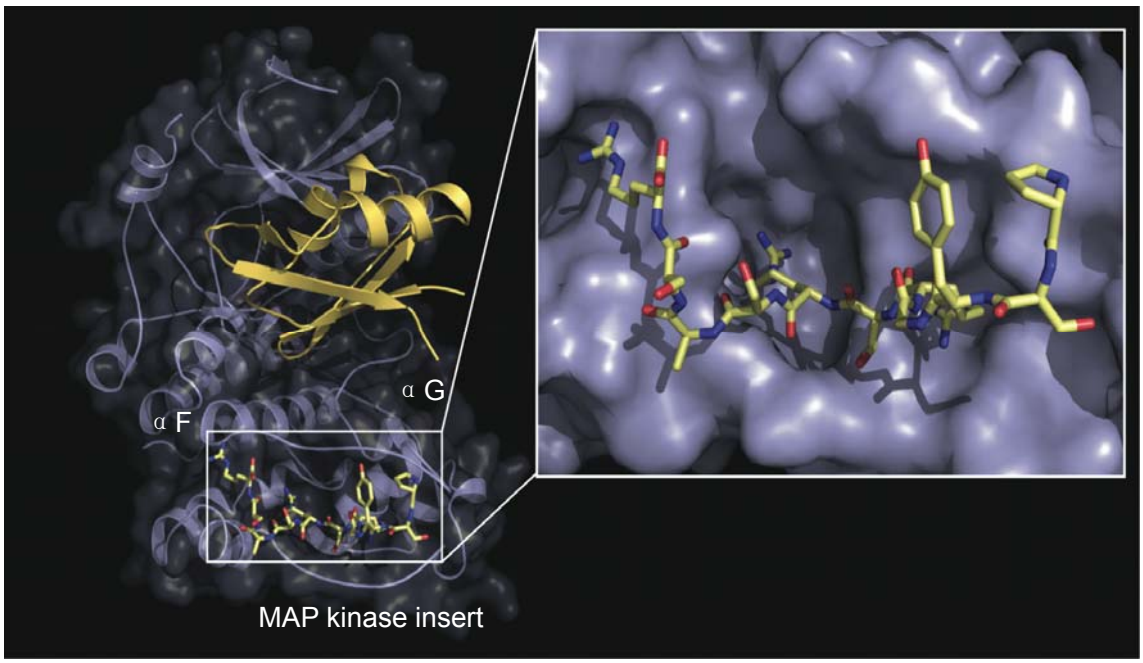


Figure 1. 7 Docking groove of SRPK1. Surface rendition of SRPK1 bound to ASF/SF2. The observed peptide is the RS1 domain of ASF/SF2, which binds to a groove formed by helices αG, αF and the MAP kinase insert of SRPK1.

1.8 IDC16 and SPRIN340

It is clear that members of the SR protein family play an essential role in the regulation of HIV-1 pre-mRNA splicing, implying that they could be potential drug targets for HIV-1. In 2007, Jamal Tazi and his team have shown proof that targeting HIV-1 splicing pathway could be a new approach of antiretroviral therapies. The team discovered a series of benzopyridoindole and pyridocarbazole derivatives that hinder the ESE dependent splicing activity of HIV virus (Bakkour et al., 2007).

After further screening of a large database of chemical compounds, the same team found an indole derivative compound (IDC16) that interferes with the exonic splicing enhancer activity of ASF/SF2. In an *ex vivo* splicing assay, IDC16 inhibits the use of several weak 3'splice sites of HIV and less splicing products were produced when treated with IDC16. This, subsequently leads to suppression of the key vital proteins and compromises the assembly of infectious particles. Furthermore, IDC16 can efficiently block HIV-1 viral production in macrophage and T cell tropic laboratory strains, or several strains with high multi-resistance to the treatment of HAART. Compared with HAART that targets essential viral proteins, this is a totally different approach where cellular factors used by the virus are being targeted and has the great likelihood to overcome drug resistance.

Besides inhibitors that directly target SR proteins, Hagiwara's group has reported an SR protein phosphorylation inhibitor 340 (SRPIN340), which is an isonicotinamide compound, that preferentially inhibits SRPK1 and SRPK2 and down-regulates the expression of SRp75, which is essential for HIV expression. SRp75 downregulates

splicing at the 5′splice site of exon 3 of HIV-1 mRNA, thereby resulting in dramatically increased level of vpr mRNA (Fukuhara et al., 2006). In addition to its potential function in HIV-1 inhibition, SRPIN340 has also been shown to interfere with the propagation of Sindbis virus in a plaque assay.

These findings together suggest that targeting host cellular proteins like SRPKs and SR proteins to inhibit HIV-1 pre-mRNA splicing might provide alternative therapies to overcome drug resistance developed from genetic variation during virus reproduction.

1.9 Structure-based drug screening

Drug designing is the inventive approach for developing new medications according to the known biological target (Liljefors, et al., 2002). Due to the rapid need for new drugs to combat emerging diseases, more and more combinational techniques, including computer modeling techniques, are developed to speed up and lower the cost of the drug discovery process. Detailed structural information on protein-protein or protein-ligand interactions obtained through biophysical methods such as x-ray crystallography or NMR spectroscopy could provide structural basis to assist the drug discovery process. Based on the knowledge of the three-dimensional structure of the biomolecules, one can adopt the structure-based drug design approach to search for small molecules that target the potential biological receptors. This method has become one of the mainstreams in medicinal chemistry after rapid development of both

computation software and hardware in the past twenty years. Computer graphics, virtual screening, automated process of crystallography and other multi-facet techniques all play essential roles in the rapidly developing field of structure-based drug discovery.

Virtual, or *in silico*, drug screening has occupied the main position at the early stage for new drug discovery in recent years. Well performed virtual screening using drug-like compound libraries is said to achieve about 1%-10% active compounds in real biological assay, while the probability using traditional methods with random library is estimated to be about 0.01% (Doman et al., 2002; Woodward et al., 2006).

Structure-based virtual screening is a methodology that is widely used today (Kroemer, 2007). Structure-based screening predicts how the tested compound docks to the target protein 3D structure, by computing and evaluating the free binding energy between the protein and compound. In structure-based drug screening, small molecules from databases are virtually docked into selected site of the target protein through a docking program, which theoretically models the ligand–protein interaction to achieve an optimal complementarity of steric and physicochemical properties. A mathematical algorithm (also referred to as scoring function) will help to identify the utmost docking position on a targeted protein. Then compounds are ranked after the docking process and selected on the basis of some user-defined criteria, like suitable structure, charge, and binding affinity. Only a small part of top-ranked compounds will be chosen as candidates for later experimental assays to reduce the cost of the drug discovery process. Such structure-based virtual screening method has great success in the discovery of inhibitors for different targets, such as the SH2 domain of the proto-oncogene

tyrosine-protein kinase Src which is implicated in the treatment of osteoporosis and other bone-related diseases (Huang et al., 2004).

1.10 AutoDock Suite

One of the most adopted virtual screening software in the research community is AutoDock. AutoDock is a suite of molecular modeling simulation software developed by the Olson Laboratory at Scripps Research institute, La Jolla. As an effective automated docking tool for protein-ligand docking, AutoDock is designed to predict the critical interaction between the small molecules and biomacromolecules with known 3D structure. Until 2009, about 1200 publications have cited using the AutoDock method. So far, the discovery of several drugs is owed to AutoDock, including the most famous well-known HIV-1 integrase inhibitor Isentress.

Currently, AutoDock suite contains two generations of docking software: AutoDock Vina and AutoDock 4. In the docking program AutoDock Vina, the search and optimization algorithm being used is termed iterated local search global optimizer. It does not take into account the atomic charges of the targets being tested but instead depends only on the atom types that are defined using an internal atom typing scheme. The scoring function is based on a machine learning approach trained against both conformation and affinity of the bound structure-ligand complex. The combination of these two algorithm results in a significant gain in performance which, in large part, due to the use of a gradient optimization method that gives it a "sense of direction" from a

single evaluation.

The other program AutoDock 4 uses a genetic algorithm that is based on natural evolution. Genetic algorithms employ a population of individuals that undergo selection in the presence of variation-inducing operators and a fitness function is used to evaluate individuals, and reproductive success varies with fitness. The results from AutoDock 4 are given as estimates of free-energy change upon ligand binding and as a prediction of the ligand-receptor complex three-dimensional conformation.

Since AutoDock 4 and AutoDock Vina use different scoring functions, it is not completely appropriate to directly compare the accuracy of the results predicted by the programs. Either program may provide a better docking result for a particular protein target, despite AutoDock Vina performs two orders of magnitude faster than AutoDock 4. Therefore, the two programs are chosen to be used in tandem in my project to improve our chance of finding true hits.

1.11 Kinase-substrate interaction inhibitors

Protein kinases are highly regulated enzymes that modify their substrates by transferring the terminal phosphate group of ATP to them. The human genome comprises about 518 protein kinase genes, constituting about 2% of all human genes (Manning et al., 2002). Up to 30% of all human proteins undergo phosphorylation by specific kinases at some point in their lifetime, which in turn mediate the majority of cellular pathways like gene expression and cell growth.

All protein kinases share a conserved bi-lobed kinase core structure (Stout et al., 2004). Between the two lobes is a deep cleft where Mg^{2+} and ATP binds through the hydrogen bonds that formed between the adenine ring and the kinase 'hinge' — a segment that connects the amino- and carboxy-terminal kinase domains (Johnson et al., 1998).

In general, kinase inhibitors are commonly designed to block the phosphorylation activity of a specific kinase and in turn interfere with the signaling pathway(s) that it regulates. As a result of the great needs to explore new drugs to treat the kinase-related cancers or inflammatory diseases, an increasing number of kinase inhibitors have been discovered in the past decades. Most existing kinase inhibitors could be divided into three types, with most of them belong to the type-I ATP-competitive compounds (Zhang et al., 2009).

Type I kinase inhibitors are small molecules that bind to the ATP-binding cleft of kinases in their active conformations and comprise the bulk of ATP-competitive inhibitors (Liu and Gray, 2006). These inhibitors typically consist of heterocyclic ring systems that bind in and around the purine-binding sites, which are occupied by the adenine ring of ATP when the kinases are in their active conformations. Since SRPK1 adopts a constitutively active conformation at all time, SRPIN340 described above thus belongs to this type of inhibitors. The major advantage of type 1 inhibitors is that they are easy to characterize and be synthesized to mimic ATP. However, all Type I inhibitors bind to the conserved ATP binding sites of active kinases, so they suffer from widespread cross-reactivity among other kinases due to lack of specificity. The

nonspecific problem leads to serious side effect and poor safety in real medicinal treatment.

Type II kinase inhibitors bind to the same area occupied by the Type I compounds but recognize the binding pocket in special conformations that exist only when the kinases are inactive. In particular, when a kinase becomes inactivated, the DFG motif, which is responsible for anchoring the activation loop of a kinase in its active conformation when the kinase is activated, adopt an “out” conformation that exposes an additional hydrophobic binding site directly adjacent to the ATP binding site. Structural comparison of different inactive kinases has unveiled that the conformations of their corresponding hydrophobic sites are mostly unique. These hydrophobic sites thus provide additional structural scaffolds for inhibitor interaction and serve as new basis for the design of more specific inhibitors. Several successful Type II kinase inhibitors such as imatinib, nilotinib (Manley et al., 2005) and sorafenib, are highly specific for their targeted kinases and have been widely used for treatment of different diseases. Therefore, due to the significantly higher specificity, type II inhibitors usually have better safety profiles as well as exerting fewer side effects unlike the type I inhibitors. However, since detailed structural information in a kinase in its inactive form is required to design a type II inhibitor, the discovery and development of this class of inhibitors is more difficult and need much longer process.

Allosteric inhibitors make up the third type of kinase inhibitors, which binds an allosteric site outside the ATP-binding site (Ohren et al., 2004). Binding of an inhibitor at an allosteric site could lead to a change in the conformation of the kinase active site,

and consequently blocks kinase activity. When allosteric inhibition happens, the substrate can no longer fit the binding region owing to the changed shape of the kinase active site. Thus it is well accepted that allosteric inhibitors possess the highest specificity among the three classes of inhibitor that have been discussed so far，because each inhibitor would perform unique inhibition mechanism towards a particular kinase. CI-1040 is one of the well-known allosteric kinase inhibitor, which inhibits the activity of MeK1 and MeK2 by occupying a pocket adjacent to the ATP binding site. However, the development of type III kinase inhibitors suffer from a major downside, which is the kinase of interest must possess an allosteric site, and its detailed structural information on the site is necessary.

Due to various problems of kinase inhibitors described above, more and more novel inhibitor design methods are being explored. It is well known that the protein-protein interactions are the basic regulation of life, including signal transduction and large molecular machines. Improper protein-protein interactions could affect many cell processes including activation or inhibition of downstream steps of metabolic or signaling pathways. Therefore design of therapeutics inhibitors for specific protein-protein interactions is an attractive approach because these inhibitors could provide new means to control or modify the pathway of interest (Zhong et al., 2007). Several successes in the development of inhibitors targeting protein-protein interaction sites including those on BCL-2 (Oltersdorf et al., 2005), S100B (Fry and Vassilev, 2005), p56lck (Huang et al., 2004) and ERK2 (Hancock et al., 2005) have proven that it is indeed feasible to search for novel drugs against these interactions. One good

example that shows this approach could also be applied to kinase inhibitor development is provided by Shapiro group, who successfully used *in silico* method to identify a small molecule inhibitor of ERK2 by targeting a unique docking domain, which in turn disrupts interactions between ERK2 and its substrates (Hancock et al., 2005).

1.12 Focus of study

It has previously been shown that the interference of the interaction between SRPK1 docking groove and ASF/SF2 results in aberrant phosphorylation and subcelluar localization of ASF/SF2. Therefore, we propose to screen for inhibitors that bind to the docking groove of SRPK1, which might block the interaction between SRPK1 and ASF/SF2 and in turn interfere the phosphorylation event, resulting in the inhibition of ASF/SF2 activity during splicing of HIV-1 mRNA.

Our aim is to use the AutoDock suite to perform structure-based virtual screening method to identify potential inhibitors that block the interaction between SRPK1 and ASF/SF2. Also we wished to evaluate the effects of the selected compounds on the splicing of HIV-1 mRNA, and subsequently viral infection and replication.

Since such inhibitors do not inhibit the kinase activity of SRPK1, they should have minimal impact on its function during the regulation of splicing of endogenous genes and thus prevent deleterious side effects. Our work will provide the platform for future development of antiretroviral therapies that could overcome the multidrug-resistance of HIV-1.

Chapter Ⅱ: Materials and Methods

2.1 Materials

2.1.1 Bacterial strain

Escherichia coli (E. coli) strain DH5 α was used for plasmid DNA amplication. *E. coli* strain BL21(DE3) pLysS was used for the expression of protein.

2.1.2 Antibodies

Mouse myc-tag primary antibody (Cell Signaling) was used for detecting the localization of ASF/SF2 with a concentration of 1:1000 in immunocytochemistry study. AlexaFlour ® 488 anti-mouse secondary antibody (Invitrogen) in a dilution of 1:1000 was used for visualization.

2.1.3 Cell line

Human cervical adenocarcinoma epithelial cell line HeLa was used in this study. Cells were routinely cultured in Dulbecco's modified Eagle's medium (Invitrogen) supplemented with 10 % fetal bovine serum (Invitrogen).

2.1.4 Plasmid

SRPK1and SRPK1ΔNS3: Expression plasmids of His-tagged SRPK1 in pRSET vector, Sumo-His-SRPK1 in pET28a(+) vector and His-SRPK1ΔNS3 in pET15bvector was provided by Prof. C.K. Ngo of School of life sciences, Chinese University of Hong Kong.

Expression plasmid of SRPK in PCMV-tag2 vector was provided by Dr. Mars Tao of

School of life sciences, Chinese University of Hong Kong.

ASF/SF2: His-tagged ASF/SF2 in pET15b vector and GST-tagged ASF/SF2 in pGEX4T2 vectorwas provided by Prof. C.K. Ngo.

pCMV-myc-tagged ASF/SF2 was generated by PCR reaction using pET15b-ASF/SF2 as template,

The N-terminal primer containing an EcoR1 restriction site:

AAGAATTCGGATGTCGGGAGGTGGTGTG

The C-terminal primer containing an Xho1 restriction site:

GTTTCTCGAGTTATGTACGAGAGCGAG

After amplification, the PCR reaction products was first double digested with the restriction enzymes, gel purified and inserted into the mammalian expression vector PCMV-Myc, provided by Dr. Mars Tao of School of life sciences, Chinese University of Hong Kong.

pΔPSP: used for the transfection and *ex vivo* splicing assay is a gift from the University of Lowa.

The plasmid was a construct of HIV-1 plasmid deleted between nt 1511 and 4551. To

generate the pΔPSP, infectious HIV-1 plasmid pNL4-3 (GenBankTM accession number: M19921) was cleaved with *Spe*I and *Bal*I nucleases to generate an 11.9-kb fragment. Then the fragment was ligated together with oligonucleotides 5′-CTAGACGCGTTTGG-3′ and 5′-CCAAACGCGT-3′, which had been previously annealed, to form a double-stranded linker.

2.1.5 Reagents

Reagents are shown in Table 2.1.

2.2 Expression and purification of Recombinant protein

His-SRPK1ΔNS3: *E.coli*BL21(DE3) pLysS cells were transformed with pET15b-SRPK1ΔNS3. Large scale culture were grown in 4L of LB with 200 μg/ml ampicillin to O.D.600 0.6 at 37℃ and induced with 0.1 mM IPTG, the shaking overnight at 16℃. After centrifuged at 8000 rpm, the cell pellet were resuspended in 120 ml of lysis buffer containing 20 mM MES pH 6.5, 250 mM NaCl, 10% glycerol, 1 mM PMSF. The resuspended cells were lysed by sonication on ice. The solution was centrifuged and the supernatant were loaded onto a fast Q-sepharose column. The column was washed with 40 ml lysis buffer. The flow through sample was combined with the wash and the concentration of NaCl was brought up to 250 mM. The mixture was loaded on to a Ni^{2+} affinity column equilibrated with 40 ml of buffer containing 20 mM MES pH 6.5, 250 mM NaCl, 10% glycerol and 5 mM imidazole, followed by 40

Table 2. 1 List of solutions and commercial products used in this study

Solutions	Composition
Molecular cloning	
Luria-Bertani (LB) medium and broth	Commercial product from USB Affymetrix
TAE buffer	40 mM Tris-acetate (pH 8.0), 1 mM EDTA
Agarose gel loading dye	0.25 % Bromophenol blue, 0.25 % Xylene Cyanol FF, 25 % Ficoll 400, 0.1 % SDS
Miniprep kit	Commercial product from Intron
Gel extraction kit	Commercial product from Intron
Tissue culture	
Phosphate Buffer Saline (PBS)	Commercial product from Sigma-Aldrich
Dulbecco's modified Eagle's medium	Commercial product from Invitrogen
Fetal Bovine Serum	Commercial product from Invitrogen
Lipofectamine^{TM} 2000 transfection reagent	Commercial product from Invitrogen
Protein expression and analysis	
Acrylamide / Bis solution	40 % Acrylamide / Bis (37.5:1) from Bio-Rad
Stacking buffer	0.5 M Tris-HCl (pH 6.8)
Stacking arcylamide gel (18x16mm)	2.2ml distilled H_2O, 0.38 ml 40 % acrylamide / bisacrylamide (37.5:1), 0.38 ml Resolving buffer, 30 µl 10% SDS, 30 µl of 20% ammonium persulphate (APS, Sigma-Aldrich) and 3 µl

	N,N,N,N,-Tetramethylethylendiamin (TEMED, Invitrogen).
Resolving buffer	1.5 M Tris-HCl (pH 8.8)
12.5 % Resolving arcylamide gel (18x16mm)	2 ml distilled H_2O, 1.56 ml 40 % acrylamide / bisacrylamide (37.5:1), 1.3 ml Resolving buffer, 50 µl 10% SDS, 50 µl of 20% APS and 5 µl TEMED
SDS-PAGE buffer	25 mM Tris-HCl (pH 8.3), 250 mM Glycine, 0.1 % SDS
Transfer buffer	25 mM Tris-HCl, 192 mM Glycine, 20 % Methanol
1 X TBS	20 mM Tris, 137 mM NaCl, pH 7.6
1 X TBST	TBS (pH 7.6), 0.1 % (v / v) Tween 20
Blocking solution for western blot	5 % (w / v) Non-fat dried milk in TBST
Stripping buffer	62.5 mM Tris, 2 % SDS, 0.1 M β-mercaptoethanol, pH 6.7
Bio-Rad Protein Assay Reagent	Commercial product from Bio-Rad
BenchMark™ Prestained Protein Ladder	Commercial product from Invitrogen

Immunocytochemistry

4 % PFA	4 % paraformaldehyde in PBS
4, 6-diamidino-2-phenylindole (DAPI)	2 µg / ml DAPI in PBS
Fluorescent Mounting Medium	Commercial product from Dako Cytomation

RT-PCR

Trizol® Reagent	Commercial product from Invitrogen
MLV-RT	Commercial product from Genesys

ml of wash buffer with 40 mM imidazole. Then the protein was eluted with 30 ml of buffer containing 200 mM imidazole. The elution was dialysed overnight in 1L of buffer consisted of 20 mM MES pH 6.5, 250 mM NaCl, 10% glycerol to remove the imidazole. His-tag protein was then cleaved with thrombin and concentrated to less than 2 ml. Protein was loaded onto a Sueprdex 200 size exclusion column equilibrated with 20 mM MES pH 6.5, 250 mM NaCl, 10% glycerol and 1 mM DTT. The pooled peak fractions were concentrated to ~6 mg, aliquoted and stored in -80 ℃.

Sumo-His-SRPK1:*E.coli*BL21(DE3) pLysS cells were transformed with pET28a(+)-SRPK1. Large scale culture were grown in 4L of LB with 100 μg/ml kanamycin to O.D.600 0.6 at 37℃ and induced with 0.1 mM IPTG, the shaking overnight at 16℃. After centrifuged at 8000 rpm, the cell pellet were resuspended in 120 ml of lysis buffer containing 20 mM MES pH 6.5, 300 mM NaCl, 10% glycerol, 3 mM PMSF. The resuspended cells were lysed by sonication on ice. The solution was centrifuged and the supernatant were loaded onto a fast Q-sepharose column. The column was washed with 40 ml lysis buffer. The flow through sample was combined with the wash and the concentration of NaCl was brought up to 300 mM. The mixture was loaded on to a Ni2+ affinity column washed with 40 ml of buffer containing 20 mM MES pH 6.5, 300 mM NaCl, 10% glycerol and 5 mM imidazole, followed by 40 ml of wash buffer with 40 mM imidazole. Then the protein is eluted with 30 ml of buffer containing 200 mM imidazole. The elution was dialysed overnight in 1L of buffer

consisted of 20 mM MES pH 6.5, 300 mM NaCl, 10% glycerol to remove the imidazole. Sumo-His-tag protein was then cleaved with SENP1C and then loaded onto the GST-column combined with Ni^{2+} affinity column to remove the sumo-his tag. The column was washed with buffer containing 20 mM MES pH 6.5, 300 mM NaCl, 10% glycerol. The GST-column was moved away and the Ni^{2+} affinity column was eluted with buffer (20mM MES pH 6.5, 300 mM NaCl, 10% glycerol and 30 mM imidazole). The eluted protein was concentrated to less than 2 ml. Protein was loaded onto a Superdex 200 size exclusion column equilibrated with 20 mM MES pH 6.5, 250 mM NaCl, 10% glycerol and 1 mM DTT. The peak eluted fractions were combined, concentrated to ~4 mg, aliquoted and flash frozen, and stored in -80 ℃.

His-SRPK1:*E.coli*BL21(DE3) pLysS cells were transformed with in pRSET-SRPK1. Large scale culture were grown in 4L of LB with 100 μg/ml kanamycin to O.D.600 0.6 at 37℃ and induced with 0.1 mM IPTG, the shaking overnight at 16℃. After centrifuged at 8000 rpm, the cell pellet were resuspended in 120 ml of lysis buffer containing 20 mM MES pH 6.5, 300 mM NaCl, 10% glycerol, 3 mM PMSF. The resuspended cells were lysed by sonication on ice. The solution was centrifuged and the supernatant were loaded onto a fast Q-sepharose column. The column was washing with 40 ml lysis buffer. The flow through sample was combined with the wash and the concentration of NaCl was brought up to 300 mM. The mixture was loaded on to a Ni2+ affinity column washed with 40 ml of buffer containing 20 mM MES pH 6.5, 300 mM

NaCl, 10% glycerol and 5 mM imidazole, followed by 40 ml of wash buffer with 40 mM imidazole. Then the protein was eluted with 30 ml of buffer containing 200 mM imidazole. The elution was dialysed overnight in 1L of buffer consisted of 20 mM MES pH 6.5, 300 mM NaCl, 10% glycerol to remove the imidazole. The eluted protein was concentrated to less than 2 ml. Protein was loaded onto a Sueprdex 200 size exclusion column equilibrated with 20 mM MES pH 6.5, 250 mM NaCl, 10% glycerol and 1 mM DTT. The peak eluted fractions were combined, concentrated to ~4 mg, aliquoted and flash frozen, and stored in -80 ℃.

GST-ASF/SF2: *E.coli*BL21(DE3) pLysS cells were transformed with pGEX4T2-ASF/SF2. Large scale culture were grown in 2L of LB with 200 μg/ml ampicillin to O.D.600 0.6 at 37℃ and induced with 0.1 mM IPTG, the shaking overnight at 25℃. After centrifuged at 8000 rpm, the cell pellet were resuspended in 120 ml of lysis buffer containing 20 mM Tris pH 7.5, 50 mM NaCl, 10% glycerol, 1 mM PMSF and 1 mM DTT. The resuspended cells were lysed by sonication on ice and centrifuged to remove insoluble fraction. The supernatant were loaded onto a fast Q-sepharose column. The column was washed with 40 ml lysis buffer. The flow through and wash from the Q-sepharose column were combined and loaded onto the GST-column overnight at 4℃. GST-column was wash with 20 ml of buffer contain 20 mM Tris pH 7.5, 500 mM NaCl, 10% glycerol and 1 mM DTT and eluted with buffer containing 100 mM glutathione. The elution was dialysed overnight in 1L of buffer

consisted of 20 mM Tris pH 7.5, 500 mM NaCl, 10% glycerol and 1 mM DTT. Finally, the protein was concentrated to ~2 mg, aliquoted and flash frozen, and stored in -80 ℃.

His- ASF/SF2: *E.coli*BL21(DE3) pLysS cells were transformed with pET15b-ASF/SF2. Large scale culture were grown in 200 ml of LB with 200 μg/ml ampicillin to O.D.600 0.6 at 37℃ and induced with 0.1 mM IPTG, the shaking overnight at 25℃. Cells were then pelleted and lysed by sonication on ice using 20 ml of lysis buffer (0.1 m MOPS, 10 mm Tris-HCl, pH 8.0, 300 mm NaCl, 10% glycerol, 1 mm PMSF). The insoluble fraction was collected by centrifugation at 20,000 rpm for 30 min, washed twice with lysis buffer containing 2 M urea, and then resuspended in denaturing buffer (lysis buffer containing 8 m urea) overnight at room temperature. The soluble fraction, collected by centrifugation at 20,000 rpm for 30 min, was loaded onto a Ni2+ affinity column and washed with denaturing buffer containing 20 mm imidazole. The protein was re-folded by passing decreasing concentrations of urea through the column (8, 6, 4, 3, 2, 1, and 0 M urea). The re-folded protein was eluted with 20 ml of elution buffer (0.1 m MOPS, 10 mm Tris-HCl, 0.3 m NaCl, and 10% glycerol, pH 3.0). The pH of the eluted protein was changed to 7.0 by using NaOH. Protein was aliquoted and flash frozen, and stored in -80 ℃.

2.3 *In silico* screening of inhibitors

The program AutoDock Vina and AutoDock 4.2 were used to screen a virtual library of approximately 500,000 low molecular weight compounds collected from natural compound database and SPECS from ZINC. Database screening targeted a region of the docking groove on SRPK1 occupied by the RS-domain of ASF/SF2.

2.4 Kinase Glo Assay

Equilibrate the Kinase GLO assay buffer to room temperature before use. 5 nM SRPK1 wild type protein and 500 nM His-ASF/SF2 were incubated in the 384-well plates with different inhibitors or DMSO in 50 mM Tris-HCl pH 7.5, 10 mM $MgCl_2$ and 5 mg/ml BSA at 22 ℃. To all wells, the reaction was initiated by adding cold ATP and then reactions were quenched by addition of equal amount Kinase GLO assay buffer. The plate is shaked and incubated for 10 minutes at room temperature. Finally, luminescence was recorded.

2.5 *In vitro* kinase assay

Five nanomoles SRPK1 wild type protein and 500 nM His-ASF/SF2 were incubated in the presence of ^{32}P-ATP and different concentration of inhibitors or DMSO in 50 mM Tris-HCl pH 7.5, 10 mM $MgCl_2$ and 5 mg/ml BSA at 22 ℃. Then reactions were quenched by addition of 4X SDS buffer and subsequent boiling. Reaction mixes were

run on a 12.5% SDS gel. The gel was dried and exposed to autoradiography film. The bands' intensity were quantitated using a Phosphorlmager and the ImageQuant software (GE Healthcare). The result will be plotted against the inhibitor concentration added to the kinase reaction. The value of IC50 for each inhibitor was determined from these data by non-linear least squares regression analysis.

2. 6 Cell Culture

HeLa (ATCC cat. No. CRL-2) was routinely cultured in Dulbecco's modified Eagle's medium (Invitrogen) with 10% Fetal Bovine Serum (Invitrogen) at 37°C incubator supplied with 5 % CO_2. Cell passage was performed in 1 to 8 ratio by trypsination.

2. 7 MTT Assay

The MTT assay (Roche) was conducted according to manufacturers' instruction. Ten thousands Hela cells were cultured on each 96-well and was incubated for 6 hours. Drugs or DMSO were added directly to each well. Cells were treated with drugs for 24h in triplicate. Ten microliters of MTT Reagent 1 was added to each well and incubated for 4 hours until purple precipitate is visible at 37°C. A hundred microliters of detergent reagent was then added and left in 37°C incubator overnight. Finally, the absorbances were measured by spectrometer (Molecular Devices) at 580 nm, with 680 nm as reference.

2. 8 Immunocytochemistry

HeLa cells were cultured on glass coverslips 12 hours before transfection. Drugs were treated 4h after tansfection. The coverslips with cells were collected 4 hours after drug-treatment and were fixed with 4 % paraformaldehyde (PFA) in PBS for 20 minutes. The coverslips were then washed with PBS for 3 times and blocked with 10 % normal donkey serum (Sigma-Aldrich) at room temperature for 1 hour. The cells on the coverslips were then incubated with primary antibody in PBS with 5 % donkey serum and 0.3% Triton-X at 4 °C overnight. All the washings were performed by PBS supplemented with 0.3% Triton-X later on. After washing for 3 times, cells were then incubated with corresponding AlexaFluor® secondary antibodies for one hour at room temperature. Finally cells were washed three times, counterstained with 2 μg / ml 4, 6-diamidino-2-phenylindole (Roche) for 5 minutes, washed again with PBS, mounted on glass slides using fluorescent-mounting medium (Dako Cytochromation) and sealed with nail enamel. Immunoflorescence were visualized under epifluorescence microscope (Olympus) and images for the same antigen were captured using the same laser settings.

2.9 *Ex vivo* splicing assay

Hela cells or HEK293 cells were grown in DMEM, supplemented with 10% FBS on

6-well plate. After 24 h, each well was co-transfected with splicing reporter constructs pΔPSP and EGFP plasmid with Lipofectamine (Invitrogen) according to the manufacturer's instructions. Four hours later, the chemicals to be tested were added in culture media. The chemicals were added from DMSO stocks (final DMSO concentration in the media 1%).

Total cellular RNA was isolated from transfected HeLa cells 48 h post-transfection using Trizol Reagent (Invitrogen), and 3 μg of RNA was reversed transcribed with random hexamers using MLV-RT (Genesys Ltd) in standard 37℃ protocol. PCR amplified with forward primer BSS (5′-GGCTTGCTGAAGCGCGCACGGCAAGAGG-3′; nt 700–727) and reverse primer SJ4.7A, which spans sites D4 and A7 (5′-TTGGGAGGTGGGTTGCTTTGATAGAG-3′; nt 8369–8381 and 6032–6044), following an incubation of 1.5 min at 95℃, and 24 cycles of 1 min at 95℃, 30 s at 59℃, and 30 s at 72℃, the reaction was ended with an extension step of 10 min at 72℃. To normalize the signals, GAPDH was used as an internal control of the PCR reactions as described. Amplified spliced products are confirmed by 10% polyacrylamide gel electrophoresis (PAGE).

2.10 Surface plasmon resonance spectroscope

GST antibody was immobilized on all flow cells of the Biacore CM5 sensor chip. Recombinant GST-ASF/SF2 protein was immobilized on a Biacore CM5 Chip with GST antibody. The final response obtained was 700 RU (response units) (1 response

unit corresponds to 1 pg of protein per mm^2) of immobilized GST-ASF/SF2. Purified SRPK1 was maintained at 50 nM in running buffer (10 mM HEPES, 0.15 M NaCl, pH 7.4). SRPK1 was preincubated with the test compounds at varying concentrations, and interaction assays was studied by a Biacore 3000 instrument.

Chapter Ⅲ: Results

3.1 *In silico* screening of inhibitors

The structure of SRPK1 (PDB ID: 1WAK) is selected for the docking experiment and 351,473 compounds from the natural products database of ZINC (http://zinc.docking.org/) and the commercially available SPECS synthetic compound library were used for screening. Natural products offer diverse chemical and structural complexity, and unmatched biological potency. It is not unusual that a natural product 'hit' from screening can become a drug directly. Therefore, we performed *in silico* screening using a natural compound database which contains nearly 90,000 compounds that are either pure natural products or simple chemical derivatives of natural products. Compounds included in the database are commercially available from 7 vendors, so compounds identified as potential hits could be easily purchased for further analysis. In addition, we have also chosen the database of SPECS compounds that are highly affordable and exhibit structural characteristics of biologically active compounds, and meet Absorption Distribution Metabolism Excretion (ADME) requirements.

Structural analyses of the SRPK1-peptide and the SRPK1-ASF/SF2 complexes both indicate that the docking groove of SRPK1 interacts with the binding partners with near perfect shape and charge complementarily. Furthermore, previous studies have confirmed that 4 residues (D548, D564, E571, and K615) within the docking groove are critical for the stability of the complexes as well as the mechanism of phosphorylation of ASF/SF2. Therefore, the docking groove was selected for *in silico* screening to identify low-molecular-weight compounds with the potential to bind to this site and subsequently block binding and phosphorylation of ASF/SF2. Inhibition of ASF/SF2

phosphorylation might in turn abrogate its function during HIV-1 splicing and prevent HIV-1 virus from completing its life cycle.

To prepare for the screening, the SRPK1 structure (PDB ID: 1WAK) was edited using the software provided by the ADT package of the AutoDock Suite to fix problems like missing atoms, added waters, and alternate locations. Polar hydrogen atoms were also added to the macromolecule and Gasteiger charges were calculated for all atom types present in the screening library. Three-dimensional affinity grid (Figure 3.1) that is used to specify the site of interest for inhibitor screening was positioned around the docking groove and calculated for all atom types present in the screening library. This defined potential grid was then used to dock all compounds in the selected libraries.

Next, the small molecule compound candidates were first filtered using criteria set out by the Lipinski's Rule of Five: MW<500, hydrophobicity (Log P) between -4 and 4, number of hydrogen bond donor <5, and acceptor <10. This selection ensures the compounds identified by the docking programs as hits would be suitable for used in experimental testing and have the potential to be developed into drugs (Figure 3.2). Compounds that did not fit any of the criteria were discarded and those that passed the filtering were converted to AutoDock format for the docking experiment.

To improve the confidence of the docking results in our experiment, we performed two rounds of docking using two different programs AutoDock Vina and AutoDock 4.2. First, we took advantage of the high docking efficiency of AutoDock Vina to screen 351,473 compounds in our selected natural products and SPECS libraries. Then,

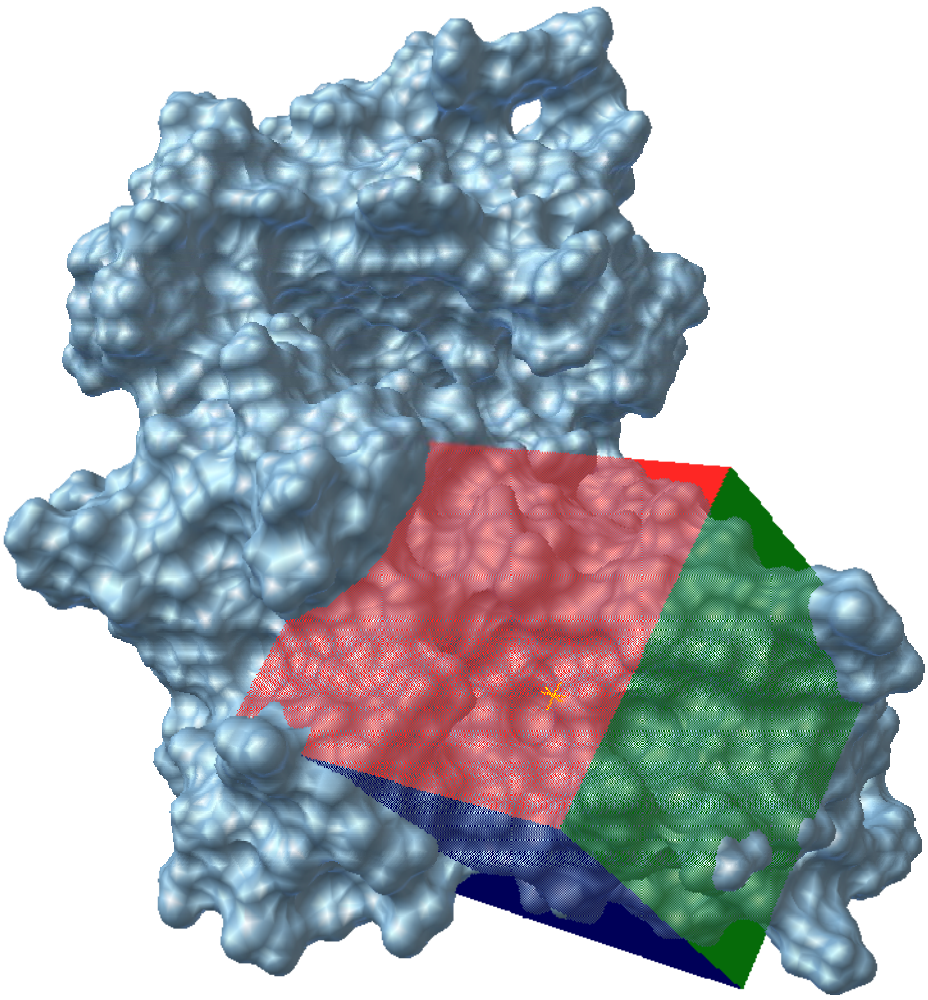
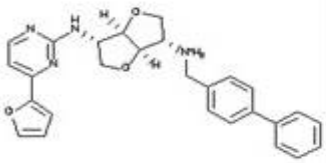


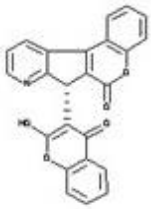
Figure 3. 1 Docking grid set on the docking groove of SRPK1 for *in silico* screening. The docking groove of SRPK1, which is necessary for the substrate binding, is located at the large lobe and is covered in the drawn docking grid. The area covered by the docking grid was used by the software for guiding the small molecules to bind at the selected site.

ZINC05433939



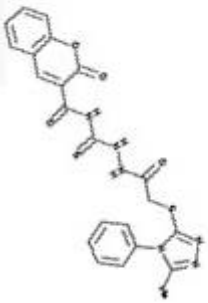
Molecular Weight	455.538 [g/mol]
XLogP3-AA	3.60
H-Bond Donor	3
H-Bond Acceptor	7

ZINC03999624



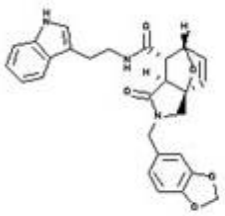
Molecular Weight	395.37 [g/mol]
XLogP3-AA	3.94
H-Bond Donor	1
H-Bond Acceptor	6

ZINC03669158



Molecular Weight	494.558 [g/mol]
XLogP3-AA	1.83
H-Bond Donor	3
H-Bond Acceptor	10

ZINC08878546



Molecular Weight	471.513 [g/mol]
XLogP3-AA	2.89
H-Bond Donor	2
H-Bond Acceptor	8

Figure 3. 2 Examples of small molecules from the natural products and SPECS libraries. The libraries were filtered by the Lipinski's Rule of Five criteria - MW<500, hydrophobicity (Log P) between -4 and 4.2, number of hydrogen bond doner <5, and acceptor <10, before being submitted to the docking program.

potential hits obtained from AutoDock Vina were subjected to a second round of docking in AutoDock 4.2, and only those compounds that are also identified by AutoDock 4.2 as hits were further analyzed and selected for experimental testing. Since the algorithms used for the docking and scoring are completely different in the two programs, this two-steps docking procedure would serve as a non-bias double-checking system for potential hits being identified. Compounds that survived both screenings should have higher potential to be true binders to SRPK1 when compare to using a single program alone.

During the first round of screening, compounds were selected if they met the following requirements: 1) the binding free energy is among the lowest 1% among all screened compounds; 2) must be lower than -8 kcal/mol; 3) the RMSDs between different docking models of the same compound are lower than 4 Å. These requirements were set to ensure that the selected compounds bind to the binding pocket stably and tightly.

In the second round of docking, 2846 compounds selected from the first screening by AutoDock Vina were inputted into the AutoDock 4.2 program. Lamarckian genetic algorithm (LGA) was employed to search for ligand conformations using docking parameters as follow: population size of 120 individuals, maximum numbers of generations and energy evaluations of 27000 and 1.7 million respectively, 50 docking runs, and random initial positions and conformations AutoDock 4.2 has a free-energy scoring function that uses the AMBER force field to estimate the free energy of binding of a ligand to its target. The binding free energies of different runs of each compound

were examined by clustering with an RMSD tolerance of 0.7Å.The docking results of target compounds were ranked in a list on the basis of the mean estimated binding free energies of their most populated clusters. The compounds with the lowest binding free energies were ranked at the top of the list. Since there was no known inhibitor for the target site can be used for control docking, no free energy cutoff was present and we simply selected the top 1-2% of ranked compounds after each screening round. The top ranking ligands were then selected for inspection visually and evaluated based on their interactions with the kinase docking groove. The whole screening procedures are showed in Figure 3.3 and the docking results of the selected compounds are listed in Table 3.1.

Nearly all screened inhibitors were predicted to bind in the vicinity of the targeted binding site and adopted a variety of orientations within the SRPK1 docking groove. After inspecting their interactions within the docking groove, 42 compounds including 8 from the Natural Product library and 34 from the SPECS synthetic library that interact with the kinase reasonably and favorably were selected for further biochemical analysis (compounds selected from SPECS library will be labeled with a numerical symbol based on the order they were identified; compounds identified from the Natural Product database will be designated starting with the letter N, followed by a numerical symbol, e.g. N-1, N-2, etc. from now on). Docking results of the selected compounds are listed in Table 3.1 and figure 3.4 illustrate examples of selected compounds docked to the SRPK1 docking groove.

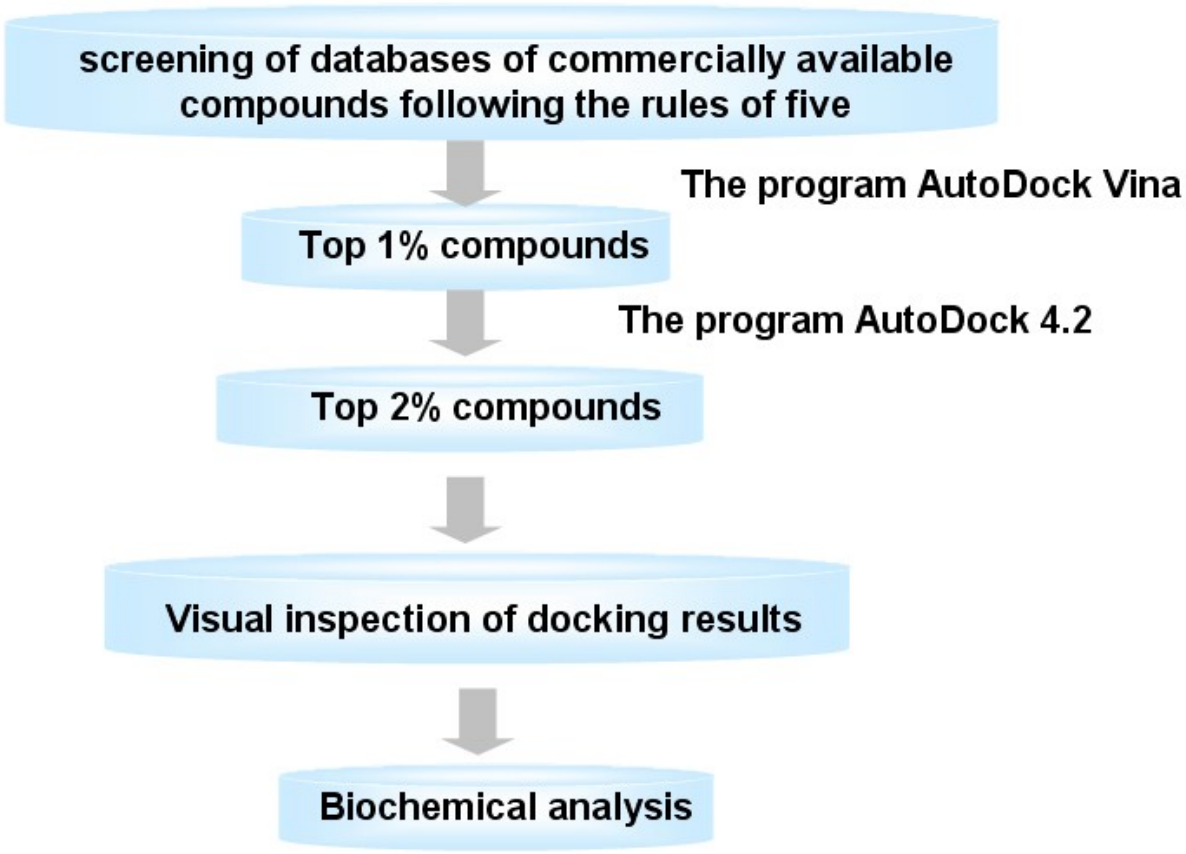


Figure 3. 3 Flow chart of the *in silico* drug screening. Small compound candidates were first filtered by the Lipinski's Rule of Five criteria, following by the two rounds of docking using the two docking programs AutoDock Vina and AutoDock 4.2.

Table 3.1: Inhibitors identified by AutoDock Vina & AutoDock 4.2

Natural Compound								
	Estimated Free Energy of Binding (kcal/mol)			Estimated Free Energy of Binding (kcal/mol)			Estimated Free Energy of Binding (kcal/mol)	
Compound ID	Autodock Vina	Autodock 4.2	Compound ID	Autodock Vina	Autodock 4.2	Compound ID	Autodock Vina	Autodock 4.2
N1	-10.2	-9.62	N2	-9.9	-9.39	N3	-10.3	-8.55
N4	-10.4	-8.19	N5	-10.0	-8.19	N6	-10.0	-8.8
N7	-10.0	-9.32	N8	-10.1	-9.32			
SPECS								
1	-9.0	-8.46	2	-9.0	-8.4500	3	-9.0	-8.14
5	-9.1	-10.2200	6	-10.9	-8.6400	7	-10.9	-8.6400
8	-9.1	-8.2300	9	-9.1	-10.1100	13	-9.0	-8.17
14	-9.0	-8.2700	16	-9.2	-8.0600	17	-9.4	-8.42
18	-9.3	-8.65	19	-9.3	-9.2200	20	-9.3	-9.3600
21	-9.9	-8.6700	23	-9.1	-9.2	24	-9.9	-8.2700
25	-9.5	-8.0700	27	-9.1	-8.38	28	-9.3	-8.85
29	-9.1	-8.5	30	-9.6	-8.7	31	-9.3	-9.21
33	-9.1	-9.35	34	-10.1	-8.56	35	-10.1	-8.39
36	-10.1	-8.3900	37	-10.0	-9.2000	38	-9.4	-8.4600
39	-9.9	-8.5	40	-9.4	-8.76	41	-10.1	-8.64
42	-9.0	-8.73						

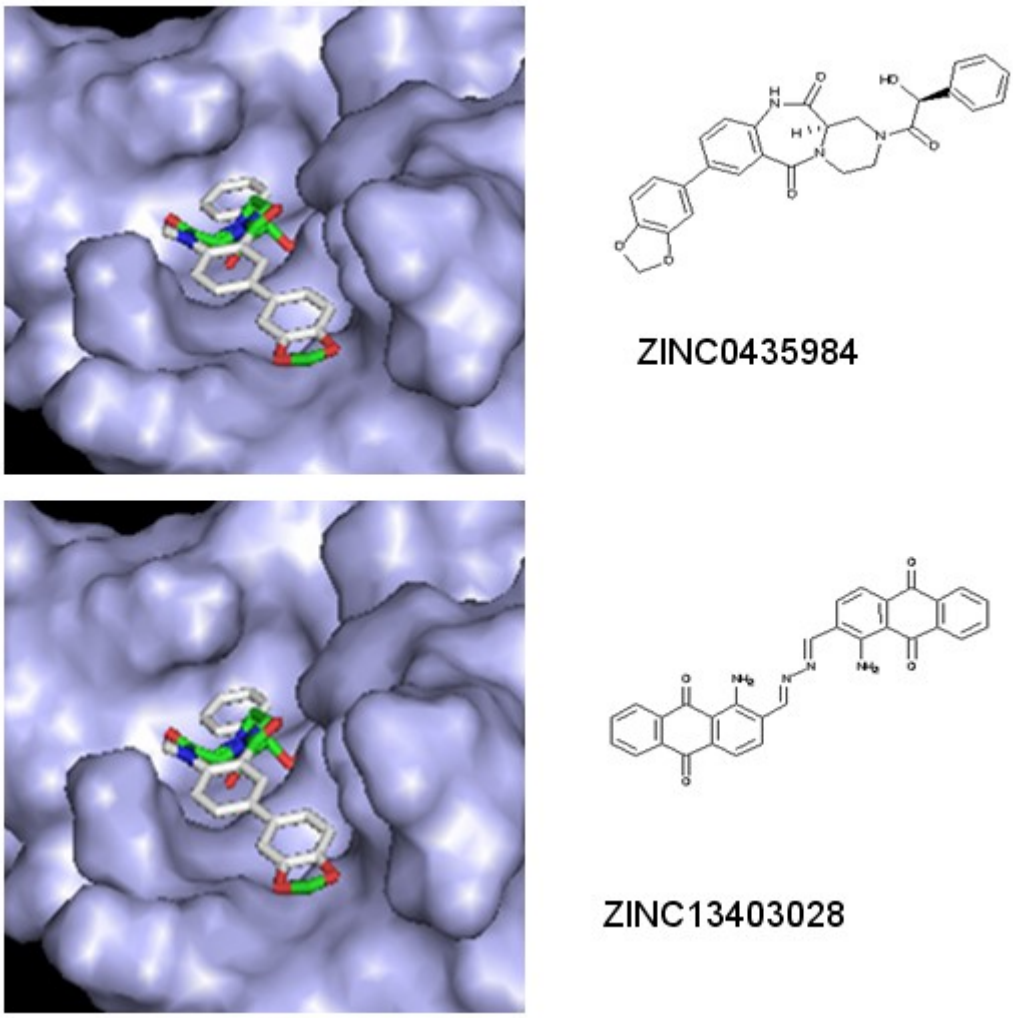


Figure 3. 4 Examples of predicted binding of active compounds to SRPK1. The binding modes of two compounds at the docking groove of SRPK1 are shown. SRPK1structure is shown in surface rendition and colored. The chemical structure of the compounds being docked are shown on the right.

3.2 Selected Compounds Inhibits SRPK1 *in Vitro*

3.2.1. Protein purification

A. SRPK1 purification

We first expressed and purified both recombinant SRPK1 and ASF/SF2 so we could perform *in vitro* assays to test the effects of different inhibitors identified in our *in silico* screening. For the expression of SRPK1, I have subcloned it into the SUMO-pET28a(+) expressing vector. This vector was provided by Prof. Shannon Au's group at the School of Life Sciences at CUHK. It was modified from pET28a vector by inserting the sequence of SUMO protein between the poly-histidine tag and the coding region. Sumo-pET28a(+)-SRPK1 were transformed into BL21(DE3) plysS cells and used to express recombinant protein under the induction of IPTG.

Protein sample was loaded on an anion exchange chromatographic column (Q-column) that is connected directly to a cation exchange chromatographic column (S-column) at a NaCl concentration of 50 mM and pH 6.5.After washing, the proteins remained on the S column was eluted using a concentration gradient of NaCl. SRPK1 was eluted at around 400 mM salt concentration (Figure 3.5).

After the first step of purification by tandem anion-cation exchange chromatography, SRPK1 was then applied to Ni^{2+}-NTA column (Figure 3.5) in the presence of 5 mM imidazole, which was expected to effectively block most non-specific binding. The loaded column was then washed by buffer containing 50 mM imidazole and the protein was finally eluted with eluting buffer that contains 200 mM imidazole.

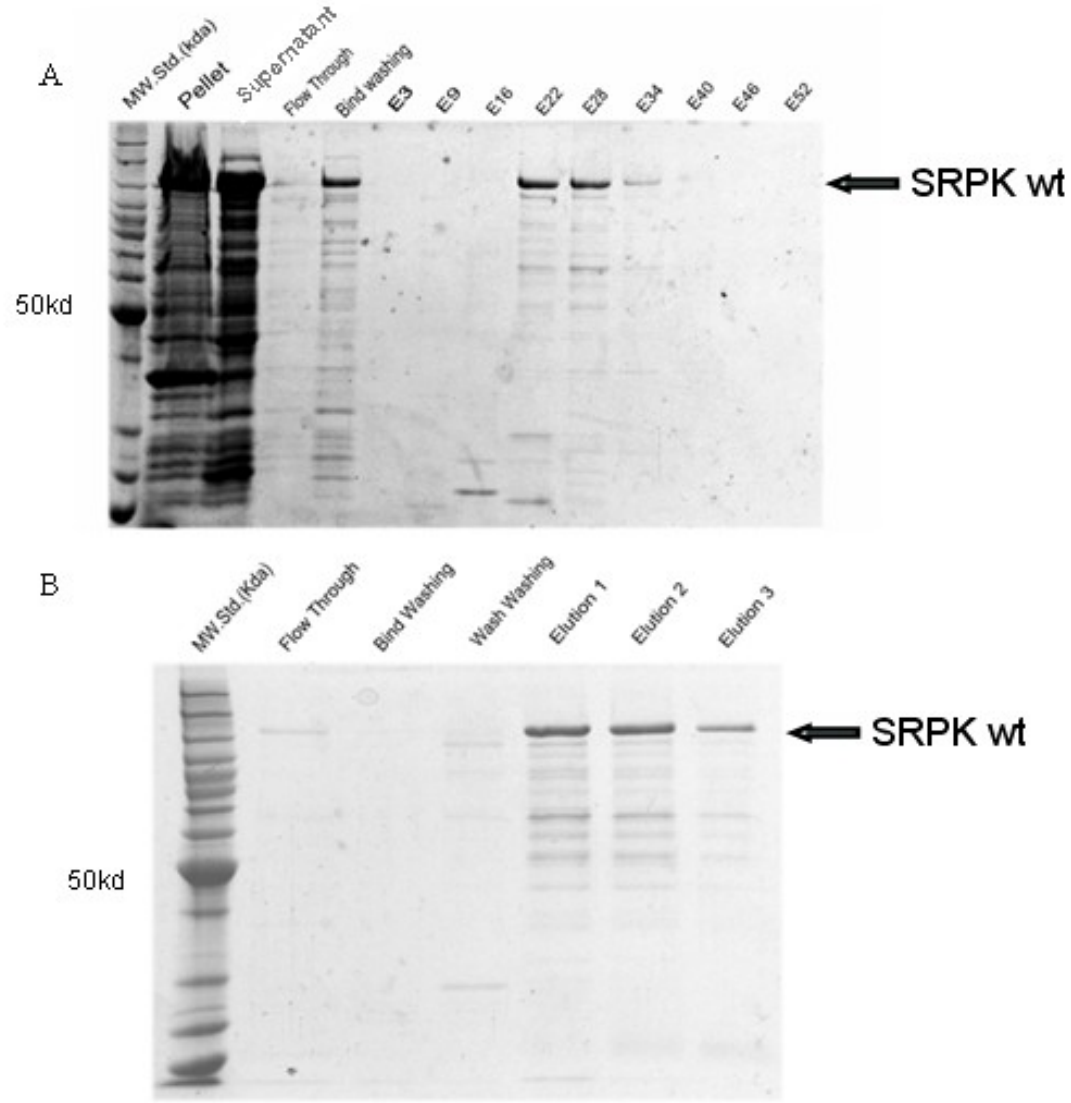


Figure 3. 5 Q & S ion-exchange chromatography and Ni^{2+}-NTA affinity chromatography purification of wild type SRPK1. A) 12.5% SDS gel analysis of SRPK1 wild type purified by Q & S columns. The flow-through fraction from the Q-sepharose column was loaded directly onto the S-sepharose column followed by the gradient elution of NaCl. The gel showed that elution 22-34 contained SRPK1 wt. B) 12.5% SDS gel analysis of SRPK1 wild type purified by Ni^{2+}-NTA Column. Loaded sample was collected from the elution fractions 20 to 37 from Q & S column purification. SRPK1 was eluted efficiently at 200 mM imidazole buffer.

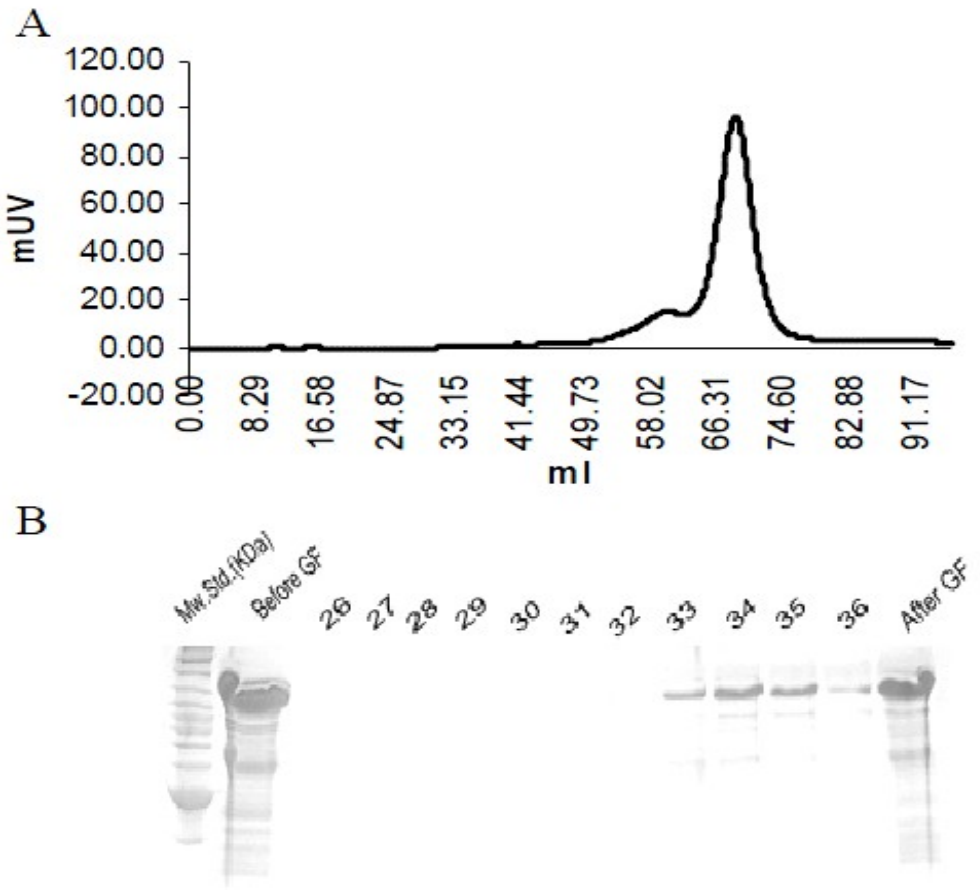


Figure 3. 6 Gel filtration chromatography purification for SRPK1. A) UV curve of elution profile. B) SDS-PAGE of the eluted fractions after gel filtration chromatography. Samples 33 to 36 were collected for concentration.

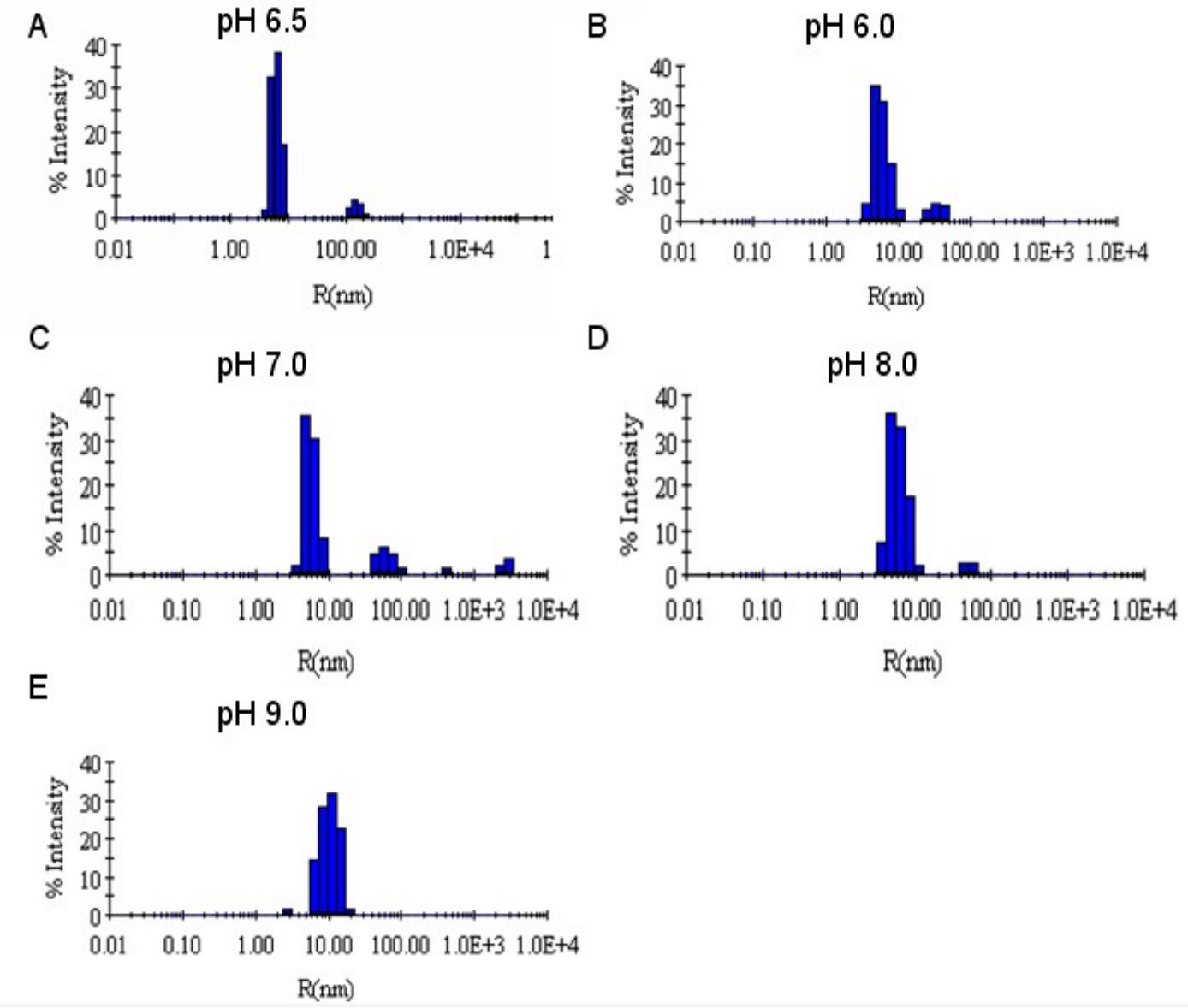


Figure 3. 7 SRPK1 wild type is monodisperse at pH 6.5. Dynamic light scattering (DLS) experiment was performed to analyze the protein. Protein was dialysed against different buffers with different pH. A) MES pH 6.5, B) MES pH 6.0, C) Tris pH 7.0 D) Tris pH 8.0 E) CAPSO pH 9.0. The results from DLS experiment suggested that MES pH 6.5 is the optimal buffer to be used during the purification process.

SRPK1 was eluted in elution fractions 1-3. SDS-PAGE analysis of the elution from Ni^{2+}-NTA column, however, showed that the purity of SRPK1wt protein has improved only slightly after this step. Next, I cleaved off the His-SUMO tag using purified SUMO protease SENP1C and reloaded the cleavage product on the Ni^{2+} column again to remove the cleaved tag. Finally, gel filtration was performed to further purified the protein. Proteins eluted from the gel filtration column were analyzed by SDS-PAGE again and flash frozen for storage. Figure 3.6 shows the final level of purity of SRPK1 wt.

A dynamic light scattering (DLS) experiment was performed to analyze if the purified protein is homogeneous by dialyzing it against buffers with different pH. Compared with the other 4 conditions, the width of the peak in condition containing MES pH 6.5 represented less broadness of the size distribution and indicated more monomodal distribution of protein particle sizes, suggesting the protein is monodisperse at this pH. These results confirmed that MES pH 6.5 is the optimal buffer to be used during the purification process (Figure 3.7).

B. ASF/SF2 purification

ASF/SF2 was subcloned into the pET15b vector with poly-histidine. *E.coli* BL21(DE3) pLysS cells were transformed with pET15b-ASF/SF2 and induced with IPTG. Cells were then pelleted and lysed by sonication and the soluble and insoluble fractions were separated by centrifugation. Recombinant His-ASF/SF2 expressed in *E. coli* as inclusion bodies. Therefore, we purified the protein under denaturing condition

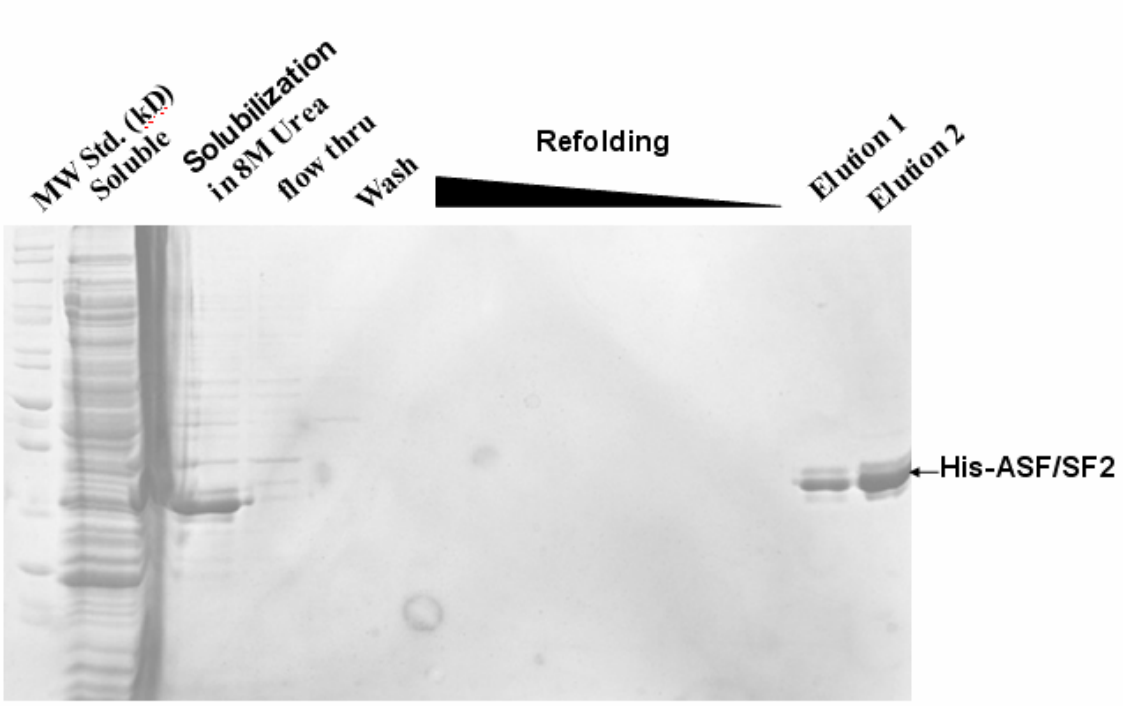


Figure 3. 8 Refolding and purification of $(His)_6$-tagged ASF/SF2 wild type. Denatured soluble fraction was load onto a Ni^{2+} affinity column and washed with denaturing buffer. The protein was refolded on column by passing decreasing concentrations of urea. The refolded protein was eluted with low pH elution buffer.

and performed refolding to obtain soluble His-ASF/SF2 protein. First, the collected insoluble fraction was first denatured using urea. Next, the denatured sample was loaded onto a Ni^{2+} affinity column and washed under denaturing condition. The protein was then refolded by passing decreasing concentration of urea through the column. The protein refolded on column was eluted using buffer at pH 2. The pH of the eluted protein was later adjusted to 7.0 by adding NaOH slowly (Figure 3.8).

3.2.2 Inhibits ASF/SF2 Phosphorylation by SRPK

A. Screening using Kinase-Glo assay

To determine whether the 34 SPECS compounds selected by AutoDock Vina and AutoDock 4.2 could inhibit the phosphorylation of ASF/SF2, we first performed the kinase-glo assay for a fast screening. The assay was performed in a 96-well white plate. Each well contains His-ASF/SF2, SRPK1wt, ATP, $MgCl_2$ and different concentrations of the selected compounds. The reactions were allowed to proceed for 5 mins and quenched by adding one volume of Kinase-Glo Reagent to the completed kinase reaction. ATP remaining at the time that the reagent was added was used as a substrate by the Ultra-Glo™ Luciferase to catalyze the mono-oxygenation of luciferin. Luminescence generated by luciferase reaction was measured using a luminometer. By comparing the luminescence of the reactions carried out in the presence of compounds and the vehicle control (dimethyl sulfoxide [DMSO]), the compound activity on SRPK1 could be determined.

Despite the candidate compounds were filtered using the criteria set in the Linpinski's Rule of Five, three compounds, N1, N3 and N6 from natural products library, and 5 compounds including compound 1, 3, 13, 17 and 19 from SPECS library remained insoluble in the presence of high concentration of DMSO.

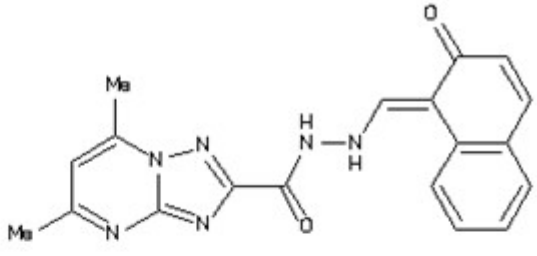
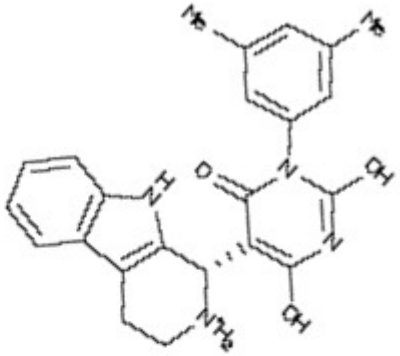
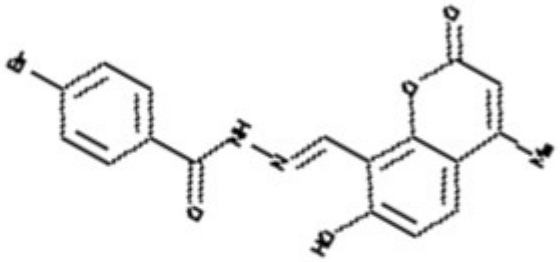
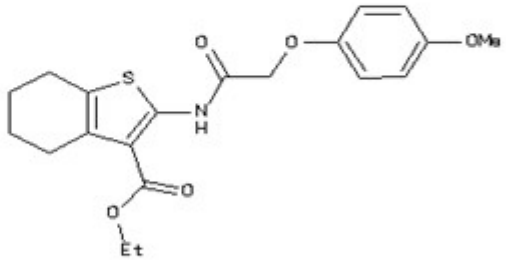
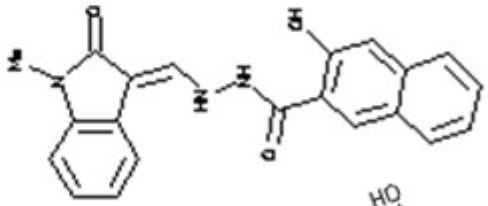
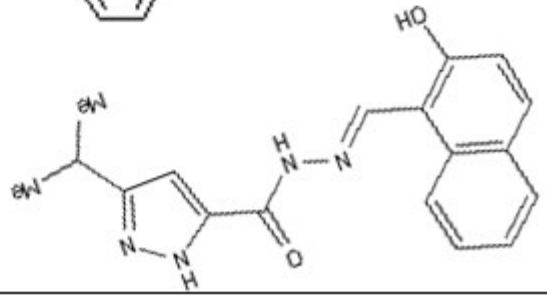
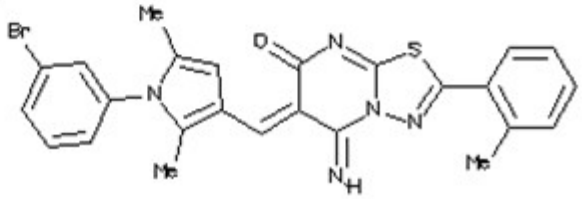
However, 34 compounds selected from screening were finally tested in the Kinase-Glo assay to determine their inhibitory activities on the phosphorylation of ASF/SF2 by SRPK1. 16 out of the 34 soluble compounds selected, including five compounds from natural product library and compound 23, 24, 25 and etc, did not produce consistent results using this method. Therefore, their activities were reassessed using radioactive phosphorylation assay. All compounds were tested two times using their highest soluble concentration, ranging from 1000μM to 25μM. Since some compounds produced inconsistent results in different trials of Kinase-Glo assays, I double checked the screening results by performing compounds that could not be determined by the Kinase-Glo assay.

Four compounds out of 34 soluble compounds tested, including N7, compound 5, 24 and 36 consistently attenuated SRPK1 phosphorylation activity compared with vehicle control (DMSO)(Figures 4.5). Structures and properties of the selected compounds are shown in the Table 3.2 and 3.3. Furthermore, since compound 5 showed strong inhibition on the phosphorylation of ASF/SF2 by SRPK1, I further identified three compounds that are structurally similar to the active compound 5, namely compound D1, D2 and D3, from the PubChem. They were purchased and subjected to the Kinase-Glo assay as well.

Table 3. 1 List of selected inhibitors after biochemical screening

N7	ZINC09330756 1-(3,5-dimethylphenyl)-6-hydroxy-5-[(1S)-2,3,4,9-tetrahydro-1H-pyrido[3,4-b]indol-2-ium-1-yl]pyrimidine-2,4-dione Molecular Formula: $C_{23}H_{23}N_4O_3$+
5	ZINC04469951 N'-[(2-hydroxy-1-naphthyl)methylene]-5,7-dimethyl[1,2,4]triazolo[1,5-a]pyrimidine-2-carbohydrazide Molecular Formula: $C_{19}H_{16}N_6O_2$
D1	ZINC00722056 4-bromo-N'-[(7-hydroxy-4-methyl-2-oxo-2H-chromen-8-yl)methylene]benzohydraz ide Molecular Formula: $C_{18}H_{13}BrN_2O_4$
D2	ZINC15885168 3-hydroxy-N'-[(1-methyl-2-oxo-1,2-dihydro-3H-indol-3-ylidene)methyl]-2-naph thohydrazide Molecular Formula: $C_{21}H_{17}N_3O_3$
D3	ZINC00351555 N'-[(2-hydroxy-1-naphthyl)methylene]-3-isopropyl-1H-pyrazole-5-carbohydrazide Molecular Formula: $C_{18}H_{18}N_4O_2$
24	ZINC00704675 ethyl 2-{[(4-methoxyphenoxy)acetyl]amino}-4,5,6,7-tetrahydro-1-benzothiophene-3-carboxylate Molecular Formula: $C_{20}H_{23}NO_5S$
36	ZINC08450637 6-{[1-(3-bromophenyl)-2,5-dimethyl-1H-pyrrol-3-yl]methylene}-5-imino-2-(2-methylphenyl)-5,6-dihydro-7H-[1,3,4]thiadiazolo[3,2-a]pyrimidin-7-one Molecular Formula: $C_{25}H_{20}BrN_5OS$

Table 3.3 Structures of selected inhibitors

Compound 5		Nature Compound N7	
Compound D1		Compound 24	
Compound D2			
Compound D3		Compound 36	

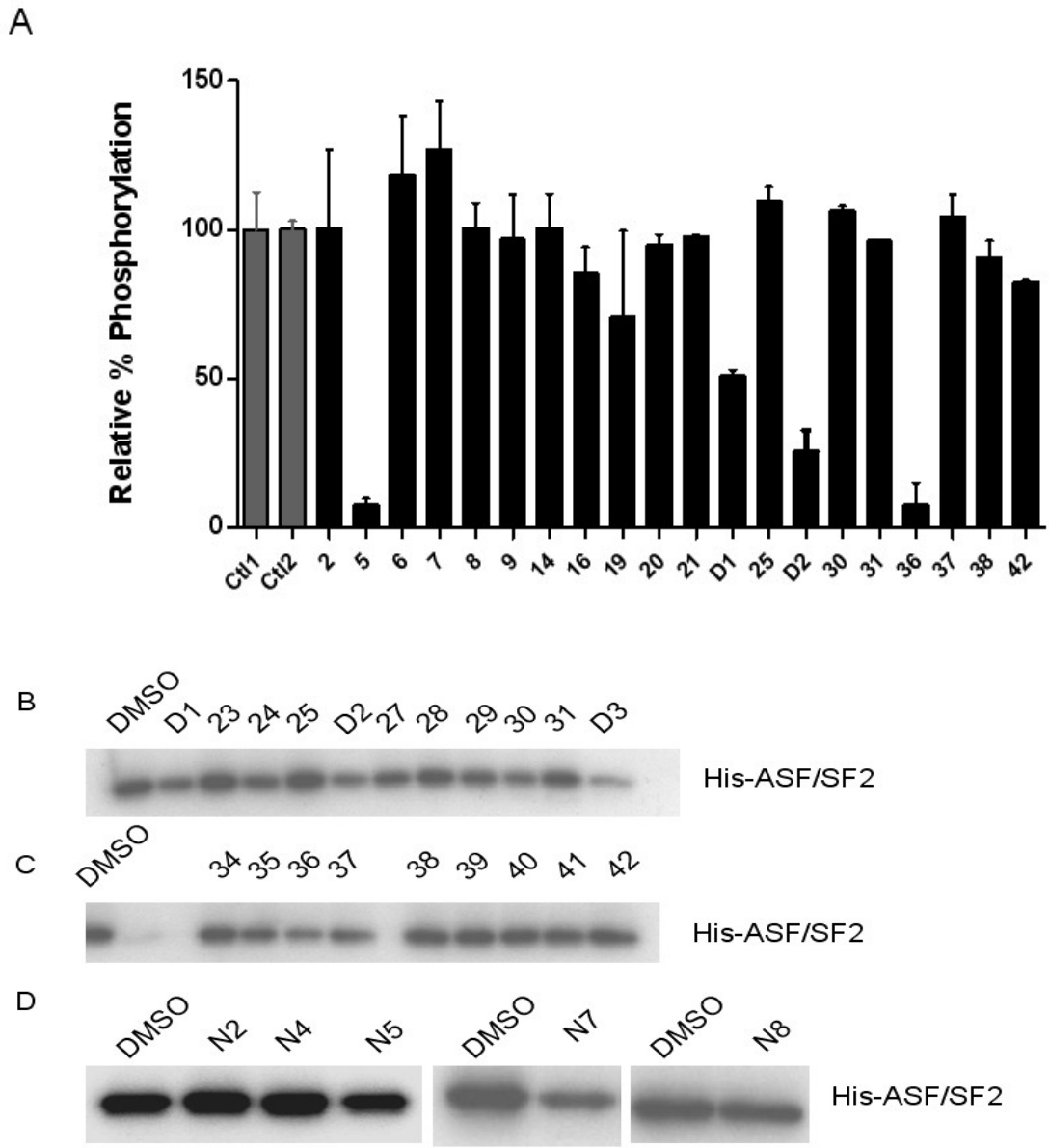


Figure 3. 9 *In vitro* kinase assays were performed to test the effect of compounds on the kinase activity of SRPK1.A) *In vitro* kinase Glo assays. The y axis shows the relative percent of phosphorylation of His-ASF/SF2 in the presence of selected compounds (black bars). Compounds tested but did not produce consistent data are excluded from this chart. B), C) and D): *In vitro* radioactive kinase assays to reassess some results from the Kinase-Glo assays as well as testing the inhibitory activities of natural compounds. Compound 5, D1, D2, D3, 24, 36 and N7 showed inhibition effect on the phosphorylation of His-ASF/SF2 by SRPK1.

While these structurally related compounds could also inhibit the phosphorylation of ASF/SF2 to some extent, compound 5 remained as the most potent inhibitor on the phosphorylation of ASF/SF2 (Figure 3.9).

Some compounds shown lower inhibition even with the highest concentration were also been weeded out, like compound 24. All the left 6 compounds showed phosphorylation repression by greater than 50%, 6 from the SPECS and one from the Natural Compound database were all selected for next step study to get the IC50.

B. *In vitro* radioactive Kinase assay to determine IC50

Since only 5 out of the 8 compounds selected from the Natural Product library could be solubilized in 5% DMSO and subjected to kinase activity assays, they were directly tested using *in vitro* radioactive kinase assays instead of the quick Kinase-Glo screening method (Figure 3.9D). The reaction was initiated with 50 μM ATP and trace amounts of γ^{32}P-ATP. The phosphorylation intensity of ASF/SF2 was quantitated using a phosphorImager and ImageQuant software (GE Healthcare). Only compounds N5 and N7 were shown to inhibit phosphorylation of ASF/SF2, with N7 exerted stronger inhibitory effect.

To evaluate the IC50s of the compounds that showed inhibition activities, the phosphorylation content of ASF/SF2 were measured in the presence of different concentrations of inhibitors. The phosphorylation content of ASF/SF2 relative to the untreated control was plotted against the inhibitor concentrations being used in the

reaction. Six compounds were shown to inhibit ASF/SF2 phosphorylation by SRPK1 in a dose-dependent manner. (Figure 3.10 and Figure 3.11). In particular, compounds 5, D2 and N7 blocked more than 80% of phosphorylation of ASF/SF2 at the highest concentration being tested. From the *in vitro* kinase activity assay results, compound 5 is calculated to have an IC50 value of 20 μM, whereas two structurally related compounds, compounds D1 and D3, showed better inhibitory activities with an IC_{50} value of 14.37 μM and 12.49 μM respectively (Figure 3.10). Results of compound 36 and natural compound N7 showed their IC50 are 7.398μM and 36.26 μM respectively (Figure 3.11).

Since natural compound N7 exhibited good inhibitory activity against the phosphorylation of ASF/SF2 by SRPK1, we obtained five structurally related derivatives from our collaborator Prof. Jiang at the Macau University of Science and Technology. Two of these natural compounds, rhynchophylline and yohimbine hydrochloride (Table 3.4), however, showed no inhibitory activity against the phosphorylation of ASF/SF2 when used at the indicated dosage (Figure 3.12). Investigations of the activities of the other three compounds are underway.

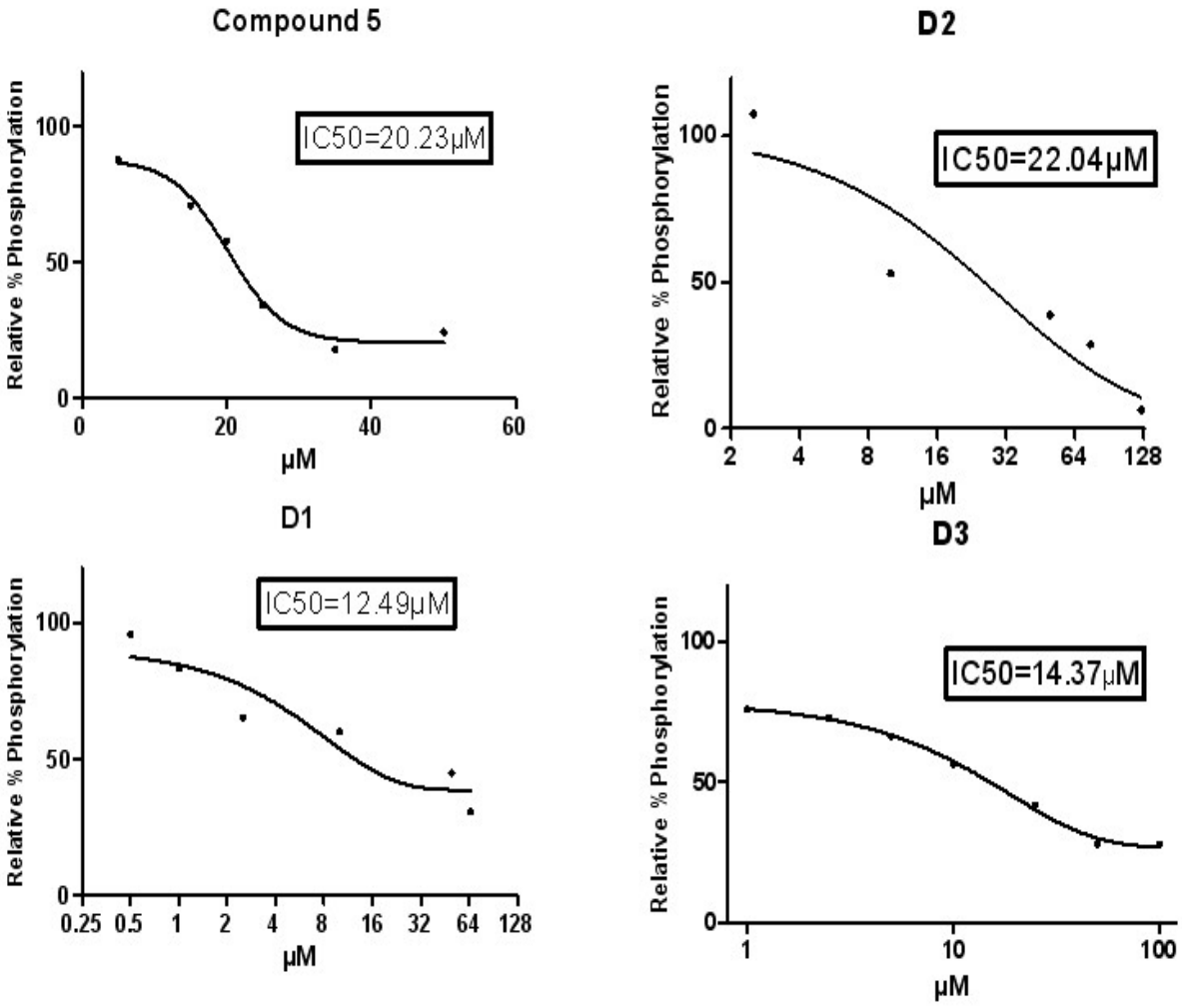


Figure 3. 10 *In vitro* kinase assay for compound 5, D1, D2 and D3. SRPK1, His-ASF/SF2 and inhibitors were incubated in 1× kinase buffer (100 mM Tris-HCl [pH 7.5], 10 mM $MgCl_2$, 5mg/ml BSA, and 1 mM DTT). The reaction was initiated with 50 uM ATP and trace amounts of γ^{32}P-ATP. The gel was dried and the bands are quantitated using a Phosphorlmager and the ImageQuant software (GE Healthcare). The results are plotted against the inhibitor concentration added to the kinase reaction. The IC50 of each compound are shown in the figure.

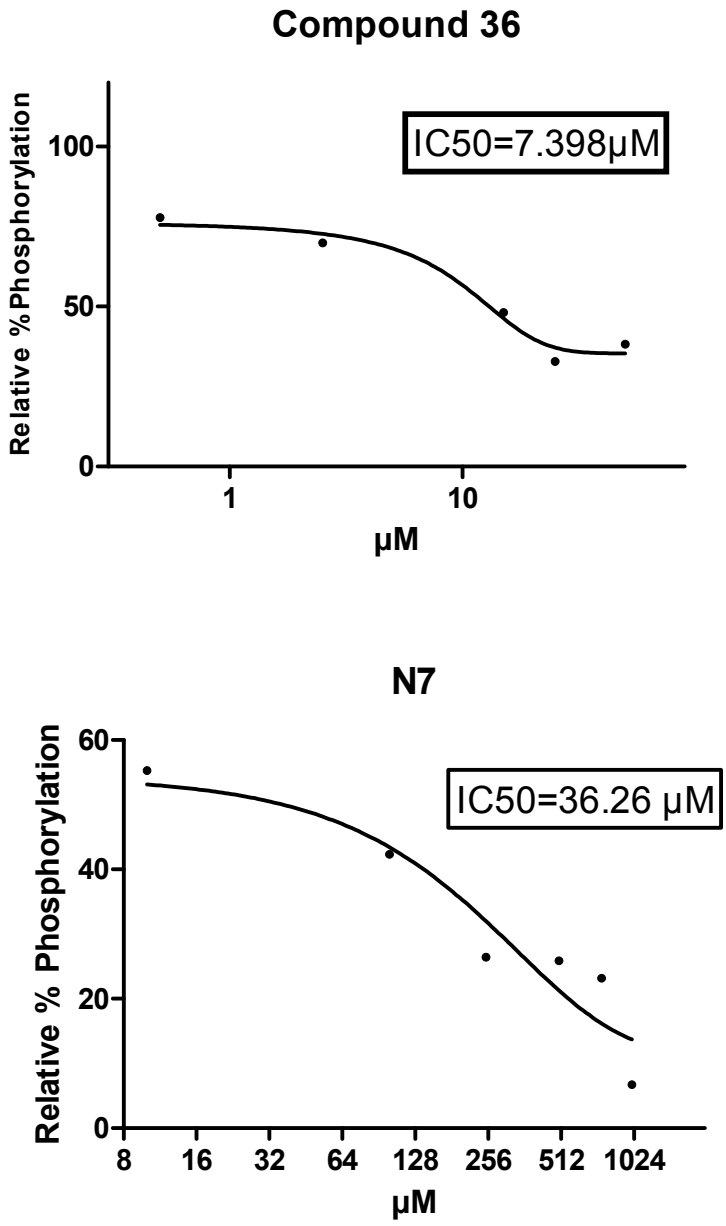
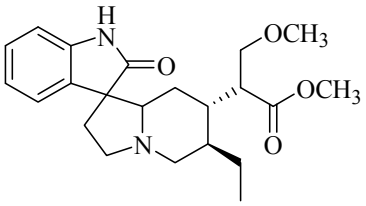
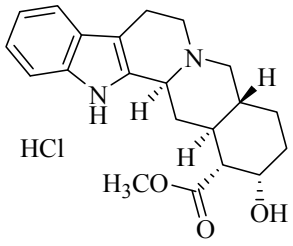


Figure 3. 11 *In vitro* kinase assay for compound 36 and natural compound N7. Experiment conditions are the same as in Figure 3.10. Compound N7 inhibited nearly 50% of ASF/SF2 phosphorylation even at the lowest concentration tested, while inhibited more than 90% of phosphorylation at 1 mM.

Table 3. 2 Two similar compounds of natural compound N7

Rhynchophylline $C_{22}H_{28}N_2O_4$	
yohimbine hydrochloride $C_{21}H_{26}N_2O_3 \cdot HCl$	

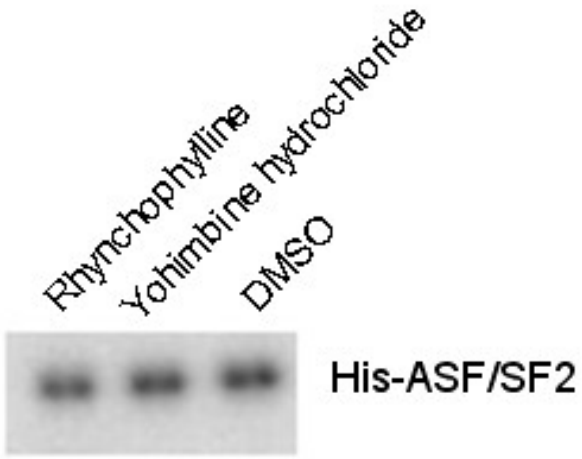


Figure 3. 12 *In vitro* kinase assay for rhynchophylline and yohimbine hydrochloride. Experiment conditions are the same as in Figure 3.10. The rhynchophylline and yohimbine hydrochloride failed to inhibit the phosphorylation of ASF/SF2.

3. 3 Surface Plasmon Resonance Binding Competition Assay

To test if the compounds inhibited the phosphorylation event by blocking protein-protein interactions between SRPK1 and ASF/SF2, I used surface plasmon resonance to measure if the inhibitors could affect SRPK1-ASF/SF2 binding. GST-tag ASF/SF2 was immobilized on a CM5 sensor chip pre-coated with GST antibodies; SRPK1 protein was pre-incubated with different compounds or DMSO for 15 mins before being applied in the mobile phase. If the selected compounds inhibit the interaction between SRPK1 and ASF/SF2, they would prevent formation of complex on the sensor chip and thus result in lower SPR signal. During the experiment, compound 23, which showed no inhibitory activity in our assays (Figure 3.13A), and a well known ATP-competing kinase inhibitor staurosporine were used as negative controls to show that the inhibition of interaction is specific and not due to binding of inhibitors at the ATP binding site.

All compounds were tested twice in the assays using the highest concentration possible, which beyond that the compounds start to precipitate. While compound 5 showed potent inhibitory effect in our phosphorylation assay, it did not appear to inhibit the interaction between SRPK1 and ASF/SF2 in the SPR experiment (Figure 3.13B). Instead, the SPR signal increased when the compound was applied in the mobile phase together with the kinase. An explanation to this observation is that the compound interacts with the protein complex at a site outside of the docking groove of SRPK1, therefore it failed to inhibit the interaction and resulted in a higher SPR signal.

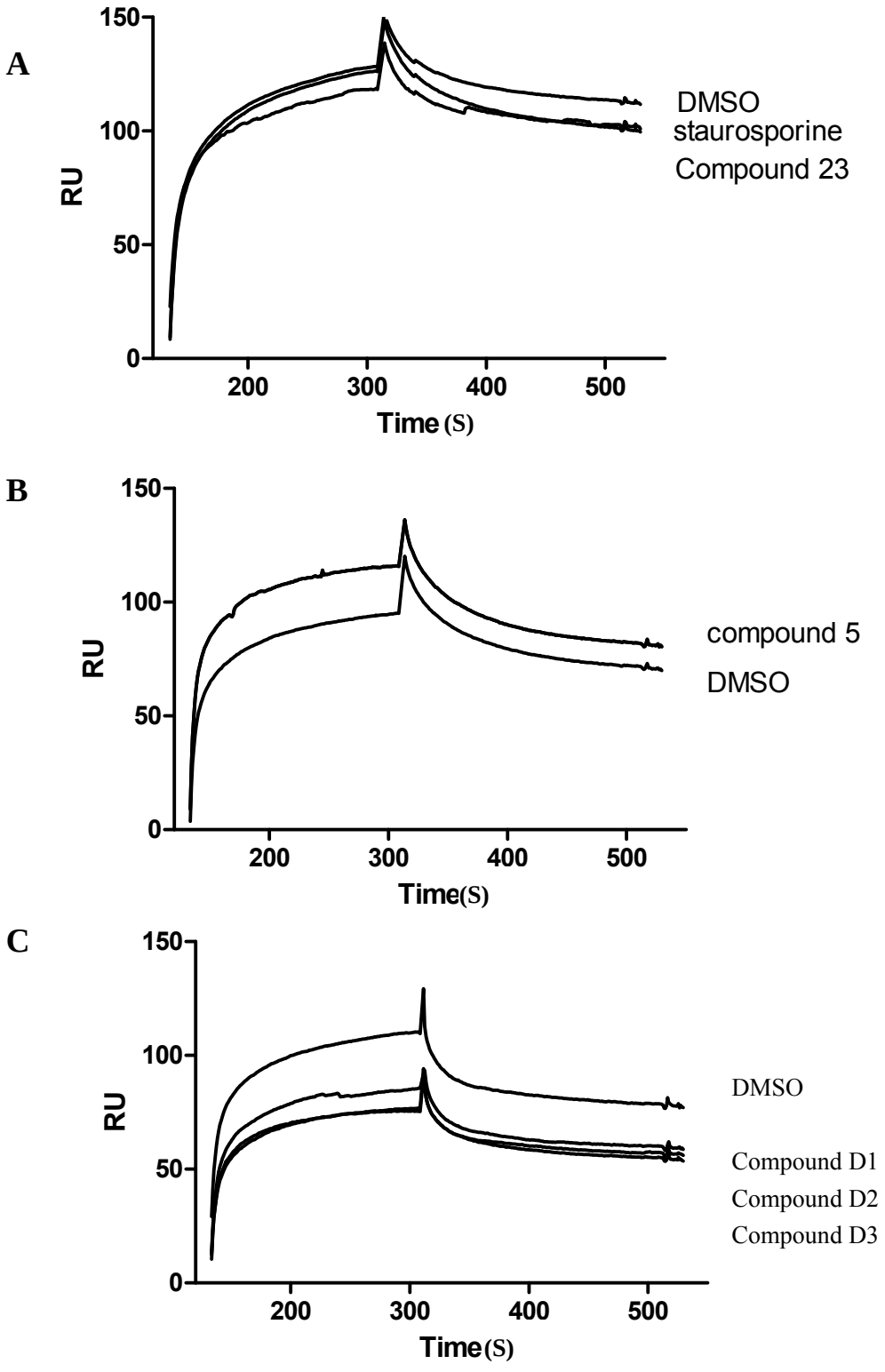


Figure 3. 13 Surface plasmon resonance sensorgrams for compound 5, D1, D2 and D3. Staurosporine and compound 23 were used as the negative control. The tested compounds do not appear to inhibit the binding between SRPK1 and ASF/SF2.

Natural Compound N7

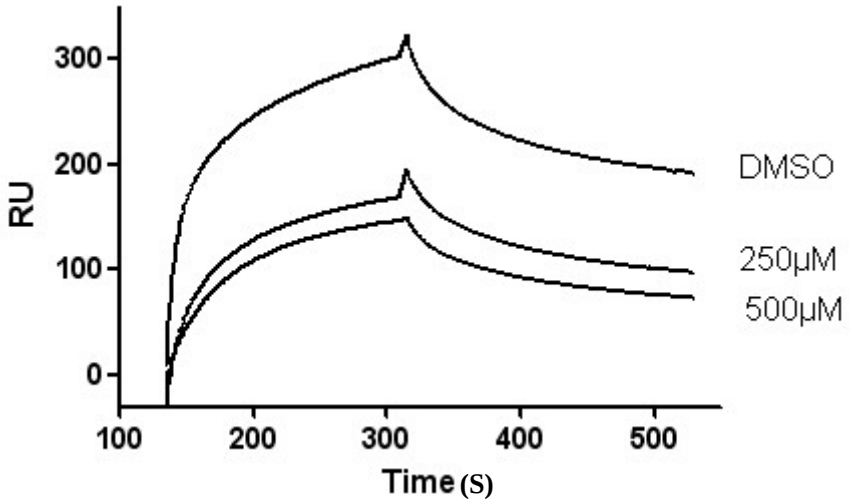


Figure 3. 14 Surface plasmon resonance sensorgram for natural compound N7. N7 appears to inhibit the binding of SRPK1 to ASF/SF2 and the IC50 for binding inhibition was estimated to be around 250μM.

Compounds D1, D2 and D3that are structurally related to compound 5 lowered the SPR signal of the complex by 20-30% (Figure 3.13C). However, due to limited solubility, the compounds could not be tested using higher concentration and thus the dose-dependent effect and IC50s of these compounds could not be determined.

Among all the compounds selected from the Natural Product library, compound N7 is the most potent in inhibiting the phosphorylation of ASF/SF2. However, while compound N7 could lower SPR signal in a dose-dependent manner, the IC50 was estimated to be around 250 μM, which is much higher than that obtained from *in vitro* kinase assay (Figure 3.14). The reason for the difference is unknown and more experiments like GST-pull down assay are needed to directly test the effect of the compound on the interaction between SRPK1 and ASF/SF2. Nonetheless, the SPR result showed that compound N7 possesses the ability to inhibit the binding of ASF/SF2 to SRPK1.

3.4 Inhibitors Alters HIV-1 Alternative Splicing *ex Vivo*

Given that ASF/SF2 plays a major role in HIV-1 pre-mRNA splicing, I tested the effects of 5 candidate compounds (natural compound N7, compound 5, D1, D2 and D3) on the HIV-1 pre-mRNA splicing using an *ex vivo* assay. pΔPSP plasmid, a HIV-1 proviral genome with deletion between nucleotides 1511 and 4550 that remove the encoding gene of gag and pol, was used for the *ex vivo* splicing assay to evaluate the inhibition efficiency of each candidates. Pre-mRNA transcribed from pΔPSP plasmid

does not differ significantly from those of wild-type HIV-1, and it contains all the splice sites and recapitulates all splicing events of HIV-1 pre-mRNA in transfected HeLa cells or infected peripheral blood mononuclear cells. RT-PCR was then performed to analyze the mRNA level produced by splicing using the forward primer BSS (5′-GGCTTGCTGAAGCGCGCACGGCAAGAGG-3′; nt 700–727) and reverse primers SJ4.7A, which spans sites D4 and A7 (5′-TTGGGAGGTGGGTTGCTTTGATAGAG-3′) to amplify all the 2 kb splicing isoforms encoding the viral proteins Nef, Rev, and Tat (Figure 3.15).

Four candidate compounds, including compounds 5, D1, D2 and D3 were tested in the *ex vivo* splicing assay using the highest soluble concentration (10μM, 15μM, 50μM, 7.5μM respectively). My results showed that while all four compounds inhibited HIV-1 pre-mRNA splicing to some extent, compounds 5 and D3 were more effective (Figure 3.16). This suggests the candidate compounds have inhibitory effects on HIV-1 splicing, however, their potencies were not consistent with the results obtained from *in vitro* kinase assays or SPR assays. Further experiments are required to determine the exact effect of these candidates on HIV-1 life cycle and their mechanisms of inhibition.

On the other hand, based on our *in vitro* results, compound N7 is a very promising candidate that might be a true SRPK1-ASF/SF2 interaction inhibitor. Compound N7 was tested *ex vivo* for its effect on HIV-1 splicing using different concentrations. Figure 3.17 shows that the synthesis of all splicing products from the transfected cells treated with compound N7 was less efficient when compared to DMSO-treated cells.

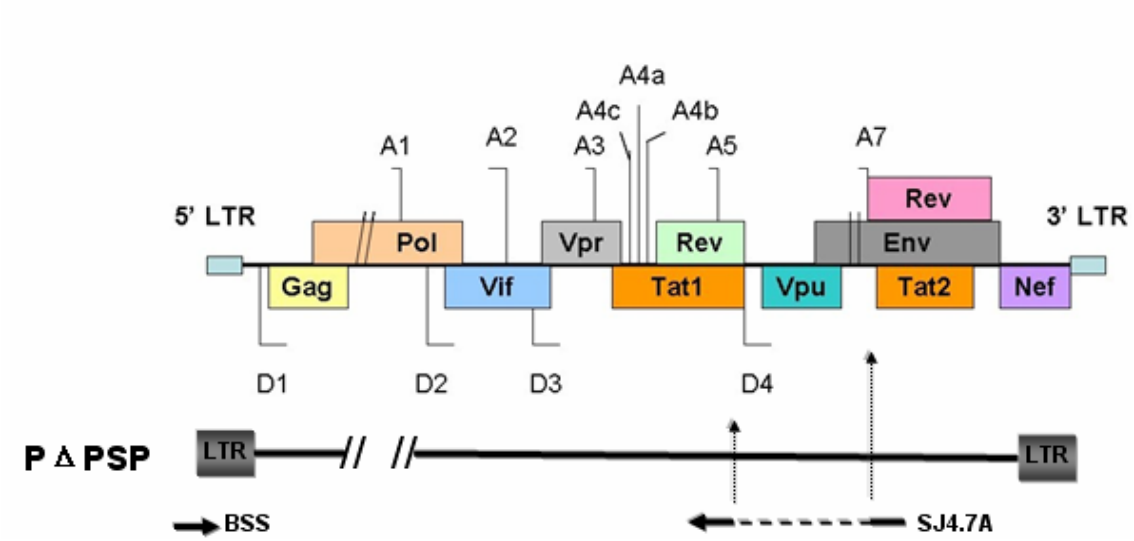


Figure 3. 15 Schematic representation of HIV-1 pΔPSP plasmid and primer design for *ex vivo* splicing assay. pΔPSP plasmid, HIV-1 proviral genome deleted between nucleotides 1511 and 4550. Forward primer BSS (5′-GGCTTGCTGAAGCGCGCACGGCAAGAGG-3′; nt 700–727) and reverse primers SJ4.7A, which spans sites D4 and A7 (5′-TTGGGAGGTGGGTTGCTTTGATAGAG-3′) were used to amplify all the 2 kb splicing isoforms.

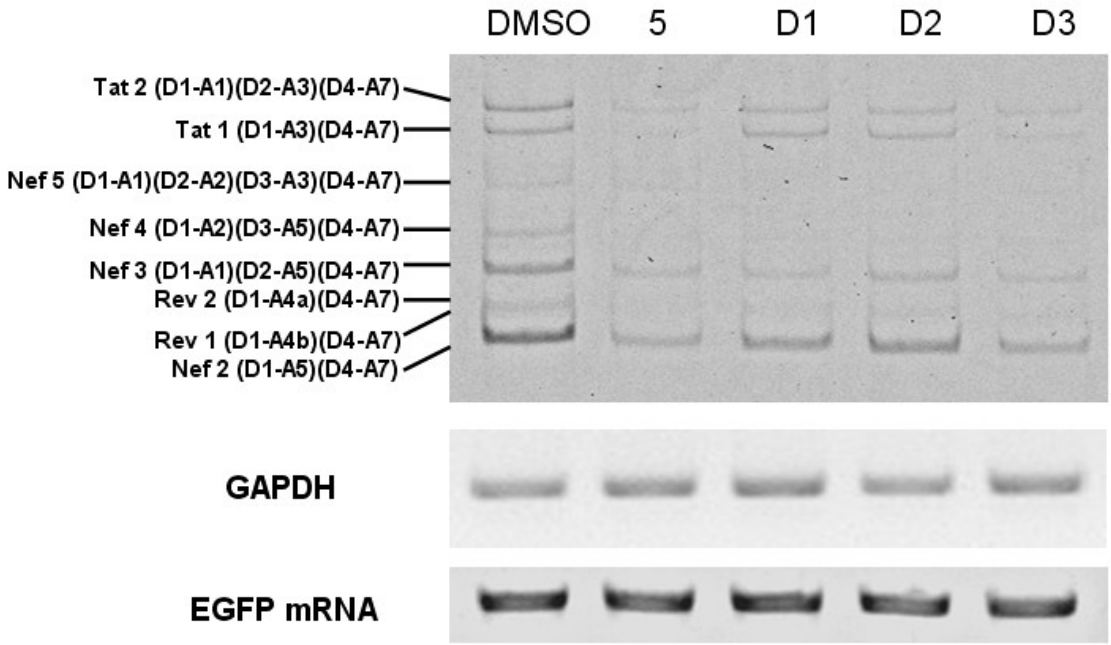


Figure 3. 16 Compounds 5, D1, D2 and D3 alter HIV alternative splicing. HEK 293 cells transfected with the pΔPSP construct and EGFP were treated with DMSO or four select compounds (lanes 1–5, respectively). Multiple spliced products of HIV-1 RNA were amplified by RT-PCR using the oligonucleotide primers BSS and SJ4.7A. The PCR products were analyzed by polyacrylamide gel electrophoresis after normalization with GAPDH and EGFP. The intensities of the HIV-1 pre-mRNA splicing products decreased in the presence of compounds, suggesting the compounds had altered the splicing activity.

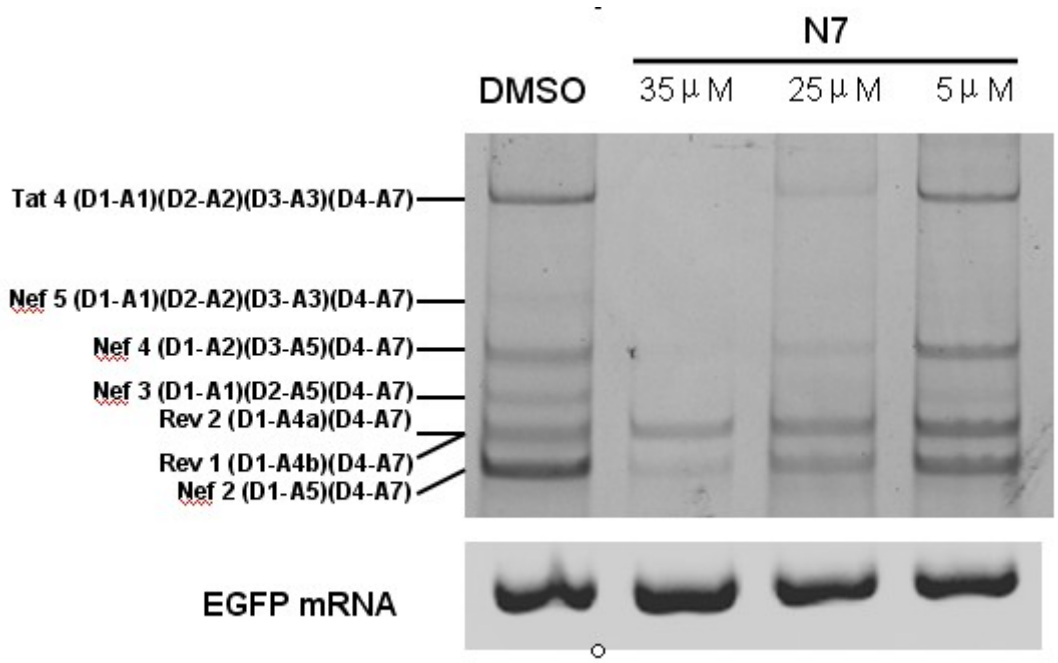


Figure 3. 17 Selective inhibition of HIV-1 RNA splicing *ex vivo* by natural compound N7. Experiment conditions are the same as in Figure 3.16. Compound N7 appears to inhibit the splicing of HIV-1 pre-mRNA in a dose-dependent manner.

Furthermore, the inhibition effect of compound N7 was dose dependent. The inhibitory effect is particular obvious for the synthesis of the larger splicing isoforms, which were completely blocked with 35 μM of compound (Figure 3.17).

This dose-dependent profile of splicing inhibition indicated that the natural compound N7 inhibits the usage of several 3'splice sites which are critically regulated by the binding of SR proteins, especially ASF/SF2. This result, together with our *in vitro* results, proved that by inhibiting the phosphorylation of ASF/SF2, it is possible to abrogate the pre-mRNA splicing of HIV-1.

3.5 Cytotoxic effect of candidate compound on HeLa cells

Since compound N7 appeared to be the sole inhibitor that blocks ASF/SF2 phosphorylation by disrupting its interaction with SRPK1, as well as abrogating the splicing of HIV-1 mRNA, I determined to continue investigating the biological activity of this compound and its effect on HIV-1 pathogenic function. In order to determine cytotoxicity of natural compound and the suitable concentration for the following *in vivo* experiment, MTT assay was first preformed to test the effect of compound N7 on HeLa cells. HeLa cells were exposed to compound N7 with concentrations ranging from 1 to 200 μM for 24 hours. Cell viability was then measured by an MTT assay and presented as a percentage of colorimetric absorbance derived from the compound-treated cells compared with that from the mock-treated cells (with 1% DMSO). The cells remained viable in the presence of 50 μM or lower concentration of

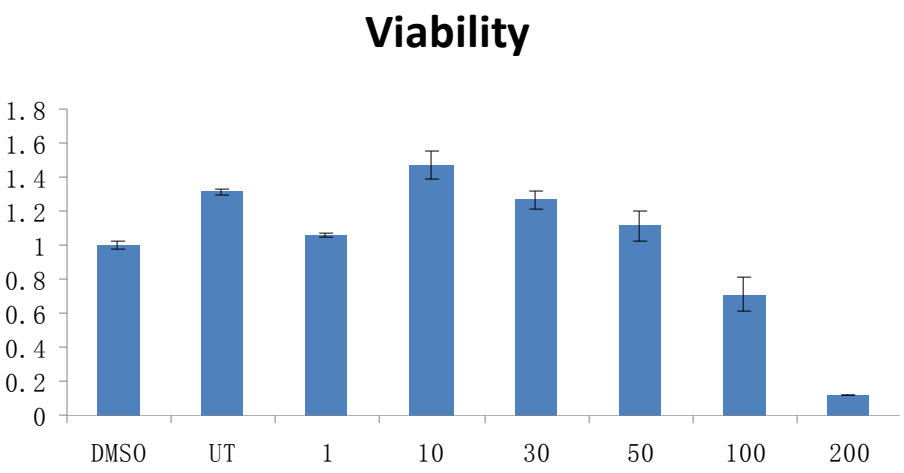


Figure 3. 18 Cytotoxic effect of natural compound N7 on HeLa cells in MTT assay. Cytotoxicity of HeLa cells after natural compound exposure in different concentrations for 24 h. The viability is normalized with 1% DMSO. Cells remained viable in the presence of 50 μM compound but started dying in the presence of higher concentration of compound.

compound (Figure 3.18). The concentration at which the compound reduced the number of viable cells to 50% was around 100μM, confirming that the concentration being used in the *ex vivo* splicing assay did not induce toxicity to the cells.

3.6 Nature compound alters ASF/SF2 localization

To explore how inhibition of ASF/SF2 phosphorylation affects its biological function, we performed immunocytochemistry to test the localization of ASF/SF2 in the presence of either the natural compound or the vehicle control (DMSO). Previous studies have indicated that ASF/SF2 mainly localizes in the nucleus in various cell types, but the protein could also be found shuttling between the nucleus and cytoplasm. In addition, localization of ASF/SF2 is depended on its phosphorylation by different kinases. We hypothesized that any inhibition in the phosphorylation mechanism of SRPK1 might disturb the cellular distribution of the SR proteins *in vivo*, especially ASF/SF2.

Wild-type myc-ASF/SF2 was transfected and overexpressed in HeLa cells and analyses for subcellular distribution by indirect immunofluorescence with anti-myc antibody. We observed that ASF/SF2 (wt) localized predominantly in the nucleus in the presence of the vehicle control DMSO as expected. However, myc-ASF/SF2 (wt) treated with 30 μM nature compound N7 was detected to diffuse in both nucleus and cytoplasm (Figure 3.19). Our result suggests that the natural compound N7 successfully altered the localization of ASF/SF2. Since phosphorylation of the RS

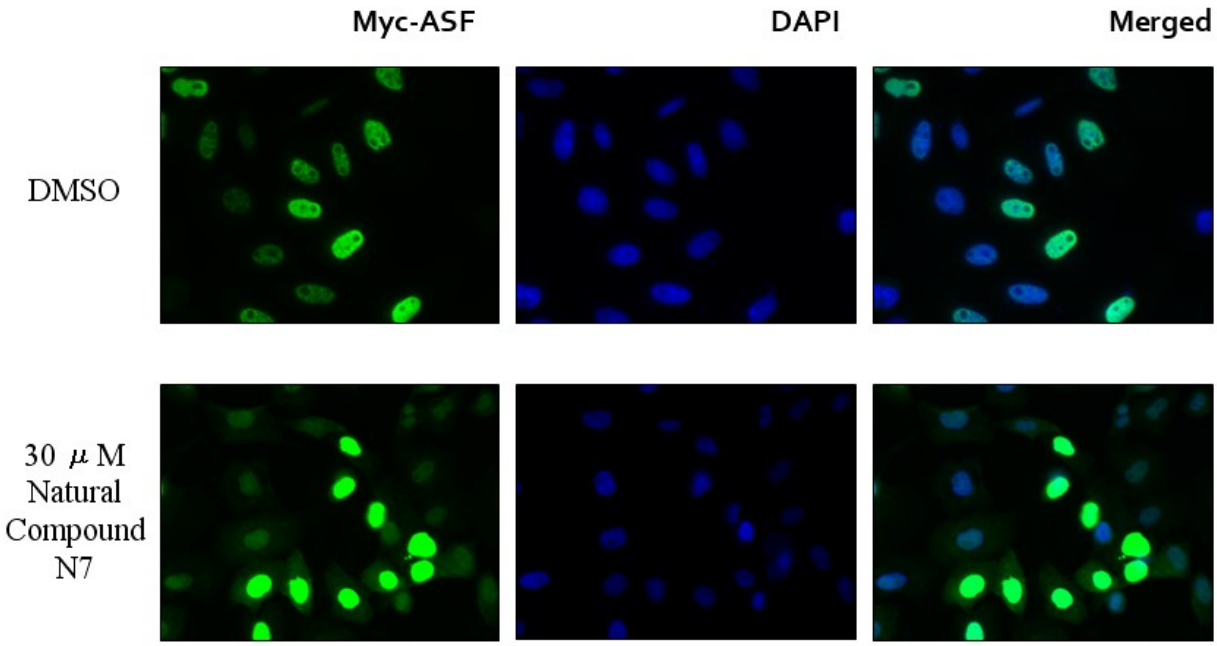


Figure 3. 19 The distribution of ASF/SF2 is altered by natural compound N7 *in vivo*. Indirect immunofluorescent staining of transfected HeLa cells using the anti-Myc antibody. Localization studies of Myc-ASF/SF2 (wt) show that ASF/SF2 are localized predominantly in the nucleus with the vehicle control (DMSO). The natural compound causes diffusion of ASF/SF2 in both of the nucleus and cytoplasm, suggesting it is no longer phosphorylated by SRPK1.

domain is critical for nuclear import of ASF/SF2, this assay demonstrates that the natural compound might inhibit specific phosphorylation-dependent localization of ASF/SF2.

Chapter Ⅳ: Discussion and Conclusion

Our aim is to develop a novel class of SRPK1 inhibitors to interfere the binding between the SRPK1 and ASF/SF2 by blocking the docking groove of SRPK1 and inhibit the phosphorylation of ASF/SF2. However, due to the high cost and long time required for drug discovery using conventional methods like high-throughput screening, we adopted computer-based drug screening method to identify potential candidates.

We have designed a two-step screening approach which combined AutoDock 4 and AutoDock Vina for new inhibitor discovery. This approach takes the full advantages of two software, by which we saved much screening time as well as achieving more accurate results through two different algorithms. AutoDock Vina greatly speeded up the screening procedures, providing the chance to screen a much large database of small molecules. On the other hand, AutoDock 4 following AutoDock Vina helped to reconfirm the docking results of each compound and increased the accuracy of the molecular modeling simulation.

351,473 compounds from the natural products database of ZINC and the commercially available SPECS were firstly screened by AutoDock Vina and 2846 compounds were selected with good binding affinity to SRPK1 docking groove. After rescreen of these 2846 candidates using AutoDock 4, 42 compounds ranked as the top candidates that have the highest binding affinities to SRPK1 in both programs were selected and purchased for further analysis. These candidates include 8 compounds from the natural products library and 34 compounds from the SPECS library. Among these 42 candidates, three compounds belonging to SPECS and one natural compound

were shown to inhibit the phosphorylation of ASF/SF2 by SRPK1 with potencies at micromolar concentrations. In particular, compound 5 and natural compound N7 were proven to be effective inhibitors. Based in this result, I further selected and purchase 3 structural-related compounds of compound 5 to test their effects on SRPK1 activity. These three compounds also appeared to exert good inhibition effects on SRPK1. However, since our aim is to develop protein-protein binding inhibitors, we needed to confirm if the selected compounds block the protein-protein interaction. While compound 5 and its analogues could inhibit the phosphorylation of ASF/SF2 by SRPK1, they failed to block the kinase-substrate interaction effectively in our SRP assays. On the other hand, the natural compound N7 successfully inhibited the binding of ASF/SF2 to SRPK1 with an IC50 around 250μM. This is inconsistent with the result obtained in the kinase activity assay, which showed that compound N7 could inhibit SRPK1 activity with an IC50 around 36μM. The reason for such inconsistency is unknown but it is possible that the effective concentrations of compound N7 in both assays are different. Since many compounds I have tested had limited solubility even in the presence of 5% DMSO, it is highly possible that some of the DMSO-dissolved compound N7 might have precipitated out from the solution phase under the different experimental conditions being used for the two assays, thus leading to large error in the calculation of IC50s. Further research is required to investigate the exact reason for the observed inconsistency.

Another interesting observation is that while 5 of the selected candidate inhibitors, namely natural compound N7, compound 5, D1, D2 and D3 could inhibit SRPK1 to

some extent, they could not completely block the kinase activity. This could due the extensive amount of RS dipeptide repeats present in the RS domain of ASF/SF2. While SRPK1 has been shown to utilize its docking groove to restrict the phosphorylation of ASF/SF2 to the N-terminal half of its RS domain, it is unpreventable that some minor non-specific phosphorylation could have happened at the C-terminal half of ASF/SF2 RS domain since both chemical and physical properties of the two halves are highly similar. Furthermore, the kinase activity *in vitro* is not abrogated in the presence of the protein-protein interaction inhibitor. Therefore, even the inhibitors could block the docking groove of SRPK1 and prevent specific phosphorylation of the N-terminal RS domain of ASF/SF2, non-specific phosphorylation elsewhere could not be prevented. Moreover, alteration in phosphorylation state of ASF/SF2, but not necessarily complete inhibition of phosphorylation, could still abrogate ASF/SF2's function in HIV-1 splicing.

The predicted binding models of the candidate compounds to SRPK1 showed that the docked compounds form direct contracts with several amino acids within the docking groove of SRPK1 (Figure 4.1 and 4.2). Particularly, the top candidate compounds were predicted to interact with at least one of the docking groove residues D548, D564, E571, and K615 that are essential for substrate binding to the docking groove. However, we could not get promising results for some of the inhibitors in the binding competition assays. One explanation is that the candidate compounds might have bound to other region(s) outside of the intended docking site, resulting in nonspecific binding on the kinase and cause allosteric inhibition on the kinase. Alternately, they might have inhibit the kinase activity by binding to the ATP-binding

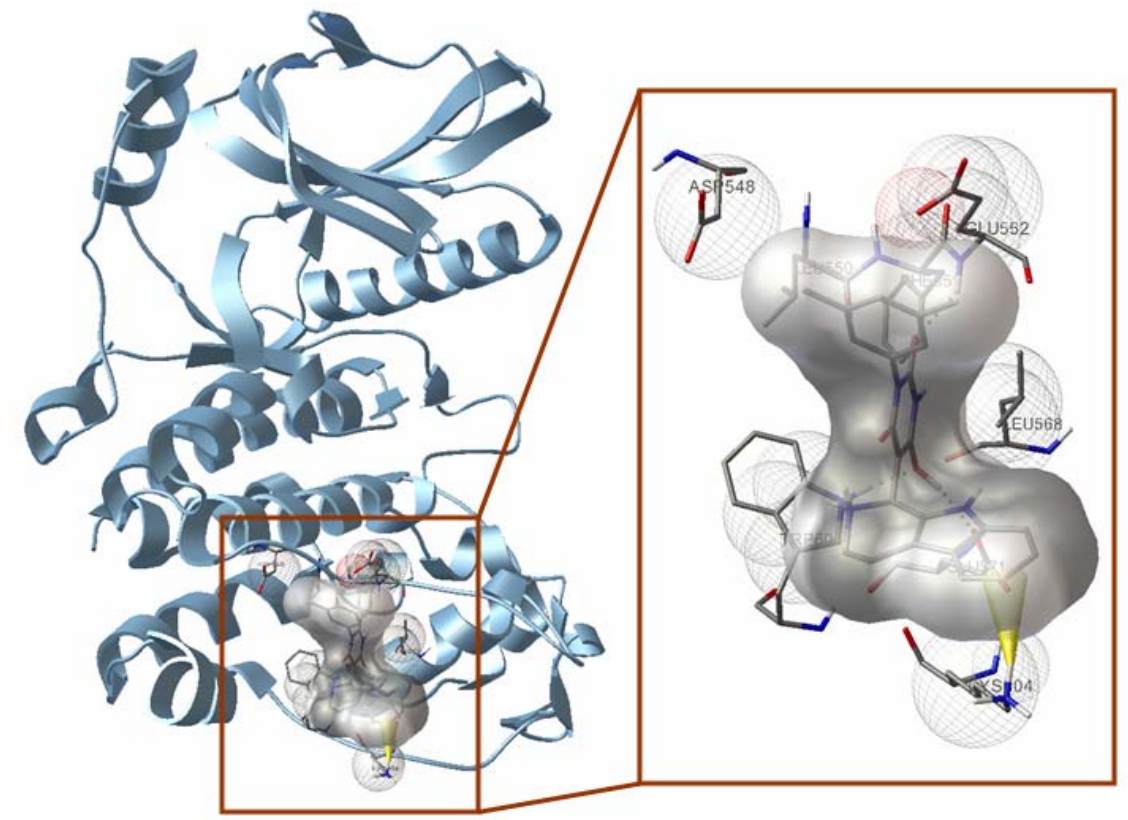


Figure 4. 1 Predicted binding model of compound 5 to SRPK1. The binding mode of compound 5 is shown. The SRPK1 structure is shown in blue. Sulfur, oxygen, or nitrogen atoms on the active compounds are indicated as yellow, red, or blue spheres, respectively. The residues that interact with the compound are denoted in the inset on the right.

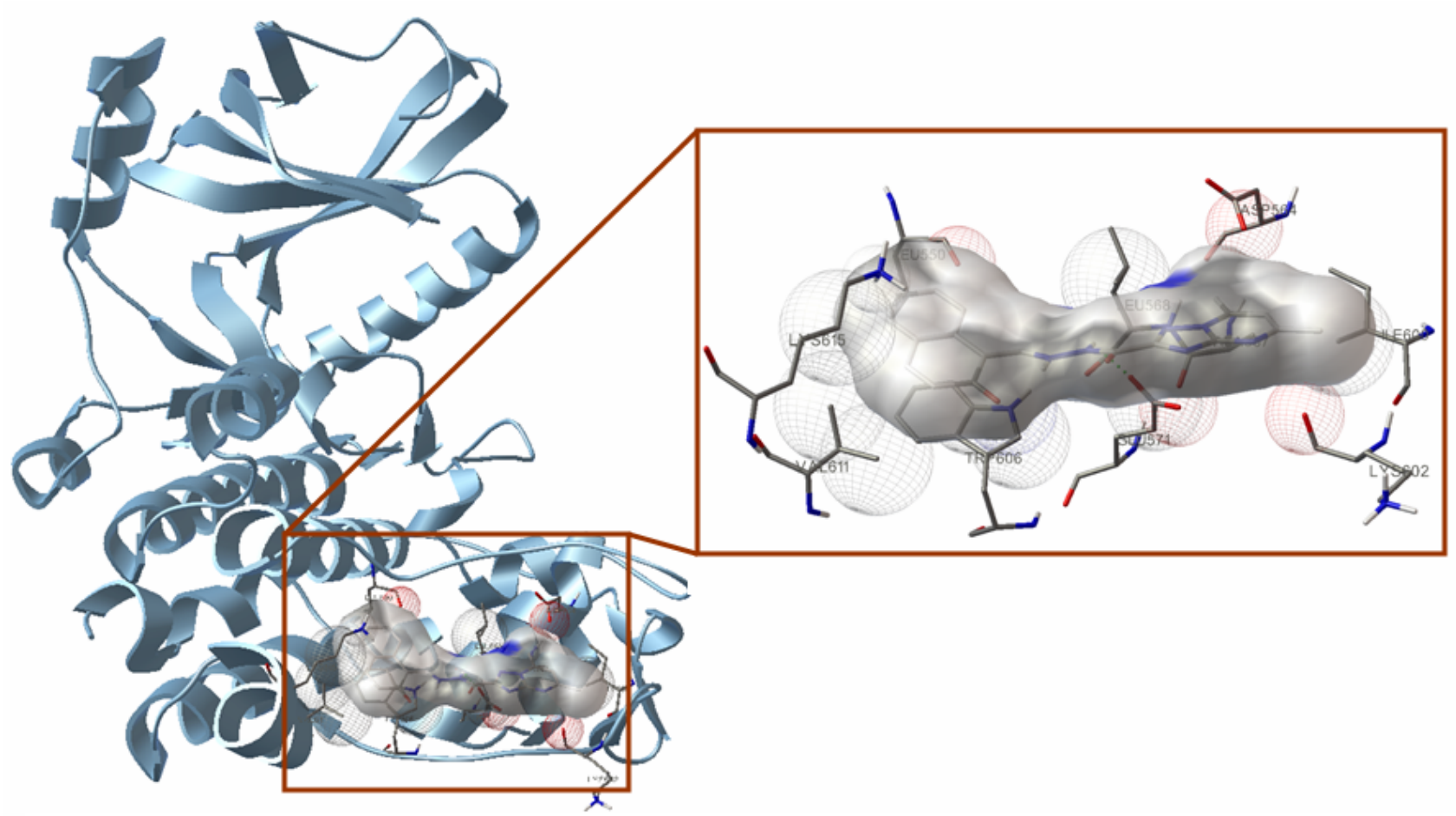


Figure 4. 2 Predicted binding model of natural compound N7 to SRPK1. The binding mode of c natural compound is shown. The SRPK1 structure is shown in blue. Sulfur, oxygen, or nitrogen atoms on the active compounds are indicated as yellow, red, or blue spheres, respectively.

cleft non-specfically. To investigate the real inhibition mechanism, further research like the resolution of the structure of SRPK1 in complex with the candidate compounds are required.

The most promising candidate obtained in this project, natural compound N7, has been proven *in vitro* to inhibit the phosphorylation of ASF/SF2 by SRPK1 through the blocking of kinase-substrate interaction. The compound has also been shown to affect the localization of ASF/SF2 in vivo, likely due to an alteration of its phosphorylation state as we have proposed. Furthermore, it displayed a good dose-dependent inhibition profile of HIV-1 pre-mRNA splicing *ex vivo*. More importantly, natural compound N7 also showed a promising effect in a syncytium reduction assays performed by our collaborators at the Kunming Institute of Zoology using HIV-1 virus infected human T cells (data not shown). Although the TI (therapeutic index) of natural compound N7 is above 11.866, which is much lower than the control drug AZT, this could serve as a lead for more potent HIV inhibitor. Therefore, while the detailed inhibition mechanism of this compound remains to be fully elucidated, my results present here strongly suggest that compound N7 could be a true hit and could be further developed into a novel HIV-1 treatment that targets its splicing pathway.

References

Bakkour, N., Lin, Y.L., Maire, S., Ayadi, L., Mahuteau-Betzer, F., Nguyen, C.H., Mettling, C., Portales, P., Grierson, D., Chabot, B. Jeanteur, P.m Branlant C., Corbeau P., Tazi J. (2007). Small-molecule inhibition of HIV pre-mRNA splicing as a novel antiretroviral therapy to overcome drug resistance. PLoS Pathog *3*, 1530-1539.

Barbaro, G., Scozzafava, A., Mastrolorenzo, A., and Supuran, C.T. (2005). Highly active antiretroviral therapy: Current state of the art, new agents and their pharmacological interactions useful for improving therapeutic outcome. Curr Pharm Design *11*, 1805-1843.

Bilodeau, P.S., Domsic, J.K., Mayeda, A., Krainer, A.R., and Stoltzfus, C.M. (2001). RNA splicing at human immunodeficiency virus type 1 3' splice site A2 is regulated by binding of hnRNP A/B proteins to an exonic splicing silencer element. J Virol *75*, 8487-8497.

Black, D.L. (2003). Mechanisms of alternative pre-messenger RNA splicing. Annu Rev Biochem *72*, 291-336.

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

Blencowe, B.J. (2000). Exonic splicing enhancers: mechanism of action, diversity and role in human genetic diseases. Trends Biochem Sci *25*, 106-110.

Blencowe, B.J., Bowman, J.A., McCracken, S., and Rosonina, E. (1999). SR-related

proteins and the processing of messenger RNA precursors. Biochem Cell Biol *77*, 277-291.

Boyer, P.L., Currens, M.J., Mcmahon, J.B., Boyd, M.R., and Hughes, S.H. (1993). Analysis of Nonnucleoside Drug-Resistant Variants of Human-Immunodeficiency-Virus Type-1 Reverse-Transcriptase. Journal of Virology *67*, 2412-2420.

Caceres, J.F., Misteli, T., Screaton, G.R., Spector, D.L., and Krainer, A.R. (1997). Role of the modular domains of SR proteins in subnuclear localization and alternative splicing specificity. J Cell Biol *138*, 225-238.

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

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

Calarco, J.A., Superina, S., O'Hanlon, D., Gabut, M., Raj, B., Pan, Q., Skalska, U., Clarke, L., Gelinas, D., van der Kooy, D., Zhen, M., Ciruna, B., Blencow, B.J. (2009). Regulation of vertebrate nervous system alternative splicing and development by an SR-related protein. Cell *138*, 898-910.

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

Caputi, M., Mayeda, A., Krainer, A.R., and Zahler, A.M. (1999). hnRNP A/B proteins are required for inhibition of HIV-1 pre-mRNA splicing. EMBO J *18*, 4060-4067.

Caputi, M., and Zahler, A.M. (2002). SR proteins and hnRNP H regulate the splicing of the HIV-1 tev-specific exon 6D. EMBO J *21*, 845-855.

Chan, D.C., and Kim, P.S. (1998). HIV entry and its inhibition. Cell *93*, 681-684.

Colwill, K., Feng, L.L., Yeakley, J.M., Gish, G.D., Caceres, J.F., Pawson, T., and Fu, X.D. (1996a). SRPK1 and Clk/Sty protein kinases show distinct substrate specificities for serine/arginine-rich splicing factors. J Biol Chem *271*, 24569-24575.

Colwill, K., Feng, L.L., Yeakley, J.M., Gish, G.D., Caceres, J.F., Pawson, T., and Fu, X.D. (1996b). SRPK1 and Clk/Sty protein kinases show distinct substrate specificities for serine/arginine-rich splicing factors. Journal of Biological Chemistry *271*, 24569-24575.

Cowper, A.E., Caceres, J.F., Mayeda, A., and Screaton, G.R. (2001). Serine-arginine (SR) protein-like factors that antagonize authentic SR proteins and regulate alternative splicing. Journal of Biological Chemistry *276*, 48908-48914.

Dagher, S.F., and Fu, X.D. (2001). Evidence for a role of Sky1 p-mediated phosphorylation in 3 ' splice involving both Prp8 and site recognition Prp17/Slu4. Rna *7*, 1284-1297.

del Rio, C. (2006). Current concepts in antiretroviral therapy failure. Top HIV Med *14*, 102-106.

Dybul, M., Fauci, A.S., Bartlett, J.G., Kaplan, J.E., and Pau, A.K. (2002). Guidelines for using antiretroviral agents among HIV-infected adults and adolescents. Recommendations of the Panel on Clinical Practices for Treatment of HIV. MMWR Recomm Rep *51*, 1-55.

Esnouf, R.M., Ren, J.S., Hopkins, A.L., Ross, C.K., Jones, E.Y., Stammers, D.K., and Stuart, D.I. (1997). Unique features in the structure of the complex between HIV-1 reverse transcriptase and the bis(heteroaryl)piperazine (BHAP) U-90152 explain resistance mutations for this nonnucleoside inhibitor. P Natl Acad Sci USA *94*, 3984-3989.

Fetzer, S., Lauber, J., Will, C.L., and Luhrmann, R. (1997). The [U4/U6 center dot U5] tri-snRNP-specific 27K protein is a novel SR protein that can be phosphorylated by the snRNP-associated protein kinase. Rna *3*, 344-355.

Fry, D.C., and Vassilev, L.T. (2005). Targeting protein-protein interactions for cancer therapy. J Mol Med-Jmm *83*, 955-963.

Fu, X.D., and Maniatis, T. (1992). Isolation of a Complementary-DNA That Encodes the Mammalian Splicing Factor Sc35. Science *256*, 535-538.

Fukuhara, T., Hosoya, T., Shimizu, S., Sumi, K., Oshiro, T., Yoshinaka, Y., Suzuki, M., Yamamoto, N., Herzenberg, L.A., and Hagiwara, M. (2006). Utilization of host SR protein kinases and RNA-splicing machinery during viral replication. Proc Natl Acad Sci U S A *103*, 11329-11333.

Geretti, A.M. (2006). Clinical implications of HIV drug resistance to nucleoside and nucleotide reverse transcriptase inhibitors. AIDS Rev *8*, 210-220.

Gorry, P.R., Howard, J.L., Churchill, M.J., Anderson, J.L., Cunningham, A., Adrian, D., McPhee, D.A., and Purcell, D.F. (1999). Diminished production of human immunodeficiency virus type 1 in astrocytes results from inefficient translation of gag, env, and nef mRNAs despite efficient expression of Tat and Rev. J Virol *73*, 352-361.

Graveley, B.R. (2000). Sorting out the complexity of SR protein functions. Rna *6*,

1197-1211.

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

Hancock, C.N., Macias, A., Lee, E.K., Yu, S.Y., MacKerell, A.D., and Shapiro, P. (2005). Identification of novel extracellular signal-regulated kinase docking domain inhibitors. Journal of Medicinal Chemistry *48*, 4586-4595.

Hedley, M.L., Amrein, H., and Maniatis, T. (1995). An amino acid sequence motif sufficient for subnuclear localization of an arginine/serine-rich splicing factor. Proc Natl Acad Sci U S A *92*, 11524-11528.

Henkel, J. (1999). Attacking AIDS with a 'cocktail' therapy? FDA Consum *33*, 12-17.

Hertel, K.J., Lynch, K.W., and Maniatis, T. (1997). Common themes in the function of transcription and splicing enhancers. Curr Opin Cell Biol *9*, 350-357.

Huang, N., Nagarsekar, A., Xia, G.J., Hayashi, J., and MacKerell, A.D. (2004). Identification of non-phosphate-containing small molecular weight inhibitors of the tyrosine kinase p56 Lck SH2 domain via in silico screening against the pY+3 binding site. Journal of Medicinal Chemistry *47*, 3502-3511.

Jacquenet, S., Decimo, D., Muriaux, D., and Darlix, J.L. (2005). Dual effect of the SR proteins ASF/SF2, SC35 and 9G8 on HIV-1 RNA splicing and virion production. Retrovirology *2*, 33.

Jacquenet, S., Mereau, A., Bilodeau, P.S., Damier, L., Stoltzfus, C.M., and Branlan, C. (2001). A second exon splicing silencer within human immunodeficiency virus type 1

tat exon 2 represses splicing of Tat mRNA and binds protein hnRNP H. Journal of Biological Chemistry *276*, 40464-40475.

Johnson, L.N., Lowe, E.D., Noble, M.E.M., and Owen, D.J. (1998). The structural basis for substrate recognition and control by protein kinases. Febs Letters *430*, 1-11.

Kammler, S., Otte, M., Hauber, I., Kjems, J., Hauber, J., and Schaal, H. (2006). The strength of the HIV-1 3' splice sites affects Rev function. Retrovirology *3*, 89.

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

Kim, S., Ikeuchi, K., Byrn, R., Groopman, J., and Baltimore, D. (1989). Lack of a negative influence on viral growth by the nef gene of human immunodeficiency virus type 1. Proc Natl Acad Sci U S A *86*, 9544-9548.

Klotman, M.E., Kim, S.Y., Buchbinder, A., DeRossi, A., Baltimore, D., and Wong-Staal, F. (1992). Kinetics of expression of multiply spliced RNA in early human immunodeficiency virus type 1 infection of lymphocytes and monocytes. Proc Natl Acad Sci U S A *89*, 1148.

Kohtz, J.D., Jamison, S.F., Will, C.L., Zuo, P., Luhrmann, R., Garcia-Blanco, M.A., and Manley, J.L. (1994). Protein-protein interactions and 5'-splice-site recognition in mammalian mRNA precursors. Nature *368*, 119-124.

Koizumi, J., Okamoto, Y., Onogi, H., Mayeda, A., Krainer, A.R., and Hagiwara, M. (1999). The subcellular localization of SF2/ASF is regulated by direct interaction with SR protein kinases (SRPKs). Journal of Biological Chemistry *274*, 11125-11131.

Kroemer, R.T. (2007). Structure-based drug design: docking and scoring. Curr Protein

Pept Sci *8*, 312-328.

Kuroyanagi, N., Onogi, H., Wakabayashi, T., and Hagiwara, M. (1998). Novel SR-protein-specific kinase, SRPK2, disassembles nuclear speckles. Biochem Biophys Res Commun *242*, 357-364.

Lai, M.C., Lin, R.I., Huang, S.Y., Tsai, C.W., and Tarn, W.Y. (2000). A human importin-beta family protein, transportin-SR2, interacts with the phosphorylated RS domain of SR proteins. J Biol Chem *275*, 7950-7957.

Lai, M.C., Lin, R.I., and Tarn, W.Y. (2001). Transportin-SR2 mediates nuclear import of phosphorylated SR proteins. P Natl Acad Sci USA *98*, 10154-10159.

Lam, B.J., and Hertel, K.J. (2002). A general role for splicing enhancers in exon definition. Rna *8*, 1233-1241.

Lim, L.P., and Burge, C.B. (2001). A computational analysis of sequence features involved in recognition of short introns. P Natl Acad Sci USA *98*, 11193-11198.

Liu, Y., and Gray, N.S. (2006). Rational design of inhibitors that bind to inactive kinase conformations. Nature Chemical Biology *2*, 358-364.

Lutzelberger, M., Reinert, L.S., Das, A.T., Berkhout, B., and Kjems, J. (2006a). A novel splice donor site in the gag-pol gene is required for HIV-1 RNA stability. Journal of Biological Chemistry *281*, 18644-18651.

Lutzelberger, M., Reinert, L.S., Das, A.T., Berkhout, B., and Kjems, J. (2006b). A novel splice donor site in the gag-pol gene is required for HIV-1 RNA stability. J Biol Chem *281*, 18644-18651.

Madhani, H.D., and Guthrie, C. (1994). Dynamic Rna-Rna Interactions in the Spliceosome. Annu Rev Genet *28*, 1-26.

Makarova, O.V., Makarov, E.M., and Luhrmann, R. (2001). The 65 and 110 kDa SR-related proteins of the U4/U6 center dot U5 tri-snRNP are essential for the assembly of mature spliceosomes. Embo Journal *20*, 2553-2563.

Manley, J.L., and Krainer, A.R. (2010). A rational nomenclature for serine/arginine-rich protein splicing factors (SR proteins). Gene Dev *24*, 1073-1074.

Manley, P.W., Cowan-Jacob, S.W., and Mestan, J. (2005). Advances in the structural biology, design and clinical development of Bcr-Abl kinase inhibitors for the treatment of chronic myeloid leukaemia. Bba-Proteins Proteom *1754*, 3-13.

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

Mayeda, A., Screaton, G.R., Chandler, S.D., Fu, X.D., and Krainer, A.R. (1999). Substrate specificities of SR proteins in constitutive splicing are determined by their RNA recognition motifs and composite pre-mRNA exonic elements. Molecular and Cellular Biology *19*, 1853-1863.

McLaren, M., Marsh, K., and Cochrane, A. (2008). Modulating HIV-1 RNA processing and utilization. Front Biosci *13*, 5693-5707.

Nakagawa, O., Arnold, M., Nakagawa, M., Hamada, H., Shelton, J.M., Kusano, H., Harris, T.M., Childs, G., Campbell, K.P., Richardson, J.A., Nishino, I., Olson, E.N. (2005). Centronuclear myopathy in mice lacking a novel muscle-specific protein kinase transcriptionally regulated by MEF2. Genes Dev *19*, 2066-2077.

Ngo, J.C., Gullingsrud, J., Giang, K., Yeh, M.J., Fu, X.D., Adams, J.A., McCammon, J.A., and Ghosh, G. (2007a). SR protein kinase 1 is resilient to inactivation. Structure *15*, 123-133.

Ngo, J.C., Huang, M., Roth, D.A., Furie, B.C., and Furie, B. (2008). Crystal structure of human factor VIII: implications for the formation of the factor IXa-factor VIIIa complex. Structure *16*, 597-606.

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

Ngo, J.C.K., Gullingsrud, J., Giang, K., Yeh, M.J., Fu, X.D., Adams, J.A., McCammon, J.A., and Ghosh, G. (2007b). SR protein kinase 1 is resilient to inactivation. Structure *15*, 123-133.

Nieratschker, V., Schubert, A., Jauch, M., Bock, N., Bucher, D., Dippacher, S., Krohne, G., Asan, E., Buchner, S., and Buchner, E. (2009). Bruchpilot in ribbon-like axonal agglomerates, behavioral defects, and early death in SRPK79D kinase mutants of Drosophila. PLoS Genet *5*, e1000700.

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

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

Ohren, J.F., Chen, H.F., Pavlovsky, A., Whitehead, C., Zhang, E.L., Kuffa, P., Yan, C.H., McConnell, P., Spessard, C., Banotai, C., Mueller, W.T., Delaney, A., Omer, C., Sebolt, L., Dudley, D.T., Leung, I.K., Flamme, C., Warmus, J., Kaufman, M., Barrett, S., Tecle, H., Hasemann, C.A. (2004). Structures of human MAP kinase kinase 1 (MEK1) and MEK2 describe novel noncompetitive kinase inhibition. Nat Struct Mol Biol *11*, 1192-1197.

Oltersdorf, T., Elmore, S.W., Shoemaker, A.R., Armstrong, R.C., Augeri, D.J., Belli, B.A., Bruncko, M., Deckwerth, T.L., Dinges, J., Hajduk, P.J., Joseph, M.K., Kitada, S., Korsmeyer, S.J., Kunzer, A.R., Letai, A., Li, C., Mitten, M.J., Nettesheim, D.G., Ng, S., Nimmer, P.M., O'Connor, J.M., Oleksijew, A., Petros, A.M., Reed, J.C., Shen, W., Tahir, S.K., Thompson, C.B., Tomaselli, K.J., Wang, B., Wendt, M.D., Zhang, H., Fesik, S.W., Fesik, S.W., Rosenberg, S.H. (2005). An inhibitor of Bcl-2 family proteins induces regression of solid tumours. Nature *435*, 677-681.

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

Pollard, V.W., and Malim, M.H. (1998). The HIV-1 Rev protein. Annu Rev Microbiol *52*, 491-532.

Purcell, D.F., and Martin, M.A. (1993). Alternative splicing of human immunodeficiency virus type 1 mRNA modulates viral protein expression, replication, and infectivity. J Virol *67*, 6365-6378.

Richman, D.D., Morton, S.C., Wrin, T., Hellmann, N., Berry, S., Shapiro, M.F., and Bozzette, S.A. (2004). The prevalence of antiretroviral drug resistance in the United States. AIDS *18*, 1393-1401.

Ropers, D., Ayadi, L., Gattoni, R., Jacquenet, S., Damier, L., Branlant, C., and Stevenin, J. (2004). Differential effects of the SR proteins 9G8, SC35, ASF/SF2, and SRp40 on the utilization of the A1 to A5 splicing sites of HIV-1 RNA. J Biol Chem *279*, 29963-29973.

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

Sakashita, E., and Endo, H. (2010). SR and SR-related proteins redistribute to segregated fibrillar components of nucleoli in a response to DNA damage. Nucleus *1*, 367-380.

Sharp, P.A. (1994). Split genes and RNA splicing. Cell *77*, 805-815.

Sonza, S., Mutimer, H.P., O'Brien, K., Ellery, P., Howard, J.L., Axelrod, J.H., Deacon, N.J., Crowe, S.M., and Purcell, D.F. (2002). Selectively reduced tat mRNA heralds the decline in productive human immunodeficiency virus type 1 infection in monocyte-derived macrophages. J Virol *76*, 12611-12621.

Staffa, A., and Cochrane, A. (1994). The Tat/Rev Intron of Human-Immunodeficiency-Virus Type-1 Is Inefficiently Spliced Because of Suboptimal Signals in the 3' Splice-Site. Journal of Virology *68*, 3071-3079.

Stoltzfus, C.M., and Madsen, J.M. (2006). Role of viral splicing elements and cellular RNA binding proteins in regulation of HIV-1 alternative RNA splicing. Curr HIV Res *4*, 43-55.

Stout, T.J., Foster, P.G., and Matthews, D.J. (2004). High-throughput structural biology in drug discovery: Protein kinases. Curr Pharm Design *10*, 1069-1082.

Tange, T.O., Damgaard, C.K., Guth, S., Valcarcel, J., and Kjems, J. (2001). The hnRNP A1 protein regulates HIV-1 tat splicing via a novel intron silencer element. EMBO J *20*, 5748-5758.

Teigelkamp, S., Mundt, C., Achsel, T., Will, C.L., and Luhrmann, R. (1997). The human U5 snRNP-specific 100-kD protein is an RS domain-containing, putative RNA helicase with significant homology to the yeast splicing factor Prp28p. Rna *3*, 1313-1326.

Wang, H.Y., Arden, K.C., Bermingham, J.R., Jr., Viars, C.S., Lin, W., Boyer, A.D., and Fu, X.D. (1999). Localization of serine kinases, SRPK1 (SFRSK1) and SRPK2 (SFRSK2), specific for the SR family of splicing factors in mouse and human chromosomes. Genomics *57*, 310-315.

Wu, J.Y., and Maniatis, T. (1993). Specific Interactions between Proteins Implicated in Splice-Site Selection and Regulated Alternative Splicing. Cell *75*, 1061-1070.

Wyatt, R., and Sodroski, J. (1998). The HIV-1 envelope glycoproteins: fusogens, antigens, and immunogens. Science *280*, 1884-1888.

Xiao, S.H., and Manley, J.L. (1998). Phosphorylation-dephosphorylation differentially affects activities of splicing factor ASF/SF2. EMBO J *17*, 6359-6367.

Xu, Y.J., Yu, W.M., Xiong, Y.Z., Xie, H.T., Ren, Z.Q., Xu, D.Q., Lei, M.G., Zuo, B., and Feng, X.T. (2011). Molecular characterization and expression patterns of serine/arginine-rich specific kinase 3 (SPRK3) in porcine skeletal muscle. Molecular Biology Reports *38*, 2903-2909.

Zahler, A.M., Damgaard, C.K., Kjems, J., and Caputi, M. (2004). SC35 and heterogeneous nuclear ribonucleoprotein A/B proteins bind to a juxtaposed exonic splicing enhancer/exonic splicing silencer element to regulate HIV-1 tat exon 2 splicing.

J Biol Chem *279*, 10077-10084.

Zahler, A.M., Lane, W.S., Stolk, J.A., and Roth, M.B. (1992). SR proteins: a conserved family of pre-mRNA splicing factors. Genes Dev *6*, 837-847.

Zhang, J.M., Yang, P.L., and Gray, N.S. (2009). Targeting cancer with small molecule kinase inhibitors. Nat Rev Cancer *9*, 28-39.

Zheng, Y.H., Lovsin, N., and Peterlin, B.M. (2005). Newly identified host factors modulate HIV replication. Immunol Lett *97*, 225-234.

Zhong, S.J., Macias, A.T., and MacKerell, A.D. (2007). Computational identification of inhibitors of protein-protein interactions. Current Topics in Medicinal Chemistry *7*, 63-82.

Zhu, J., Mayeda, A., and Krainer, A.R. (2001). Exon identity established through differential antagonism between exonic splicing silencer-bound hnRNP A1 and enhancer-bound SR proteins. Molecular Cell *8*, 1351-1361.