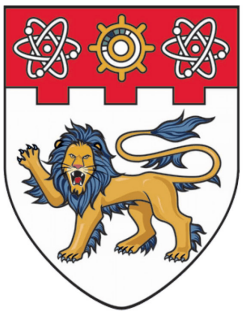


NANYANG TECHNOLOGICAL UNIVERSITY
SINGAPORE

STRUCTURAL CHARACTERIZATION OF VACCINIA-RELATED KINASE 1 (VRK1), A HISTONE MITOTIC KINASE

NGOW YEEN SHIAN

SCHOOL OF BIOLOGICAL SCIENCES

2019

STRUCTURAL CHARACTERIZATION OF VACCINIA-RELATED KINASE 1 (VRK1), A HISTONE MITOTIC KINASE

NGOW YEEN SHIAN

School of Biological Sciences

A thesis submitted to the Nanyang Technological University in partial fulfilment of the requirement for the degree of Doctor of Philosophy

2019

Statement of Originality

I hereby certify that the work embodied in this thesis is the result of original research done by me except where otherwise stated in this thesis. The thesis work has not been submitted for a degree or professional qualification to any other university or institution. I declare that this thesis is written by myself and is free of plagiarism and of sufficient grammatical clarity to be examined. I confirm that the investigations were conducted in accord with the ethics policies and integrity standards of Nanyang Technological University and that the research data are presented honestly and without prejudice.

01 Feb 2019

.....
Date

.....
NGOW YEEN SHIAN

Supervisor Declaration Statement

I have reviewed the content and presentation style of this thesis and declare it of sufficient grammatical clarity to be examined. To the best of my knowledge, the thesis is free of plagiarism and the research and writing are those of the candidate's except as acknowledged in the Author Attribution Statement. I confirm that the investigations were conducted in accord with the ethics policies and integrity standards of Nanyang Technological University and that the research data are presented honestly and without prejudice.

01 Feb 2019

.....
Date

.....
PROFESSOR YOON HO SUP

Authorship Attribution Statement

This thesis contains material from 1 paper published in the following peer-reviewed journal where I was the first co-author.

Chapter 3 is published as NGOW, Y. S., RAJAN, S., YE, H. & YOON, H. S. 2018. Crystal structure of human vaccinia-related kinase 1 in complex with AMP-PNP, a non-hydrolyzable ATP analog. Protein Sci. DOI: 10.1002/pro.3552

The contributions of the co-authors are as follows:

- Professor Yoon Ho Sup provided the initial project direction, guidance and corrected the manuscript drafts.
- I prepared the manuscript drafts together with Dr Sreekanth Rajan.
- I co-designed the study with Professor Yoon and Dr Sreekanth and performed all the laboratory work at the School of Biological Sciences.
- All sample preparation and NMR-TROSY data analysis were conducted by me.
- Dr Sreekanth Rajan collected the X-ray crystallography diffraction data.
- Dr Ye Hong collected the NMR-TROSY experiment spectra data.
- Dr Sreekanth Rajan and myself carried out the data processing and interpretation.

01 Feb 2019

....

Date NGOW YEEN SHIAN

Acknowledgements

First, and most of all, I am hugely indebted to Professor Yoon Ho Sup for the opportunity to work on this project under his supervision. I am very grateful for his insightful advice, patience and encouragement throughout the period of my project.

I am deeply grateful and thankful to Dr Sreekanth Rajan for his time and effort in the collection of the crystal diffraction data, helping me in the crystallographic study of VRK1 with ATP analog, and proofreading my thesis and giving me lots of constructive suggestions. I would like to express my sincere thanks and gratitude to Dr Ye Hong for her time and help with performing NMR experiments.

I am very thankful to Professor Lars Nordenskiold for his kindness and generosity for providing the constructs for nucleosome studies and the opportunity to learn from his laboratory members. I would like to specially express my gratitude to Dr Chen Qinming for her the time and guidance on the reconstitution of nucleosome core complex.

I am grateful to have such a great opportunity to learn and work with a group of excellent researchers from different cultural backgrounds. I would like to express my deepest thanks and gratitude to Dr Serap Beldar, Ms Toh Hui Ting, Mr Jonathan Gan and Mr Yoo Jun Yeob for their time and effort in proofreading my thesis drafts and giving me positive comments. My sincere appreciation goes to them and my other group members, Dr Choi Minjoo, Dr Harikishore, Dr Ajit Prakash, Dr Nguyen Quoc Toan and Dr Shin Joon for their kind support and assistance throughout this research period.

Lastly, I would like to thank my family members for their unconditional love, endless support and tolerance over the last five years, in my pursuit of this Ph.D study.

Tables of Contents

Lists of Figures

List of Tables

List of Abbreviations

Abbreviation	Full name
6xHis	Hexa-Histidine
6XHis-SUMO	Hexa-Histidine-SUMO
AMP-PNP	Adenylyl-imidodiphosphate
APS	Ammonium Persulfate
ATP	Adenosine Triphosphate
BAF	Barrier-to-autointegration factor
BME	2-mercaptoethanol
bp	basepair
cAMP	Cyclic Adenosine Monophosphate
Camp	Chloramphenicol
cAPK	cAMP-dependent protein kinase
Car	Carbenicillin
CIA	Chloroform: Isoamyl alcohol
CK	Casein Kinase
cryst	Crystal
CV	Column Volume
DDR	DNA Damage Response
DNA	Deoxyribonucleic acid
DNA-PAGE	Deoxyribonucleic acid-Polyacrylamide Gel Electrophoresis
dNTP	Deoxynucleotide
DTT	Dithiothreitol
DMSO	Dimethyl Sulfoxide
EDTA	Ethylenediaminetetraacetic acid
FPLC	Fast Protein Liquid Chromatography
G1	Gap1
G2	Gap2
HO	Histone Octamer
IPTG	Isopropyl β-D-1-thiogalactopyranoside
kb	kilobase
LB	Luria-Bertani
M	Mitotic
M.W.	Molecular Weight
NCP	Nucleosome Core Particle
NEB	New England BioLabs
NMR	Nuclear Magnetic Resonance
PCR	Polymerase Chain Reaction
PDB	Protein Data Bank
PEG	Polyethylene glycol
PES	Polyethersulfone

PMSF	Phenylmethylsulfonyl Fluoride
PTM	Post-translational Modification
RMSD	Root-Mean-Square Deviation
RNA	Ribonucleic acid
S	Synthesis
SDS	Sodium Dodecyl Sulfate
SDS-PAGE	Sodium Dodecyl Sulfate - Polyacrylamide Gel Electrophoresis
SEC	Size Exclusion Chromatography
SER	Surface Entropy Reduction
SP	SUMO Protease
TB	Terrific Broth
TBE	Tris-Borate-EDTA
TBS	Tris-buffered Saline
TBS-T	Tris-buffered Saline – Tween 20
TE	Tris- Ethylenediaminetetraacetic acid
TROSY	Transverse Relaxation Optimized Spectroscopy
TY	Tryptone-Yeast
VRK	Vaccinia-related kinase
WB	Western Blot
wt	wild-type

Abstract

The human vaccinia-related kinase 1 (VRK1), is a mitotic kinase, involved in cell division regulation and carcinogenesis. VRK1 phosphorylates the histone H3 at Thr3 and Ser10, which is a critical event for initiating chromosome condensation during mitosis. The deregulation of VRK1 can affect chromatin architecture and remodeling during mitosis, which could result in aneuploidy and genome instability, leading to cancer. Recent studies have shown that high levels of VRK1 is directly correlated to poor clinical outcomes in breast cancer patients, apart from colorectal, prostate and lung cancers among others. Hence, inhibiting VRK1 is considered a viable approach toward developing cancer drugs. VRK1 can be targeted by ATP-competitive inhibitors, commonly known as Type I inhibitors, as it possesses a conserved ATP-binding domain. Currently, there are ~37 FDA-approved kinase inhibitors available, and most of them belong to this category, implying the importance of kinase inhibitor discovery and at the same time posing a challenge toward designing specific kinase inhibitors. Therefore, a three-dimensional portrait of VRK1 with its substrate will provide structural insights into developing specific inhibitors for VRK1 which lacks cross-reactivity with other kinases. To-date, the structural basis of VRK1-ATP interaction is not available, hindering the progress toward designing VRK1 specific Type I inhibitors. Moreover, the molecular basis of histone H3 substrate recognition on VRK1 remain poorly understood, which is also impeding our efforts in developing alternative inhibitors against VRK1. Thus, in this thesis, I have focused on the structural characterization of (a) ATP-binding to VRK1 and (b) VRK1's interaction with the nucleosome, for the understanding of VRK1-mediated histone H3 regulation. In this direction, first the crystallographic

structure of VRK1 in complex with AMP-PNP was determined at a resolution of 2.07Å, revealing several key residues such as Asp132, Phe134, Met131, Ser181 and Val196 of VRK1 important for ATP interaction, supported by NMR studies. A structural comparison with the inhibitor-bound VRK1, its paralogs and other mitotic kinases enabled us to identify hotspots to design more specific ATP-competitive VRK1 inhibitor. In the second part of thesis, the molecular interaction of VRK1 with histone H3 was characterized. To this end, attempts were made to obtain nucleosome core particle (NCP) in complex with VRK1. While our preliminary studies using mobility shift electrophoretic assay indicated complex formation, nonetheless, obtaining the homogenous NCP-VRK1 complex remained a challenge. To achieve this purpose, further optimization is required to unravel the VRK1 and histone H3 interaction mechanism, which together with the ATP binding studies, can offer more strategies for the development of novel VRK1 inhibitors with better specificity and potency in the future.

Chapter 1

Introduction

1.1 Cell signaling regulation and Protein Phosphorylation

In multicellular organisms, cells communicate and transmit biological information to regulate activities that are essential for growth and development, differentiation, immunity, apoptosis and so on. Cells naturally receive signals from extracellular environment and relay them into the cells via various signaling molecules that trigger signal transduction pathways (Alberts et al., 2002). These pathways are tightly regulated by changes in the signaling proteins such as protein amount, their localization, interacting partners and post-translational modifications (Lee and Yaffe, 2016). Post-translational modifications (PTMs) are the covalent addition of chemical groups to the amino acids of proteins after translation. These modifications specifically alter the structure, localization, activity and protein-protein interactions network of the target proteins, thereby regulating signal transduction events (Karve and Cheema, 2011). Among the post-translational modifications, protein phosphorylation, the addition of a phosphate group to the sidechain of amino acids such as serine, threonine and tyrosine, is the most common mechanism which causes functional and conformational changes on the target proteins (Manning et al., 2002). Protein phosphorylation and dephosphorylation are reversible processes regulated by kinases and phosphatases respectively. Phosphorylation can alter the functional status of a protein (active or inactive) via structural conformational changes that mainly initiates the phosphorylation of many other proteins in a signaling cascade, which significantly affects many cellular activities in prokaryotes and eukaryotes (Cheng et al., 2011).

In the human genome, there are over 500 protein kinases, constituting nearly 1.7% of the total human genes (Manning et al., 2002). These protein kinases are

classified into seven major families such as (i) Protein kinase A, G and C (AGC); (ii) CDK, MAPK, GSK3 and CLK protein kinases (CMGC); (iii) Calmodulin/Calcium regulated kinases (CAMK); (iv) Casein Kinase 1 (CK1); (v) Sterile serine/threonine kinases (STE); (vi) Tyrosine kinase-like kinases (TKL) and (vii) Tyrosine kinases (TK), based on sequence homology analysis (Hanks & Hunter, 1995; Manning et al., 2002). As the function of a protein directly related with its protein sequence and three-dimensional structure, the enzymes within the same family possess similar substrate specificity and regulation mechanism (Manning et al., 2002). The regulatory roles of protein kinases in many cellular activities are important for survival, and mutations or deregulation of protein kinases can result in genetic disorders or diseases such as cancer. Hence, the understanding of the catalytic domain structure of the eukaryotic protein kinases can help in the development of treatment with high specificity and potency.

Subsequent sections of this thesis will introduce and discuss the structural and functional features of kinase catalytic subunits with a focus on human vaccinia-related kinase 1 (VRK1).

1.2 The structure of conserved catalytic subunit of kinase

Protein kinases may differ in sizes, substrate recognition, specificity and regulatory mechanism, yet they share a conserved catalytic domain. Based on the sequence alignment analysis of many protein kinases, the kinase domain can be divided into 12 subdomains (indicated by Roman numerals), with some subdomains containing the important conserved residues (Hanks and Hunter, 1995) (Figure 1-1A). The crystal structure of cAMP-dependent protein kinase

(cAPK) solved at a resolution of 2.2 Å (PDB ID 1ATP) (Zheng et al., 1993) became a scaffold for mapping the subdomains determined earlier (Hanks and Hunter, 1995) onto the structure, which greatly aided in the understanding of the contribution of these conserved motifs to the kinase ATP-binding site.

The cAPK is folded into a two-lobe structure, where the N-terminal lobe consists of five-stranded β-sheet and a conserved αC helix, whereas the C-terminal lobe are mainly helices and a β-sheet (Figure 1-1B). The I-IV subdomains located in the N-terminal lobe are important for binding and positioning of the ATP while VI-XI subdomains in the C-terminal lobe are involved in substrate binding and catalyzing the transfer of phosphate group from ATP to substrate (Hanks and Hunter, 1995, Taylor and Kornev, 2011). The two lobes are connected by a loop, known as the hinge in the kinase, and the backbone of the residues in the hinge region usually make contacts with the adenine ring of ATP (Tanramluk et al., 2009). In the subdomain I, the first consensus motif (GxGxxGxV), is a Glycine-rich loop (G-loop) linking the strand β1 to β2 (Figure 1-1B). This loop interacts with the γ-phosphate group of ATP and orient it in place for catalysis. For this reason, it is also named as the phosphate binding loop, P-loop. The G-loop is the most flexible region in the N-terminal lobe (Patel and Doerksen, 2010), and the tip of the loop is observed as a closed conformation when its backbone attached to the γ-phosphate of ATP (Madhusudan et al., 2002). Generally, they adopt an 'Open-conformation' and 'Closed conformation' before and after binding the ATP, respectively to stabilize the phosphate groups. The conserved residue Lys72 (cAPK numbering) present in β3 strand (subdomain II) interacts with α- and β- phosphates to position and stabilize the ATP molecule. Moreover, the conserved residue Glu91 located in αC-helix (subdomain III) forms a salt bridge

with Lys72 and aids in stabilizing the interaction between Lys72 and the phosphate groups of ATP (Hanks and Hunter, 1995). The Lys72 and Glu91 salt bridge is highly conserved in the catalytic domain and is a distinct characteristic of an active kinase. In subdomain VIB, the conserved motif located on the catalytic loop (HRDLKxxN) consists of important interacting residues. During phosphotransfer, Asp166 on this loop can act as a catalytic base to accept the proton donated by the substrate while Lys168 can neutralize the negative charge on the γ-phosphate. The Asn171, together with Asp184 (on the DFG motif in subdomain VII), can interact with the phosphate groups of ATP through coordination with manganese ions, which aids in the positioning and stabilizing ATP for catalysis (Hanks and Hunter, 1995, Taylor and Kornev, 2011, Endicott et al., 2012). Magnesium ions are also known to substitute manganese, to provide similar ionic interactions. The activation segment is the region extended from the DFG motif to the APE motif. The consensus APE motif in subdomain VIII has a glutamine residue that can form an ion pair with the conserved arginine in subdomain XI to stabilize the C-terminal lobe (Hanks and Hunter, 1995). The key motifs of the catalytic domain discussed are highly conserved in most of the eukaryotic protein kinases. The presence of these signature motifs can be used to categorize a kinase as an Active-kinase, else they are coined as Inactive or Pseudo kinase.

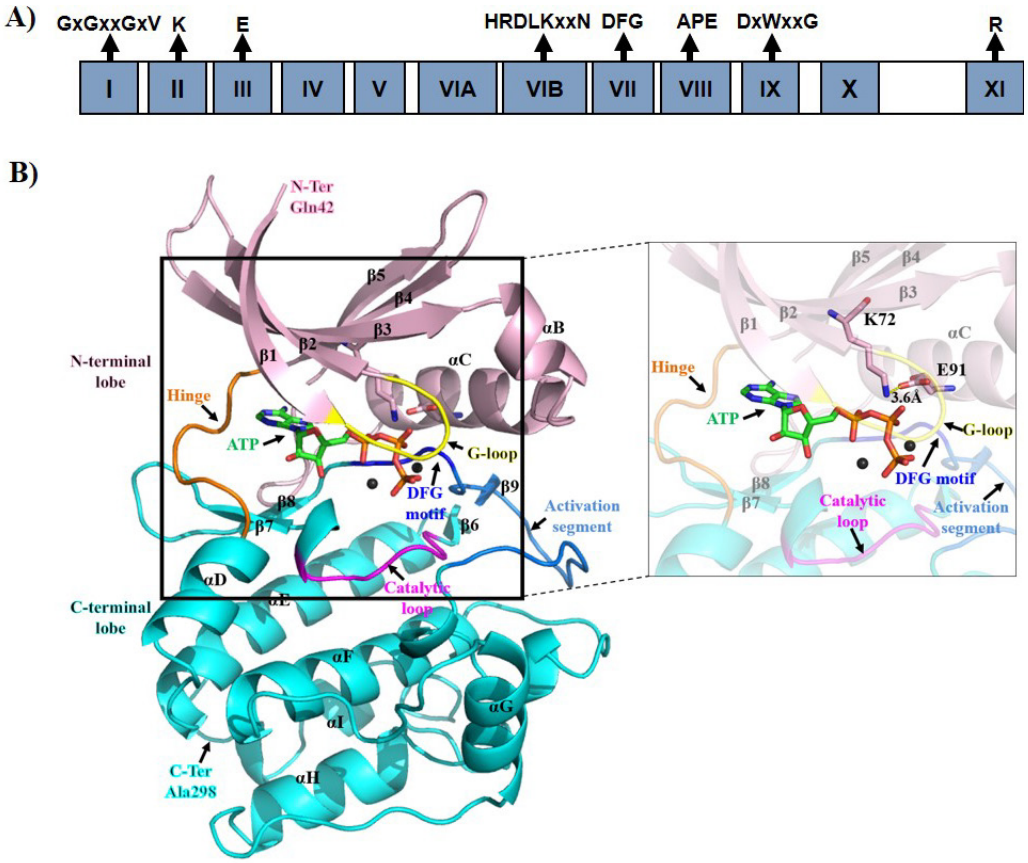


Figure 1-1 Illustration of the catalytic subunit of a protein kinase.

A) 12 subdomains of the protein kinase and the conserved motifs are marked by black arrow. B) Cartoon representation of the cAMP-dependent kinase structure (PDB ID 1ATP). The five β-sheet, αB and αC helix in the N-terminal lobe are colored in pink while the helices and β-sheet in the C-terminal lobe are colored in cyan. The secondary structure elements are labeled for reference. The ATP is shown in stick mode while the manganese (Mn^{2+}) ions are represented by the black sphere. The zoomed-in view of the kinase's catalytic domain is shown on the right panel. The inset shows the interaction of the G-loop (in yellow), the hinge (in orange), the DFG motif (in blue), K72 and E91 (in sticks) with the ATP and Mn^{2+} ions. The K72-E91 forms a salt bridge at 3.6Å (shown as a yellow dotted lines). The catalytic loop and activation segment are highlighted in magenta and light blue.

1.3 Mitotic regulation by kinases

Cell division is a cycle of events in which a parent cell is split into two identical daughter cells. This is a vital process as it ensures continuous cell proliferation and tissue repair in multicellular organisms. The cell cycle consists of two major phases: interphase and mitosis (M phase). During interphase, the chromatin DNA is replicated in the S phase prior to mitosis. During mitosis, chromatins are

condensed to chromosomes, then aligned on the metaphase plate, and later segregated to produce two daughter cells, with each carrying a set of identical genetic materials (Alberts et al., 2002).

The cell cycle is tightly controlled by the DNA damage and metaphase checkpoints (Figure 1-2A). The DNA damage checkpoint ensures that the damaged DNA is repaired before entering mitosis while the metaphase checkpoint makes sure that all chromosomes are attached properly to the mitotic spindle before anaphase to achieve proper segregation (Nigg, 2001). The regulation of cell cycle processes is crucial as errors can result in aneuploidy and instability of genome leading to cancer and other diseases (Ma and Poon, 2011, Kops et al., 2005). Mitotic kinases such as cyclin-dependent kinase 1 (CDK1), Aurora A (AURA), Aurora B (AURB), Polo-like kinase 1 (PLK1), Haspin and VRK1 can phosphorylate substrates involved in the chromosome condensation, segregation, and spindle assembly during cell division to ensure fidelity in the process of transmitting genetic information to the next generation for survival (Nigg, 2001, Ma and Poon, 2011) (Figure 1-2A).

During cell division, the condensation of chromatin DNA into higher order chromosome structure before mitosis is critical for the smooth chromosome movement and correct segregation (Xu et al., 2009, Wei et al., 1999). A chromatin is an array of nucleosomes, in which each nucleosome is made up of the 146-bp DNA wound around the histone octamer (Luger et al., 1997). The histone octamer is composed of 2 sets of histones H2A, H2B, H3 and H4. The nucleosomes are bound together by H1 protein (Zhou et al., 2013) and form a 'beads-on-a-string' structure, which wind up to form the 30-nm chromatin fiber and organizes into higher ordered chromatin structures (Thoma et al., 1979).

Recent studies have questioned the presence of these 30-nm fibers (Maeshima et al., 2010, Eltsov et al., 2008, Tremethick, 2007), identifying highly disordered structures lacking the 30-nm fibers while observing 10-nm fibers (Razin and Gavrilov, 2014, Eltsov et al., 2008). The condensed chromatin is then tightly coiled to produce the chromatid of a chromosome (Annunziato, 2008). PTMs such as phosphorylation, methylation, acetylation and ubiquitination of the histone protein can play a role in gene transcription and DNA damage regulation. Of these PTMs, the phosphorylation of histones is involved in modulating the chromatin into higher order structures during mitosis (Xu et al., 2009, Wang and Higgins, 2013).

The phosphorylation on the histone H3 tail was shown to affect chromosome condensation regulation during mitosis (Baek, 2011) (Figure 1-2 B). Aurora B phosphorylates histone H3 at Ser10 during the late G2 phase and Ser28 during the prophase (Goto et al., 1999), which later plays an important role in initiating chromosome condensation during mitosis (Goto et al., 2002). On the other hand, the phosphorylation of histone H3 at Thr3 by Haspin is essential for the proper alignment of chromosome on the metaphase plate during mitosis and meiosis (Dai et al., 2005). Recently, VRK1 was discovered as another chromatin mitotic kinase that can phosphorylate both Thr3 and Ser10 of histone H3 that also aid in the chromatin condensation during cell division (Kang et al., 2007). This will be further discussed in section 1.5.1.

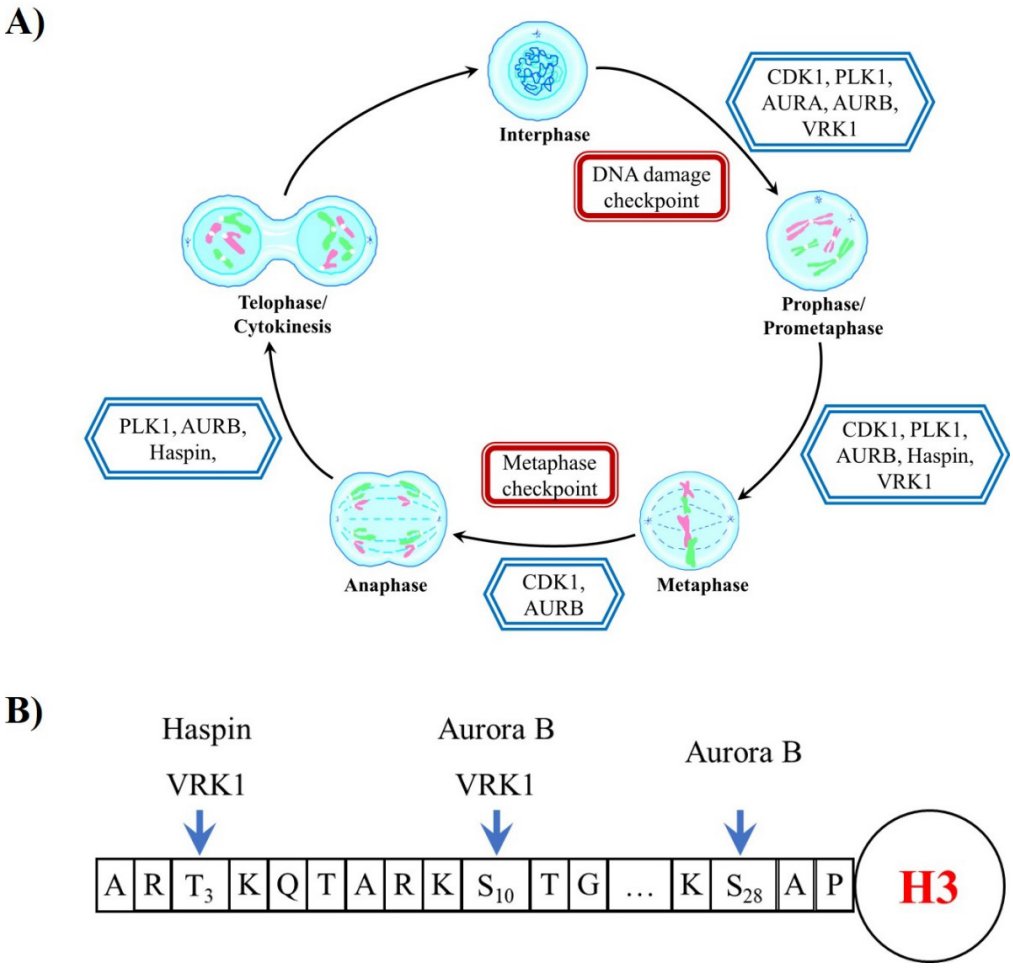


Figure 1-2 Illustration on the regulation of cell cycle events.

A) Many mitotic kinases have more than one role during cell division (blue box), and the cell cycle is controlled by checkpoints (red box) that ensures the fidelity of genetic material transmission to next generation. B) Diagram showing the phosphorylation of histone H3 tail by mitotic kinases to regulate chromosome condensation during mitosis.

1.4 Vaccinia-related kinase 1 and its family

In the human kinome, there is a group of Ser/Thr kinase that branched off early from the casein kinase I (CKI) family, known as the Vaccinia-related kinase (VRK) (Manning et al., 2002). Vaccinia-related kinase family members were discovered to have close sequence homology to vaccinia virus B1 kinase (vvB1) (Nezu et al., 1997). In mammals, there are three paralogs, VRK1, VRK2 and VRK3 present in the family. When comparing the amino acids sequence of vvB1

with the Homo sapiens VRK family, it was observed that the vvB1 shows an overall identity of 39%, 34% and 27% with the human VRK1, VRK2 and VRK3, respectively. In addition, the comparison among the paralogs showed that human VRK1 shares a 44% and 33% sequence identity with human VRK2 and VRK3, respectively, while human VRK2 shows a sequence identity of 23% with human VRK3 (Nichols and Traktman, 2004). Hereafter, the VRK1, VRK2 and VRK3 refers to those belonging to Homo sapiens.

VRK1 is highly expressed in tissues with high proliferation rates like the testis, thymus and fetal liver. On the other hand, VRK2 is highly expressed in tissues such as skeletal muscles, heart, fetal liver and pancreas, while VRK3 is expressed in almost all human tissues (Nezu et al., 1997). VRK1 has a protein sequence length of 396 amino acids where its N-terminal consists of a kinase domain and the C-terminal contains a nuclear localization signal (NLS) (Figure 1-3), hence it usually localizes within the nucleus, although it is also found in the cytosol of some cell lines (Nichols and Traktman, 2004, Valbuena et al., 2007a). VRK2 is spliced into two isoforms, 1 and 2, with 508 and 397 amino acids, respectively. The two proteins have a kinase domain in the N-terminal end but differ at the C-terminal end. VRK2 isoform 1 has a transmembrane domain located at the C-terminal end (Figure 1-3), which enables it to localize at the endoplasmic reticulum (ER) and the nuclear envelope, while isoform 2, despite lacking the NLS region, is still found in the nucleus and cytoplasm (Nichols and Traktman, 2004, Blanco et al., 2006). Lastly, VRK3 is made up of 474 amino acids, and the NLS at its N-terminal enables to localize in the nucleus (Nichols and Traktman, 2004) (Figure 1-3). However, the VRK3 is known as a pseudokinase as it contains an inactive kinase domain (Scheeff et al., 2009).

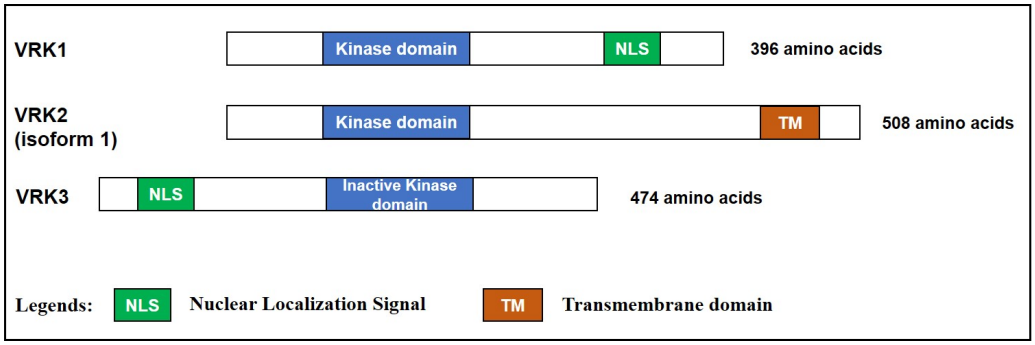


Figure 1-3 Schematic diagram illustrating the domain organization of Vaccinia-related kinase (VRK) family.

The domain organization of VRK1, VRK2 and VRK3 are shown, in which only the isoform 1 of VRK2 is presented here. The kinase domain, nuclear localization signal (NLS), and transmembrane domain are colored in blue, green and orange respectively.

The solution structure of VRK1 (from amino acids 1-361) (PDB ID 2LAV) determined by our lab using nuclear magnetic resonance (NMR) revealed that the protein adopts a bilobate kinase fold, consisting of eight α-helices, nine β-sheets and an extended C-terminal tail (Figure 1-4) (Shin et al., 2011). The signature motifs in the subdomains of protein kinase (Figure 1-5A) (Hanks and Hunter, 1995) are also observed in the structure of VRK1 (Figure 1-4). These signature motifs important for ATP-binding are identified as (i) the P-loop (GxGxxGxV of subdomain I); (ii) the Lys71 in β3 strand and Glu83 in αC helix (conserved lysine and glutamate in subdomain II and III respectively); (iii) DRF motif (the hinge region); (iv) the catalytic loop (HRDLKxxN motif present in subdomain VIB); (v) the DYG motif in the activation loop, conserved in vvB1 kinase and VRK family (Nichols and Traktman, 2004), has a tyrosine replacing the phenylalanine of DFG motif in subdomain VII). The additional alpha helix (αC4 helix), commonly observed in the VRK family, is also seen in the VRK1 structure, in which the key residues Ile96 and Trp99 were observed forming hydrophobic contacts with Tyr167 of αE helix (Shin et al., 2011). Interestingly, VRK1 has a unique C-terminal tail (amino acids 341-361), a non-catalytic domain but can

orient around the catalytic site and interact with the important conserved motifs of VRK1. It was discovered that the removal of the C-terminal tail will result in a reduction of VRK1 autocatalytic activity (Shin et al., 2011) which led to the conclusion that the extended C-terminal tail can regulate the enzymatic activity of VRK1.

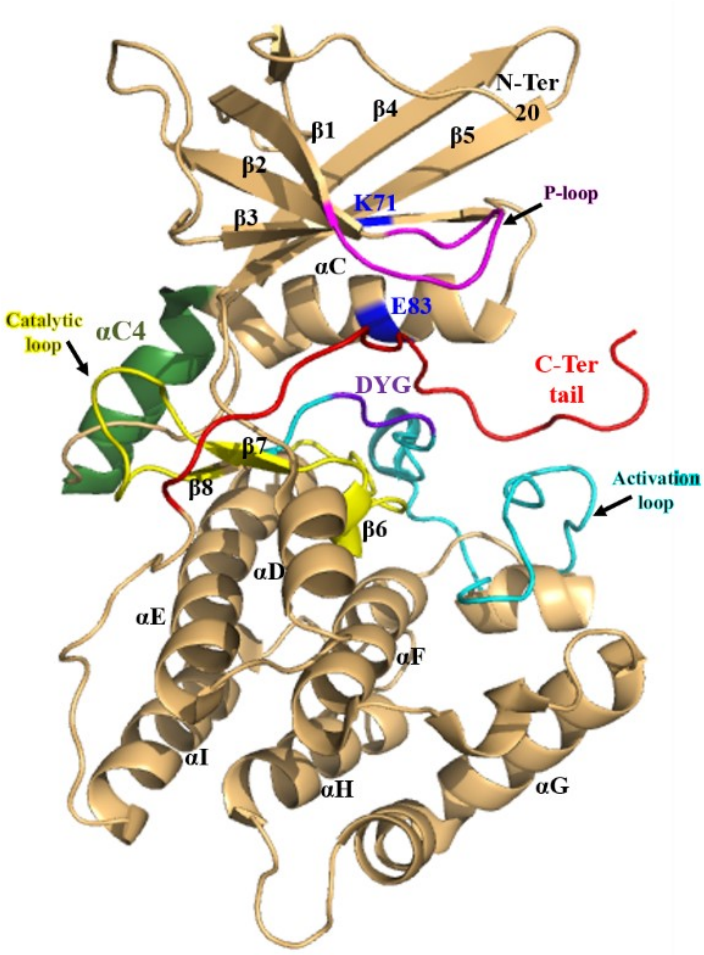


Figure 1-4 The NMR solution structure of VRK1.

Cartoon representation of VRK1 structure (PDB ID 2LAV) in pale orange, showing from amino acids 20-361, and its secondary structure elements labeled for reference. The signature motifs important for ATP-binding are labeled and colored as (i) P-loop (in magenta); (ii) residue K71 and E83 (in blue); (iii) αC4 helix (in green); (iv) catalytic loop (in yellow), (v) the DYG motif (in purple); (vi) activation loop (in cyan) and (vii) extended C-terminal tail (amino acid 341-361) (in red).

Of the VRK family, only the VRK1 and VRK2 are known as the active kinases (Nichols and Traktman, 2004) while VRK3 is characterized as an inactive one evident by the degraded catalytic subunit (Scheeff et al., 2009). The sequence

alignment of VRK1 with VRK2 and VRK3, using Clustal Omega (Sievers et al., 2011) and ESPript (Gouet et al., 2003) (Figure 1-5A), revealed that the signature motifs important for ATP-binding and catalysis are conserved in VRK1 and VRK2. In VRK3, the key residues important for catalytic activity were replaced, resulting in an inactive catalytic domain (Scheeff et al., 2009) (Figure 1-5A). The overall structures of VRK1, VRK2 and VRK3 are very similar (Figure 1-5B), and the superposition of the VRK1 (PDB ID 2LAV) with VRK2 (PDB ID 2V62), and VRK3 (PDB ID 2JII) structures, reveals a Root-Mean-Square Deviation (RMSD) values of 0.99 Å (245 equivalent C^{α} atoms) and 1.55 Å (233 equivalent C^{α} atoms) respectively. The αC4 helix in VRK1 is conserved in both VRK2 and VRK3, but the C-terminal tail (amino acids 341-361) is unique only to VRK1.

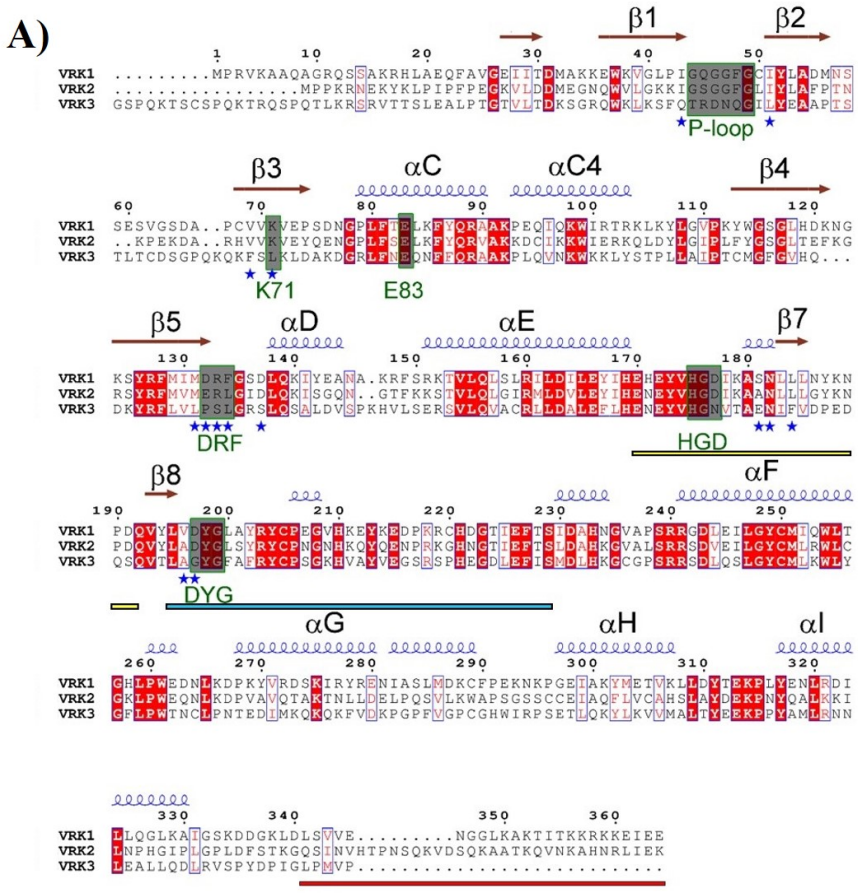


Figure 1-5 The sequence and structure comparison of VRK family members.

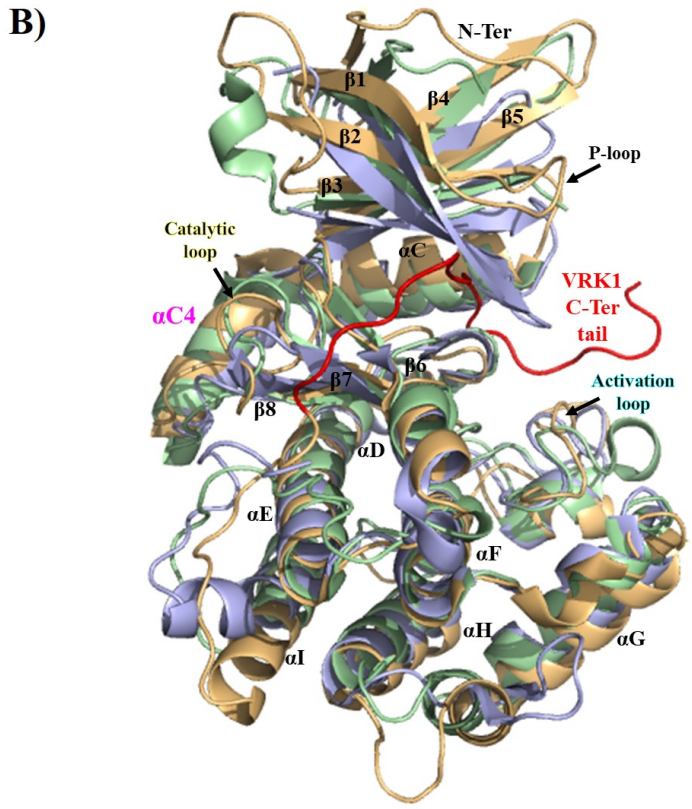


Figure 1-5 (cont.)

A) Sequence alignment of VRK1, 2 and 3 using Clustal omega and ESPript. The signature motifs of a protein kinase are indicated and shown within the grey colored boxes (labeled using VRK1 numbering below). The secondary structure corresponding to VRK1 is shown above the alignment. The ATP binding residues in VRK1 are marked by blue colored asterisk. The catalytic loop, activation loop and C-terminal tail are highlighted in yellow, cyan and red respectively. B) Superposition of VRK1 (PDB ID 2LAV) with VRK2 (PDB ID 2V62) and VRK3 (PDB ID 2JII) shown in cartoon mode. The structure of VRK1, 2 and 3 are colored in pale orange, light green and light purple respectively. The secondary structure elements are numbered according to VRK1. The αC4 helix (indicated in magenta) is unique to the VRK family members. The extended C-terminal tail from amino acids 341-361 (shown in red) is uniquely present in VRK1.

1.5 Regulatory roles of VRK1 in cell division

In the recent years, VRK1 was discovered to be involved in the regulation of biological events of the cell division cycle (Valbuena et al., 2011). Valbuena et al. determined that VRK1, an early response gene similar to *MYC* and *FOS* genes, is necessary for initiating mitotic entry during cell division, and the loss of VRK1 can cause a block in cell cycle progression (Valbuena et al., 2008). Specifically, VRK1 was discovered to have important regulatory roles in (i) chromatin

remodeling during mitosis and (ii) signaling DNA damage response (DDR) for sustaining cell survival. VRK1 can phosphorylate (i) histone H3's N-terminal tail to initiate chromosome condensation during mitosis; (ii) barrier-to-autointegration factor (BAF) to regulate the disassembly and reassembly of nuclear envelope; (iii) DDR proteins such as the H2AX and NBS1 to initiate an early response to DNA damage induced by ionizing radiation; and (iv) p53 protein that can mediate cellular protection through the maintenance of the autoregulatory loop system between VRK1 and p53 (Figure 1-6).

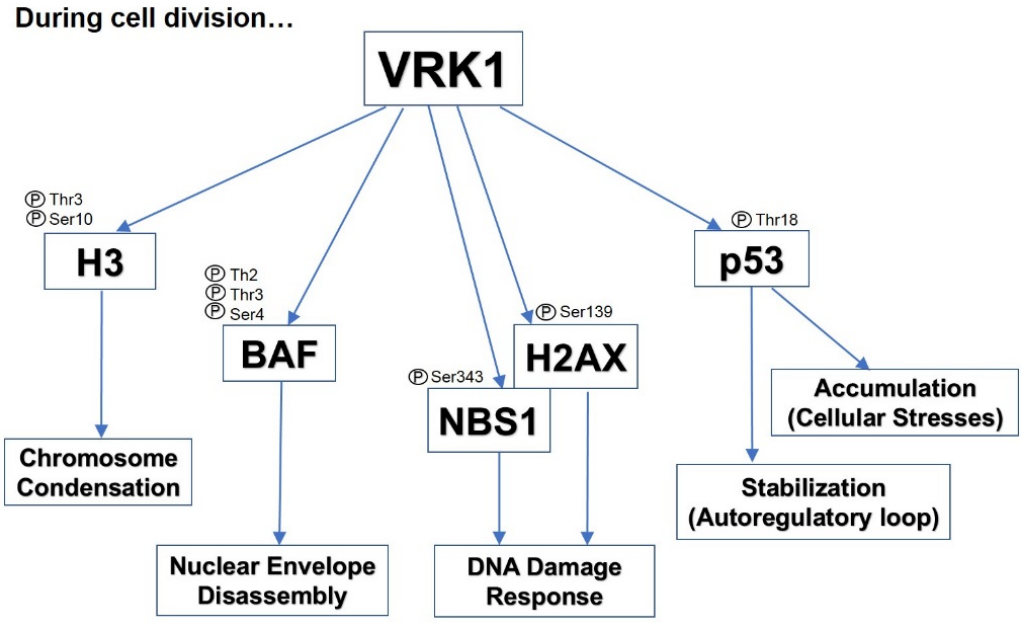


Figure 1-6 Summary of VRK1 regulatory roles during cell division.

VRK1 can phosphorylate (i) Histone H3 at Thr3 and Ser10 which regulates chromosome condensation, (ii) BAF at Thr2/3 and Ser4 which controls nuclear envelope disassembly, (iii) H2AX at Ser139 and NBS1 at Ser343 that is responsible for the early initiation of DNA Damage Response, and (iv) p53 at Thr18 under normal cell conditions to stabilize the protein at its basal level through an autoregulatory loop and accumulates p53 when exposed to cellular stresses resulting in DNA damage.

Previously, our lab discovered the colocalization of VRK1 with the chromatin and showed that the expression and regulatory roles of VRK1 during mitosis is oscillated during cell cycle (Kang et al., 2007) and is controlled by a histone variant, the Macro histone H2A1.2 (MacroH2A1) (Kim et al., 2012). The high

expression of MacroH2A1 during the interphase, was observed to be binding to VRK1 and inhibiting its phosphorylation activity. At the late G2 to M phase, MacroH2A1 protein decreases and releases from VRK1, leading to an increase in VRK1 which allows the regulation of chromosome condensation, nuclear envelope reassembly and DNA damage response signaling during cell cycle progression (Kim et al., 2012).

1.5.1 Chromosome condensation

Chromatin condensation is important for cell division, and the phosphorylation of histone H3 tail at Ser10 is involved in the initiation of chromosome condensation (Van Hooser et al., 1998, Wei et al., 1999). Aurora B can phosphorylate histone H3 at Ser10 during the late G2 phase (Hendzel et al., 1997, Crosio et al., 2002), and phosphorylates Ser28 during the prophase (Goto et al., 1999), which plays important role in initiating chromosome condensation during mitosis (Goto et al., 2002). Interestingly, another phosphorylation at histone H3 Thr3 by Haspin, was revealed to be essential for the proper alignment of chromosome on the metaphase plate during mitosis and meiosis (Dai et al., 2005). Recently VRK1, a chromatin mitotic kinase, was discovered to phosphorylate histone H3 at Thr3 and Ser10, mainly occurring during the transition from late G2 to mitotic (M) phase (Kang et al., 2007). The phosphorylation of VRK1 on H3 is oscillated during cell cycle and is controlled by its negative regulator, MacroH2A1 (Kim et al., 2012). It was discovered that VRK1-mediated H3 phosphorylation during mitosis, plays regulatory roles in chromosome condensation and chromosome alignment at metaphase plate, similar to Aurora B and Haspin, respectively. The loss of VRK1 can cause a downregulation of H3

phosphorylation, which can lead to abnormal chromosome morphology and segregation during mitosis, while the overexpression of VRK1 causes the hypercondensation of nuclei (Kang et al., 2007). Hence, VRK1 together with Aurora B and Haspin, in mediating H3 regulation, are important for chromosome condensation and segregation during mitosis, and errors to either part can result in severe cellular catastrophe.

1.5.2 Nuclear envelope reassembly

In prophase, the disassembly of the nuclear envelope is necessary for chromosome to move and attach to mitotic spindle. BAF, a highly conserved dimeric protein (Zhuang et al., 2014, Cai et al., 1998), can bind to double stranded DNA (Zheng et al., 2000) and the LEM (Lamina-associated-polypeptide 2 (LAP2)/emerin/MAN1) domain (Furukawa, 1999, Mansharamani and Wilson, 2005, Lee et al., 2001). BAF, interacting with chromatin DNA and LEM domain at the nuclear lamina, is critical for nuclear lamina organization and nuclear envelope formation (Haraguchi et al., 2008, Furukawa et al., 2003). During mitosis, VRK1 was discovered to phosphorylate BAF at Ser4 and/or Thr2/3 (Nichols et al., 2006) and thereby regulates the nuclear envelope disassembly through the disruption of BAF interaction with DNA and LEM domain (Gorjánácz et al., 2007).

During interphase, the BAF dimers bind to chromatin DNA and the LEM domain located mainly on the inner nuclear membrane (Gorjánácz et al., 2007), while VRK1 located evenly in the nucleoplasm is bound to and inhibited by MacroH2A1 (Kim et al., 2012). During mitotic entry, MacroH2A1 is degraded, and the activated VRK1 can travel from the nucleoplasm to the nuclear periphery,

to phosphorylate BAF dimers which releases it from interacting partners (Gorjánácz et al., 2007). As a result, the nuclear envelope disassembles and BAF is uniformly distributed in the cytoplasm. During mitotic exit, the enzymatic activity of VRK1 is inhibited by LEM-4 protein and LEM-4, which can also bind to the protein phosphatase 2A (PP2A) complex to dephosphorylate BAF dimers (Asencio et al., 2012). The BAF dimers can again bind to the DNA and LEM domain to reassemble the nuclear envelope. Therefore, VRK1 plays an essential role in maintaining the nuclear envelope morphology through its interaction with BAF and LEM4 (Gorjanacz, 2013). Studies have shown that the absence of VRK1 results in, the reduction of BAF phosphorylation which can cause abnormal nuclear envelope morphology and the BAF detained on the chromosome can affect the fidelity of the genome (Molitor and Traktman, 2014).

1.5.3 DNA damage response

Cells exposed to extrinsic stresses such as ionizing radiation or ultraviolet light (UV) can result in DNA damage, where the accumulation of damaged DNA can cause genome instability leading to cancer and other diseases. Hence, it is important for the cells to sense DNA damage and later trigger a cascade of signaling pathways that leads to downstream cellular events such as cell cycle arrest, DNA repair or apoptosis. This mechanism is termed as the DNA damage response (DDR) pathway (Wan et al., 2014, Jackson and Bartek, 2009). The DNA Damage Response (DDR) signaling pathway consists of signal sensors, transducers, and effectors. The sensors are proteins that can recognize DNA damage caused by UV or ionizing radiation (Marechal and Zou, 2013). The transducers of DDR pathway are kinases such as Ataxia Telangiectasia Mutated

(ATM), VRK1 and their downstream kinases. The DNA double-stranded breaks (DSBs) induced by ionizing radiation activates ATM (Marechal and Zou, 2013) or VRK1 (Sanz-Garcia et al., 2012, Monsalve et al., 2016), thereby phosphorylating H2AX or NSB1, which are proteins involved in initiating DNA repair or apoptosis. VRK1 was discovered to phosphorylate H2AX at Ser139, similar to ATM, and the formation of γH2AX foci at the DNA damaged site, can bind other DDR sensors such as NBS1 (Kobayashi et al., 2002) and recruits ATM, ataxia telangiectasia and Rad3 related (ATR) and DNA-dependent protein kinase (DNA-PK) for downstream cellular events such as cell cycle arrest, DNA repair or apoptosis (Marechal and Zou, 2013, Wan et al., 2014, Salzano et al., 2015). Later, VRK1 was discovered to phosphorylate NBS1 at Ser343 to initiate the formation of NBS1 foci in response to DNA damage (Monsalve et al., 2016). The phosphorylation activity of VRK1 on H2AX or NBS1 can function by itself or cooperate with the ATM. The loss of VRK1 was observed to prevent the formation of γH2AX or NBS1 foci at DNA damaged regions in ionizing radiation induced cells, which is important for early DDR (Monsalve et al., 2016, Salzano et al., 2015). The absence of the foci formation can prevent downstream DDR pathway for DNA repair or apoptosis. Hence, the presence of VRK1, and its roles in phosphorylating these important DDR sensors is critical for DNA repair mechanisms.

1.5.4 Cellular Protection

p53, a tumor suppressor protein, is activated upon DNA damage or other cellular stresses, which is involved in the regulation of major cellular events such as DNA repair (Williams and Schumacher, 2016), G1 arrest (Pellegata et al., 1996,

Kuerbitz et al., 1992) or apoptosis (Schuler et al., 2000). When cells are under stress-free condition, the amount and stability of p53 are tightly controlled by the interactions with its negative regulator hdm2/Mdm2. Mdm2 binds to p53 and causes the ubiquitination of p53 which leads to degradation by the proteasome (Kubbutat et al., 1997, Moll and Petrenko, 2003). Recently, VRK1 was discovered to be involved in the stabilization of p53 in stress free cells, by phosphorylating p53 at Thr18 and prevents the downregulation of p53 by Mdm2 (Vega et al., 2004). Moreover, it was discovered that p53 and VRK1 forms a stable complex at basal levels, which can act as an early response system for detecting cellular stresses and get activated by DNA damage induced by UV light (Lopez-Sanchez et al., 2014). During DNA damage, VRK1 phosphorylating p53 at Thr18, inhibits p53 and Mdm2 interaction and contributes to the recruitment of p300 which causes the accumulation of p53 for transcription activities important for DNA repair or apoptosis (Vega et al., 2004). After cellular repair, the activated p53 has to return back to the basal level and it was discovered that p53 can regulate its activator, VRK1, by developing an autoregulatory loop, important for cellular maintenance (Valbuena et al., 2006). The high levels of p53 will inhibit VRK1-mediated p53 phosphorylation (Valbuena et al., 2006), and the presence of the unphosphorylated Thr18 can enhance the binding between p53 and Mdm2 and upregulates the degradation of p53 (Vega et al., 2004). Interestingly, p53 mutation in tumor cells was observed to disrupt the autoregulatory loops formed between VRK1 and p53, resulting in a drastic increase in VRK1 level, such as in human lung cancer cells (Valbuena et al., 2007b). Therefore, VRK1-mediated p53 regulation is important for promoting cellular protection during external stresses.

In these VRK1-mediated phosphorylation activities that occur during cell division, VRK1 was observed to interact with the chromatin DNA in close proximity to maintain genomic stability and integrity. Therefore, it is important to control and stabilize the amount of VRK1 to prevent the proliferation of cancer cells.

1.6 VRK1 as a therapeutic target for cancer

Recent studies on different types of carcinomas showed that there is a positive correlation between the overexpression of VRK1 and cancer development (Molitor and Traktman, 2013, Huang et al., 2016). In the study of human lung carcinomas, high levels of VRK1 was observed in tumors with p53 mutation. This implied that deficiency in p53 function can affect the autoregulatory loop between VRK1 and p53, which results in the increase of VRK1 during cancer cell proliferation (Valbuena et al., 2007b). In another study, high levels of VRK1 was observed to be expressed in breast tumor positive for estrogen and progesterone receptors. The study further showed the importance of VRK1 in regulating DNA damage response induced by ionizing radiation, through the triggering of p53-binding protein 1 (53BP1) foci formation (Sanz-Garcia et al., 2012). The knockdown of VRK1 downregulates the formation of 53BP1 foci necessary for DNA-damage repair. As VRK1 is involved in initiating the DDR pathway, high levels of VRK1 in cancer patients, may cause the tumor to be more resistant to DNA-damage based therapies such as ionizing radiation or doxorubicin (Salzano et al., 2014). Studies targeting VRK1-mediated regulation revealed that small natural compounds such as luteolin (Kim et al., 2014), brazilin (Kim et al., 2015a) and ursolic acid (Kim et al., 2015b), could lead to the

suppression of cancer cell proliferation. Hence, VRK1 is shown to be a promising therapeutic target against cancer.

The important regulatory roles of mitotic kinases in cell cycle progression and their implications in cancer development, suggests that the designing of inhibitors against these kinases can aid in controlling their phosphorylation activities in cancer cell proliferation. Till-date, there are about 37 kinase inhibitors approved by the Food and Drug Administration (FDA) (Zhang et al., 2017), among them more than 50% targets tyrosine kinases (Wu et al., 2015). However, most of these kinase inhibitors have shown to target the ATP-binding site and can cross-react with more than one target (Wu et al., 2015). This is due to the highly conserved catalytic domain among the protein kinase superfamilies (Hanks and Hunter, 1995). VRK1, belonging to the CK1 family, also contains a highly conserved catalytic domain. VRK1, like other mitotic kinases, have important regulatory roles in cell division and deregulation of these kinases can cause overwhelming cell proliferation or genome instability leading to carcinogenesis. Therefore, the development of inhibitors for VRK1 to control its regulation in cancer cells is important. More importantly, identifying highly specific inhibitor against VRK1, which will not cross-react with other mitotic kinases with the same functional roles, would be of utmost priority. To design specific kinase inhibitors, the understanding and comparison of the catalytic domain of different protein kinases are important and necessary. Recently, the studies on mitotic kinases and the comparison of their catalytic domain provided insights for developing kinase inhibitors with better specificity (Bayliss et al., 2012). Mitotic kinases such as Aurora B and Haspin, which have similar functional roles as VRK1, has their ATP-bound crystal structures available for comparison (Sessa and Villa, 2014,

Eswaran et al., 2009). The structure of VRK1 bound to ATP is not yet available and the determination of the complex structure can provide more insights toward designing ATP-competitive inhibitor with better specificity.

1.7 Aims of study

The aim of this thesis focuses on the molecular characterization of the ATP-binding domain of VRK1. In particular the molecular basis of VRK1 interaction with (a) ATP and (b) nucleosome.

a) Recently, several synthetic inhibitors identified from the PKIS library (Elkins et al., 2016), were shown to bind VRK1 at the ATP binding site (Counago et al., 2017), but their specificity is questionable. Hence, we hope that solving the ATP-bound VRK1 structure, and its comparison with the inhibitor-bound structures, and with its own VRK family members and other mitotic kinases, can provide insights into designing more potent and specific inhibitors.

b) Besides the understanding of VRK1 catalytic domain, the comprehension of the VRK1- substrate recognition is also essential. As mentioned, VRK1 has several regulatory roles involved in the remodeling of chromatin during mitosis, especially the phosphorylation of histone H3 important for chromosome condensation. Therefore, for the next part of our study we would like to determinate the complex structure of VRK1 with nucleosome core particle (NCP) using X-ray crystallography for the understanding of VRK1-mediated H3 regulation. Toward this direction, we would like to reconstitute the 145-bp 601 nucleosome core particle and complex it with VRK1. Following this, attempts to crystallize the complex and solve the structure would be carried out, to look at how histone H3 interacts with the catalytic domain of VRK1.

Chapter 2

Materials and Methods

2.1 Materials

2.1.1 Plasmid constructs and primers

For the VRK1 wild-type construct (VRK1^{wt}), the human *vrk1* (h*vrk1*) full-length gene (residues 1-396) was cloned into pet29b plasmid vector (Novagen, USA), and for the VRK1 NMR solution structure construct (VRK1$^{\Delta 35}$), 35 amino acids were deleted from the C-terminus of the h*vrk1* gene (residues 1-361). The VRK1 crystallization construct (VRK1^{cryst}) was purchased from GenScript, USA. The VRK1^{cryst} plasmid contains a deletion of 32 amino acid residues from the C-terminus of the *hvrk1* (residues 1-364) and 11-point mutations of Lys/Gln to Ala residues. The 11-point mutations are K34A, K35A, E36A, E212A, K214A, E215A, E292A, K293A, K295A, K359A and K360A. The purchased VRK1^{cryst} construct was then cloned into pE-SUMO plasmid vector (Life Sensors, USA) for downstream studies.

The plasmid constructs for the nucleosome study were kindly given by Professor Lars Nordenskiold's group, SBS, NTU. The human histone H2A type 1B/E (Zhong et al., 1983), H2B type 1K (Dobner et al., 1991) and H4 (Pauli et al., 1987) genes were cloned into pet3a vector, while human histone H3 type H3.1(Kardalinou et al., 1993) gene was cloned into pet3d vector. Multiple repeats of the Widom 601 145-bp (145-bp 601) DNA sequence (Vasudevan et al., 2010) was cloned into pUC19 vector.

All plasmid constructs in this study are summarized in Table 2-1. Their respective gene sequences are shown in Appendix I and Appendix II.

Table 2-1 List of plasmid constructs used in this study.

No.	Plasmid Constructs	Gene	Vector properties
1.	pet29b–VRK1^{wt}	Human VRK1 wild type	C-terminal 6xHis-tag
2.	pet29b– VRK1$^{\Delta 35}$	Human VRK1 with 35 amino acids deleted from C-terminus end	C-terminal 6xHis-tag
3.	pE-SUMO–VRK1^{cryst}	Human VRK1 with 32 amino acids removed from C-terminus end and 11 K/E mutated to A	N-terminal 6xHis-SUMO tag
4.	pet3a–H2A	Human histone H2A	Untagged
5.	pet3a–H2B	Human histone H2B	Untagged
6.	pet3d–H3	Human histone H3	Untagged
7.	pet3a–H4	Human histone H4	Untagged
8.	pUC19–145-bp 601	Widom 601 145-bp DNA (with multiple repeats)	Untagged

Primers used in this study were purchased from Sigma-Aldrich, USA (Table 2-2).

Table 2-2 List of primers used for molecular cloning.

No.	Description	Sequences
1	Forward primer of VRK1^{wt} for pet29b (NdeI)	5’ CAC TCA CAT ATG CCT CGT GTA AAA GCA GCT CAA 3’
2	Reverse primer of VRK1^{wt} for pet29b (XhoI)	5’ GCG AAG CTC GAG CTT CTG GAC TCT CTT TCT GGT TCT 3’
3	Reverse primer of VRK1$^{\Delta 35}$ for pet29b (XhoI)	5’ GGATCA CTC GAG TTT CTT TCG CTT CTT TGT 3’
4	Forward primer of VRK1^{cryst} for pE-SUMO (BsaI)	5' T TAGGT CTC TAG GTA TGC CGC GTG TGA AAG CGG CG 3'
5	Reverse primer of VRK1^{cryst} for pE-SUMO (XhoI)	5' TA CTC GAG TTC TTC AAT CTC CGC CGC ACG CTT 3'

2.1.2 Bacterial strains and antibiotic stocks

Different strains of *Escherichia coli* (*E. coli)* competent cells were used based on the experimental purposes. The DH5α strain was selected for the amplification of plasmid DNA in *E. coli* cells. Both the BL21 (DE3) and BL21 (DH3) pLysS strains contain T7 RNA polymerase that is commonly used for the

overexpression of proteins. BL21 (DH3) pLysS competent cells were selected over BL21 (DH3) as it is better suited for expression of toxic proteins and has an additional selection marker using chloramphenicol.

Antibiotics stocks such as kanamycin (Gibco, USA), carbenicillin (Gold Biotechnology, USA) and chloramphenicol (Sigma-Aldrich, USA) were used accordingly against the specific antibiotic resistance gene in each plasmid. Kanamycin and carbenicillin stock solutions were prepared at a concentration of 30 mg/mL and 100 mg/mL respectively, to a volume of 10 mL in autoclaved water. Both antibiotic stocks were filter sterilized using the 0.2 µM syringe filter (Sartorius, GER). Chloramphenicol stock was prepared at concentration of 25 mg/mL to a total volume of 10 mL in 100% ethanol. All the antibiotics were aliquoted into smaller volumes and stored at 4ºC/-20ºC.

2.1.3 Bacterial culture media

Luria-Bertani (LB) broth, Luria-Bertani (LB) agar, 2xTY medium and Terrific broth were used for the growth of *E. coli* bacterial strains in DNA and protein production. The minimal M9 medium was used for producing labeled protein for NMR. The composition of each medium is included in Appendix III-2.

2.1.4 Buffers for purification, extraction and reconstitution

The compositions of the buffers used for a) VRK1 protein purification, b) Histones protein purification, c) DNA extraction and purification and d) Nucleosome core particle reconstitution is included in Appendix III-3. The compositions of stock solutions used are shown in Appendix III-4.

2.1.5 SDS- and DNA- PAGE reagents and gel composition

The gel recipes and the reagents used for SDS- and DNA-PAGE are described in Appendix III-5.

2.2 Methods

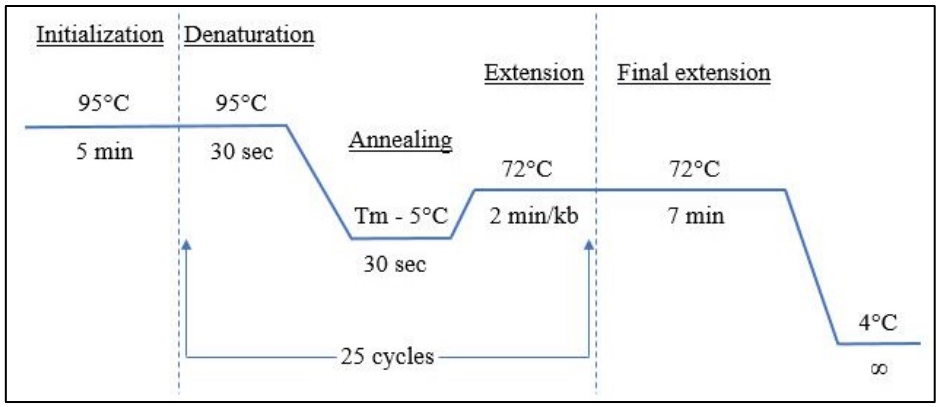
2.2.1 Molecular cloning

The gene of interest to be cloned into the specific vector was first amplified by polymerase chain reaction (PCR) using *pfu* DNA polymerase (Thermo Scientific, USA). 25 µL of PCR reaction was set up as following:

Components	Amount (µL)	Final Concentration
Autoclaved water	18.5	
10x Pfu buffer + $MgSO_4$	2.5	1x
*Forward primer (10 µM)	1	0.4 µM
*Reverse primer (10 µM)	1	0.4 µM
dNTP (10 mM)	0.5	0.2 mM
Template DNA (~100 ng/µL)	1	~100 ng
pfu polymerase (2.5 U/µL)	0.5	1.25 U
Total	25	

* The primers used for amplifying the different DNA constructs were listed in Table 2-2.

The mixture was loaded into a 0.2 mL PCR tube and placed into GeneAmp® PCR system 9700 fast thermal cycler (Applied Biosystems, USA). The reaction was carried out with the following cycling parameters:

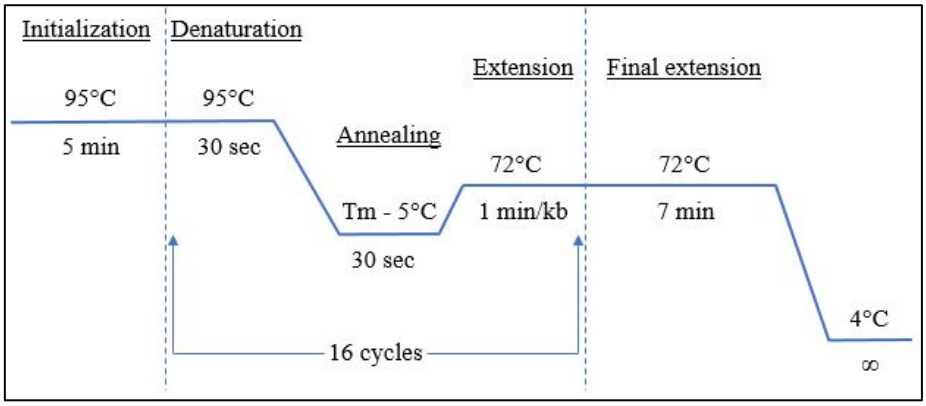


The PCR product was analyzed on 1% agarose gel. Following the New England Biolabs protocol, the cloning vector and insert were subjected to restriction enzyme digestions. The digested vector was analyzed on 1% agarose gel and the pure band was gel purified using the AxyPrep DNA Gel Extraction kit (Axygen Biosciences, USA). Conversely, the PCR amplified DNA insert was purified using the AxyPrep PCR Clean-up kit (Axygen Biosciences, USA). A 20 µl ligation mixture with a molar ratio of 1:3 vector to DNA was mixed with T4 DNA ligase (New England Biolabs, USA) and incubated at 16°C for overnight. The ligation products were then transformed into DH5α *E. coli* competent cells. The colonies obtained were screened by colony PCR using *Taq* DNA polymerase (Thermo Scientific, USA). Each colony selected was diluted in 30 µL of LB broth. A 50 µL of colony PCR reaction mixture was set up as followed:

Components	Amount (µL)	Final Concentration
Autoclaved water	35.75	
Taq buffer + KCl (10x)	5	1x
$MgCl_2$ (25 mM)	1	2 mM
*Forward primer (10 µM)	1	0.2 µM
*Reverse primer (10 µM)	1	0.2 µM
dNTP (10 mM)	1	0.2 mM
Template DNA	5	
Taq polymerase (5 U/µL)	0.25	1.25 U
Total	50	

*The primers used for amplifying the different DNA constructs were listed in Table 2-2.

The PCR reactions were placed into PCR thermal cycler and carried out with the following procedure:



The PCR product was analyzed on a 1% agarose gel. Positive clone was sent for DNA sequencing performed by 1st Base company.

2.2.1.1 pE-SUMO–VRK1^{cryst} construct

For protein expression, VRK1^{cryst} gene was sub-cloned into pE-SUMO vector. The pE-SUMO vector contained a T7 promoter and expressed recombinant protein that carried an N-terminal 6xHis-SUMO tag. The VRK1^{cryst} DNA insert was amplified by polymerase chain reaction (PCR), and the PCR products were analyzed on 1% agarose gel. The VRK1^{cryst} DNA insert with a sequence length of 1,100 bp was observed on the gel (Appendix IV). Both the gene (VRK1^{cryst} DNA insert) and the vector (pE-SUMO) were double digested with BsaI and XhoI restriction enzymes. The digested single stranded pE-SUMO vector DNA (~5,600 bp) was seen on the 1% agarose gel (Appendix IV). The VRK1^{cryst} DNA was ligated with pE-SUMO vector and positive clones were screened by colony PCR. The DNA sequence of the positive pE-SUMO–VRK1^{cryst} construct was further confirmed by sequencing (Appendix VI).

2.2.2 Competent cells preparation

The respective *E. coli* competent cell stock was streaked onto a LB agar plate without antibiotics added and incubated at 37°C overnight. On the next day, a single colony was picked and inoculated into 5 mL of LB broth without antibiotics. The culture was grown overnight in the 37°C shaking incubator at 200 rpm. Then, the 5 mL overnight-culture was added into 100 mL of fresh LB broth and continued to incubate at 37°C until the absorbance at 600 nm reached between 0.4-0.5. The bacterial culture was chilled on ice for 10 min and cells were harvested by centrifugation at 2000x g for 15 min at 4ºC. The supernatant was discarded, and the cells were resuspended with 50 mL of ice-cold autoclaved 100 mM $CaCl_2$ and kept on ice for 40 min to 2 h. The cell culture was centrifuged at 2000x g for 15 min at 4ºC. The harvested cell pellet was then resuspended in 2.5 mL of 0.1 M $CaCl_2$ (ice-cold autoclaved) with 20% v/v glycerol and aliquoted into tubes of 50 µL volume each. The tubes of competent cells were snap-frozen with liquid nitrogen and stored in the -80°C freezer.

2.2.3 Transformation of *Escherichia coli* cell with targeted plasmid

The desired plasmid (~ 50 ng) was added into 50 µL of competent cell and incubated on ice for 30 min. The cells were then heat shocked at 42°C for 50 sec, and immediately transferred onto ice to be chilled for 2 min. 250 µL of LB broth was added to the cells and incubated in a 37°C shaking incubator at 200 rpm for an hour. After that, the incubated sample was plated onto the LB agar plate with appropriate antibiotics and incubated in the 37°C incubator for about 14 h.

2.2.4 Plasmid DNA extraction

A single colony of DH5α *E. coli* cells transformed with plasmid DNA containing targeted gene was picked and inoculated into 5 mL of LB broth with appropriate antibiotics, and then incubated in a 37°C shaking incubator for 12-16 h. On the next day, the plasmid DNA was isolated using the AxyPrep Plasmid Miniprep Spin kit (Axygen Biosciences, USA) following the protocol provided. The concentration of the isolated plasmid DNA was quantified using the NanoDrop 2000 spectrophotometer (Thermo Scientific, USA) and later stored in the -20ºC freezer. A good-quality DNA sample will have an $A_{260/280}$ reading between 1.8 – 2.

2.2.5 Protein expression using bacterial cell system

A single BL21 (DH3) cell colony transformed with plasmid DNA containing the gene of interest was picked and inoculated into the fresh LB broth with antibiotic added. The starter-culture was incubated overnight at 37°C with shaking. The next morning, the overnight-culture was added into the 1 L fresh LB broth with antibiotic in a 1:100 dilution. The bacterial culture was incubated at 37°C and shaking at 180 rpm until the absorbance at 600 nm reached between 0.5-0.7. Then, the cells were induced with 1 mM of IPTG and continuously incubated in the shaking incubator at 25ºC for 4 h to express the protein of interest. However, some cells were induced at lowered temperatures of 16-20 °C for 12-14 h to enhance the expression of soluble proteins. Finally, the cells were harvested by centrifugation at 5500x g for 7 min using the rotor JLA 10.5 (Beckman Coulter, USA) and the pellet was kept in a -80ºC freezer.

2.2.5.1 VRK1^{cryst} protein expression optimization

The pE-SUMO–VRK1^{cryst} plasmid was transformed into *E. coli* BL21 (DE3) competent cells, using the bacterial expression system to overexpress the target protein. To optimize the growth conditions for the bacterial cell to express the target protein, primarily expressing it in its soluble form, the transformed bacterial cells were cultured in small scale and induced with Isopropyl β-D-1-thiogalactopyranoside (IPTG) at different temperatures. The pE-SUMO–VRK1^{cryst} transformed *E. coli* cells were first induced with IPTG at 25°C, in which the target proteins expressed in its insoluble state. This was observed from the SDS-PAGE analysis, as a band at the molecular weight (M.W.) at approximately 50 kDa in the cell pellet fraction and not in the supernatant (Appendix IV). The induction temperature was decreased to 16°C and cultured overnight to give the cells a longer period to express the protein in the cold environment (San-Miguel et al., 2013). A band was observed in the supernatant fraction at about 50 kDa on the SDS-PAGE gel which reflected that most of the VRK1^{cryst} was expressed in its soluble state (Appendix IV).

2.2.6 Protein purification using Ni^{2+}-NTA affinity chromatography

The 1 L cell culture pellet was resuspended in 30 mL of cell lysis buffer and sonicated on ice for 30-45 min with 3 sec on - 1 sec off pulse and an amplitude of 25%. The cells were sonicated until the cell lysate was significantly clarified and by centrifuged at 38,000x g for 30 min using the JA 25.5 rotor (Beckman Coulter, USA). The soluble fraction was then collected and incubated with 2-4 mL bed volume of Ni^{2+}-NTA agarose beads (Qiagen, DE) pre-equilibrated with 10x bed volume of cell lysis buffer. The protein was incubated with the beads on

a nutator at 4°C for an hour. The incubated mix was poured into a chromatography column (Bio-Rad, USA) to filter out the unbound proteins. The flow-through was collected and the beads were washed with 10x bed volume of wash buffer to remove non-specific binding proteins. Elution buffer was then added at a volume of 15 mL and incubated with the beads for 10 min to enhance the elution of 6xHis/6xHis-SUMO tagged protein. The soluble, flow-through, wash and elution fractions were collected and analyzed using SDS-PAGE.

6xHis-SUMO tagged protein was further incubated with sumo protease (SP) at 4°C overnight. On the next day, the overnight sample was incubated with pre-equilibrated Ni^{2+}-NTA beads for 10 min and filtered by the chromatography column. The flow-through containing the protein (separated from cleaved 6xHis-SUMO tag) was collected.

All the desired proteins purified by nickel purification were further purified by size exclusion chromatography.

2.2.7 Ni^{2+} - NTA agarose beads regeneration

The used Ni^{2+}-NTA agarose beads can be reused by recharging the beads with fresh Ni^{2+}. Firstly, the used beads were poured into the chromatography column to drain out the 20% ethanol storage buffer. Then, 2 CV of 8 M Urea was added to the beads, followed by 5 CV of autoclaved water. The beads were first washed with an increasing concentration of ethanol - 1 CV of 25%, 50%, 75% and 100%, followed by a decreasing concentration of ethanol - 1 CV of 75%, 50% and 25%. Then, the beads were rinsed with 5 CV of autoclaved water before stripping off the nickel ions. The Ni^{2+} ions were removed by chelation using 5 CV of 100 mM EDTA and then washed with 5 CV of autoclaved water to remove excess EDTA.

The beads were recharged with 2 CV of 100 mM of Ni^{2+} solution followed by washing with 5 CV of autoclaved water. Lastly, the beads were washed with 2 CV of 20% ethanol and stored in 20% ethanol at 50% slurry.

2.2.8 Buffer exchange using PD-10 desalting column

The protein of interest present in any buffer was exchanged into desired buffer using the PD-10 desalting column (GE Healthcare, UK), following the protocol provided.

2.2.9 Concentration of protein using Amicon concentrator

Protein was concentrated to desired amount using the Amicon ultra-15 centrifugal filter units with a molecular weight (M.W.) cut-off of 10 kDa (Merck Millipore, USA). The protein with a M.W. higher than the cut-off was trapped by the membrane hence allowing the protein sample to be concentrated.

2.2.10 Protein purification by Size Exclusion Chromatography (SEC)

SEC/gel filtration is a method that uses a liquid phase to separate the proteins through a resin column according to their molecular sizes. The HiLoad® 16/600 Superdex 200 pg column (GE Healthcare, UK) was first mounted onto the ÄKTA FPLC system (GE Healthcare, UK). The column was then equilibrated with 1.2 column volume (CV) of gel filtration buffer while the 2 mL sample loop was equilibrated with 5 mL of gel filtration buffer. After that, the protein sample concentrated to 1.5 mL was manually injected into the sample loop. The program was set up to inject the protein sample into the column and after which the column was pumped with 1 CV of gel filtration buffer at a flow rate of 0.5 mL/min with

an alarm pressure set at 0.5 MPa. The eluent was collected in fraction of 2 mL per tube and the fractions containing the protein of interest were analyzed by SDS-PAGE.

2.2.11 Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis (SDS-PAGE) analysis

The fractions collected from the protein purification processes were analyzed by SDS-PAGE, especially the SEC sample to check the purity of the protein of interest. A small amount of 10-20 µL of each fraction was mixed with 6x SDS loading dye and heated for 5 min at 95°C to denature the proteins. The percentage of the polyacrylamide gel was prepared according to the molecular size of the target protein that was to be resolved. The amount of protein sample to be loaded into the well was determined by the concentration of the protein, depending on the maximum sample volume the well can contain. The protein samples were loaded into the gel wells along with the Precision PlusProtein^{TM} All Blue Pre-stained protein standard and ran in 1x SDS running buffer at a voltage of 120 V for 70-90 min. Afterwards, the gel was stained with Coomassie blue staining solution for 15 min and destained in Destaining buffer for an hour. Then, the gel was stored in distilled water and the gel image was captured using the gel documentation machine (Bio-Rad, USA).

2.2.12 Western blot analysis

The protein samples were resolved by SDS-PAGE and the protein bands were later transferred onto a nitrocellulose membrane using the wet transfer method. The SDS-PAGE gel was sandwiched between the nitrocellulose membrane and

filter papers, then the filter pads were added before the filter paper. The gel together with the membrane were compressed in a cassette and submerged in a tank filled with WB transfer buffer (prepared for SDS-PAGE) and subjected to an electrical field. The transfer was done in 60-70 min at a voltage of 100 V.

The membrane containing the protein bands was incubated in 5% w/v skim milk in TBS-T for an hour at room temperature to block non-specific protein binding. The membrane was then blot with primary antibody prepared in 5% w/v skim milk in TBS-T for overnight at 4°C. On the next day, the membrane was washed with TBS-T thrice, for 10 min at each time, after which it was probed with secondary antibody in 5% w/v skim milk in TBS-T for an hour at room temperature, followed by 3 rounds of TBS-T wash. The visualization of the protein was carried out by chemiluminescent Western blots later imaged using X-ray films. The membrane was incubated with the ECL mix prepared from WesternBright ECL HRP substrate kit (Advansta, USA) and imaged on the X-ray film at different exposure times.

2.2.13 Protein concentration determination by NanoDrop spectrophotometer

The concentration of the protein was determined by NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific, USA) using the Protein A280 program provided in the software. In this program, the extinction coefficients (E.W) and the molecular weight (M.W.) of the protein to be measured were input into program. The NanoDrop 2000 was first blank with the buffer the protein is in. Then 2 μL of protein sample was loaded onto the NanoDrop 2000. The

measurement given as the concentration of protein was in mg/mL. The $A_{260/280}$ reading of a pure proteins is below 0.7.

2.2.14 ADP-Glo kinase assay

The ADP-Glo kinase assay (Promega, USA) is an assay that measures the ADP produced after kinase activity by luminescence method. The kinase activity experiment can be performed by following the procedure provided by Promega. Using the 384-well flat bottom white plate (Greiner, USA), a 5 µL of kinase reaction mix was set up containing 1 µM of VRK1 protein in 40 mM Tris pH 7.5, 20 mM $MgCl_2$, 0.1 mg/ml BSA, 2.5 mM $MnCl_2$, 50 µM DTT assay buffer. The kinase was titrated with 0 – 250 µM of ATP and the reactions were carried out at room temperature for an hour. After that, 5 µL of ADP-Glo reagent was added into the kinase reaction and incubated for 40 minutes. The reagent was used to eliminate any unused ATP that remained in the kinase reaction. Lastly, 10 µL of kinase detection reagent was added to the mix to convert the ADP produced from the kinase reaction to ATP. The converted ATP can then be converted to a signal through luciferase reaction. The luminescence signal was later measured with a plate-reading luminometer, with integration time set to 1000 ms for each well. The data collected was analyzed using GraphPad Prism 5 software (La Jolla, CA).

2.2.15 Preparation of ^{15}N-labeled protein

The uniformly ^{15}N-labeled VRK1$^{\Delta 35}$ protein was prepared by culturing the cells in M9 minimal medium containing ^{15}N-labeled ammonium chloride. VRK1$^{\Delta 35}$ plasmid DNA was first transformed into BL21 (DH3) *E. coli* cells. A single colony was picked and inoculated into 5 mL of M9 medium starter-culture. On

the next day, the starter-culture was added into 500 mL of M9 medium in the dilution of 1:100 and was cultured at 37ºC with shaking at 180 rpm for 6 hours to reach an absorbance at 600 nm between 0.6-0.8. The culture was later induced with 1 mM of IPTG and further incubated at 20ºC with shaking for overnight. The proteins were purified following the protocol in Section 2.2.6. The NMR samples were prepared in 20 mM sodium phosphate (pH 7.0), 150 mM NaCl, 1 mM DTT, 0.01% NaN_3 and 5% D_2O buffer.

2.2.16 ^{1}H–^{15}N Transverse Relaxation Optimized Spectroscopy (TROSY) experiment

600 µL of 0.2 mM of ^{15}N-labeled VRK1$^{\Delta35}$ with 5% D_2O added, was serial titrated with 6 µL of 100 mM of Adenylyl-imidodiphosphate (AMP-PNP), a non-hydrolyzable ATP (Sigma-Aldrich, USA). The ^{1}H-^{15}N TROSY titration data for the protein to ligand in the molar ratio of 1:0.5, 1:1, 1:3 and 1:5 were collected at 25°C by Dr Ye Hong. Data were acquired by performing 24 accumulative scans per titration on the Bruker 700 MHz NMR spectrometer equipped with a 5 mm z-gradient TXI cryoprobe with a resolution of 2048 and 256 points in the direct and indirect dimensions. The spectra acquired were processed through Topspin version 2.1 (Bruker BioSpin) and analyzed using the NMRFAM Sparky (Lee et al., 2015).

The chemical shift perturbation was calculated using the formula $\sqrt{[(\Delta\delta_H)^2 + 0.156 \times (\Delta\delta_N)^2]}$. $\Delta\delta_H$ is the ^{1}H ppm difference between the free VRK1$^{\Delta35}$ and ligand-bound VRK1$^{\Delta35}$, while $\Delta\delta_N$ is the ^{15}N ppm difference between the free VRK1$^{\Delta35}$ and ligand-bound VRK1$^{\Delta35}$. The chemical shift perturbation was plotted.

2.2.17 Crystallization and X-ray diffraction data collection

Purified VRK1^{cryst} in 20 mM Tris-HCl (pH 7.5), 150 mM NaCl and 1 mM DTT, concentrated to 10 mg/mL, was used for the initial co-crystallization trials. The protein was incubated with AMP-PNP in molar ratio of 1:6 and supplemented with 10 mM of magnesium chloride ($MgCl_2$) for overnight at 4°C. Prior to crystallization screening, the incubated protein was spun down in a cold centrifuge at 12,000 xg for 5 min.

Crystallization trials were performed using the hanging drop vapor diffusion method at 18 °C, using the crystallization screening kits available in our lab such as the HR-Index (Hampton Research, USA), Wizard I & II (Rigaku Reagents, USA) and MemGold (Molecular Dimensions, USA). Concurrently, we also tried the crystallization conditions published for VRK1-inhibitor complexes (PDB ID's 3OP5, 5UVF and 5UKF). The VRK1^{cryst}–AMP-PNP complex was mixed with the reservoir solution in equal volume (1 µL each) and sealed with 250 µL of reservoir solution. After a week, small crystals were observed growing under 35 % w/v PEG 3350, 0.1 M Tris-HCl (pH 8.5) and 0.2 M ammonium sulphate. Hence this condition was selected for optimization.

To improve the crystal size, we increased the VRK1^{cryst} concentration to 20-30 mg/mL. The protein was incubated with AMP-PNP in the molar ratio of 1:6 and supplemented with 10 mM $MgCl_2$. To improve protein nucleation, we optimized the precipitant condition by varying the crystallant, buffer and salt as shown in Table 2-3. In addition, we also varied the protein to reservoir solution ratios, such as 1:1, 1:2 and 2:1.

Table 2-3 Optimization of VRK1^{cryst}-bound AMP-PNP crystallization screening condition.

Precipitant component	Hit condition	Optimization
Crystallant	35% w/v PEG 3350	18-30 % PEG 3350
Buffer (0.1 M)	Tris (pH 8.5)	1. Tris (pH 7.0-9.0) 2. HEPES (pH 6.7-7.5)
Salt (0.2 M)	Ammonium sulphate	1. Ammonium sulphate 2. Magnesium chloride

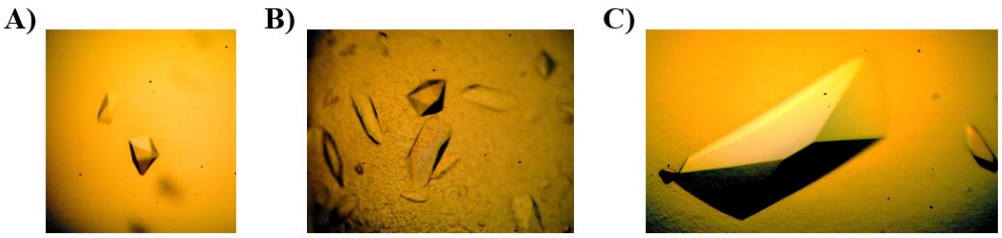


Figure 2-1 Protein crystals from crystallization screening process.

The best crystals of VRK1^{cryst}–AMP-PNP complex (Figure 2-1) appeared in 27.5% w/v PEG 3350, 0.1 M of HEPES (pH 7.0) and 0.2 M of ammonium sulphate, after 12-15 days at 18 °C. The crystals were cryo-protected with reservoir solution supplemented with 15 % w/v glycerol.

The diffraction data were collected at National Synchrotron Radiation Research Center (Hsinchu City, Taiwan) by Dr Sreekanth Rajan. As the crystal was big (~ 0.5 x 0.5 x 0.3 mm^3), three datasets were collected from the same crystal.

2.2.18 Crystal structure determination

The images obtained were indexed and integrated using iMosflm (Battye et al., 2011), and later merged and scaled by SCALA (Evans, 2006), both of which are a part of the CCP4 suite programs (Winn et al., 2011). Based on the cell parameters and systematic absences, the crystal was identified to belong to the orthorhombic system with the space group P $2_12_12_1$. Furthermore, based on the intensity statistics (Table 3-3) parameters such as R_{merge}, mean ($I/\sigma I$), completeness, etc, the resolution cut-off was justified to be 2.07Å. To identify the number of molecules in the asymmetric unit and enable molecular replacement solution, the Mathews co-efficient module under the CCP4 suite, which estimates the crystal volume per unit of protein molecular weight, was used. This suggested the presence of four molecules/chains in the asymmetric unit, where the initial phases were acquired through molecular replacement using the program PHASER (McCoy et al., 2007) by utilizing VRK1-inhibitor structure (PDB ID 5UKF) as the starting model, resulting in a LLG/TFZ score of 17296/32.8. Moreover, it could be observed that our present structure is comparable to the model (PDB ID 5UKF or any other VRK1-inhibitor crystal complexes), which also had four chains in the asymmetric unit, but many residues at the N-terminal end of chain C and D were missing. Hence, we used the program BALBES (Long et al., 2008), and the missing regions at the N-terminal of chain C and D were determined to better accuracy than the model, resulting with the Qfactor of 0.869 (an ideal Qfactor is 1.0) and R_{factor}/R_{free} of 0.24/0.28. BALBES is a program used for solving the crystal structure of protein by Molecular Replacement (MR) method. When the protein sequence and structural factor files are provided to the system, it will search through the internal database

for homologous model then generate a variety of template models. The refinement were performed by using BUSTER (Smart et al., 2008, Smart et al., 2012) with the autoBUSTER mode limiting the resolution between 20-2.07Å, resulting in a final R_{factor}/R_{free} of 0.18/0.22. For electron density map fitting the program COOT (Emsley and Cowtan, 2004) was used. After structure refinement and map fitting, the electron density of AMP-PNP, with the magnesium ion, present in the active site were clearly visible in all the four chains in the asymmetric unit. In addition, Poly-ethylene glycol (PEG), sulphate and chloride ions, from the crystallization precipitant, were also observed in the electron density map. The Fo-Fc and 2Fo-Fc electron density maps were contoured at 3.0 and 1.0σ cut-offs, respectively, for the identification of water molecules. The Ramachandran statistics were calculated using Molprobity (Chen et al., 2010, Davis et al., 2007). The protein ligand interactions were analyzed by Ligplot (Wallace et al., 1995) and manual examination. The sequence alignment was done in Clustal Omega (Sievers et al., 2011) and EsPript (Gouet et al., 2003). All figures presented were prepared using the PyMOL (DeLano Scientific, Palo Alto, CA, USA). The detailed refinement statistics were tabulated in Table 3-3.

2.2.19 Nucleosome core particle reconstitution

The preparation of 145-bp 601nucleosome core particle was shown as a schematic diagram in Figure 2-2, which is made up of two parts, 1) the DNA preparation and 2) the Histone octamer refolding (Luger et al., 1999).

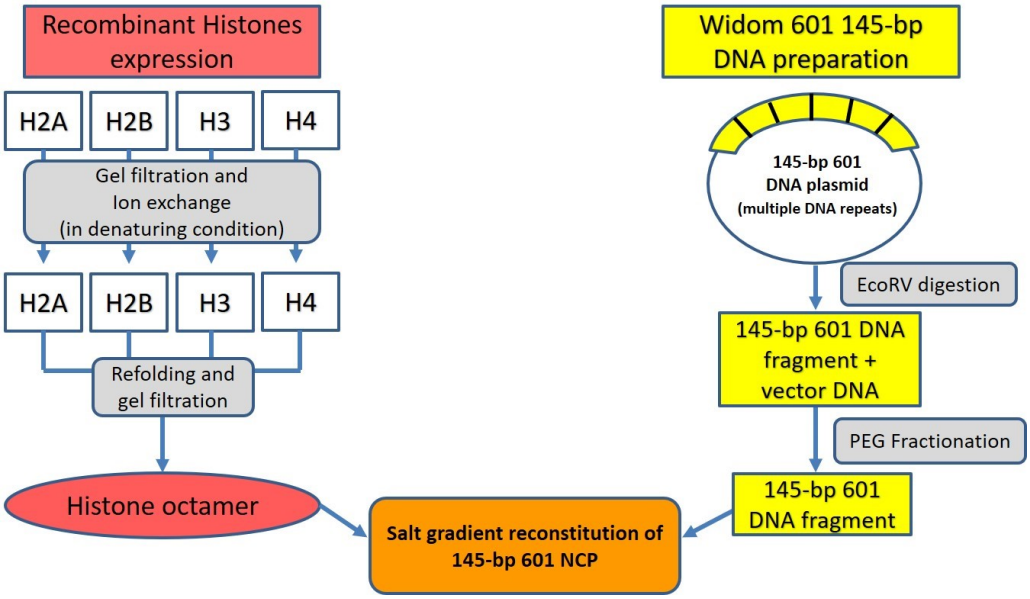


Figure 2-2 A schematic diagram of 145-bp 601 nucleosome core particle reconstitution.

2.2.19.1 145-bp 601 DNA fragment preparation

The pUC19–145-bp 601 plasmid DNA was transformed into DH5α *E. coli* cells. A single colony of successful transformants was picked from the agar plate, inoculated into a 5 mL Terrific broth (TB) containing carbenicillin, and incubated in a shaking incubator at 37°C at 200 rpm overnight. On the following day, the 5 mL overnight culture was added into 100 mL of fresh TB with carbenicillin and incubated at 37°c with shaking until the absorbance at 600 nm reached 0.8. The 100 mL culture was then distributed evenly into twelve flasks of TB with carbenicillin added (500 mL of TB per flask), and incubated at 37°C, shaking at 220 rpm for 19 h.

On the next day, the 6 L of bacterial cultures was split evenly into six 500 mL centrifuge bottles. The culture was harvested by centrifugation at 5500x g for 7 min using the JLA 10.5 rotor. The cell pellet in each bottle was resuspended with 60 mL of alkaline lysis buffer I and homogenized using a sterile glass rod. After resuspending the cells, 120 mL of alkaline buffer II was added, and the mixture was shaken vigorously for 2 min followed by a 20 min incubation on ice. During the incubation, the bottles of cells were shaken vigorously every 5 min. Next, 210 mL of cold alkaline lysis buffer III was added and mixed immediately and gently by inverting the bottle for 2 min. After which, the bottles of cells were incubated on ice for 20 min and were inverted gently every 5 min during the incubation. The cell lysate was centrifuged at 10,000x g for 20 min at 4°C.

The supernatant collected was filtered through the sterile gauze pads (8-fold thick). For plasmid DNA precipitation, isopropanol was added at 0.52x volume to the supernatant, mixed and incubated for 30 min at room temperature. The solution was centrifuged at 16,000x g for 30 min at 20°C. The supernatant was discarded, and the pellet was air-dried before dissolving it with 40 mL of TE 10/50 buffer. 240 µL of 10 mg/mL RNase A (Thermo Scientific, USA) was added, the solution was incubated at 37°C with shaking for overnight to digest any RNA contaminants present in the sample.

Following RNA digestion, the solution was centrifuged at 10,000x g for 15 min at 20°C, to remove any precipitant. The supernatant was extracted with 10 mL of Tris-equilibrated phenol for every 20 mL of solution and centrifuged at 27,000x g for 20 min at 20°C. The phenol extraction step was repeated twice, after which the aqueous phase was mixed with 10 mL of CIA (Chloroform-isoamyl alcohol 24:1 mixture) and centrifuged at 12,000x g for 10 min at 20°C.

The plasmid DNA was then precipitated by adding 0.2x volume of 4 M NaCl and 0.4x volume of 40 % PEG 6000 to the aqueous phase. The solution was mixed and incubated on ice for 30 min, then later centrifuged at 27,000x g for 10 min at 4°C. The supernatant containing the RNA was discarded and the pellet was resuspended with 15 mL of TE 10/0.1 buffer. The resuspended solution was washed twice with 10 mL of CIA to remove the PEG. The plasmid DNA was again precipitated out by adding 0.1 volume of 4 M NaCl and 2.5 volume of 100 % ice cold ethanol to the aqueous phase, followed by 30 min of incubation on ice. After which, the solution was centrifuged at 27,000x g for 20 min at 4°C. The pellet was air-dried for 15 min at room temperature and resuspended in 10 mL of TE 10/0.1 buffer. The plasmid DNA concentration was measured using a NanoDrop 2000 spectrophotometer. A blank run with TE 10/0.1 buffer was first carried out to calibrate the spectrophotometer before measuring the DNA concentration of the sample. 2 µL of DNA sample was then loaded onto the spectrophotometer for measurement in ng/µL. The purity of the sample could be analyzed by observing the $A_{260/280}$ reading.

2.2.19.2 *EcoRV* digestion and PEG precipitation

The 145-bp 601 DNA fragments were cleaved from the plasmid by using *EcoRV-HF* restriction enzyme (New England Biolabs, USA). An *EcoRV-HF* digestion (following the NEB protocol) was first performed on 100 µg of plasmid DNA and analyzed on a 0.8% agarose gel to check the DNA quality and digestibility. A large scale of *EcoRV-HF* digestion was then performed by setting up a reaction mixture consisting of 2 mg/mL of plasmid DNA, 1x cut smart buffer from NEB (50 mM Potassium acetate, 20 mM Tris-acetate, 10 mM Magnesium acetate and

100 µg/mL BSA at pH 7.9) and *EcoRV-HF* enzyme at 150 U per mg of plasmid. The reaction mixture for each batch of 50 mg plasmid DNA was carried out at 37°C for 24 h. The digested sample was analyzed on 8% DNA-PAGE gel to check for complete restriction enzyme digestion.

After *EcoRV-HF* digestion, the DNA was precipitated out by adding of 0.1x volume of 4 M NaCl and 2.5x volume of 100% ice cold ethanol to solution, then after incubated on ice for 30 min. The solution was centrifuged at 27,000x g for 20 min at 4°C, then the pellet was air-dried for 15 min and dissolved with TE 10/0.1 buffer.

The 145-bp 601 DNA fragments can be separated from the empty vector by PEG fractionation method. 40 % w/v of PEG 6000 and 4 M of NaCl were added into the digested sample to a final concentration of 7.5% w/v PEG 6000 and 0.5 M NaCl. The solution was incubated on ice for an hour followed by centrifugation at 27,000x g for 20 min at 4°C. The supernatant was ethanol precipitated and the pellet was dissolved in TE 10/0.1 buffer. The sample was subjected to CIA extraction twice to remove PEG. The aqueous phase was again subjected to ethanol precipitation and the air-dried pellet was then dissolved with TE 10/0.1 buffer. The 145-bp 601 DNA concentration was measured using a NanoDrop 2000 spectrophotometer and the purity was determined by analyzing on a 0.8% agarose gel.

2.2.19.3 Transformation of Escherichia coli with histone genes

The plasmids consisting of Histone H2A, H2B, H3 or H4 gene were each transformed into BL21 (DH3) pLysS competent cells respectively. The individual plasmid DNA (50 ng/µL) was added into a tube of 50 µL competent

cells and incubated on ice for 30 minutes. The cells were subjected to heat shock at 42°C for 50 sec and immediately chilled in ice for 2 min. 250 µL of fresh LB with chloramphenicol (Camp) was added to the cells and incubated in a shaking incubator at 37°C and 200 rpm for an hour. The cells were then plated onto the LB agar plate with carbenicillin (Car) and Camp and incubated at 37°C for overnight.

2.2.19.4 Histones protein expression

A single cell colony containing the histone plasmid was picked up and inoculated into fresh 2xTY medium containing Car, Camp and 0.1% w/v D-glucose. The bacterial culture was incubated at 37°C in a shaking incubator for overnight. On the next day, the overnight culture was added into the 500 mL of fresh 2xTY medium containing Car, Camp and 0.1% w/v D-glucose in 1:100 dilution (12 flasks of 500 mL bacterial culture). The 6 L culture was incubated at 37°C and shaking at 200 rpm, until the absorbance at 600 nm reached 0.4-0.6. The cells were induced with 0.4 mM IPTG and continued to incubate at 37°C for 2-3 h. The cells were later harvested by centrifugation at 5500x g for 7 min using JLA 10.5 rotor. The cell pellet collected was subjected to inclusion bodies preparation.

2.2.19.5 Protein inclusion bodies preparation

The cell pellet from 6 L of bacterial culture was resuspended in 100 mL of Histone wash buffer. The cells were sonicated on ice at an amplitude of 25%, with a pulse of 1 sec on and 1 sec off for 30 min until the viscosity of solution was reduced. Afterwards, the sample was centrifuged at 12,000x g for 15min at 4^{o}C in a JA 25.5 rotor. The cell pellet (containing histones protein) was

resuspended in 160 mL of Histones wash buffer with Triton X-100. A glass rod was used to help dissolve the pellet. Then, the resuspended sample was spun down at 12,000x g for 15 min at 4ºC. The supernatant was discarded, and the pellet was washed with Triton X-100 containing buffer by repeating the step above. Afterwards, it was washed twice with 160 mL Histone wash buffer to remove the Triton X-100. The inclusion bodies prepared was stored at -20ºC until needed. The "crude histone" can be stored at -20ºC for a longer time.

2.2.19.6 Histones purification by FPLC and ion exchange

The "crude histone" was first homogenized with 1 mL of dimethyl sulfoxide (DMSO), followed by the addition of 40 mL of S-200 unfolding buffer to homogenize as much as possible. The suspension was centrifuged 3 times at 18,000x g for 30 min at 20°C to remove all aggregates. The HiLoad 16/60 Superdex 75 PG column (GE healthcare, USA) was mounted onto the ÄKTA FPLC machine. The column was first equilibrated with 1.5 CV of SAUDE-1000 buffer. The denatured histone sample was injected into the column at a flow rate of 0.5 mL/min and a maximum pre-column alarm pressure of 0.5 MPa. The size exclusion purification was carried out at room temperature, and the histones were observed to be eluted after 60 mL. The purity of the collected fractions was checked by analyzing on a 15% SDS-PAGE gel. The fractions with the most histone protein, were pooled together and dialyzed against 4 L of Milli-Q water containing 5 mM of BME for 3 times over a period of 24 h. The dialyzed sample was centrifuged at 12,000x g, 4ºC for 10 min to remove aggregates. The histone concentration was measured at absorbance at 276 nm using the NanoDrop 2000

spectrophotometer. The aliquoted histones (100 mg per tube) were lyophilized, then labeled as "S-200" and stored at -80ºC until needed.

The "S-200" labeled pellet was dissolved with SAUDE-200 buffer. The histone was injected into the pre-equilibrated 6 mL Resource S ion exchange column (GE healthcare, USA) connected to ÄKTA FPLC machine. The column was first washed with SAUDE-200 at a flow rate of 1 mL/min for 10 min to remove non-binding molecules, followed by gradually increasing the SAUDE-200 buffer to 65% of SAUDE-1000 at a flow rate of 2 mL/min over a period of 20 min, ensuring that the pressure do not exceed 0.5 MPa. The collected fractions were analyzed on a 15% SDS-PAGE gel. The fractions containing the pure histones (which has a $A_{260/280}$ reading below 0.7) were pulled together and dialyzed with water containing 5 mM BME. The dialyzed sample was aliquoted and lyophilized before storing at -80ºC.

2.2.19.7 Histone octamer refolding

The purified and lyophilized histones were dissolved separately in histone octamer (HO) unfolding buffer at 37°C for 30 min. The histone H2A, H2B, H3 and H4 were combined in equimolar stoichiometry of 1:1:1:1 and diluted to give a final protein concentration of 1 mg/mL. The solution was placed in the dialysis bag with a cut-off of 6-8 kDa and dialyzed twice against 2 L of refolding buffer at room temperature for 4-5 h per round and repeated one last round overnight. After dialysis, the solution was centrifuged twice at 9,000x g for 10 min at 20°C to remove any precipitation. Given that 1 mg/mL of HO solution gives an absorbance at 276 nm of 0.45, the approximate total yield and concentration of HO refolded was calculated and then used to determine the target volume for

concentrating the sample. The HO was concentrated using the Amicon concentrator with a cut-off 10 kDa to a final concentration of about 10 mg/mL or more.

The refolded HO sample was injected into the pre-equilibrated HiLoad 16/60 Superdex 200 PG column (GE healthcare, USA) mounted on the ÄKTA FPLC machine. The buffer was injected at a flow rate of 1 mL/min with an alarm pressure set at 0.5 MPa. The HO was eluted after 55 mL and the fractions collected were analyzed on 15% SDS-PAGE gel. The fractions containing the pure HO were pooled together and concentrated. The HO was mixed with 100% of glycerol to a final glycerol concentration of 50% and the concentration of the HO was determined at absorbance at 276 nm using the NanoDrop 2000 spectrophotometer. The sample was stored in -20ºC freezer.

2.2.19.8 Salt gradient dialysis reconstitution

To reconstitute the nucleosome core particle (NCP), the final reaction set up contained 6 µM of 145-bp 601 DNA, 2 M of KCl, 10 mM of DTT and autoclaved water. The HO is mixed with the DNA in different molar ratio such as 0.5:1, 0.8:1, 1:1, 1.05:1, 1.08:1, 1.1:1 and 1.3:1. A microscale of 100 µL of reconstitution was set up to obtain a desired reconstitution molar ratio before proceeding with a large-scale reconstitution. The reaction was then loaded into a dialysis tube to be dialyzed against 100 mL of TCS buffers with decreasing molar of KCl, changed every 2.5 h later at room temperature. The last round of dialysis was performed in TCS buffer without KCl for overnight at room temperature. The next day, the samples were collected and spun down twice at maximum

speed for 10 min at 4°C to remove precipitates. The samples were then analyzed on a 6% DNA-PAGE gel to check the quality of reconstitution.

After obtaining the desired reconstitution molar ratio, the reaction can be upscaled to milliliter quantities, depending on the amount of NCP needed for experiment.

2.2.20 Mobility Shift Electrophoretic Assay

This gel shift assay was applied to study the binding of the nucleosome core particle (NCP) with VRK1. The nucleosome core particle (NCP) was first incubated with VRK1 protein ($VRK1^{wt}$, $VRK1^{\Delta 35}$ and $VRK1^{cryst}$) at different molar ratios, in buffer containing 20 mM Tris-HCl (pH7.5), 1 mM EDTA, 1 mM DTT and varying concentrations of sodium chloride.

The 20 µL reaction was carried out with 1 µM of NCP and incubated with VRK1 proteins in different molar ratios such as 1:0.5, 1:1, 1:2, 1:3, 1:5, 1:8, 1:10 and 1:15. The VRK1 and NCP were complexed in buffers with different concentration of NaCl (from 0-250 mM). The samples were incubated at different durations such as 15 min, 1 h and overnight at 4°C. Later the incubated samples were mixed with 10% sucrose and 10 µL of the samples were loaded into the 6% DNA-PAGE gel and ran at a voltage of 100 for 80 minutes in the cold room. The gel was later stained with gel red and visualized under the Syngene T:Genius gel documentation system.

Chapter 3

Structural Characterization of VRK1 in Complex with an ATP Analog

3.1 Motivation

Mitosis is a cell division process in which a parent cell undergoes a series of regulated events and divide into two genetically identical daughter cells. This is an important cell proliferation progression that is tightly controlled by a network of mitotic kinases, as the deregulation and mutation of these kinases can lead to carcinogenesis (Schmit and Ahmad, 2007). Hence, mitotic kinases have been important cancer drug targets, with the main focus on their catalytic domain, as with any other kinase. Most of the FDA-approved small molecule kinase inhibitors are competitive in nature, binding directly to the active site pocket to compete with the ATP molecule (Wu et al., 2015). To design ATP-competitive inhibitors with high specificity and sensitivity, the structural knowledge of the ATP binding mode in kinases is necessary. As the signature motifs of the ATP binding pocket in many protein kinases are highly conserved (Bayliss et al., 2012), one of the key factors towards designing specific kinase inhibitors is the comparison of the catalytic domain of kinases to identify and exploit subtle sequence / structural differences (Rabiller et al., 2010, Eswaran and Knapp, 2010, Norman et al., 2012).

The mitotic kinase that our group is focusing on is the human Vaccinia-related kinase 1 (VRK1), which has important regulatory roles in cell division progression. Among our current projects, the utmost objective is to design inhibitors that target the ATP-binding moiety of kinases. The structure of several mitotic kinases has been determined by X-ray crystallography and the comparison of the active site region of these kinases contribute to the design of highly specific structure-based kinase inhibitors (Bayliss et al., 2012). It is pivotal to understand the unique interactions between ATP and a particular kinase,

although its general mode of binding is usually similar. Hence, we are interested in determining the structure of VRK1 in complex with ATP to identify their unique interactions. Moreover, a few VRK1-inhibitor bound complex structures were recently determined (Counago et al., 2017) and are available in the PDB database. The screening of these inhibitors was carried out using the PKIS library (Elkins et al., 2016), which consists of known kinase inhibitors. The specificity of the identified inhibitors is debatable, as only the binding was characterized, using thermal shift assay, but no functional assay was carried out. Nevertheless, the insights from these bound structures provided few important clues toward VRK1 inhibitor binding. By comparing the ATP-bound VRK1 structure with the inhibitor-bound structures, and with its own VRK family members and other mitotic kinases, we hope to obtain insights into designing more potent and specific inhibitors for VRK1.

3.2 Results and Discussion

3.2.1 VRK1 protein expression and purification

Both VRK1^{wt} and VRK1$^{\Delta 35}$ constructs were cloned into pet29b vector (Shin et al., 2011) and transformed into BL21 DE3 *E. coli* cells. The protein expression was carried out by following the conditions described in section 2.2.5 and protein purification by nickel affinity chromatography method in section 2.2.6. After nickel purification, the fractions collected were analyzed on a 12% SDS-PAGE gel.

In Figure 3-1A, the coomassie-stained gel showed the presence of VRK1^{wt} protein (M.W. 45 kDa) in the eluate fraction after nickel affinity column purification. It can be observe that these fractions were contaminated with other protein impurities, hence the sample was further purified by size exclusion chromatography using the HiLoad 16/60 Superdex 200 PG column. From the chromatogram shown in Figure 3-1B, it appeared that VRK1^{wt} protein eluted from 76-85 mL. For VRK1$^{\Delta 35}$ protein (M.W. 40 kDa), collected in the eluate after nickel affinity purification (Figure 3-1C), was observed to elute from 79-88 mL during size-exclusion chromatography (Figure 3-1D).

From each litre of bacterial culture, we obtained ~ 10-12 mg and 8-10 mg of VRK1^{wt} and VRK1$^{\Delta 35}$ proteins, respectively with high purity.

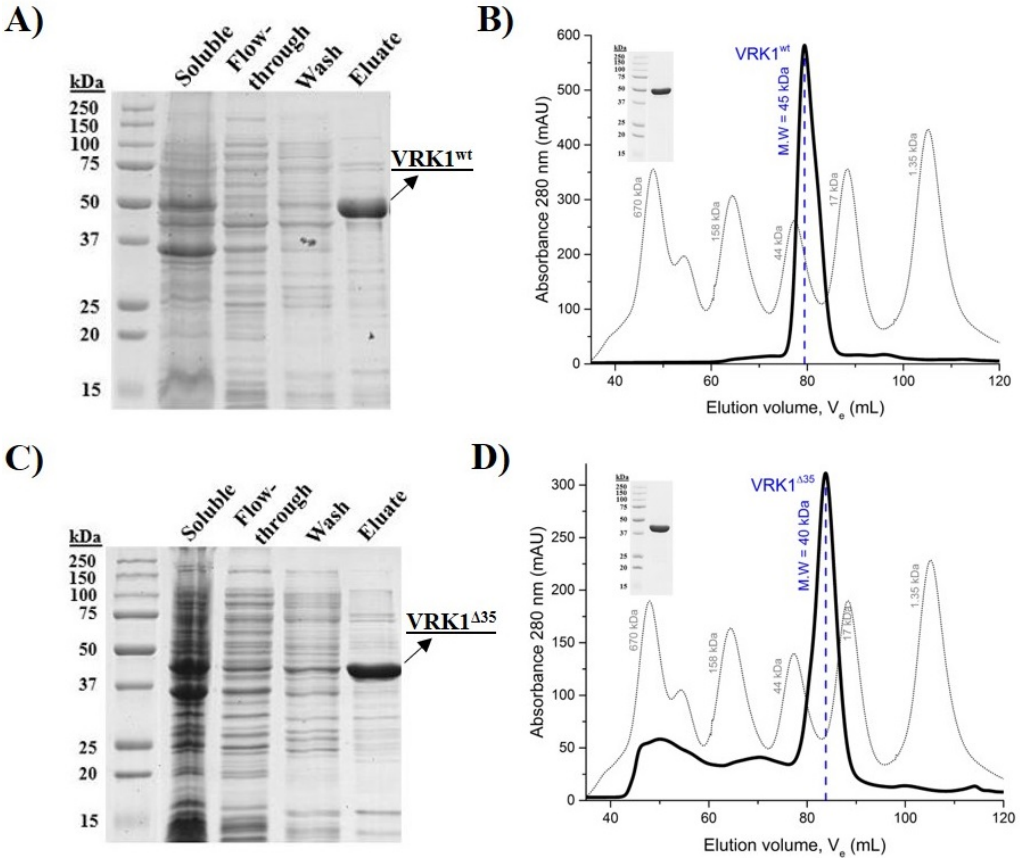


Figure 3-1 VRK1^{wt} and VRK1$^{\Delta 35}$ protein purification analysis.

A, C) SDS-PAGE analysis of VRK1^{wt} (A) and VRK1$^{\Delta 35}$ (C) after nickel affinity chromatography. The soluble, flow-through, wash and eluate fractions collected after nickel column purification are loaded along with the protein marker on a 12% SDS-PAGE gel. VRK1^{wt} protein, M.W. 45 kDa, and VRK1$^{\Delta 35}$ protein, M.W. 40 kDa, are present in the eluate fraction, respectively (indicated by black arrow). B, D) Size exclusion chromatography purification of VRK1^{wt} (B) and VRK1$^{\Delta 35}$ (D). The eluate collected from the nickel purification was further purified using the Superdex S200 16/600 column. VRK1^{wt} protein eluted from 76-85 mL, is observed as a predominant peak in the black graph (B). While, VRK1$^{\Delta 35}$ protein is observed to elute from 78-88 mL shown in the black graph (D). The inset on the top left of VRK1^{wt} (B) and VRK1$^{\Delta 35}$ (D) chromatogram, showed the respective purified protein analyzed on a 12% SDS-PAGE gel after purification. The grey dotted graph represent the Bio-rad protein standards analyzed on the same column included as a reference.

3.2.2 VRK1 and AMP-PNP interaction studies by NMR spectroscopy

The interaction of ATP with VRK1 was first characterized by NMR titration studies using ^{15}N-labeled VRK1$^{\Delta 35}$ protein and AMP-PNP, a non-hydrolyzable ATP analogue. The ^{15}N-labeled VRK1$^{\Delta 35}$ was expressed by following the method described in section 2.2.15 and ^{15}N-labeled protein carrying the 6xHis-tag were also purified according to the protocol in section 2.2.6. The purification analysis

results had similar profile as the un-labeled VRK1$^{\Delta 35}$ protein shown in Figure 3-1C and D. From each litre of M9-minimal medium cultured bacterial cells, 8-10 mg of ^{15}N-labeled VRK1$^{\Delta 35}$ protein with high purity was obtained. The labeled proteins were titrated serially with AMP-PNP in molar ratios of 1:0.5, 1:1, 1:3 and 1:5. The ^{1}H-^{15}N TROSY titration experiments were carried out using the parameters described in section 2.2.16. The spectra were analyzed using the NMRFAM Sparky (Lee et al., 2015) software.

The spectra of ^{15}N-labeled VRK1$^{\Delta 35}$ titrated with AMP-PNP at different molar ratios were overlaid to examine the ATP interaction region on the kinase (Figure 3-2A). It could be observed that the peaks of VRK1$^{\Delta 35}$ were well-dispersed, indicating that the protein was folded orderly. From the spectra of titrations (1:0.5, 1:1, 1:3 and 1:5), it was observed that some of the peaks possessed the fast exchange characteristics such as the "peak walking" pattern. While, some peaks showed a disappearance of intensity when titrated with 0.5x of AMP-PNP to VRK1$^{\Delta 35}$ (Figure 3-2A). These observations showed a fast-intermediate exchange, indicative of a transient interaction between VRK1$^{\Delta 35}$ and AMP-PNP. In comparison with other residues, those residues having stronger perturbations and peaks that disappeared were identified. The chemical shift perturbation for VRK1$^{\Delta 35}$ to AMP-PNP in the ratio of 1:5 was calculated using the formula, $\sqrt{[(\Delta\delta_H)^2 + 0.156 \times (\Delta\delta_N)^2]}$. The resulting chemical shift (δΔ) of VRK1$^{\Delta 35}$ upon addition of AMP-PNP was plotted and shown in Figure 3-3.

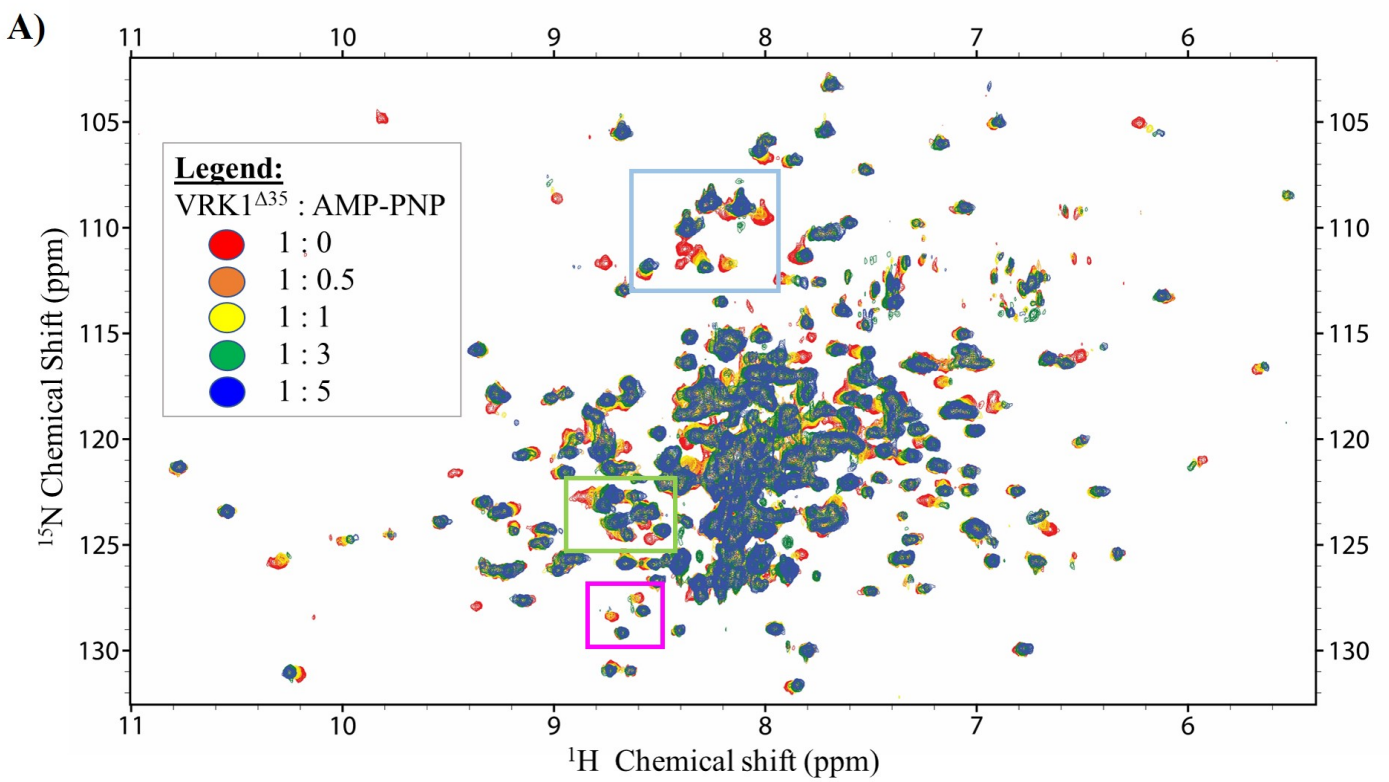


Figure 3-2 NMR titration analysis of VRK1 $^{\Delta 35}$ with AMP-PNP.

A) The 2D ^{1}H-^{15}N TROSY overlaid spectra of ^{15}N-labeled VRK1$^{\Delta 35}$ in free state (red) with the spectra of ^{15}N-labeled VRK1$^{\Delta 35}$ mixed with AMP-PNP in the molar ratio of VRK1$^{\Delta 35}$ to AMP-PNP at 1:0.5 (orange), 1:1 (yellow), 1:3 (green) and 1:5 (blue). B) The zoomed-in views (in pale blue, light green and magenta colored boxes) shown are the regions showing chemical shift changes upon the addition of AMP-PNP. The direction of the shifts from free VRK1$^{\Delta 35}$ (in red) toward AMP-PNP bound complex 1:5 (in blue) is indicated by the black arrows and all residues are labeled using the single letter amino acid code.

B)

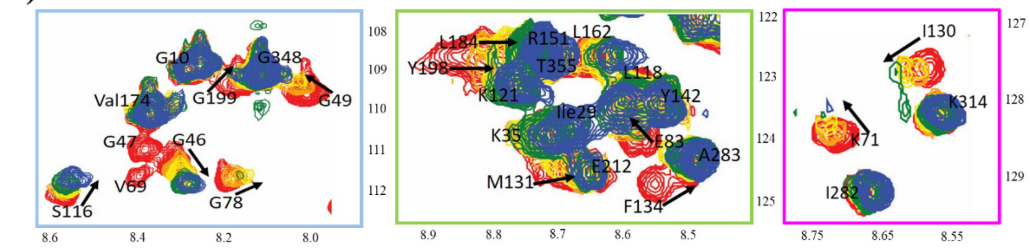


Figure 3-2 (cont.)

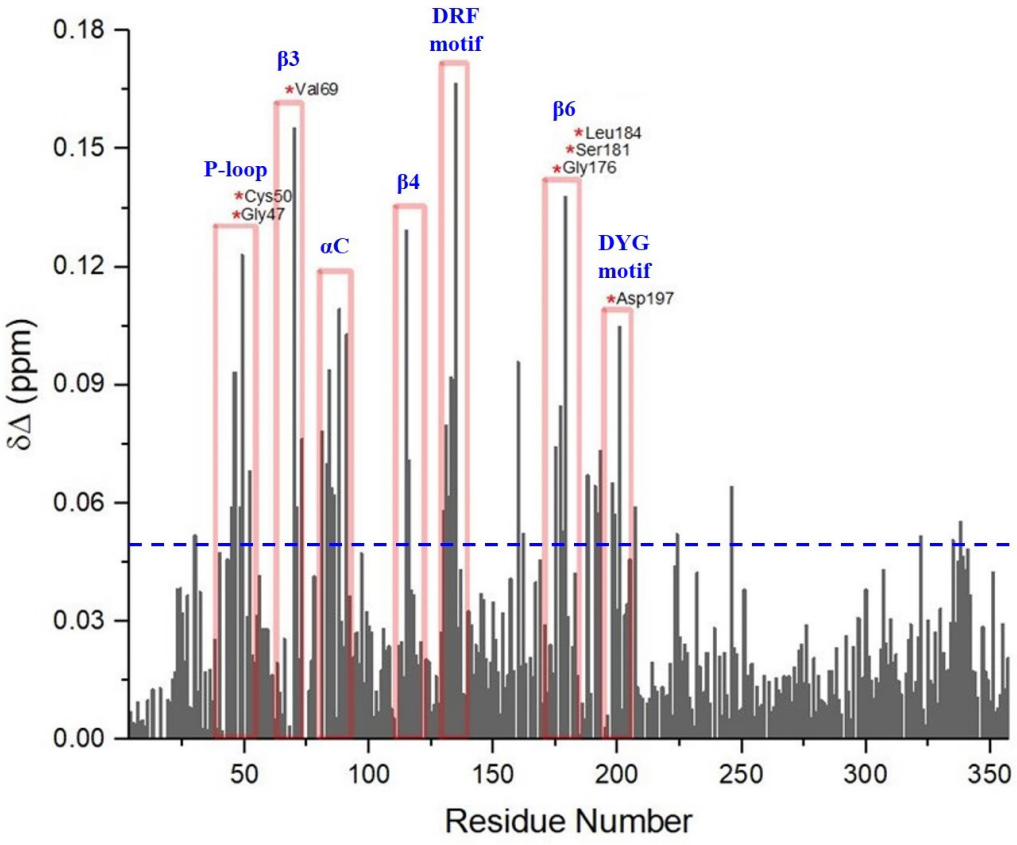


Figure 3-3 Chemical shift perturbation of VRK1$^{\Delta 35}$ with the addition of AMP-PNP.

The histogram exhibited the chemical shift changes of ^{15}N-labeled VRK1$^{\Delta 35}$ caused by the addition of AMP-PNP in the molar ratio of 1 (VRK1$^{\Delta 35}$) to 5 (AMP-PNP). The chemical shift perturbation (δΔ) plotted is calculated by the formula $\sqrt{[(\boldsymbol{\Delta\delta_H})^2 + \mathbf{0.156} \times (\boldsymbol{\Delta\delta_N})^2]}$. The blue dotted line indicates the threshold level of the chemical shift perturbation at 0.05. The residues showing chemical shift perturbations above the threshold value of 0.05 (δΔ > 0.05) are identified and highlighted in the pale red boxes, numbered with VRK1 secondary structure elements. The residues with diminished intensity are indicated by red asterisk. The residues of VRK1 are represented by 3-letter amino acid code.

The binding constant (K_D) for AMP-PNP was calculated using the one site-specific binding model on the Graphpad software. The NMR dissociation constant values were based on the chemical perturbation of 4 residues (Figure 3-4). The K_D values calculated for the four residues is shown in Table 3-1, and the average K_D calculated was 0.39 ± 0.06 mM. The interaction between VRK1 and ATP analog showed a low affinity binding which further indicated a transient interaction between the kinase and ATP.

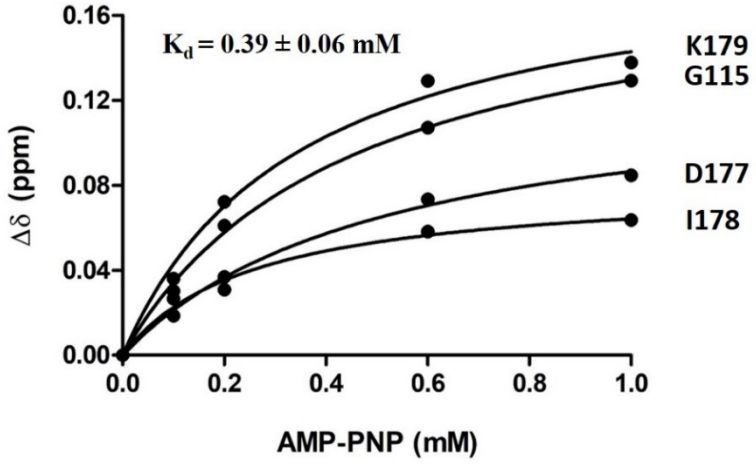


Figure 3-4 Determination of dissociation constants using NMR titration experiments.

The chemical shift perturbation (δΔ) is calculated by the formula $\sqrt{[(\boldsymbol{\Delta\delta_H})^2 + \mathbf{0.156} \times (\boldsymbol{\Delta\delta_N})^2]}$, where $\Delta\delta_H$ and $\Delta\delta_N$ are the observed chemical shift changes for ^1H and ^{15}N, respectively. For determination of dissociation constant (K_D), Δδ is plotted as a function of the concentration of AMP-PNP (mM) and the data for multiple peaks are fitted using the Graphpad – One site-specific binding method.

Table 3-1 Dissociation constants values of AMP-PNP determined for 4 residues of VRK1.

Peptides	**K_D (mM)**
G115	0.45 ± 0.05
D177	0.52 ± 0.08
I178	0.26 ± 0.06
K179	0.35 ± 0.08

The VRK1$^{\Delta 35}$ residues identified for the interaction with AMP-PNP was found to correspond to the conserved motifs of a typical kinase. The backbone amides of residues Gln45, Gly46, Gly47, Gly49 and Cys50 of the phosphate-binding loop (P-loop), (also known as the G-loop), revealed strong perturbation when titrated with AMP-PNP, as this region was known to interact with the phosphate groups of ATP to stabilize the ligand in the active site (Madhusudan et al., 2002). The P-loop sandwiched between the β1 and β2 strands (Figure 3-5A), was well-conserved among the kinase family and showed high conformational flexibility (Patel and Doerksen, 2010). In general, they adopted 'Open-conformation' and 'Closed-conformation' before and after binding of ATP, respectively to stabilize the phosphate groups. The perturbation observed in these conserved residues, Lys71 (on the β3 strand) and Glu83 (αC loop), suggested that they might be involved in the salt-bridge formation which participates in linking the β3 strand and αC loop of the N-lobe of kinase. In addition, the Lys71 could better interact with the α- and β- phosphate of the ATP and stabilize it in a position, enhancing catalysis (Johnson et al., 1996, Hanks and Hunter, 1995). The residues Met131, Asp132, Arg133, Phe134 and Gly135 of the hinge region (DRF motif) formed the bridge that linked the N- and C-lobes of the kinase (Figure 3-5A). The flexible hinge, was found to be important for the interaction with the adenosine group of ATP (Cowan-Jacob, 2006). The critical residues for VRK1$^{\Delta 35}$ and AMP-PNP interaction were identified and summarized in Table 3-2 and Figure 3-5.

Table 3-2 A summary of interacting residues identified from the NMR titration of VRK1 $^{\Delta 35}$ with AMP-PNP.

VRK1$^{\Delta 35}$ residues	**Secondary structure elements**
Gln45, Gly46, Gly47*, Gly49, Cys50*	P-loop
Val69*, Val70, Lys 71, Glu73	β3
Phe81, Glu83, Leu84, Lys85, Phe86, Gln88, Ala91	αC
Gly115, Ser116	β4
Met131, Asp132, Arg133, Phe134, Gly135	DRF motif (Hinge region)
His175, Gly176*, Asp177, Ile178, Lys179, Ser181*, Leu184*	β6
Asp197*, Tyr198, Gly199	DYG motif

The residues showing chemical shift perturbations above the threshold value of 0.05 ($\delta\Delta > 0.05$) and residues (indicated by asterisk) displaying diminished intensity are identified. The secondary structure elements of VRK1 are numbered accordingly. The residues of VRK1 are represented by 3-letter amino acid code.

The interacting residues were mapped onto the VRK1$^{\Delta 35}$ structure (PDB ID 2LAV) and have been labeled to indicate their positions (Figure 3-5B). From the NMR titration experiments, the ATP binding site on VRK1 was confirmed and corroborated to the expected canonical active site. However, these experiments could only help to identify residues / region involved in the interaction, but not the understanding of the network of interactions between ATP with VRK1. For example, the aspartate of DYG motif was known to be involved in an interaction with the magnesium ion and ATP molecule (Knighton et al., 1991). In our NMR titrations, we observed that Asp197 (in DYG motif) perturbed away when titrated with AMP-PNP, implying that it might be interacting with the AMP-PNP. However, we do not know whether Asp197 is interacting with the magnesium ion (from $MgCl_2$ added in buffer) and AMP-PNP. Hence, it is necessary to obtain the structure of VRK1 bound to AMP-PNP, to better understand the molecular basis of this interaction at atomic level.

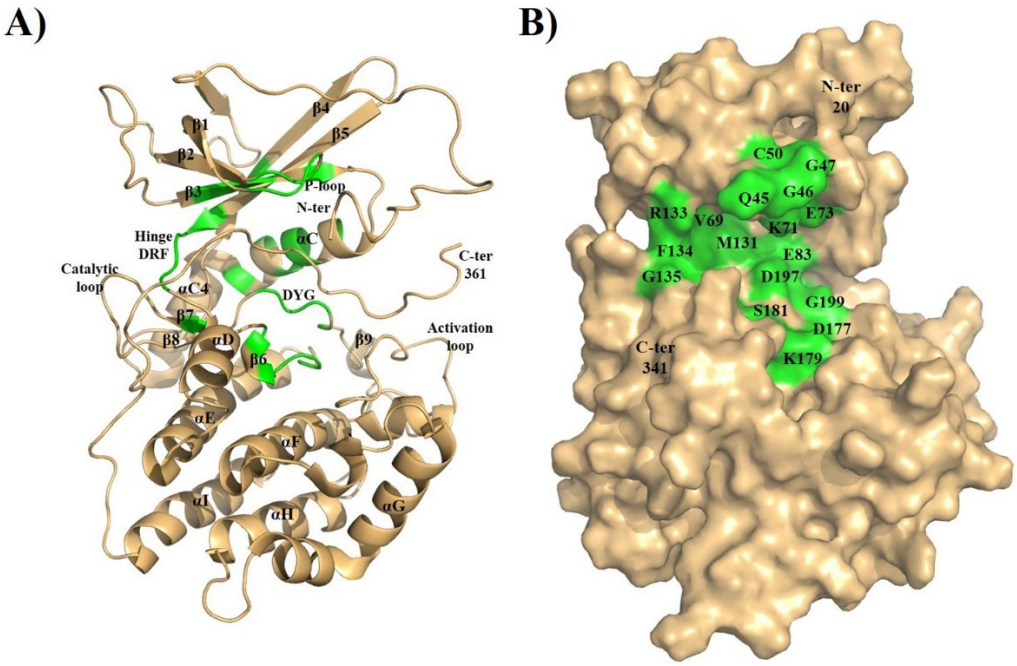


Figure 3-5 Mapping of the interacting residues on VRK1 structure.

A) VRK1$^{\Delta 35}$ structure (PDB ID 2LAV), residues 1-361, are shown in cartoon mode (pale orange), with the secondary structural features of VRK1 labeled. The regions involved in the VRK1 and AMP-PNP interaction are highlighted in green. B) VRK1$^{\Delta 35}$ structure from residues 20-341 is shown in surface mode (in pale orange). The residues from N-terminal (N-ter) and C-terminal (C-ter) ends are hidden to reveal the ATP binding site mapped onto VRK1$^{\Delta 35}$. The residues ($\delta\Delta >$ 0.05) are identified and mapped onto the surface of the VRK1 ATP-binding pocket (shown in green). The residues are labeled in single amino acid code.

3.2.3 VRK1 and AMP-PNP interaction studies by X-ray crystallization

3.2.3.1 pE-SUMO–VRK1^{cryst} construct preparation

From the NMR titration analysis, we identified the residues that are involved in the interaction between VRK1 and AMP-PNP. Next, we would like to obtain the crystal structure of VRK1 bound to AMP-PNP, to further understand their interactions on an atomic level.

Previously, we tried to crystallize VRK1^{wt} and VRK1$^{\Delta 35}$ using various crystallization screens, both in the presence and absence of ATP, but all attempts were unsuccessful. These constructs were somehow resistant to crystallization, though their expression and purification profile were good. According to literature, two approaches are usually adopted to enhance protein crystallization (a) substitution/deletion of disordered region and/or (b) mutation of charged residues (Muller, 2017). In the VRK1 construct, the presence of the dynamic C-terminus tail (residues 341-396), a disordered region might hinder crystallization. Reports have shown that charged residues, lysines (Lys) and glutamates (Glu), mostly localized on protein surface (Baud and Karlin, 1999) are less preferred at protein-protein interface (Lo Conte et al., 1999) as they hinder the protein's ability to crystallize. The Surface Entropy Reduction (SER) method is a strategy where two to three high entropy surface residues, such as Lys and Glu are replaced with the lower one like alanine (Ala). It has been shown that the substitution of Lys/Glu with Ala can reduce the entropic barrier necessary to overcome crystal packing forces (Goldschmidt et al., 2007, Derewenda, 2004). By using the SER prediction server (Goldschmidt et al., 2007),four clusters of residues on VRK1^{wt} protein were selected to be suitable for mutation to increase the chances of crystallization (Appendix V). These Lys/Glu residues were Lys34,

Lys35, Glu36, Glu212, Lys214, Glu215, Glu292, Lys293, Lys295, Lys359 and Lys360. Moreover, the flexible C-terminal tail of $VRK1^{wt}$ protein can be truncated to help in the packing of the crystals in a more ordered manner. This approach was adopted to obtain the crystal structures of the inhibitor bound complexes of VRK1 (Counago et al., 2017).

Hence, we generated the $VRK1^{cryst}$ construct as published earlier (Counago et al., 2017), with residues 365 to 396 deleted and 11 Lys/Glu residues mutated to Ala. The pUC57 plasmid containing the desired $VRK1^{cryst}$ gene was then purchased from GenScript, USA.

3.2.3.2 $VRK1^{cryst}$ protein purification

The 6xHis-SUMO tag on the $VRK1^{cryst}$ protein was first cleaved off by SUMO protease (SP cleavage). The $VRK1^{cryst}$ protein was then purified by gel filtration and the chromatogram showed that it eluted from 78-88 mL (Figure 3-6). The elution fraction was analyzed on 12% SDS-PAGE gel and a purified $VRK1^{cryst}$ protein band (M.W. 40 kDa) was observed. The $VRK1^{cryst}$ identity was confirmed using the western blot experiment, probing with anti-His primary antibody purchased from Santa Cruz (Figure 3-7).

Each litre of bacterial culture can yield 25-30 mg of $VRK1^{cryst}$ protein. From the results, it could be observed that the $VRK1^{cryst}$ construct overexpressed more soluble protein than the $VRK1^{wt}$ construct. This is an advantage as protein with better solubility is more likely to get crystallized (Muller, 2017, Abdalla et al., 2016).

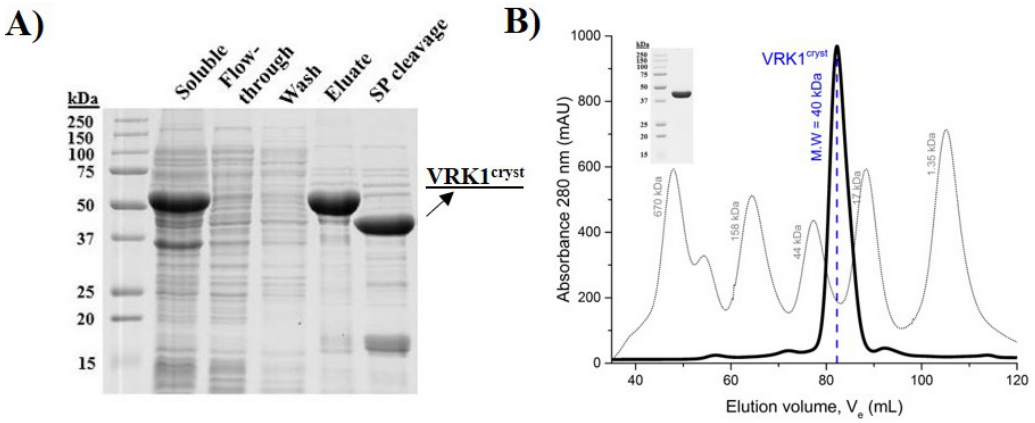


Figure 3-6 VRK1^{cryst} protein purification analysis.

A) SDS-PAGE analysis after nickel affinity chromatography. The soluble, flow-through, wash, eluate and SP cleavage fractions collected after nickel column purification are loaded along with the protein marker on a 12% SDS-PAGE gel. VRK1^{cryst} protein, M.W. 40 kDa, is observed after sumo protease (SP) cleavage (indicated by the black arrow). B) Size exclusion chromatography purification of VRK1^{cryst} by Superdex S200 16/600 column and it is observed to elute from 78-88 mL (in black). The inset on the top left showed the purified protein ran on a 12% SDS-PAGE gel after purification. The grey dotted graph represent the Bio-rad protein standards analyzed on the same column is included as a reference.

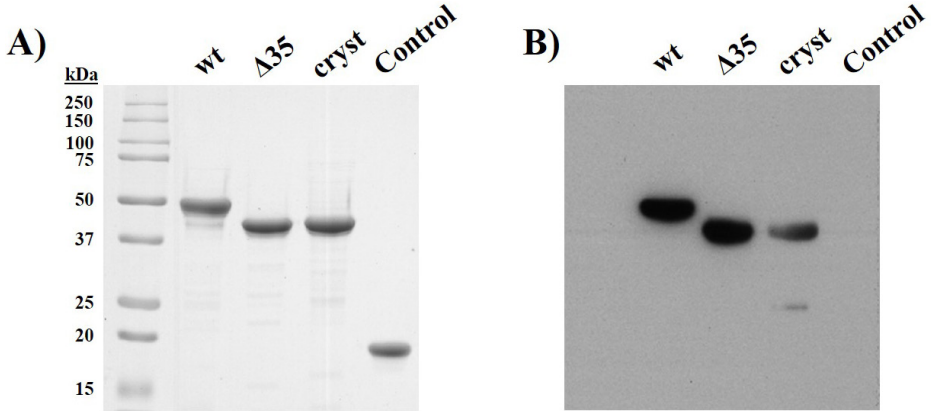


Figure 3-7 Western blot analysis of VRK1^{cryst}.

A) SDS-PAGE analysis of VRK1 proteins. In the 12% SDS-PAGE gel, VRK1^{wt}, VRK1$^{\Delta 35}$ and VRK1^{cryst} His-tagged proteins and a negative control (a non-His tagged protein, Macrodomain H2A1.2 protein, M.W. 20 kDa) were loaded along with protein marker. B) Western blot analysis of VRK1 proteins. The proteins on the SDS-PAGE gel were transferred onto a nitrocellulose membrane, then probed with anti-His primary antibody. The chemiluminescent signal was captured onto the X-ray blue film with 10 sec exposure time.

3.2.3.3 VRK1^{cryst} protein kinase activity

Both VRK1^{wt} and VRK1$^{\Delta 35}$ kinases were shown to be active by our lab (Shin et al., 2011). Here, the ADP-Glo kinase assay was used to determine the catalytic activity of the mutant kinase, VRK1^{cryst}. This experiment measures the ADP produced after kinase activity using a luminescence method. The ADP produced is converted to ATP, which is then converted to a light signal through luciferase reaction. The luminescence signal (Relative Light Unit, RLU) measured is equivalent to the ADP produced by kinase reaction. In the assay, the kinase activity is determined by the titration of the kinase with one of its substrates, ATP. Two controls, positive and negative, are included in the assay for analysis. For the positive control the functional VRK1^{wt} is used, while the negative control is performed without adding a kinase. From the graph shown in Figure 3-8, it can be observed that the mutant, VRK1^{cryst} is a catalytically active kinase, as the increase of light signal detected is equivalent to the amount of ADP produced from the kinase reaction.

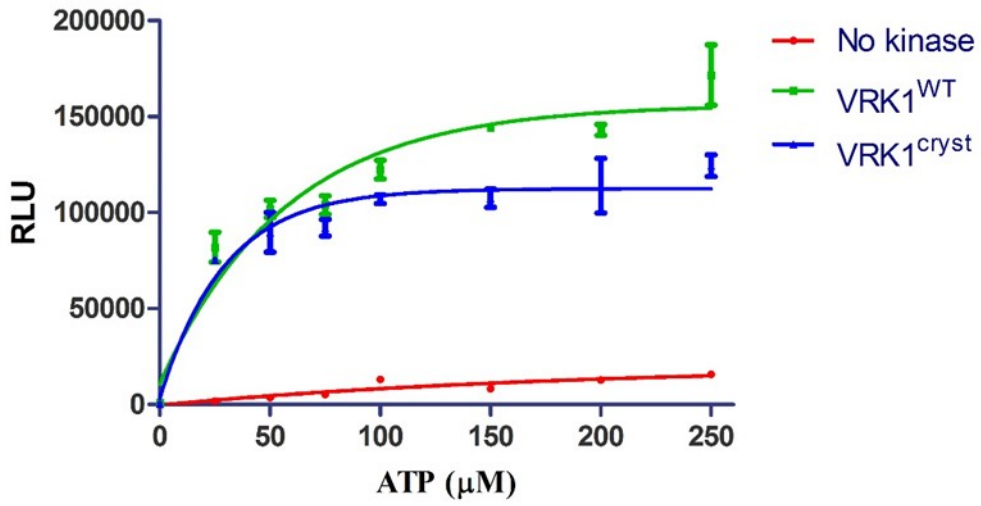


Figure 3-8 VRK1^{cryst} kinase activity.

The ADP-Glo assay shows the stimulation of VRK1^{wt} and VKR1^{cryst} kinase by ATP. The experiments are carried in triplicates and the data collected is analyzed using GraphPad Prism 5 software. The signal is measured in relative light units (RLU).

3.2.3.4 Determination of VRK1^{cryst}–AMP-PNP co-crystal structure

The VRK1^{cryst}–AMP-PNP co-crystal structure determination was performed according to the protocol described in section 2.2.18. Briefly, the crystal belongs to orthorhombic group, $P2_12_12_1$, with four molecules/chains in the asymmetric unit, was determined up to a resolution 2.07 Å. The AMP-PNP ligand was clearly observed to be attached to all four chains. The detailed refinement statistics were tabulated in Table 3-3.

Table 3-3 Crystallography data collection and refinement statistics for VRK1^{cryst}–AMP-PNP complex structure.

Data Collection	
Beamline	TPS05A
Wavelength (Å)	0.9998
Detector	MX300HS
No. of crystals	1
Crystal-Detector Distance (mm)	250
No of data sets	3
Phi Range (º)	0-180 / 90-270 / 180-360
Exposure (seconds / frame)	0.15 / 0.10 / 0.10
Oscillation / Frame (º)	0.5
Data Reduction	
Space Group	P 2_1 2_1 2_1
Unit Cell	
a; b; c (Å)	92.59; 96.84; 192.52
α = β = γ (º)	90
Resolution (Å)	66.00 - 2.07
R_{merge}	0.106 (0.701) †
Rpim	0.032 (0.215)
Unique Reflections	106113 (5210)
Mean [(I)/σ(I)]	21.3 (5.7)
Completeness	100 (100)
Multiplicity	22.4 (22.5)
CC (1/2)	0.999 (0.968)
Refinement	
Number of Reflections	105361
Resolution (Å)	20.00 – 2.07
R-Value	0.187
R-Free	0.223
No. of atoms	
Total / VRK1 atoms / AMP-PNP atoms / Water molecules / Hetatoms	11084 / 9986 / 124 / 889 / 81
Mean B-Value (Å^2)	
Total / VRK1 atoms / AMP-PNP atoms / Water molecules / Hetatoms	41.2 / 39.9 / 76.9 / 47.1 / 80.1
R.m.s.d. from ideal values	
Bond Lengths (Å)	0.010
Bond Angles (º)	1.00
Torsion Angles (º)	16.73
Ramachandran Statistics (%)	
Preferred Regions	96.8
Allowed Regions	3.2
Outliers	0

† ***Values in parenthesis correspond to those in highest resolution bin.***

3.2.3.5 Structural analysis of VRK1^{cryst}–AMP-PNP complex

The crystal structure of VRK1^{cryst} complex with AMP-PNP was solved to a resolution of 2.07Å (Figure 3-9A). The crystals belonged to the orthorhombic space group $P2_12_12_1$, with four molecules in the asymmetric unit and each chain bound to an AMP-PNP molecule and a magnesium ion. The superposition of the four chains of VRK1^{cryst}–AMP-PNP in Figure 3-9B, and the Root-mean-square-deviation (RMSD) calculated in Table 3-4, revealed that the structures of the four chains are similar to each other. As expected, the adenosine moiety of AMP-PNP docked into the active site pocket while the tri-phosphate tail was positioned towards the P-loop. The electron density map disclosed that all the atoms of AMP-PNP are well-resolved and the composite omit map showing the electron density of AMP-PNP further validated the claim (Figure 3-9C).

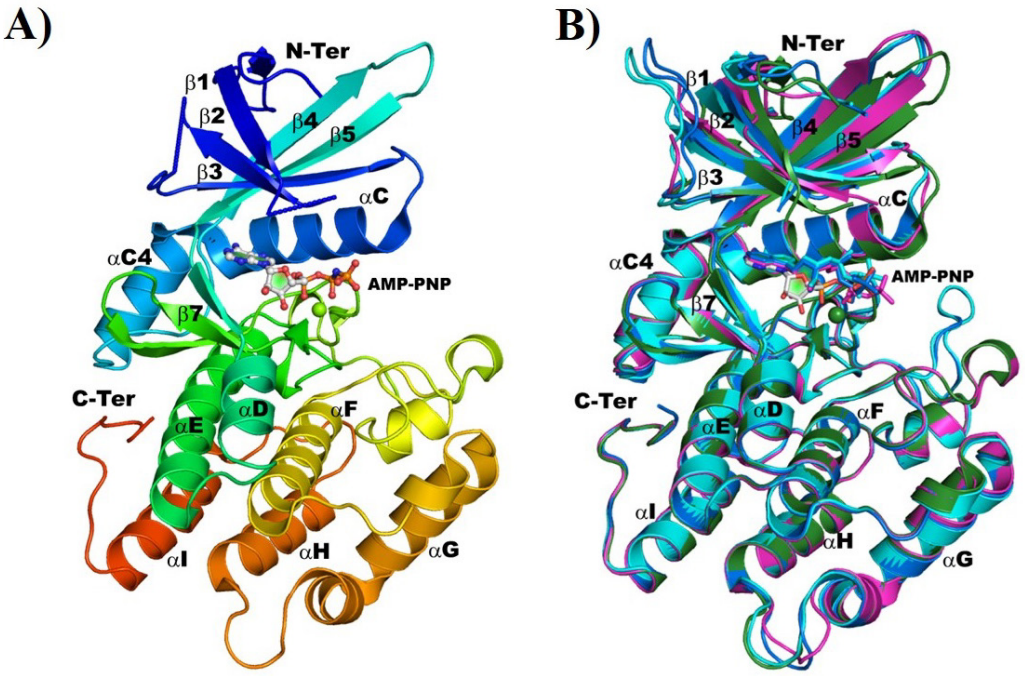


Figure 3-9 Analysis of VRK1^{cryst}–AMP-PNP crystal structure.

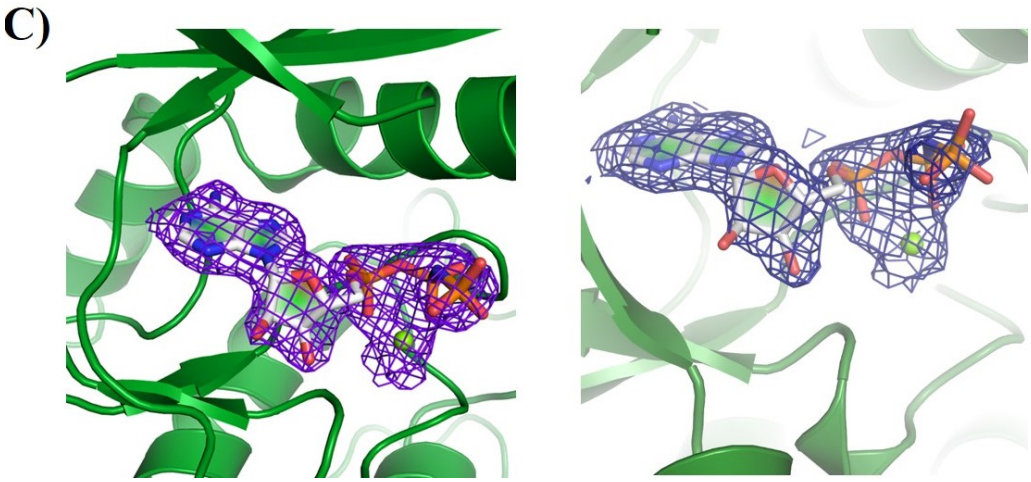


Figure 3-9 (cont.)

A) The structure of VKR1^{cryst} bound to AMP-PNP represented in cartoon, with AMP-PNP shown in ball and stick mode and the magnesium ion (Mg^{2+}) displayed as a green sphere. B) Superposition of the VRK1^{cryt}-AMP-PNP chain A, B, C and D structures shown in cartoon mode. The 4 chains, A, B, C and D are colored in marine blue, cyan, light magenta and green respectively. The AMP-PNP (shown in stick mode) for chain A, B, C and D are shown in marine blue, cyan, light magenta and white color. The magnesium ion is displayed as a green sphere. The secondary structure elements are numbered for reference. C) The 2Fo-Fc electron density map (left) and the composite omit map (right), contoured at 1σ cut-off, exhibiting a well resolved electron density for AMP-PNP (displayed in ball and stick mode) and the coordinated Mg^{2+} ion (shown as a green sphere).

Table 3-4 Superposition of the four chains of VRK1^{cryst}–AMP-PNP complex structures and their RMSD values.

	Chain A	Chain B	Chain C	Chain D
Chain A	-	0.174 (249 C^{α} atoms)	0.380 (260 C^{α} atoms)	0.471 (245 C^{α} atoms)
Chain B		-	0.438 (251 C^{α} atoms)	0.556 (245 C^{α} atoms)
Chain C			-	0.175 (233 C^{α} atoms)
Chain D				-

The hydrogen bonds and nonbonded interactions stabilizing the AMP-PNP in the VRK1 active site were tabulated and illustrated in Table 3-5 and Figure 3-10A respectively. The nitrogen atoms (N1/N6) of adenosine of AMP-PNP form

hydrogen bond with backbone oxygen and nitrogen of Asp132 and Phe134, respectively. The oxygen atoms (O1A/O2A and O1B/O2B/O2G) of the phosphate tail are hydrogen bonded to the side chain atoms of Lys71 (NZ) and Asp197 (OD2), respectively (Table 3-5). In addition, Ser181 formed a hydrogen bond with the ribose sugar oxygen in chain D, but not in others. Apart from these, several non-bonded contacts formed by residues Ile51, Val69, Phe134 and Leu184 with the adenosine and ribose rings, secure the AMP-PNP's adenosine group in the active site. Overall, the DRF motif interact with the adenine ring, while the DYG motif interacts with the phosphate groups of AMP-PNP, linked by a magnesium ion. Comparing the AMP-PNP orientation within the four chains in the asymmetric unit, the adenosine moiety adopts similar position, while the phosphate tail adopts similar orientation in all three chains (A, B, C), except in D chain as shown in Figure 3-10B (left inset). This could be attributed to the magnesium ion position, which makes ionic interaction with the neighboring residues at approximate distances 2.8Å (Figure 3-10B – middle inset), in chain A, B, C, whereas the magnesium ion bound to chain D forms coordination bonding (at distances of 2.1Å) with the neighboring residues (Figure 3-10B – right inset). In our structure, the electron density of the residues of P-loop was not detected, probably due to the flexibility of the region (Patel and Doerksen, 2010). Nevertheless, the trace was seen to mimic the P-loop orientation in the inhibitor bound structure (PDB ID 5UKF). In addition, the electron density for the C-terminal tail residues (343-364) was not observed, owning to the flexibility of this region in the presence of AMP-PNP, similar to the inhibitor-bound VRK1 structures (Counago et al., 2017).

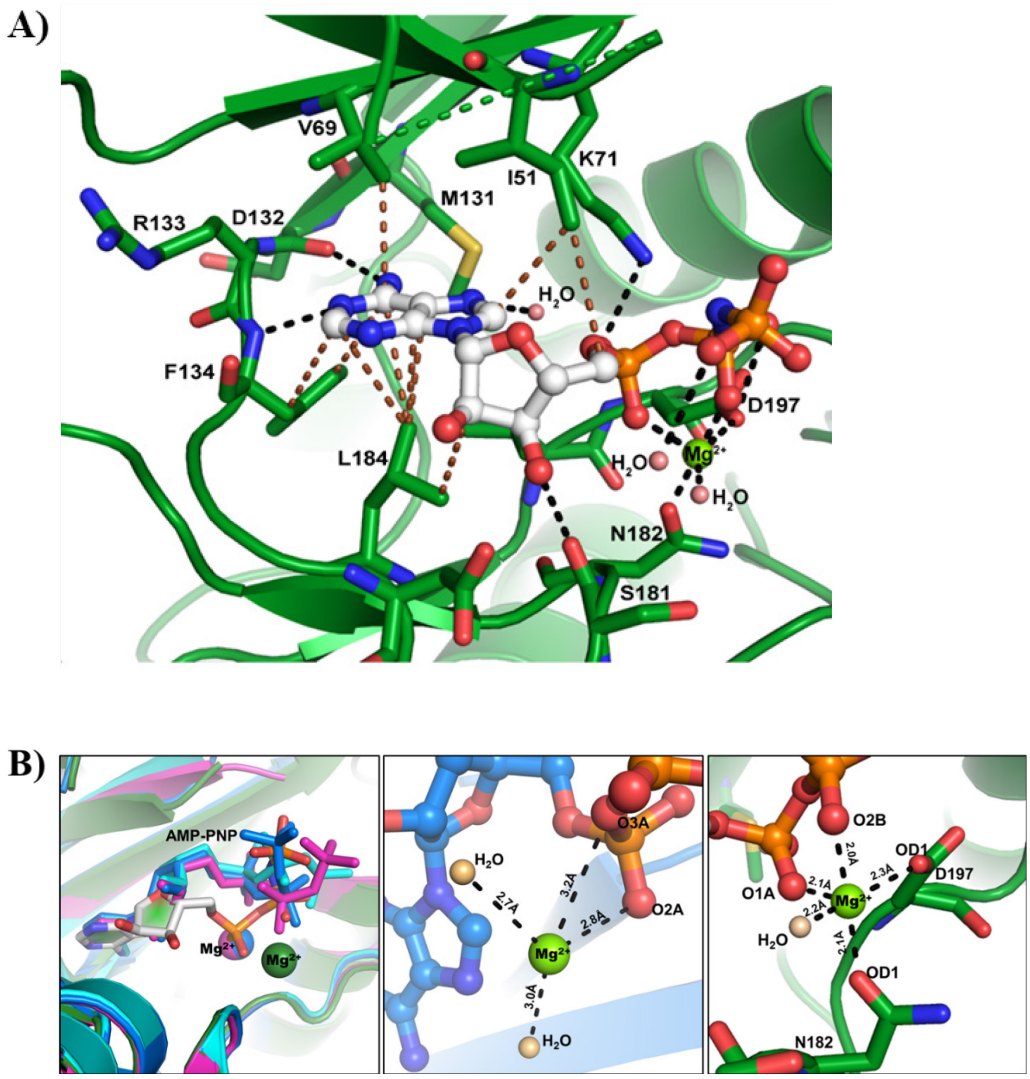


Figure 3-10 Interaction map of AMP-PNP in VRK1^{cryst} active site.

A) Interactions formed by AMP-PNP (shown in ball and stick mode) and the residues at the active site of VRK1^{cryst} (shown in stick mode) are shown. The hydrogen bonds and non-bonded contacts are shown as black and brown dotted lines, respectively. The coordination geometry formed between the Mg^{2+} ion (shown as a green sphere) with the AMP-PNP phosphate groups and the nearby D197 and N182 is shown. The amino acids are represented in single letter code. B) In the left panel, the orientations of AMP-PNP (shown in stick mode) and Mg^{2+}ion (shown in sphere) for all the four chains are showed, and the AMP-PNP in chain A, B, C and D is represented in marine blue, cyan, light magenta and white color respectively. The interactions formed between Mg^{2+} ion (shown as a green sphere) with the AMP-PNP phosphate groups and the nearby water molecules (shown in light orange sphere) is shown in the middle panel. Since, the orientation of the AMP-PNP in chain A, B and C are similar, the interaction map in chain A is presented as a representative. The coordination geometry formed between the Mg^{2+} ion (shown as a green sphere) with the AMP-PNP phosphate groups and the nearby D197 and N182 in chain D is shown in the right panel. The amino acids are represented in single letter code.

Table 3-5 Interactions made by VRK1^{cryst} with AMP-PNP (ANP).

Hydrogen bonds (Chain A, B, C, D)		
ANP Atom	VRK1/ Mg^{2+} Atoms	D-A Distance (Å)
ANP N6	Asp 132 O	2.8 ± 0.1
ANP O1A/O2A *	Lys 71 NZ	3.5 ± 0.3
ANP O1B/O2B/O2G *	Asp 197 OD2	2.9 ± 0.2
ANP N1	Phe 134 N	2.9 ± 0.1
ANP O3'	Ser 181 O	3.9 ± 0.7
ANP N7	H_2O O	2.8 ± 0.1
ANP O3'	H_2O O	4.5 ± 0.8
Mg interaction (in Chain A, B and C)		
ANP O1A	Mg^{2+}	2.7 ± 0.2
ANP O3A	Mg^{2+}	3.7 ± 0.5
Mg coordination (in Chain D)		
Mg^{2+}	ANP O1A	2.1
Mg^{2+}	ANP O2B	2.0
Mg^{2+}	Asp 197 OD1	2.3
Mg^{2+}	Asn 182 OD1	2.1
Mg^{2+}	H_2O O	2.2
Nonbonded contacts		
ANP Atom	VRK1 Residues	
ANP C2	Phe134, Leu184	
ANP C4	Leu184	
ANP C5	Leu184	
ANP C6	Val69, Asp132, Phe134, Leu184	
ANP C8	Ile51	
ANP C5'	Ile51	
ANP C2'	Leu184	
ANP PA	Ile51	
ANP O1A	Ile51, Lys71	
ANP O2B	Asp197	
ANP O2'	Leu184	
ANP O4'	Ile51	
ANP O5'	Ile51	
ANP N1	Asp132, Arg133, Phe134	
ANP N6	Val69, Met131, Asp132, Phe134	

* The phosphate (PO_4) oxygens nearest to the corresponding residues were considered for the measurement (since the PO_4 orientation varies in all the 4 chains, as they are known to be flexible part of ATP, hence these changes are expected).

3.2.3.6 Comparison of VRK1^{cryst}–AMP-PNP structure with the apo- and inhibitor bound VRK1 complex

For comparison, the VRK1^{cryst}–AMP-PNP complex is superposed onto the apo (PDB ID 2LAV) and inhibitor bound VRK1 (PDB ID 5UKF) structures (Figure 3-11) and showed a RMSD of 1.70Å (270 equivalent C^{α} atoms) and 0.56Å (243 equivalent C^{α} atoms), respectively (Table 3-6).This implies that our structure is more similar to the inhibitor bound structure in comparison to the apo form.

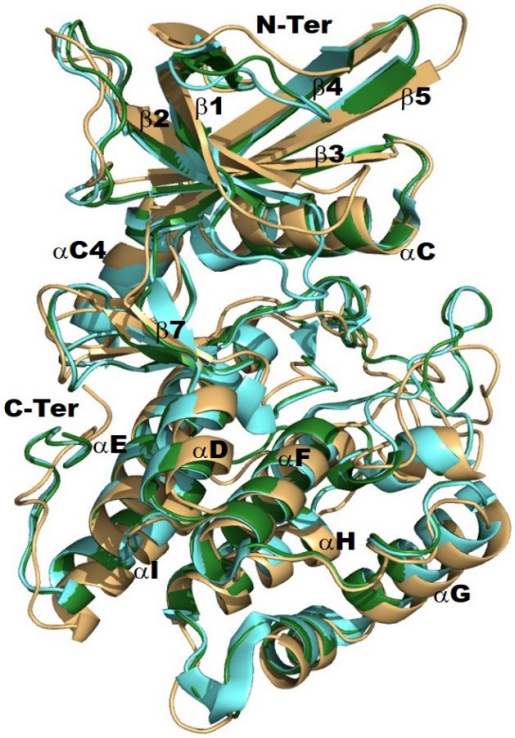


Figure 3-11 Superposition of VRK1^{cryst}–AMP-PNP complex on Apo and inhibitor bound VRK1 structures.

A cartoon representation of the structural alignment of Apo VRK1 (PDB ID 2LAV) (in pale orange) with the crystal structure of VRK1^{cryst}–AMP-PNP (in green) and inhibitor bound VRK1 (PDB ID 5UKF) (in cyan).

Table 3-6 Superposition of VRK1^{cryst}–AMP-PNP complex on Apo and inhibitor bound VRK1 structures and their RMSD values.

	AMP-PNP	**Inhibitor**
Apo	1.70Å (270 C^{α} atoms)	1.74Å (292 C^{α} atoms)
AMP-PNP		0.56Å (243 C^{α} atoms)

The comparison of the hydrogen bond and non-bonded interactions made by VRK1^{cryst} with AMP-PNP and inhibitor (GW297361X) is shown in Table 3-7. The chemical structures of the ligands are shown in Figure 3-12. For illustration, the interaction map of AMP-PNP or GW297361X in VRK1 active site are shown in Figure 3-10A and Figure 3-13.

Table 3-7 Comparison of interactions made by VRK1 with AMP-PNP (ANP) and GW297261X (8E1) structure (PDB ID 5UKF).

Hydrogen bonds					
VRK1-AMP-PNP			VRK1- GW297361X		
ANP Atom	VRK1/ Mg^{2+} Atoms	D-A Distance (Å)	8E1 Atom	VRK1 Atoms	D-A Distance (Å)
ANP N6	Asp 132 O	2.8	8E1 N2	Asp 132 O	3.1
ANP O2A	Lys 71 NZ	3.4	8E1 O	Phe 134 N	2.9
ANP O2B	Asp 197 OD2	3.3	8E1 N3	Asp 137 N	3.1
ANP N1	Phe 134 N	2.9			
ANP O3	Ser 181 O	2.9			
ANP O1A	Mg^{2+}	2.8			
ANP O3A	Mg^{2+}	3.1			
ANP N7	H_2O O	2.8			
ANP O3	H_2O O	3.3			
Mg coordination (in Chain D alone)					
Mg^{2+}	ANP O1A	2.1			
Mg^{2+}	ANP O2B	2.0			
Mg^{2+}	Asp 197 OD1	2.3			
Mg^{2+}	Asn 182 OD1	2.1			
Mg^{2+}	H_2O O	2.2			
Nonbonded contacts					
VRK1-AMP-PNP			VRK1- GW297361X		
ANP Atom	VRK1 Residues		8E1 Atom	VRK1 Residues	
ANP C2	Phe134, Leu184		8E1 C	Phe134, Gly135	
ANP C4	Leu184		8E1 C1	Gly135	
ANP C5	Leu184		8E1 C2	Gly135	
ANP C6	Val69, Asp132, Phe134, Leu184		8E1 C5	Gly135	
ANP C8	Ile51		8E1 C6	Leu184	
ANP C5’	Ile51		8E1 C7	Leu184	
ANP C2’	Leu184		8E1 C9	Val69, Phe134	
ANP PA	Ile51		8E1 C10	Met131, Phe134	
ANP O1A	Ile51, K71		8E1 C14	Phe48	
ANP O2B	Asp197		8E1 C15	Phe134	
ANP O2’	Leu184		8E1 S	Phe48, Ile51	
ANP O4’	Ile51		8E1 O	Val69, Arg133, Phe134	
ANP O5’	Ile51		8E1 N1	Val196	
ANP N1	Asp132, Arg133, Phe134		8E1 N2	Val69, Asp132, Arg133, Phe134	
ANP N6	Val69, Met131, Asp132, Phe134		8E1 N3	Ser136, Asp137, Lys140	

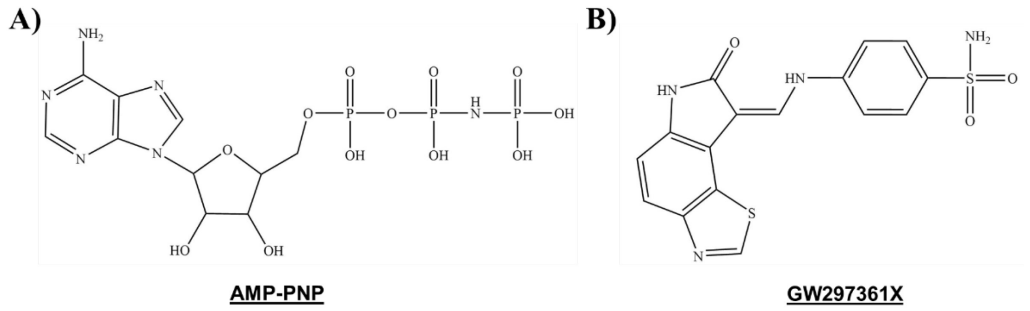


Figure 3-12 Chemical structure of AMP-PNP and GW297361X.

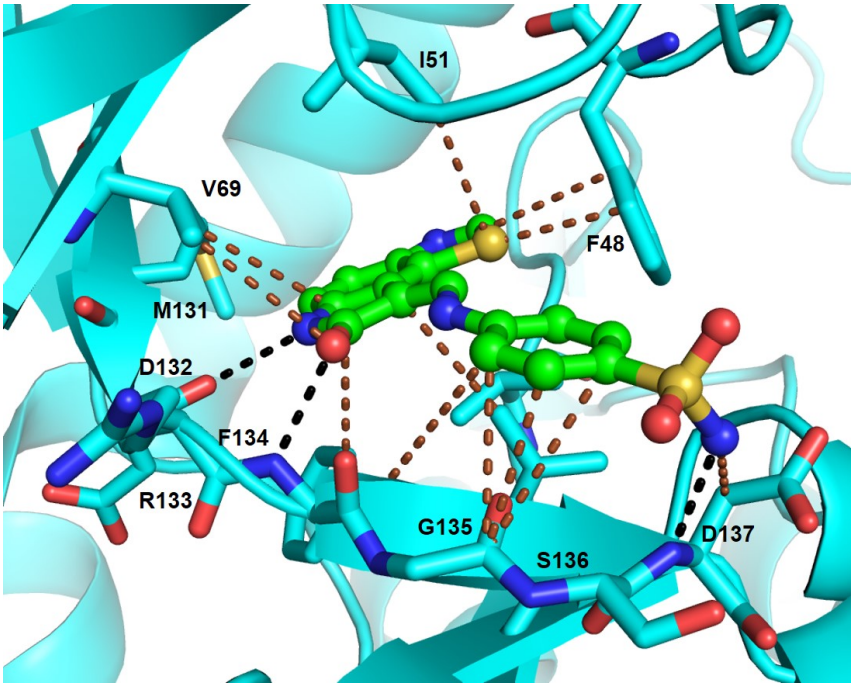


Figure 3-13 Interaction map of GW297361X in VRK1 active site.

Interactions formed by the GW297361X (shown in ball and stick mode) and the residues at the active site of VRK1 (shown in stick mode). The hydrogen bonds and non-bonded contacts are shown as black and brown dotted lines, respectively. The amino acids are represented in single letter code.

Next, comparative analysis of these structures was carried out, with special emphasis on the signature regions of a kinase domain, as given below.

a) P-loop

It was observed that the P-loop of the apo and inhibitor bound forms adopted "Open" and "Close" conformations, respectively (Figure 3-14A). Although the residues of the P-loop of AMP-PNP bound complex were not traced (Figure 3-14B), it could be observed that the arrangement of the residues around this missing loop were similar to that of the inhibitor bound conformation. Hence, it can be speculated that the P-loop in the AMP-PNP bound complex might also adopt a "Closed-conformation", though this needs further investigation. The P-loop of an active kinase usually interacts with the phosphate groups of ATP molecule (Madhusudan et al., 2002). In all the VRK1 crystal structures, it was observed that the P-loop is a highly flexible region, and the electron density of the P-loop could only be traced in one of the chain in the inhibitor bound VRK1 complex (in chain A of PDB ID 5UKF).

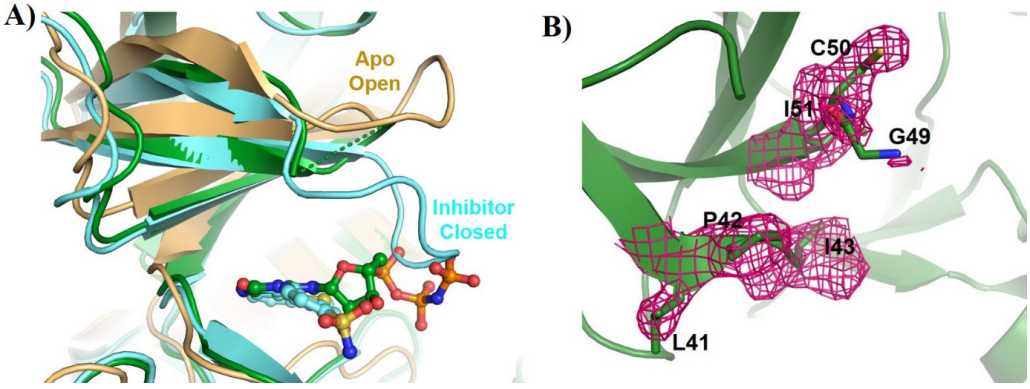


Figure 3-14 Comparison of P-loop conformation in AMP-PNP or inhibitor – bound and Apo VRK1 structures.

A) The P-loop of the Apo VRK1 structure is oriented in an open conformation (Apo open) but observed as a closed conformation in the inhibitor structure (Inhibitor closed). The ligands are shown in ball and stick mode. B) The composite omit map showing the lack of electron density for the atoms of the P-loop residues (G44 to F48). The amino acids are represented in single letter code.

b) Salt bridge between Lys71 and Glu83

A distinct feature of an active kinase is the formation of a salt bridge between the lysine and glutamate, which bridges the β3 strand and αC helix, providing structural stability. The lysine residue can come in contact with the phosphate groups of ATP to stabilize and position it prior to catalysis (Hanks and Hunter, 1995). In VRK1^{cryst}–AMP-PNP bound structure, Lys71 and Glu83 interacts significantly tighter (2.79 ± 0.01 Å) than that observed in the GW297361-VRK1 (3.19 ± 0.69 Å). In the apo NMR ensemble structure, this salt bridge distance was 7 ± 2.18 Å (Table 3-8, Figure 3-15A). Hence, the presence of ligand in VRK1 can enhance the salt bridge formation by Lys71 and Glu83.

Table 3-8 Comparison of the salt bridge distances of VRK1^{cryst}–AMP-PNP complex with inhibitor bound VRK1 and Apo structure.

Distance between the Nε of Lys71 and O of Glu83	
VRK1^{cryst}–AMP-PNP	2.79 ± 0.01Å
Inhibitor (PDB ID 5UKF) (GW297361X bound VRK1)	3.19 ± 0.69Å
Apo (PDB ID 2LAV) (20 structures)	7.00 ± 2.18Å

*The apo structure is an NMR structures with 20 ensembles. The distance corresponds to the average and its standard deviation. The minimum (2.6Å) and maximum (10.4Å) distances are also shown for reference.

c) DRF and DYG motifs

The DRF (Figure 3-15B) and DYG (Figure 3-15C) motifs were observed to have rather similar orientations in the AMP-PNP and inhibitor bound complexes which is in contrast with the apo structure, where some of the residues' side-chains were oriented differently. Nonetheless, it is observed that all the three structures followed a 'DYG-in' orientation. Moreover, it can be observed that the α-C helix of the apo form is adopting an 'out-like' conformation as demonstrated by that the distance between the Glu83 and Tyr198 (of DYG motif) at 7.2Å, which opposed the ligand bound conformation (at 4.7Å), implying a 'in' conformation adopted (Kooistra et al., 2016). This was identified using the KLIFS webserver (http://klifs.vu-compmedchem.nl).

d) Met131 residue

The arrangement of Met131, located next to the DRF motif, also showed flexibility depending on the ligand bound in its active site. In the apo structure, the Met131 side chain flips toward the active site, in contrast with the inhibitor bound form, as it forms non-bonded interactions with the ligand. It is obvious that the inhibitor was positioned deeper into the pocket opposed to AMP-PNP, and it has to reorient the Met131 further away when compared to the AMP-PNP complex (Figure 3-15D). A similar flexibility was observed when comparing the VRK2 apo structure with its corresponding inhibitor bound complex. These Met residues located in the active site of kinases are highly conserved among the Ser/Thr kinase family (Veredas et al., 2017), and more studies are required to better understand their functional roles in VRK1.

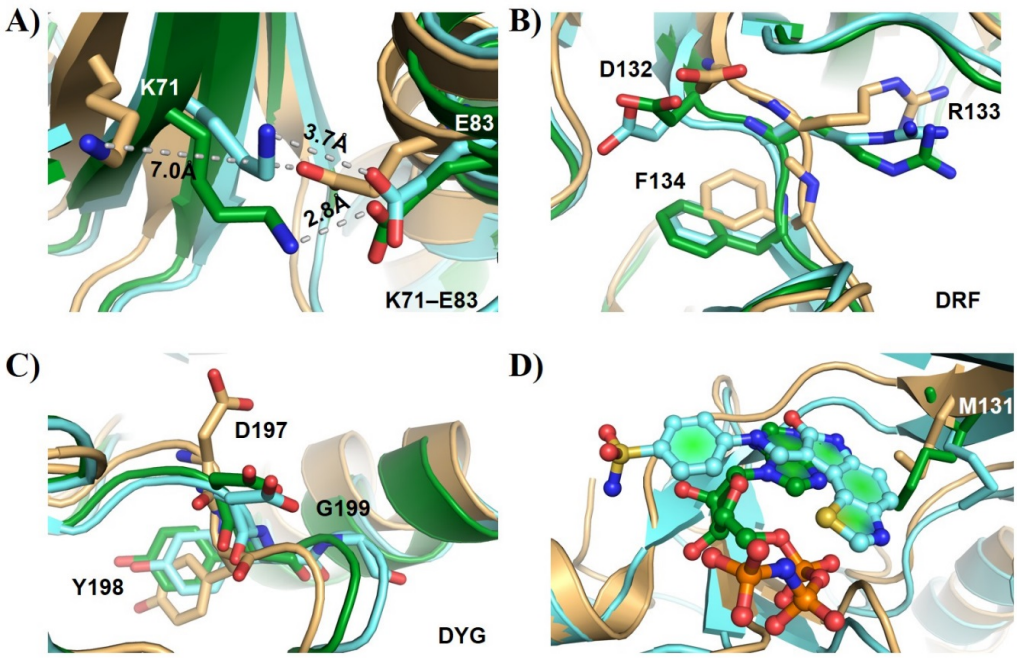


Figure 3-15 Comparison of salt bridge, DRF, DYG and M131 orientations in AMP-PNP or inhibitor – bound and Apo VRK1 structures.

A) The salt bridge formation between the K71 and E83 (residues' side chain shown in stick mode) is observed in the AMP-PNP (2.8Å) and inhibitor (3.7Å) structures but not in the apo form (7.9Å). The distances between K71 and E83 is displayed as yellow dotted lines. B, C) The orientation of the DRF and DYG motifs (in stick mode) are similar in both the AMP-PNP and inhibitor bound structure but differ in the Apo form. D) Residue M131 (displayed in stick mode) is observed to have a flexible orientation which depend on the type of ligand bound in the active site. The inhibitor bound structure which has its ligand located deeper into the pocket, has its Met131 oriented furthest away from the active site opposing the AMP-PNP bound and Apo form. The ligands are shown in ball and stick mode. In (A)-(D), the amino acids are represented in single letter code.

In summary, from the comparison of these structures, it is observed that (i) the P-loop might adopt a "Closed- conformation" similar to inhibitor-bound, though this needs further investigation; (ii) the presence of ligands can enhance the Lys71-Glu83 salt bridge formation; (iii) the DYG adopts an 'DYG-in' conformation and (iv) the inhibitor which docks deeper into the active site than the AMP-PNP, results in a change in the orientation of Met131 side chain to accommodate the ligand.

3.2.3.7 Insights into inhibitor design from VRK1^{cryst}–AMP-PNP structure

3.2.3.7.1 Comparison with VRK1-bound inhibitor complex

From the superposition and comparison of the AMP-PNP bound VRK1^{cryst} complex with the inhibitor bound VRK1 (GW297361X – PDB ID 5UKF) discussed earlier, a few moieties of the inhibitor are observed to be distinctive from AMP-PNP and can provide insights in designing more potent and specific inhibitors for VRK1. The docking of AMP-PNP and the inhibitor (GW297361X) in VRK1 active site are shown in Figure 3-16A and B respectively.

a) The thiazole moiety of the inhibitor is observed to insert deeper into the active site pocket, in contrast to the adenosine ring of the AMP-PNP molecule, which is a contributing feature of a good ATP competitive inhibitor (Figure 3-16A and B). In addition, the sulfur-π interaction with Phe48 aromatic sidechain was also indicated as a stabilizing factor (Counago et al., 2017). However, there is no interaction observed between the nitrogen atom in the thiazole ring with the neighboring atoms, which is located 4.2Å away from the nearest Asp197 (part of DYG motif) backbone (site S1 shown in Figure 3-16D). Hence, the modification to this thiazole nitrogen, which enable the formation of hydrogen bond with Asp197 can helps increase the inhibitor's binding affinity.

b) The oxindole ring of the inhibitor mimics the interactions made by the adenosine group of AMP-PNP, where it can interact with the Asp132 and Phe134 through hydrogen bonding and with Ile51 by forming non-polar contacts (site S2 shown in Figure 3-16D).

c) The benzene sulfonamide moiety of the inhibitor is observed to be stabilize by fewer interactions when compared with the equivalent ribose and

phosphate groups in AMP-PNP (site S3 shown in Figure 3-16D). Hence, modifications to this region can aid in improving the interactions with the side chain of Asp197 and residue Ser181, similar to AMP-PNP (Figure 3-16C). Moreover, the application of Serine specific war-heads can increase the specificity of inhibitor construct for VRK1.

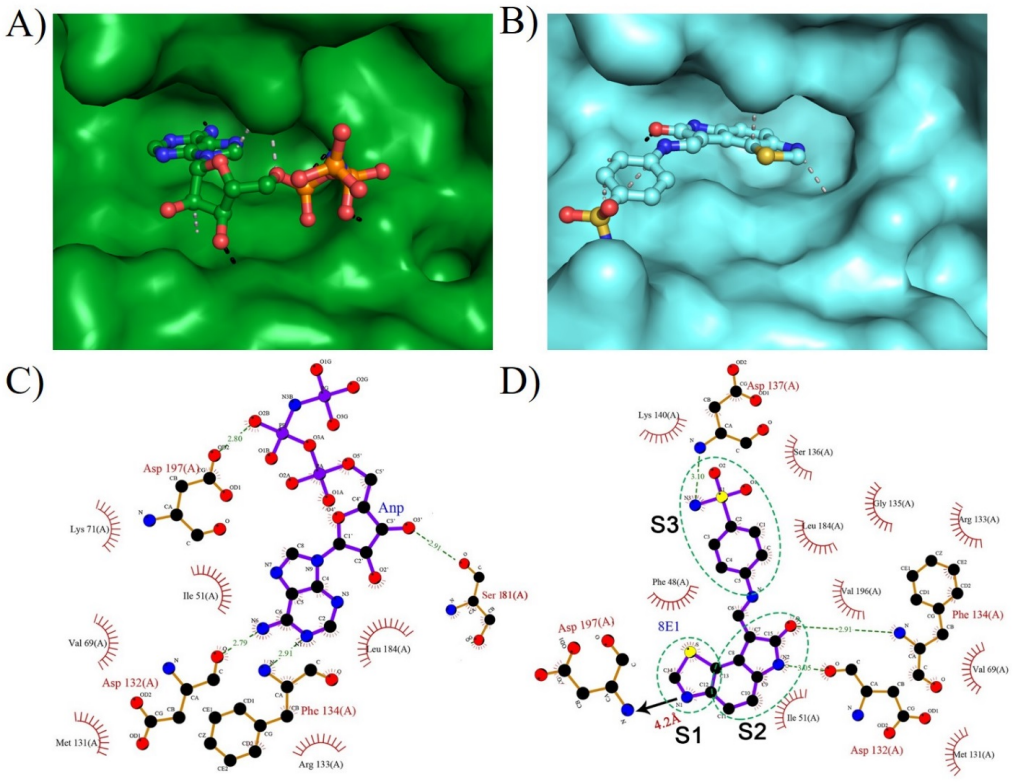


Figure 3-16 Comparison of AMP-PNP-and inhibitor-bound VRK1 structures

A, B) The surface representation showing the docking of AMP-PNP (in green) (A) and inhibitor (GW297361X, in cyan) (B) in the active site of VRK1. The P-loop of the inhibitor-bound structure (B) is removed for clarity. It is observed that the inhibitor (B) is located deeper in the active site in comparison with AMP-PNP. The hydrogen bonds and non-bonded contacts formed by the AMP-PNP and inhibitor at the active site are indicated black and grey dotted lines, respectively. C, D) Illustration of the hydrogen bonds (shown in green dotted lines) and non-polar contacts made by AMP-PNP (C) and inhibitor (D) with the residues at the active site of VRK1. The interactions can be grouped into three moieties (denoted by S1, S2 and S3) in the inhibitor. The thiazole moiety nitrogen (S1) is approximately 4Å away from the backbone of Asp197, when a region corresponding to that in AMP-PNP is missing. The oxindole ring of the inhibitor (S2) mimics the contacts made by the adenosine ring in AMP-PNP, with residues Ile51, Asp132 and Phe134. The benzene sulfonamide moiety of the inhibitor (S3) is substituted by the ribose and phosphate groups in AMP-PNP and is stabilized by more interactions compared to the inhibitor.

3.2.3.7.2 Comparison with VRK2 and Aurora B kinase structures

Next the superposition and comparison of the AMP-PNP bound VRK1^{cryst} complex with VRK2 (from VRK family) and Aurora B (other mitotic kinase), was carried out to provide more clues in the development of VRK1 inhibitor with higher specificity.

The superposition of AMP-PNP bound VRK1^{cryst} structure onto the inhibitor bound VRK2 (PDB ID 5UU1) and AMP-PNP bound Aurora B (PDB ID 4C2W) structures, revealed an RMSD value of 0.87Å (239 equivalent Cα atoms) and 4.75Å (171 equivalent Cα atoms) respectively. From the results, it can be observed that VRK1 and VRK2 are similar, while VRK1 and Aurora B showed that they have some differences and the most distinct different is the presence of αC4 helix in VRK1 (Figure 3-18). The αC4 helix is unique to the VRK family and is believed to be important for maintaining VRK protein in a closed conformation, which is a feature of an active kinase (Scheeff et al., 2009) (Figure 3-17). The sequence alignment of VRK1 with VRK2 and Aurora B are shown in Figure 3-17 and Figure 3-18 respectively. It is observed that the signature motifs of an active kinase are highly conserved in all three kinases. The residues involved in interaction with the ligand in VRK1, VRK2 and Aurora B are indicated in blue, orange and green asterisks respectively.

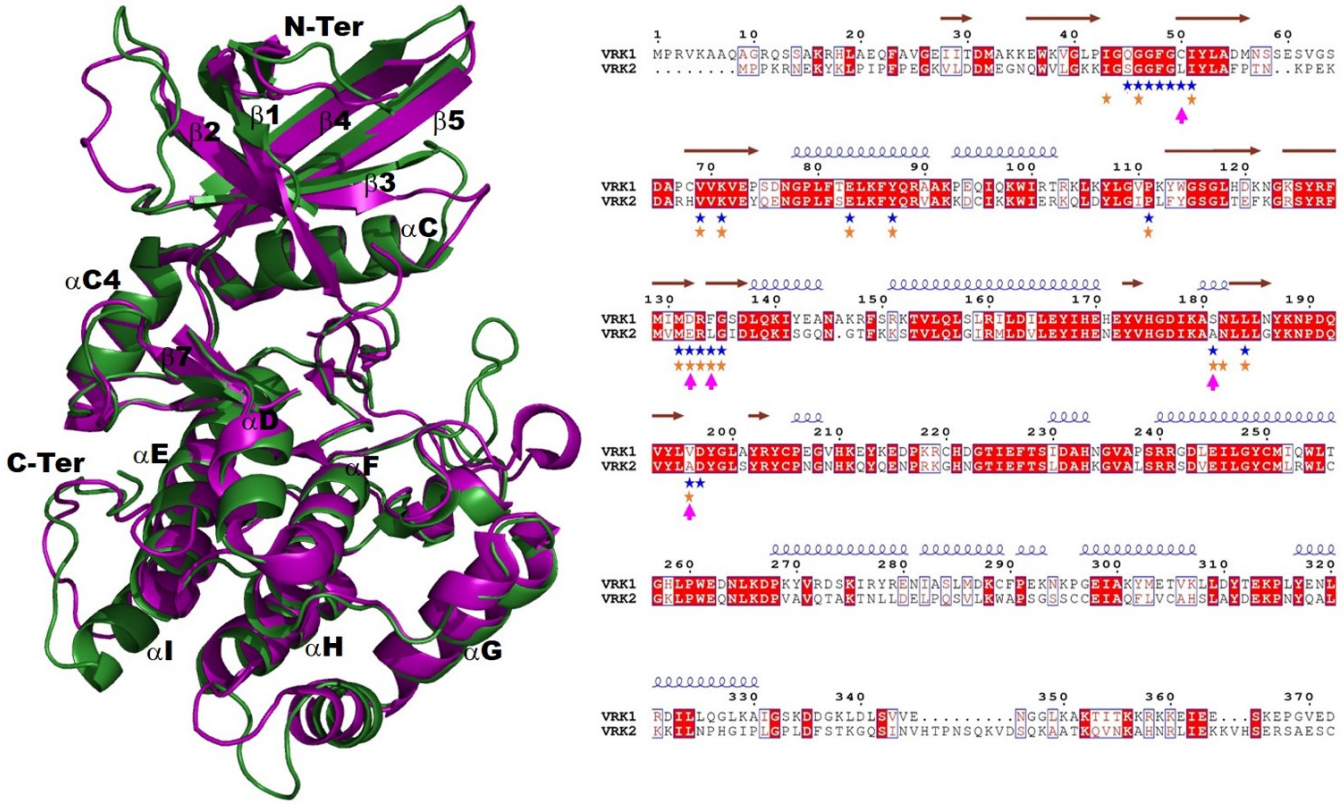


Figure 3-17 Structure alignment of VRK1 and VRK2.

In the left panel, the superposition of VRK1^{cryst}-AMP-PNP (in green) and VRK2-inihibitor bound (in purple) (PDB ID 5UU1) in cartoon representation, revealed an RMSD value of 0.87Å (239 equivalent Cα atoms). The secondary structure elements are labeled with reference to VRK1. In the right panel, sequence alignment of VRK1 and VRK2 using the Clustal omega and ESPript is shown. The secondary structure corresponding to VRK1 is shown above the alignment. The residues involved in ligand binding in VRK1 and VRK2 are marked by blue and orange asterisk, respectively. The five unique residues in VRK1 and the corresponding residues in VRK2 are indicated by magenta arrows.

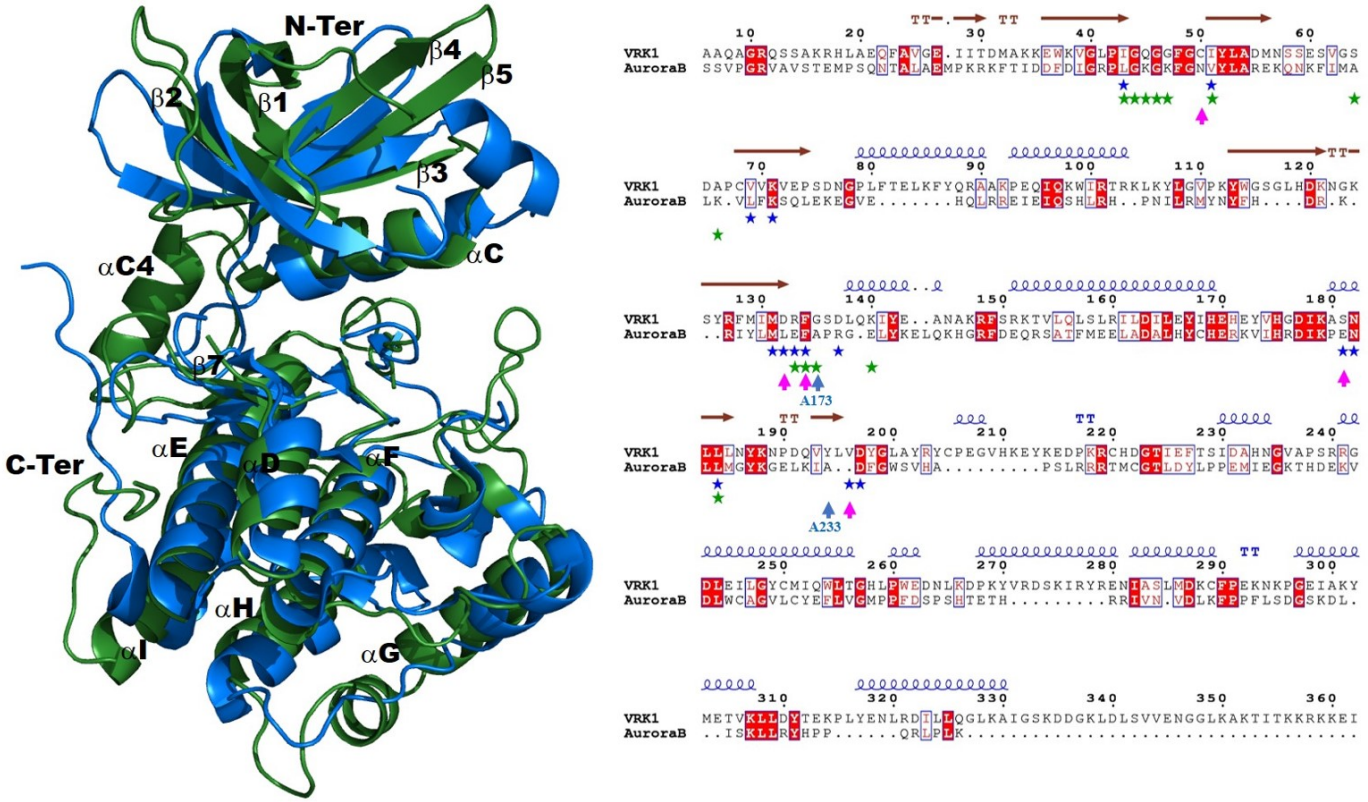


Figure 3-18 Structure alignment of VRK1 and Aurora B.

In the left panel, the superposition of VRK1^{cryst}-AMP-PNP (in green) and Aurora-AMP-PNP bound (in blue) (PDB ID 4C2W) in cartoon representation, revealed an RMSD value of 4.75Å (171 equivalent Cα atoms). The secondary structure elements are labeled with reference to VRK1. In the right panel, sequence alignment of VRK1 and Aurora B using the Clustal omega and ESPript is shown. The secondary structure corresponding to VRK1 is shown above the alignment. The residues involved in ligand binding in VRK1 and Aurora B are marked by blue and green asterisk, respectively. The five unique residues in VRK1 and the corresponding residues in Aurora B are indicated by magenta arrows.

By comparing VRK1 with VRK2 and Aurora B, five residues of VRK1 (Cys50, Asp132, Phe134, Ser181 and Val196), involved in interaction with ligand at the active site, are identified to be unique to VRK1. The corresponding five residues in VRK2 and Aurora B are shown in Table 3-9.

Table 3-9 VRK1 unique residues and corresponding residues in VRK2 and Aurora B.

VRK1	VRK2	Aurora B
Cys50	Leu42	Asn106
Asp132	Glu122	Leu170
Phe134	Leu124	Phe172
Ser181	Ala170	Glu220
Val196	Ala185	Ala233

The five unique residues of VRK1 involved in interaction and their corresponding residues in VRK2 and Aurora B are compared and shown in Figure 3-19.

In regard to the development of VRK1 inhibitor, the substitutions of Cys50 and Ser181 in VRK1, in comparison with VRK2 and Aurora B, are observed to be more significant than the Asp132, Phe134 and Val196 substitutions. As these substitutions (Asp132 and Phe134) occurring at DRF motif does not have much effect on ligand interaction, as only the backbone atoms of these residues are involved in interaction, and the mutation from valine (Val196) to alanine is not critical as both have hydrophobic characteristics. On the other hand, the Cys50 located at the end of P-loop motifs is unique in VRK1, and the position of cystine at the end of P-loop can be a target site for developing an allosteric covalent inhibitor for VRK1 (Zhao et al., 2017). Interestingly, the Ser181 of VRK1 which showed interaction with the AMP-PNP, supported by our structure analysis discussed in section 3.2.3.5, revealed to be a positive target site for inhibitor design. Serine has a side chain oxygen that can acts as a good hydrogen donor or

aid in covalent attachment with ligand, which is very promising for designing potent inhibitor toward the active site of VRK1. As Ser181 is unique to VRK1 and is also shown to interact with AMP-PNP at the active site, can be useful for the development of specific inhibitors for VRK1.

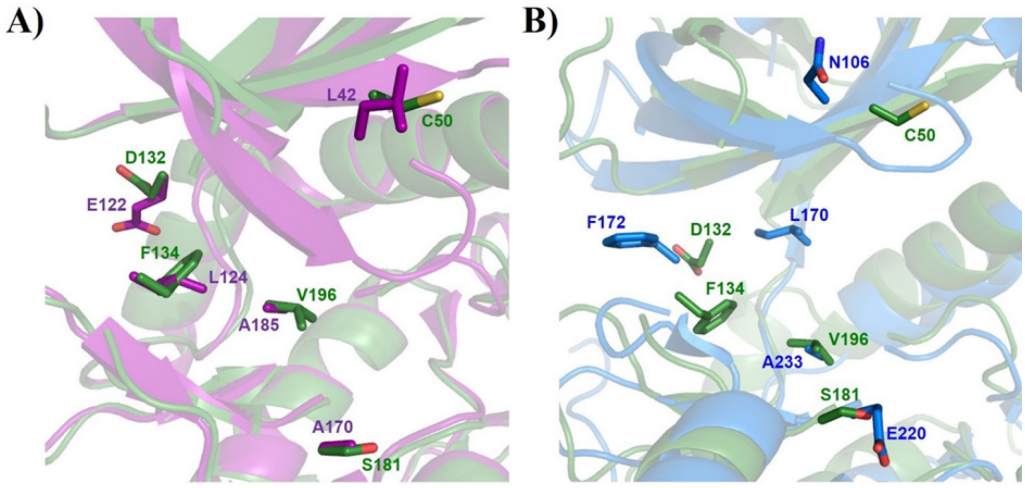


Figure 3-19 Comparison of VRK1 unique residues with the corresponding residues in VRK2 and Aurora B.

A) The unique residues at the active site of VRK1 (shown in green) is compared with the corresponding residues in VRK2 (shown in purple). B) The unique residues at the active site of VRK1 (shown in green) is compared with the corresponding residues in Aurora B (shown in blue). For clarification, only the side chain of the residues is shown for comparison, and the residues are labeled in single amino acid letter.

By using NMR spectroscopy titration and X-ray crystallography, the ATP-binding site on VRK1 was identified. The crystal structure of VRK1 bound to AMP-PNP was solved at a resolution of 2.07Å, and the structural analysis revealed the molecular basis of how ATP binds and stabilizes VRK1, and the active conformation of VRK1 is prepared for the downstream catalytic activity. On the next chapter, we would like to understand regulatory mechanism between VRK1 and histone H3 through the determination of VRK1 bound nucleosome complex by X-ray crystallography.

Chapter 4

Molecular Characterization of VRK1 with Histone H3

4.1 Motivation

VRK1 phosphorylates various substrates such as BAF, histone H3 and p53 which participates in cell proliferation progression. It will be interesting to discover the mechanism behind how VRK1 phosphorylates its substrates, and this can be understood through studying the interaction of substrates at the catalytic site of the kinase. Currently, the structure of VRK1 with any of its substrate proteins is not available yet. Since the interaction of VRK1 with its substrates is modelled mostly by docking, it would be interesting to experimentally determine the structure of VRK1 bound to its substrate. In this study, we would like to focus on understanding the mechanism behind VRK1-mediated histone H3 phosphorylation activity. Previously, our lab has discovered that VRK1 can phosphorylate histone H3's N-terminal tail at Thr3 and Ser10 and regulate the condensation of chromosomes during mitosis. The overexpression of VRK1 resulted in the hyper-condensation of nuclei during the cell division cycle (Kang et al., 2007). This implied that the presence of VRK1 plays an important role in mitotic regulation. Moreover, the phosphorylation of histone H3 by VRK1 oscillates during cell cycle and is controlled by macroH2A1, a variant histone (Kim et al., 2012). The interactions of VRK1 with histone H3 and chromatin were previously shown (Kang et al., 2007), but the interaction site is not yet discovered. Understanding the molecular basis of VRK1 – histone H3 recognition can further provide insights into the regulation of VRK1 phosphorylation mechanism and assist in designing inhibitors. Hence, the aim of this study is to determine the complex structure of VRK1 with nucleosome core particle (NCP) using X-ray crystallography, to aid in understanding the molecular basis of VRK1 and histone H3 interaction during phosphorylation. Toward this direction, we would like to

reconstitute the 145-bp 601 nucleosome core particle and crystallize it with VRK1. Following this, attempts to crystallize the complex and solve the structure would be carried out, to look at how histone H3 interacts with the catalytic domain of VRK1.

4.2 Results and Discussion

4.2.1 Amplification and purification of 145-bp 601 DNA fragment

The 145-bp 601 DNA is used for the reconstitution of the nucleosome core particle (NCP). The pUC19–145-bp 601 plasmid construct was made by the Davey's lab (Vasudevan et al., 2010) based on the Lowary & Widom, 1998 report of the 601 sequence (Lowary and Widom, 1998). The 145-bp 601 plasmid containing multiple copies of 145-bp 601 DNA fragment was amplified in large quantities using the *E.coli* DH5α bacterial cells. The plasmid DNA was extracted from the *E.coli* cells using the alkaline lysis method and purified by phenol-chloroform isoamyl alcohol (PCI) DNA extraction protocol described in section 2.2.19.1.

Before proceeding to large scale digestion of the plasmid DNA, a 100 µg of DNA digestion was performed to check on the DNA quality and digestibility. This is to ensure that the *EcoRV* restriction enzyme is active and the purified plasmid DNA can be well digested under the given condition and time. The *EcoRV* digested 145-bp 601 plasmids were analyzed on a 0.8% agarose gel shown in Figure 4-1A. Later, the plasmid DNA was digested in batches of 50 mg each, with 120-150 U of *EcoRV* per mg of plasmid. The 145-bp 601 DNA was purified from the vector DNA by PEG fractionation method described in section 2.2.19.2. After PEG fractionation, the supernatant and pellet fractions were collected and analyzed on 8% DNA-PAGE gel. From the gel image shown in Figure 4-1B, it can be observed that the 145-bp 601 DNA was fractionated into the supernatant, while the vector DNA (~2,600 bp) remained in the pellet. The 145-bp 601 DNA was later extracted by ethanol-precipitation. For every 50 mg of plasmid DNA, approximately 15 mg of 145-bp 601 DNA was purified.

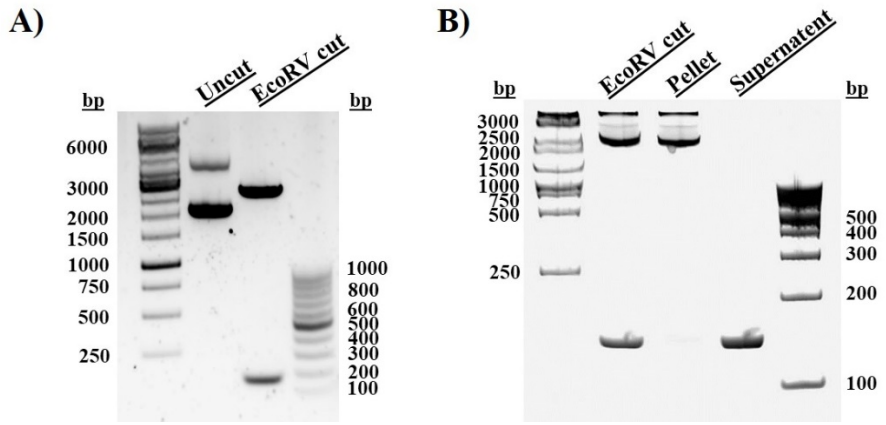


Figure 4-1 145-bp 601 DNA extraction and purification.

A) Microscale digestion of 145-bp 601 plasmid DNA by *EcoRV*. The *EcoRV* digested plasmid DNA is analyzed on the 0.8% agarose gel, along with the uncut pUC19–145-bp 601 plasmid, 100-bp and 1-kb DNA markers. In the digested plasmid sample (*EcoRV* cut), two bands representing the vector DNA (~2,600 bp) and DNA fragment (145 bp) are observed. B) Large scale digestion of 145-bp 601 plasmid DNA by *EcoRV*. The digested 145-bp 601 plasmid (*EcoRV* cut), pellet and supernatant collected from PEG fractionation are analyzed on an 8% DNA-PAGE, along with 100-bp and 1-kb DNA markers. The vector DNA (~2,600 bp) and 145-bp 601 DNA are PEG fractionated into supernatant and pellet fractions respectively.

4.2.2 Expression and purification of wild-type human histones

The wild-type human histone H2A, H2B, H3 and H4 were purified for the refolding of histone octamer. The four plasmids containing each histone gene were kindly given by Prof Lars' group within our School. The histone proteins were purified according to the protocol described in section 2.2.19.6. From all the histone chromatographs observed (Figure 4-2), the first peak eluted from 40-60 mL mainly contained high DNA contaminants, showed by high $A_{260/280}$ ratio. The second peak observed on the chromatograph was then the histone protein peak. The eluate fractions collected from the second peak were analyzed on 15% SDS-PAGE gel.

For histone H2A, the gel analysis showed that the fractions (A1-A4) collected from elution volume 64-80 mL (Figure 4-2A), contained the histone H2A (M.W. 14 kDa) and some impurities. While the histone H2B fractions (A1-A4) collected from elution volume 72-88 mL (Figure 4-2B), contained the histone H2B (M.W.

14 kDa) with little amount of impurities. In the histone H3 chromatogram (Figure 4-2C), it was observed that the histone H3 peak (V_e: 58-74 mL) was overlapping with the first peak, and from the gel analysis, fraction A4 to A7 contained mostly histone H3 (M.W. 15 kDa) with some impurities. Histone H4 was eluted from 76-92 mL (Figure 4-2D), and fraction A1 to A4 contained the histone H4 (M.W. 11 kDa) with little amount of impurities, observed from the gel analysis. All the fractions collected for each histone after gel filtration were checked on the NanoDrop spectrophotometer for DNA contamination. All the fractions of each histone had an $A_{260/280}$ ratio between 0.75-0.85.

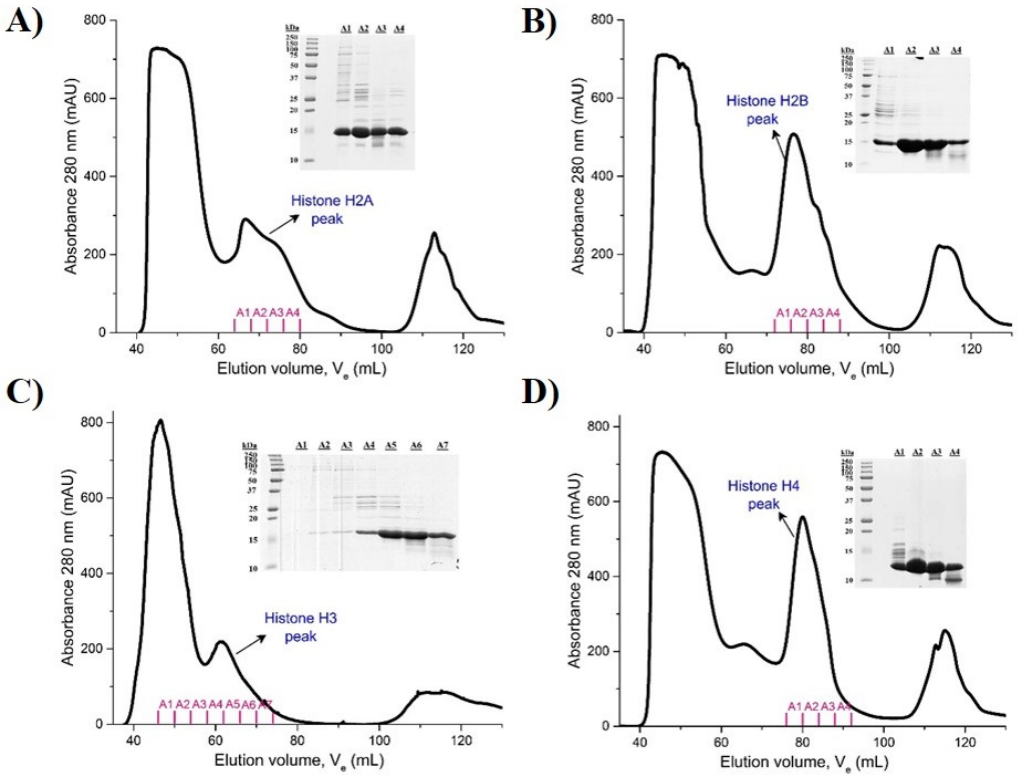


Figure 4-2 Size exclusion chromatography analysis of Histone proteins.

A-D) The chromatogram of histone H2A (A), H2B (B), H3 (C) and H4 (D) obtained after size exclusion chromatography (SEC). The histone proteins are purified under denaturing conditions using the Superdex S75 column. The protein peak corresponding to each histone is indicated by the black arrow and blue labeling. The fractions collected for each histone are shown in pink numbering and the SDS gel analysis of these fractions collected is shown as an inset on the right of respective histone chromatogram.

Each histone was further purified by Resource S cation exchange column, to remove other protein impurities. The chromatogram of each histone cation exchange chromatography has been shown in Figure 4-3. The eluted fractions were analyzed on a 15% SDS-PAGE gel and the image has been shown as an inset on the left of respective histone chromatogram. The pure fractions of each histone were checked on the NanoDrop spectrophotometer for DNA contamination. The fractions of each histone showed an $A_{260/280}$ ratio between 0.5-0.6. The pure fractions were then pooled together and dialyzed. The histones were lyophilized and stored in -80°C freezer and were ready to be use for histone octamer refolding.

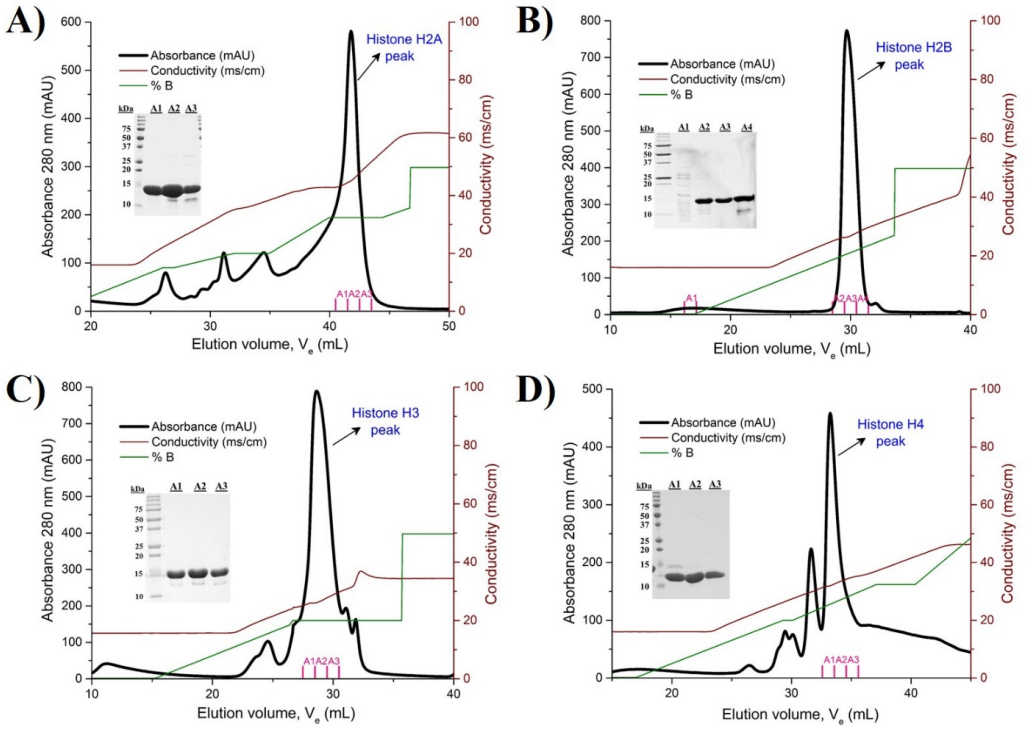


Figure 4-3 Cation exchange chromatography analysis of histone proteins.

A-D) The chromatogram of histone H2A (A), H2B (B), H3 (C) and H4 (D) obtained after cation exchange chromatography. The histone proteins are purified under denaturing conditions using the Resource S (cation) column. The protein peak corresponding to each histone is indicated by the black arrow and blue labeling. The fractions collected for each histone are shown in pink numbering and the SDS gel analysis of these fractions collected is shown as an inset on the left of respective histone chromatogram.

4.2.3 Refolding of Histone octamer

For the refolding of histone octamer, histone H2A, H2B, H3 and H4 proteins were mixed in the molar ratio of 1:1:1:1 and dialyzed with refolding buffer for a day. The histone octamer sample was concentrated using the Amicon concentrator and loaded into the Superdex S200 16/600 column. The histone octamer complex was eluted out using refolding buffer and the fractions collected were analyzed on the 15% SDS-PAGE gel. The chromatogram of histone octamer gel filtration analysis and the SDS-PAGE gel image are shown in Figure 4-4. The purified histone octamer fractions were pooled together and concentrated. The histone octamer was mixed with 50% glycerol and the final protein concentration was about 10-15 mg/mL. The protein was stored at -20°C and is ready to be used for nucleosome core particle reconstitution.

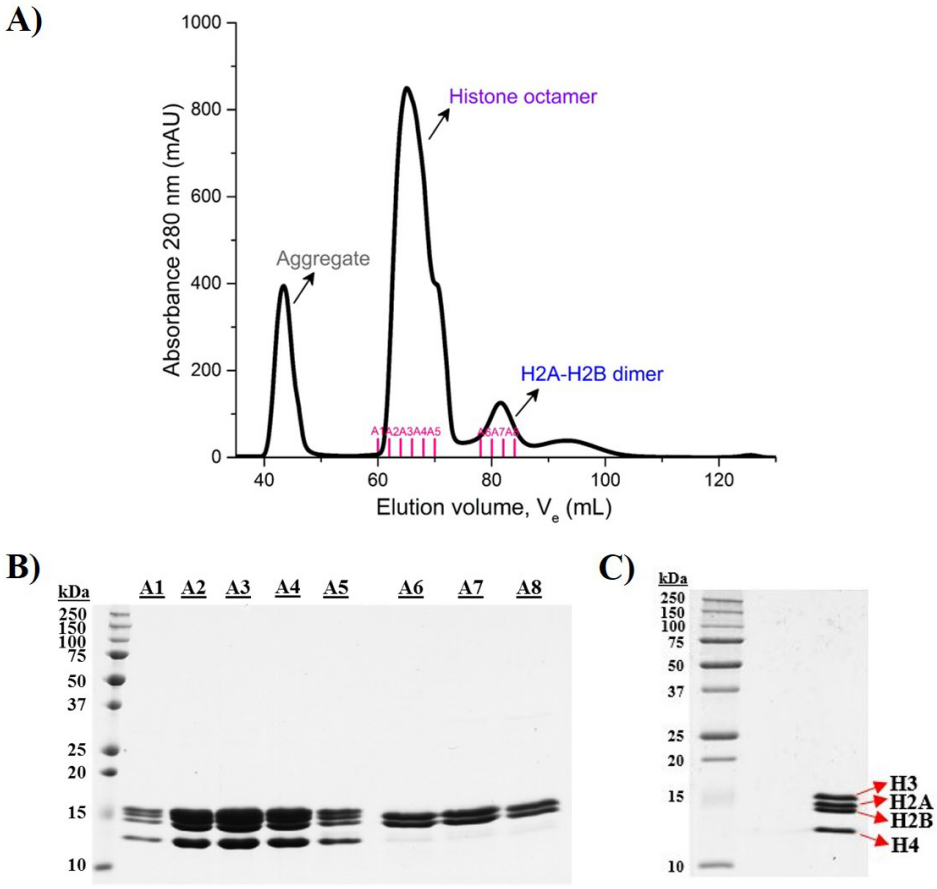


Figure 4-4 Histone octamer purification analysis.

Figure 4-4 (cont.)

A) Size exclusion chromatography analysis of Histone octamer. The refolded histone octamer was purified by Superdex S200 column. The protein aggregate peak (Ve: 40-50 mL), histone octamer peak (Ve: 60-70 mL) and H2A-H2B dimer peak (Ve: 79-85 mL) are labeled in grey, purple and blue, respectively. The fractions collected are labeled in pink numbering. B, C) SDS-PAGE analysis of Histone octamer after gel filtration. The fractions (A1-A8) collected were analyzed on 15% SDS-PAGE gel, along with the protein marker. Fractions A1 to A5 contained the histone octamer while A6 to A8 contained the H2A-H2B dimer. The M.W. of each histone is H3 = 15 kDa, H2A = 13.9 kDa, H2B = 13.7 kDa and H4 = 11 kDa.

4.2.4 Reconstitution of 145-bp 601 nucleosome core particle

The 145-bp 601 nucleosome core particle (NCP) was reconstituted in 100 µL volume, by mixing the 145-bp 601 DNA and histone octamer in 2M KCl and dialyzed against stepwise decrease in salt concentration (protocol described in section 2.2.19.8). 7 different molar ratio's of histone octamer to DNA (0.5:1, 0.8:1, 1:1, 1.05:1, 1.08:1, 1.1:1 and 1.3:1) were used to reconstitute the NCP. The products were then analyzed on 6% DNA-PAGE gel. In Figure 4-5, the 145-bp 601 NCP was observed as a band between 300-400 bp. It can also be observe that the amount of free 145-bp 601 DNA decreased as more histone octamer was added. The optimum ratio was determined by the least amount of free 145-bp 601 DNA left after reconstitution. This was observed in the sample with the molar ratio of histone octamer to DNA at 1.05:1 which was confirmed using the 6% DNA-PAGE gel (Figure 4-5). Beyond the ratio of 1.1:1, precipitation was observed after reconstitution.

An optimum ratio of histone octamer to DNA was acquired experimentally and the ratio varies between batches of histone octamer used. The reconstituted NCP was kept at 4°C and ready to be used for its interaction studies with VRK1.

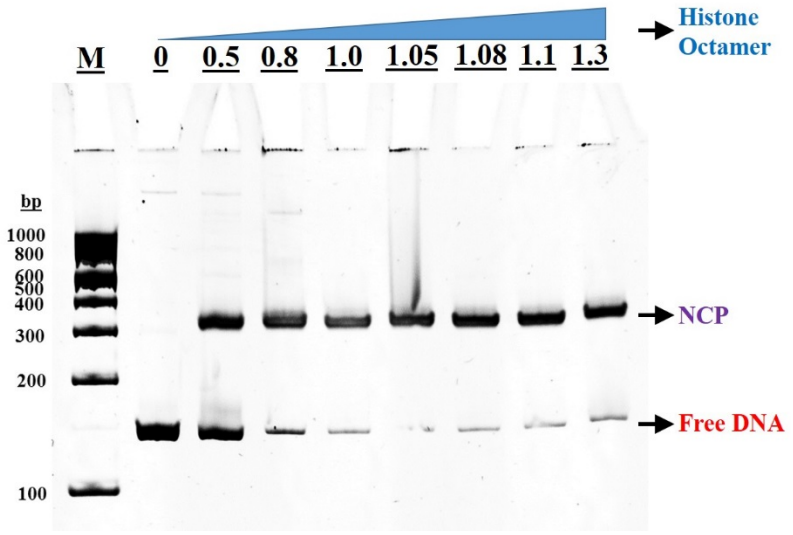


Figure 4-5 Reconstitution of 145-bp 601 nucleosome core particle.

7 molar ratios of histone octamer to 145-bp 601 DNA (0.5:1, 0.8:1, 1.0:1, 1.05:1, 1.08:1, 1.1:1, 1.3:1) was reconstitute with the nucleosome core particle, and the products were loaded, together with the 100 bp DNA marker on the 6% DNA-PAGE gel. The free DNA was observed at 145 bp and the reconstituted NCP was between 300-400 bp.

4.2.5 Optimization of VRK1-bound 145-bp 601 NCP complex construction

To understand the regulatory mechanisms of VRK1 on histone H3's N-terminal tail modification during mitosis, our main objective was to solve the structure of VRK1 bound to the nucleosome core particle (NCP) and study the interaction between histone H3 and the catalytic domain of VRK1. The first part of this study is to obtain the VRK1-bound NCP complex in stable conformation with high homogeneity and quantity before attempting to carry out crystallization. Accordingly, the binding ability of VRK1 on NCP was investigated by Mobility Shift Electrophoretic assay to obtain the optimum ratio to obtain the VRK1-NCP complex with high stability. Increasing concentrations (from 0.5 µM up to 15 µM) of VRK1^{wt} was added to 1 µM of NCP and after 15 min of incubation, precipitates were observed in the mixtures with 5 µM to 15 µM of VRK1^{wt} added. The mixtures were spun down and the supernatant was resolved on the 6% DNA-PAGE gel. From Figure 4-6A, it was observed that the NCP band shifted (the

band showing the bound form) with the increasing amount of $VRK1^{wt}$ protein added. Moreover, the amount of free NCP decreases simultaneously with an increase in the amount of $VRK1^{wt}$ and NCP in the bound form. It was observed that the complexing of $VRK1^{wt}$ to NCP reached a saturation point at 10 μM of $VRK1^{wt}$ to 1 μM of NCP.

The binding affinity of $VRK1^{\Delta 35}$ and $VRK1^{cryst}$ constructs with NCP were also tested. Interestingly, it was observed that both constructs bind weaker to NCP, compared with $VRK1^{wt}$. In Figure 4-6B, a lesser amount of the bound form (VRK1-NCP complex) was observed with the addition of $VRK1^{\Delta 35}$ protein, and the protein mixture was observed to be unresolved on the gel. In the case of $VRK1^{cryst}$ protein, a lesser amount of the bound form, which implied more presence of free NCP, was seen with increasing concentration of VRK1 (from 5 μM to 15 μM) added. The differences between the VRK1 constructs ($VRK1^{wt}$, $VRK1^{\Delta 35}$ and $VRK1^{cryst}$) are the absence of the C-terminal tail from residues 362 to 396 (in $VRK1^{\Delta 35}$ and $VRK1^{cryst}$) and the mutation of several Lys/Glu residues on the protein surface of VRK1 (in $VRK1^{cryst}$) (Appendix I-4). It can be observed that the absence of C-terminal tail and the mutation of surface charged residues, have shown a reduction in VRK1 protein stability and binding affinity to NCP. The C-terminal tail of VRK1, is a flexible region with many charged residues and was found to be important for its structure stability and catalytic activity (Shin et al., 2011).

As mentioned earlier, precipitation was observed in the mixture of NCP with increasing addition of $VRK1^{wt}$ from 5 to 15 μM. During incubation, a small amount of salt, NaCl (0-10 mM) was added into the buffer. The low amount of salt used could affect the solubility of the protein, which might be one of the

causes of the precipitation. Since, it was observed that the interaction of NCP with VRK1 reaches saturation at a molar ratio of 1:10 (NCP to VRK1) (Figure 4-6A), we mixed NCP to VRK1 in the molar ratio of 1:10 in the presence of increasing concentration of NaCl (from 0 mM to 250 mM), so as to increase protein solubility and protein-protein interactions to prevent precipitation. However, the addition of NaCl to the mixtures did not resolve the precipitation problem, and it was observed that increasing NaCl concentrations (from 100 mM to 250 mM) caused the bound form to dissociate (free NCP observed) (Figure 4-6C). The buffer system was experimentally changed to HEPES or phosphate, as some has suggested that Tris might not be suitable for protein-protein interaction studies due to its temperature-dependent pH property. Tris pH value can increase by approximately 0.03 pH unit per degree Celsius decrease in temperature (Good et al., 1966). However, the change of buffer systems did not help to resolve or even reduce the precipitation. Since the construction of $VRK1^{wt}$ with NCP to a saturation point of 1:10 for obtaining a stable and homogenous sample was not possible, the NCP and $VRK1^{wt}$ were mixed at a lower molar ratio of 1:5 to obtain a mixture of bound and free NCP sample, which was then separated by using size exclusion chromatography. However, when the NCP was mixed with $VRK1^{wt}$ in the molar ratio of 1:5 in a larger quantity, the mixture of protein was not stable enough to withstand the downstream purification processes. Simultaneously, the NCP-VRK1 samples, incubated at 1:5 molar ratio, were subjected to crystallization screening without success. Apart from this, we tried negative staining electron microscopy (EM) to observe the binding of VRK1 with NCP in solution. From negative staining only the NCP alone particles (~200 kDa) at a diameter of 11 nm were observed (Finch et al., 1977). Any other particles

that look different from the 11 nm NCP, could not be observed. The most probable reason for this could be that the sample used for negative staining is heterogeneous, containing VRK1 bound and unbound NCPs. The VRK1 (~ 45kDa) might be too small for imaging using negative staining. In general, the size of the target protein as well as the homogeneity of the sample can affect the quality of the data for negative staining or Cryo-EM imaging. Since both these factors, size and homogeneity, are presently a constraint for VRK1-NCP complex used here, it would be important to obtain a homogenous VRK1-NCP complex sample for imaging or crystallization.

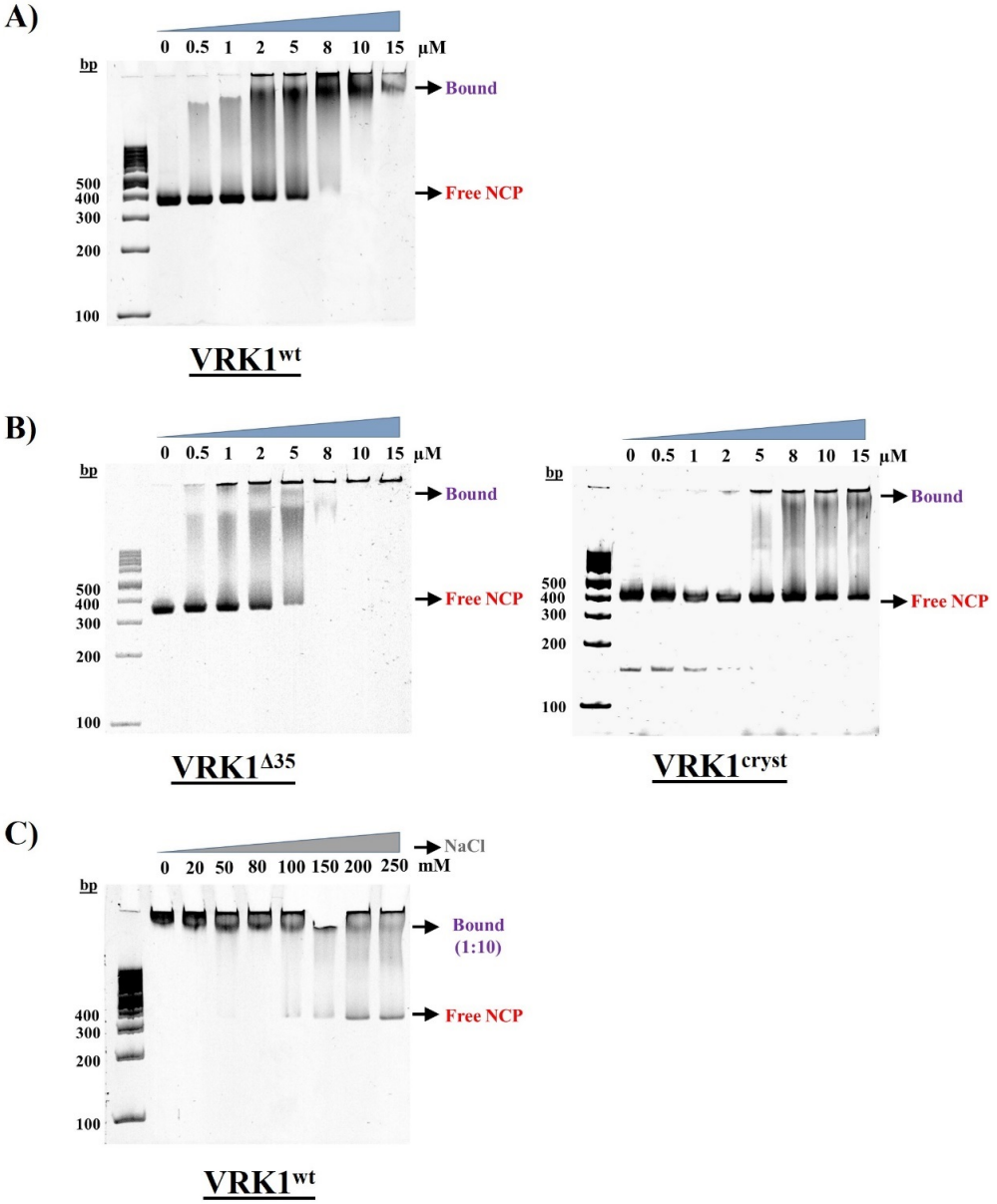


Figure 4-6 Mobility Shift Electrophoretic assay of NCP with the addition of VRK1 protein.

A, B) Binding assay of NCP with increasing concentration of VRK1 added. The 6% DNA-PAGE gel analysis of 1 µM of NCP mixed with increasing concentrations (from 0.5 µM up to 15 µM) of VRK1 is shown in A (for $VRK1^{wt}$) and B (for $VRK1^{\Delta 35}$ and $VRK1^{cryst}$). The mixtures were loaded along with the 100 bp DNA marker. C) Binding assay of NCP with VRK1 at molar ratio 1:10 in the presence of increasing salt concentration. 1 µM of NCP was mixed with 10 µM of $VRK1^{wt}$ in buffers containing increasing concentration of NaCl (from 0 to 250 mM). The mixtures were loaded along with the 100bp DNA marker on the 6% DNA-PAGE gel. The increasing concentration of VRK1 protein (in µM) is indicated by the blue triangle while the increasing concentration of NaCl (mM) is indicated by the grey triangle. The free NCP and the bound form of NCP with VRK1 are shown using red and purple words respectively.

4.2.6 Crystallization of VRK1 with histone H3's N-terminal peptide

As the attempt to obtain a homogenous sample of VRK1-NCP complex in stable form for X-ray crystallography or negative staining study was not successful, we attempted to study the H3 tail interaction with VRK1 by synthesizing the histone H3's N-terminal tail fragment (from amino acid 1 to 15 – hereafter referred as H3 peptide).

To start with, it is important to note that the study of protein-protein interactions can be affected by buffer conditions such as the pH, salt and ionic strength. One could predict the possible binding probability based on the pI values of the target proteins. In this case, the pI values of $VRK1^{wt}$, $VRK1^{\Delta 35}$ and $VRK1^{cryst}$ are 9.02, 9.10 and 8.75, respectively, while that of the H3 peptide is 11.33 indicating that all are basic in nature. Nonetheless the probable substrate binding interface in VRK1 contains a lot of acidic patches which could provide a complimentary interface for the highly basic H3 peptide recognition. Though these are just predictions, one needs to have the complex structure to reveal the true nature of these interactions. Toward this end, attempts to co-crystallize the ternary complex of VRK1 with AMP-PNP and the H3 peptide, using the known crystallization conditions for obtaining the $VRK1^{cryst}$ bound AMP-PNP crystal, were made. Further, the VRK1-AMP-PNP crystal structure indicated that upon crystal packing, the probable substrate binding interface was still accessible which encouraged us to attempt soaking method apart from the co-crystallization efforts. So, the $VRK1^{cryst}$ - AMP-PNP crystals were also soaked with H3 peptide, using varying concentrations (2 mM to 20 mM) and time (few seconds to 1hr). Crystals from co-crystallization and soaking trials were diffracted at TPS05A beamline at NSRRC, Taiwan. The best crystals (0.3-0.4 mm in length) (Figure

4-7) from either trial diffracted ~2.7 – 3.0Å. However, analysis of these data revealed the absence of H3 peptide density in the crystal. The reason for the exclusion of H3 peptide from VRK1–AMP-PNP structure, may be due to the flexible nature of the short H3 peptide, which does not have any defined secondary structure regions. Since, the VRK1-AMP-PNP-H3 ternary complex could not be obtained using crystallography method, in future we plan to study the interaction of the H3 peptide with VRK1, using NMR titration and X-ray crystallography, by 1) using the post-translationally modified H3 peptide and 2) varying the length of H3 peptide.

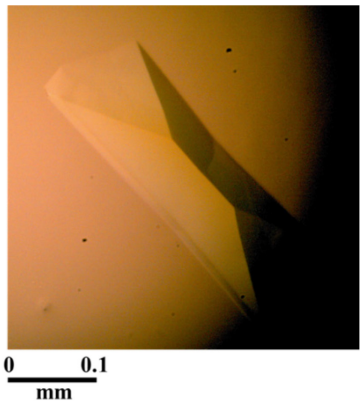


Figure 4-7 Crystal image of VRK1-bound AMP-PNP with the addition of H3 peptide.

VRK1-bound AMP-PNP complex co-crystallized with the H3 peptide in 27.5% PEG 3350, 0.1 M HEPES (pH7.0) and 0.2 M ammonium sulphate.

Chapter 5

Conclusion and Future Perspectives

5.1 Conclusion

In conclusion, the ATP-binding residues identified on VRK1 by NMR titration studies are matching to the catalytic site of VRK1, discovered from the AMP-PNP bound VRK1 crystal structure solved at a resolution of 2.07Å. The crystal structure revealed the presence of conserved motifs such as the lysine-glutamate salt bridge formation, DRF and DYG motifs, which are distinct features of an active conformation. The comparisons of AMP-PNP bound VRK1, with the apo and inhibitor counterparts, revealed that (i) the P-loop (of the AMP-PNP bound VRK1) might adopt a "Closed-conformation" similar to inhibitor bound which is reminiscent of ATP binding, though this needs further investigation; (ii) the presence of ligand can enhance the Lys71-Glu83 salt bridge formation; (iii) the DYG adopts a 'DYG-in' conformation and (iv) the inhibitor which docks deeper into the active site than the AMP-PNP, results in a change in the orientation of Met131 side chain (next to DRF motif) which exhibits flexibility to accommodate the ligand. Moreover, the comparisons of AMP-PNP bound VRK1 with the ligand bound VRK1, VRK2 and Aurora B structures, revealed that Ser181 in VRK1 which show interaction at the active site, can be a promising hotspot for developing ATP-competitive inhibitors specific to VRK1, due to the nature of serine side chain oxygen that is a good hydrogen donor and covalent attachment site. The above analysis and findings can provide great insights in the designing of effective and specific ATP competitive inhibitors for VRK1. After the discovery of the ATP-binding site of VRK1, the understanding of how VRK1 can catalyze the transfer of phosphate to its interacting substrate for downstream signaling activities is also important. The study of how VRK1 interacts with its substrate histone H3, to facilitate the phosphorylation of H3 tail for chromosome

condensation during mitosis is also helpful for the designing of inhibitor which can target VRK1-mediated regulation. With this aim, the 145-bp 601 DNA was purified and later wrapped around the histone octamer refolded from human wild-type core histones H2A, H2B, H3 and H4, to reconstitute the 145-bp 601 nucleosome core particle (NCP). The NCP was successfully reconstituted as confirmed by 6% DNA-PAGE analysis. The 145-bp 601 NCP binding to VRK1 was observed by using the mobility shift electrophoresis assay. The NCP migrated to a higher molecular weight with the addition of VRK1. However, the construction of homogenous VRK1 bound to NCP complex was unsuccessful, despite trying different conditions (salt, buffer and pH), due to the instability of complex. It is to be noted that only a few structures of NCP-protein complexes are available to-date, such as the NCP in complex with RCC1 chromatin factor (PDB ID 3MVD) (Makde et al., 2010) and Snf2, a transcription regulatory protein (PDB ID 5X0X) (Liu et al., 2017), implying that this is a challenging project which needs more optimization and characterization. As an alternative approach, crystallizing of the H3 peptide (histone H3's N-terminal tail – amino acid 1-15) with the VRK1-bound AMP-PNP was attempted to study the interaction between VRK1 and histone H3 during phosphorylation. The best crystals were diffracted at ~2.7 – 3.0Å but the analysis of these data revealed the absence of H3 peptide density. Henceforth, the determination of the structure of VRK1 bound NCP by crystallography could not be accomplished. It would be critical to further examine the histone H3 recognition site on VRK1 which is important for understanding the VRK1-mediated chromatin remodeling during mitosis. This discovery can provide insights for developing inhibitors targeting VRK1-mediated phosphorylation in mitosis.

5.2 Future Perspectives

From the crystal structure of AMP-PNP bound VRK1, several residues of VRK1 which are involved in the interaction with AMP-PNP were identified. These interacting residues identified from VRK1 can be mutated to other residues through site-directed mutagenesis. Further testing of the functionality of these mutated VRK1 residues using in vivo or in vitro kinase assays, can aid in identifying the critical residue(s) important for VRK1 catalytic function.

For the chemistry-based studies, two approaches could be considered. Firstly, novel VRK1 inhibitors could be designed based on our findings. Secondly, ligand screening based on VRK1 specificity, could be performed to identify possible 'hits'. The potential ligands identified can then be tested for their inhibitory effects using cell-based assays on relevant cancer cell lines. In addition, the potential ligands can be complexed with VRK1 and studied using structure-based experiments to identify the possible ligand interactions with VRK1. The study of these ligand bound VRK1 structures can provide insights for developing novel and highly specific VRK1 inhibitors.

For the analysis of VRK1-histone H3 regulation, we would like to continue optimizing the crystal screening of AMP-PNP bound VRK1 with the addition of H3 peptide, to determine how the histone H3 tail interacts with VRK1. We would also like to further optimize the existing protocols of VRK1 and nucleosome core particle (NCP) complex construction, to obtain homogenous and stable protein complex. This complex can be used in future studies to understand VRK1-NCP interaction using biophysical techniques such as X-ray crystallography, Cryo-electron microscopy or solid-state nuclear magnetic resonance (NMR).

Chapter 6

References

ABDALLA, M., ALI ELTAYB, W., SAMAD, A., SHM, E. & DAFAALLA, T. I. M. 2016. *Important Factors Influencing Protein Crystallization*.
ALBERTS, B., JOHNSON, A., LEWIS, J., RAFF, M., ROBERTS, K. & WALTER, P. 2002. Molecular Biology of The Cell. 4th ed. New York: Garland Science Taylor and Francis Group.
ANNUNZIATO, A. 2008. DNA packaging: Nucleosomes and chromatin. 1**,** 26.
ASENCIO, C., DAVIDSON, I. F., SANTARELLA-MELLWIG, R., LY-HARTIG, T. B., MALL, M., WALLENFANG, M. R., MATTAJ, I. W. & GORJANACZ, M. 2012. Coordination of kinase and phosphatase activities by Lem4 enables nuclear envelope reassembly during mitosis. *Cell,* 150**,** 122-35.
BAEK, S. H. 2011. When signaling kinases meet histones and histone modifiers in the nucleus. *Mol Cell,* 42**,** 274-84.
BATTYE, T. G., KONTOGIANNIS, L., JOHNSON, O., POWELL, H. R. & LESLIE, A. G. 2011. iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallogr D Biol Crystallogr,* 67**,** 271-81.
BAUD, F. & KARLIN, S. 1999. Measures of residue density in protein structures. *Proc Natl Acad Sci U S A,* 96**,** 12494-9.
BAYLISS, R., FRY, A., HAQ, T. & YEOH, S. 2012. On the molecular mechanisms of mitotic kinase activation. *Open Biol,* 2**,** 120136.
BLANCO, S., KLIMCAKOVA, L., VEGA, F. M. & LAZO, P. A. 2006. The subcellular localization of vaccinia-related kinase-2 (VRK2) isoforms determines their different effect on p53 stability in tumour cell lines. *Febs j,* 273**,** 2487-504.
CAI, M., HUANG, Y., ZHENG, R., WEI, S. Q., GHIRLANDO, R., LEE, M. S., CRAIGIE, R., GRONENBORN, A. M. & CLORE, G. M. 1998. Solution structure of the cellular factor BAF responsible for protecting retroviral DNA from autointegration. *Nat Struct Biol,* 5**,** 903-9.
CHEN, V. B., ARENDALL, W. B., 3RD, HEADD, J. J., KEEDY, D. A., IMMORMINO, R. M., KAPRAL, G. J., MURRAY, L. W., RICHARDSON, J. S. & RICHARDSON, D. C. 2010. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr,* 66**,** 12-21.
CHENG, H. C., QI, R. Z., PAUDEL, H. & ZHU, H. J. 2011. Regulation and function of protein kinases and phosphatases. *Enzyme Res,* 2011**,** 794089.
COUNAGO, R. M., ALLERSTON, C. K., SAVITSKY, P., AZEVEDO, H., GODOI, P. H., WELLS, C. I., MASCARELLO, A., DE SOUZA GAMA, F. H., MASSIRER, K. B., ZUERCHER, W. J., GUIMARAES, C. R. W. & GILEADI, O. 2017. Structural characterization of human Vaccinia-Related Kinases (VRK) bound to small-molecule inhibitors identifies different P-loop conformations. *Sci Rep,* 7**,** 7501.
COWAN-JACOB, S. W. 2006. Structural biology of protein tyrosine kinases. *Cell Mol Life Sci,* 63**,** 2608-25.
CROSIO, C., FIMIA, G. M., LOURY, R., KIMURA, M., OKANO, Y., ZHOU, H., SEN, S., ALLIS, C. D. & SASSONE-CORSI, P. 2002. Mitotic phosphorylation of histone H3: spatio-temporal regulation by mammalian Aurora kinases. *Mol Cell Biol,* 22**,** 874-85.
DAI, J., SULTAN, S., TAYLOR, S. S. & HIGGINS, J. M. 2005. The kinase haspin is required for mitotic histone H3 Thr 3 phosphorylation and normal metaphase chromosome alignment. *Genes Dev,* 19**,** 472-88.
DAVIS, I. W., LEAVER-FAY, A., CHEN, V. B., BLOCK, J. N., KAPRAL, G. J., WANG, X., MURRAY, L. W., ARENDALL, W. B., 3RD, SNOEYINK, J.,

RICHARDSON, J. S. & RICHARDSON, D. C. 2007. MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res,* 35**,** W375-83.
DEREWENDA, Z. S. 2004. Rational protein crystallization by mutational surface engineering. *Structure,* 12**,** 529-35.
DOBNER, T., WOLF, I., MAI, B. & LIPP, M. 1991. A novel divergently transcribed human histone H2A/H2B gene pair. *DNA Seq,* 1**,** 409-13.
ELKINS, J. M., FEDELE, V., SZKLARZ, M., ABDUL AZEEZ, K. R., SALAH, E., MIKOLAJCZYK, J., ROMANOV, S., SEPETOV, N., HUANG, X. P., ROTH, B. L., AL HAJ ZEN, A., FOURCHES, D., MURATOV, E., TROPSHA, A., MORRIS, J., TEICHER, B. A., KUNKEL, M., POLLEY, E., LACKEY, K. E., ATKINSON, F. L., OVERINGTON, J. P., BAMBOROUGH, P., MULLER, S., PRICE, D. J., WILLSON, T. M., DREWRY, D. H., KNAPP, S. & ZUERCHER, W. J. 2016. Comprehensive characterization of the Published Kinase Inhibitor Set. *Nat Biotechnol,* 34**,** 95-103.
ELTSOV, M., MACLELLAN, K. M., MAESHIMA, K., FRANGAKIS, A. S. & DUBOCHET, J. 2008. Analysis of cryo-electron microscopy images does not support the existence of 30-nm chromatin fibers in mitotic chromosomes in situ. *Proc Natl Acad Sci U S A,* 105**,** 19732-7.
EMSLEY, P. & COWTAN, K. 2004. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr,* 60**,** 2126-32.
ENDICOTT, J. A., NOBLE, M. E. & JOHNSON, L. N. 2012. The structural basis for control of eukaryotic protein kinases. *Annu Rev Biochem,* 81**,** 587-613.
ESWARAN, J., PATNAIK, D., FILIPPAKOPOULOS, P., WANG, F., STEIN, R. L., MURRAY, J. W., HIGGINS, J. M. & KNAPP, S. 2009. Structure and functional characterization of the atypical human kinase haspin. *Proc Natl Acad Sci U S A,* 106**,** 20198-203.
EVANS, P. 2006. Scaling and assessment of data quality. *Acta Crystallogr D Biol Crystallogr,* 62**,** 72-82.
FINCH, J. T., LUTTER, L. C., RHODES, D., BROWN, R. S., RUSHTON, B., LEVITT, M. & KLUG, A. 1977. Structure of nucleosome core particles of chromatin. *Nature,* 269**,** 29-36.
FURUKAWA, K. 1999. LAP2 binding protein 1 (L2BP1/BAF) is a candidate mediator of LAP2-chromatin interaction. *J Cell Sci,* 112 (Pt 15)**,** 2485-92.
FURUKAWA, K., SUGIYAMA, S., OSOUDA, S., GOTO, H., INAGAKI, M., HORIGOME, T., OMATA, S., MCCONNELL, M., FISHER, P. A. & NISHIDA, Y. 2003. Barrier-to-autointegration factor plays crucial roles in cell cycle progression and nuclear organization in Drosophila. *J Cell Sci,* 116**,** 3811-23.
GASTEIGER, E., HOOGLAND, C., GATTIKER, A., DUVAUD, S. E., WILKINS, M. R., APPEL, R. D. & BAIROCH, A. 2005. Protein Identification and Analysis Tools on the ExPASy Server. *In:* WALKER, J. M. (ed.) *The Proteomics Protocols Handbook.* Totowa, NJ: Humana Press.
GOLDSCHMIDT, L., COOPER, D. R., DEREWENDA, Z. S. & EISENBERG, D. 2007. Toward rational protein crystallization: A Web server for the design of crystallizable protein variants. *Protein Sci,* 16**,** 1569-76.
GOOD, N. E., WINGET, G. D., WINTER, W., CONNOLLY, T. N., IZAWA, S. & SINGH, R. M. M. 1966. Hydrogen Ion Buffers for Biological Research*. *Biochemistry,* 5**,** 467-477.
GORJANACZ, M. 2013. LEM-4 promotes rapid dephosphorylation of BAF during mitotic exit. *Nucleus,* 4**,** 14-7.

GORJÁNÁCZ, M., KLERKX, E. P. F., GALY, V., SANTARELLA, R., LÓPEZ-IGLESIAS, C., ASKJAER, P. & MATTAJ, I. W. 2007. Caenorhabditis elegans BAF-1 and its kinase VRK-1 participate directly in post-mitotic nuclear envelope assembly. *Embo j,* 26**,** 132-43.
GOTO, H., TOMONO, Y., AJIRO, K., KOSAKO, H., FUJITA, M., SAKURAI, M., OKAWA, K., IWAMATSU, A., OKIGAKI, T., TAKAHASHI, T. & INAGAKI, M. 1999. Identification of a novel phosphorylation site on histone H3 coupled with mitotic chromosome condensation. *J Biol Chem,* 274**,** 25543-9.
GOTO, H., YASUI, Y., NIGG, E. A. & INAGAKI, M. 2002. Aurora-B phosphorylates Histone H3 at serine28 with regard to the mitotic chromosome condensation. *Genes Cells,* 7**,** 11-7.
GOUET, P., ROBERT, X. & COURCELLE, E. 2003. ESPript/ENDscript: Extracting and rendering sequence and 3D information from atomic structures of proteins. *Nucleic Acids Res,* 31**,** 3320-3.
HANKS, S. K. & HUNTER, T. 1995. Protein kinases 6. The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification. *Faseb j,* 9**,** 576-96.
HARAGUCHI, T., KOJIDANI, T., KOUJIN, T., SHIMI, T., OSAKADA, H., MORI, C., YAMAMOTO, A. & HIRAOKA, Y. 2008. Live cell imaging and electron microscopy reveal dynamic processes of BAF-directed nuclear envelope assembly. *J Cell Sci,* 121**,** 2540-54.
HENDZEL, M. J., WEI, Y., MANCINI, M. A., VAN HOOSER, A., RANALLI, T., BRINKLEY, B. R., BAZETT-JONES, D. P. & ALLIS, C. D. 1997. Mitosis-specific phosphorylation of histone H3 initiates primarily within pericentromeric heterochromatin during G2 and spreads in an ordered fashion coincident with mitotic chromosome condensation. *Chromosoma,* 106**,** 348-60.
HUANG, W., CUI, X., CHEN, Y., SHAO, M., SHAO, X., SHEN, Y., LIU, Q., WU, M., LIU, J., NI, W., LU, C. & WAN, C. 2016. High VRK1 expression contributes to cell proliferation and survival in hepatocellular carcinoma. *Pathol Res Pract,* 212**,** 171-8.
JACKSON, S. P. & BARTEK, J. 2009. The DNA-damage response in human biology and disease. *Nature,* 461**,** 1071-8.
JOHNSON, L. N., NOBLE, M. E. & OWEN, D. J. 1996. Active and inactive protein kinases: structural basis for regulation. *Cell,* 85**,** 149-58.
KANG, T. H., PARK, D. Y., CHOI, Y. H., KIM, K. J., YOON, H. S. & KIM, K. T. 2007. Mitotic histone H3 phosphorylation by vaccinia-related kinase 1 in mammalian cells. *Mol Cell Biol,* 27**,** 8533-46.
KARDALINOU, E., EICK, S., ALBIG, W. & DOENECKE, D. 1993. Association of a human H1 histone gene with an H2A pseudogene and genes encoding H2B.1 and H3.1 histones. *J Cell Biochem,* 52**,** 375-83.
KARVE, T. M. & CHEEMA, A. K. 2011. Small Changes Huge Impact: The Role of Protein Posttranslational Modifications in Cellular Homeostasis and Disease. *Journal of Amino Acids,* 2011.
KIM, S. H., LYU, H. N., KIM, Y. S., JEON, Y. H., KIM, W., KIM, S., LIM, J. K., LEE, H. W., BAEK, N. I., CHOI, K. Y., LEE, J. & KIM, K. T. 2015a. Brazilin Isolated from Caesalpinia sappan suppresses nuclear envelope reassembly by inhibiting barrier-to-autointegration factor phosphorylation. *J Pharmacol Exp Ther,* 352**,** 175-84.
KIM, S. H., RYU, H. G., LEE, J., SHIN, J., HARIKISHORE, A., JUNG, H. Y., KIM, Y. S., LYU, H. N., OH, E., BAEK, N. I., CHOI, K. Y., YOON, H. S. &

KIM, K. T. 2015b. Ursolic acid exerts anti-cancer activity by suppressing vaccinia-related kinase 1-mediated damage repair in lung cancer cells. *Sci Rep,* 5**,** 14570.
KIM, W., CHAKRABORTY, G., KIM, S., SHIN, J., PARK, C. H., JEONG, M. W., BHARATHAM, N., YOON, H. S. & KIM, K. T. 2012. Macro histone H2A1.2 (macroH2A1) protein suppresses mitotic kinase VRK1 during interphase. *J Biol Chem,* 287**,** 5278-89.
KIM, Y. S., KIM, S. H., SHIN, J., HARIKISHORE, A., LIM, J. K., JUNG, Y., LYU, H. N., BAEK, N. I., CHOI, K. Y., YOON, H. S. & KIM, K. T. 2014. Luteolin suppresses cancer cell proliferation by targeting vaccinia-related kinase 1. *PLoS One,* 9**,** e109655.
KNIGHTON, D. R., ZHENG, J. H., TEN EYCK, L. F., ASHFORD, V. A., XUONG, N. H., TAYLOR, S. S. & SOWADSKI, J. M. 1991. Crystal structure of the catalytic subunit of cyclic adenosine monophosphate-dependent protein kinase. *Science,* 253**,** 407-14.
KOBAYASHI, J., TAUCHI, H., SAKAMOTO, S., NAKAMURA, A., MORISHIMA, K., MATSUURA, S., KOBAYASHI, T., TAMAI, K., TANIMOTO, K. & KOMATSU, K. 2002. NBS1 localizes to gamma-H2AX foci through interaction with the FHA/BRCT domain. *Curr Biol,* 12**,** 1846-51.
KOOISTRA, A. J., KANEV, G. K., VAN LINDEN, O. P., LEURS, R., DE ESCH, I. J. & DE GRAAF, C. 2016. KLIFS: a structural kinase-ligand interaction database. *Nucleic Acids Res,* 44**,** D365-71.
KOPS, G. J., WEAVER, B. A. & CLEVELAND, D. W. 2005. On the road to cancer: aneuploidy and the mitotic checkpoint. *Nat Rev Cancer,* 5**,** 773-85.
KUBBUTAT, M. H., JONES, S. N. & VOUSDEN, K. H. 1997. Regulation of p53 stability by Mdm2. *Nature,* 387**,** 299-303.
KUERBITZ, S. J., PLUNKETT, B. S., WALSH, W. V. & KASTAN, M. B. 1992. Wild-type p53 is a cell cycle checkpoint determinant following irradiation. *Proc Natl Acad Sci U S A,* 89**,** 7491-5.
LEE, K. K., HARAGUCHI, T., LEE, R. S., KOUJIN, T., HIRAOKA, Y. & WILSON, K. L. 2001. Distinct functional domains in emerin bind lamin A and DNA-bridging protein BAF. *J Cell Sci,* 114**,** 4567-73.
LEE, M. J. & YAFFE, M. B. 2016. Protein Regulation in Signal Transduction. *Cold Spring Harb Perspect Biol,* 8.
LEE, W., TONELLI, M. & MARKLEY, J. L. 2015. NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy. *Bioinformatics,* 31**,** 1325-7.
LIU, X., LI, M., XIA, X., LI, X. & CHEN, Z. 2017. Mechanism of chromatin remodelling revealed by the Snf2-nucleosome structure. *Nature,* 544**,** 440-445.
LO CONTE, L., CHOTHIA, C. & JANIN, J. 1999. The atomic structure of protein-protein recognition sites. *J Mol Biol,* 285**,** 2177-98.
LONG, F., VAGIN, A. A., YOUNG, P. & MURSHUDOV, G. N. 2008. BALBES: a molecular-replacement pipeline. *Acta Crystallographica Section D,* 64**,** 125-132.
LOPEZ-SANCHEZ, I., VALBUENA, A., VAZQUEZ-CEDEIRA, M., KHADAKE, J., SANZ-GARCIA, M., CARRILLO-JIMENEZ, A. & LAZO, P. A. 2014. VRK1 interacts with p53 forming a basal complex that is activated by UV-induced DNA damage. *FEBS Lett,* 588**,** 692-700.

LOWARY, P. T. & WIDOM, J. 1998. New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. *J Mol Biol,* 276**,** 19-42.
LUGER, K., MADER, A. W., RICHMOND, R. K., SARGENT, D. F. & RICHMOND, T. J. 1997. Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature,* 389**,** 251-60.
LUGER, K., RECHSTEINER, T. J. & RICHMOND, T. J. 1999. Preparation of nucleosome core particle from recombinant histones. *Methods Enzymol,* 304**,** 3-19.
MA, H. T. & POON, R. Y. 2011. How protein kinases co-ordinate mitosis in animal cells. *Biochem J,* 435**,** 17-31.
MADHUSUDAN, AKAMINE, P., XUONG, N. H. & TAYLOR, S. S. 2002. Crystal structure of a transition state mimic of the catalytic subunit of cAMP-dependent protein kinase. *Nat Struct Biol,* 9**,** 273-7.
MAESHIMA, K., HIHARA, S. & ELTSOV, M. 2010. Chromatin structure: does the 30-nm fibre exist in vivo? *Curr Opin Cell Biol,* 22**,** 291-7.
MAKDE, R. D., ENGLAND, J. R., YENNAWAR, H. P. & TAN, S. 2010. Structure of RCC1 chromatin factor bound to the nucleosome core particle. *Nature,* 467**,** 562-6.
MANNING, G., WHYTE, D. B., MARTINEZ, R., HUNTER, T. & SUDARSANAM, S. 2002. The protein kinase complement of the human genome. *Science,* 298**,** 1912-34.
MANSHARAMANI, M. & WILSON, K. L. 2005. Direct binding of nuclear membrane protein MAN1 to emerin in vitro and two modes of binding to barrier-to-autointegration factor. *J Biol Chem,* 280**,** 13863-70.
MARECHAL, A. & ZOU, L. 2013. DNA damage sensing by the ATM and ATR kinases. *Cold Spring Harb Perspect Biol,* 5.
MCCOY, A. J., GROSSE-KUNSTLEVE, R. W., ADAMS, P. D., WINN, M. D., STORONI, L. C. & READ, R. J. 2007. Phaser crystallographic software. *J Appl Crystallogr,* 40**,** 658-674.
MOLITOR, T. P. & TRAKTMAN, P. 2013. Molecular genetic analysis of VRK1 in mammary epithelial cells: depletion slows proliferation in vitro and tumor growth and metastasis in vivo. *Oncogenesis,* 2**,** e48-.
MOLITOR, T. P. & TRAKTMAN, P. 2014. Depletion of the protein kinase VRK1 disrupts nuclear envelope morphology and leads to BAF retention on mitotic chromosomes. *Mol Biol Cell,* 25**,** 891-903.
MOLL, U. M. & PETRENKO, O. 2003. The MDM2-p53 interaction. *Mol Cancer Res,* 1**,** 1001-8.
MONSALVE, D. M., CAMPILLO-MARCOS, I., SALZANO, M., SANZ-GARCIA, M., CANTARERO, L. & LAZO, P. A. 2016. VRK1 phosphorylates and protects NBS1 from ubiquitination and proteasomal degradation in response to DNA damage. *Biochim Biophys Acta,* 1863**,** 760-9.
MULLER, I. 2017. Guidelines for the successful generation of protein-ligand complex crystals. *Acta Crystallogr D Struct Biol,* 73**,** 79-92.
NEZU, J., OKU, A., JONES, M. H. & SHIMANE, M. 1997. Identification of two novel human putative serine/threonine kinases, VRK1 and VRK2, with structural similarity to vaccinia virus B1R kinase. *Genomics,* 45**,** 327-31.
NICHOLS, R. J. & TRAKTMAN, P. 2004. Characterization of three paralogous members of the Mammalian vaccinia related kinase family. *J Biol Chem,* 279**,** 7934-46.

NICHOLS, R. J., WIEBE, M. S. & TRAKTMAN, P. 2006. The vaccinia-related kinases phosphorylate the N' terminus of BAF, regulating its interaction with DNA and its retention in the nucleus. *Mol Biol Cell,* 17**,** 2451-64.
NIGG, E. A. 2001. Mitotic kinases as regulators of cell division and its checkpoints. *Nat Rev Mol Cell Biol,* 2**,** 21-32.
PATEL, R. Y. & DOERKSEN, R. J. 2010. Protein kinase–inhibitor database: Structural variability of and inhibitor interactions with the protein kinase P-loop. *J Proteome Res,* 9**,** 4433-42.
PAULI, U., CHRYSOGELOS, S., STEIN, G., STEIN, J. & NICK, H. 1987. Protein-DNA interactions in vivo upstream of a cell cycle-regulated human H4 histone gene. *Science,* 236**,** 1308-11.
PELLEGATA, N. S., ANTONIONO, R. J., REDPATH, J. L. & STANBRIDGE, E. J. 1996. DNA damage and p53-mediated cell cycle arrest: a reevaluation. *Proc Natl Acad Sci U S A,* 93**,** 15209-14.
RAZIN, S. V. & GAVRILOV, A. A. 2014. Chromatin without the 30-nm fiber: constrained disorder instead of hierarchical folding. *Epigenetics,* 9**,** 653-7.
SALZANO, M., SANZ-GARCIA, M., MONSALVE, D. M., MOURA, D. S. & LAZO, P. A. 2015. VRK1 chromatin kinase phosphorylates H2AX and is required for foci formation induced by DNA damage. *Epigenetics,* 10**,** 373-83.
SALZANO, M., VAZQUEZ-CEDEIRA, M., SANZ-GARCIA, M., VALBUENA, A., BLANCO, S., FERNANDEZ, I. F. & LAZO, P. A. 2014. Vaccinia-related kinase 1 (VRK1) confers resistance to DNA-damaging agents in human breast cancer by affecting DNA damage response. *Oncotarget,* 5**,** 1770-8.
SAN-MIGUEL, T., PEREZ-BERMUDEZ, P. & GAVIDIA, I. 2013. Production of soluble eukaryotic recombinant proteins in E. coli is favoured in early log-phase cultures induced at low temperature. *Springerplus,* 2**,** 89.
SANZ-GARCIA, M., MONSALVE, D. M., SEVILLA, A. & LAZO, P. A. 2012. Vaccinia-related kinase 1 (VRK1) is an upstream nucleosomal kinase required for the assembly of 53BP1 foci in response to ionizing radiation-induced DNA damage. *J Biol Chem,* 287**,** 23757-68.
SCHEEFF, E. D., ESWARAN, J., BUNKOCZI, G., KNAPP, S. & MANNING, G. 2009. Structure of the pseudokinase VRK3 reveals a degraded catalytic site, a highly conserved kinase fold, and a putative regulatory binding site. *Structure,* 17**,** 128-38.
SCHULER, M., BOSSY-WETZEL, E., GOLDSTEIN, J. C., FITZGERALD, P. & GREEN, D. R. 2000. p53 induces apoptosis by caspase activation through mitochondrial cytochrome c release. *J Biol Chem,* 275**,** 7337-42.
SESSA, F. & VILLA, F. 2014. Structure of Aurora B–INCENP in complex with barasertib reveals a potential transinhibitory mechanism. *Acta Crystallogr F Struct Biol Commun,* 70**,** 294-8.
SHIN, J., CHAKRABORTY, G., BHARATHAM, N., KANG, C., TOCHIO, N., KOSHIBA, S., KIGAWA, T., KIM, W., KIM, K. T. & YOON, H. S. 2011. NMR solution structure of human vaccinia-related kinase 1 (VRK1) reveals the C-terminal tail essential for its structural stability and autocatalytic activity. *J Biol Chem,* 286**,** 22131-8.
SIEVERS, F., WILM, A., DINEEN, D., GIBSON, T. J., KARPLUS, K., LI, W., LOPEZ, R., MCWILLIAM, H., REMMERT, M., SODING, J., THOMPSON, J. D. & HIGGINS, D. G. 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol,* 7**,** 539.

SMART, O., BRANDL, M., FLENSBURG, C., KELLER, P. & PACIOREK, W. 2008. *Refinement with local structure similarity restraints (LSSR) enables exploitation of information from related structures and facilitates use of NCS.*
SMART, O. S., WOMACK, T. O., FLENSBURG, C., KELLER, P., PACIOREK, W., SHARFF, A., VONRHEIN, C. & BRICOGNE, G. 2012. Exploiting structure similarity in refinement: automated NCS and target-structure restraints in BUSTER. *Acta Crystallogr D Biol Crystallogr,* 68**,** 368-80.
SÖRENSON, S. 1909. Enzyme studies II. The measurement and meaning of hydrogen ion concentration in enzymatic processes. *Biochemische Zeitschrift,* 21**,** 131-200.
TANRAMLUK, D., SCHREYER, A., PITT, W. R. & BLUNDELL, T. L. 2009. On the origins of enzyme inhibitor selectivity and promiscuity: a case study of protein kinase binding to staurosporine. *Chem Biol Drug Des,* 74**,** 16-24.
TAYLOR, S. S. & KORNEV, A. P. 2011. Protein Kinases: Evolution of Dynamic Regulatory Proteins. *Trends Biochem Sci,* 36**,** 65-77.
THOMA, F., KOLLER, T. & KLUG, A. 1979. Involvement of histone H1 in the organization of the nucleosome and of the salt-dependent superstructures of chromatin. *J Cell Biol,* 83**,** 403-27.
TREMETHICK, D. J. 2007. Higher-order structures of chromatin: the elusive 30 nm fiber. *Cell,* 128**,** 651-4.
VALBUENA, A., LOPEZ-SANCHEZ, I. & LAZO, P. A. 2008. Human VRK1 is an early response gene and its loss causes a block in cell cycle progression. *PLoS One,* 3**,** e1642.
VALBUENA, A., LOPEZ-SANCHEZ, I., VEGA, F. M., SEVILLA, A., SANZ-GARCIA, M., BLANCO, S. & LAZO, P. A. 2007a. Identification of a dominant epitope in human vaccinia-related kinase 1 (VRK1) and detection of different intracellular subpopulations. *Arch Biochem Biophys,* 465**,** 219-26.
VALBUENA, A., SANZ-GARCIA, M., LOPEZ-SANCHEZ, I., VEGA, F. M. & LAZO, P. A. 2011. Roles of VRK1 as a new player in the control of biological processes required for cell division. *Cell Signal,* 23**,** 1267-72.
VALBUENA, A., SUAREZ-GAUTHIER, A., LOPEZ-RIOS, F., LOPEZ-ENCUENTRA, A., BLANCO, S., FERNANDEZ, P. L., SANCHEZ-CESPEDES, M. & LAZO, P. A. 2007b. Alteration of the VRK1-p53 autoregulatory loop in human lung carcinomas. *Lung Cancer,* 58**,** 303-9.
VALBUENA, A., VEGA, F. M., BLANCO, S. & LAZO, P. A. 2006. p53 Downregulates Its Activating Vaccinia-Related Kinase 1, Forming a New Autoregulatory Loop. *Mol Cell Biol,* 26**,** 4782-93.
VAN HOOSER, A., GOODRICH, D. W., ALLIS, C. D., BRINKLEY, B. R. & MANCINI, M. A. 1998. Histone H3 phosphorylation is required for the initiation, but not maintenance, of mammalian chromosome condensation. *J Cell Sci,* 111 (Pt 23)**,** 3497-506.
VASUDEVAN, D., CHUA, E. Y. & DAVEY, C. A. 2010. Crystal structures of nucleosome core particles containing the '601' strong positioning sequence. *J Mol Biol,* 403**,** 1-10.
VEGA, F. M., SEVILLA, A. & LAZO, P. A. 2004. p53 Stabilization and accumulation induced by human vaccinia-related kinase 1. *Mol Cell Biol,* 24**,** 10366-80.
VEREDAS, F. J., CANTON, F. R. & ALEDO, J. C. 2017. Methionine residues around phosphorylation sites are preferentially oxidized in vivo under stress conditions. *Sci Rep,* 7**,** 40403.

WALLACE, A. C., LASKOWSKI, R. A. & THORNTON, J. M. 1995. LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. *Protein Eng,* 8**,** 127-34.
WAN, G., LIU, Y., HAN, C., ZHANG, X. & LU, X. 2014. Noncoding RNAs in DNA Repair and Genome Integrity. *Antioxid Redox Signal,* 20**,** 655-77.
WANG, F. & HIGGINS, J. M. 2013. Histone modifications and mitosis: countermarks, landmarks, and bookmarks. *Trends Cell Biol,* 23**,** 175-84.
WEI, Y., YU, L., BOWEN, J., GOROVSKY, M. A. & ALLIS, C. D. 1999. Phosphorylation of histone H3 is required for proper chromosome condensation and segregation. *Cell,* 97**,** 99-109.
WILLIAMS, A. B. & SCHUMACHER, B. 2016. p53 in the DNA-Damage-Repair Process. *Cold Spring Harb Perspect Med,* 6.
WINN, M. D., BALLARD, C. C., COWTAN, K. D., DODSON, E. J., EMSLEY, P., EVANS, P. R., KEEGAN, R. M., KRISSINEL, E. B., LESLIE, A. G., MCCOY, A., MCNICHOLAS, S. J., MURSHUDOV, G. N., PANNU, N. S., POTTERTON, E. A., POWELL, H. R., READ, R. J., VAGIN, A. & WILSON, K. S. 2011. Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr,* 67**,** 235-42.
WU, P., NIELSEN, T. E. & CLAUSEN, M. H. 2015. FDA-approved small-molecule kinase inhibitors. *Trends Pharmacol Sci,* 36**,** 422-39.
XU, D., BAI, J., DUAN, Q., COSTA, M. & DAI, W. 2009. Covalent modifications of histones during mitosis and meiosis. *Cell Cycle,* 8**,** 3688-94.
ZHANG, J., SALMINEN, A., YANG, X., LUO, Y., WU, Q., WHITE, M., GREENHAW, J., REN, L., BRYANT, M., SALMINEN, W., PAPOIAN, T., MATTES, W. & SHI, Q. 2017. Effects of 31 FDA approved small-molecule kinase inhibitors on isolated rat liver mitochondria. *Arch Toxicol,* 91**,** 2921-2938.
ZHAO, Z., LIU, Q., BLIVEN, S., XIE, L. & BOURNE, P. E. 2017. Determining Cysteines Available for Covalent Inhibition Across the Human Kinome. *J Med Chem,* 60**,** 2879-2889.
ZHENG, J., TRAFNY, E. A., KNIGHTON, D. R., XUONG, N. H., TAYLOR, S. S., TEN EYCK, L. F. & SOWADSKI, J. M. 1993. 2.2 A refined crystal structure of the catalytic subunit of cAMP-dependent protein kinase complexed with MnATP and a peptide inhibitor. *Acta Crystallogr D Biol Crystallogr,* 49**,** 362-5.
ZHENG, R., GHIRLANDO, R., LEE, M. S., MIZUUCHI, K., KRAUSE, M. & CRAIGIE, R. 2000. Barrier-to-autointegration factor (BAF) bridges DNA in a discrete, higher-order nucleoprotein complex. *Proc Natl Acad Sci U S A,* 97**,** 8997-9002.
ZHONG, R., ROEDER, R. G. & HEINTZ, N. 1983. The primary structure and expression of four cloned human histone genes. *Nucleic Acids Res,* 11**,** 7409-25.
ZHOU, B. R., FENG, H., KATO, H., DAI, L., YANG, Y., ZHOU, Y. & BAI, Y. 2013. Structural insights into the histone H1-nucleosome complex. *Proc Natl Acad Sci U S A,* 110**,** 19390-5.
ZHUANG, X., SEMENOVA, E., MARIC, D. & CRAIGIE, R. 2014. Dephosphorylation of barrier-to-autointegration factor by protein phosphatase 4 and its role in cell mitosis. *J Biol Chem,* 289**,** 1119-27.

Appendices

Appendix I VRK1 proteins sequences and parameters

I-1 VRK1 wild-type (VRK1^{wt}) sequence and its protein parameters *

```
        10         20         30         40         50         60
MPRVKAAQAG RQSSAKRHLA EQFAVGEIIT DMAKKEWKVG LPIGQGGFGC IYLADMNSSE

        70         80         90        100        110        120
SVGSDAPCVV KVEPSDNGPL FTELKFYQRA AKPEQIQKWI RTRKLKYLGV PKYWGSGLHD

       130        140        150        160        170        180
KNGKSYRFMI MDRFGSDLQK IYEANAKRFS RKTVLQLSLR ILDILEYIHE HEYVHGDIKA

       190        200        210        220        230        240
SNLLLNYKNP DQVYLVDYGL AYRYCPEGVH KEYKEDPKRC HDGTIEFTSI DAHNGVAPSR

       250        260        270        280        290        300
RGDLEILGYC MIQWLTGHLP WEDNLKDPKY VRDSKIRYRE NIASLMDKCF PEKNKPGEIA

       310        320        330        340        350        360
KYMETVKLLD YTEKPLYENL RDILLQGLKA IGSKDDGKLD LSVVENGGLK AKTITKKRKK

       370        380        390
EIEESKEPGV EDTEWSNTQT EEAIQTRSRT RKRVQK
```

Number of amino acids: 396

Molecular weight: 45476.04

Theoretical pI: 9.02

Amino acid composition:

Ala (A)	22	5.6%
Arg (R)	24	6.1%
Asn (N)	14	3.5%
Asp (D)	25	6.3%
Cys (C)	6	1.5%
Gln (Q)	15	3.8%
Glu (E)	32	8.1%
Gly (G)	29	7.3%
His (H)	9	2.3%
Ile (I)	24	6.1%
Leu (L)	35	8.8%
Lys (K)	43	10.9%
Met (M)	8	2.0%
Phe (F)	9	2.3%
Pro (P)	17	4.3%
Ser (S)	22	5.6%
Thr (T)	16	4.0%
Trp (W)	6	1.5%
Tyr (Y)	20	5.1%
Val (V)	20	5.1%
Pyl (O)	0	0.0%
Sec (U)	0	0.0%
(B)	0	0.0%
(Z)	0	0.0%
(X)	0	0.0%

Total number of negatively charged residues (Asp + Glu): 57

Total number of positively charged residues (Arg + Lys): 67

Atomic composition:

Carbon	C	2025
Hydrogen	H	3212
Nitrogen	N	564
Oxygen	O	598
Sulfur	S	14

Formula: $C_{2025}H_{3212}N_{564}O_{598}S_{14}$
Total number of atoms: 6413

Extinction coefficients:

Extinction coefficients are in units of M^{-1} cm^{-1}, at 280 nm measured in water.

Ext. coefficient 63175
Abs 0.1% (=1 g/l) 1.389, assuming all pairs of Cys residues form cystines

Ext. coefficient 62800
Abs 0.1% (=1 g/l) 1.381, assuming all Cys residues are reduced

Estimated half-life:

The N-terminal of the sequence considered is M (Met).

The estimated half-life is: 30 hours (mammalian reticulocytes, in vitro).
>20 hours (yeast, in vivo).
>10 hours (Escherichia coli, in vivo).

Instability index:

The instability index (II) is computed to be 39.20
This classifies the protein as stable.

Aliphatic index: 78.31

Grand average of hydropathicity (GRAVY): -0.718

***VRK1^{wt} protein parameters calculated using ExPASy (Bioinfomatics Resource Portal) – ProtParam (Gasteiger et al., 2005).**

I-2 VRK1$^{\Delta 35}$ sequence and its protein parameters*

```
       10          20          30          40          50          60
MPRVKAAQAG  RQSSAKRHLA  EQFAVGEIIT  DMAKKEWKVG  LPIGQGGFGC  IYLADMNSSE

       70          80          90         100         110         120
SVGSDAPCVV  KVEPSDNGPL  FTELKFYQRA  AKPEQIQKWI  RTRKLKYLGV  PKYWGSGLHD

      130         140         150         160         170         180
KNGKSYRFMI  MDRFGSDLQK  IYEANAKRFS  RKTVLQLSLR  ILDILEYIHE  HEYVHGDIKA

      190         200         210         220         230         240
SNLLLNYKNP  DQVYLVDYGL  AYRYCPEGVH  KEYKEDPKRC  HDGTIEFTSI  DAHNGVAPSR

      250         260         270         280         290         300
RGDLEILGYC  MIQWLTGHLP  WEDNLKDPKY  VRDSKIRYRE  NIASLMDKCF  PEKNKPGEIA

      310         320         330         340         350         360
KYMETVKLLD  YTEKPLYENL  RDILLQGLKA  IGSKDDGKLD  LSVVENGGLK  AKTITKKRKK

E
```

Number of amino acids: 361

Molecular weight: 41346.57

Theoretical pI: 9.10

Amino acid composition:

Ala (A)	21	5.8%
Arg (R)	20	5.5%
Asn (N)	13	3.6%
Asp (D)	24	6.6%
Cys (C)	6	1.7%
Gln (Q)	12	3.3%
Glu (E)	25	6.9%
Gly (G)	28	7.8%
His (H)	9	2.5%
Ile (I)	22	6.1%
Leu (L)	35	9.7%
Lys (K)	40	11.1%
Met (M)	8	2.2%
Phe (F)	9	2.5%
Pro (P)	16	4.4%
Ser (S)	19	5.3%
Thr (T)	11	3.0%
Trp (W)	5	1.4%
Tyr (Y)	20	5.5%
Val (V)	18	5.0%
Pyl (O)	0	0.0%
Sec (U)	0	0.0%
(B)	0	0.0%
(Z)	0	0.0%
(X)	0	0.0%

Total number of negatively charged residues (Asp + Glu): 49
Total number of positively charged residues (Arg + Lys): 60

Atomic composition:

Carbon	C	1853
Hydrogen	H	2929
Nitrogen	N	509
Oxygen	O	535
Sulfur	S	14

Formula: $C_{1853}H_{2929}N_{509}O_{535}S_{14}$
Total number of atoms: 5840

Extinction coefficients:

Extinction coefficients are in units of M^{-1} cm^{-1}, at 280 nm measured in water.

Ext. coefficient 57675
Abs 0.1% (=1 g/l) 1.395, assuming all pairs of Cys residues form cystines

Ext. coefficient 57300
Abs 0.1% (=1 g/l) 1.386, assuming all Cys residues are reduced

Estimated half-life:

The N-terminal of the sequence considered is M (Met).

The estimated half-life is: 30 hours (mammalian reticulocytes, in vitro).
>20 hours (yeast, in vivo).
>10 hours (Escherichia coli, in vivo).

Instability index:

The instability index (II) is computed to be 34.43
This classifies the protein as stable.

Aliphatic index: 81.86

Grand average of hydropathicity (GRAVY): -0.618

VRK1$^{\Delta 35}$ protein parameters calculated using ExPASy (Bioinfomatics Resource Portal) – ProtParam.

I-3 VRK1^{cryst} sequence and its protein parameters*

```
   10          20          30          40          50          60
MPRVKAAQAG  RQSSAKRHLA  EQFAVGEIIT  DMAAAAWKVG  LPIGQGGFGC  IYLADMNSSE

         70          80          90         100         110         120
SVGSDAPCVV  KVEPSDNGPL  FTELKFYQRA  AKPEQIQKWI  RTRKLKYLGV  PKYWGSGLHD

        130         140         150         160         170         180
KNGKSYRFMI  MDRFGSDLQK  IYEANAKRFS  RKTVLQLSLR  ILDILEYIHE  HEYVHGDIKA

        190         200         210         220         230         240
SNLLLNYKNP  DQVYLVDYGL  AYRYCPEGVH  KAYAADPKRC  HDGTIEFTSI  DAHNGVAPSR

        250         260         270         280         290         300
RGDLEILGYC  MIQWLTGHLP  WEDNLKDPKY  VRDSKIRYRE  NIASLMDKCF  PAANAPGEIA

        310         320         330         340         350         360
KYMETVKLLD  YTEKPLYENL  RDILLQGLKA  IGSKDDGKLD  LSVVENGGLK  AKTITKKRAA


EIEE
```

Number of amino acids: 364

Molecular weight: 41086.15

Theoretical pI: 8.75

Amino acid composition:

Ala (A)	32	8.8%
Arg (R)	20	5.5%
Asn (N)	13	3.6%
Asp (D)	24	6.6%
Cys (C)	6	1.6%
Gln (Q)	12	3.3%
Glu (E)	23	6.3%
Gly (G)	28	7.7%
His (H)	9	2.5%
Ile (I)	23	6.3%
Leu (L)	35	9.6%
Lys (K)	33	9.1%
Met (M)	8	2.2%
Phe (F)	9	2.5%
Pro (P)	16	4.4%
Ser (S)	19	5.2%
Thr (T)	11	3.0%
Trp (W)	5	1.4%
Tyr (Y)	20	5.5%
Val (V)	18	4.9%
Pyl (O)	0	0.0%
Sec (U)	0	0.0%
(B)	0	0.0%
(Z)	0	0.0%
(X)	0	0.0%

Total number of negatively charged residues (Asp + Glu): 47
Total number of positively charged residues (Arg + Lys): 53

Atomic composition:

Carbon	C	1840
Hydrogen	H	2897
Nitrogen	N	505
Oxygen	O	534
Sulfur	S	14

Formula: $C_{1840}H_{2897}N_{505}O_{534}S_{14}$
Total number of atoms: 5790

Extinction coefficients:

Extinction coefficients are in units of M^{-1} cm^{-1}, at 280 nm measured in water.

Ext. coefficient 57675
Abs 0.1% (=1 g/l) 1.404, assuming all pairs of Cys residues form cystines

Ext. coefficient 57300
Abs 0.1% (=1 g/l) 1.395, assuming all Cys residues are reduced

Estimated half-life:

The N-terminal of the sequence considered is M (Met).

The estimated half-life is: 30 hours (mammalian reticulocytes, in vitro).
>20 hours (yeast, in vivo).
>10 hours (Escherichia coli, in vivo).

Instability index:

The instability index (II) is computed to be 37.30
This classifies the protein as stable.

Aliphatic index: 85.27

Grand average of hydropathicity (GRAVY): -0.452

***VRK1^{cryst} protein parameters calculated using ExPASy (Bioinfomatics Resource Portal) – ProtParam.**

I-4 VRK1 constructs multiple sequence alignment by Clustal Omega

```
VRK1wt          MPRVKAAQAGRQSSAKRHLAEQFAVGEIITDMAKKEWKVGLPIGQGGFGCIYLADMNSSE    60
VRK1Δ35         MPRVKAAQAGRQSSAKRHLAEQFAVGEIITDMAKKEWKVGLPIGQGGFGCIYLADMNSSE    60
VRK1cryst       MPRVKAAQAGRQSSAKRHLAEQFAVGEIITDMAAAAWKVGLPIGQGGFGCIYLADMNSSE    60
                ********************************   ************************

VRK1wt          SVGSDAPCVVKVEPSDNGPLFTELKFYQRAAKPEQIQKWIRTRKLKYLGVPKYWGSGLHD    120
VRK1Δ35         SVGSDAPCVVKVEPSDNGPLFTELKFYQRAAKPEQIQKWIRTRKLKYLGVPKYWGSGLHD    120
VRK1cryst       SVGSDAPCVVKVEPSDNGPLFTELKFYQRAAKPEQIQKWIRTRKLKYLGVPKYWGSGLHD    120
                ************************************************************

VRK1wt          KNGKSYRFMIMDRFGSDLQKIYEANAKRFSRKTVLQLSLRILDILEYIHEHEYVHGDIKA    180
VRK1Δ35         KNGKSYRFMIMDRFGSDLQKIYEANAKRFSRKTVLQLSLRILDILEYIHEHEYVHGDIKA    180
VRK1cryst       KNGKSYRFMIMDRFGSDLQKIYEANAKRFSRKTVLQLSLRILDILEYIHEHEYVHGDIKA    180
                ************************************************************

VRK1wt          SNLLLNYKNPDQVYLVDYGLAYRYCPEGVHKEYKEDPKRCHDGTIEFTSIDAHNGVAPSR    240
VRK1Δ35         SNLLLNYKNPDQVYLVDYGLAYRYCPEGVHKEYKEDPKRCHDGTIEFTSIDAHNGVAPSR    240
VRK1cryst       SNLLLNYKNPDQVYLVDYGLAYRYCPEGVHKAYAADPKRCHDGTIEFTSIDAHNGVAPSR    240
                ******************************* *  ************************

VRK1wt          RGDLEILGYCMIQWLTGHLPWEDNLKDPKYVRDSKIRYRENIASLMDKCFPEKNKPGEIA    300
VRK1Δ35         RGDLEILGYCMIQWLTGHLPWEDNLKDPKYVRDSKIRYRENIASLMDKCFPEKNKPGEIA    300
VRK1cryst       RGDLEILGYCMIQWLTGHLPWEDNLKDPKYVRDSKIRYRENIASLMDKCFPAANAPGEIA    300
                ************************************************  * *****

VRK1wt          KYMETVKLLDYTEKPLYENLRDILLQGLKAIGSKDDGKLDLSVVENGGLKAKTITKKRKK    360
VRK1Δ35         KYMETVKLLDYTEKPLYENLRDILLQGLKAIGSKDDGKLDLSVVENGGLKAKTITKKRKK    360
VRK1cryst       KYMETVKLLDYTEKPLYENLRDILLQGLKAIGSKDDGKLDLSVVENGGLKAKTITKKRAA    360
                ************************************************************

VRK1wt          EIEESKEPGVEDTEWSNTQTEEAIQTRSRTRKRVQK 396
VRK1Δ35         E---------------------------------- 364
VRK1cryst       EIEE------------------------------- 364
                *
```

All the amino acids identical in all three VRK1 constructs are marked by the asterisk.

VRK1^{wt} C-terminal tail from amino acids 362 to 396, highlighted in blue.

The 11 lysine (K) or glutamate (E) residues substituted into alanine (A) in VRK1^{cryst} construct are shown in red.

Appendix II Histone protein and DNA sequences

The sequences of widom 601 – 145-bp DNA, histone H2A, H2B, H3 and H4.

Widom 601 – 145-bp DNA sequence

ATCAGAATCCCGGTGCCGAGGCCGCTCAATTGGTCGTAGACAGCTCTAGCACC
GCTTAAACGCACGTACGCGCTGTCCCCCGCGTTTTAACCGCCAAGGGGATTAC
TCCCTAGTCTCCAGGCACGTGTCAGATATATACATCGAT

Histone H2A [Homo sapiens] 130 amino acids

MSGRGKQGGKARAKAKTRSSRAGLQFPVGRVHRLLRKGNYSERVGAGAPVYLA
AVLEYLTAEILELAGNAARDNKKTRIIPRHLQLAIRNDEELNKLLGRVTIAQG
GVLPNIQAVLLPKKTESHHKAKGK

Histone H2B [Homo sapiens] 126 amino acids

MPEPAKSAPAPKKGSKKAVTKAQKKDGKKRKRSRKESYSVYVYKVLKQVHPDT

GISSKAMGIMNSFVNDIFERIAGEASRLAHYNKRSTITSREIQTAVRLLLPGE

LAKHAVSEGTKAVTKYTSAK

Histone H3 [Homo sapiens] 136 amino acids

MARTKQTARKSTGGKAPRKQLATKAARKSAPATGGVKKPHRYRPGTVALREIR

RYQKSTELLIRKLPFQRLVREIAQDFKTDLRFQSSAVMALQEACEAYLVGLFE

DTNLCAIHAKRVTIMPKDIQLARRIRGERA

Histone H4 [Homo sapiens] 103 amino acids

MSGRGKGGKGLGKGGAKRHRKVLRDNIQGITKPAIRRLARRGGVKRISGLIYE

ETRGVLKVFLENVIRDAVTYTEHAKRKTVTAMDVVYALKRQGRTLYGFGG

Appendix III Materials used in this thesis

III-1 Chemicals

Chemicals	Company
2-mercaptoethanol (BME)	Sigma-Aldrich, USA
2-Propanol	Merck, USA
Absolute ethanol (100%)	Merck, USA
Acetic acid glacial	Sigma-Aldrich, USA
Acrylamide/Bis solution (29:1)	Bio-Rad, USA
Amberlite Mb-20 resin	Sigma-Aldrich, USA
Ammonium chloride	Sigma-Aldrich, USA
Ammonium persulfate (APS)	Sigma-Aldrich, USA
Ammonium sulfate	Sigma-Aldrich, USA
Bacto agar	BD Difco, USA
Bacto tryptone	BD Difco, USA
Bradford assay dye solution (5x)	Bio-Rad, USA
Bromophenol blue	Affymetrix /USB, USA
Calcium chloride	Merck, USA
Chloroform	Fisher Scientific, USA
Concentrated hydrochloric acid (37% HCl)	Merck, USA
Coomassie brilliant blue G-250	Bio-Rad, USA
D-Glucose	Sigma-Aldrich, USA
Dimethyl sulfoxide (DMSO)	Sigma-Aldrich, USA
Dithiothreitol (DTT)	Gold Biotechnology, USA
Ethylenediaminetetraacetic acid (EDTA)	Sigma-Aldrich, USA
Gel-red	Biotium, USA
Glycerol	Promega, USA
Glycine	1st Base Biochemicals, SG
Guanidine hydrochloride	Sigma-Aldrich, USA
Imidazole	Sigma-Aldrich, USA
Isoamyl alcohol	Sigma-Aldrich, USA
Isopropyl β-D-1-thiogalactopyranoside (IPTG)	Gold Biotechnology, USA
Manganese (II) chloride	Merck, USA
Magnesium sulfate	Sigma-Aldrich, USA
Phenylmethylsulfonyl fluoride (PMSF)	Sigma-Aldrich, USA
Polyethylene glycol (PEG) 3350	Sigma-Aldrich, USA
Polyethylene glycol (PEG) 6000	Sigma-Aldrich, USA
Potassium acetate	Sigma-Aldrich, USA
Potassium cacodylate	Sigma-Aldrich, USA
Potassium chloride	Sigma-Aldrich, USA

Potassium hydroxide	Merck, USA
Potassium phosphate monobasic	Merck, USA
Potassium phosphate dibasic	Merck, USA
Sodium azide	Sigma-Aldrich, USA
Sodium acetate	Sigma-Aldrich, USA
Sodium chloride	Merck, USA
Sodium dodecyl sulfate (SDS)	Hoefer, USA
Sodium hydroxide	Merck, USA
Sodium phosphate monobasic	Merck, USA
Sodium phosphate dibasic	Merck, USA
Sucrose	Affymetrix /USB, USA
Tris-acetate-EDTA (TAE) 10x	Invitrogen, USA
Tris-borate-EDTA (TBE) 10x	Invitrogen, USA
TEMED	Sigma-Aldrich, USA
Thiamine-HCl	Sigma-Aldrich, USA
Tris base	Promega, USA
Triton X-100	Affymetrix /USB, USA
Tween-20 (10%)	Bio-Rad, USA
UltraPure^{TM} buffer-saturated phenol	Invitrogen, USA
Urea	Affymetrix /USB, USA
Yeast extract	Affymetrix /USB, USA
^{15}N-labeled ammonium chloride	Cambridge Isotope Laboratories, USA
Deuterium oxide, 99.9% (D_2O)	Cambridge Isotope Laboratories, USA

III-2 Bacterial culture media

a. Luria-Bertani (LB) broth

5 g of yeast extract, 5 g of sodium chloride (NaCl), and 10 g of bacto tryptone were weighed into 1 L of milli-Q water and autoclaved before use.

b. Luria-Bertani (LB) agar

2.5 g of yeast extract, 2.5 g of NaCl, 5 g of bacto tryptone and 8.5 g of bacto agar were mixed in 500 mL of milli-Q water and then autoclaved. After autoclaved, the LB agar was cooled to about 50°C and respective antibiotic was added. 20 mL of LB agar was pour into each petri dish and left to solidify, after then the agar plates were kept at 4°C fridge.

c. 2x TY media

5 g of NaCl, 10 g of yeast extract and 16 g of bacto tryptone were weighted into 1 L of milli-Q water and autoclaved before use.

d. Terrific broth (TB)

The TB medium recipe was shown in 1.

Table III-1 Terrific broth recipe

Components	500 mL	Working concentration
Milli Q water	450 ml	
Bacto Tryptone	6 g	
Yeast extracts	12 g	
KH_2PO_4 (1 M) *	8.5 mL	17 mM
K_2HPO_4 (1 M) *	36 mL	72 mM
Glycerol (50% v/v) *	4 mL	0.4%

*1 M of KH_2PO_4, 1 M of K_2HPO_4 and 50% v/v of glycerol solutions were prepared and autoclaved beforehand.

1 M of potassium phosphate monobasic solution (KH_2PO_4) or 1 M of potassium phosphate dibasic solution (K_2HPO_4) was prepared by dissolving 136.09 g of KH_2PO_4 powder or 174.18 g of K_2HPO_4 powder in milli-Q water, then top up to 1 L volume, respectively. 50% v/v glycerol was prepared by

dissolving 250 mL of glycerol in 250 mL of milli-Q water. 6 g of bacto tryptone and 12 g of yeast extract were mixed in 450 mL of milli-Q water in a 2 L culture flask and autoclaved separately. After the 450 mL of medium was cooled, the above ingredients were added into it in the laminar flow hood.

e. M9 minimal

The M9 minimal medium used for expressing labeled protein was prepared according to Table III-2.

Table III-2 M9 minimal medium recipe

Reagents	500 mL	Working conc.
$Na_2HPO_4.H_2O$ (0.5 M)	52.6 mL	52.6 mM
KH_2PO_4 (1 M)	13.3 mL	26.6 mM
NaCl (5 M)	1 mL	10 mM
$MgSO_4$ (1 M)	0.6 mL	1.2 mM
Thiamine-HCl (0.5% w/v)	0.6 mL	0.0006%
$CaCl_2$ (0.1 M)	0.6 mL	0.12 mM
D-Glucose (20% w/v)	11.3 mL	0.452%
Ammonium Chloride	0.5 g	0.1 %
Autoclave water to make up	500 mL	

All the reagents above except ammonium chloride, were prepared and autoclaved before use. Calcium chloride ($CaCl_2$) solution was added last to the M9 medium and vigorously shaken to dissolve to prevent precipitation. To prepare ^{15}N-labeled protein samples, 1 g of ^{15}N-labeled ammonium chloride was used in place of the unlabeled ammonium chloride.

III-3 Buffer for purification, extraction and reconstitution

a. VRK1 protein purification

The composition of the buffers used for protein purification were listed in Table III-3. All the buffers were prepared in milli-Q water to a final volume of 500 mL and sterilized through 0.2 μM PES membrane filter unit before use. The reducing agent was only added right before each experiment to minimize degradation during storage.

Table III-3 VRK1 protein purification buffer recipes

Buffers for protein purification	Composition
Cell lysis buffer	20 mM Sodium phosphate (pH 7.2), 500 mM NaCl, 1 mM PMSF, 5 mM BME
Wash buffer	20 mM Sodium phosphate (pH 7.2), 1 M NaCl, 1 mM PMSF, 20 mM imidazole, 5 mM BME
Elution buffer	20 mM Sodium phosphate (pH 6.0), 1M NaCl, 1 mM PMSF, 500 mM imidazole, 5 mM BME
Gel filtration buffer	20 mM Sodium phosphate (pH 7.0), 150 mM NaCl, 1 mM DTT, 0.01% NaN_3 or 20 mM Tris-HCl (pH 7.5), 150 mM NaCl, 1 mM DTT, 0.01% NaN_3

b. Histones protein purification

The composition of the buffers for histones protein purification were listed in Table III-4. All the buffers were prepared in milli-Q water to a final volume of 1 L and sterilized by 0.2 μM PES membrane filter unit before use. The reducing agent was added prior each experiment.

Table III-4 Histone protein purification buffer recipes.

Buffers for histone purification	Composition
Histone wash buffer	50 mM Tris-HCl (pH 7.5), 100 mM NaCl, 1 mM BME
Histone wash buffer with Triton X-100	50 mM Tris-HCl (pH 7.5), 100 mM NaCl, 1 mM BME, 1% Triton X-100
S_{200} unfolding buffer	7 M Guanidine-HCl, 20 mM sodium acetate (pH 5.2), 10 mM DTT
SAUDE 1000	7 M Urea, 20 mM sodium acetate, 1 M NaCl, 5 mM BME, 1 mM EDTA
SAUDE 200	7 M Urea, 20 mM sodium acetate, 200 mM NaCl, 5 mM BME, 1 mM EDTA
Histone octamer (HO) unfolding buffer	7 M Guanidine-HCl, 10 mM Tris-HCl (pH 7.5), 10 mM DTT
Refolding buffer	10 mM Tris-HCl (pH 7.5), 2 M NaCl, 1 mM EDTA, 5 mM BME

c. 145-bp 601 DNA fragment extraction and purification

The composition of the buffers used for 601-145 bp DNA extraction and purification was listed in Table III-5. All the buffers were prepared in autoclaved water to a final volume of 1 L. Alkaline lysis buffer III was autoclaved and cooled in the 4 °C fridge prior to DNA extraction.

Table III-5 145-bp 601 DNA extraction and purification buffer recipes.

Buffers for DNA purification	Composition
Alkaline lysis buffer I	25 mM Tris-HCl (pH 8.0), 10 mM EDTA, 50 mM Glucose
Alkaline lysis buffer II	200 mM NaOH, 1% w/v SDS
Alkaline lysis buffer III	4 M Potassium acetate, 2 M Acetic acid
TE 10/50	10 mM Tris-HCl (pH 8.0), 50 mM EDTA
TE 10/0.1	10 mM Tris-HCl (pH8.0), 0.1 mM EDTA

d. Nucleosome core particle (NCP) reconstitution

The NCP reconstitution buffers were prepared with different constitution of KCl. All the buffers were prepared in milli-Q water to a final volume of 1 L and sterilized through 0.2 μM filter unit before use. The DTT was added only during NCP reconstitution. The composition of each buffer was listed in Table III-6.

Table III-6 NCP reconstitution buffer recipes.

Buffers for NCP reconstitution	Composition
TCS-2	20 mM Tris-HCl (pH 7.5), 1 mM EDTA, 1 mM DTT, 2 M KCl
TCS-0.85	20 mM Tris-HCl (pH 7.5), 1 mM EDTA, 1 mM DTT, 0.85 M KCl
TCS-0.65	20 mM Tris-HCl (pH 7.5), 1 mM EDTA, 1 mM DTT, 0.65 M KCl
TCS-0.45	20 mM Tris-HCl (pH 7.5), 1 mM EDTA, 1 mM DTT, 0.45 M KCl
TCS-0	20 mM Tris-HCl (pH 7.5), 1 mM EDTA, 1 mM DTT

III-4 Stock solutions

All stock solutions with a total volume of less than 50 mL were sterilized by 0.2 μM syringe filters. Stock solutions above 100 mL in volume were filter sterilized using the 0.2 μM polyethersulfone (PES) membrane filter units (Nalgene Labware, USA). Stock solutions prepared for DNA extraction and purification experiments were sterilized by autoclaving.

a. 1 M Isopropyl β-D-1-thiogalactopyranoside (IPTG)

   2.38 g of IPTG powder was dissolved in milli-Q water and top up to 10 mL. It was aliquoted into 1 mL per eppendorf tube and stored at -20°C.

b. 0.1 M Phenylmethylsulfonyl fluoride (PMSF)

   0.17 g of PMSF powder was dissolved in 10 mL of 100% isopropanol, then aliquoted into 1 mL per tube and kept at -20ºC.

c. 1 M Dithiothreitol (DTT)

   1.54 g of DTT powder was prepared in milli-Q water, then top up to 10 mL. It was aliquoted into 1 mL per tube and store at -20°C.

d. 10% w/v Sodium azide (NaN_3)

   5 g of NaN_3 powder was weighed and dissolved in milli-Q water, then top up to a final volume of 50 mL.

e. Sodium phosphate buffer

   Sodium phosphate buffer was prepared according to the formula written by (Sörenson, 1909). The X buffer, 0.2 M of dibasic sodium phosphate ($Na_2HPO_4.2H_2O$), was prepared by dissolving 35.6 g of $Na_2HPO_4.2H_2O$ powder in 1 L of milli-Q water. The Y buffer, 0.2 M of monobasic sodium phosphate ($NaH_2PO_4.H_2O$), was prepared by dissolving 27.6 g of $NaH_2PO_4.H_2O$ powder in 1 L of milli-Q water.

Following Table III-7, 100 mL of 0.2 M of sodium phosphate buffer at different pH can be prepared by mixing *a* ml of X with *b* ml of Y.

Table III-7 0.2 M Sodium phosphate buffer at different pH formula.

X (*a* mL)	Y (*b* mL)	pH	X (*a* mL)	Y (b mL)	pH
93.5	6.5	5.7	45	55	6.9
92	8	5.8	39	61	7.0
90	10	5.9	33	67	7.1
87.5	12.5	6.0	28	72	7.2
85	15	6.1	23	77	7.3
81.5	18.5	6.2	19	81	7.4
77.5	22.5	6.3	16	84	7.5
73.5	26.5	6.4	13	87	7.6
68.5	31.5	6.5	10.5	90.5	7.7
62.5	37.5	6.6	8.5	91.5	7.8
56.5	43.5	6.7	7	93	7.9
51	49	6.8	5.3	94.7	8.0

f. 5 M Sodium chloride (NaCl)

292.2 g of NaCl powder was dissolved in milli-Q water and top up to 1 L.

g. 1 M Tris-HCl (pH 7.5 or 8.0)

121.14 g of Tris base powder was dissolved in milli-Q water and the pH was adjusted to 7.5 or 8.0 using 37% concentrated hydrochloric acid (HCl) solution. The pH adjusted Tris-HCl solution was then top up to a final volume of 1 L.

h. 1 M HEPES (pH 7.5)

238.3 g of HEPES powder, was weighed and dissolved in milli-Q water, the pH of the solution was adjusted to 7.5 by using 10 M of sodium hydroxide (NaOH) solution. 1 M HEPES (pH 7.5) solution was then top up to a final volume of 1 L.

i. 1 M Imidazole (pH 7.0)

68.07 g of imidazole powder was dissolved in milli-Q water and the pH was adjusted to 7.0 using 37% concentrated HCl solution. The imidazole at pH 7.0 was top up with milli-Q water to 1 L.

j. 10% Ammonium Persulfate (APS)

10 g of APS powder, was dissolved in milli-Q water, later top up to 10 mL with milli-Q water.

k. 9 M Urea (de-ionized)

1081 g of urea powder was dissolved and top up to 2 L with milli-Q water. The urea solution was further de-ionized using the Amberlite MB-20 resin before use.

l. 8 M Guanidine hydrochloride (Gdn-HCl)

152.84 g of Gdn-HCl powder was dissolved and top up to 500 mL with milli-Q water.

m. 3 M Sodium acetate (pH 5.2)

123.05 g of sodium acetate powder was dissolved in milli-Q water and the pH was adjusted to 5.2 using acetic acid glacial solution. The pH adjusted solution was top up with milli-Q water to 500 mL.

n. 0.5 M Ethylenediaminetetraacetic acid (EDTA) (pH 8.0)

0.5 M EDTA at pH 8.0 used for DNA extraction experiments was prepared by dissolving 73.06 g of EDTA powder in milli-Q water. To dissolve the powder, the pH of the solution had to be adjusted to 8.0 using 10 M NaOH. After then, the solution was top up to 500 mL using milli-Q water.

o. 1 M Magnesium sulfate ($MgSO_4$)

6.02 g of $MgSO_4$ powder was dissolved in milli-Q water and top up to 50 mL and autoclaved before use.

p. 0.1 M calcium chloride ($CaCl_2$)

5.55 g of $CaCl_2$ of powder was weighted and dissolved in milli-Q water and top up to 500 mL. The solution was autoclaved before use.

q. 0.5% w/v Thiamine hydrochloride

0.25 g of Thiamine-HCl powder was dissolved in autoclaved water and top up to 50 mL. The solution was sterilized by syringe filter.

r. 1 M D-Glucose or 20% w/v D-glucose

1 M D-glucose or 20% w/v D-glucose was prepared by dissolving 90.07 g or 100g of glucose powder in milli-Q water and top up to 500 mL, respectively. The solutions were autoclaved before use.

s. 5 M Sodium hydroxide (NaOH)

199.98 g of NaOH powder was weighed and dissolved in autoclaved water in the cold room. The powder was added slowly into the water with stirring to dissolve. The solution was top up to 1 L with autoclaved water.

t. 10% w/v Sodium dodecyl sulfate (SDS)

50 g of SDS powder was added slowly in autoclaved water with stirring to dissolve. The solution was top up to 500 mL with autoclaved water.

u. 40% w/v PEG 6000 and 50% w/v PEG 3350

20 g of Polyethylene glycol (PEG) 6000 or 25 g of PEG 3350 was dissolved in autoclaved water respectively. The solutions were top up to 50 mL using autoclaved water.

III-5 SDS- and DNA- PAGE reagents and gel composition

a. SDS-PAGE and DNA-PAGE reagents

The 6x SDS loading dye was prepared in 10 mL with 300 mM Tris-HCl (pH6.8), 50% v/v glycerol, 12% w/v SDS, 0.2% w/v Bromophenol blue and 600 mM DTT. The loading dye was aliquoted in small volume and kept in -20°C freezer. 6x purple DNA gel loading dye was supplied by NEB, USA.

The formula for the buffers used for gel running were listed in Table III-8. All the buffers were prepared in 1 L volume.

Table III-8 DNA- or SDS- PAGE gel running buffer recipes.

Buffers for gel running	Composition
5x SDS running buffer	125 mM Tris base, 960 mM glycine, 0.5% w/v SDS
WB Transfer buffer (for SDS-PAGE)	25 mM Tris, 192 mM glycine, 0.1% w/v SDS, 20% v/v methanol
WB Transfer buffer (for DNA-PAGE)	25 mM Tris, 192 mM glycine, 20% v/v methanol
20x TBS	500 mM Tris, 60 mM KCl, 2.8 M NaCl
TBS-T	25 mM Tris, 3 mM KCl, 140 mM NaCl, 0.1% v/v Tween-20
Coomassie blue staining buffer	0.25% w/v Coomassie-Brillant Blue G250, 50% v/v methanol, 10% v/v acetic acid glacial
Destaining buffer	45% v/v methanol, 10% v/v acetic acid glacial

b. Gel recipes

The composition of different percentage of DNA-PAGE or SDS-PAGE gel was shown in Table III-9**.**

Table III-9 DNA- or SDS- PAGE gel recipes.

Gel (for 5 mL)	Composition
6% DNA-PAGE	0.25x TBE, 0.16% APS, 6% Acryl/Bis, 0.08% TEMED
8% DNA-PAGE	0.25x TBE, 0.16% APS, 8% Acryl/Bis, 0.08% TEMED
10% SDS-PAGE	390 mM Tris-HCl (pH 8.8), 0.1% APS, 0.1% SDS, 10% Acryl/Bis mix, 0.04% TEMED
12% SDS-PAGE	390 mM Tris-HCl (pH 8.8), 0.1% APS, 0.1% SDS, 12% Acryl/Bis mix, 0.04% TEMED
15% SDS-PAGE	390 mM Tris-HCl (pH 8.8), 0.1% APS, 0.1% SDS, 15% Acryl/Bis mix, 0.04% TEMED
Stacking gel	126 mM Tris-HCl (pH 6.8), 0.1% APS, 0.1% SDS, 5% Acryl/Bis mix, 1% TEMED

Appendix IV Protein Expression and purification results

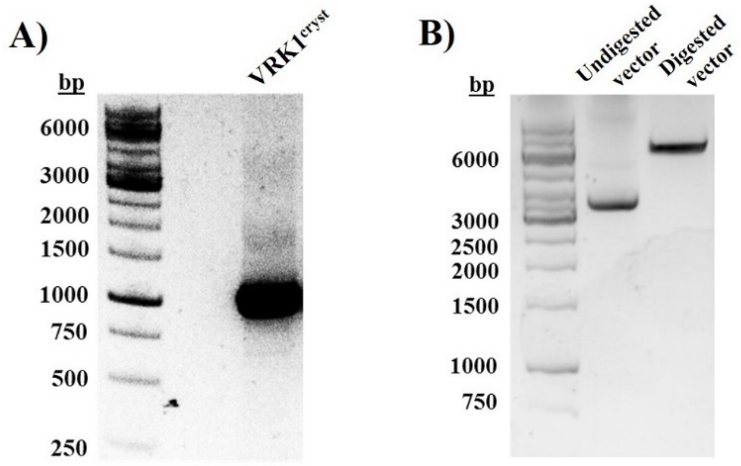


Figure IV-1 Molecular cloning of VRK1^{cryst} into pE-SUMO vector

A) Amplification of VRK1^{cryst} gene by PCR reaction. The PCR amplified VRK1^{cryst} gene with a sequence length of 1,101 bp, is loaded along with the 1 kb DNA marker on 1% agarose gel. B) Digestion of pE-SUMO vector by restriction enzymes. The BsaI/XhoI double digested pE-SUMO vector (~5,600 bp) is ran on the 1% agarose gel, along with the 1 kb DNA marker. Lane 2 shows the undigested vector while lane 3 displays the digested vector.

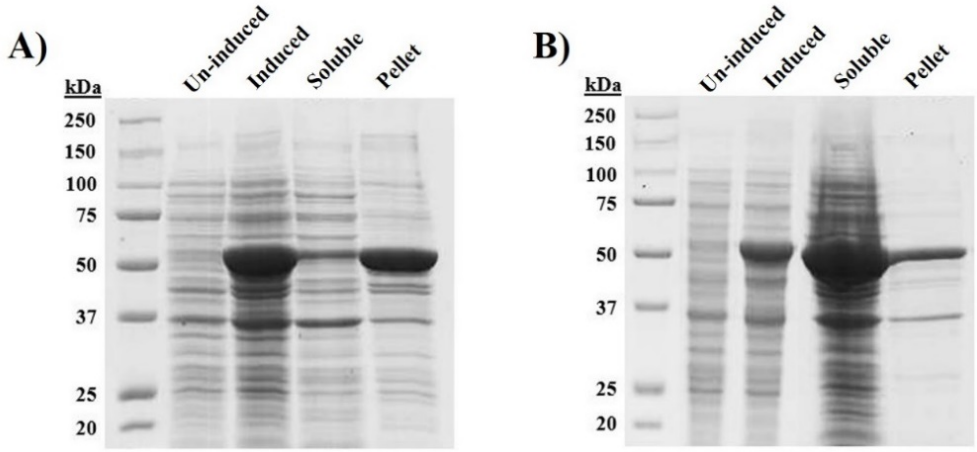
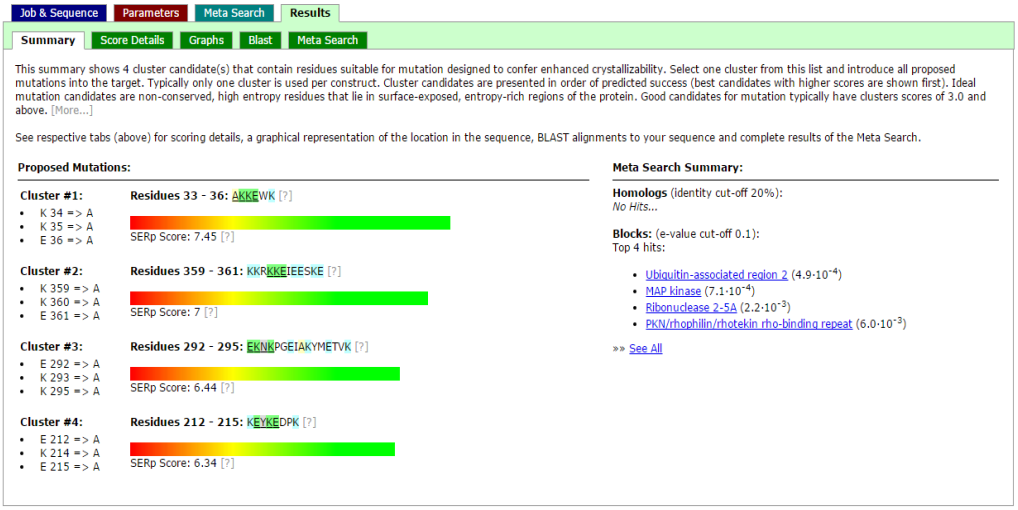


Figure IV-2 VRK1^{cryst} protein expression and solubility optimization.

A) SDS-PAGE analysis of VRK1^{cryst} protein induction at 25°C. The un-induced and induced samples, together with the soluble and pellet fractions obtained after cell lysis are loaded along with protein marker onto the 10% SDS-PAGE gel. Majority of the expressed protein (M.W. 50 kDa) is observed in the pellet fraction. B) SDS-PAGE analysis of VRK1^{cryst} protein induction at 16°C. The un-induced, induced, soluble and pellet fractions collected are analyzed on 10% SDS-PAGE gel, along with protein marker. Most of the expressed protein is observed in the soluble fraction.

Appendix V VRK1 SER results

Four clusters of residues suitable for mutation in VRK1 are identified by using the Surface Entropy Reduction (SER) prediction server (). The results are shown as below.



Appendix VI Sequencing result of VRK1^{cryst} construct

VRK1^{cryst} sequence is shown in blue.

MGHHHHHHGSDSEVNQEAKPEVKPEVKPETHINLKVSDGSSEIFFKIKKTTPL
RRLMEAFAKRQGKEMDSLRFLYDGIRIQADQAPEDLDMEDNDIIEAHREQIGG
MPRVKAAQAGRQSSAKRHLAEQFAVGEIITDMAAAAWKVGLPIGQGGFGCIYL
ADMNSSESVGSDAPCVVKVEPSDNGPLFTELKFYQRAAKPEQIQKWIRTRKLK
YLGVPKYWGSGLHDKNGKSYRFMIMDRFGSDLQKIYEANAKRFSRKTVLQLSL
RILDILEYIHEHEYVHGDIKASNLLLNYKNPDQVYLVDYGLAYRYCPEGVHKA
YAADPKRCHDGTIEFTSIDAHNGVAPSRRGDLEILGYCMIQWLTGHLPWEDNL
KDPKYVRDSKIRYRENIASLMDKCFPAANAPGEIAKYMETVKLLDYTEKPLYE
NLRDILLQGLKAIGSKDDGKLDLSVVENGGLKAKTITKKRAAEIEELEHHHHH
H