Mutational Analysis of Isoform Selectivity and Conformational Equilibria in Protein Kinase Inhibition

Leila Tamara Alexander

Supervised by:
Prof. Stefan Knapp and Prof. Charlotte Deane (University of Oxford)
Dr. Sandra Jacob and Dr. Henrik Möbitz (Novartis AG)

Nuffield Department of Medicine and Wolfson College
University of Oxford

A thesis submitted in partial fulfilment for the degree of
Doctor of Philosophy

Michaelmas Term 2013

To my mother

Mutational Analysis of Isoform Selectivity and Conformational Equilibria in Protein Kinase Inhibition,
Leila Tamara Alexander
Nuffield Department of Medicine and Wolfson College, University of Oxford
A thesis submitted in partial fulfilment for the degree of Doctor of Philosophy, MT 2013.

Abstract

Deregulation of protein kinases is associated with many diseases making them important targets for therapeutic intervention. Kinases can switch between active and inactive conformations that can be targeted by type 1 or type 2 inhibitors respectively. One of the most relevant conformational switches is the 'in' and 'out' movement of the ATP/Mg^{2+} binding motif DFG. Factors modulating the conformational equilibria such as the residue environment of regulatory motifs remain poorly understood despite their importance for drug discovery.

In this thesis, the first model system tested the hypothesis that accessibility of the DFG-out conformation is restricted by the energetic cost of transition between the in and out states. CDK2 was chosen as a target that was thought to have an inaccessible DFG-out conformation, and several point mutations were introduced to promote this conformational transition. Detailed biochemical and biophysical characterisation illustrated that the mutants bound type 2 inhibitors more potently than the wild type. In addition, the wild-type CDK2 was shown to bind type 2 inhibitors in the absence, but not in the presence, of cyclin. The first known CDK2 co-crystal structure in the DFG-out conformation was solved, opening the door to a new class of CDK2 inhibitors.

In the second project, site-directed mutagenesis was used to explore the residues determining inhibitor selectivity between PIM1 and PIM2. Evaluation of ligand binding to the variants and comparison of PIM1 and PIM2 crystal structures showed that flexibility of the phosphate-binding loop was the dominant factor determining the differences in their affinities for ATP and small molecule inhibitors.

These studies illustrate that residues contributing to kinase conformational equilibria can be just as important for inhibitor binding as contact residues formed in the ligand complex.

Acknowledgements

With deep gratitude, I express my thanks to my supervisors, Professor Stefan Knapp, Professor Charlotte Deane, Dr Henrik Möbitz and Dr Sandra Jacob for their guidance, support and advice.

I would like to acknowledge the support of all the past and present members of the Structural Genomics Consortium and the Oxford Protein Bioinformatics Group at the University of Oxford, with whom I had great pleasure working with. In particular, I would like to thank Tracy Keates, Oleg Fedorov, Ildiko Felletar, Sarah Martin, Seb Kelm, Jon Elkins, Pavel Savitsky, Alex Bullock and JP Ebejer for valuable discussions and for introducing me to a variety of methods.

I am extremely grateful for the practical help given to me by the Novartis team. I would like to express my gratitude to Peter Drueckes, Jörg Trappe, Doriano Fabbro, Sandrine Di Bello, Anne-Sophie Mangold, Sylvie Lehmann, Anke Blechschmidt and many others, from whom I have learnt a lot.

I am very grateful to Jola Kopec and Vernon Bailey, who devoted their time and effort to proof read my thesis.

This project was funded by EPSRC and Novartis AG through the Systems Approaches in Bio-Medical Science Industrial Doctorate Centre (SABS IDC).

Finally, I would like to thank my family and friends for their endless love and support and encouragement.

Leila Tamara Alexander
October 2013

Table of contents

Table of figures

Table of tables

Abbreviations

αC-Glu	conserved glutamate residue located in the helix αC
ATP	adenosine triphosphate
BLI	Biolayer interferometry
CAK	CDK-activating kinase
CAMK	Calmodulin-dependent protein kinases
CDK	Cyclin-dependent protein kinase
CML	Chronic myeloid leukaemia
DFG	Aspartate-Phenylalanine-Glycine
ΔG	Gibbs free energy
ΔH	Enthalpy change
ΔS	Entropy change
DSF	Differential scanning fluorimetry
ΔT_m	Change in protein melting temperature
DTT	Dithiothreitol
E.coli	*Escherichia coli*
ESI-TOF MS	Electrospray ionisation time-of-flight mass spectrometry
FDA	Food and Drug Administration
HRD	Histidine-Arginine-Aspartate
IC_{50}	Half-maximal inhibitory concentrations
IPTG	Isopropyl 1-thio-β-D-galactopyranoside
ITC	Isothermal titration calorimetry
K_a	Equilibrium association constant
k_{cat}	Catalytic rate of reaction
K_d	Equilibrium dissociation constant
K_m^{ATP}	ATP concentration that results in half maximal reaction velocity

K_m ^{peptide}	Peptide concentration that results in half maximal reaction velocity
k_{off}	Dissociation rate
k_{on}	Association rate
LB	Lysogeny broth
LIC	Ligation independent cloning
n	Stoichiometry
NMR	Nuclear magnetic resonance
nsSNP	Nonsynonymous single nucleotide polymorphism
PCR	Polymerase chain reaction
PDB	Protein data bank
PEG	Polyethylene glycol
PIM	Proviral integration site of moloney virus in murine lymphomas
RMSD	Root-mean-square deviation
Sf9	*Spodoptera frugiperda* cell line
SGC	Structural Genomics Consortium
SPR	Surface plasmon resonance
TCEP	Tris(2-carboxyethyl)phosphine
TEV	Tobacco etch virus
TLS	Translation/libration/screw
T_m	Protein melting temperature
WT	Wild-type

1. Introduction

1.1. Protein kinases and their role in diseases

The human kinome comprises 518 protein kinases, which control and modulate a wide range of biological processes (Manning et al., 2002). These enzymes all catalyse the same reaction in eukaryotic cells, the transfer of the γ-phosphate group of adenosine triphosphate (ATP) to the hydroxyl group of a serine, threonine or tyrosine residue on a substrate.

Hanks and Hunter performed sequence conservation and phylogeny analysis of the protein kinases and revealed many conserved motifs, which were classified into 12 subdomains (Hanks et al., 1988). Today the kinase family is structurally well defined and by August 2013, 198 unique human kinase domain structures had been determined and deposited in the protein data bank (PDB) (Berman et al., 2000). Even though they are diverse in their amino-acid sequence, all typical protein kinases share a conserved bilobal fold, formed by 250-300 amino acids (Figure 1.1). The N-terminal lobe is composed of β-sheets and the conserved regulatory 'helix αC', whereas the larger C-terminal lobe consists mainly of α-helical secondary structure; the two lobes are connected by a loop called the 'hinge'. The ATP binding pocket is located in a deep cavity formed by the N- and C-terminal lobes and is lined by structural elements responsible for the catalytic activity of the kinase.

One of these elements, the 'activation loop' located in the C-terminal lobe, is typically formed by 20-30 residues and functions as a docking site for the substrate. The N-terminus of the activation loop is marked by the highly conserved sequence motif DFG (Asp-Phe-Gly), where the aspartate is critical for protein catalysis by interacting with a magnesium ion that coordinates the γ-phosphate groups of ATP (Hubbard, 1997, Taylor and Kornev, 2011). Protein kinase activity is controlled by multiple mechanisms

and commonly involves large conformational changes of the activation loop when the kinase switches between its inactive and active forms.

Another conserved motif, HRD (His-Arg-Asp), is located in the catalytic loop that directly precedes the activation loop. The HRD motif's aspartate acts as a catalytic base that deprotonates the hydroxyl group of the substrate residue during the catalysis (Gibbs and Zoller, 1991, Jura et al., 2011).

An important structural element, the 'P-loop', is located in the N-terminal lobe between the β2 and β3 strands and contains a conserved glycine-rich sequence motif (GXGXϕG), where ϕ is usually phenylalanine or tyrosine. The flexible P-loop plays a crucial role in catalysis by forming a part of the ATP binding pocket and by making interactions with the β and γ-phosphates of ATP.

Another important structural element is helix αC in the N-terminal lobe of the kinase. This very dynamic helix contains a conserved glutamate residue ('αC-Glu'), which forms a salt bridge with a conserved lysine residue in strand β3 ('cat-K') in order to coordinate α- and β-phosphate groups of the ATP molecule. Indeed, the mutation of catalytic lysine is known to generate catalytically dead kinases (Bachmann and Möröy, 2005, Iyer et al., 2005).

Protein kinases are turned on and off by very precise biological mechanisms. During reversible inactive-to-active conformational transitions, the structural elements, such as the helix αC, the P-loop and the DFG motif play important regulatory roles (Möbitz and Fabbro, 2012, Huse and Kuriyan, 2002). Certain features might facilitate the flip of the DFG motif by providing space for the phenylalanine and by coordinating the aspartate side chain when it rotates out of the active site (Huse and Kuriyan, 2002, Levinson et al., 2006). Moreover, a network of conserved hydrophobic interactions between the P-loop, helix αC and the activation segment have been suggested to be central in kinase activation by linking the ATP and substrate binding regions. Spatially conserved hydrophobic motifs, termed 'spines', reveal the internal architecture of protein kinase

core (Kornev et al., 2006). First is the regulatory spine, which consists of four non-consecutive hydrophobic residues from β4 strand, the helix αC, the activation loop and the catalytic loop. The regulatory spine is assembled in active kinases and disassembled if a kinase is inactive. Another is the catalytic spine, which comprises eight conserved residues from both N- and C- terminal lobes and is completed by the ATP molecule. Both spines are connected to the helix αF, which is deeply buried in the C-lobe.

In most cases the kinase and substrate have complementary sequences and there interactions are based on charge, hydrogen bond formation or hydrophobic effect (Section 1.4.2).

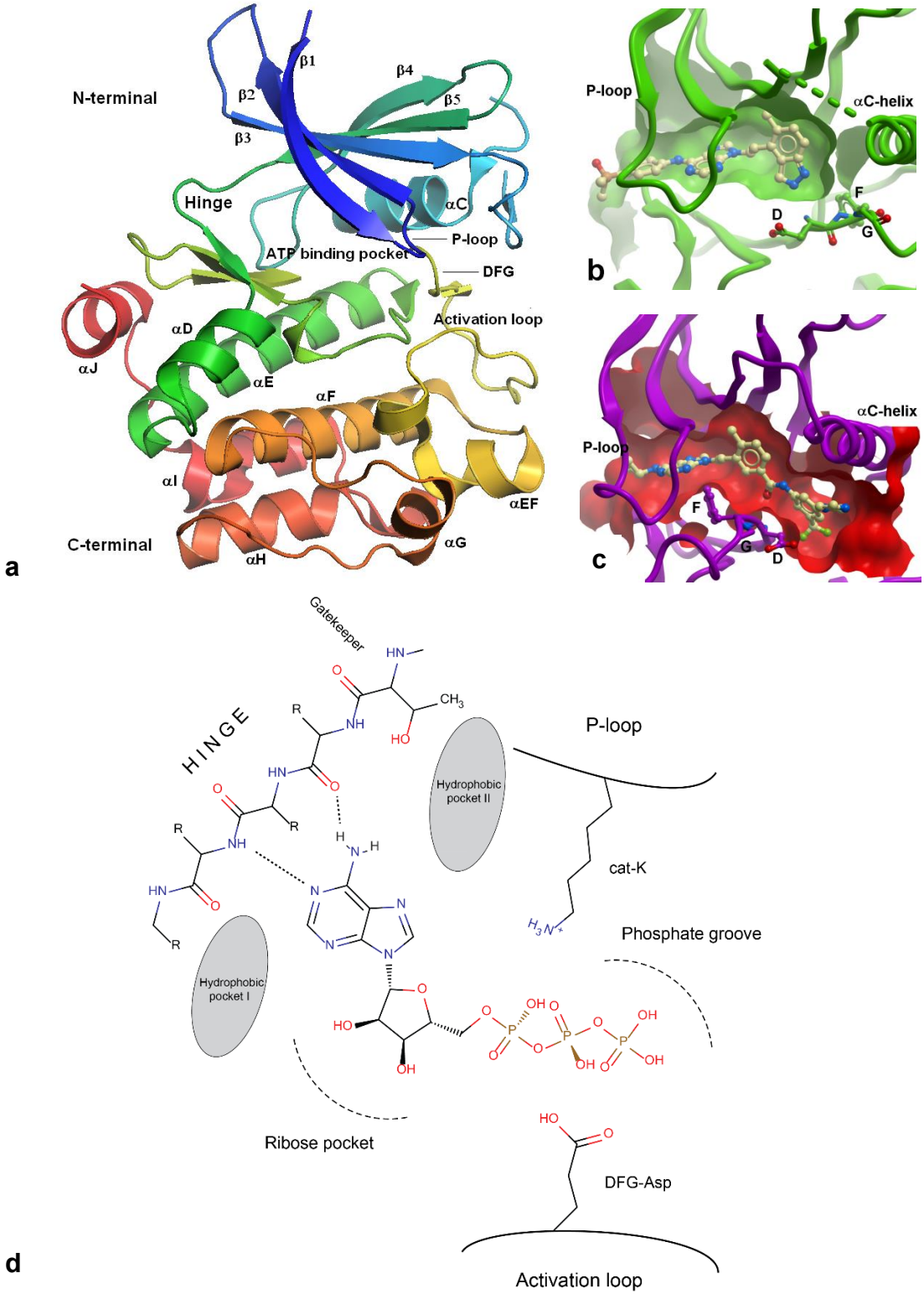


Figure 1.1 Protein kinases and their conformations. (a) Structure of a typical kinase, PDB: 2c3i. The chain is coloured beginning with blue at the N-terminus going through the rainbow to red at the C terminus with structural elements labelled. (b) Comparison of different conformational states of ABL kinase in complex with the DFG-out inhibitor, PDB: 3kfa and (c) with the DFG-in inhibitor, PDB: 3kf4. The effect of the inhibitor on the size and shape of the binding pocket is shown with a surface representation of the binding site. (d) Schematic representation of ATP-binding pocket with the key elements indicated.

1.2. Protein kinases as drug targets

Protein kinases are involved in many cell signalling networks and coordinate basic cellular functions such as gene expression, cell division and differentiation. Dysfunction of kinases, mainly overexpression and upregulation, has been implicated in many diseases including cancer, diabetes and inflammation (Noble et al., 2004, Zhang et al., 2009). Protein kinases therefore have been attractive targets for drug development, and small molecule inhibitors have been extensively pursued for potential therapeutic benefits.

The inhibitors that target the ATP-binding pocket are commonly classified into type 1 and type 2 (Liu and Gray, 2006). Type 1 inhibitors occupy the ATP site of the active form of a kinase and have at least one hydrogen bond donor or acceptor which enables binding to the hinge region (Figure 1.1b). Type 2 inhibitors also bind to the ATP site, but recognise the inactive DFG-out conformation and occupy the hydrophobic side pocket vacated by the phenylalanine of the DFG motif (Figure 1.1c and Figure 1.2b). Not all type 2 inhibitors form a hydrogen bond with the hinge; one of their distinct features is a central urea or amide linker, which serves as a facilitator of an essential hydrogen bond between the αC-Glu side-chain with the backbone amide of aspartate of the DFG motif. Hydrophobic moieties of type 2 inhibitors extend into the deep DFG-out pocket providing additional interactions with the target, whereas an orthogonally arranged aromatic scaffold enhances the affinity and ensures an extended conformation of the small molecule (Figure 1.2).

All active kinases are designed to bind ATP and hence have a similar active site architecture, while inactivation of a protein kinase often involves large movements of its subdomains, making this state structurally diverse. The targeting of distinct inactive conformations of kinases has emerged as a promising strategy for inhibitor design after the first kinase inhibitor, imatinib, was approved by the Food and Drug Administration (FDA) in 2001. The success of imatinib was phenomenal and this drug revolutionised

the treatment of chronic myeloid leukaemia (CML). According to the clinical studies reports, "complete hematologic responses were observed in 53 of 54 patients with CML treated with daily dosage of 300 mg or more and typically occurred in the first four weeks of therapy" (Druker et al., 2001), whereas the five-year follow up study estimated overall survival rate for patients to be 89%, with a relapse rate of only about 17% (Druker et al., 2006). Crystallographic studies have revealed that imatinib binds to a distinct inactive conformation of its target, the ABL kinase domain (Schindler et al., 2000, Nagar et al., 2002). In this DFG-out binding mode, rotation of phenylalanine of the DFG motif creates a large hydrophobic pocket that is efficiently targeted by type 2 inhibitors (Figure 1.1c). Strikingly, imatinib fails to potently inhibit the highly related Src kinase family, suggesting that a DFG-out conformation of Src is associated with high energetic barrier (Dar et al., 2008).

Unfortunately, induced mutations cause imatinib resistance, leading to relapse in some patients (Daub et al., 2004). Dasatinib is a potent kinase inhibitor approved in 2006 as a second-line treatment for those patients (Daub, 2010). Unlike imatinib, dasatinib binds to the active DFG-in conformation of the ABL kinase domain (Figure 1.1b) and inhibits not only the wild-type, but also most of the reported ABL mutants (Kéri et al., 2011). However, being a multi-targeted type 1 kinase inhibitor, dasatinib interacts with a larger number of kinases and is known to cause serious adverse effects.

Overall, targeting the inactive conformations alone does not guarantee more favourable selectivity profiles, and some type 2 inhibitors bind to multiple kinases while some type 1 inhibitors are relatively selective. Addressing unique active site features, like the additional hydrophobic pocket next to the gatekeeper residue (Zuccotto et al., 2010), and allosteric binding sites located outside the ATP pocket (Emery et al., 2009) have also been shown to be successful in designing selective inhibitors (Figure 1.1d).

In total, 17 kinase inhibitors and five monoclonal antibodies have already been approved for clinical use and many more are at various stages of clinical testing, mainly for oncology indications (Fabbro et al., 2012, Möbitz and Fabbro, 2012).

Unfortunately, cross-reactivity, adverse effects and developed resistance result in high attrition rates of kinase inhibitors in the clinic and a low rate of new therapeutic discovery.

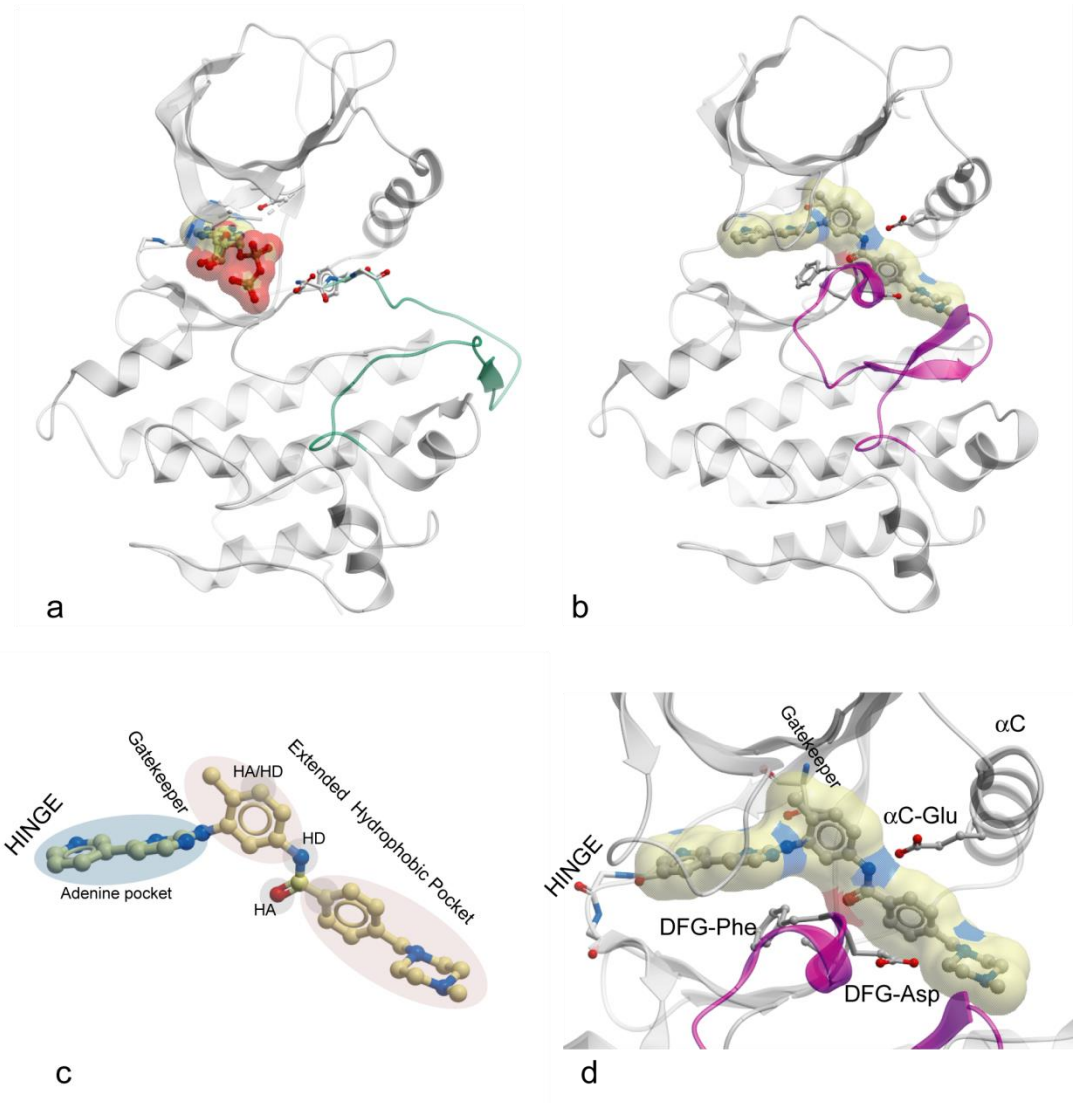


Figure 1.2 Abl protein kinase in the active and inactive state. (a) Crystal structure of Abl kinase in complex with ADP, PDB: 1c2i. The activation loop is coloured in green and side chains of the gatekeeper and the DFG-motif are shown as sticks. (b) Crystal structure of Abl kinase in complex with Imatinib, PDB: 2hyy, the activation loop is coloured in purple. (c) Pharmacophore model for type 2 inhibitor (Imatinib) with the approximate positions of the hinge region and the gatekeeper residues indicated. (d) Positioning of Imatinib within the ATP-binding site of Abl kinase, important resides are indicated.

1.3. The equilibrium nature of protein conformations

In general, proteins are dynamic molecules and can be described in terms of their dynamically fluctuating conformational ensembles and ensemble distributions (Kar et al., 2010). The concept of equilibrium shift of the ensemble was proposed nearly 15 years ago and its basic idea is that proteins in solution pre-exist in a number of conformational substates (Ma et al., 1999). The equilibrium of the substates is not static and can be shifted by mutations of amino-acid sequence, post-translational modifications, ligand binding or other environmental factors.

Reversible switching between active and inactive conformations of a protein kinase is a central feature in regulation of signalling networks, hence active, inactive and numerous intermediate kinase conformations exist in a dynamic equilibrium. Indeed, large conformational changes have already been described by population shift, where in addition to active and inactive states many intermediate states have also been sampled using experimental and computational techniques (Grant et al., 2009, Shan et al., 2009, Endicott et al., 2012). To date, however, a limited number of protein kinase conformations has been identified and while the key structural features of kinases such as the DFG motif, αC-helix and P-loop have been well characterised (Cowan-Jacob et al., 2009, Kornev and Taylor, 2010), the mechanisms leading to the displacement of these elements resulting in kinase inactivation remain poorly understood. In theory, every kinase can adopt any conformation at an energetic cost, which is dictated by its primary sequence, post-translational modifications and binding to protein partners. Quantitative evaluation of energetic penalties and the mechanisms of dynamic conformational changes that occur during protein binding remain the subject of intense interest (Csermely et al., 2010). At present, our vast amount of information results from structural determinations by means of X-ray crystallography. However, conventional crystallographic studies only reveal static snapshots of protein conformations, a space average over all molecules in the crystal and a time average over the period required

for data collection. In order to learn about functionally important transient intermediate states and to encompass the dynamic nature of conformational fluctuations due to binding events with interaction partners, crystallographic analysis should be complemented with other approaches. Comprehensive biochemical and biophysical analysis, nuclear magnetic resonance (NMR) technologies and molecular dynamics simulations are amongst methods that can effectively reflect the dynamics of the conformational distributions. Alternatively, targeted amino-acid substitutions have been shown to shift the conformational equilibrium, which can be characterised by evaluation of the binding specificity of activating and inactivating mutants (Csermely et al., 2010, Möbitz and Fabbro, 2012).

Currently, only a small fraction of the kinome has been selectively targeted by small molecule inhibitors (Fedorov et al., 2010) and there is a big demand for the development of better drug design strategies addressing not only clinically validated targets, but also novel kinases. Therefore, gaining detailed knowledge of dynamics within the kinase catalytic domain is important for understanding kinase-ligand interactions and for designing selective inhibitors.

1.4. Function and regulation of CDK2

1.4.1. Cyclin dependent kinases govern the cell cycle

Most events in the eukaryotic cycle are controlled by the transfer of the γ-phosphate group of ATP to the hydroxyl of serines and threonines in regulatory proteins. Cyclin-dependent protein kinases (CDKs) can be described as machines that run the cell cycle programme since they are the key enzymes that catalyse this reaction (Morgan, 2006).

Different CDKs are activated at different stages of the cell cycle (Figure 1.3a) and their activities are highly controlled through a number of mechanisms. Most fundamentally, activation of a respective CDK is mediated through association with a regulatory cyclin subunit (Figure 1.3c), the levels of which are tightly controlled by transcription and ubiquitin-mediated degradation in a time-based manner (Morgan, 2006). Other regulatory mechanisms include positive and negative phosphorylation of CDKs, as well as interactions with inhibitors and accessory proteins (Harper and Elledge, 1998, Moore, 2013, Malumbres and Barbacid, 2005).

In human cells, 21 genes encoding CDKs have been identified and classified on the basis of similarities in sequence and function (Malumbres et al., 2009). At least five constitute the core cell cycle engine: CDK1, CDK2, CDK3, CDK4 and CDK6 (Figure 1.3a). Transition of cell cycle from the G0 to G1 phase is assigned to the CDK3-Cyclin C complex, while cyclin D-dependent kinases CDK4 and CDK6 govern progression through G1.

The activity of retinoblastoma tumour suppressor gene product (Rb), which controls the G1/S transition, is modulated by the sequential phosphorylation by CDK4/6-Cyclin D and CDK2-cyclin E (Malumbres and Barbacid, 2001). Once Rb is phosphorylated by CDK4/6, it releases the E2F-DP transcription factors required for transcription of cyclin E. Consequent partnering of cyclin E with CDK2 results in additional phosphorylation and full inactivation of Rb, while the E2F is responsible for production of numerous S-phase proteins (Schwartz and Shah, 2005). During the G1/S transition the CDK2-cyclin E also phosphorylates other substrates involved in chromosomal DNA replication.

While cyclins D and E are subjected to ubiquitin-mediated proteasome degradation early in S-phase (Elledge and Harper, 1998), cyclin A is being synthesised. The CDK2-cyclin A complex then governs S-phase progression and DNA synthesis, enabling production histones, proliferating cell nuclear antigen and other crucial proteins. Importantly, E2F-DP phosphorylation by CDK2-cyclin A inhibits its DNA-binding ability,

and thus E2F-DP-dependent transactivation abolishes (Xu et al., 1994). Active CDK1-cyclin A complex is necessary for DNA damage checkpoint control, followed by association of CDK1 with cyclin B to drive the cell into mitosis.

1.4.2. Activation of CDK2

Like other CDKs, CDK2 activity is tightly regulated by a complex mechanism. Monomeric CDK2 is inactive and its activation *in vitro* requires at least two steps: binding to a regulatory cyclin and activation loop phosphorylation (Figure 1.3). Cyclin A, cyclin B and cyclin E are amongst the identified activators of CDK2 (Brown et al., 2007). Binding to cyclin induces significant conformational modifications within the CDK2 subunit.

The most dramatic change is the reorientation of the αC-helix that induces interaction between the catalytic αC-Glu (E51) and the ATP-coordinating lysine (K33) for catalysis (Figure 1.3d). In contrast, the activation loop of the inactive unphosphorylated CDK2 contains a short helical segment (αL12), which displaces the αC helix and exposes the αC-Glu to solvent.

Cyclin binding also results in changes of the activation loop positioning, ensuring appropriate exposure of the threonine phosphorylation site (T160) to the solvent. Consequent phosphorylation of the CDK2-cyclin complex at T160 by Cdk7/cyclin H (CAK) rigidifies the activation loop forming a platform for substrate recognition, further stabilising the complex and enhancing the kinase activity (Figure 1.3e) (Brown et al., 2007). Dual phosphorylation of the P-loop (T14 and Y15) by the inhibitory kinases Wee1 and Myt1, negatively regulates activity of CDK2 (Gu et al., 1992, Welburn et al., 2007). Unlike activation loop phosphorylation, P-loop phosphorylation does not induce major conformational changes, but reduces CDK2 affinity for the peptide substrate, thereby inhibiting the kinase activity. In addition, phosphorylation of Thr14 could

probably prevent correct positioning of ATP in the binding pocket (Echalier et al., 2010).

The Cip/Kip family of cyclin-dependent kinases inhibitors (CKIs) bind to both the cyclin and CDK2 subunits, disrupting the CDK2 fold and interacting with the ATP binding pocket (Kaldis, 1999). Strikingly, association with the Cip/Kip family also promotes the association of cyclin D with CDK4 and CDK6, which in turn activates CDK2 (Johnson et al., 2002).

This activation model is not universal for all CDKs. For example, CDK7 partners with cyclin H and the complex is stabilised by phosphorylation of the two residues within the activation segment of the kinase, or alternatively by association with a third protein MAT1 (Kaldis, 1999).

CDK2 and CDK7 have different substrate specificities. The consensus peptide substrate in CDK2, S/T-P-X-K/R (where X is any amino acid), is characterised by a proline immediately after the phosphorylatable residue (p+1). This proline fits into the hydrophobic cavity near the catalytic site, available only when T160 is phosphorylated. CDK2 has a strong preference for a charged residue at the next position (p+2), which is exposed to the solvent, and a basic residue in the p+3 position. Structural studies revealed that the latter forms a hydrogen bond with T160 and to the main chain of the cyclin subunit, I270 in cyclin A (Brown et al., 1999).

CDK7, however, lacks this preference for a proline and can tolerate other residues at the p+3 position. Like other transcriptional CDKs, CDK7 is able to phosphorylate different substrates that a dissimilar in sequence, demonstrating more relaxed specificity compared to CDK2 (Larochelle et al., 2001, Echalier et al., 2010). This recognition of a substrate is thought to result not from the sequence features surrounding the phosphorylatable residue, but rather from the remote protein-protein interaction sites. Supposedly, the presence of MAT1 could allow for specificity for proline-directed phosphorylation sites (Echalier et al., 2010).For many years CDK2 has been regarded as an attractive drug target because deregulation of CDK2 is frequently

associated with cancer. Genetic studies revealed that CDK2$^{-/-}$ mice were viable but sterile, and indicated that CDK2 was not essential for G1/S transition in cells, including murine embryogenesis and post-natal development (Ortega et al., 2003, Berthet et al., 2003). Moreover, experiments with cultured cancer cells illustrated that inhibition of CDK2 prevents cell proliferation (Tetsu and McCormick, 2003).

For these reasons, extensive effort from both pharmaceutical industry and academic groups lead to the development of hundreds, if not thousands, of small-molecule CDK2 inhibitors, several of which are currently in clinical trials as anti-tumour agents (Johnson, 2009). Numerous CDK2 crystal structures revealed different conformations of the ATP-binding site, however, the inactive DFG-out state has not yet been demonstrated. This made CDK2 an ideal model to elucidate the effect of sequence modifications on energetic and structural determinants that govern the conformational plasticity of protein kinases.

Many studies have been done to exploit conserved features around the ATP-binding site and their role in inhibitor binding. Over the last decade, Shokat and co-workers have made significant contributions to the field, and their work on chemical genetic strategies for inhibiting protein kinases is of particular relevance to this project (Garske et al., 2011a). Exploration of the hydrophobic residue preceding the hinge region called the gatekeeper demonstrated that inhibitor specificity can be achieved via shape complementarity. A protein kinase with a small gatekeeper (glycine or alanine) could be selectively inhibited by a bulky inhibitor, while a protein kinase with a cysteine gatekeeper could be selectively targeted by an electrophilic compound. Inhibitor sensitivity could also be affected by secondary mutations near the engineered cysteine residue within the ATP-binding pocket.

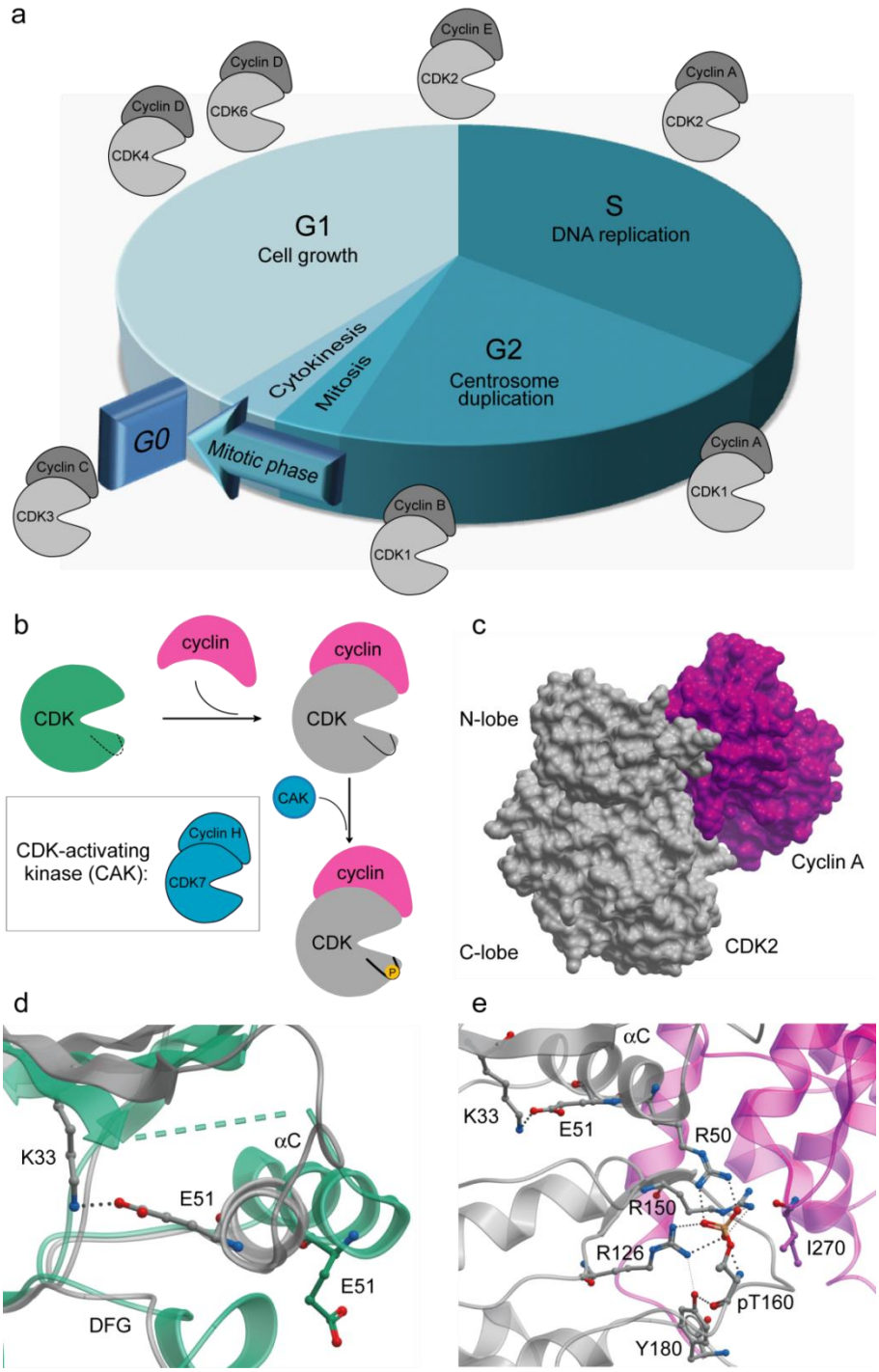


Figure 1.3 Function and regulation of CDK2. (a) Overview of core cell cycle regulation: different CDK/cyclin heterodimer complexes are highly specialised for promoting particular cell cycle transitions (Harper and Elledge, 1998, Moore, 2013, Malumbres and Barbacid, 2005). (b) Schematic diagram demonstrating two CDK activation steps: binding to a regulatory cyclin and activation loop phosphorylation by a variety of kinases collectively called CAK (CDK7-Cyclin H for CDK2). (c) Surface representation of CDK2-cyclin A complex, where the cyclin docks against the αC-helix of the CDK2 causing significant conformational changes. (d) An overlay of apo-CDK2 (green) and CDK2 from the complex (grey) highlighting change in orientation of the αC-helix and contact formation between catalytic residues E51 and K33. (e) Activation segment stabilisation motif in CDK2 (grey) – cyclinA (pink) complex, the residues involved in the activation segment stabilisation and the hydrogen bond network are shown. The figures (c)-(e) were created using ICM (Molsoft, LLC) with the following structures: apo-CDK2 (PDB: 1b38), CDK2-cyclin A complex (PDB: 1qmz).

1.4.3. Project description

The aim of this project was the identification of residues that have an effect on the energetics of the DFG movements and thereby the ability of CDK2 to bind type 2 inhibitors. CDK2 was chosen as a model to study DFG-in/out conformational changes because this protein kinase family member is easily expressed and has the biggest number of PDB depositions. Most importantly, despite this large amount of structural data, no DFG-out conformation has yet been observed. Therefore this kinase served as a prototype for a kinase that cannot easily assume a DFG-out conformation due to the large conformational penalty. By substituting amino-acids that may influence DFG movement, our goal was to convert CDK2 into a kinase that can assume DFG-out and can be targeted by type 2 inhibitors. Detailed structural and sequence comparisons, site-directed mutagenesis, biochemical, biophysical testing and crystallography were implemented to elucidate the structural mechanisms that modulate transitions of kinase regulatory elements from active to inactive states. Essentially, this information could be useful to probe sequence and structural features that define the DFG-in/out movement of the activation loop and assist in the design of more specific CDK2 inhibitors.

1.5. PIM kinase family

1.5.1. Biological role of PIM kinases

PIM kinase was identified nearly 30 years ago as a Proviral Integration site of Moloney virus in murine lymphomas (Cuypers et al., 1984). Integration of the viral DNA in the 5' encoding region of the PIM locus leads to stabilisation of the messenger RNA, overexpression of PIM and formation of tumours (Bachmann and Möröy, 2005, Bullock et al., 2005a). In humans, the PIM family comprises three serine/theronine protein kinases, PIM1, PIM2 and PIM3. Despite being classified as calmodulin-dependent protein kinases (CAMK) (Manning et al., 2002), PIM proteins are constitutively active and do not require association with a regulatory partner or phosphorylation for activity. Members of this family are highly similar in primary structure and despite the fact that all three human PIM isozymes are located on different chromosomes and have distinct expression patterns, they have been suggested to be at least partially functionally redundant (Mikkers et al., 2004, Brault et al., 2010). PIM proteins are key players in cell cycle progression and apoptosis (Bachmann and Möröy, 2005, Bullock et al., 2005a). Expression of PIM proteins is regulated by cytokines (Wang et al., 2001) and high levels of these kinases confer a growth advantage through a variety of mechanisms, including inactivation of the proapoptotic protein BAD (Aho et al., 2004) and the cell cycle inhibitor $p21^{cip1/waf1}$ (Wang et al., 2002) via phosphorylation . Overall, PIM1 has been shown to prevent the normal process of apoptosis and to promote cell survival, thereby contributing to the development of malignancies.

Importantly, after substrate stimulation PIM1 protein levels are transient with a short half-life in primary cells being rapidly degraded by heat shock proteins and the ubiquitin–proteasome pathway (Shay et al., 2005), and this could be the main regulation mechanism of PIM activity.

All three PIM kinases are aberrantly expressed in a wide variety of tumours (Nawijn et al., 2011) although considerable attention has been paid to PIM1 in particular. After

studies with PIM1-deficient mice, which did not show any anatomic or developmental abnormalities (Laird et al., 1993), PIM1 kinase emerged as a target for cancer therapy. In addition, analysis of PIM1 gene in cancer tissues led to identification of nonsynonymous single nucleotide polymorphisms (nsSNPs), which result in polypeptide sequence with amino-acid substitutions (Forbes et al., 2011, Greenman et al., 2007, Yuan et al., 2006). The effects of these nsSNPs on stability, function and interactions of PIM1 with other proteins have been studied extensively (Lori et al., 2013, Zhang et al., 2012).

1.5.2. Structures of PIM kinases

In 2005 the crystal structure of human PIM1 was determined (Bullock et al., 2005a, Qian et al., 2005, Kumar et al., 2005) (Figure 1.1a), followed by human PIM2 in 2009 (Bullock et al., 2009), while the crystal structure of PIM3 kinase has not yet been deposited in the PDB. The available PIM1 and PIM2 structures revealed a number of unique structural features of PIM kinases and offered new opportunities for the development of selective inhibitors.

Both isozymes have the conserved protein kinase fold that forms a bilobal structure (Figure 1.1a). While all active kinases bind the cofactor ATP in their binding pocket and have similar active site architecture, PIM proteins have distinct features. Firstly, due to the presence of a proline residue in the hinge region, PIM kinases are not able to form a second hydrogen bond between the hinge backbone and the adenine ring of ATP (Pogacic et al., 2007). Furthermore, a structure-based alignment of PIM proteins with other characterised kinases reveals at least one amino-acid insertion following that conserved proline residue (Figure 1.4), which provides a larger hydrophobic pocket for binding inhibitors. PIM kinases are also characterised by a unique β-hairpin insert in the N-terminal lobe linking β3 to helix αC, which is partially disordered in the PIM2 structure (Bullock et al., 2009). As the ATP-binding site represents the target site for the inhibitor, these unique characteristics make the PIM family a promising target for structure-based drug design, since they offer the potential for enhanced selectivity.

Indeed, there has been a massive interest in PIM inhibitors both in academic and industrial groups. The most advanced drug candidate SGI-1776 has been subjected to clinical trials sponsored by Astex Pharmaceuticals (Chen et al., 2009, Mumenthaler et al., 2009, Chen et al., 2011). Unfortunately in 2010 the trial was discontinued due to dose-limiting toxicity (http://clinicaltrials.gov/show/NCT00848601). Another clinical candidate, AZD1208 by AstraZeneca, is currently undergoing a phase II clinical trial. (http://clinicaltrials.gov/show/NCT01489722). Very potent pan-PIM inhibitors with a

similar chemical scaffold have been reported recently (Wang et al., 2013, Burger et al., 2013) and have also been developed by collaborators at the University of Oxford Chemistry department.

From a comparison of the ATP binding pockets of PIM1 and PIM2 it is evident that the few non-conserved residues point their side chains away from the binding site, creating pockets with the same structural topology. Despite this high similarity, the substantial majority of reported inhibitors show considerable selectivity for PIM1 and PIM3 over PIM2 (Brault et al., 2010, Pogacic et al., 2007). The mechanisms leading to the observed inhibitor selectivity have not yet been studied and the reasons for differences in inhibitor binding remain unclear. Most likely, dynamic features such as P-loop and/or hinge flexibility, which result in different conformations of structural elements, could play a role. Indeed, some investigators have suggested that flexibility of this glycine-rich phosphate binding loop is an important factor in ligand binding affinity and specificity (Doudou et al., 2009, Gohda and Hakoshima, 2008).

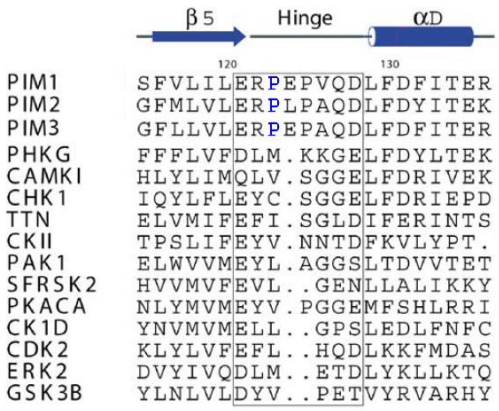


Figure 1.4. Sequence alignment of the hinge region of selected human kinases. Secondary structure elements are indicated and PIM-characteristic proline residue is coloured in blue (adapted from (Kumar et al., 2005)).

1.5.3. P-loop flexibility

P-loop flexibility has been observed and documented in many structures and characterised by molecular dynamics studies (Doudou et al., 2010). Six P-loop residues (46-51, PIM1 numbering) link β1 and β2 strands and form a part of the ATP binding pocket interacting with the adenine ring. From structural observations it is evident that the P-loop changes its conformation when the ligand binds, and its orientation could be an important part of the inhibitor binding mechanism (Figure 1.5a).

As suggested by a number of AMP-PNP inhibited structures (Bullock et al., 2005a, Qian et al., 2005, Kumar et al., 2005), the P-loop is displaced by ATP and takes on the so-called extended P-loop conformation. In contrast, apo-enzymes and co-structures with some small ligands adopt a folded P-loop conformation (Figure 1.5b).

Since PIM kinases are constitutively active, they do not adopt an inactive DFG-out or αC-out conformation to prevent ATP binding, although a range of P-loop fluctuations can be observed (Figure 1.5a). Depending on the compound in the co-structure, distinct active and inactive P-loop conformations are recognised, where phenylalanine (F49 in PIM1) at the tip of the P-loop points ‘in’ or ‘out’ of the binding pocket (Figure 1.5b). The ‘in’ or ‘folded’ position of F49 results in $\pi-\pi$ interaction with aromatic rings of the inhibitor, which is incompatible with ATP binding. Phenylalanine stacking similar to this has also been observed in a few other kinases (Guimarães et al., 2011) and the propensity of a kinase to adopt a ‘folded’ P-loop conformation could probably be explained by specific combinations of the residues within the P-loop area. Improved potency and selectivity of compounds could be the potential advantages of targeting this inactive conformation (Guimarães et al., 2011, Pogacic et al., 2007).

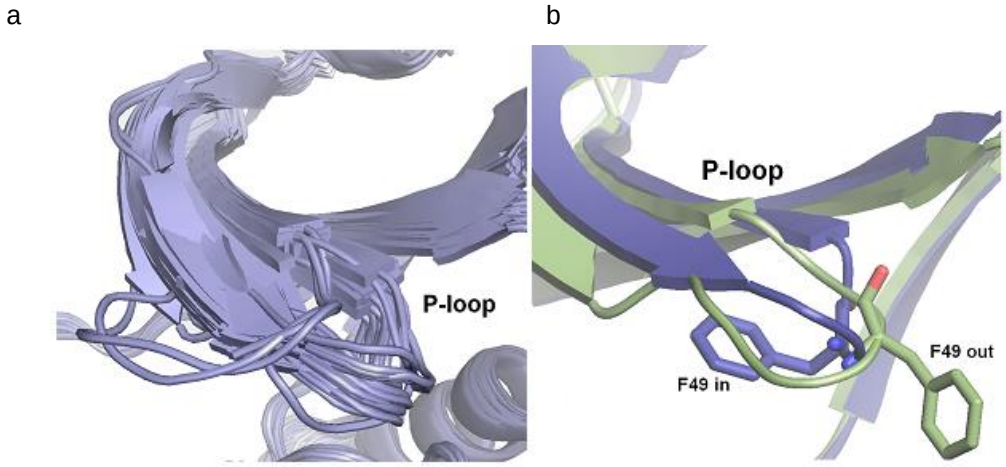


Figure 1.5. P-loop flexibility in PIM kinases. (a) An overlay of all available PIM1 structures and co-structures, showing P-loop fluctuations, (b) P-loop in active (green, PDB: 1yxt) and inactive (blue, PDB: 2c3i) conformations.

1.5.4. Project description

The aim of this project is the identification of residues that confer the specificity of protein-ligand interaction resulting in different binding profiles of PIM1 and PIM2 proteins when targeted by small molecule inhibitors. To address this, we adopted a bioinformatic approach taking into account the phylogenetic history of PIM kinases and the structures of PIM1 conformations upon binding to various ligands. Additionally, a number of point mutations, located at the hinge and the P-loop, were introduced into the PIM1 sequence. PIM1 was selected for the following reasons: it is known to be a stable protein and has been thoroughly characterised both structurally and biochemically compared to PIM2. The PIM2 residues were therefore introduced into the PIM1 sequence to engineer mutants that have inhibitor binding profiles similar to that of PIM2. Experiments involving screening of representative chemical libraries and X-ray crystallography provided insights into the structural flexibility of the P-loop and the hinge. The potential reasons for the inability to obtain consequent PIM2 and PIM3 structures were also investigated by predicting and comparing disorder of PIM family proteins.

2. Methods

2.1. Site-directed mutagenesis

Site-directed mutagenesis was performed with the assistance of Dr Pavel Savitsky. The cDNA sequences available at the Structural Genomics Consortium (SGC) were used as templates for site-directed mutagenesis. All primers used (Table 2.1) were purchased from Sigma–Aldrich (Oakville, ON). Mutant sequences were generated by site-directed mutagenesis of a wild-type cDNA using the primer extension method (Reikofski and Tao, 1992, Heckman and Pease, 2007). The method involves incorporating mutagenic primers in independent PCRs before combining them into the final product. The reaction required two flanking master primers (A and D) that mark the 5′ ends of both strands, and two internal primers (B and C) that contain the mutation of interest and also create overlapping nucleotide sequences (Figure 2.1).

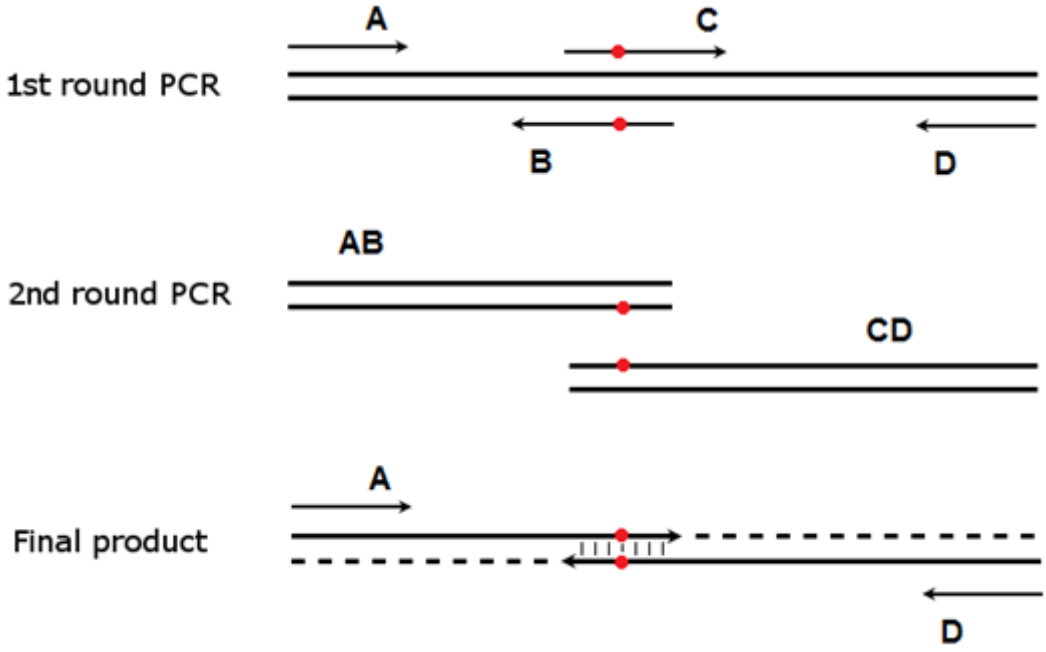


Figure 2.1 Site-directed mutagenesis using the primer extension method. During the 1st PCR round, two flanking master primers, A and D, and two internal primers, B and C containing the mutation of interest (indicated by the red dot) are used to generate intermediate PCR products AB and CD. In the 2nd PCR round, strands of each product anneal at their overlapping complementary regions containing the mutation, and amplification of the product AD is driven by primers A and D.

Table 2.1 Construct information.

Construct ID	Modification	Vector Name	Primers	Expected Mass	Estimated Extinction Coefficient
PIM1A-c001	wt	pLIC-SGC1	Fwd: TACTTCCAATCCATGCTCTTGTCCAAAATCAACTCGCTTGC Rev: TATCCACCTTTACTGTCAGCTGGGCCCCGGCGACAGGCTGTGG	38011.2	50420
PIM1Z-c001	S46K/S51T	pNIC28-Bsa4	Obtained from *	38066.4	50420
PIM1Z-c002	E124L	pNIC28-Bsa4	Obtained from *	37995.3	50420
PIM1Z-c003	S46K	pNIC28-Bsa4	Obtained from *	38052.4	50420
PIM1Z-c004	V126A	pNIC28-Bsa4	Obtained from *	37983.2	50420
PIM1Z-c005	S51T	pNIC28-Bsa4	Obtained from *	38025.3	50420
PIM1Z-c009	H68V	pNIC28-Bsa4	Fwd: GTGGCCATCAAAGTCGTGGAGAAGGACCGGATTTCCGA Rev: TCCTTCTCCACGACTTTGATGGCCACCGGCAAGTTGTC	37973.3	50420
PIM1Z-c010	I56H/P63Q	pNIC28-Bsa4	Fwd: CACCGCGTCTCCGACAACTTGCAGGTGGCCATCAAACACGTGGAGAA Rev: CTGCAAGTTGTCGGAGACGCGGTGGCCTGAGTAGACCGAGCCGAAGC	38066.3	50420
PIM2A-c007	wt	pNIC28-Bsa4	Fwd: TACTTCCAATCCATGTTGACCAAGCCTCTACAGG Rev: TATCCACCTTTACTGTCAGGGTAGCAAGGACCAGGC	36743.2	48930
PIM2A-c008	wt truncated	pNIC28-Bsa4	Fwd: TACTTCCAATCCATGGGCAAGGATCGGGAAGCGTTCGAGGCC Rev: TATCCACCTTTACTGTTAGGGTAGCAAGGACCAGGCCAAAGG	34841.0	48930
PIM3A-c030	wt	pNIC28-Bsa4	Made at the SGC	39295.8	44920
CDK2A-c002	wt	pNIC28-Bsa4	Fwd: TACTTCCAATCCATGGAGAACTTCCAAAAGGTGGAAAAG Rev: TATCCACCTTTACTGTCAGAGTCGAAGATGGGGTACTG	36482.3	38390
CDK2A-c004	Biotinylation	pNIC-Bio3	Fwd: TACTTCCAATCCATGGAGAACTTCCAAAAGGTGGAAAAG Rev: TATCCACCTTTACTGCTGAGTCGAAGATGGGGTACTGG	38872.9	45380
CDK2Z-c001	F80T	pNIC28-Bsa4	Fwd: TCTACCTGGTTACTGAATTTCTGCACCAAGATCTCA Rev: TGCAGAAATTCAGTAACCAGGTAGAGTTTATTTTCT	36436.3	38390
CDK2Z-c002	F80M	pNIC28-Bsa4	Fwd: TCTACCTGGTTATGGAATTTCTGCACCAAGATCTCA Rev: TGCAGAAATTCCATAACCAGGTAGAGTTTATTTTCT	36466.3	38390
CDK2Z-c003	C118L	pNIC28-Bsa4	Fwd: GCCTAGCTTTCTTACATTCTCATCGGGTCCTCCAC Rev: CCGATGAGAATGTAAGAAAGCTAGGCCCTGGAGCA	36492.3	38390
CDK2Z-c004	C118I	pNIC28-Bsa4	Fwd: GCCTAGCTTTCATCCATTCTCATCGGGTCCTCCAC Rev: CCGATGAGAATGGATGAAAGCTAGGCCCTGGAGCA	36492.3	38390
CDK2Z-c005	A144C	pNIC28-Bsa4	Fwd: CCATCAAGCTATGTGACTTTGGACTAGCCAGAGCTTTT Rev: AGTCCAAAGTCACATAGCTTGATGGCCCCCTCTGTGTT	36514.4	38390
CDK2Z-c006	A144G	pNIC28-Bsa4	Fwd: ATCAAGCTAGGAGACTTTGGACTAGCCAGAGCT Rev: TCCAAAGTCTCCTAGCTTGATGGCCCCCTCTGT	36468.3	38390
CDK2Z-c007	C118L/A144C	pNIC28-Bsa4	Sequential use of CDK2Z-c003 and CDK2Z-c005 primers	36524.4	38390
CDK2Z-c008	C118L/A144G	pNIC28-Bsa4	Sequential use of CDK2Z-c003 and CDK2Z-c005 primers	36478.3	38390
CDK2Z-c009	Biotinylation	pNIC-Bio3	Fwd: TACTTCCAATCCATGGAGAACTTCCAAAAGGTGGAAAAG Rev: TATCCACCTTTACTGCTGAGTCGAAGATGGGGTACTGG	38915.0	45380
CCNA2A-c003	wt	pNIC28-Bsa4	Obtained from **	43402.9	43320
CCNB1A-c005	wt	pCOEX-1	Made at the SGC	34452.2	31860

* Department of Biochemical Sciences “A. Rossi Fanelli”, Sapienza University of Rome, Rome, Italy.
** Novo Nordisk Foundation Center for Protein Research, Copenhagen, Denmark.

In the first round of PCR, two separate reactions were performed, firstly using primer pairs A and B, and secondly using primers C and D. The PCR products were treated overnight with Dpn I restriction enzyme (Stratagene) to digest the methylated, non-mutated parental DNA template. The resulting fragments AB and CD were mixed for the second round of PCR with primers A and D to create the final product, which contains the desired mutated sequence.

A standard PCR reaction was performed using the Platinum Pfx proofreading enzyme (Invitrogen, Carlsbad, CA, USA) in a volume of 18 µl containing 1-2 ng template, 1.6 µM of each primer was mixed with 5 µl 10 x Pfx amplification buffer, 0.5 µl 50 mM $MgSO_4$, 0.75 µl 10mM dNTP mix, 13.5 µl H_2O and 2.5 µl Platinum Pfx. Cycles generally consisted of initial denaturation (95°C, 5 min), then 30-35 cycles of denaturation (98°C, 30 seconds), primer annealing and extension (1 min per DNA kilobase-pair) at 68 °C. Cycling was performed in thin-walled plastic PCR tubes (ABgene) and a PTC-225 thermocycler (Bio-Rad).

2.2. Ligation independent cloning

The prepared PCR products were purified (QIAquick PCR Purification Kit, QIAGEN Ltd. UK) and subcloned into pET28 derived expression vectors, pNIC28-Bsa4 (gi|124015065) or pNIC-Bio3 (gi|355331699) (Keates et al., 2012), using ligation independent cloning (LIC) (Aslanidis and de Jong, 1990). The LIC method is based on annealing short complementary single-stranded 5′-tails generated by the 3′-exonuclease activity of T4 DNA polymerase, where 12-16 base overhangs in cDNA were ensured by building appropriate extensions into the primers according to the expression vector (Table 2.1) (Savitsky et al., 2010). For the T4 DNA polymerase treatment, 5 µl PCR product was mixed with 1.5 µl 10 x T4 Polymerase buffer (Novagen), 1 µl 25 mM dCTP, 0.5 µl 100 mM dithiothreitol (DTT), 0.5 µl BSA (NEB), 0.5 µl T4 Polymerase (Novagen) and 1.5 µl H_2O and incubated for 30 min at 22°C, followed by 30 min at 80°C in a PTC-225 thermocycler (Bio-Rad). The vector was treated in the same manner, but with dGTP instead of dCTP. Annealing was performed by mixing 1 µl treated plasmid with 3 µl treated insert, incubated 5-10 min at 22°C and placed on ice. The cloning vectors for *E. coli* included fusion tags for affinity purification, typically N-terminal His_6 tags that can be cleaved with Tobacco etch virus

(TEV) protease (Appendix A). Competent Mach1 cells (Invitrogen, UK) were transformed with the constructs to yield the final plasmid DNA.

2.3. Protein production

2.3.1. Transformation into *E.coli* protein expression strain

E.coli strains of BL21(DE3)-R3-pRARE cells (phage-resistant derivative with a pRARE plasmid encoding rare codon tRNAs) (Savitsky et al., 2010) or Rosetta-R3-BirA (encoding the BirA gene for biotin-protein ligase) (Keates et al., 2012) were used as expression hosts. For PIM proteins, an additional co-expression plasmid encoding lambda protein phosphatase was incorporated (Gileadi et al., 2008). Competent cells were transformed by heat shock for 45 seconds in a water bath at 42°C, then immediately placed on ice for 2 min, followed by addition of 200 µl LB medium and incubation for 1 hour at 37°C. 50-100 µl transformed cells were plated onto LB-agar media, supplemented with appropriate antibiotics.

2.3.2. Protein expression in *E.coli*

The expression culture was grown in lysogeny broth (LB) medium at 37°C until OD_{600} of 0.8 was reached. At this point the temperature was reduced to 18°C and expression was induced by the addition of 0.5 mM isopropyl 1-thio-β-D-galactopyranoside (IPTG). Expression was continued for 16 to 18 h. The cells were harvested by centrifugation at 4000*g* for 20 min at 4°C and resuspended in lysis buffer comprising 50 mM HEPES pH 7.5, 300 mM NaCl, 5 mM imidazole, 5% glycerol, and 0.5 mM tris(2-carboxyethyl)phosphine (TCEP).

2.3.3. Expression in baculovirus-infected insect cells

The baculovirus used to produce CDK2 for successful structural studies (Rosenblatt et al., 1993, De Bondt et al., 1993) was provided by Novartis (Emeryville). The baculovirus was amplified in *Spodoptera frugiperda* (Sf9) insect cells in Sf900 II media (Invitrogen) and used to infect Sf9 cells grown in suspension to a density of 2×10^6 cells/ml. At 48 h postinfection the cells were harvested by centrifugation at 800*g* for 20 min at 4°C and cell pellets were stored at −20 °C. Cells were re-suspended in a buffer consisting of 10 mM TRIS pH 7.4 and 25 mM NaCl.

2.3.4. Protein purification

The cell suspension was lysed using high-pressure cell homogenization (Avestin C3 Emulsiflex) and cleared by centrifugation at 56000*g* for 40 min at 4 °C. The cleared cell lysate was loaded onto a column containing Ni Sepharose High Performance media (GE Healthcare) equilibrated with lysis buffer. The column was then washed with buffer typically consisting of 50 mM HEPES pH 7.5, 300 mM NaCl, 40 mM imidazole, 5% glycerol and 0.5 mM TCEP. Proteins were eluted with 50 mM HEPES pH 7.5, 300 mM NaCl, 250 mM imidazole, 5% glycerol and 0.5 mM TCEP. The eluted proteins were treated with tobacco etch virus (TEV) protease for 18 h at 4°C (1:20 mass ratio) to remove the hexa-histidine tag, leaving an additional serine residue at the N terminus. The proteins were further purified by size exclusion chromatography by using a Superdex S200 16/60 column (GE Healthcare) buffered in 50 mM HEPES pH 7.5, 300 mM NaCl, 5% glycerol, and 0.5 mM TCEP, at a flow rate of 1 ml/min.

For purification of the untagged protein, cell extracts were clarified by centrifugation and DNA was removed using a diethylaminoethyl cellulose (DE52, Whatmann) column. The cleared lysate was loaded onto a 5-ml HiTrap Q HP ion-exchange column

(GE Healthcare) that was previously equilibrated with buffer containing 10 mM TRIS pH 7.4 and 25 mM NaCl. The flow-through contained the protein, whereas impurities were retarded by the column. The flow-through was subsequently loaded onto 5 ml HiTrap™ Blue HP column (GE Healthcare), and eluted using a linear gradient of up to 1M NaCl. The positive fractions were pooled and dialysed against 10mM HEPES, pH 7.4, 25mM NaCl, and 5mM DTT.

For all proteins, positive fractions as analysed by SDS-PAGE (Bio-Rad or Novex) were pooled and concentrated with Amicon-Ultra centrifugal concentrators (Millipore) with a molecular weight cut-off of 10 or 30 kDa. Protein aliquots were flash frozen in liquid nitrogen and stored at –80ºC.

2.3.5. Mass spectometry

The correct mass of each protein was confirmed by electrospray ionisation mass spectrometry. Electrospray ionisation time-of-flight (ESI-TOF) mass spectrometry (Agilent) was used to determine the intact mass of the protein. The protein was diluted to 0.1 mg/ml in 0.1% formic acid and 10 µl of the solution was injected into the LC-ESI-TOF. The LC component of the system is a C2 reverse phase column which allows the system to effectively desalt the protein prior to application to the ESI-TOF. The system washes the column in 0.1% formic acid, 5% acetonitrile and then elutes the column with a gradient from 5% to 60% acetonitrile. The resulting data was interpreted by QS Analyst (Applied Biosystems) and MagTran software (Zhang and Marshall, 1998).

2.3.6. Absorbance spectra and measurement of protein concentration

The concentrations were determined using a NanoDrop spectrometer (Thermo Scientific, Waltham, MA) against a buffer blank, using molar extinction coefficient and molecular weight of a protein derived from the primary sequence as shown in Table 2.1 (ExPASy server (Gasteiger et al., 2003)). Protein concentrations were determined using the Beer's law (Equation 2.1), where A_{280} is the absorbance of the protein, ε_{280} is the molar extinction coefficient ($M^{-1}\,cm^{-1}$), C is the protein concentration (M), l is the path length (cm).

Equation 2.1 $$A_{280} = \varepsilon_{280}\, C\ \ l$$

2.4. Differential scanning fluorimetry

Differential scanning fluorimetry (DSF) involves the progressive heating of a protein, while a dye fluoresces upon contact with the exposed hydrophobic core of the denatured protein. With the rise of temperature, the protein unfolds inducing increase in fluorescence levels and therefore allowing detection of the protein melting temperature (T_m). Since potent inhibitors stabilise the protein in its folded conformation thereby increasing T_m, shifts in melting temperature serve as indicators of a compound-protein binding strength (Vedadi et al., 2006).

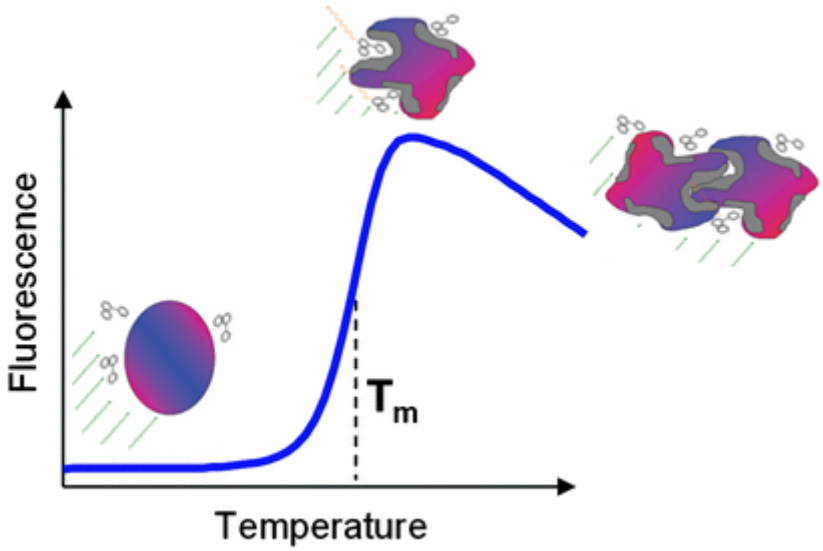


Figure 2.2 Schematic representation of a typical recording of during a DSF experiment. The fluorescence of the dye Sypro Orange is dramatically increased upon binding to hydrophobic patches during protein unfolding (adapted from (Niesen et al., 2007)).

Each protein was buffered in 10 mM HEPES, 500 mM NaCl, pH 7.0 at a final concentration of 2 µM and mixed with an environmentally sensitive fluorescent dye SYPRO-Orange (Molecular probes, 1000x concentrated stock). Ligands were added at a final concentration of 10 µM. The samples were placed in a 96-well plate, heated from 25 °C to 96 °C at 3 °C per minute and fluorescence levels were recorded at each step by a Stratagene Mx3005P real-time PCR instrument. The change in fluorescence was monitored using excitation and emission filters set to 465 and 590 nm respectively. Data were processed using Prism v5.01 (GraphPad Software) and the inflection point of the transition curve (T_m) was calculated using Boltzmann (Equation 2.2), where *LL* and *UL* are the values of minimum and maximum intensities respectively, *d* is the slope of the curve at T_m (Niesen et al., 2007).

Equation 2.2

$$y = LL + \frac{(UL-LL)}{1+e^{\frac{Tm-x}{d}}}$$

2.5. Enzymatic assays

Microfluidic electrophoretic mobility-shift assay was used to compare activities of kinases in a panel, observe the effects of mutations and establish kinetic parameters of binding. The assay is based on the electrophoretic separation of phosphorylated (product) and non-phosphorylated (substrate) peptides based on their charge difference [11] (Figure 2.3).

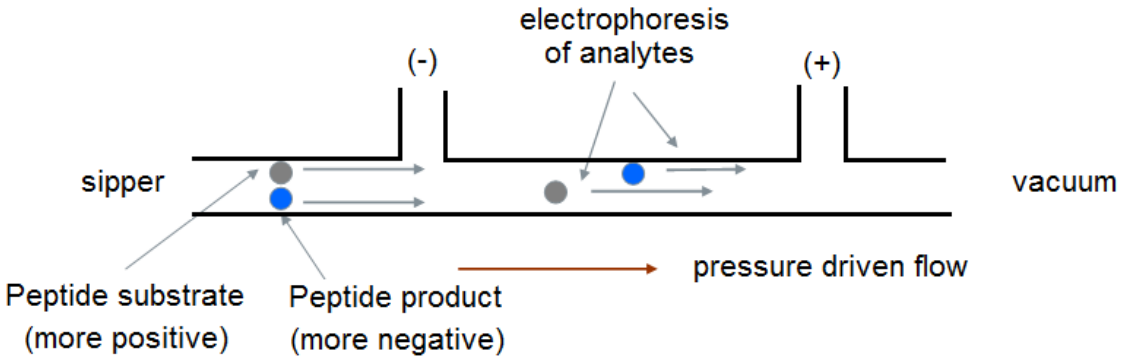


Figure 2.3 Principles of microfluidic electrophoretic mobility-shift assay technology. Due to difference in charge, the phosphorylated peptide (product) migrates through the chip faster than the non-phosphorylated (substrate). The signals from the fluorescent tag attached to the peptide are detected and their relative ratio serves as a measure of enzyme activity.

The peptides are tagged with fluorescent marker and the shift in mobility was detected via laser-induced fluorescence allowing for direct and real-time measurements of kinase activities (sequences: 5-FAM-QSPKKG-CONH2 for CDK2 and 5-FAM-AKRRRLSSLRA-COOH for PIM). Results were expressed as a conversion ratio by measuring peak heights for both the substrate and product and dividing the product peak height by the sum of peak heights for both substrate and product.

Kinase activity against peptides was measured using the Caliper protein kinase profiling system (Caliper Life Sciences, Hopkinton, MA, U.S.A.). Concentrations for the enzyme and ATP were assay specific and optimised for each kinase individually during assay optimisation. The assay optimisation started with determination of the lowest

protein concentration that provides sufficient signal for analysis under conditions of saturating peptide and ATP. This was done to ensure a steady linear reaction with substrate consumption less than 20% so that the initial enzymatic velocity (*V)* follows the Michaelis-Menten equation (Equation 2.3), where V_{max} is the maximal velocity at saturating substrate conditions.

Equation 2.3 $$V = \frac{V_{max}[S]}{K_m + [S]},$$

Next, the Michaelis constants for ATP (K_m^{ATP}) and peptide ($K_m^{peptide}$) were obtained from measurements conducted at constant enzyme concentration and a range of ATP and peptide concentrations respectively. Equation 2.3 was applied to the curve obtained by plotting the measured enzyme velocities versus the corresponding substrate concentration, leading to the values for K_m and V_{max} by a nonlinear regression fit (GraphPad Software). The k_{cat} values were then defined as shown in Equation 2.4, where $[E_0]$ is the total enzyme concentration.

Equation 2.4 $$k_{cat} = {V_{max}}/{[E_0]}$$

For IC_{50} measurements, a 384 assay well microtiter plate contained 8-point serial dilutions of inhibitors and 16-point serial dilutions of staurosporine as a reference compound. First, 4.5 µl of 2 x kinase solution was added to 0.05 µl compound (maximum concentration 1.8 mM in 90% DMSO and 10% H_2O). To allow for possible slow association, plates with type 2 inhibitors were incubated for 60 min at 30°C. The assay started after addition of 4.5 µl of 2 x peptide with ATP and run for 60 min at 30°C, after which 16 µl of stop solution was added (100 mM HEPES pH 7.5, 5% DMSO, 0.1% coating reagent (Caliper Lifescience) 10 mM EDTA pH 8.0, 0.015% BRIJ35). All reactions were performed in 50 mM HEPES, pH 7.5, 1 mM DTT, 0.02%

Tween 20, 0.02 % BSA, 10 mM beta-glycerophosphate, 0.01 mM Na_3VO_4, 0.6 % DMSO and 2 µM peptide.

Liquid handling and incubation steps were done on a Thermo CatX workstation equipped with Innovadyne Nanodrop Express. Stopped kinase reactions were analysed in a LC3000 reader (Caliper Lifescience), where both substrate and product were quantified by measuring the laser-induced fluorescence intensities of the peptide's fluorescein labels and calculating the turnover. The dose-dependent influence of the compound on kinase activity was as an IC_{50} value. Each IC_{50} value was calculated from the plot of percentage of inhibition relative to controls versus inhibitor concentration according to Equation 2.5 using GraphPad Software or Helios.

Equation 2.5

$$y = A_2 + \frac{A_1 - A_2}{1 + \left({}^{x}/_{IC_{50}}\right)^{P}},$$

where y is the inhibition fraction at the inhibitor concentration x, A_1 is the lowest inhibition value, A_2 is the maximum inhibition value and *P* is the Hill coefficient.

2.6. Isothermal titration calorimetry

Direct measurement of protein binding energetics was performed by isothermal titration calorimetry (ITC). This technique was used as a benchmark to validate hits established during DSF and Caliper experiments. In a single experiment ITC provides binding affinity constant (K_d) and the stoichiometry (n) of bimolecular interactions, and in addition allows resolution of thermodynamic terms that define the binding affinity (K_a), the enthalpy (ΔH) and entropy (ΔS) changes.

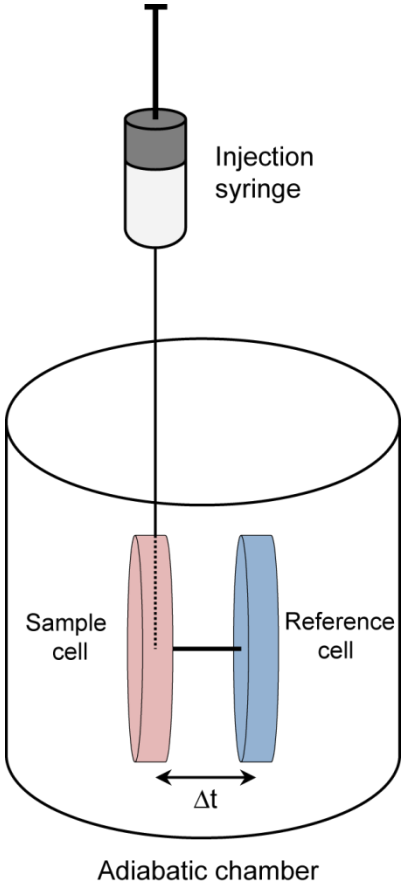


Figure 2.4. Schematic representation of a power compensation in an isothermal titration calorimeter.

All ITC experiments were performed using a VP-ITC (Microcal, Northampton, MA) instrument, which is based on the principle of isothermal power compensation (Leavitt and Freire, 2001, Holdgate and Ward, 2005). Two identical cells are surrounded in an adiabatic chamber in order to minimise any heat flow to or from the system (Figure 2.4). The binding process occurs in a sample cell containing protein, which is accessible for delivery of ligand solution from the injection syringe, while the reference cell serves as a thermal blank and filled with water. When binding occurs, the heat (Q_p) is generated (exothermic reaction) or absorbed (endothermic reaction) by the sample cell, resulting in a temperature difference between the two cells. Once the temperature difference is detected by the heat sensing device, a feedback system adjusts and records the amount of power supplied to the sample cell over time (*P(t)*), minimising the temperature difference between the sample and reference cells. In case of an

exothermic reaction, the applied power decreases, while for endothermic reactions the feedback power to the sample cell increases, thereby maintaining a close-to-zero temperature difference between the cells.

This way the electrical energy (W_e) crossing the boundary of the calorimetric system compensates the temperature changes caused by the chemical reaction. Because the system is adiabatic and isobaric, change in enthalpy after each injection equals the heat associated with the binding interaction, which is obtained by calculating the area under each peak according to Equation 2.6, where ΔH is the amount of heat released per mole of ligand bound.

Equation 2.6 $$\Delta H = Qp = W_{el} = \int_{t_1}^{t_2} P(t)dt$$

Decreasing magnitude of peaks upon each injection indicate saturation of binding. When complete saturation is achieved, subsequent injections produce small consistent peaks corresponding primarily to dilution of protein and heat effect due to mixing; these heats need to be subtracted from all the injection peaks before analysis.

All experiments were carried out at 10 or 15°C in 50 mM HEPES pH 7.5, 300 mM NaCl and 1 mM DTT and all protein samples were dialysed against this buffer in advance. The concentration of the proteins was chosen between 0.1 mM and 0.2 mM, while the ligand was 10 to 15 times less. A typical titration was conducted using an initial injection of 2 µl, followed by 25 identical injections of 10 µl at time intervals of at least 240 s.

Thermodynamic parameters were calculated using origin 7.0 software (OriginLab, Northampton MA) provided with the equipment and in all cases a single binding site model was employed. With knowledge of the volume and concentration of both the reactants, the experimental data were analysed using a non-linear least squares curve-fitting algorithm with three floating variables: stoichiometry (n), association constant (K_a), and the change in enthalpy (ΔH) of the interaction. From these parameters the

change in both entropy (ΔS) and Gibbs free energy (ΔG_b) of association can be calculated using the Equation 2.7 and Equation 2.8.

Equation 2.7 $$K_d = {}^{1}/_{K_a}$$

Equation 2.8 $$\Delta G_b = \Delta H - T\Delta S = RT\, ln\, K_d$$

2.7. Kinetics experiments based on biolayer interferometry

Biolayer interferometry (BLI) is a technology based on interferometry developed by FortéBio to monitor molecular interactions. The biosensors utilised in this technique are made of glass fibre with an additionally engineered biocompatible surface where molecules attach. When white light is directed down the fibre, it partially reflects from the two surfaces: one between the glass fibre and biocompatible surface and the second between the layer of biomolecules and the solution (Figure 2.5). Once the surface tip of the glass fibre contains immobilised protein, the thickness of the bimolecular layer increases and causes a shift in the path length of the reflection wave from the interface with the solution. Specific interaction of the immobilised molecule with the ligand additionally changes the biomolecular layer thickness on the sensor tips, leading to the spectral shift ($\Delta\lambda$). The reflected light is detected by the sensor and the shifts of spectra in the interference pattern are monitored in a real-time manner, providing kinetic data of binding events.

The binding of CDK2 to inhibitors was analysed on an Octet RED 384 instrument (FortéBio), which offers fast, and cost-effective, compared to surface plasmon resonance (SPR), analyses for multiple real-time characterisations of biomolecular interaction. Each experiment was performed at 25°C using 50 mM TRIS pH 7.5, 300 mM NaCl, 0.005% P20 Tween and 1 mM DTT as buffer. The buffer-equilibrated super streptavidin biosensors (FortéBio) were loaded with 3 µg/ml biotinylated kinase

for 300 s, washed in buffer for 300 s, incubated with inhibitors for at least 300 s, followed by washing in buffer for at least 500 s. Typically, small molecule stock solutions were serially diluted (2.5-fold, 8-point serial dilutions) in buffer. The binding data were fitted using the Octet analysis software 7.0, where the association and dissociation curves were corrected for non-specific binding by subtracting two references: sensors with no kinase incubated with inhibitors and sensors with kinase incubated in buffer.

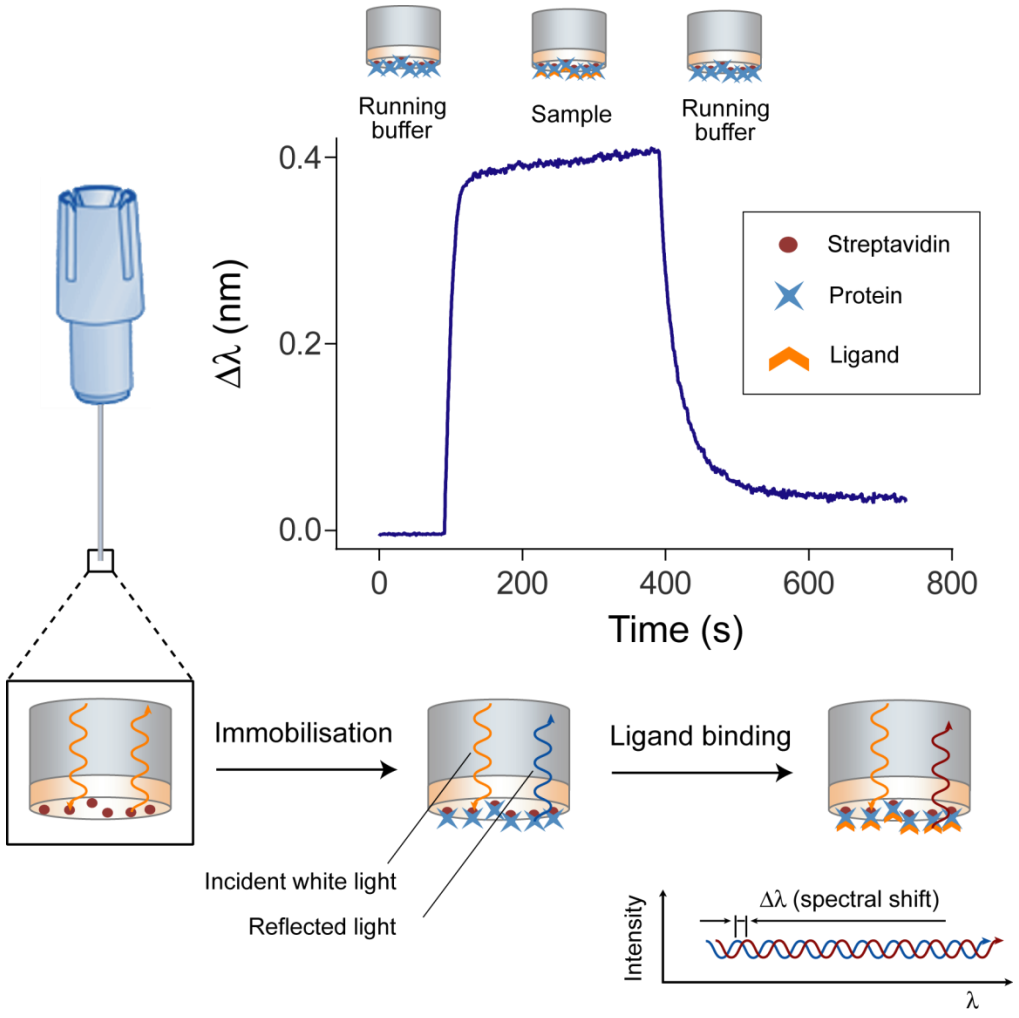


Figure 2.5 Principles of the FortéBio biolayer interferometry. White light is reflected through the glass fibre optic biosensor and capturing of the ligands by the protein changes the thickness of molecular layer resulting in shift of the light spectra that can be measured in real time.

2.8. Crystallisation

All crystallization was performed using the sitting drop vapour diffusion method at 4°C in 150 nl drops. CDK2 was buffered in 10 mM Tris-HCl pH7.4, 25 mM NaCl, 5 mM DTT with 1mM compound added, and concentrated in a 10-kDa-cutoff Amicon Ultra-15 concentrator (Millipore) to 10 mg/ml. Crystals were grown in drops in a reservoir solution containing 10% PEG4000, 0.1 M HEPES pH 7.5 and 0.05 M ammonium sulphate. On mounting, CDK2 crystals were cryo-protected with an additional 25% glycerol before vitrification in liquid nitrogen.

Each of the PIM proteins buffered in 50 mM HEPES pH 7.5, 300 mM NaCl, 5 mM DTT was mixed with 0.6 mM consensus substrate peptide Pimtide (ARKRRRHPSGPPTA) and 1 mM inhibitor, and then concentrated to 6-7 mg/ml.

Crystals of V126A PIM1 mutant were grown in a precipitant solution containing 0.2 M Na_2SO_4, 0.1 M bis-trispropane at pH 6.5, 20% polyethylene glycol (PEG) 3350 and 10% ethylene glycol. Larger crystals were obtained by using follow up screening plates with similar conditions.

PIM2 was crystallised in drops in a reservoir solution containing 0.1 M bis-tris-propane pH 7.0, 0.25 M potassium thiocyanate, 20% PEG 3350 and 10% ethylene glycol. After PIM2 crystals were formed, they were soaked in a solution containing the inhibitor in order to obtain PIM2 bound with the inhibitor. Prior to vitrification in liquid nitrogen, all PIM crystals were briefly soaked in mother liquor supplemented with 25% ethylene glycol.

2.9. Structure determination and refinement

All native diffraction data were collected at 100 K at the Diamond Light Source. Data processing and refinement were performed with the assistance of Dr Jon Elkins and Dr Apirat Chaikuad using the CCP4 suite of software (1994). Diffraction images were indexed and integrated using iMOSFLM (Battye et al., 2011) and data were scaled using SCALA (Evans, 2006). Phases were calculated by molecular replacement in Phaser (McCoy et al., 2007) with the corresponding kinase structure as a search model. Model rebuilding was carried out using COOT (Emsley et al., 2010) interspersed with cycles of restrained refinement with REFMAC5 (Murshudov et al., 1997). During later stages of refinement, translation/libration/screw (TLS) parameters were introduced. Progress of the refinement for each structure was judged throughout by following a reduction in R_{free} (calculated from 5% of the data that was excluded from the refinement). The stereochemical properties and quality of the final model were assessed with the MOLPBROBITY software (Davis et al., 2007). Structural figures and graphical renderings were made with ICM (Molsoft LLC, San Diego, CA) or PyMOL Molecular Graphics System (Schrödinger, LLC).

3. DFG-in/out transitions in CDK2

3.1. Design of point mutations

CDK2 has been chosen as a model to study DFG-in/out conformational transitions. This kinase could be easily expressed, had the largest structural coverage (over 180 PDB depositions) and most importantly, despite this large amount of structural data, no DFG-out conformation has yet been observed. Therefore this kinase served as a prototype of a kinase that could not easily assume a DFG-out conformation. By substituting amino-acids that may influence DFG movement, the goal was to convert CDK2 into a kinase that can more easily assume DFG-out and can therefore be targeted by type 2 inhibitors.

A large scale sequence and structural comparison of the human kinome has been performed at Novartis (Möebitz, unpublished data) and used to elucidate and propose mechanisms that govern the conformational plasticity of protein kinases. To start with, a unified nomenclature was introduced in order to match the conserved features across different sequences. The major kinase conformations were then classified followed by comparison of multiple sequence alignments between the conformations. To prospectively identify residues that could be involved in stabilising the inactive DFG-out state, the analysis of DFG-in vs. DFG-out conformations was combined with type 1 and type 2 inhibition data. As a result, we focused our attention on the three prime candidates for mutagenesis.

The first and the most obvious place to make a point mutation was the gatekeeper, a residue preceding the hinge region (F80 in CDK2). The size of the gatekeeper side chain has been recognised to control the accessibility to the deeper hydrophobic pocket inside the ATP-binding site and controlling kinase selectivity for ATP-competitive inhibitors (Zuccotto et al., 2010, Garske et al., 2011b). Mutations of this residue are frequently associated with the development of resistance mechanisms when bulkier amino-acids occupying this position prevent drugs from binding to their

target kinase (KIT (Tamborini et al., 2004), PDGFRA (Cools et al., 2003), EGFR (Kobayashi et al., 2005), EML4-ALK (Choi et al., 2010), and FLT3 (Smith et al., 2013) kinases as examples). Several potential explanations of this resistance have been suggested, including simple steric blocking of the deeper hydrophobic pocket and stabilisation of the hydrophobic spine, which is necessary for kinase activity (Kornev et al., 2008, Kornev and Taylor, 2010). In addition, large aromatic gatekeepers are known to form contacts with the αC-helix, thereby shifting the equilibrium to the active conformation (Taylor and Kornev, 2011, Wang et al., 2006) and numerous gatekeeper mutations leading to enhanced kinase activity have been reported (Zuccotto et al., 2010, Choi et al., 2012). The analysis revealed that in general kinases with threonine or methionine at the gatekeeper position have the ability to assume a DFG-out conformation, unlike those with phenylalanine, leucine or valine in that position, perhaps due to stabilisation of the active state. Therefore, mutants F80T and F80M would be expected to make it easier to adopt a DFG-out conformation.

The second position for point mutations was cysteine 118 located at the αE helix (Figure 3.1). We observed that in CDK2 the DFG-Phe146 is ‘sandwiched’ between the large gatekeeper (F80) and C118. Besides, the side chain of the C118 is at a distance of 3.8Å from the aromatic ring of the DFG-Phe (closest atom to atom contact) and complies with the geometry known to form a strong sulphur-arene interaction in protein structures (Salonen et al., 2011, Chakrabarti and Bhattacharyya, 2007). At the time of the kinome analysis, the only non-tyrosine kinases observed in DFG-out conformations were p38α, Aurora A, CDK6 and CDK7. Sequence alignment of the nine CDK2 family members indicated that both CDK6 and CDK7 had leucine at the position of C118 (Figure 3.1c). Interestingly, about 90% of tyrosine kinases, which account for the vast majority of deposited DFG-out structures (Figure 3.1d), have leucine at this position. Regulation of the tyrosine kinases often relies on a native inactive state, unlike the other kinase families which tend to have native active conformations (Möbitz and

Fabbro, 2012). All these observations resulted in the hypothesis that the interaction of C118 with F146 stabilises the DFG motif in an active 'in' conformation. Cysteine was therefore mutated to leucine (C118L) to remove sulphur-aromatic interaction and allow for rotation of the DFG-Phe into the ATP-binding pocket, and to isoleucine to investigate the effect of this residue on αC-helix displacement.

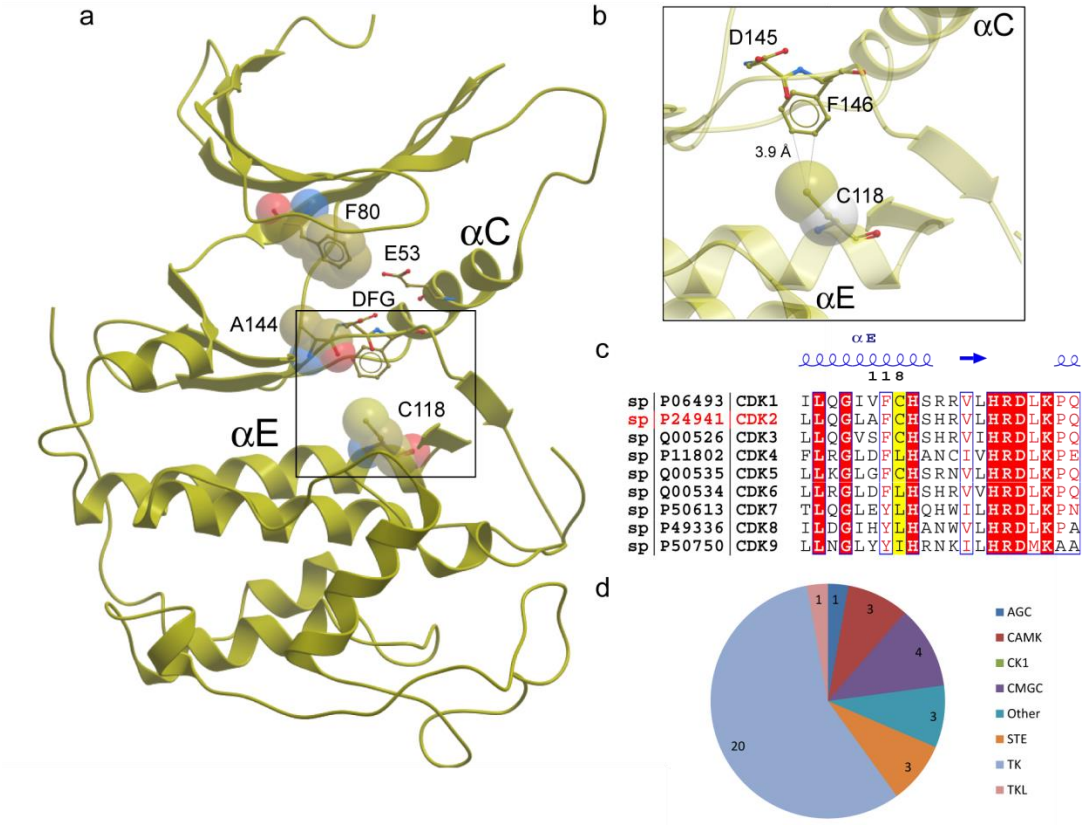


Figure 3.1 Location of CDK2 mutations. (a) Structure of CDK2 (PDB: 1qmz) shown as a ribbon diagram with mutated residues highlighted and important structural elements indicated. (b) Magnified view of the DFG motif and the favourable interaction between F146 and C118 is indicated with dashed lines. (c) Sequence alignment of nine human CDK family members. The position corresponding to Cys118 of CDK2 located at the αE helix is highlighted in yellow. (d) Distribution of DFG-out structures in individual kinases across human kinome families as deposited to the PDB by November 2011.

The third position for point mutations was the residue preceding the DFG motif (termed DFG-1) where the side chain points towards the ATP-binding pocket. This residue often makes direct contacts with the bound ligand and like the gatekeeper, stabilises the regulatory hydrophobic spine (Taylor 2010). While alanine is the most commonly observed residue across the kinome in this position, glycine and cysteine appear to play a role in lowering the torsional barrier for the DFG backbone rotation and stabilising the end state of the DFG-out. Our strategy involved mutating the DFG-1 position to cysteine (A144C) and glycine (A114G) followed by generation of double mutants (C118L/A144C and C118L/A144G) to combine the potential effects.

3.2. Screening procedure

The work plan depicted in Figure 3.2 has been applied to test the predictions experimentally. Residues suggested by structural comparisons to facilitate the adoption of a DFG-out conformation were mutated by site-directed mutagenesis and the mutated protein was expressed and purified (Appendix G). The sensitivity of the mutants to the type 1 and type 2 inhibitors was tested according to the hypothesis that type 2 inhibitors would be inactive on the wild-type CDK2, while binding to a mutant would suggest that a DFG-out conformation has been successfully induced. The assays were carefully selected to investigate the effect of the compound on enzymatic activity and ligand binding.

First, techniques with rapid turnover and low associated cost (DSF, Caliper) were used to identify compound hits and provide basic information about binding. Second, hits were validated with ITC experiments, which allowed for accurate determination of binding affinity and thermodynamic parameters. Since each ITC run requires large amounts of purified proteins, only a limited number of interactions were feasible to study. Third, BLI methods were implemented to investigate association and

dissociation rates of compound binding. And finally, protein crystallography was used to confirm the compound binding mode.

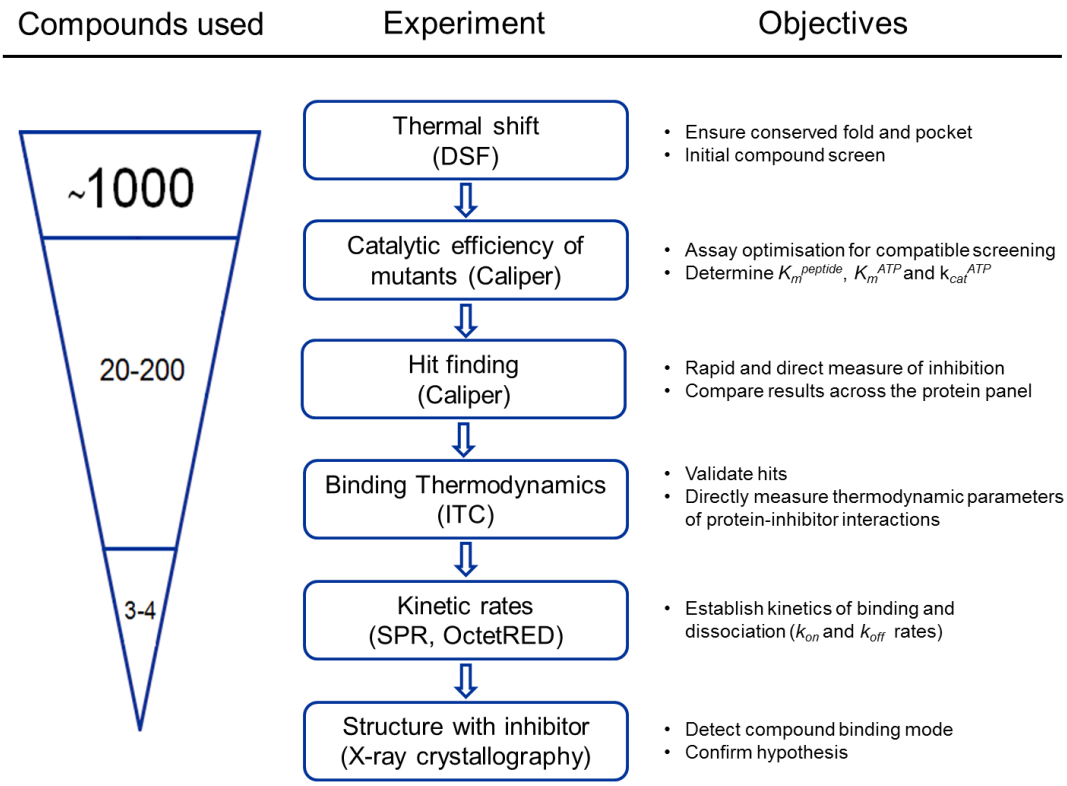


Figure 3.2 Flow diagram showing the experimental procedures for compound screening. Each experimental activity, objectives and estimated number of compounds used during the activity are illustrated.

3.3. Evaluation of structural integrity and stability of mutants

The impact of mutations on the structural integrity and thermal stability of CDK2 was investigated using DSF, which involves the progressive heating of a protein while a dye fluoresces upon contact with the exposed hydrophobic core of the denatured protein. Therefore with the rise in temperature, the protein unfolds inducing an increase in fluorescence levels the measurement of which allows determination of the protein melting temperature (T_m).

The wild-type enzyme exhibited a melting temperature of 52.4 ± 0.5 °C (Table 3.1). Overall, the mutants demonstrated similar behaviour with unfolding temperatures ranging from 48 °C to 55 °C, confirming that introduced mutations had no strong effect on the overall protein fold and stability. It should be noted, however, that the average T_m of the A144C mutant was about 4 °C lower (48.0 ± 0.3 °C), suggesting that the cysteine introduced at the binding site could affect N- and C-lobe interactions resulting in slightly earlier unfolding of the kinase. This effect is suppressed when a second point mutation is introduced as for the C118L/A144C and C118L/A144G double mutants, which had similar melting temperatures to the wild-type (51.8 ± 0.2 °C and 53.0 ± 0.2 °C respectively).

Table 3.1 Melting temperatures (T_m) obtained for the wild-type CDK2 and mutants using DSF show that mutations introduced have no strong effect on CDK2 stability. Mean values and standard deviations were derived from at least 40 independent experiments in each case.

Protein	T_m (°C)
Wild-type	52.4 ± 0.5
F80T	54.3 ± 0.2
F80M	52.3 ± 0.3
C118L	55.1 ± 0.3
C118I	54.0 ± 0.2
A144C	48.0 ± 0.3
A144G	50.6 ± 0.2
C118L/A144C	51.8 ± 0.2
C118L/A144G	53.0 ± 0.2

3.4. Binding of type 2 inhibitors to WT was suggested by DSF

The thermal shift based assay was also used as a rapid and low-cost screening method to identify inhibitors that bind to CDK2 variants. Ligand binding can stabilise the protein in its folded conformation, leading to an increase in the temperature at which unfolding begins. Changes in T_m therefore serve as indicators of a compound-protein binding strength.

Scatter plots (Figure 3.3) illustrate differences in binding of inhibitors to wild-type CDK2 and the mutants. The change in melting temperature (ΔT_m) was calculated as the difference between the melting temperatures of the protein in the presence of ligand compared with the protein alone. Once ΔT_m values for all proteins were calculated, the data were sorted from smallest to largest ΔT_m for CDK2. Then the compounds were subdivided in two groups: type 1 and type 2 inhibitors.

In general, a bigger temperature shift indicates higher stabilisation of a protein against thermal denaturation and thus corresponds to stronger binding. As seen from Figure 3.3a, the gatekeeper substitution F80T caused a dramatic reduction of type 1 inhibitor binding. For example, for the compound used as a positive control, staurosporine, ΔT_m changed from 13.6 °C to 0.9 °C, indicating a complete lack of binding to the F80T mutant. Despite that, an inhibitor with a type 2 scaffold, K03859 was noticed to bind to the F80T mutant giving a ΔT_m of 11.9 °C, while for the wild-type ΔT_m was only 2.7 °C. The F80M (Figure 3.3b) and C118I mutants (Figure 3.3d) did not demonstrate any notable difference in compound binding, while the inhibitor interactions with the A144G (Figure 3.3f) and C118L/A144G mutants (Figure 3.3h) were slightly weaker than that of the wild-type, particularly for type 2 binders.

Introduction of a cysteine to the ATP-binding pocket (DFG-1 position) resulted in higher affinity of type 1 inhibitors for the A144C variant (Figure 3.3e) with milder effect for the double mutant C118L/A144C (Figure 3.3g). This can be explained by a possible S-H/π interaction of the cysteine with an aryl ring frequently present in kinase inhibitors

coupled with the aromatic-aromatic edge to face interaction with the phenylalanine gatekeeper (F80) (Furet et al., 2006). The type 2 compounds demonstrated higher affinity for the C118L (Figure 3.3c) and A144C mutants (Figure 3.3e), while the binding data for the C118L/A144C mutant suggest a potential synergistic effect of these amino-acid substitutions (Figure 3.3g).

Overall, the thermal studies confirmed the conserved protein fold and pocket shape for each mutant and suggested type 2 binding to the CDK2 variants. The melting temperature shift assay is a good first indication, however the data produced are hard to quantify and could be subject to artefacts (Hau et al., 2011). Therefore these thermodynamic studies were complemented with enzymology (Figure 3.2).

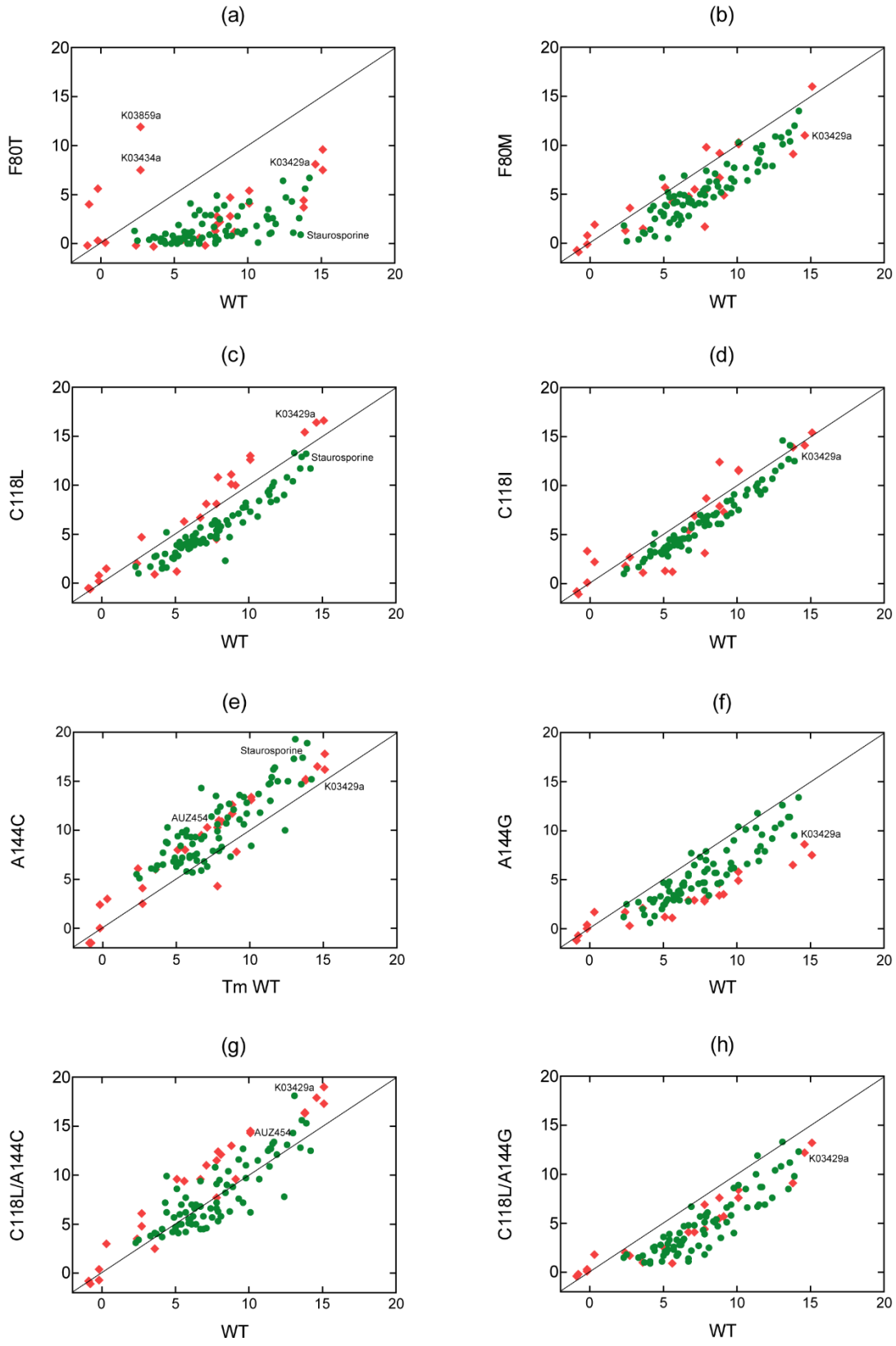


Figure 3.3 Scatter plots of relative compound binding for each CDK2 mutant with respect to the wild-type (WT) as established by DSF. Each value represents a shift in melting temperature (ΔT_m in °C) calculated as the difference between the melting temperatures of the protein in the presence of ligand compared with the apo-protein. Type 1 compounds are shown as green circles and type 2 compounds are shown as red rhombuses. Compounds of interest are labelled.

3.5. Reconstitution of CDK2 and cyclin results in active CDK2/cyclin complexes

The objective of this project was to demonstrate that certain amino-acid positions have effect on the orientation of the DFG motif. Hence, identification of type 2 compounds that bind to a CDK2 mutant(s) but not to the wild-type would provide experimental evidence for a shift in conformational equilibrium. Since 'inactivating' mutations were introduced, we also characterised the catalytic efficiency for all CDK2 variants by following the kinetics under steady-state conditions. All enzymatic assays were conducted using microfluidic electrophoretic mobility shift assay (Caliper) described in Chapter 2.

As an enzyme, CDK2 alone is inactive and at least two steps are necessary for its full activity (Chapter 1). First, activation of CDK2 requires complex formation with a cyclin protein, like cyclin A or cyclin B, leading to significant conformational changes in and around the ATP binding pocket of the kinase. Second, phosphorylation of the T160 residue by an upstream kinase results in organisation of the substrate recognition site (Figure 1.3).

We discovered that CDK2 can be activated to a level sufficient for enzymology experiments by simply mixing CDK2 with either cyclin A or cyclin B *in vitro*, with no need for co-expression or addition of other protein-activators. A thorough optimisation process was needed to have a reliable and reproducible biochemical screening assay. First, titrations of each CDK2 variant were performed to define enzyme concentrations, which result in a sufficiently high assay signal and at the same time ensure steady-state conditions during the initial phase of the reaction, when about 10% of the substrate has been converted into product. When the conditions were maintained in the steady state, the initial velocity of the reaction was defined by the slope of the amount of product formed versus time plot making the reaction rate compliant with the Michaelis-Menten model for enzyme kinetics. The optimum concentration for each

variant was established under conditions of saturating peptide substrate, ATP and cyclin A. We observed that the mutations clearly reduced the kinases' ability to phosphorylate the substrate peptide and therefore larger amounts of mutants compared to the wild-type were necessary for a detectable signal (Table 3.2).

Second, cyclin A titration experiments were performed for each CDK2 variant in order to establish the optimum CDK2/cyclin A ratio when the kinase is fully saturated. Figure 3.4 shows that without cyclin A the kinase is inactive and addition of cyclin A results in the formation of an active complex. Once a 1:1 molar ratio was reached, CDK2 was fully saturated with cyclin A and demonstrates maximum activity. For all the following enzymology experiments, equimolar amounts of CDK2 and cyclin A were used.

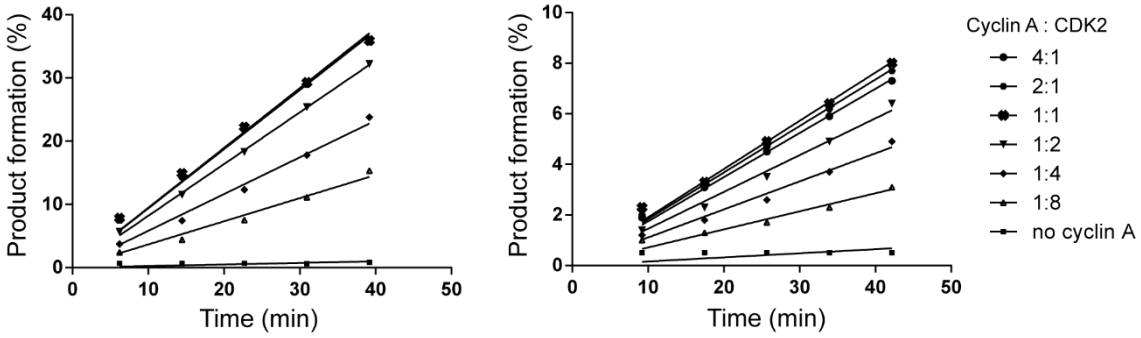


Figure 3.4 Cyclin A titration experiments using the Caliper assay demonstrate that addition of cyclin A activates CDK2. Both the wild type (left) and the A118L/A144C mutant (right) reach a maximum rate of peptide substrate phosphorylation at a 1:1 molar ratio with the cyclin A. The mutant has a lower rate of substrate conversion than the wild type confirming the inactivating nature of the amino-acid substitution.

3.6. Impact of mutations on enzymatic activities

A key step in the evaluation of the CDK2 variants involved studying the effect of the mutations on ATP and peptide binding by establishing corresponding K_m values under initial velocity conditions.

The $K_m^{peptide}$ value is defined by the peptide concentration which results in half maximal reaction velocity. Peptide titration experiments resulted in similar $K_m^{peptide}$ values for all

CDK2 kinases, indicating that the mutations did not affect peptide binding, so the same peptide concentration has been used for all the following assays.

However, K_m^{ATP} values, defined as ATP concentration that results in half maximal reaction velocity, were observed to be increased for all mutants meaning that inactivating mutations reduced affinity for ATP. The ratios of Michaelis–Menten constants, k_{cat}/K_m^{ATP}, were used as descriptors of catalytic efficiency of CDK2 variants, where k_{cat} is the catalytic rate of reaction (Table 3.2).

Comparison of catalytic rates revealed that the mutants' ability to phosphorylate the peptide substrate was decreased, confirming the inactivating nature of the mutations. On average, the efficiency of the C118I mutant was decreased 10 times compared to that of the wild-type, while the F80M, C118L and A144G mutants demonstrated nearly 70-fold decrease. In contrast, the DFG-1 mutant A144C had a similar K_m^{ATP} value to the wild-type, however, the rate of catalysis was almost halved. Interestingly, the double-substitution variant C118L/A144C seemed to balance the effects of the single mutants. The extreme cases of F80T and C118L/A144G amino-acid substitutions demonstrated very low affinities for ATP and exact establishment of K_m^{ATP} was limited by the assay capacity: we observed that in kinetic mode, increase of ATP concentration beyond 1 M resulted in reduction of catalytic activities. The data for these mutants therefore were not suitable for curve fitting since there was no convergence to a plateau.

We noted that the complex of wild-type CDK2 with cyclin A mixed *in vitro* was less efficient in product formation than the co-expressed complex provided by Novartis, showing the effects on both the apparent ATP binding constant (K_m^{ATP}) and on the catalytic rate (k_{cat}). The insect-cell produced protein is likely to be phosphorylated on T160, which explains higher activity. Because T160 phosphorylation was absent in all other CDK2 variant, partnering with cyclin only partially activated the kinase. This was key to the success of the screen to identify type II inhibitors that are designed to target the inactive state.

Table 3.2 Steady-state kinetic constants for the ATP titration experiments determined using Caliper microfluidic mobility shift assay.

CDK2 variant	Kinase concentration (nM)	K_m (μM)	k_{cat} (s^{-1})	k_{cat}/K_m (M^{-1} s^{-1} x 10^4)
WT	20	44.0 ± 14.5	18.1 ± 5.5	41.2 ± 18.5
F80T	100	Not converged*		
F80M	100	432.2 ± 21.8	2.6 ± 0.4	0.6 ± 0.1
C118L	50	328.1 ± 76.8	1.9 ± 0.2	0.6 ± 0.1
C118I	100	96.1 ± 16.1	3.3 ± 1.9	3.5 ± 2.0
A144C	100	51.6 ± 13.3	9.8 ± 0.6	18.9 ± 5.0
A144G	100	302.0 ± 74.6	2.1 ± 0.6	0.7 ± 2.5
C118L/A144C	100	52.4 ± 16.5	1.6 ± 0.3	3.1 ± 1.2
C118L/A144G	100	Not converged*		
co-expressed WT	10	27.2 ± 19.3	40.8 ± 8.4	149.9 ± 110.7

* A curve could not be fitted because the data did not converge to a plateau.

3.7. Enzymatic assay detected type 2 compound binding to the mutant but not the wild-type CDK2

The sensitivities of CDK2 and the mutants to kinase inhibitors were established and presented as half-maximal inhibitory concentrations (IC_{50}). Since all screened inhibitors were ATP-competitive, the ATP concentration is an important factor in the ability of a compound to inhibit the kinase. Therefore, for comparable quantitative analysis of the IC_{50} values across the kinase panel, the ATP concentration for each variant was chosen to be equal to K_m^{ATP}. For the F80T and C118L/A144G mutants ATP concentration was set to be 450 μM, corresponding to the maximum K_m^{ATP} value for the CDK2 panel (Table 3.2). Unfortunately data collection for the C118L/A144G mutant was unsuccessful due to undetectable activity of the kinase.

Several compounds that bind to the mutants but not the wild-type have been detected (Table 3.3). For example, binding of the inhibitor K03861a to the wild-type was not identified at 10 μM concentration (considered as 'no binding'), however, this compound clearly inhibited the mutant A144C with IC_{50} value of 7.2 μM and even more potently the mutant C118L/A144C with IC_{50} value of 0.8 μM (Figure 3.5). The results additionally confirmed exclusive binding of the compound K03859a to the F80T mutant, which was suggested by our DSF experiments (Figure 3.3). Interestingly, the chemical structure of this inhibitor is very similar to the registered type 2 inhibitor sorafenib (Chang et al., 2007), and while the gatekeeper substitution F80T resulted in dramatic affinity reduction of all typical CDK2 inhibitors, including staurosporine (Table 3.3 and Figure 3.3), further investigation of this interaction could provide useful insight into the role of polar residues in the active site of protein kinases.

Type 2 inhibitors are known to have slow association rates and to accommodate for this we tested whether compounds bind to the CDK2/cyclin A complex in a different manner to those pre-incubated with the kinase domain and cyclin A added at the start of the reaction. There was, however, no clear distinction in IC_{50} values resulting from the two experimental setups (Appendix B).

Table 3.3 The IC_{50} values (μM) for hits with type 2 scaffold and the positive control staurosporine determined by Caliper assay. The data for the C118L/A144G mutant could not be determined due to inability of the mutant to phosphorylate the substrate at a sufficient level for assay detection.

CDK2/cyclin A variant	IC_{50} (μM)					
	Staurosporine	K03861a	K03859a	Comp1	Comp2	Comp3
WT	0.01	>10	>10	>10	1.9	>10
F80T	1.94	>10	7.1	>10	8.9	>10
F80M	0.38	>10	>10	>10	>10	>10
C118L	0.06	7.2	>10	>10	1.2	3.5
C118I	0.07	>10	>10	>10	2.1	>10
A144C	0.03	>10	>10	>10	3.0	>10
A144G	0.19	>10	>10	>10	>10	>10
C118L/A144C	0.05	0.8	>10	6.5	0.4	1.7

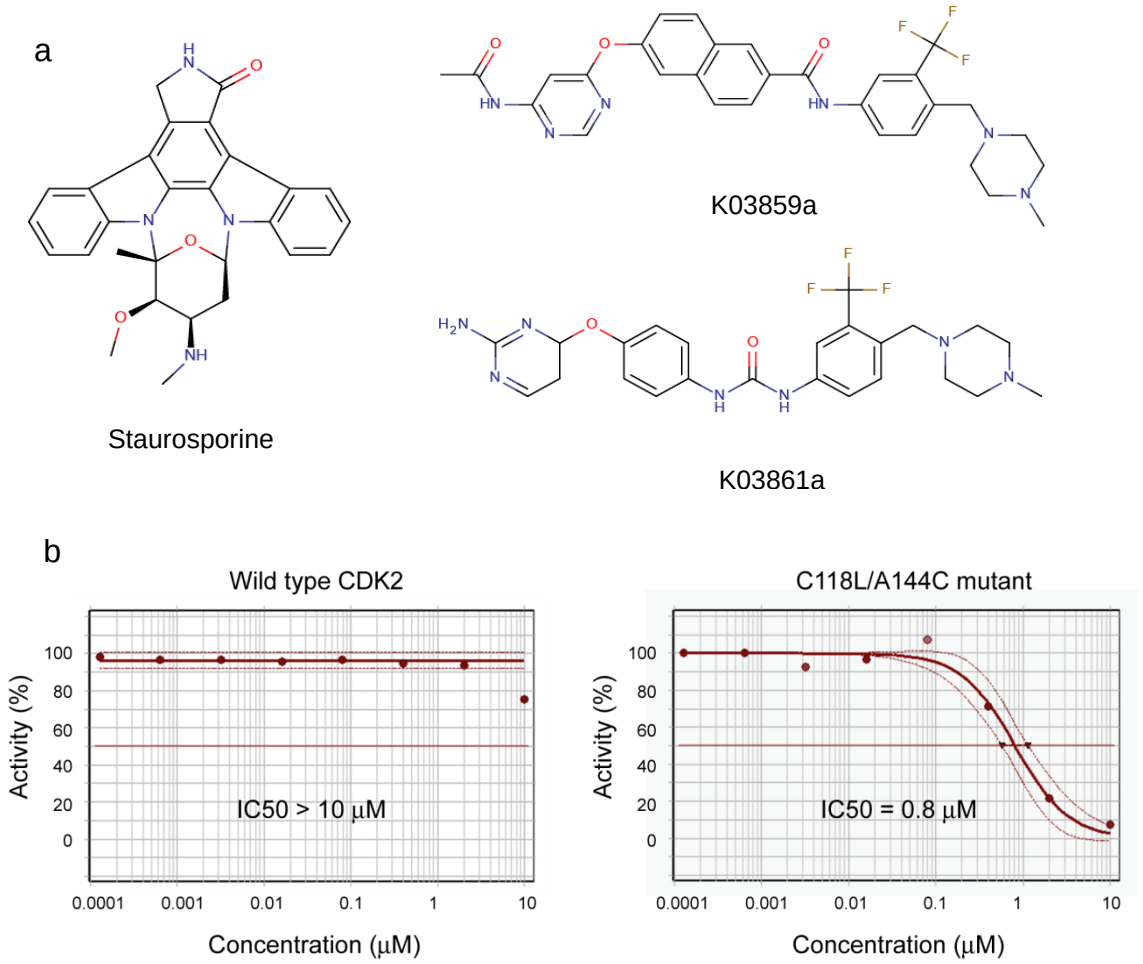


Figure 3.5 Hit finding using Caliper microfluidic mobility shift assay. (a) Chemical structures of selected compounds described in Table 3.3. (b) Dose response curves determined by Caliper assay for the wild-type complex with cyclin A the C118L/A144C complex with cyclin A mutant with K03861a inhibitor.

3.8. Type 2 inhibitors can bind to the wild-type CDK2

Caliper microfluidic mobility shift and thermal shift based assays are good pre-screening steps for identification of hits and nomination of mutants for more detailed analysis. ITC has been used as a benchmark to evaluate data from those assays. A set of hit compounds was taken through to ITC to quantitatively compare binding to the wild-type and the mutants (C118L, A144C, C118L/A144C). The limiting factor is that each ITC run requires large amounts of purified protein and the procedure is time demanding. Nonetheless, the ITC method provides a unique and most direct means for determining the thermodynamic properties of protein-ligand binding, and this technique is capable of measuring not only the magnitude of the binding affinity but also the magnitude of the two thermodynamic terms which define the binding affinity, the enthalpy (ΔH) and entropy (ΔS) changes.

The representative fitted curves (Figure 3.6) and parameter values along with errors in least-squares fitting (Table 3.4) are reported. All interactions are consistently exothermic since the peaks are below the baseline and enthalpic contributions are the dominating factor for the reaction. Titrations of C118L/A144C mutant into K03861a validated the hit, which appeared to be very potent with K_d value 8.2±1.4 nM. The interactions of single point mutants C118L and A144C were slightly weaker, with K_d values of 20.0±2.8 nM and 14.3±1.6 nM respectively, supporting the hypothesis of synergistic effect of these mutations. Because inhibition of the wild-type CDK2 was not detected during enzymology experiments (Caliper assay), there was no expectation of binding and the experiment was carried out as a negative control for completeness. Surprisingly, strong enthalpic signal was detected during titrations and K_d of 52.7±3.4 nM was established, indicating potent binding of type 2 inhibitor to the kinase domain of the wild-type CDK2.

We therefore decided to test if the presence of cyclin in the Caliper assay was the reason for no apparent inhibition of the wild-type enzyme. The ITC experiments were

repeated with CDK2/cyclin B complexes titrated into K03861a. After initial trials, the spacing interval parameter between each injection was increased from 240 s to 400 s to accommodate for the potential conformational changes during binding. No binding of the wild-type CDK2/cyclin B complex was detected but the complex of C118L/A144C mutant with cyclin B was still interacting with the compound with K_d of 135.3 ± 14.7 μM. These findings indicate that the addition of cyclin activated the CDK2 kinase domain and thus made the inactive DFG-out conformation inaccessible for type 2 binding. In contrast, inactivating mutations in the C118L/A144C variant shifted the conformational equilibrium and still allowed for a DFG-out conformation even in the presence of cyclin, showing reduced affinity (~16 fold reduction) compared to affinity when cyclin was not added. These results fully validate the findings of the previous assays, where binding of K03861a to the wild-type CDK2 was not detected. Besides, the relatively weak binding to the C118L/A144C mutant can be explained by the presence of cyclin, which was necessary for running the activity assay. In addition, potency and selectivity for the C118L/A144C mutant by K03861a and other compounds with type 2 scaffold have been previously indicated by DSF screens (Figure 3.3).

In all cases, the binding reactions were enthalpically driven and characterised by unfavourable or only slightly favourable entropy changes. The big enthalpic contribution reflects the strength of the protein-ligand interactions relative to those existing with the solvent (Leavitt and Freire, 2001), thereby the negative sign of large changes in enthalpy indicates a net gain of a number of energetically favourable contacts or an improvement of their strengths. Changes in conformational entropy as well as changes in solvation entropy are reflected by the entropy change of the reaction. Negative change of entropy values are often linked to conformational penalties, associated with loss of freedom by the ligand and the protein.

Our results allowed accurate quantification of energetic penalties associated with the intrinsic (point-mutations) and external (association with another protein) alterations. Hence, the energy required for binding type 2 inhibitor to the wild-type CDK2 was

determined to be -9.6 kcal and the double amino-acid substitution C118L/A144C reduced this cost by 1.0 kcal, resulting in more potent binding as illustrated by differences in binding affinity (K_d) and biochemical data (IC_{50} values). Moreover, for this mutant we estimated that the conformational penalty associated with binding to cyclin B was increased by 1.5 kcal, thereby reducing the strength of the interaction with the inhibitor (Table 3.4).

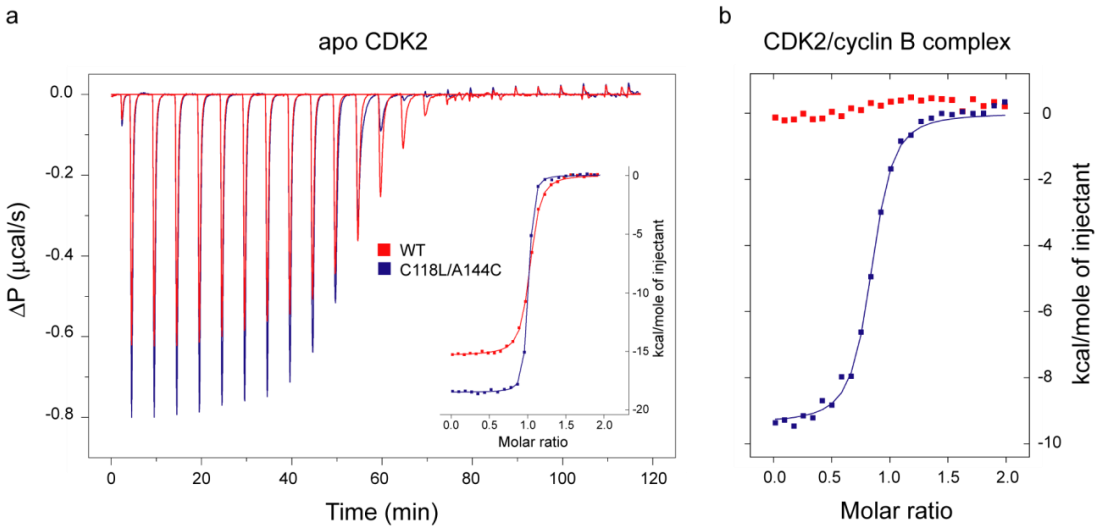


Figure 3.6 Binding of K03861a to the wild-type and to the C118L/A144C mutant as determined by ITC. The exothermic reaction is evident in a row thermogram as heat is released and decreasing heat signals upon each injection indicate saturation of binding. The data were fit to a single site model using chi-square minimization algorithm in Origin 7.0 software revealing the binding isotherms. (a) Titrations of the kinase domain into K03861a indicate potent binding to the inhibitor. (b) Formation of the kinase-cyclin complexes resulted in reduced binding of inhibitor to the C118L/A144C and impaired binding to the wild-type.

Table 3.4 Thermodynamic parameters of K03861a (type 2) inhibitor binding to CDK2 variants as determined by ITC. All errors shown for the ITC are taken from the fit of the titration determined by the Origin software using a single site model with chi-square minimization algorithm.

CDK2 variant	Kd (nM)	ΔHobs (kcal/mol)	TΔS (kcal/mol)	n	ΔG (kcal)	ΔΔG (kcal)
WT	50.0±3.7	-14.9	-5.3	0.96	-9.6	
C118L	18.6±2.9	-15.0	-4.8	0.99	-10.2	0.6[b]
A144C	15.4±2.3	-15.9	-5.6	0.98	-10.3	0.7 [b]
C118L/A144C	9.7±0.8	-18.0	-7.4	0.94	-10.6	1.0 [b]
WT with cyclin B	N/b[a]					
C118L/A144C with cyclin B	134.1±15.8	-9.4	-0.3	0.82	-9.1	1.5 [c]

[a] No binding detected under current assay conditions.

[b] $\Delta\Delta G = \Delta G^{WT} - \Delta G^{mutant}$

[c] $\Delta\Delta G = \Delta G^{with\ cyclin\ B} - \Delta G^{without\ cyclin\ B}$

3.9. Kinetic characterisation confirms long compound residency typical for DFG-out binding mode

While binding affinity describes how well a molecule binds to its target, this measure gives no information about how rapidly a molecule associates and dissociates from its target. When interacting with the ATP-binding pockets of protein kinases, type 2 inhibitors often display slower rates of association and dissociation compared to type 1 inhibitors (Hari et al., 2013) and the mechanism that could explain this is the conformational selection (Chapter 1). The basic assumption is that a protein kinase slowly isomerises between active, inactive and perhaps other states in solution, and inter-conversion between these states is relatively slow. Upon addition of a type 2 inhibitor, CDK2 kinases in the DFG-out conformation bind the compound and are therefore removed from the equilibrium between the free states. The inter-conversion between the conformations is likely to be the rate limiting step and is expected to make association slower. If the interaction is very strong and association rate is limited, extremely slow dissociation may be observed.

Direct monitoring of binding via real-time kinetics measurements was performed on a 16-channel ForteBio Octet RED384 instrument, which is based on BioLayer Interferometry (BLI) technology (Chapter 2). The data generated for the wild-type CDK2 established the slower binding kinetics of DFG-out inhibitors compared to that of type 1 inhibitor flavopiridol (Carlson et al., 1996).

The binding phase of the curves reflects the fact that the off rate is particularly low for the K03429a compound, implying its high potency and type 2 binding mode (Figure 3.7). The determined equilibrium constants (K_d) closely agreed with values determined using ITC (Table 3.4 and Table 3.5).

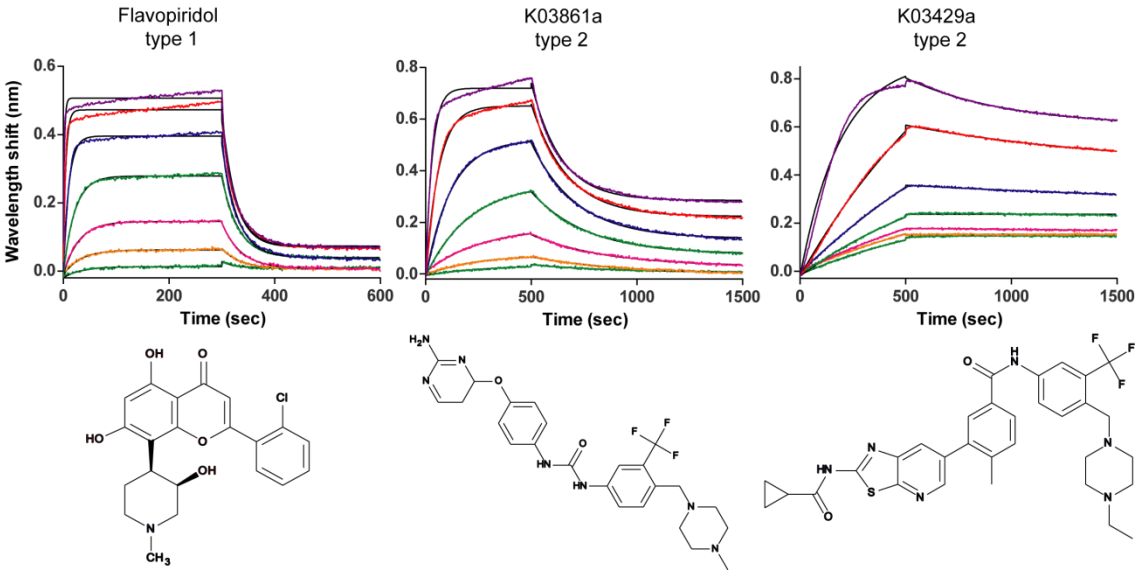


Figure 3.7 Biolayer interferometry real-time kinetic studies of the interaction between immobilised wild type CDK2 and three inhibitors. Dose-response curves are shown in colours and nonlinear least square fitting are shown in black. Type 1 inhibitor flavopiridol demonstrated typical fast on- and off-rates, while type 2 inhibitor K03861a had slower association and dissociation, and another type 2 inhibitor K03429a had very slow dissociation. The corresponding chemical structures are presented below.

Table 3.5 Kinetic interaction of CDK2 with type 1 and type 2 inhibitors. Reported values are representative of a single biolayer interferometry experiment as shown in Figure 3.7.

Parameter	Flavopiridol	K03861a	K03429a
k_{on} ($M^{-1}s^{-1}$)	5.0×10^{4}	2.3×10^{4}	
k_{off} (s^{-1})	3.5×10^{-2}	3.2×10^{-3}	not fitted
K_d (nM)	500	60	

3.10. Structural evidence for a DFG-out conformation of CDK2

After confirming potent binding of the type 2 inhibitors to the wild-type CDK2, it was crucial to identify the mode of binding. Co-crystallisation of the recombinant CDK2 from the *E.coli* expression system was unsuccessful after repeated attempts. This could be due to the fact that the expressed protein was heterogeneous with mass shifts of 178 Da and 258 Da, indicating potential gluconoylation and phosphogluconoylation of CDK2 (Figure 3.8). These common post-translational modifications have been frequently observed in recombinant proteins and the heterogeneity created by gluconoylation is known to interfere with crystallisation of proteins (Kim et al., 2001, Geoghegan et al., 1999, Aon et al., 2008). The additional bands were consistently present in tagged, untagged CDK2 and all the mutants, suggesting that the modification is within the protein sequence itself and not at the N-terminal His-tag as observed previously (Kim et al., 2001). We therefore expressed wild-type CDK2 using the baculovirus strain (De Bondt et al., 1993), and determined the co-crystal structure of CDK2 with type 2 inhibitor K03861a by X-ray crystallography.

The resulting structure provided definitive evidence that CDK2 is able to adopt a DFG-out conformation, where DFG-Phe (F146) undergoes large translocation compared to all previous CDK2 crystal structures which are in the DFG-in conformation (Figure 3.9). The compound K03168a occupies the adenine-binding pocket of the ATP-binding site and makes two H-bonds within the hinge region: one between the pyramidine-N and backbone NH of L83, and another between the amino group of the aminopyrimidine and the backbone carbonyl of L83. The carboxyl side-chain of the conserved αC-helix glutamate (E51) is engaged in at least one H-bond interaction with the NH from urea, while maintaining its usual salt bridge with catalytic leucine K33. The carbonyl from the urea linker also accepts the H-bond from the backbone NH of DFG-Asp (D145). The compound K03861a has a phenyl ring which is substituted with a trifluoromethyl group at the meta position. This typical type 2 scaffold motif projects into the hydrophobic

pocket vacated by the rotation of the DFG-Phe side chain around the C-N bond of the DFG-Asp into the ATP-binding pocket. There was a clear break in the density of the activation loop between residues 151-164 indicating that the activation loop is disordered as it is often observed in inactive kinase structures.

Surprisingly, a second K03861a molecule was found outside the ATP-binding pocket, packed between the C-lobe as shown and αC-helix of the symmetry related molecule. Very similar examples with other protein kinases have been reported previously (Eswaran et al., 2008). Because the ITC data indicated 1:1 stoichiometry of binding, we concluded that the K03861a molecule located outside the active site has no biological significance, and only serves to stabilise crystal contacts at the interface between symmetry-related molecules.

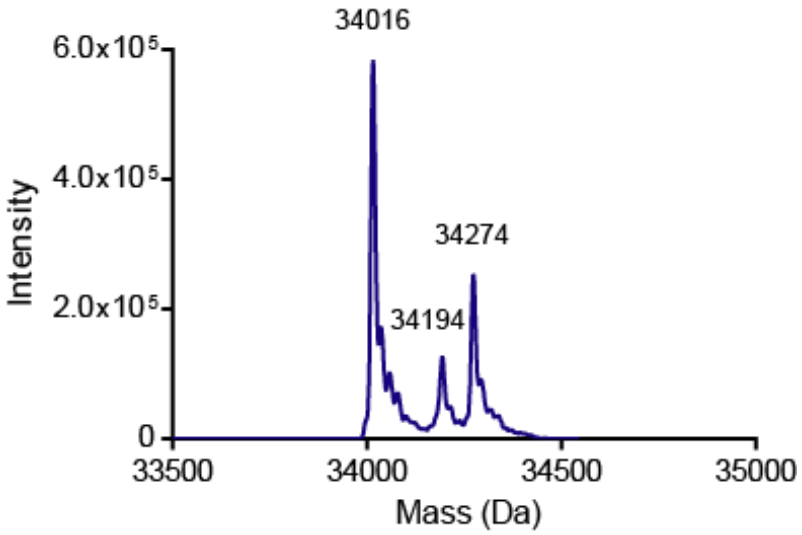


Figure 3.8 Mass spectrometry analysis of cleaved wild-type CDK2 from *E.coli* expression system. ESI-TOF spectrum showing molecular ions corresponding to the expected mass of untagged CDK2 (34016 Da), 34194 Da and 34274 Da, which are 178 Da and 258 Da larger than the expected mass and could be gluconoylated and phosphogluconoylated derivatives.

Table 3.6 Data collection, processing and refinement statistics for the CDK2 complex with K03861a inhibitor.

Data collection	
Space group	$P2_12_12_1$
Unit cell [a b c (Å)]	64.41 67.27 89.13
Resolution range[a] (Å)	67.27-2.00 (2.11-2.00)
No. of unique observations[a]	26499 (3591)
⟨I/σ(I)⟩[a]	10.8 (1.8)
CC(1/2)	0.998 (0.742)
Completeness[a] (%)	98.8 (93.4)
Multiplicity[a]	3.7 (2.6)
Wilson B factor (Å)	35.6
*R*merge[a] (%)	6.3 (52.9)
*R*pim[a] (%)	3.7 (37.5)
Refinement	
R_{work} (R_{free}) [b] (%)	20.20 (24.76)
RMSD. bond length (Å)	0.0087
RMSD. bond angle (°)	1.3278
Number of atoms	2413
Number of residues	295
Water molecules	135
Compound	36
Ramachandran	
Preferred (%)	97.43
Allowed (%)	2.57
Disallowed (%)	0.0

[a] Values in parentheses show the statistics for the highest resolution shell.

[b] Rfree was calculated using 5% of the diffraction data assigned randomly and not used throughout refinement.

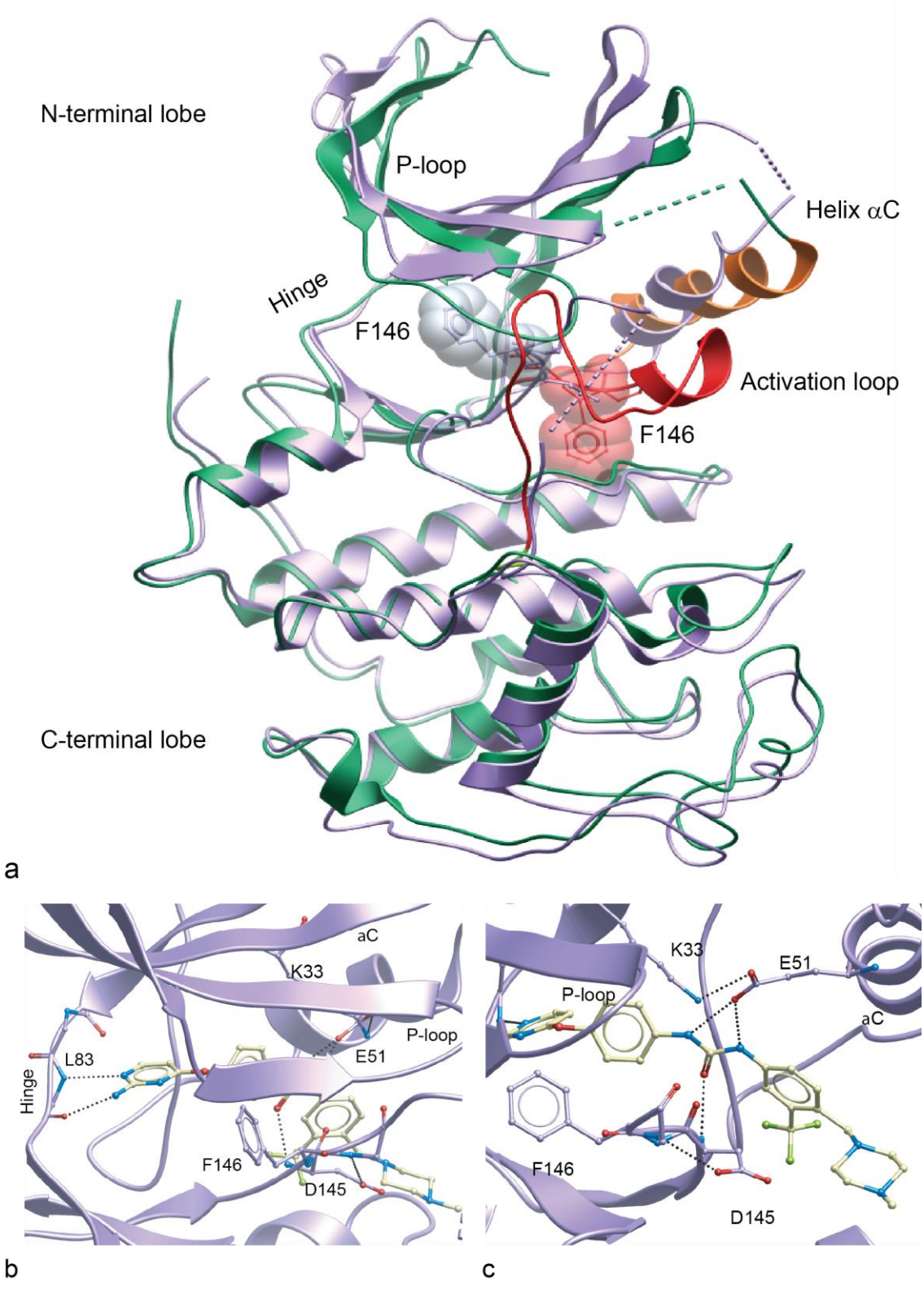


Figure 3.9 Crystal structure of CDK2 in complex with K03861a determined by X-ray crystallography confirms that inhibitor binding occurs to the DFG-out conformation of the kinase. (a) An overlay of the DFG-in (green, PDB: 1b38) and DFG-out (purple) CDK2 structures. Large translocation of the DFG-Phe (F146, highlighted) is revealed, while the activation loop of the DFG-in structure is coloured in red and the αC helix is in orange. (b) Two typical hydrogen bonds formed with the CDK2 hinge. (c) Hydrogen bond network formed between urea linker, catalytic lysine (K33) and αC-Glu (E51). The figures were prepared using ICM (Molsoft, LLC).

3.11. Discussion and conclusions

This project tested the hypothesis that primary sequence alteration of a protein kinase can shift the conformational equilibrium between its active and inactive states. CDK2 was chosen as a target that was thought to have inaccessible DFG-out conformation due to the large conformational penalty between the 'in' and 'out' states. Residues suggested by structural comparisons were mutated to promote this transition. Biochemical and biophysical assays were used to investigate if the mutations affected kinase's sensitivity to small molecule inhibitors, including type 2 compounds that bind to the DFG-out conformation. A number of type 2 inhibitors were identified as hits for the mutants, however, binding of type 2 inhibitors to the wild-type CDK2 was also detected. The detailed characterisation revealed that type 2 compounds were potent inhibitors of the monomeric CDK2, but not the CDK2 complex with cyclin. The co-crystal structure of CDK2 with type 2 inhibitor K03861a confirmed the DFG-out binding mode.

Additionally, we demonstrated that the large conformational changes required to accommodate type 2 inhibitor binding were responsible for slow binding kinetics. In particular this conformational adaptation plays a key role in stabilising the compound-target complex, thereby increasing the compound residence time. While effectiveness of an inhibitor is traditionally quantified using K_d or IC_{50} values, which are related to the concentration of inhibitor needed for a certain level of target occupancy under equilibrium conditions, inhibitor-target binding *in vivo* is not wholly governed by equilibrium conditions. Hence, the prolonged compound residency time could be important for its pharmacological selectivity and efficacy, and compounds with a very slow dissociation could have a significant clinical advantage over rapidly reversible inhibitors since they inhibit the activity of a kinase for a significant time period.

The gatekeeper and the residue preceding the DFG motif were already pointed out as residues that influence the ability of a kinase to adopt a DFG-out conformation (Hari et

al., 2013). The authors demonstrated that mutating leucine to cysteine at the DFG-1 position prevents Erk2 from binding to the indicator probe with a type 2 scaffold, implying that the introduced cysteine disrupts the DFG-out conformation. This contradicts our results, which suggest that cysteine at this position actually assists in stabilisation of the DFG-out state. Moreover, a number of protein kinases that bind type 2 inhibitors including CDK2 and FLT3 have a cysteine at this position, suggesting that the authors' predictions could be limited to the specific case of the studied model Erk2.

However, we demonstrated that similarly to disease-related mutations, amino-acid substitutions can exert a dramatic effect by altering the ability of a kinase to dynamically switch between near-native conformations. Mutations that are located within the active site of CDK2 (A144C), as well as those in the outer interaction shell (C118L), show a synergistic effect on both catalytic activity of the CDK2 and its ability to bind type 2 inhibitors.

We determined free energies of binding associated with these interactions and quantified energetic penalties associated with the intrinsic (point-mutations) and external (association with another protein) modifications, providing a unique context for interpretation of conformational ensembles and how shifts in the distributions of conformations are linked to the complex thermodynamic process of ligand binding.

These findings support the assumption that a kinase can adopt all major conformations at an energy cost, and further strengthen the concept of a dynamic conformational equilibrium. Essentially, the differences in affinities between the authentic CDK2 and the C118L/A144C mutant imply differences in population of conformers coupled with their inter-conversion rates, and the availability of a DFG-out conformation is the determining factor for the type 2 inhibition. In the absence of cyclin, the C188L/A144C mutant can more readily form the DFG-out conformation than the wild-type, which is reflected by more potent binding of the K03861a inhibitor.

Upon addition of cyclin, the equilibrium shifts dramatically towards the active DFG-in conformation and therefore binding of this inhibitor to the wild-type could no longer be

detected under the same experimental conditions. The C118L/A144C mutant was still binding type 2 inhibitors, although affinities of these interactions were significantly reduced compared to those without cyclin. Arguably the introduced point-mutations have only a minor effect, if any, on the overall fold of the kinase or a particular conformational state. The basic question remains to be answered regarding the structural alterations caused by the mutations: is the mutant's ability to bind cyclin affected or is a ternary complex with the inhibitor formed? One way or another, this work highlights the importance of conformational adaptation in the evaluation of drug-target interactions. Failure to properly accommodate this can mislead screening campaigns and not only underestimate the affinity and residency time of a compound, but overlook the fact of binding. The inhibition of CDK2 by type 2 inhibitors is an excellent example of this.

3.12. Future work

Our studies illustrate that residues contributing to kinase conformational equilibria and protein-protein interactions can be just as important for inhibitor binding as contact residues formed in the ligand complex. The traditional view that contacts residues alone define inhibitor potency and selectivity should therefore be updated.

An important question to address is how type 2 inhibition of CDK2 can be tested in cellular systems. Since cyclin expression is transient, these data may improve the understanding of cell cycle regulation and assist in therapeutic research.

Regarding a dynamic interplay between compound potency and conformational dynamics of its target, it seems reasonable to suggest that more potent type 2 inhibitors like K03429a and K03860a bind to the wild-type CDK2 even if cyclin is

present. We therefore plan to perform additional ITC experiments to test this hypothesis.

Even though it seems clear that the introduced point-mutations have only a minor effect on the binding mode of type 2 inhibitors, obtaining co-crystal structures of mutants would be useful to support this statement.

An additional way to quantify conformational equilibrium shifts could be evaluation of binding affinities between cyclin and CDK2 variants. Our attempts to establish these K_d values using ITC or BLI were unsuccessful, and tentative microscale thermophoresis (MST) experiments did not detect differences in affinities for cyclin A between the wild type, the F80T and C188L/A144C mutants. However, SPR experiments could be useful in orthogonal characterisation of cyclin binding.

Another area for future exploration is explaining the reasons of exclusive binding of the type 2 compound K03859a to the gatekeeper mutant F80T. The typical drug-resistant mutations described in patients are ‘reverse’ substitutions of gatekeeper threonine to a bulkier residue, such as isoleucine (Gorre et al., 2001), or phenylalanine (Warmuth et al., 2003). The detailed understanding of this interesting interaction could therefore provide useful insight into involvement of this mutation in the resistance mechanism and to more fully exploit the role of polar residues in the active site of protein kinases.

Moreover, our discovery of the fact that CDK2 was able to adopt a DFG-out may trigger medicinal chemistry efforts. There is a strong potential for design and development of a CDK2 inhibitor that targets the cysteine (C118) located next to the side pocket (Figure 3.1). Since the cysteine at this position is uncommon in protein kinases, this inhibitor could have an excellent selectivity profile.

As additional work in the field of protein kinases continues, further descriptions of the structural determinants of conformational preferences provide useful perspectives into activation and inhibition of these important drug targets.

4. Isoform selectivity in PIM kinases

PIM kinases are promising drug targets in the treatment of cancer (Chapter 1). Despite highly conserved active sites, most reported compounds inhibit PIM1 and PIM3 more strongly by a factor of at least ten (Bullock et al., 2005b), which has been suggested to be related to differences in dynamic properties of the different PIM isoforms (Brault et al., 2010). Additionally, PIM2 has considerably higher affinity for ATP than PIM1 and PIM3, which could make cellular inhibition of this protein kinase much more challenging (Burger et al., 2013).

To elucidate the reasons for this differential compound binding behaviour, we performed bioinformatic analysis of PIM sequences and structures, coupled with experiments involving site-directed mutagenesis, screening of representative chemical libraries, and X-ray crystallography.

4.1. Bioinformatic analysis of sequences and conformations of PIM kinases

4.1.1. Evolution of the PIM protein kinase family

To compare PIM protein kinases and to gain insights into their evolution, we used principles of the evolutionary trace analysis. This approach relies on the extraction of functionally important residues from sequence conservation patterns in homologous proteins (Lichtarge et al., 1996) and could assist in determination of mutational effects on biological function and ligand binding.

The following human PIM protein sequences, obtained from the UniProt database (Bairoch et al., 2005), were used as queries: PIM1|P11309, PIM2|Q9P1W9 and PIM3|Q86V86 (Appendix C). A PSI-BLAST (Altschul et al., 1997) search was then performed against the non-redundant protein database downloaded from the NCBI

website (ftp://ftp.ncbi.nih.gov/blast/db) on 17 July 2010, with default settings and the maximum Expect value (E) set to 1. In order to obtain the most comprehensive hit set we performed three searches using each query sequence individually and afterwards combined the results, yielding 3167 unique sequences. Only 173 sequences met the criteria of being relevant PIM homologues and were retained for further analysis (Table 4.1).

Sequences of all accepted hits in FASTA format were obtained from the NCBI protein sequence database after which a multiple sequence alignment of PIM homologues was performed using MUSCLE (Edgar, 2004). Since the accuracy of phylogenetic analysis is directly influenced by the alignment quality, the results were visualised in Jalview (Waterhouse et al., 2009) and any irrelevant sequences were excluded. We applied a maximum likelihood tree building algorithm using the *proml* program of the Phylip software package (Felsenstein, 1989) and visualised the generated tree using iTOL (Letunic and Bork, 2007), with GenBank IDs and species names as primary keys of the sequences identification (Appendix D). The vast majority of identified hits belonged to two organisms, *Danio rerio* (zebrafish) and *Taeniopygia guttata* (zebra finch). Recently, Kong and co-workers (Kong et al.) investigated the reasons for an unusual expansion of PIM kinases in *Taeniopygia guttata* and concluded that these duplications occurred as the result of mutational bias. Based on that conclusion, we conjecture that the same statement may be true for *Danio rerio*. Therefore we excluded PIM genes from these two species, which also clustered into separate branches, from our analysis as they are probably irrelevant to the selectivity problem. The excluded clades were collapsed, and amongst the remaining 52 examined sequences, 10 formed a PIM1-like family that was closely related to 19 PIM3-like sequences, while 12 PIM2-like genes clustered separately and 11 sequences did not belong to any of the above (Figure 4.1). The hinge region appeared to be highly conserved across the species with only two changes from PIM1 and PIM3 to PIM2: E124 to leucine (hereafter PIM1 numbering) and V126 to alanine (Figure 4.2).

After the examination of a region containing the P-loop (residues 38-77), sequence logos for each subgroup were constructed using WebLogo (Crooks et al., 2004), and a number of conserved changes between amino acids with very different chemical or structural properties were identified (Figure 3.1a). We focused our attention on the most interesting and unexpected, but conserved across each clade amino-acid substitutions, and those located on the surface of PIM1 protein were nominated as the key substitutions that could play a role in ligand binding (Figure 3.1b).

Table 4.1. Sequence homology analysis of PIM protein kinases

Filtering criterion	Justification	Software tool
Sequence length < 200 residues	To discard sequence fragments since the catalytic domain of PIM proteins is 280 amino acids (Kumar et al., 2005, Selten et al., 1986)	Custom Perl script
Sequences with identity over 99% (redundant hits)	To eliminate very close homologues and duplicates that might confound results	CD-Hit (Li and Godzik, 2006). Human PIM proteins kept as representatives of their respective clusters
Sequences without PIM-specific hinge pattern [ML]ER**P**....D[LM].*DFG†	Literature review and manual inspection of multiple sequence alignments	Regular expressions in Perl script

† Sequence pattern was refined after manual inspection ensuring that true hits are not rejected, while every accepted sequence contains the DFG motif found in any functional kinase and PIM-characteristic hinge insertion along with proline residue (in bold) as described in the Introduction, (.) matches any residue, (.*) matches any residue zero or more times.

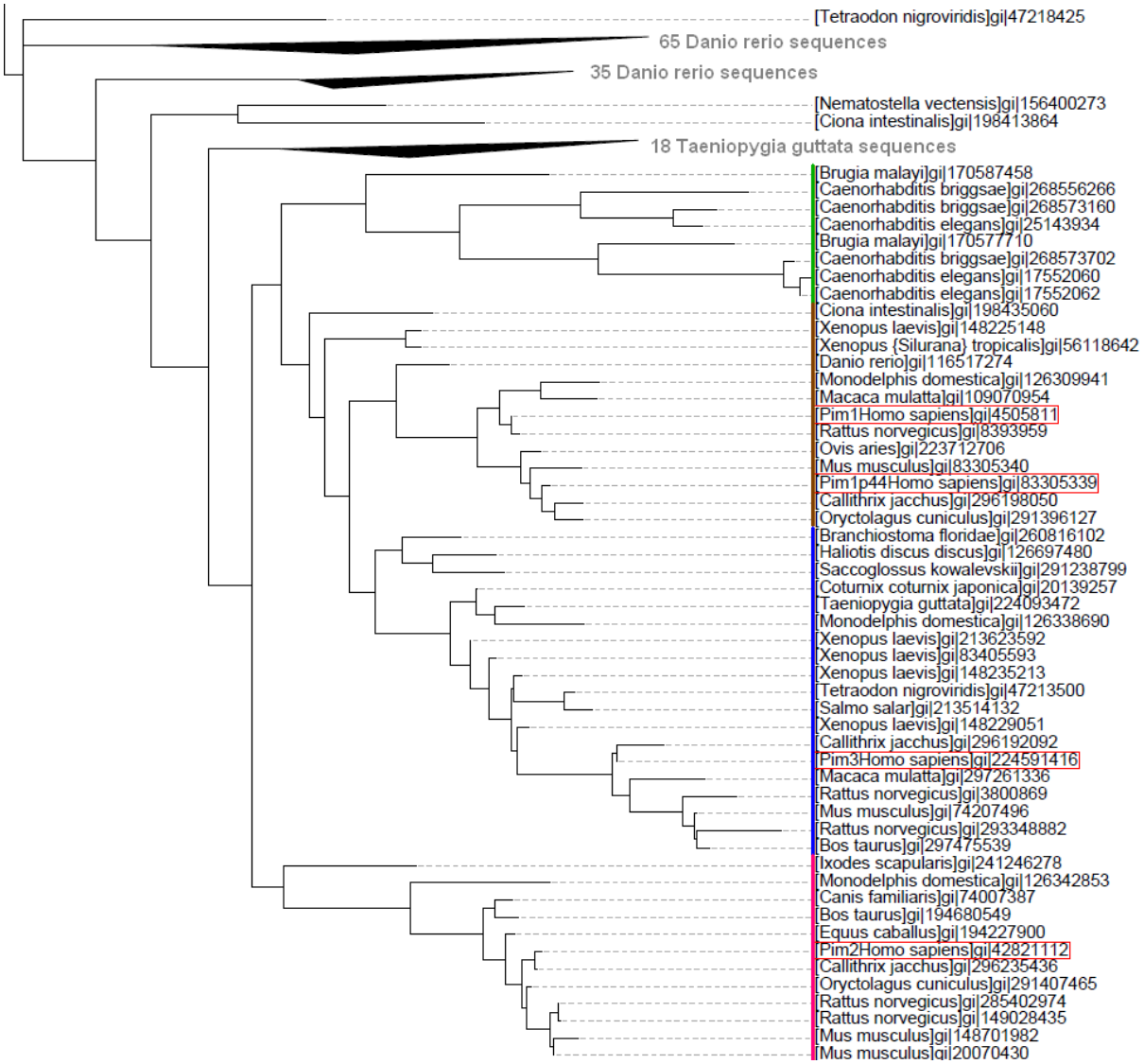


Figure 4.1. Reduced phylogenetic tree, black triangles indicate collapsed branches. PIM-like families are marked with colour lines and human PIM proteins are highlighted with red rectangles

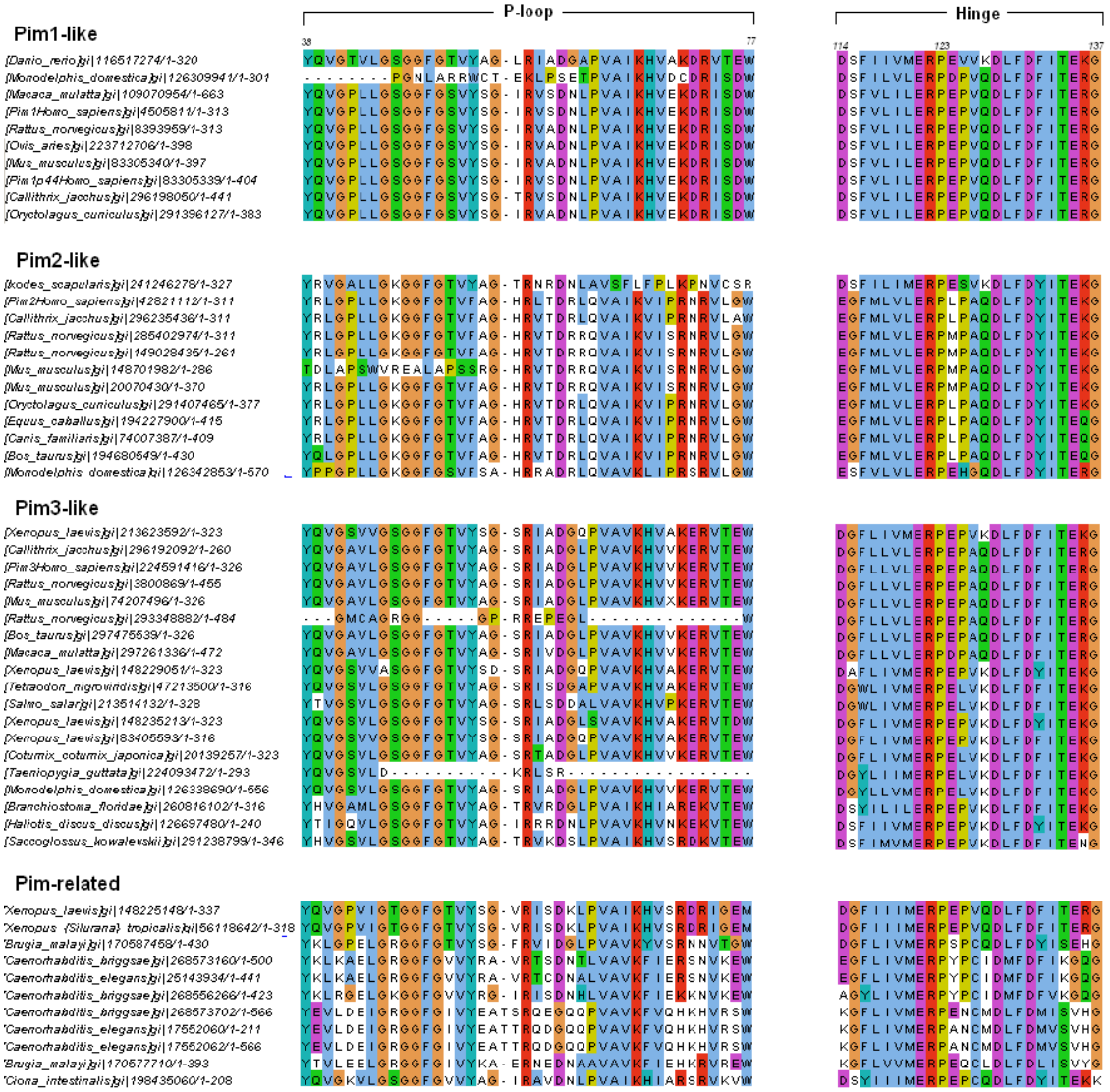
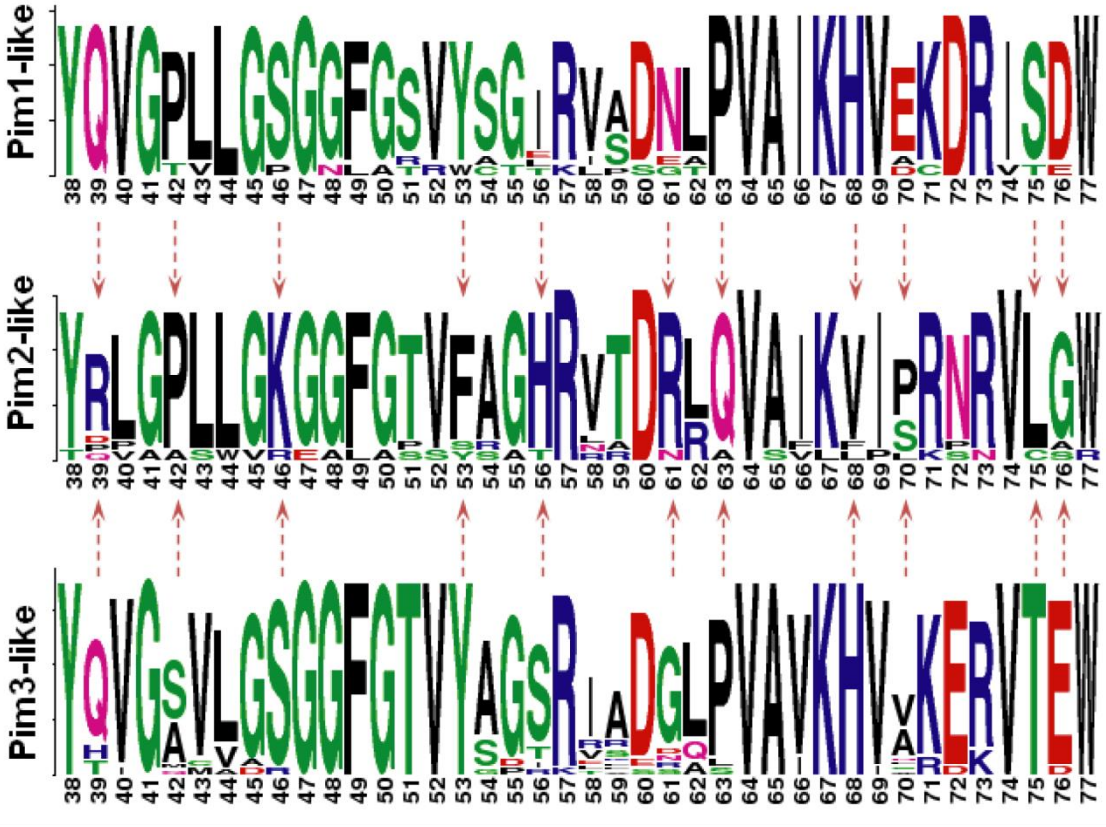
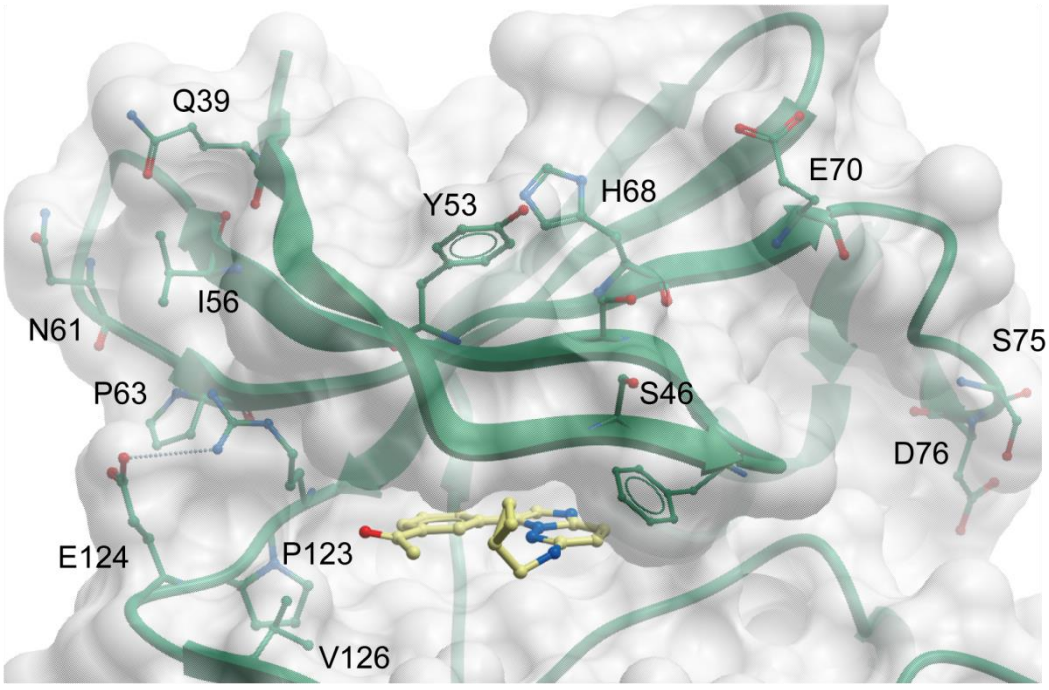


Figure 4.2. Multiple sequence alignment of P-loop and hinge regions of PIM homologues. Conserved residues and their positions are coloured with the ClustalX scheme using Jalview.



a



b

Figure 4.3 An evolutionary analysis of PIM kinases identified protein functional determinants. (a) Sequence logos of P-loop region from phylogenetic subtrees. The key conserved changes are marked with arrows. (b) Residues mapped on the PIM1 surface, PDB: 2c3i.

4.1.2. Structural analysis of PIM1 kinase

In order to study P-loop residue interaction patterns, we investigated the set of all known X-ray structures of PIM1. At the time of the analysis, 41 PIM1 kinase structures had been deposited in the PDB, consisting of two apo-forms and 39 liganded structures. As in our sequence study, we focused our attention on the interactions of protein residues within the P-loop region. Based on the observed configurations of the P-loop, we subdivided the representative PIM1 structures into four subclasses (Table 4.2).

Table 4.2. Classification of the selected PIM1 structures

No	Subclass	Ligand	PDB codes and references
1	Apo-enzymes	None	1xqz (Qian et al., 2005) 1ywv (Kumar et al., 2005)
2	Complexes with ATP analogues	AMP-PNP	1xr1 (Qian et al., 2005) 1yxt (Kumar et al., 2005) 2bzk (partial P-loop) (Bullock et al., 2005b)
3	Co-structures with PIM1-specific inhibitors	Imidazo-Pyridazine Benzothienopyrimidone (6e) Benzothienopyrimidone (3b)	2c3i (Pogacic et al., 2007) 3jxw (Tao et al., 2009) 3jya (Tao et al., 2009)
4	Co-structures with dual inhibitors	Benzothienopyrimidone (12b) Naphto-difuran	3ju0 (Tao et al., 2009) 3ma3 (partial P-loop) (Lopez-Ramos et al., 2010)

A custom MATLAB 7.10 (The MathWorks Inc., Natick, MA) script was developed to transform protein structures into residue interaction networks that can be visualised as a graph. In our network each node corresponds to an amino acid, while edges represent residue interactions. Two residues were defined as ‘interacting’ if any atom of the first residue is located within 4.5 Å of any atom of the second residue (Hamer et al., 2010). The trivial intra- amino-acid contacts were eliminated and the observed multiple edges indicate how strongly a particular residue interacts with its counterparts.

Similarly, we indicated protein-ligand contacts, accounting for presence or absence of interactions between the P-loop residues and the PIM inhibitor reported in co-crystal structures. Furthermore, for every given residue, we analysed the composition of surrounding residues in terms of their location at the sequence level. We defined short-range interactions as those between residues which are within three sequence positions, whereas long-range interactions are those between residues further away from one other. Only long-range interactions were considered in the analysis and the matrices of different PIM1 structures were compared to establish the common interaction patterns as well as key differences in residue contacts (Figure 4.4a). A sanity check was conducted to ensure that the maps support the protein models.

Given the initial observation of the P-loop fluctuations, we established two major conformations (Figure 4.4b). Complexes with PIM1-specific inhibitors (2c3i and 3jya) shared the same P-loop orientation as the apo-enzyme (1xqz) and we called it ‘conformation A’. The structural comparison of PIM1 structures with the ATP-mimetic inhibitor AMP-PNP was intriguing since the P-loop orientation in 1yxt was in conformation A, which is considerably different from the analogous 1xr1. Moreover, the P-loop has not been fully resolved in the 2bzk structure, but the surrounding modelled residues perfectly aligned with 1xr1. Remarkably, PIM1 complexes with dual inhibitors 3jy0 and 3ma3 have the same P-loop orientation as 1xr1 and 2bzk. These four structures formed a group which we termed ‘conformation B’.

From a pairwise comparison analysis of the contact maps derived from these structures, we deduced that residue interactions 63-56, 73-68 and 73-69 were present in conformation A, but not in B. Correspondingly, residue pairs 46-50 and 46-51 were interacting in each structure having a P-loop in conformation B, but not identified in those with conformation A.

The pair of residues 56-63 was investigated further, since they were also identified in sequence analysis. The partially conserved position 56 is isoleucine in PIM1 and serine in PIM3 kinases, however, PIM2 has histidine at this site. In addition, a

conserved proline found in PIM1 and PIM3 at position 63 is substituted by polar glutamine in PIM2. If PIM2 is able to adopt conformation B and this glutamine-histidine pair forms a hydrogen bond, the P-loop movement could be restricted resulting in weaker binding of the inhibitors (Figure 3.1b). The full set of results of this structural analysis is provided in the Appendix F.

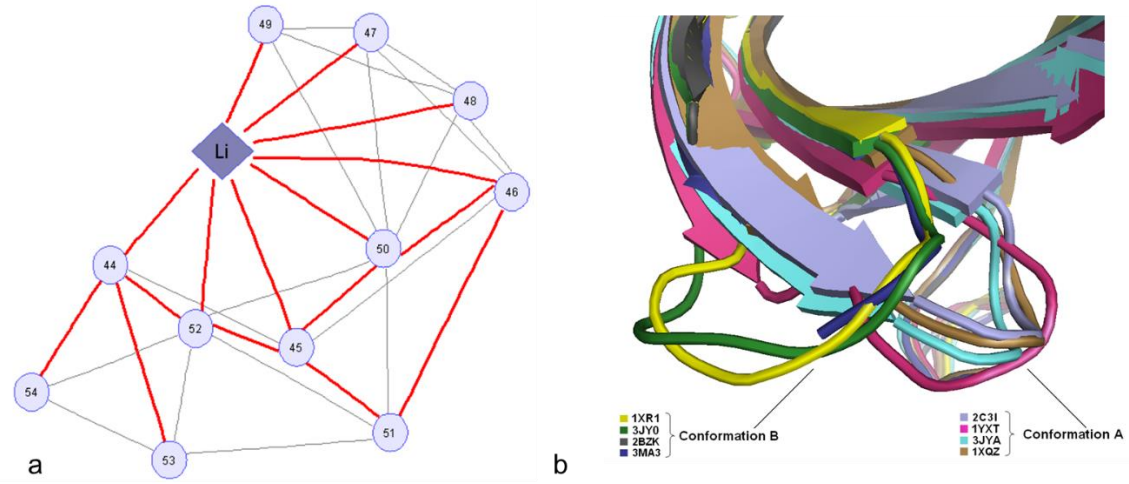
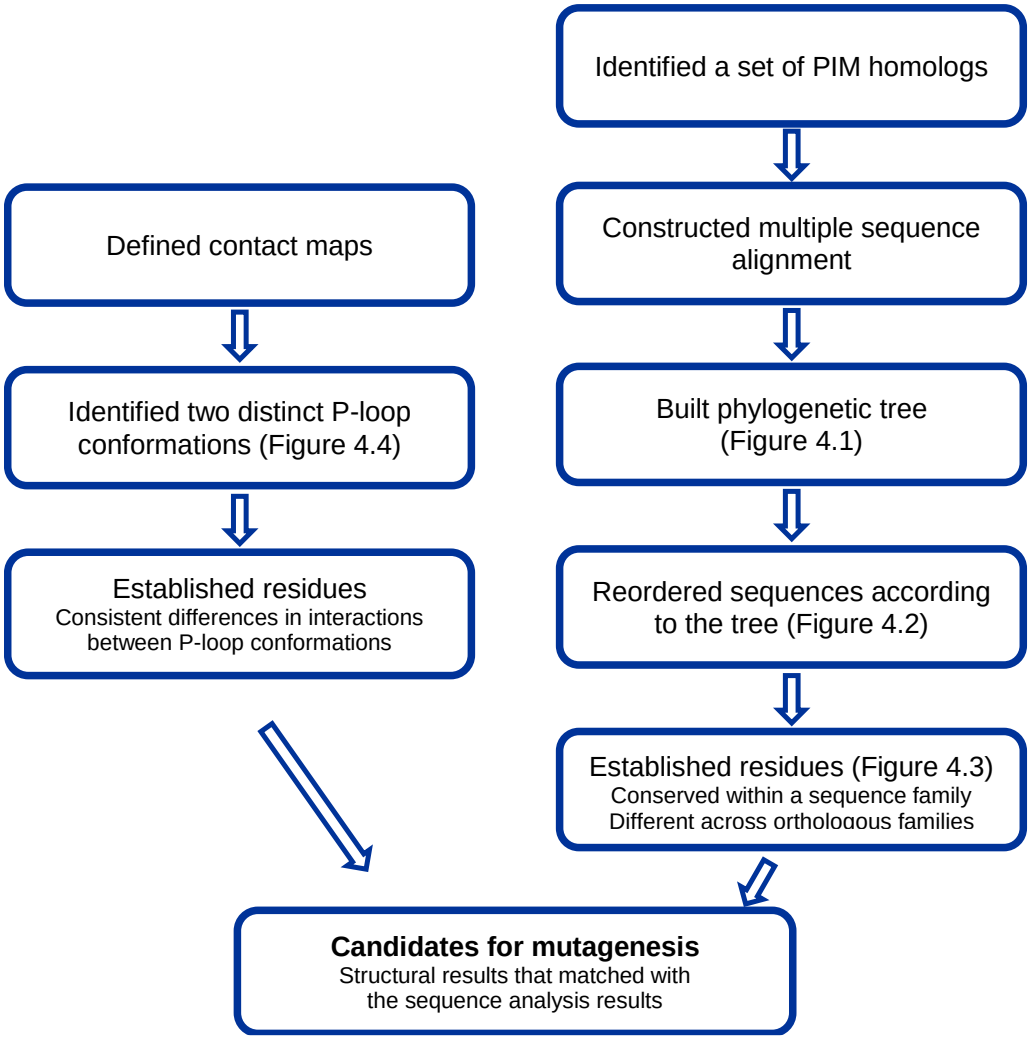


Figure 4.4. Structural analysis of PIM1 kinases. (a) Contact map visualisation with nodes representing amino acids (numbered according to a sequence) or ligand (Li), and edges representing residue interactions. Long-range interactions, i.e. between residues that are further than three sequence positions from each other, are coloured in red. (b) Classification of PIM1 protein structures described in Table 4.2. Conformations A and B are defined depending on the P-loop orientation. 2bzk and 3ma3 are classified as conformation B, despite their incomplete P-loop (the decision was based on the orientation of surrounding residues).

4.1.3. A combination of sequence and structural studies provided the basis for rational mutagenesis

Our sequence conservation analysis revealed important clues about the protein. We collected all available PIM homologues from different species and created a multiple sequence alignment of these sequences, which was then subjected to phylogenetic analysis (Figure 4.5). Once the phylogenetic tree was constructed, comparative studies of the orthologous clades were performed. As a result, residues that are highly

conserved within subtrees but differ among orthologous groups (PIM1 and PIM3 versus PIM2) were identified. In order to improve our understanding of PIM kinase selectivity we also considered the available structural data. Using a MATLAB script we identified interacting residue pairs in PIM1 co-structures with selective ligands, which are not in contact in PIM1 complexes with dual inhibitors. A number of conserved substitutions, S46K, S51T, I56H, P63Q, H68V, E124L and V126A were identified in both sequence and structural studies. Therefore we conclude that these substitutions are likely to be the PIM ligand specificity determinants and are the best candidates for site-directed mutagenesis.



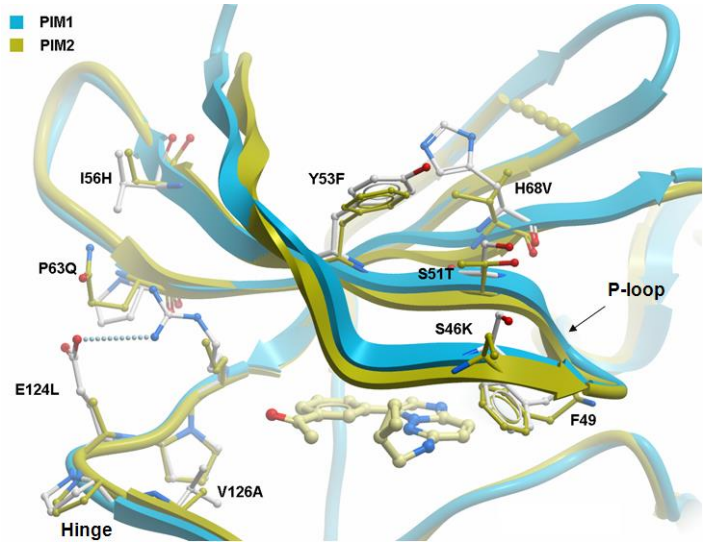


Figure 4.5 Flow diagram showing the analysis of PIM structures (left) and sequences (right). Below is an overlay of PIM1 and PIM2 structures (P-loop and hinge regions) with point mutations shown as sticks and labelled, PDB: 2c3i and 2iwi. The P-loop folds into the ATP site making aromatic contacts with the conserved residue F49 and the inhibitor.

4.2. New PIM2 construct – a truncated variant with improved crystallisation properties

As stated in the introduction, a crystal structure of the PIM3 kinase is not yet known. Moreover, only one PIM2 structure is available publicly (Bullock et al., 2009) and further SGC attempts to co-crystallise PIM2 with inhibitors have been unsuccessful. It has also been noted that PIM2 is less thermally stable and has a much shorter half-life than PIM1. To find reasons for the inability of PIM2 and PIM3 to crystallise, and whether the differences in molecular recognition are guided by structural variability, flexibility or disorder-to-order transitions, we evaluated PIM proteins using the regional order neural network server (RONN) (Yang et al., 2005). This method takes protein sequence as an input and detects natively disordered regions using a neural network. The output is a graph showing the probability of disorder versus residue position in the protein sequence (Figure 4.6). The results for PIM1 and PIM3 prediction show little disorder, whereas two PIM2 segments, one at the beginning of the sequence (residues 1-30) and the second towards the end (residues 265-290) show 60-90% probability of being disordered.

The comparison of all three protein disorder graphs provided no evidence that disorder-to-order transitions alter molecular recognition and selectivity profiles of PIM kinases. A small disordered segment (residues 265-280) shared by both PIM2 and PIM3 is unlikely to be the cause of PIM3 instability. However, an additional large disordered region in the PIM2 kinase might explain the poor tendency of PIM2 to crystallise. Coupled with the fact that the first 22 residues in the available PIM2 structure were not defined by electron density, we designed a new PIM2 construct with truncated N-terminal fragment (residues 1-22). The truncated PIM2 protein was then successfully crystallised.

There was no apparent explanation for failure of PIM3 to crystallise, implying that other factors apart from disorder contribute to this protein’s crystallisation properties.

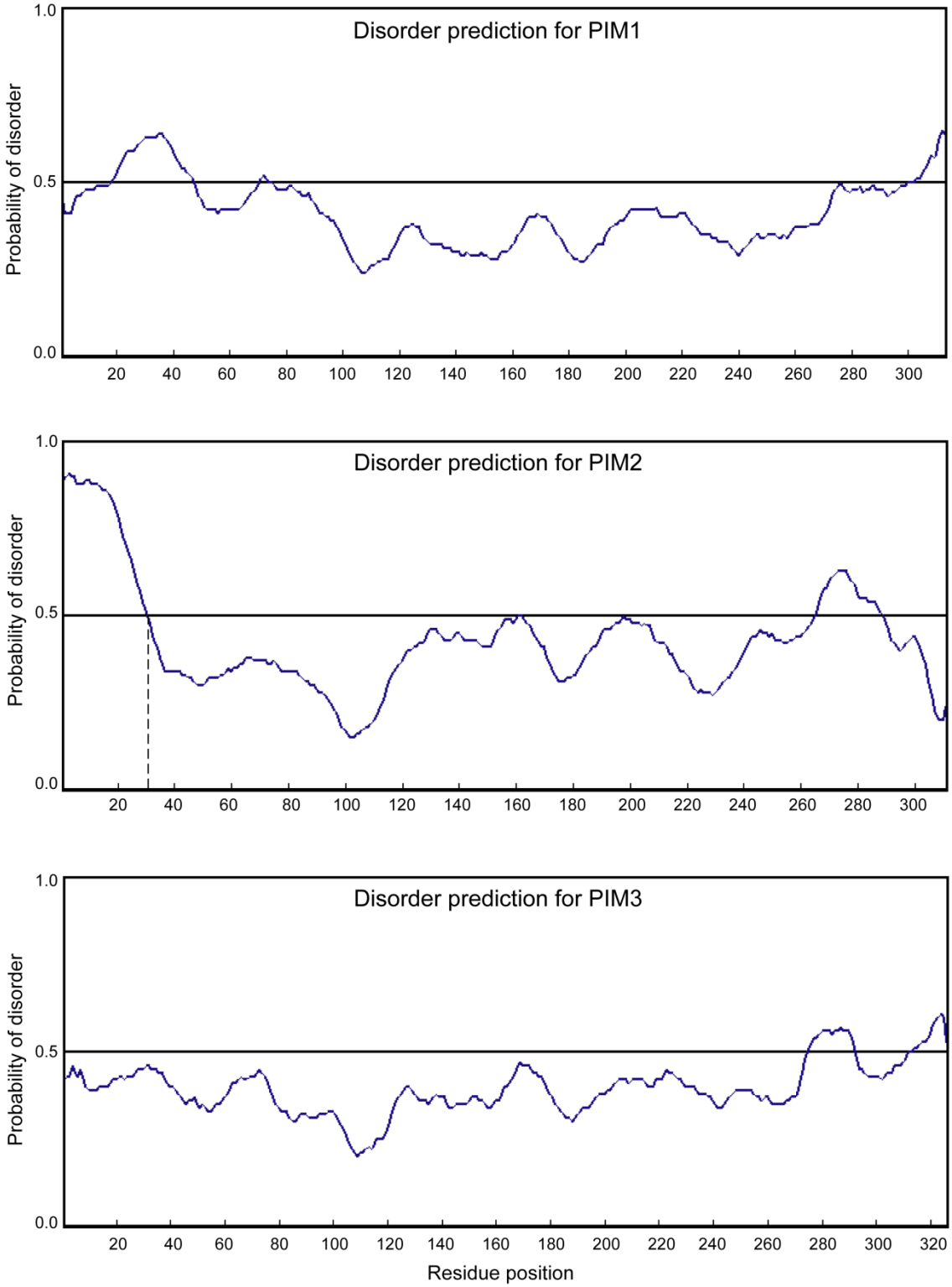


Figure 4.6. RONN plots of disorder probability per residue for PIM kinases. The horizontal lines mark the threshold for disorder predictions (corresponding to probability of 0.5). While PIM1 and PIM3 proteins are predicted to have little disorder, the segment at the beginning of the PIM2 sequence (residues 1-30, marked with vertical dashed line) shows the high probability of disorder.

4.3. Thermal stability and structural integrity of PIM variants

The role played by the residues highlighted in our analysis (Section 4.1) was further investigated by site-directed mutagenesis, along with the effect of two cumulative mutations (Table 4.3). DSF experiments were conducted to compare PIM isoforms/mutants to each other and identify potent inhibitors for further studies. As explained in the Methods chapter, the DSF technique involves the progressive heating of a protein, while a dye fluoresces upon contact with the exposed hydrophobic core of the unfolded protein. With the rise of temperature, the protein unfolds inducing increase in fluorescence levels and therefore allowing detection of the protein melting temperature (T_m). Since potent inhibitors stabilise the protein in its folded conformation thereby increasing T_m, shifts in melting temperature serve as indicators of a compound-protein binding strength.

As a prerequisite for this measurement, the melting temperature of each variant without compounds added was established (Table 4.3). PIM2 displayed a lower melting temperature than PIM1 and PIM3 suggesting that this enzyme is less thermally stable. The introduction of point mutations did not significantly affect the structural integrity of the kinase since all mutants had melting temperatures similar to those of wild-type PIM kinases. The biggest effect was observed in E124L and I56H/P63Q mutants, for which T_m was reduced by up to 5 °C. In the 124L mutant this effect can be explained by deletion of the salt bridge with R122 (Figure 4.5) and the observed T_m value was established to be similar to that of PIM2.

Table 4.3 Melting temperatures (T_m) obtained for the PIM panel using DSF show that PIM2 is less thermally stable than PIM and PIM3. Most PIM1 mutations have no strong effect on PIM1 stability. Mean values and standard error were derived from at least 40 independent experiments.

Protein	T_m (°C)
PIM1	46.3 ± 1.0
S46K/S51T	45.9 ± 1.2
E124L	42.9 ± 1.0
S46K	47.2 ± 1.3
V126A	48.0 ± 0.9
S51T	46.7 ± 1.7
H68V	46.7 ± 0.3
I56H/P63Q	42.1 ± 0.5
PIM2	41.8 ± 1.2
PIM3	44.6 ± 0.4

4.4. Thermal shift assay allows nomination of potent inhibitors for quantitative analysis

Thermal stability shift assay, which identifies binding compounds by a shift in the protein's melting temperature, T_m, was applied to rapidly screen a collection of available kinase inhibitors. About 1000 compounds were screened, and a selection of potent inhibitors demonstrating differences in binding to PIM kinases and the mutants is presented as a colour-coded map (Table 4.4).

Changes in melting temperatures (ΔT_m) were calculated as the differences between the melting temperatures of the protein in the presence of ligand compared with the protein alone. Once ΔT_m values for all proteins were computed, the data were sorted from smallest to largest ΔT_m for PIM1. Then the compounds were subdivided in four groups according to their scaffolds: imidazopyridazines, carbolines, furanes (with

related di-substitutions) and novel kinase inhibitors synthesised by GlaxoSmithKline (GSK) in 2011 (Table 4.4).

The majority of the compounds, in particular imidazopyridazines (e.g. K00135) and carbolines (e.g. K01844a), show 2-3 °C difference in ΔT_m, when binding to PIM2, which corresponds to 20-100 fold selectivity of these ligands towards the PIM1 kinase (Pogacic et al., 2007). Also, a number of ligands from the furane class have weaker affinity for the P-loop mutants, including the compound K00906, which completely failed to bind to those with single S46K and S51T substitutions. The double mutant S46/S51 shows a similar effect, but to a lesser extent.

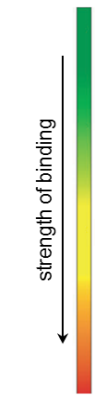
The ΔT_m values of mutants V126A and I56H/P63Q were reduced compared to PIM1 for most of the presented compounds, accounting for most of the observed effect in melting temperature between PIM1 and PIM2. Other mutations did not show substantial reduction in binding affinities, given that their ΔT_m values are close to those of PIM1.

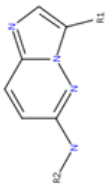
Some GSK compounds (e.g. K02803, K02908) selectively bind to PIM2, which has not been reported previously. Although the mutants did not replicate this type of 'reverse' behaviour, the compounds were selected for further investigation to see if this effect could be explained and documented.

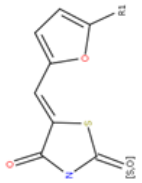
Overall, the thermal studies confirmed conserved protein fold and pocket shape for each mutant. After comparison of ΔT_m values across the PIM variants we selected a range of potent inhibitors with different chemical scaffolds for quantitative analysis of their binding affinities.

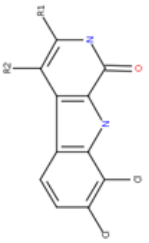
Table 4.4 Melting temperature shifts (°C) for all PIM variants and selection of the most interesting compounds as determined by DSF. Higher values correspond to stronger ligand binding and are shown in deeper red.

	SGC ID	PIM1	S46K/S51T	E124L	S46K	V126A	S51T	H68V	I56H/P63Q	PIM2	PIM3
Novel GSK compounds	K02994a	4.6	4.9	4.1	4.8	5.8	4.3	4.4	5.2	7.8	6.7
	K03077a	4.8	4.8	4.3	4.9	5.9	4.5	3.3	4.0	7.6	5.9
	K02978a	4.8	2.5	5.2	5.2	4.1	6.0	5.9	3.8	10.1	7.6
	K00228b	5.2	7.0	5.9	7.0	6.0	7.2	2.7	1.9	10.0	7.7
	K03020a	5.9	5.9	4.3	7.5	5.1	5.4	5.3	3.7	5.1	6.1
	K03149a	6.0	5.9	5.9	4.7	4.2	6.0	6.1	3.8	4.7	4.9
	K02908a	6.0	6.9	7.2	1.0	7.0	6.4	6.5	5.1	11.1	10.1
	K03146a	6.3	5.9	6.3	5.1	5.1	5.7	6.2	4.5	1.4	4.9
	K03137a	6.5	6.2	6.1	5.6	4.6	5.9	6.3	4.5	0.9	4.5
	K02803a	7.0	8.7	6.8	9.9	8.4	7.7	7.1	6.6	11.7	8.9
	K02967a	7.0	6.5	6.1	5.8	7.6	6.9	6.3	5.7	8.6	7.4
	K03128a	7.2	6.7	6.2	5.4	3.8	6.5	7.0	4.8	0.6	3.5
	K03129a	7.4	7.0	7.0	5.9	4.2	6.7	7.3	5.8	1.8	5.0
	K03036a	7.7	7.3	5.7	6.2	6.6	7.8	6.2	5.3	7.6	7.0
	K03141a	8.1	7.3	7.8	6.3	5.7	6.7	7.0	5.6	5.2	6.5
	K02985a	8.3	8.1	7.9	7.9	6.7	8.2	8.9	6.6	8.6	7.6
	K02931a	8.5	8.8	8.9	9.7	7.1	8.4	7.9	6.7	11.4	10.6
	K03140a	9.6	9.2	9.3	8.9	6.2	8.8	9.6	7.9	2.5	6.7
	K02888a	10.2	9.5	9.9	10.2	8.8	10.0	9.4	7.8	8.0	10.1
Imidazopyridazines	K00493	5.7	5.3	5.5	5.6	3.9	5.3	5.9	3.9	5.4	5.0
	K00499	5.7	5.1	4.7	5.2	3.0	4.8	5.7	3.6	2.6	4.1
	K00135	5.9	5.6	5.1	5.9	4.8	5.9	6.6	4.6	4.3	6.8
	K00514	6.4	5.5	5.4	6.1	4.5	5.3	6.3	4.4	5.3	6.3
	K00488	6.6	6.0	5.8	6.7	4.8	6.3	6.7	4.7	5.9	6.5
	K00512	7.7	6.7	6.5	7.2	4.8	6.4	8.0	6.0	6.0	6.2
	K00487	7.8	6.8	7.4	7.6	5.0	7.5	8.1	5.7	5.6	6.4
	K00516	7.8	7.0	6.9	7.2	5.0	6.8	8.3	6.7	5.9	6.1
	K00500	8.1	7.2	7.4	7.4	5.2	7.1	8.6	5.5	4.9	7.1
	K00486	8.2	7.4	7.4	7.9	5.7	7.7	7.9	6.0	6.0	7.2
Furanes	K00869a	5.7	5.8	6.5	5.1	4.9	4.7	6.0	5.3	6.0	6.9
	K00827a	6.0	6.1	7.3	6.2	7.2	7.0	5.7	3.9	6.2	5.2
	K00884a	6.3	6.2	6.6	5.8	6.7	3.3	6.4	4.7	8.1	9.5
	K00875a	6.3	6.6	7.2	5.9	5.8	5.3	6.5	5.5	5.9	7.3
	K00893a	6.4	6.8	6.1	5.7	5.7	5.2	1.5	1.4	7.9	8.2
	K00828a	6.7	6.8	6.3	7.2	6.3	5.4	6.7	4.4	5.3	7.6
	K00913a	6.8	7.2	7.6	6.5	6.3	3.2	8.0	6.1	7.0	8.4
	K00878a	7.0	7.6	7.2	6.7	7.3	5.0	6.6	6.0	7.9	10.2
	K00894a	7.1	3.4	7.3	6.1	4.5	5.0	1.3	0.8	7.7	11.6
	K00872a	7.1	7.1	7.6	7.1	4.8	5.9	7.0	5.7	4.6	5.8
	K00780a	7.3	7.3	6.2	8.5	8.2	7.1	5.2	4.1	7.5	13.4
	K00910a	7.4	6.9	4.8	5.9	6.5	2.8	7.7	2.1	5.8	10.8
	K00865a	7.5	7.9	7.5	6.3	6.7	5.2	1.9	4.8	7.8	9.9
	K00880a	7.6	6.5	8.2	6.6	8.5	4.2	6.0	3.4	4.1	12.6
	K00879a	7.7	7.4	8.0	6.4	7.4	6.3	8.4	7.3	8.4	9.5
	K00881a	8.1	6.7	6.5	6.7	9.0	5.7	0.8	1.2	9.7	13.0
	K00909a	9.3	5.8	7.7	7.3	6.8	5.9	2.2	0.9	6.1	11.9
	K00906a	9.9	3.5	8.0	-0.2	8.0	0.7	1.7	2.4	12.0	14.4
	K00889a	10.3	9.9	10.6	9.3	7.5	7.7	9.5	7.1	10.7	14.3
Carbolines	K01763a	2.8	3.8	1.7	2.8	1.9	1.6	3.3	1.4	2.7	3.2
	K01740a	2.9	3.8	2.5	2.9	1.8	2.8	3.4	0.5	1.9	3.4
	K01845a	3.0	3.6	2.0	2.6	3.1	2.4	1.8	1.5	2.7	3.8
	K01736a	3.0	3.6	2.9	3.2	3.4	3.8	3.2	1.1	1.9	4.0
	K01739a	3.3	3.8	2.6	3.5	2.2	2.7	3.8	1.8	0.9	3.8
	K01789a	3.5	2.5	3.0	3.8	2.5	2.9	4.0	1.8	3.3	4.6
	K01847a	3.9	3.9	1.1	3.8	2.6	2.7	1.7	3.1	1.6	4.9
	K01741a	4.0	4.9	3.5	4.2	3.4	4.0	2.7	0.3	1.6	4.7
	K01737a	4.8	5.8	4.6	5.2	5.0	4.9	5.2	2.9	2.8	6.2
	K01768a	4.8	6.0	4.2	5.2	3.9	4.7	5.1	2.5	4.1	4.9
	K01764a	5.1	5.9	4.8	5.0	3.7	4.8	4.9	3.5	4.1	5.0
	K01786a	6.2	6.5	4.5	6.2	5.3	6.4	6.6	4.7	3.6	7.1
	K01742a	6.5	7.4	6.3	6.9	5.5	6.4	7.2	5.0	4.7	7.2
	K01844a	9.6	9.4	8.2	8.6	8.0	8.2	8.2	7.8	6.6	9.6









4.5. PIM variants demonstrate different enzyme kinetics

Melting temperature shift assay is a good first indication of protein-ligand interaction, but the data produced may be subject to errors and potentially to artefacts (Hau et al., 2011). To obtain comprehensive and reliable results, the biophysical thermodynamic studies were complemented with enzymology experiments. Peptide phosphorylation was monitored using mobility shift microfluidic assay (Caliper), which is based on the electrophoretic separation of phosphorylated (product) and non-phosphorylated (substrate) peptides from each well, based on their charge difference. The peptides are tagged with a fluorescent marker and can be easily detected allowing direct and real-time measurements of kinase activities (see Methods section for more details).

A key step prior to evaluation of protein-inhibitor binding is the establishment of a relationship between rate of reaction and concentration of a substrate, i.e the affinity of the enzyme for its substrate. Due to previous experimental experiences where substrate affinities for PIM proteins were determined to be very similar, we avoided high consumption of the fluorescent peptide substrate by omitting the $K_m^{peptide}$ determination step and assumed that the substrate affinities for all PIM variants were comparable.

In contrast, K_m^{ATP} values, defined as ATP concentration that results in half maximal reaction velocity demonstrated that PIM2 had much higher affinity for ATP than PIM1 and PIM3 with K_m^{ATP} value about 10 fold lower (Table 3.2). We monitored the extent to which PIM1 variants mimic the behaviour of PIM2 kinase and discovered that the P-loop mutants S46K and S46K/S51T had intermediate K_m^{ATP} values between PIM1 and PIM2, indicating that lysine at the position 46, located at the tip of the P-loop, could be responsible for stronger ATP binding in PIM2 by playing a role in the P-loop flexibility (Figure 4.5).

The hinge mutation E124L resulted in elevated K_m^{ATP} indicating that the removal of the E124-R122 salt bridge not only affects stability as suggested by the DSF results (Table

3.1) but also lobe dynamics of the kinase, which influences ATP binding. The P-loop double mutant I56H/P63Q had the highest value of K_m^{ATP}, 718.0 ± 225.6 μM, illustrating that the potential hydrogen bond formed by polar histidine and glutamate has a strong effect on ATP affinity.

We also used the ratios of Michaelis-Menten constants (k_{cat}/K_m) as descriptors for the catalytic efficiency of each variant, where k_{cat} is the rate of reaction (Table 3.2). Overall, PIM3 has the highest catalytic efficiency with k_{cat}/K_m of 10.7 ± 1.7 (M^{-1} s^{-1} x 10^6), whereas the mutants E124L and I56H/P63Q had decreased at least two-fold on average compared to that of PIM1, which had k_{cat}/K_m of 2.2 ± 0.3 (M^{-1} s^{-1} x 10^6). Interestingly, this double mutant together with the hinge mutation (E124L) demonstrated the turnover rate (k_{cat}) to be very similar to PIM2. In addition, we observed that the ability of the P-loop mutants S46K, S51T and S46K/S51T to phosphorylate the substrate increased by 2-3 that of PIM1, confirming the fact that these mutations were replicating enzymatic behaviour of PIM2, for which k_{cat}/K_m was determined to be 9.5 ± 1.7 (M^{-1} s^{-1} x 10^6).

Table 4.5 Steady state kinetic constants for the ATP titration experiments as determined by Caliper microfluidic mobility shift assay. The values are reported as the mean ± standard deviation resulting from three separate measurements.

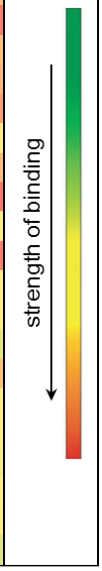
PIM variant	K_m (μM)	k_{cat} (s^{-1} x 10^4)	k_{cat}/K_m (M^{-1} s^{-1} x 10^6)
PIM1	440.4 ± 41.7	9.7 ± 0.6	2.2 ± 0.3
S46K/S51T	203.5 ± 30.7	11.8 ± 0.1	6.3 ± 1.0
E124L	534.2 ± 105.8	4.4 ± 0.3	0.8 ± 0.2
S46K	294.3 ± 54.8	11.7 ± 0.5	4.2 ± 0.8
V126A	300.4 ± 52.7	10.8 ± 0.5	3.9 ± 0.7
S51T	277.8 ± 65.9	13.6 ± 1.5	5.6 ± 1.5
H68V	396.3 ± 62.1	7.7 ± 0.8	2.3 ± 0.4
I56H/P63Q	718.0 ± 225.6	4.1 ± 0.4	0.6 ± 0.2
PIM2	34.4 ± 5.6	3.6 ± 0.3	9.5 ± 1.7
PIM3	308.7 ± 49.4	25.5 ± 1.0	10.7 ± 1.7

4.6. Evaluation of IC_{50} values

After the initial DSF screening of PIM kinases, 23 compounds were selected for the enzyme inhibition assay (Table 4.6). All experiments were conducted with ATP concentration corresponding to K_m^{ATP} for each kinase (Table 3.2). As anticipated, PIM1 and PIM3 appeared to be more susceptible to inhibition than PIM2, typically showing several fold selectivity complying with the previously published results (Pogacic et al., 2007). The selectivity of the GSK compounds detected during DSF screening has been confirmed: as an example, the average IC_{50} value for the compound K03128a was determined to be 0.07 μM for PIM1 and > 10 μM for PIM2. And like in DSF, this compound has weaker affinity for the hinge mutants E124L and V126A with mean IC_{50} values of 0.27 μM and 0.51 μM respectively. This result further indicates that substitution of residues located on the hinge could be key determinants of the inhibitor binding affinity. The mutant I56H/P63Q with the double substitution at the back of the P-loop appeared to mimic the PIM2 behaviour with intermediate IC_{50} values for most compounds. The S46K and S51T P-loop substitutions, however, resulted in mutants with similar binding profile to that of PIM1.

Table 4.6 Table showing IC_{50} values (μM) for hits with type 2 scaffold and the positive control staurosporine determined by Caliper microfluidic mobility shift assay. The values are means of three independent experiments. Standard deviations are presented in Appendix E.

SGC ID	PIM1	S46K/S51T	E124L	S46K	V126A	S51T	H68V	I56H/P63Q	PIM2	PIM3A
K00997a	0.01	0.01	0.02	0.01	0.02	0.01	0.00	0.02	0.34	0.02
K03140a	0.01	0.02	0.04	0.02	0.08	0.02	0.01	0.05	>10	0.10
K01844a	0.02	0.04	0.06	0.03	0.04	0.03	0.02	0.06	1.48	0.05
K03129a	0.03	0.05	0.10	0.04	0.17	0.04	0.02	0.12	6.26	0.20
K00746a	0.04	0.04	0.11	0.06	0.05	0.04	0.04	0.11	2.65	0.06
K02698a	0.05	0.03	0.13	0.08	0.15	0.03	0.04	0.15	0.25	0.20
K03527a	0.05	0.09	0.13	0.08	0.10	0.07	0.02	0.13	0.43	0.18
K01742a	0.05	0.09	0.14	0.08	0.09	0.07	0.04	0.16	5.85	0.13
K03128a	0.07	0.14	0.27	0.12	0.51	0.09	0.05	0.29	>10	0.57
K03137a	0.07	0.12	0.18	0.10	0.15	0.09	0.07	0.21	>10	0.19
Staurosporine	0.16	0.21	0.23	0.18	0.05	0.18	0.12	0.26	0.22	0.04
K00070a	0.17	0.19	0.58	0.23	0.55	0.20	0.30	0.49	4.13	0.81
K00906a	0.18	0.19	0.40	0.22	0.16	0.23	0.18	0.33	0.34	0.05
K00487	0.21	0.30	0.40	0.24	1.01	0.22	0.13	0.30	0.45	0.42
K00135	0.26	0.41	0.55	0.37	0.37	0.34	0.25	0.70	6.23	0.40
K00512	0.51	0.72	1.25	0.71	1.86	0.82	0.36	1.19	2.43	4.17
K00780a	0.60	0.34	2.27	0.40	0.49	0.82	0.70	1.34	0.96	0.20
K02723a	0.60	0.89	1.03	0.64	1.19	0.88	0.34	1.77	>10	0.42
K02908a	0.68	0.57	0.98	0.66	1.09	0.79	0.45	0.74	0.45	0.30
K00486	0.69	1.33	1.22	1.15	2.00	0.95	0.53	1.93	9.53	2.20
K02978a	1.12	1.66	2.72	1.87	3.48	2.01	1.42	1.92	0.32	0.38
K00072	1.20	1.11	3.20	1.05	1.12	1.21	0.87	1.69	0.98	0.83
K01779a	3.48	5.57	>10	4.55	7.89	7.08	4.86	6.15	>10	4.98



4.7. ITC experiments determine binding constants and thermodynamic properties of protein-ligand binding

The T_m and Caliper assays are good pre-screening steps for nomination of mutants and compounds for more detailed analysis of protein-ligand interactions via thermodynamic studies. Isothermal titration calorimetry (ITC) is the most direct method for determining the thermodynamic properties of protein-ligand binding and comparing those properties could help to understand the reasons for compound selectivity. However, the limiting factor is that each ITC run requires large amounts of purified proteins and the problem scales if many compounds are being tested.

The experimental characterisation of PIM1 interaction with the imidazo [1,2-b] pyridazine inhibitor (K00135) by ITC has already been performed (Bullock et al., 2005b). The Gibbs free energy of binding was determined to be -9.8 kcal/mol with a moderate binding enthalpy of -5.3 kcal/mol. To compare the inhibition of the two proteins directly and to understand what interactions govern the free energies of binding, we performed ITC measurements for the PIM2 kinase with this inhibitor under the same conditions (Table 4.7). As anticipated, the K00135 compound has a greater (more than 10-fold) affinity for PIM1 than PIM2 with dissociation binding constants of 27.7 ± 4.1 nM and 306.9 ± 55.9 nM respectively. The total free energy of binding for the PIM2 interaction was established to be 1.4 kcal/mol less than for PIM1 with ΔG of -8.4 kcal/mol. Moreover, while similarity in the inhibitor binding mode and the active site environment in both kinases was suggested by very close values of enthalpic contributions, the variation in compound affinity could be explained by differences in entropic contributions, such as conformations of the P-loop and presence of solvent within the binding site.

Table 4.7 Thermodynamic parameters of K00135 inhibitor binding to PIM kinases as determined by ITC. All errors shown for the ITC are taken from the fit of the titration determined by the Origin software.

PIM kinase	K_d (nM)	ΔHobs (kcal/mol)	TΔS (kcal/mol)	n	ΔG (kcal/mol)
PIM1*	27.7±4.1	-5.3	4.5	0.93	-9.8
PIM2	306.9±55.9	-5.1	3.3	0.88	-8.4

* PIM1 experiments were conducted previously (Bullock et al., 2005b) and are included here for comparison.

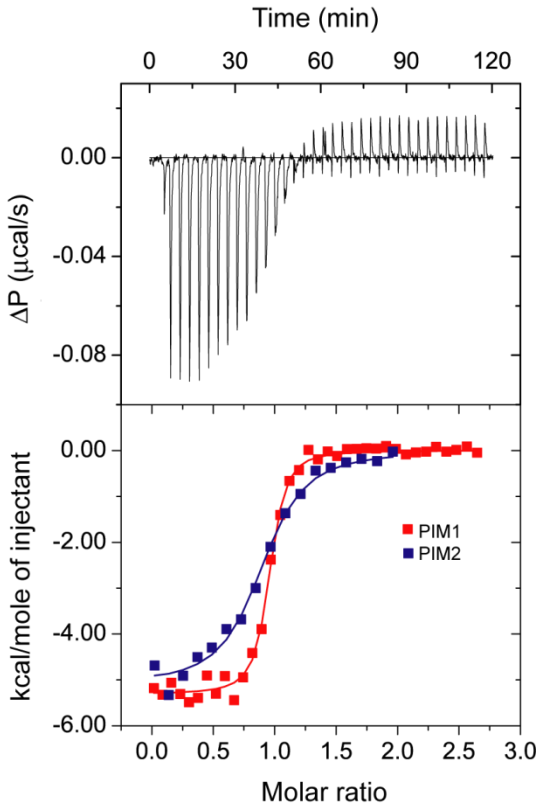


Figure 4.7 Isothermal calorimetry data of K00135 binding to PIM1 and PIM2 kinases. The protein has been titrated into a solution of the ligand in 50 mM HEPES pH 7.5, 150 mM NaCl and 1 mM DTT at 10 °C. Row binding data reveal exothermic binding while decreasing heat signals upon each injection indicate saturation of binding (PIM1 data shown as an example in the upper panel). The normalised binding isotherms are shown in the lower panel. Data were fit to a single site model using Origin 7.0. The best fitting functions are shown as solid lines, and fitting parameters of these experiments are summarised in Table 4.7.

4.8. Crystal structures highlight that flexibility plays a role in compound binding

4.8.1 Comparison of PIM1 and PIM2 structures

Although the PIM kinase family has been extensively studied prior to this work, only one PIM2 structure has been deposited in the PDB (Pogacic et al., 2007). When a series of imidazo[1,2-b]pyridazine inhibitors was identified with anti-leukemic activity, PIM2 was markedly less susceptible to inhibition than PIM1 (Pogacic et al., 2007). A crystal structure of PIM1 with the compound K00135a has been available since 2005 (PDB: 2c3i) and the PIM2 complex with the same compound would directly allow assessment of the differences in binding. We therefore crystallised the truncated PIM2 construct (Section 4.2) in apo-form and after soaking experiments obtained a structure of the PIM2 complex with K00135a. The PIM2 protein crystallised in space group $P2_1$ with four protein molecules per asymmetric unit (Table 4.8).

Superposition of main chain atoms of PIM1 and PIM2 reveals high similarity between the two kinases with all atom root-mean-square deviation (RMSD) of 0.472 Å, and with nearly identical binding mode of the inhibitor (Figure 4.8). The compound K00135 binds to the ATP binding pocket and the P-loop adopts an inactive or ‘folded’ conformation where phenylalanine (F49, hereafter PIM1 numbering) points into the active site, making a kinase incompatible with ATP binding. The inhibitor accepts a hydrogen bond from the side chain of catalytic K67 and forms a number of hydrophobic contacts, including L44, F49, I104, and L120, which stabilise the interaction (Figure 4.8b). Presumably, a weaker contact is formed with the hinge region by donation of a hydrogen bond to the backbone carbonyl of hinge residue (E121 in PIM1) (Pogacic et al., 2007). As stated in Section 4.1.1, the hinge region is highly conserved across PIM kinases with the exception of E124 and V126. The latter is an alanine in PIM2 and is the only one conservative change amongst residues that make interactions with the inhibitor. Furthermore, the side chain of E124 forms a salt bridge with the neighbouring

arginine. The substitution of E124 for leucine in PIM2 destroys this salt bridge and may change dynamics of the N-terminal and C-terminal lobes (Figure 4.8c). The new PIM2 structure reveals that this arginine forms a strong bond with a glutamine, which his not present in PIM1 (P63Q substitution). These changes may determine mobility of the N- and C-terminal lobes and hence protein stability and ligand binding.

The P-loop in both PIM1 and PIM2 complexes with K00135 displays a folded conformation. In PIM1, all P-loop atoms are well defined by the electron density and have low B-factors revealing high loop rigidity and tight F49-inhibitor stacking (Figure 4.9). In PIM2, however, the corresponding atoms either have a high B-factor compared to the globular protein core or could not be resolved. These differences are consistent with our findings on compound binding, and while the B-factor analysis could perhaps be treated with some caution, since mobility of the P-loop in PIM1 could potentially be affected by crystal packing thereby locking it in one position, there are PIM1 structures that resulted from the same crystal packing and have P-loop disordered (Table 4.2). Moreover, in apo-PIM2 structure the P-loop residues were disordered, indicating that the P-loop is free to move in between different conformations and lack electron density even though it could be structured. This confirms that there is a link between the P-loop flexibility and the differential affinities of PIM inhibitors and ATP.

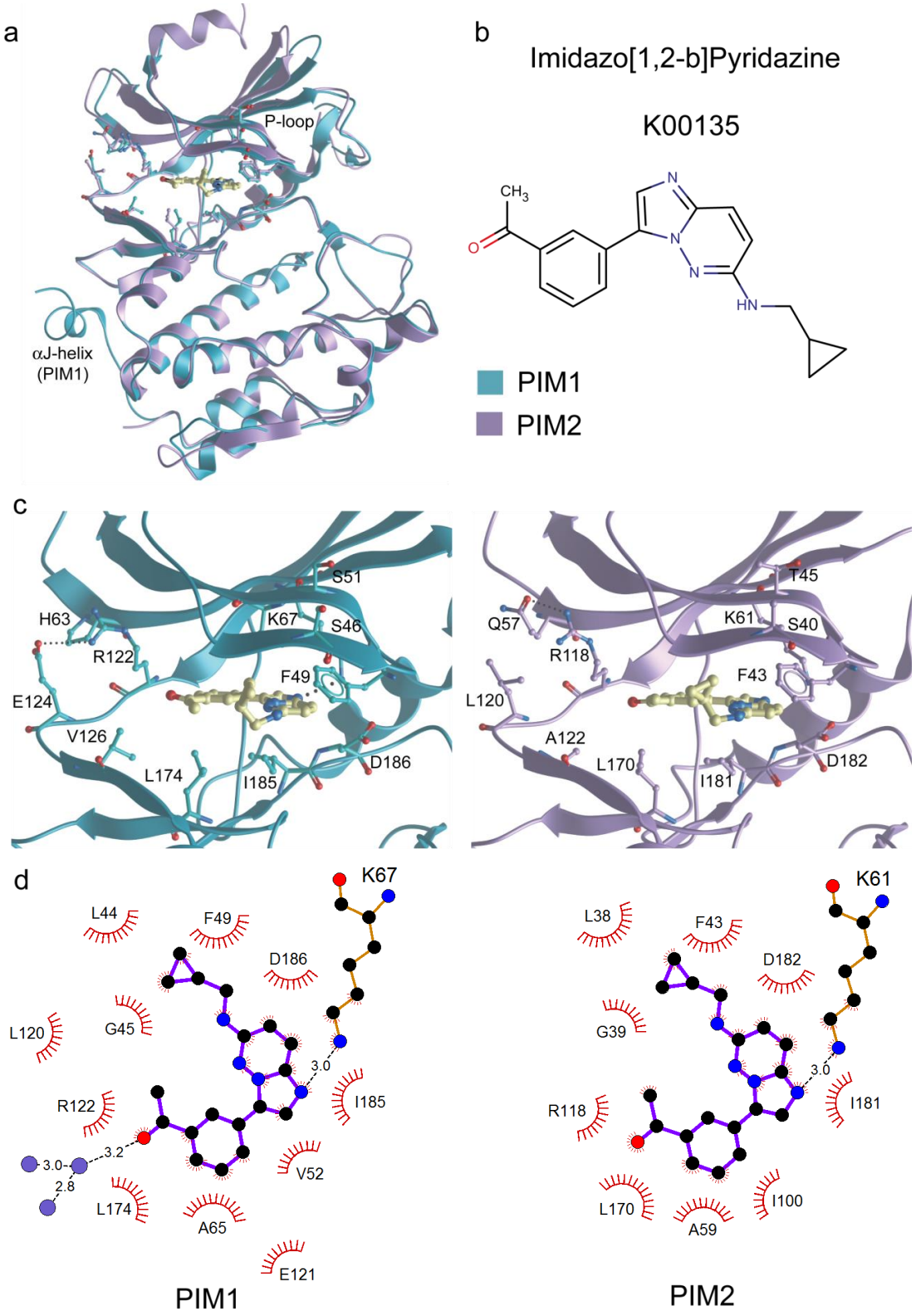


Figure 4.8 Comparison of PIM1 and PIM2 complexes with K00135. (a) Overlay of the two proteins is shown in ribbon representation reveals the strong conservation of the kinase fold, with the exception of the C-terminal αJ helix, which is not present in PIM2. (b) Chemical structure of the K00135 compound. (c) A close-up view of the binding site region with the side chain making contacts with the inhibitor, and relevant mutated residues are shown as sticks. Dotted lines represent hydrogen bonds between R122-E124 in PIM1 and Q57-R118, as described in the text. (d) Interaction maps of K00135 in the active site of PIM1 and PIM2 kinases generated by LigPlot+ (Laskowski and Swindells, 2011).

Table 4.8 Crystallographic data and refinement statistics.

Protein	Apo PIM2	PIM2 with K00135	PIM1 V126A mutant with K00487a
Data collection			
Space group	$P2_1$	$P2_1$	$P6_5$
Unit cell parameters [a, b, c (Å)] [α, β, γ (°)]	72.3, 62.4, 144.0 90.0, 93.1, 90.0	72.9, 62.6, 144.4 90.0, 93.1, 90.0	98.0, 98.0, 81.0 90.0, 90.0, 120.0
Resolution range [a] (Å)	49.63 – 2.15 (2.27 – 2.15)	49.91 – 2.00 (2.11 – 2.00)	49.01 – 2.05 (2.11 – 2.05)
No. of unique observations [a]	67888 (9867)	84670 (11844)	27847(2143)
Completeness [a] (%)	97.0 (97.3)	96.2 (93.1)	100.0(100.0)
<I/σ(I)> [a]	6.9 (2.0)	6.2 (1.9)	13.5(1.9)
CC(1/2)	0.99 (0.52)	0.99 (0.68)	0.99(0.57)
Wilson B	20.07	20.51	28.48
Multiplicity [a]	6.3 (6.2)	3.3 (3.4)	9.3(9.4)
R_{merge} [a] (%)	21.7 (104.5)	11.7 (63.3)	13.7(140.3)
R_{pim} [a] (%)	9.3 (45.2)	7.5 (40.2)	7.1(72.7)
Refinement			
R_{work} (R_{free}) [b] (%)	22.09(27.56)	22.50(26.78)	17.77(19.86)
RMSD bond length (Å)	0.0114	0.0114	0.0092
RMSD bond angle (°)	1.499	1.4817	1.2900
Number of atoms	8469	8566	2423
Protein (residues)	1060	1064	274
Water molecules	281	302	130
Ramachandran			
Preferred (%)	96.53	97.61	96.38
Allowed (%)	3.47	2.39	3.62
Disallowed (%)	0.0	0.0	0.0

[a] Values in parentheses show the statistics for the highest resolution shell.

[b] Rfree was calculated using 5% of the diffraction data assigned randomly and not used throughout refinement.

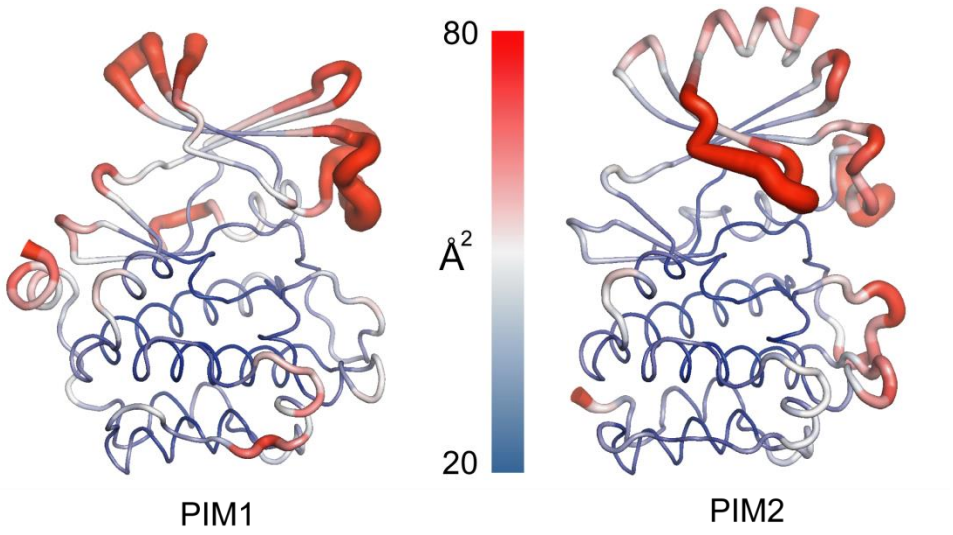


Figure 4.9 B-factors analysis of co-crystal structures with the same inhibitor K00135 reveals that the atomic fluctuations of the P-loop in PIM2 are much higher than in PIM1. The structures are drawn in 'cartoon putty' representation and the colour is ramped by residue from blue as the lowest B-factor value to red as the highest B-factor value; additionally, thicket tubes represent larger B-factors.

4.8.2 Valine 126 contributes to strong interactions between PIM1 and imidazopyridazine inhibitors

The only non-conserved residue in the active site interacting with the imidazopyridazine inhibitor was valine at position 126 in PIM1, which is replaced by alanine in PIM2. To address the role of this substitution we co-crystallised the V126A mutant with imidazopyridazine inhibitor K00487a. Providing that the compound K00487 has the same binding mode with the wild-type PIM1 (Figure 4.10c), the mutation resulted in distance between the inhibitor and the side chain of this hinge residue of 4.9Å compared to 3.4Å in PIM1 (PDB: 2c3i). We therefore suggest that additional interaction of V126 in PIM1 with the inhibitor results in stronger affinity. This complies with the DSF and IC_{50} results (Table 3.1 and Table 4.6), where the 2.8 °C difference in ΔT_m and 5-fold variation in IC_{50} indicated that this V126A mutation reduced the potency of the inhibitor.

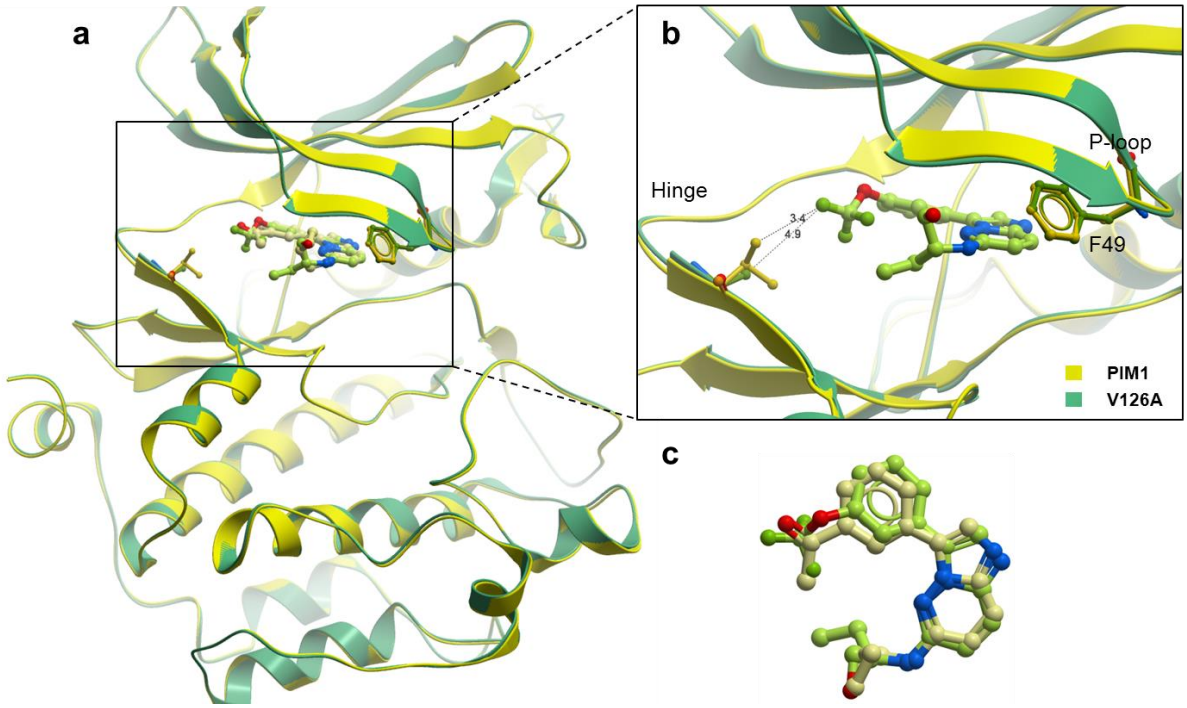


Figure 4.10 Crystal structure of V126A mutant of PIM1 kinase in complex with K00487 inhibitor. (a) An overlay with wild-type PIM1 structure (PDB: 2c3i), F49 and V126A are shown as sticks. In both structures F49 residue is in 'folded' conformation (b) ATP-binding pocket occupied by the inhibitor, shortest contact distances between K00487 and the side chain of the residue at position 126 are shown, (c) an overlay of inhibitors from the co-structures illustrating no differences in the binding mode, K00487 (green) and K00135 (yellow). Both compounds belong to the imidazopyridazine class.

4.9. Discussion and conclusions

The PIM2 protein kinase is known to demonstrate different ligand binding behaviour compared to PIM1 and PIM3 despite highly conserved active sites. To date, most experimental data have been generated using PIM1, while very little is known about the regulation of PIM2 and PIM3. Nonetheless, recent biological studies indicate that PIM2 could be a more attractive therapeutic target (Gómez-Abad et al., 2011, Lu et al., 2013) and the origin of this ligand selectivity for PIM1 and PIM3 over PIM2 remains an important question in the field of PIM protein kinases.

The effects of divergent evolution on the three-dimensional protein structures and ligand binding of the PIM protein kinase family members has been addressed in this

project. A set of residues has been identified as candidates for mutagenesis and followed by characterisation of the thermodynamic, kinetic and structural properties of the protein panel.

A simple explanation for the difference in compound affinity could be a difference in the global stabilities of the proteins, which were evaluated through DSF. We established that the melting temperature of PIM2 was about 5 °C lower than that of PIM1 and PIM3, suggesting that this enzyme is thermally and conformationally less stable. Hence, the differences in binding could be a result of a difference in the stability and the dynamic properties defining the individual proteins.

Enzymatic experiments showed that PIM2 has much higher affinity for ATP than PIM1 and PIM3, with about 10-fold increase in K_m^{ATP}, in agreement with the published results (Burger et al., 2013). As an essential part in the catalytic cycle, protein kinases undergo large conformational transitions, which involve movements of flexible loops and rotations of domains thereby reshaping the active site and allowing ATP and downstream substrate to enter and leave. Structural studies have demonstrated that in PIM kinases rearrangements of the P-loop are of particular significance. As suggested by apo- and AMP-PNP inhibited PIM1 structures, this P-loop movement may have important consequences, and ATP-compatible binding may involve energetic cost due to rotation of F49 out of the binding pocket. This energetic penalty could be lower in PIM2 if the P-loop adopts multiple conformations with similar probabilities, or if ATP binds to the near-native intermediate state of this protein.

Evolutionary analysis and structural comparisons indicated that the P-loop is the prime candidate responsible for differences in ATP-binding and affinities for inhibitors. Another line of evidence was provided by ITC experiments, which suggested that the difference in ligand binding could be explained by conformational selection. We compared binding of the K00135 inhibitor to PIM1 and PIM2 and discovered that despite 10-fold variation difference in dissociation constants (K_d of 27.7 ± 4.1 nM and 306.9 ± 55.9 nM respectively), the binding enthalpies were very similar. Therefore, the

reduced binding affinity for PIM2 could be explained by the entropic penalty, suggesting a conformational search for the complex state.

Furthermore, the observed differences in binding mode of this inhibitor were not apparent from the high-resolution structural studies of PIM1 and PIM2 and we therefore rule out major structural variations as a cause for the difference in inhibition. Instead, the P-loop in PIM2 was much more flexible and loss of intermolecular associations due to this flexibility could be one of the reasons of weaker compound affinity. The degrees of freedom of P-loop motion were even higher in the PIM2 apo-structure, where some P-loop residues could not be resolved. While the folded conformation was better defined with the inhibitor, there still was relatively little induced strain due to binding, given that the P-loop remained poorly defined in the electron density map compared to PIM1. These differences match with the experimental determination of binding affinity of the inhibitor to PIM1 and PIM2 proteins by DSF, IC_{50} values, and ITC.

Several mutants have highlighted that reduction of flexibility caused by point mutations results in affinity variations for some ligands. For example, wild-type PIM1 was potently inhibited by K00135 and K00487 compounds, yet the V126A and I56H/P63Q mutants demonstrated weaker binding, thereby replicating PIM2 affinity profiles.

On the other hand, we discovered that enzymatic abilities of the P-loop mutants S46K, S51T and S46K/S51T were mimicking PIM2 by showing lower K_m^{ATP} values and improved catalytic efficiencies compared to PIM1. It is therefore likely that these P-loop modifications reduce the energetic penalty that is required to displace the P-loop and to accommodate ATP into the binding site. Once highly optimised PIM1 inhibitors like K00135 are present, binding of ATP could be disfavoured, since the former do not change the structure significantly.

Recently reported low-picomolar range pan-PIM inhibitors no longer demonstrate preference for PIM1 and PIM3 over PIM2 (Burger et al., 2013, Wang et al., 2013),

illustrating the fine interplay between compound potency and conformational dynamics of its target.

While our findings sheds some light on the observed differences in PIM1 and PIM2 behaviour, much more work still needs to be done to address this complex thermodynamic problem involving ligand binding.

4.10. Future work

To complete this work and to enhance understanding of how ligands interact with individual PIM kinases by considering association and dissociation rates of inhibitor binding, SPR experiments will be performed at GSK as a part of the K4DD initiative (http://www.imi.europa.eu/content/k4dd).

What remains to be developed is a detailed understanding of how flexibility underlies the functionality and ligand binding in PIM kinases. Therefore, molecular dynamics (MD) simulations coupled with NMR hydrogen-deuterium exchange studies could be used to determine the differences in flexibility of specific regions of these proteins, such as the P-loop and hinge.

Thus far we have shown that each PIM kinase family member is highly conserved throughout animal evolution. The basic questions are yet to be answered: what are the reasons for nature producing the three family members and what are the differences in their functions. While a number of biological studies about the importance of PIM kinases in different cancers have been made (Mikkers et al., 2004, Lu et al., 2013, Brault et al., 2010), differentiation of the specific functions still remain to be investigated systematically. Chemical probes for each individual PIM kinase could potentially reveal the functional roles of each protein. Development of chemical probes in collaboration with the University of Strathclyde has begun, and understanding of which part of an inhibitor is important for binding affinity is an important aspect of that

project. In addition, since PIM kinases have an atypical hinge region characterised by an insertion of one additional residue and the presence of a proline, PIM family-specific inhibitors could be developed.

Since motions and flexibility of protein kinases are essential for their function, it is not surprising that in response to environmental changes, evolutionary pressure could tune protein dynamics instead of altering the structural fold. Homologues from the PIM protein kinase family, which share high sequence identity and structural similarity, provide an excellent system to explore the mechanisms by which this dynamics is affected. Further understanding of the residues contributing to flexibility and ligand binding will add to the amount of knowledge about protein kinases in general. Other kinase systems demonstrating differential binding behaviour, like DYRK1 and DYRK2 (Soundararajan et al., 2013) could also be useful for future investigations of the relationships between sequence, structure and ligand interactions.

References

1994. The CCP4 suite: programs for protein crystallography. *Acta crystallographica. Section D, Biological crystallography,* 50, 760-763.

AHO, T., SANDHOLM, J., PELTOLA, K., MANKONEN, H., LILLY, M. & KOSKINEN, P. 2004. Pim-1 kinase promotes inactivation of the pro-apoptotic Bad protein by phosphorylating it on the Ser112 gatekeeper site. *FEBS letters,* 571, 43-49.

ALTSCHUL, S., MADDEN, T., SCHÄFFER, A., ZHANG, J., ZHANG, Z., MILLER, W. & LIPMAN, D. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research,* 25, 3389-3402.

AON, J., CAIMI, R., TAYLOR, A., LU, Q., OLUBOYEDE, F., DALLY, J., KESSLER, M., KERRIGAN, J., LEWIS, T., WYSOCKI, L. & PATEL, P. 2008. Suppressing posttranslational gluconoylation of heterologous proteins by metabolic engineering of Escherichia coli. *Applied and environmental microbiology,* 74, 950-958.

ASLANIDIS, C. & DE JONG, P. J. 1990. Ligation-independent cloning of PCR products (LIC-PCR). *Nucleic acids research,* 18, 6069-6074.

BACHMANN, M. & MÖRÖY, T. 2005. The serine/threonine kinase Pim-1. *The international journal of biochemistry & cell biology,* 37, 726-730.

BAIROCH, A., APWEILER, R., WU, C., BARKER, W., BOECKMANN, B., FERRO, S., GASTEIGER, E., HUANG, H., LOPEZ, R., MAGRANE, M., MARTIN, M., NATALE, D., O'DONOVAN, C., REDASCHI, N. & YEH, L.-S. 2005. The Universal Protein Resource (UniProt). *Nucleic Acids Research,* 33, D154-D159.

BATTYE, G., KONTOGIANNIS, L., JOHNSON, O., POWELL, H. & LESLIE, A. 2011. iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta crystallographica. Section D, Biological crystallography,* 67, 271-281.

BERMAN, H., WESTBROOK, J., FENG, Z., GILLILAND, G., BHAT, T. N., WEISSIG, H., SHINDYALOV, I. & BOURNE, P. 2000. The Protein Data Bank. *Nucleic Acids Research,* 28, 235-242.

BERTHET, C., ALEEM, E., COPPOLA, V., TESSAROLLO, L. & KALDIS, P. 2003. Cdk2 knockout mice are viable. *Current biology : CB,* 13, 1775-1785.

BRAULT, L., GASSER, C., BRACHER, F., HUBER, K., KNAPP, S. & SCHWALLER, J. 2010. PIM serine/threonine kinases in the pathogenesis and therapy of hematologic malignancies and solid cancers. *Haematologica,* 95, 1004-1015.

BROWN, N., LOWE, E., PETRI, E., SKAMNAKI, V., ANTROBUS, R. & JOHNSON, L. 2007. Cyclin B and cyclin A confer different substrate recognition properties on CDK2. *Cell cycle (Georgetown, Tex.),* 6, 1350-1359.

BROWN, N. R., NOBLE, M. E., ENDICOTT, J. A. & JOHNSON, L. N. 1999. The structural basis for specificity of substrate and recruitment peptides for cyclin-dependent kinases. *Nature cell biology,* 1, 438-443.

BULLOCK, A., DEBRECZENI, J., AMOS, A., KNAPP, S. & TURK, B. 2005a. Structure and substrate specificity of the Pim-1 kinase. *The Journal of biological chemistry,* 280, 41675-41682.

BULLOCK, A., DEBRECZENI, J., FEDOROV, O., NELSON, A., MARSDEN, B. & KNAPP, S. 2005b. Structural Basis of Inhibitor Specificity of the Human Protooncogene Proviral Insertion Site in Moloney Murine Leukemia Virus (PIM-1) Kinase. *Journal of Medicinal Chemistry,* 48, 7604-7614.

BULLOCK, A., RUSSO, S., AMOS, A., PAGANO, N., BREGMAN, H., DEBRECZENI, J., LEE, W., DELFT, F., MEGGERS, E. & KNAPP, S. 2009. Crystal Structure of the PIM2 Kinase in Complex with an Organoruthenium Inhibitor. *PLoS ONE,* 4, e7112.

BURGER, M., HAN, W., LAN, J., NISHIGUCHI, G., BELLAMACINA, C., LINDVAL, M., ATALLAH, G., DING, Y., MATHUR, M., MCBRIDE, C., BEANS, E., MULLER, K., TAMEZ, V., ZHANG, Y., HUH, K., FEUCHT, P., ZAVOROTINSKAYA, T., DAI, Y., HOLASH, J., CASTILLO, J., LANGOWSKI, J., WANG, Y., CHEN, M. & GARCIA, P. 2013. Structure Guided Optimization, in Vitro Activity, and in Vivo Activity of Pan-PIM Kinase Inhibitors. *ACS Med. Chem. Lett.*

CARLSON, B. A., DUBAY, M. M., SAUSVILLE, E. A., BRIZUELA, L. & WORLAND, P. J. 1996. Flavopiridol induces G1 arrest with inhibition of cyclin-dependent kinase (CDK) 2 and CDK4 in human breast carcinoma cells. *Cancer Research,* 56, 2973-2978.

CHAKRABARTI, P. & BHATTACHARYYA, R. 2007. Geometry of nonbonded interactions involving planar groups in proteins. *Progress in Biophysics and Molecular Biology,* 95, 83-137.

CHANG, Y., ADNANE, J., TRAIL, P., LEVY, J., HENDERSON, A., XUE, D., BORTOLON, E., ICHETOVKIN, M., CHEN, C., MCNABOLA, A., WILKIE, D., CARTER, C., TAYLOR, I., LYNCH, M. & WILHELM, S. 2007. Sorafenib (BAY 43-9006) inhibits tumor growth and vascularization and induces tumor apoptosis and hypoxia in RCC xenograft models. *Cancer chemotherapy and pharmacology,* 59, 561-574.

CHEN, L., REDKAR, S., BEARSS, D., WIERDA, W. & GANDHI, V. 2009. Pim kinase inhibitor, SGI-1776, induces apoptosis in chronic lymphocytic leukemia cells. *Blood,* 114, 4150-4157.

CHEN, L., REDKAR, S., TAVERNA, P., CORTES, J. & GANDHI, V. 2011. Mechanisms of cytotoxicity to Pim kinase inhibitor, SGI-1776, in acute myeloid leukemia. *Blood,* 118, 693-702.

CHOI, H. G., ZHANG, J., WEISBERG, E., GRIFFIN, J., SIM, T. & GRAY, N. 2012. Development of 'DFG-out' inhibitors of gatekeeper mutant kinases. *Bioorganic & Medicinal Chemistry Letters,* 22, 5297-5302.

CHOI, Y. L., SODA, M., YAMASHITA, Y., UENO, T., TAKASHIMA, J., NAKAJIMA, T., YATABE, Y., TAKEUCHI, K., HAMADA, T., HARUTA, H., ISHIKAWA, Y., KIMURA, H., MITSUDOMI, T., TANIO, Y. & MANO, H. 2010. EML4-ALK mutations in lung cancer that confer resistance to ALK inhibitors. *The New England journal of medicine,* 363, 1734-1739.

COOLS, J., STOVER, E., BOULTON, C., GOTLIB, J., LEGARE, R., AMARAL, S., CURLEY, D., DUCLOS, N., ROWAN, R., KUTOK, J., LEE, B., WILLIAMS, I., COUTRE, S., STONE, R., DEANGELO, D., MARYNEN, P., MANLEY, P., MEYER, T., FABBRO, D., NEUBERG, D., WEISBERG, E., GRIFFIN, J. & GILLILAND, G. 2003. PKC412 overcomes resistance to imatinib in a murine model of FIP1L1-PDGFRα-induced myeloproliferative disease. *Cancer cell,* 3, 459-469.

COWAN-JACOB, S., MÖBITZ, H. & FABBRO, D. 2009. Structural biology contributions to tyrosine kinase drug discovery. *Current Opinion in Cell Biology,* 21, 280-287.

CROOKS, G., HON, G., CHANDONIA, J.-M. & BRENNER, S. 2004. WebLogo: A Sequence Logo Generator. *Genome Research,* 14, 1188-1190.

CSERMELY, P., PALOTAI, R. & NUSSINOV, R. 2010. Induced fit, conformational selection and independent dynamic segments: an extended view of binding events. *Trends in Biochemical Sciences,* 35, 539-546.

CUYPERS, H. T., SELTEN, G., QUINT, W., ZIJLSTRA, M., MAANDAG, E. R., BOELENS, W., VAN WEZENBEEK, P., MELIEF, C. & BERNS, A. 1984. Murine leukemia virus-induced T-cell lymphomagenesis: integration of proviruses in a distinct chromosomal region. *Cell,* 37, 141-150.

DAR, A., LOPEZ, M. & SHOKAT, K. 2008. Small molecule recognition of c-Src via the Imatinib-binding conformation. *Chemistry & Biology,* 15, 1015-1022.

DAUB, H. 2010. Kinase inhibitors: Narrowing down the real targets. *Nature Chemical Biology,* 6, 249-250.

DAUB, H., SPECHT, K. & ULLRICH, A. 2004. Strategies to overcome resistance to targeted protein kinase inhibitors. *Nature reviews. Drug discovery,* 3, 1001-1010.

DAVIS, I., LEAVER-FAY, A., CHEN, V., BLOCK, J., KAPRAL, G., WANG, X., MURRAY, L., ARENDALL, B., SNOEYINK, J., RICHARDSON, J. & RICHARDSON, D. 2007. MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Research,* 35, W375-W383.

DE BONDT, H., ROSENBLATT, J., JANCARIK, J., JONES, H., MORGANT, D. & KIM, S.-H. 1993. Crystal structure of cyclin-dependent kinase 2. *Nature,* 363, 595-602.

DOUDOU, S., BURTON, N. & HENCHMAN, R. 2009. Standard Free Energy of Binding from a One-Dimensional Potential of Mean Force. *Journal of Chemical Theory and Computation,* 5, 909-918.

DOUDOU, S., SHARMA, R., HENCHMAN, R., SHEPPARD, D. & BURTON, N. 2010. Inhibitors of PIM-1 Kinase: A Computational Analysis of the Binding Free Energies of a Range of Imidazo [1,2-b] Pyridazines. *Journal of Chemical Information and Modeling,* 50, 368-379.

DRUKER, B., GUILHOT, F., O'BRIEN, S., GATHMANN, I., KANTARJIAN, H., GATTERMANN, N., DEININGER, M., SILVER, R., GOLDMAN, J., STONE, R., CERVANTES, F., HOCHHAUS, A., POWELL, B., GABRILOVE, J., ROUSSELOT, P., REIFFERS, J., CORNELISSEN, J., HUGHES, T., AGIS, H., FISCHER, T., VERHOEF, G., SHEPHERD, J., SAGLIO, G., GRATWOHL, A., NIELSEN, J., RADICH, J., SIMONSSON, B., TAYLOR, K., BACCARANI, M., SO, C., LETVAK, L. & LARSON, R. 2006. Five-Year Follow-up of Patients Receiving Imatinib for Chronic Myeloid Leukemia. *New England Journal of Medicine,* 355, 2408-2417.

DRUKER, B. J., TALPAZ, M., RESTA, D. J., PENG, B., BUCHDUNGER, E., FORD, J. M., LYDON, N. B., KANTARJIAN, H., CAPDEVILLE, R., OHNO-JONES, S. & SAWYERS, C. L. 2001. Efficacy and safety of a specific inhibitor of the BCR-ABL tyrosine kinase in chronic myeloid leukemia. *The New England journal of medicine,* 344, 1031-1037.

ECHALIER, A., ENDICOTT, J. & NOBLE, M. 2010. Recent developments in cyclin-dependent kinase biochemical and structural studies. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics,* 1804, 511-519.

EDGAR, R. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic acids research,* 32, 1792-1797.

ELLEDGE, S. J. & HARPER, J. W. 1998. The role of protein stability in the cell cycle and cancer. *Biochimica et biophysica acta,* 1377.

EMERY, C., VIJAYENDRAN, K., ZIPSER, M., SAWYER, A., NIU, L., KIM, J., HATTON, C., CHOPRA, R., OBERHOLZER, P., KARPOVA, M., MACCONAILL, L., ZHANG, J., GRAY, N., SELLERS, W., DUMMER, R. & GARRAWAY, L. 2009. MEK1 mutations confer resistance to MEK and B-RAF inhibition. *Proceedings of the National Academy of Sciences*.

EMSLEY, P., LOHKAMP, B., SCOTT, W. G. & COWTAN, K. 2010. Features and development of Coot. *Acta Crystallographica Section D,* 66, 486-501.

ENDICOTT, J., NOBLE, M. & JOHNSON, L. 2012. The structural basis for control of eukaryotic protein kinases. *Annual review of biochemistry,* 81, 587-613.

ESWARAN, J., BERNAD, A., LIGOS, J., GUINEA, B., DEBRECZENI, J., SOBOTT, F., PARKER, S., NAJMANOVICH, R., TURK, B. & KNAPP, S. 2008. Structure of the human protein kinase MPSK1 reveals an atypical activation loop architecture. *Structure (London, England : 1993),* 16, 115-124.

EVANS, P. 2006. Scaling and assessment of data quality. *Acta crystallographica. Section D, Biological crystallography,* 62, 72-82.

FABBRO, D., COWAN-JACOB, S., MÖBITZ, H. & MARTINY-BARON, G. 2012. Targeting cancer with small-molecular-weight kinase inhibitors. *Methods in molecular biology (Clifton, N.J.),* 795, 1-34.

FEDOROV, O., MULLER, S. & KNAPP, S. 2010. The (un)targeted cancer kinome. *Nature Chemical Biology,* 6, 166-169.

FELSENSTEIN, J. 1989. PHYLIP - Phylogeny Inference Package (Version 3.2). *Cladistics,* 5, 164-166.

FORBES, S., BINDAL, N., BAMFORD, S., COLE, C., KOK, C. Y., BEARE, D., JIA, M., SHEPHERD, R., LEUNG, K., MENZIES, A., TEAGUE, J., CAMPBELL, P., STRATTON, M. & FUTREAL, A. 2011. COSMIC: mining complete cancer genomes in the Catalogue of Somatic Mutations in Cancer. *Nucleic Acids Research,* 39, D945-D950.

FURET, P., BOLD, G., MEYER, T., ROESEL, J. & GUAGNANO, V. 2006. Aromatic interactions with phenylalanine 691 and cysteine 828: a concept for FMS-like tyrosine kinase-3 inhibition. Application to the discovery of a new class of potential antileukemia agents. *Journal of Medicinal Chemistry,* 49, 4451-4454.

GARSKE, A., PETERS, U., CORTESI, A., PEREZ, J. & SHOKAT, K. 2011a. Chemical genetic strategy for targeting protein kinases based on covalent complementarity. *Proceedings of the National Academy of Sciences,* 108, 15046-15052.

GARSKE, A. L., PETERS, U., CORTESI, A. T., PEREZ, J. L. & SHOKAT, K. M. 2011b. Chemical genetic strategy for targeting protein kinases based on covalent complementarity. *Proc Natl Acad Sci U S A,* 108, 15046-52.

GASTEIGER, E., GATTIKER, A., HOOGLAND, C., IVANYI, I., APPEL, R. & BAIROCH, A. 2003. ExPASy: The proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Research,* 31, 3784-3788.

GEOGHEGAN, K. F., DIXON, H. B., ROSNER, P. J., HOTH, L. R., LANZETTI, A. J., BORZILLERI, K. A., MARR, E. S., PEZZULLO, L. H., MARTIN, L. B., LEMOTTE, P. K., MCCOLL, A. S., KAMATH, A. V. & STROH, J. G. 1999. Spontaneous alpha-N-6-phosphogluconoylation of a "His tag" in Escherichia coli: the cause of extra mass of 258 or 178 Da in fusion proteins. *Analytical biochemistry,* 267, 169-184.

GIBBS, C. S. & ZOLLER, M. J. 1991. Rational scanning mutagenesis of a protein kinase identifies functional regions involved in catalysis and substrate interactions. *Journal of Biological Chemistry,* 266, 8923-8931.

GILEADI, O., BURGESS-BROWN, N., COLEBROOK, S., BERRIDGE, G., SAVITSKY, P., SMEE, C., LOPPNAU, P., JOHANSSON, C., SALAH, E. & PANTIC, N. 2008. High throughput production of recombinant human proteins for crystallography. *Methods in molecular biology (Clifton, N.J.),* 426, 221-246.

GOHDA, K. & HAKOSHIMA, T. 2008. A molecular mechanism of P-loop pliability of Rho-kinase investigated by molecular dynamic simulation. *Journal of computer-aided molecular design,* 22, 789-797.

GÓMEZ-ABAD, C., PISONERO, H., BLANCO-APARICIO, C., RONCADOR, G., GONZÁLEZ-MENCHÉN, A., MARTINEZ-CLIMENT, J., MATA, E., RODRÍGUEZ, M., MUÑOZ-GONZÁLEZ, G., SÁNCHEZ-BEATO, M., LEAL, J., BISCHOFF, J. & PIRIS, M. 2011. PIM2 inhibition as a rational therapeutic approach in B-cell lymphoma. *Blood,* 118, 5517-5527.

GORRE, M. E., MOHAMMED, M., ELLWOOD, K., HSU, N., PAQUETTE, R., RAO, P. N. & SAWYERS, C. L. 2001. Clinical resistance to STI-571 cancer therapy caused by BCR-ABL gene mutation or amplification. *Science (New York, N.Y.),* 293, 876-880.

GRANT, B., GORFE, A. & MCCAMMON, A. 2009. Ras Conformational Switching: Simulating Nucleotide-Dependent Conformational Transitions with Accelerated Molecular Dynamics. *PLoS Comput Biol,* 5, e1000325.

GREENMAN, C., STEPHENS, P., SMITH, R., DALGLIESH, G., HUNTER, C., BIGNELL, G., DAVIES, H., TEAGUE, J., BUTLER, A., STEVENS, C., EDKINS, S., O'MEARA, S., VASTRIK, I., SCHMIDT, E., AVIS, T., BARTHORPE, S., BHAMRA, G., BUCK, G., CHOUDHURY, B., CLEMENTS, J., COLE, J., DICKS, E., FORBES, S., GRAY, K., HALLIDAY, K., HARRISON, R., HILLS, K., HINTON, J., JENKINSON, A., JONES, D., MENZIES, A., MIRONENKO, T., PERRY, J., RAINE, K., RICHARDSON, D., SHEPHERD, R., SMALL, A., TOFTS, C., VARIAN, J., WEBB, T., WEST, S., WIDAA, S., YATES, A., CAHILL, D., LOUIS, D., GOLDSTRAW, P., NICHOLSON, A., BRASSEUR, F., LOOIJENGA, L., WEBER, B., CHIEW, Y.-E., DEFAZIO, A., GREAVES, M., GREEN, A., CAMPBELL, P., BIRNEY, E., EASTON, D., CHENEVIX-TRENCH, G., TAN, M.-H., KHOO, S. K., TEH, B. T., YUEN, S. T., LEUNG, S. Y., WOOSTER, R., FUTREAL, A. & STRATTON, M. 2007. Patterns of somatic mutation in human cancer genomes. *Nature,* 446, 153-158.

GU, Y., ROSENBLATT, J. & MORGAN, D. O. 1992. Cell cycle regulation of CDK2 activity by phosphorylation of Thr160 and Tyr15. *The EMBO journal,* 11, 3995-4005.

GUIMARÃES, C., RAI, B., MUNCHHOF, M., LIU, S., WANG, J., BHATTACHARYA, S. & BUCKBINDER, L. 2011. Understanding the impact of the P-loop conformation on kinase selectivity. *Journal of chemical information and modeling,* 51, 1199-1204.

HAMER, R., LUO, Q., ARMITAGE, J., REINERT, G. & DEANE, C. 2010. i-Patch: Interprotein contact prediction using local network information. *Proteins,* 78, 2781-2797.

HANKS, S. K., QUINN, A. M. & HUNTER, T. 1988. The protein kinase family: conserved features and deduced phylogeny of the catalytic domains. *Science (New York, N.Y.),* 241, 42-52.

HARI, S., MERRITT, E. & MALY, D. 2013. Sequence determinants of a specific inactive protein kinase conformation. *Chemistry & Biology,* 20, 806-815.

HARPER, J. W. & ELLEDGE, S. J. 1998. The role of Cdk7 in CAK function, a retro-retrospective. *Genes & Development,* 12, 285-289.

HAU, J. C., FONTANA, P., ZIMMERMANN, C., DE POVER, A., ERDMANN, D. & CHÈNE, P. 2011. Leveraging the contribution of thermodynamics in drug discovery with the help of fluorescence-based thermal shift assays. *Journal of biomolecular screening,* 16, 552-556.

HECKMAN, K. & PEASE, L. 2007. Gene splicing and mutagenesis by PCR-driven overlap extension. *Nat. Protocols,* 2, 924-932.

HOLDGATE, G. & WARD, W. 2005. Measurements of binding thermodynamics in drug discovery. *Drug Discovery Today,* 10, 1543-1550.

HUBBARD, S. R. 1997. Crystal structure of the activated insulin receptor tyrosine kinase in complex with peptide substrate and ATP analog. *The EMBO journal,* 16, 5572-5581.

HUSE, M. & KURIYAN, J. 2002. The Conformational Plasticity of Protein Kinases. *Cell,* 109, 275-282.

IYER, G., GARROD, S., WOODS, V. & TAYLOR, S. 2005. Catalytic Independent Functions of a Protein Kinase as Revealed by a Kinase-dead Mutant: Study of the Lys72His Mutant of cAMP-dependent Kinase. *Journal of Molecular Biology,* 351, 1110-1122.

JOHNSON, L. 2009. Protein kinase inhibitors: contributions from structure to clinical compounds. *Quarterly Reviews of Biophysics,* 42, 1-40.

JOHNSON, L., DE MOLINER, E., BROWN, N., SONG, H., BARFORD, D., ENDICOTT, J. & NOBLE, M. 2002. Structural studies with inhibitors of the cell cycle regulatory kinase cyclin-dependent protein kinase 2. *Pharmacology & therapeutics,* 93, 113-124.

JURA, N., ZHANG, X., ENDRES, N., SEELIGER, M., SCHINDLER, T. & KURIYAN, J. 2011. Catalytic control in the EGF receptor and its connection to general kinase regulatory mechanisms. *Molecular Cell,* 42, 9-22.

KALDIS, P. 1999. The cdk-activating kinase (CAK): from yeast to mammals. *Cellular and molecular life sciences : CMLS,* 55, 284-296.

KAR, G., KESKIN, O., GURSOY, A. & NUSSINOV, R. 2010. Allostery and population shift in drug discovery. *Current Opinion in Pharmacology,* 10, 715-722.

KEATES, T., COOPER, C., SAVITSKY, P., ALLERSTON, C., PHILLIPS, C., HAMMARSTRÖM, M., DAGA, N., BERRIDGE, G., MAHAJAN, P., BURGESS-BROWN, N., MÜLLER, S., GRÄSLUND, S. & GILEADI, O. 2012. Expressing the human proteome for affinity proteomics: optimising expression of soluble protein domains and in vivo biotinylation. *New biotechnology,* 29, 515-525.

KÉRI, G., ŐRFI, L. & NÉMETH, G. 2011. Kinase Inhibitors in Signal Transduction Therapy. *Protein Kinases as Drug Targets*, 115-144.

KIM, K. M., YI, E. C., BAKER, D. & ZHANG, K. Y. 2001. Post-translational modification of the N-terminal His tag interferes with the crystallization of the wild-type and mutant SH3 domains from chicken src tyrosine kinase. *Acta crystallographica. Section D, Biological crystallography,* 57, 759-762.

KOBAYASHI, S., BOGGON, T., DAYARAM, T., JÄNNE, P., KOCHER, O., MEYERSON, M., JOHNSON, B., ECK, M., TENEN, D. & HALMOS, B. 2005. EGFR mutation and resistance of non-small-cell lung cancer to gefitinib. *The New England journal of medicine,* 352, 786-792.

KONG, L., LOVELL, P., HEGER, A., MELLO, C. & PONTING, C. 2010. Accelerated evolution of PAK3- and PIM1-like kinase gene families in the zebra finch, Taeniopygia guttata. *Molecular biology and evolution,* 27, 1923-1934.

KORNEV, A., HASTE, N., TAYLOR, S. & EYCK, L. T. 2006. Surface comparison of active and inactive protein kinases identifies a conserved activation mechanism. *Proceedings of the National Academy of Sciences of the United States of America,* 103, 17783-17788.

KORNEV, A. & TAYLOR, S. 2010. Defining the conserved internal architecture of a protein kinase. *Biochimica et biophysica acta,* 1804, 440-444.

KORNEV, A., TAYLOR, S. & TEN EYCK, L. 2008. A helix scaffold for the assembly of active protein kinases. *Proceedings of the National Academy of Sciences,* 105, 14377-14382.

KUMAR, A., MANDIYAN, V., SUZUKI, Y., ZHANG, C., RICE, J., TSAI, J., ARTIS, D., IBRAHIM, P. & BREMER, R. 2005. Crystal Structures of Proto-oncogene Kinase Pim1: A Target of Aberrant Somatic Hypermutations in Diffuse Large Cell Lymphoma. *Journal of Molecular Biology,* 348, 183-193.

LAIRD, P. W., VAN DER LUGT, N. M., CLARKE, A., DOMEN, J., LINDERS, K., MCWHIR, J., BERNS, A. & HOOPER, M. 1993. In vivo analysis of Pim-1 deficiency. *Nucleic Acids Research,* 21, 4750-4755.

LAROCHELLE, S., CHEN, J., KNIGHTS, R., PANDUR, J., MORCILLO, P., ERDJUMENT-BROMAGE, H., TEMPST, P., SUTER, B. & FISHER, R. P. 2001. T-loop phosphorylation stabilizes the CDK7-cyclin H-MAT1 complex in vivo and regulates its CTD kinase activity. *The EMBO journal,* 20, 3749-3759.

LASKOWSKI, R. & SWINDELLS, M. 2011. LigPlot+: multiple ligand-protein interaction diagrams for drug discovery. *Journal of Chemical Information and Modeling,* 51, 2778-2786.

LEAVITT, S. & FREIRE, E. 2001. Direct measurement of protein binding energetics by isothermal titration calorimetry. *Current opinion in structural biology,* 11, 560-566.

LETUNIC, I. & BORK, P. 2007. Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation. *Bioinformatics (Oxford, England),* 23, 127-128.

LEVINSON, N. M., KUCHMENT, O., SHEN, K., YOUNG, M. A., KOLDOBSKIY, M., KARPLUS, M., COLE, P. A. & KURIYAN, J. 2006. A Src-like inactive conformation in the abl tyrosine kinase domain. *PLoS Biol,* 4, e144.

LI, W. & GODZIK, A. 2006. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics,* 22, 1658-1659.

LICHTARGE, O., BOURNE, H. R. & COHEN, F. E. 1996. An evolutionary trace method defines binding surfaces common to protein families. *Journal of Molecular Biology,* 257, 342-358.

LIU, Y. & GRAY, N. 2006. Rational design of inhibitors that bind to inactive kinase conformations. *Nature Chemical Biology,* 2, 358-364.

LOPEZ-RAMOS, M., PRUDENT, R., MOUCADEL, V., SAUTEL, C., BARETTE, C., LAFANECHÈRE, L., MOUAWAD, L., GRIERSON, D., SCHMIDT, F., FLORENT, J.-C., FILIPPAKOPOULOS, P., BULLOCK, A., KNAPP, S., REISER, J.-B. & COCHET, C. 2010. New potent dual inhibitors of CK2 and Pim kinases: discovery and structural insights. *FASEB J.,* 24, 3171-3185.

LORI, C., LANTELLA, A., PASQUO, A., ALEXANDER, L., KNAPP, S., CHIARALUCE, R. & CONSALVI, V. 2013. Effect of Single Amino Acid Substitution Observed in Cancer on Pim-1 Kinase Thermodynamic Stability and Structure. *PLoS ONE,* 8, e64824.

LU, J., ZAVOROTINSKAYA, T., DAI, Y., NIU, X.-H., CASTILLO, J., SIM, J., YU, J., WANG, Y., LANGOWSKI, J., HOLASH, J., SHANNON, K. & GARCIA, P. 2013. Pim2 is required for maintaining multiple myeloma cell growth through modulating TSC2 phosphorylation. *Blood,* 122, 1610-1620.

MA, B., KUMAR, S., TSAI, C.-J. & NUSSINOV, R. 1999. Folding funnels and binding mechanisms. *Protein Engineering,* 12, 713-720.

MALUMBRES, M. & BARBACID, M. 2001. To cycle or not to cycle: a critical decision in cancer. *Nature reviews. Cancer,* 1, 222-231.

MALUMBRES, M. & BARBACID, M. 2005. Mammalian cyclin-dependent kinases. *Trends in Biochemical Sciences,* 30, 630-641.

MALUMBRES, M., HARLOW, E., HUNT, T., HUNTER, T., LAHTI, J., MANNING, G., MORGAN, D., TSAI, L.-H. & WOLGEMUTH, D. 2009. Cyclin-dependent kinases: a family portrait. *Nature cell biology,* 11, 1275-1276.

MANNING, G., WHYTE, D. B., MARTINEZ, R., HUNTER, T. & SUDARSANAM, S. 2002. The Protein Kinase Complement of the Human Genome. *Science,* 298, 1912-1934.

MCCOY, A., GROSSE-KUNSTLEVE, R., ADAMS, P., WINN, M., STORONI, L. & READ, R. 2007. Phaser crystallographic software. *Journal of applied crystallography,* 40, 658-674.

MIKKERS, H., NAWIJN, M., ALLEN, J., BROUWERS, C., VERHOEVEN, E., JONKERS, J. & BERNS, A. 2004. Mice Deficient for All PIM Kinases Display Reduced Body Size and Impaired Responses to Hematopoietic Growth Factors. *Mol. Cell. Biol.,* 24, 6104-6115.

MÖBITZ, H. & FABBRO, D. 2012. Conformational Bias: A key concept for protein kinase inhibition. *European Pharmaceutical Review.*

MOORE, J. 2013. In the wrong place at the wrong time: does cyclin mislocalization drive oncogenic transformation? *Nature reviews. Cancer,* 13, 201-208.

MORGAN, D. 2006. *The Cell Cycle*, Oxford University Press.

MUMENTHALER, S., NG, P., HODGE, A., BEARSS, D., BERK, G., KANEKAL, S., REDKAR, S., TAVERNA, P., AGUS, D. & JAIN, A. 2009. Pharmacologic inhibition of Pim kinases alters prostate cancer cell growth and resensitizes chemoresistant cells to taxanes. *Molecular cancer therapeutics,* 8, 2882-2893.

MURSHUDOV, G. N., VAGIN, A. A. & DODSON, E. J. 1997. Refinement of macromolecular structures by the maximum-likelihood method. *Acta crystallographica. Section D, Biological crystallography,* 53, 240-255.

NAGAR, B., BORNMANN, W., PELLICENA, P., SCHINDLER, T., VEACH, D., MILLER, T., CLARKSON, B. & KURIYAN, J. 2002. Crystal Structures of the Kinase Domain of c-Abl in Complex with the Small Molecule Inhibitors PD173955 and Imatinib (STI-571). *Cancer Research,* 62, 4236-4243.

NAWIJN, M., ALENDAR, A. & BERNS, A. 2011. For better or for worse: the role of Pim oncogenes in tumorigenesis. *Nature reviews. Cancer,* 11, 23-34.

NIESEN, F., BERGLUND, H. & VEDADI, M. 2007. The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nat. Protocols,* 2, 2212-2221.

NOBLE, M., ENDICOTT, J. & JOHNSON, L. 2004. Protein Kinase Inhibitors: Insights into Drug Design from Structure. *Science,* 303, 1800-1805.

ORTEGA, S., PRIETO, I., ODAJIMA, J., MARTÍN, A., DUBUS, P., SOTILLO, R., BARBERO, J. L., MALUMBRES, M. & BARBACID, M. 2003. Cyclin-dependent kinase 2 is essential for meiosis but not for mitotic cell division in mice. *Nature genetics,* 35, 25-31.

POGACIC, V., BULLOCK, A., FEDOROV, O., FILIPPAKOPOULOS, P., GASSER, C., BIONDI, A., MEYER-MONARD, S., KNAPP, S. & SCHWALLER, J. 2007. Structural Analysis Identifies Imidazo[1,2-b]Pyridazines as PIM Kinase Inhibitors with In vitro Antileukemic Activity. *Cancer Res,* 67, 6916-6924.

QIAN, K., WANG, L., HICKEY, E., STUDTS, J., BARRINGER, K., PENG, C., KRONKAITIS, A., LI, J., WHITE, A., MISCHE, S. & FARMER, B. 2005. Structural Basis of Constitutive Activity and a Unique Nucleotide Binding Mode of Human Pim-1 Kinase. *Journal of Biological Chemistry,* 280, 6130-6137.

REIKOFSKI, J. & TAO, B. Y. 1992. Polymerase chain reaction (PCR) techniques for site-directed mutagenesis. *Biotechnology advances,* 10, 535-547.

ROSENBLATT, J., DE BONDT, H., JANCARIK, J., MORGAN, D. & KIM, S.-H. 1993. Purification and Crystallization of Human Cyclin-dependent Kinase 2. *Journal of Molecular Biology,* 230, 1317-1319.

SALONEN, L., ELLERMANN, M. & DIEDERICH, F. 2011. Aromatic Rings in Chemical and Biological Recognition: Energetics and Structures. *Angew. Chem. Int. Ed.,* 50, 4808-4842.

SAVITSKY, P., BRAY, J., COOPER, C., MARSDEN, B., MAHAJAN, P., BURGESS-BROWN, N. & GILEADI, O. 2010. High-throughput production of human proteins for crystallization: the SGC experience. *Journal of structural biology,* 172, 3-13.

SCHINDLER, T., BORNMANN, W., PELLICENA, P., MILLER, W. T., CLARKSON, B. & KURIYAN, J. 2000. Structural mechanism for STI-571 inhibition of abelson tyrosine kinase. *Science.,* 289, 1938-1942.

SCHWARTZ, G. & SHAH, M. 2005. Targeting the cell cycle: a new approach to cancer therapy. *Journal of clinical oncology,* 23, 9408-9421.

SELTEN, G., CUYPERS, H. T., BOELENS, W., ROBANUS-MAANDAG, E., VERBEEK, J., DOMEN, J., VAN BEVEREN, C. & BERNS, A. 1986. The primary structure of the putative oncogene pim-1 shows extensive homology with protein kinases. *Cell,* 46, 603-611.

SHAN, Y., SEELIGER, M., EASTWOOD, M., FRANK, F., XU, H., MORTEN, DROR, R., KURIYAN, J. & SHAW, D. 2009. A conserved protonation-dependent switch controls drug binding in the Abl kinase. *Proceedings of the National Academy of Sciences,* 106, 139-144.

SHAY, K., WANG, Z., XING, P.-X., MCKENZIE, I. & MAGNUSON, N. 2005. Pim-1 Kinase Stability Is Regulated by Heat Shock Proteins and the Ubiquitin-Proteasome Pathway. *Molecular Cancer Research,* 3, 170-181.

SMITH, C., LASATER, E., ZHU, X., LIN, K., STEWART, W., DAMON, L., SALERNO, S. & SHAH, N. 2013. Activity of ponatinib against clinically-relevant AC220-resistant kinase domain mutants of FLT3-ITD. *Blood,* 121, 3165-3171.

SOUNDARARAJAN, M., ROOS, A., SAVITSKY, P., FILIPPAKOPOULOS, P., KETTENBACH, A., OLSEN, J., GERBER, S., ESWARAN, J., KNAPP, S. & ELKINS, J. 2013. Structures of Down Syndrome Kinases, DYRKs, Reveal Mechanisms of Kinase Activation and Substrate Recognition. *Structure,* 21, 986-996.

TAMBORINI, E., BONADIMAN, L., GRECO, A., ALBERTINI, V., NEGRI, T., GRONCHI, A., BERTULLI, R., COLECCHIA, M., CASALI, P., PIEROTTI, M. & PILOTTI, S. 2004. A new mutation in the KIT ATP pocket causes acquired resistance to imatinib in a gastrointestinal stromal tumor patient. *Gastroenterology,* 127, 294-299.

TAO, Z.-F., HASVOLD, L., LEVERSON, J., HAN, E., GUAN, R., JOHNSON, E., STOLL, V., STEWART, K., STAMPER, G., SONI, N., BOUSKA, J., LUO, Y., SOWIN, T., LIN, N.-H., GIRANDA, V., ROSENBERG, S. & PENNING, T. 2009. Discovery of 3H-Benzo[4,5]thieno[3,2-d]pyrimidin-4-ones as Potent, Highly Selective, and Orally Bioavailable Inhibitors of the Human Protooncogene Proviral Insertion Site in Moloney Murine Leukemia Virus (PIM) Kinases. *Journal of Medicinal Chemistry,* 52, 6621-6636.

TAYLOR, S. & KORNEV, A. 2011. Protein kinases: evolution of dynamic regulatory proteins. *Trends in biochemical sciences,* 36, 65-77.

TETSU, O. & MCCORMICK, F. 2003. Proliferation of cancer cells despite CDK2 inhibition. *Cancer cell,* 3, 233-245.

VEDADI, M., NIESEN, F., ALLALI-HASSANI, A., FEDOROV, O., FINERTY, P., WASNEY, G., YEUNG, R., ARROWSMITH, C., BALL, L., BERGLUND, H., HUI, R., MARSDEN, B., NORDLUND, P., SUNDSTROM, M., WEIGELT, J. & EDWARDS, A. 2006. Chemical screening methods to identify ligands that promote protein stability, protein crystallization, and structure determination. *Proceedings of the National Academy of Sciences,* 103, 15835-15840.

WANG, X., MAGNUSON, S., PASTOR, R., FAN, E., HU, H., TSUI, V., DENG, W., MURRAY, J., STEFFEK, M., WALLWEBER, H., MOFFAT, J., DRUMMOND, J., CHAN, G., HARSTAD, E. & EBENS, A. 2013. Discovery of novel pyrazolo[1,5-a]pyrimidines as potent pan-Pim inhibitors by structure- and property-based drug design. *Bioorganic & medicinal chemistry letters*.

WANG, Z., BHATTACHARYA, N., MIXTER, P., WEI, W., SEDIVY, J. & MAGNUSON, N. 2002. Phosphorylation of the cell cycle inhibitor p21Cip1/WAF1 by Pim-1 kinase. *Biochimica et biophysica acta,* 1593, 45-55.

WANG, Z., BHATTACHARYA, N., WEAVER, M., PETERSEN, K., MEYER, M., GAPTER, L. & MAGNUSON, N. S. 2001. Pim-1: a serine/threonine kinase with a role in cell survival, proliferation, differentiation and tumorigenesis. *Journal of veterinary science,* 2, 167-179.

WANG, Z., LIU, J., SUDOM, A., AYRES, M., LI, S., WESCHE, H., POWERS, J. & WALKER, N. 2006. Crystal structures of IRAK-4 kinase in complex with inhibitors: a serine/threonine kinase with tyrosine as a gatekeeper. *Structure (London, England : 1993),* 14, 1835-1844.

WARMUTH, M., SIMON, N., MITINA, O., MATHES, R., FABBRO, D., MANLEY, P., BUCHDUNGER, E., FORSTER, K., MOAREFI, I. & HALLEK, M. 2003. Dual-specific Src and Abl kinase inhibitors, PP1 and CGP76030, inhibit growth and survival of cells expressing imatinib mesylate-resistant Bcr-Abl kinases. *Blood,* 101, 664-672.

WATERHOUSE, A., PROCTER, J., MARTIN, D., CLAMP, M. & BARTON, G. 2009. Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics,* 25, 1189-1191.

WELBURN, J., TUCKER, J., JOHNSON, T., LINDERT, L., MORGAN, M., WILLIS, A., NOBLE, M. & ENDICOTT, J. 2007. How tyrosine 15 phosphorylation inhibits the activity of cyclin-dependent kinase 2-cyclin A. *The Journal of biological chemistry,* 282, 3173-3181.

XU, M., SHEPPARD, K. A., PENG, C. Y., YEE, A. S. & PIWNICA-WORMS, H. 1994. Cyclin A/CDK2 binds directly to E2F-1 and inhibits the DNA-binding activity of E2F-1/DP-1 by phosphorylation. *Molecular and cellular biology,* 14, 8420-8431.

YANG, Z. R., THOMSON, R., MCNEIL, P. & ESNOUF, R. 2005. RONN: the bio-basis function neural network technique applied to the detection of natively disordered regions in proteins. *Bioinformatics (Oxford, England),* 21, 3369-3376.

YUAN, H.-Y., CHIOU, J.-J., TSENG, W.-H., LIU, C.-H., LIU, C.-K., LIN, Y.-J., WANG, H.-H., YAO, A., CHEN, Y.-T. & HSU, C.-N. 2006. FASTSNP: an always up-to-date and extendable service for SNP function analysis and prioritization. *Nucleic Acids Research,* 34, W635-W641.

ZHANG, J., YANG, P. & GRAY, N. 2009. Targeting cancer with small molecule kinase inhibitors. *Nat Rev Cancer,* 9, 28-39.

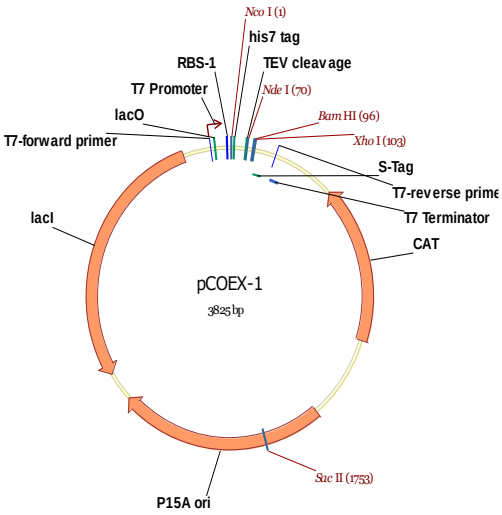
ZHANG, Z. & MARSHALL, A. G. 1998. A universal algorithm for fast and automated charge state deconvolution of electrospray mass-to-charge ratio spectra. *Journal of the American Society for Mass Spectrometry,* 9, 225-233.

ZHANG, Z., MITEVA, M., WANG, L. & ALEXOV, E. 2012. Analyzing effects of naturally occurring missense mutations. *Computational and mathematical methods in medicine,* 2012.

ZUCCOTTO, F., ARDINI, E., CASALE, E. & ANGIOLINI, M. 2010. Through the "gatekeeper door": exploiting the active kinase conformation. *Journal of medicinal chemistry,* 53, 2681-2694.

Appendices

Appendix A. Plasmids used for *E.coli* expression.

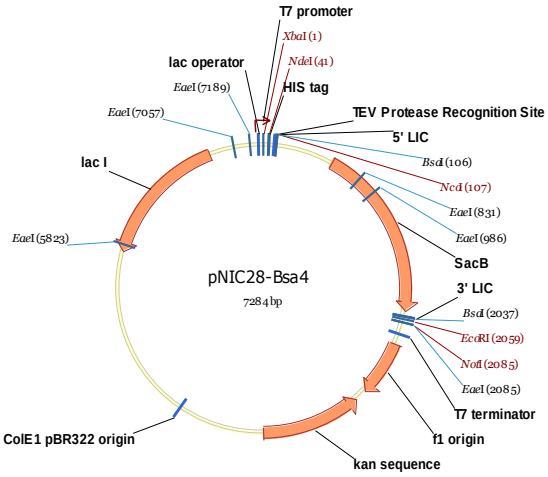


Accession: N/A

N-tag:
MGSSHHHHHHHSSGRENLYFQ*GHM

C-tag: None

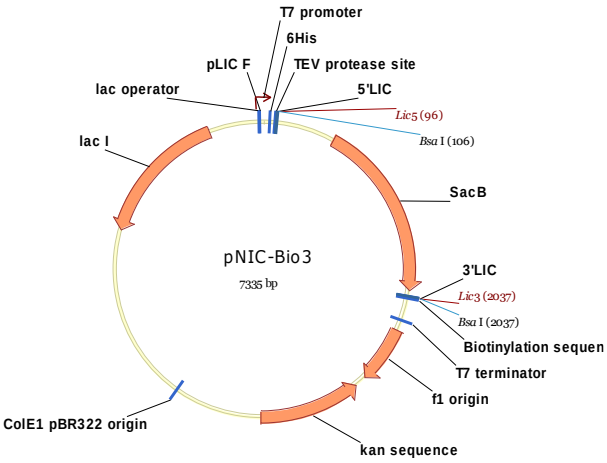
Proteins: PIM1, PIM1 mutants



Accession: EF198106
N-tag:
MHHHHHHSSGVDLGTENLYFQ*SM

C-tag: None

Proteins: CDK2 panel, PIM2, PIM3



Accession: JN792439
N-tag:
MHHHHHHSSGVDLGTENLYFQ*SM

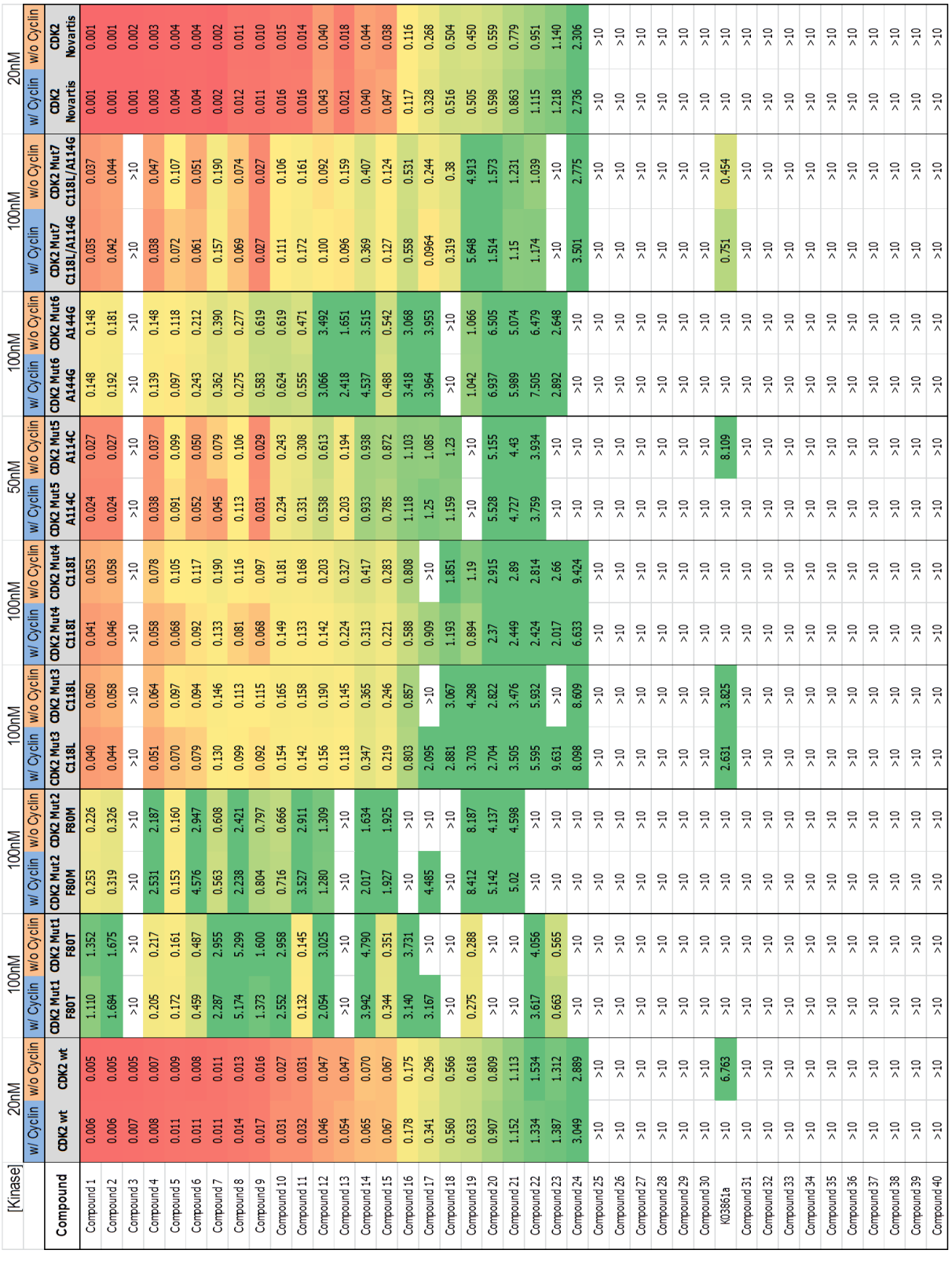
C-tag:
GSKGGYGLNDIFEAQ(K^)IEWHE

Proteins: CDK2

* TEV protease cleavage site
K^ biotin acceptor lysine

Appendix B. Comparison of the IC_{50} values (µM) for CDK2 variants obtained during enzymology studies.

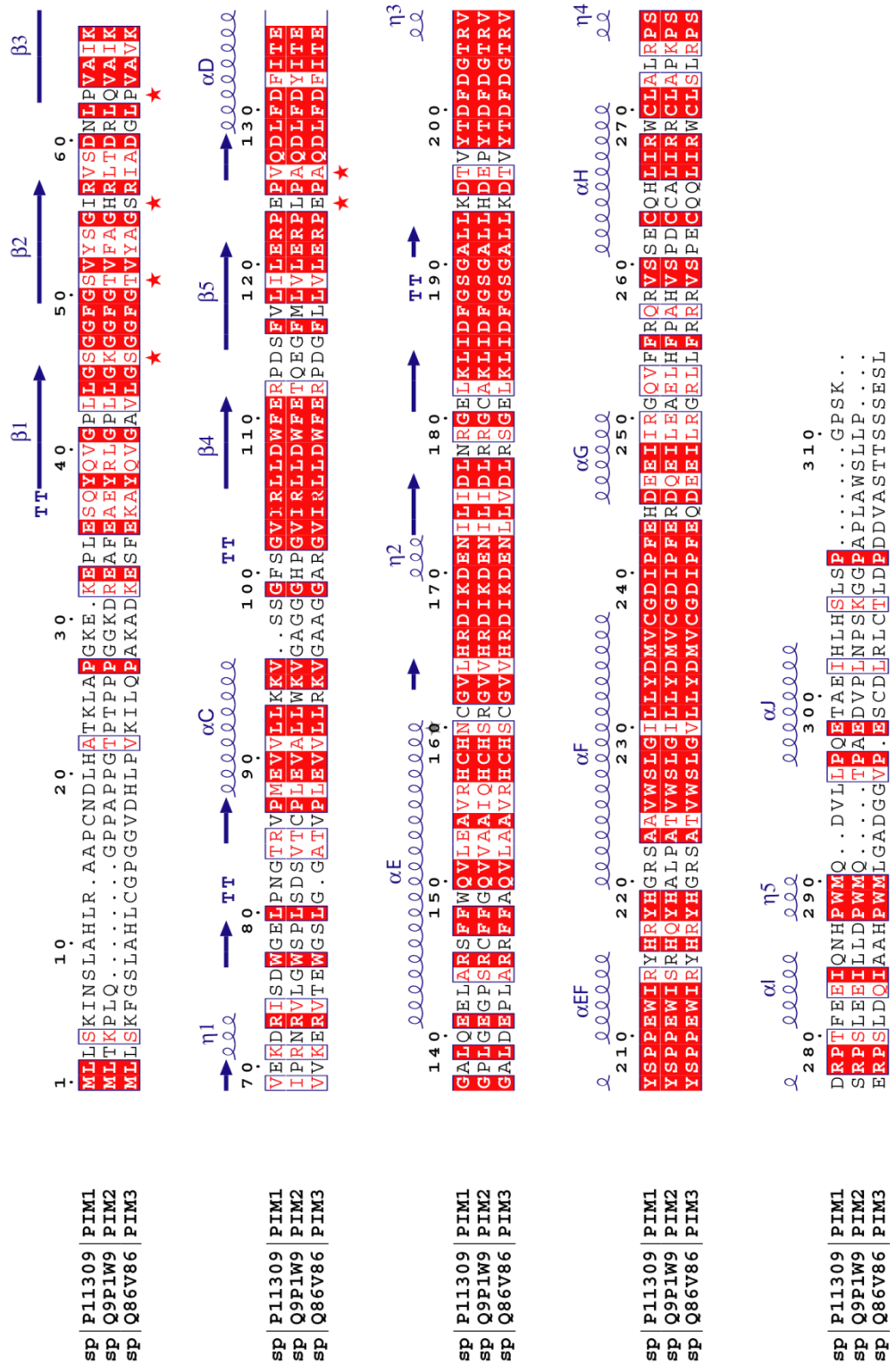
For each protein two starting conditions were tested: pre-incubated with cyclin A for 1 hour and adding cyclin A immediately before reaction begins. Red (low) values indicate stronger and green (higher) weaker binding. No clear differences in results were observed.



[Kinase]	20nM		100nM			100nM			100nM		50nM		100nM		100nM		20nM	
	w/ Cyclin	w/o Cyclin	w/ Cyclin	w/o Cyclin	w/ Cyclin	w/o Cyclin	w/ Cyclin	w/o Cyclin	w/ Cyclin	w/o Cyclin	w/ Cyclin	w/o Cyclin	w/ Cyclin	w/o Cyclin	w/ Cyclin	w/o Cyclin	w/ Cyclin	w/o Cyclin
Compound	**CDK2 wt**	**CDK2 wt**	**CDK2 Mut1 F80T**	**CDK2 Mut1 F80T**	**CDK2 Mut2 F80M**	**CDK2 Mut2 F80M**	**CDK2 Mut3 C118L**	**CDK2 Mut3 C118L**	**CDK2 Mut4 C118I**	**CDK2 Mut4 C118I**	**CDK2 Mut5 A114C**	**CDK2 Mut5 A114C**	**CDK2 Mut6 A144G**	**CDK2 Mut6 A144G**	**CDK2 Mut7 C118L/A114G**	**CDK2 Mut7 C118L/A114G**	**CDK2 Novartis**	**CDK2 Novartis**
Compound 1	0.006	0.005	1.110	1.352	0.253	0.226	0.040	0.050	0.041	0.053	0.024	0.027	0.148	0.148	0.035	0.037	0.001	0.001
Compound 2	0.006	0.005	1.684	1.675	0.319	0.326	0.044	0.058	0.046	0.058	0.024	0.027	0.192	0.181	0.042	0.044	0.001	0.001
Compound 3	0.007	0.005	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	0.001	0.002
Compound 4	0.008	0.007	0.205	0.217	2.531	2.187	0.051	0.064	0.058	0.078	0.038	0.037	0.139	0.148	0.038	0.047	0.003	0.003
Compound 5	0.011	0.009	0.172	0.161	0.153	0.160	0.070	0.097	0.068	0.105	0.091	0.099	0.097	0.118	0.072	0.107	0.004	0.004
Compound 6	0.011	0.008	0.459	0.487	4.576	2.947	0.079	0.094	0.092	0.117	0.052	0.050	0.243	0.212	0.061	0.051	0.004	0.004
Compound 7	0.011	0.011	2.287	2.955	0.563	0.608	0.130	0.146	0.133	0.190	0.045	0.079	0.362	0.390	0.157	0.190	0.002	0.002
Compound 8	0.014	0.013	5.174	5.299	2.238	2.421	0.099	0.113	0.081	0.116	0.113	0.106	0.275	0.277	0.069	0.074	0.012	0.011
Compound 9	0.017	0.016	1.373	1.600	0.804	0.797	0.092	0.115	0.068	0.097	0.031	0.029	0.583	0.619	0.027	0.027	0.011	0.010
Compound 10	0.031	0.027	2.552	2.958	0.716	0.666	0.154	0.165	0.149	0.181	0.234	0.243	0.624	0.619	0.111	0.106	0.016	0.015
Compound 11	0.032	0.031	0.132	0.145	3.527	2.911	0.142	0.158	0.133	0.168	0.331	0.308	0.555	0.471	0.172	0.161	0.016	0.014
Compound 12	0.046	0.047	2.054	3.025	1.280	1.309	0.156	0.190	0.142	0.203	0.538	0.613	3.066	3.492	0.100	0.092	0.043	0.040
Compound 13	0.054	0.047	>10	>10	>10	>10	0.118	0.145	0.224	0.327	0.203	0.194	2.418	1.651	0.096	0.159	0.021	0.018
Compound 14	0.065	0.070	3.942	4.790	2.017	1.634	0.347	0.365	0.313	0.417	0.933	0.938	4.537	3.515	0.369	0.407	0.040	0.044
Compound 15	0.067	0.067	0.344	0.351	1.927	1.925	0.219	0.246	0.221	0.283	0.785	0.872	0.488	0.542	0.127	0.124	0.047	0.038
Compound 16	0.178	0.175	3.140	3.731	>10	>10	0.803	0.857	0.588	0.808	1.118	1.103	3.418	3.068	0.558	0.531	0.117	0.116
Compound 17	0.341	0.296	3.167	>10	4.485	>10	2.095	>10	0.909	>10	1.25	1.085	3.964	3.953	0.0964	0.244	0.328	0.268
Compound 18	0.560	0.566	>10	>10	>10	>10	2.881	3.067	1.193	1.851	1.159	1.23	>10	>10	0.319	0.38	0.516	0.504
Compound 19	0.633	0.618	0.275	0.288	8.412	8.187	3.703	4.298	0.894	1.19	>10	>10	1.042	1.066	5.648	4.913	0.505	0.450
Compound 20	0.907	0.809	>10	>10	5.142	4.137	2.704	2.822	2.37	2.915	5.528	5.155	6.937	6.505	1.514	1.573	0.598	0.559
Compound 21	1.152	1.113	>10	>10	5.02	4.598	3.505	3.476	2.449	2.89	4.727	4.43	5.989	5.074	1.15	1.231	0.863	0.779
Compound 22	1.334	1.534	3.617	4.056	>10	>10	5.595	5.932	2.424	2.814	3.759	3.934	7.505	6.479	1.174	1.039	1.115	0.951
Compound 23	1.387	1.312	0.663	0.565	>10	>10	9.631	>10	2.017	2.66	>10	>10	2.892	2.648	>10	>10	1.218	1.140
Compound 24	3.049	2.889	>10	>10	>10	>10	8.098	8.609	6.633	9.424	>10	>10	>10	>10	3.501	2.775	2.736	2.306
Compound 25	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10
Compound 26	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10
Compound 27	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10
Compound 28	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10
Compound 29	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10
Compound 30	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10
K03861a	>10	6.763	>10	>10	>10	>10	2.631	3.825	>10	>10	>10	8.109	>10	>10	0.751	0.454	>10	>10
Compound 31	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10
Compound 32	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10
Compound 33	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10
Compound 34	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10
Compound 35	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10
Compound 36	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10
Compound 37	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10
Compound 38	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10
Compound 39	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10
Compound 40	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10	>10

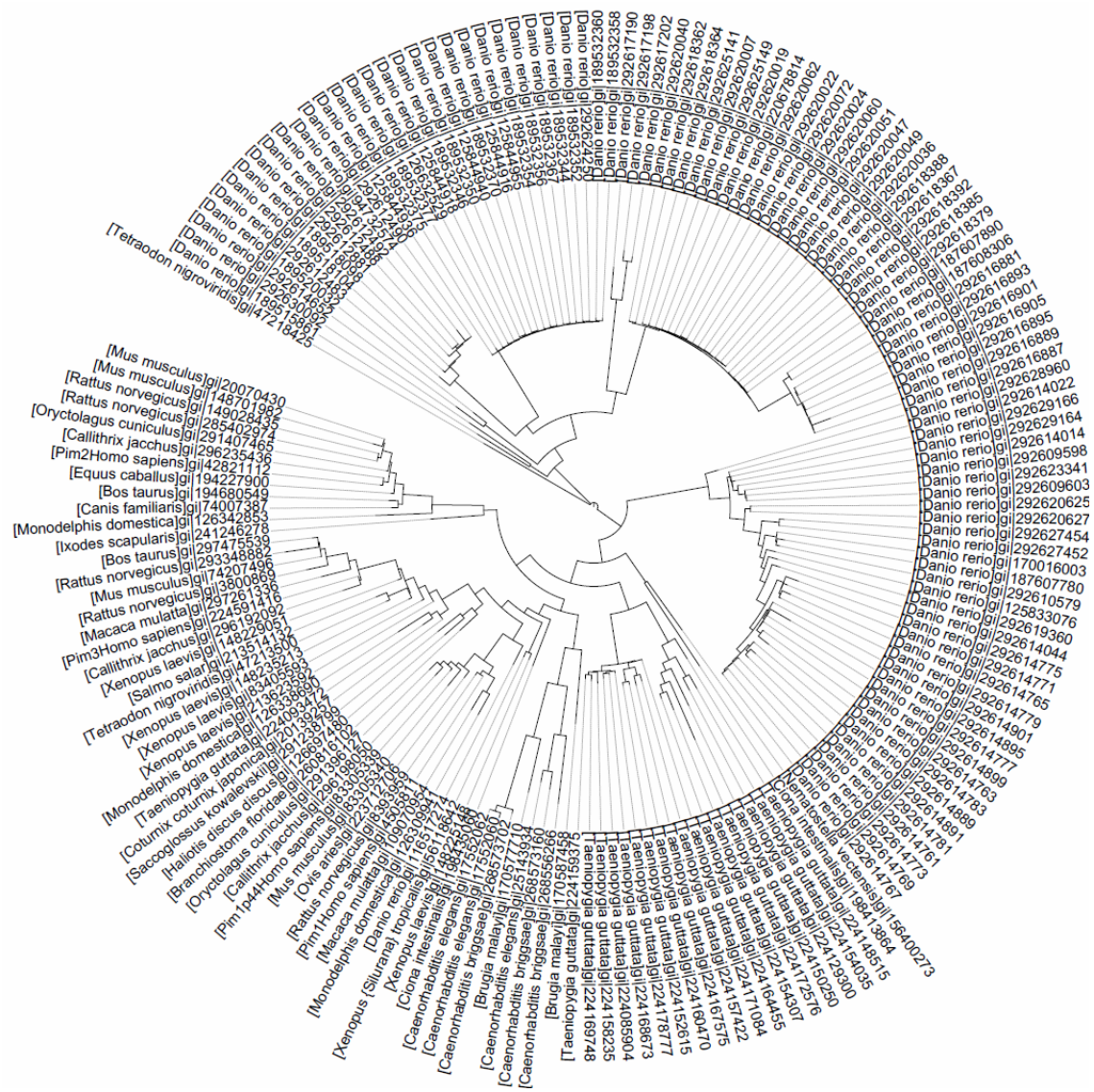
Appendix C. Human PIM kinases amino acid sequences alignment.

ClustalW2 alignment with secondary structure elements indicated.



Appendix D. A phylogenetic tree of the PIM protein kinase family.

GenBank IDs and species names serve as primary keys of the sequences identification.



Appendix E. The IC_{50} values for PIM variants

Tables showing IC_{50} values (μM) for hits with type 2 scaffold and the positive control staurosporine determined by Caliper microfluidic mobility shift assay. Values in the top table are means of three independent experiments. Standard deviations are presented in the bottom table.

SGC ID	PIM1	S46K/S51T	E124L	S46K	V126A	S51T	H68V	I56H/P63Q	PIM2	PIM3A
K01742a	0.052	0.090	0.142	0.079	0.093	0.070	0.038	0.163	5.847	0.132
K01844a	0.022	0.038	0.056	0.030	0.037	0.029	0.017	0.064	1.485	0.047
K01779a	3.481	5.567	>10	4.549	7.894	7.079	4.859	6.154	>10	4.983
K00135	0.265	0.408	0.547	0.370	0.374	0.342	0.253	0.700	6.235	0.400
K00487	0.208	0.300	0.405	0.236	1.009	0.216	0.128	0.296	0.447	0.421
K00512	0.506	0.717	1.252	0.709	1.859	0.817	0.364	1.188	2.429	4.170
K00997a	0.006	0.010	0.020	0.009	0.019	0.008	0.004	0.020	0.339	0.024
K00780a	0.600	0.343	2.269	0.397	0.491	0.820	0.699	1.345	0.964	0.197
K00906a	0.176	0.190	0.402	0.217	0.160	0.233	0.183	0.331	0.342	0.053
K02698a	0.051	0.029	0.129	0.076	0.148	0.029	0.043	0.154	0.247	0.205
K00072	1.204	1.106	3.199	1.050	1.123	1.215	0.871	1.693	0.977	0.830
K00746a	0.044	0.043	0.106	0.061	0.046	0.041	0.039	0.111	2.652	0.057
K02723a	0.603	0.888	1.032	0.642	1.186	0.878	0.339	1.768	>10	0.424
K00070a	0.166	0.189	0.583	0.233	0.550	0.202	0.301	0.489	4.131	0.808
K03137a	0.072	0.117	0.185	0.101	0.152	0.090	0.070	0.207	>10	0.190
K03128a	0.065	0.138	0.272	0.124	0.511	0.094	0.055	0.293	>10	0.571
K03129a	0.026	0.046	0.096	0.043	0.175	0.037	0.016	0.117	6.260	0.201
K03140a	0.009	0.017	0.042	0.017	0.075	0.015	0.010	0.048	>10	0.097
K02978a	1.118	1.663	2.724	1.869	3.483	2.005	1.421	1.919	0.320	0.376
K02908a	0.676	0.568	0.982	0.663	1.089	0.786	0.453	0.744	0.454	0.299
K00486	0.686	1.325	1.217	1.147	1.995	0.949	0.527	1.926	9.527	2.202
K03527a	0.051	0.088	0.128	0.077	0.105	0.072	0.019	0.135	0.426	0.181
Staurosporine	0.160	0.205	0.232	0.181	0.049	0.183	0.123	0.256	0.223	0.038

SGC ID	PIM1	S46K/S51T	E124L	S46K	V126A	S51T	H68V	I56H/P63Q	PIM2	PIM3A
K01742a	0.004	0.009	0.024	0.005	0.007	0.005	0.007	0.037	1.547	0.021
K01844a	0.002	0.005	0.006	0.004	0.005	0.003	0.003	0.006	0.356	0.006
K01779a	0.879	2.267		2.878	3.171	3.295	4.002	0.486		2.214
K00135	0.037	0.017	0.078	0.016	0.045	0.024	0.091	0.124	3.976	0.048
K00487	0.012	0.051	0.055	0.030	0.920	0.006	0.008	0.021	0.012	0.016
K00512	0.056	0.191	0.761	0.141	1.091	0.105	0.079	0.452	0.811	4.175
K00997a	0.000	0.001	0.005	0.000	0.001	0.000	0.002	0.003	0.089	0.001
K00780a	0.033	0.029	0.224	0.036	0.052	0.090	0.012	0.079	0.198	0.022
K00906a	0.021	0.042	0.069	0.064	0.037	0.092	0.054	0.089	0.079	0.018
K02698a	0.003	0.003	0.017	0.004	0.020	0.001	0.007	0.024	0.041	0.038
K00072	0.120	0.194	2.546	0.196	0.273	0.316	0.128	0.821	0.215	0.191
K00746a	0.003	0.006	0.018	0.003	0.017	0.005	0.011	0.013	1.636	0.035
K02723a	0.099	0.094	0.548	0.048	0.288	0.128	0.069	0.807		0.058
K00070a	0.006	0.018	0.088	0.006	0.034	0.010	0.315	0.054	1.630	0.094
K03137a	0.008	0.005	0.021	0.004	0.014	0.006	0.008	0.017		0.011
K03128a	0.003	0.018	0.057	0.005	0.074	0.010	0.021	0.059		0.113
K03129a	0.002	0.007	0.025	0.004	0.034	0.003	0.014	0.030	3.006	0.030
K03140a	0.000	0.006	0.020	0.003	0.015	0.002	0.003	0.009		0.016
K02978a	0.117	0.768	2.604	1.207	1.587	1.049	0.942	1.131	0.167	0.165
K02908a	0.133	0.023	0.411	0.134	0.346	0.106	0.169	0.085	0.063	0.032
K00486	0.050	0.117	0.181	0.036	0.160	0.053	0.038	0.288	0.374	0.422
K03527a	0.002	0.003	0.010	0.004	0.013	0.003	0.018	0.009	0.101	0.014
Staurosporine	0.005	0.006	0.013	0.012	0.001	0.000	0.054	0.019	0.028	0.002

Appendix F. Analysis of residual interactions in selected PIM1 structures.

Table illustrating the results were obtained using custom made MATLAB scripts as described in section 4.1.2. A ligand is numbered as a residue 78.

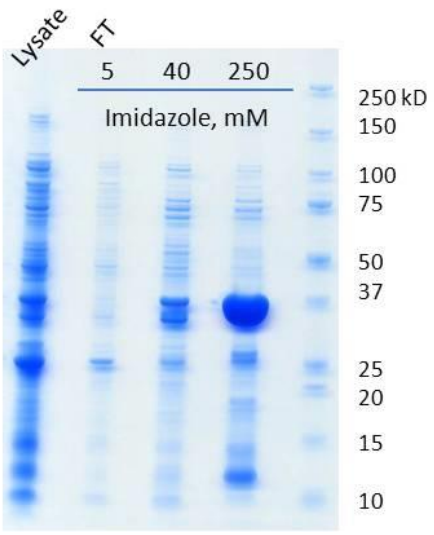
In 3JYA but not in 3JY0	In 3JY0 but not in 3JYA	3JYA (3b) Ki values PIM1 63nM PIM2 160nM 3JY0 (12b) Ki values PIM1 - 1nM PIM2 - 2nM
66 39 52 42 49 45 50 45 51 45 52 45 78 45 50 46 51 46 67 49 68 49 69 49 78 49 68 50 66 54	63 56 73 68 73 69	
in 3JXW but not in 3JY0	in 3JY0 but not in 3JXW	3JXW (6e) Ki values PIM1 5nM PIM2 29nM 3JY0 (12b) Ki values PIM1 1nM PIM2 2nM
52 42 50 45 51 45 52 45 50 46 51 46 73 48 67 49 68 49 69 49 78 49 68 50	63 56 73 68 73 69	
in 2C3I but not in 3JY0	in 3JY0 but not in 2C3I	2C3I (ImidazoPyridazines) IC50 values PIM1 0.12 μM PIM2 1.8 μM
49 45 50 45 51 45 52 45 78 45 50 46 51 46 67 49 68 49 69 49 78 49 68 50 69 50 66 54	61 56 63 56 73 68 73 69	

In 1XR1 but not in 3JY0 55 40 65 44 52 45 78 45 78 47 78 49 73 50 78 50 73 51 78 73	In 3JY0 but not in 1XR1 None.	1XR1 (AMP-PNP) 3JY0 (Dual)
in 1YXT but not in 3JY0 52 42 50 45 51 45 52 45 78 45 50 46 51 46 78 46 78 47 73 48 78 48 69 49 73 49 75 49 78 49 67 50 68 50 69 50 78 50	in 3JY0 but not in 1YXT 73 68 73 69	1YXT (AMP-PNP) 3JY0 (Dual)
in 2BZK but not in 3JY0 78 68	in 3JY0 but not in 2BZK 63 56 73 68 73 69	2BZK (AMP-PNP) no P-loop, residues 52-77 considered 3JY0 (Dual)
in 1YXT but not in 1XR1 52 42 50 45 51 45 50 46 51 46 78 46 73 48 78 48 69 49 73 49 75 49 67 50	in 1XR1 but not in 1YXT 55 40 65 44 73 50 73 51 73 68 73 69 78 73	1XR1 (AMP-PNP) 1YXT (AMP-PNP)

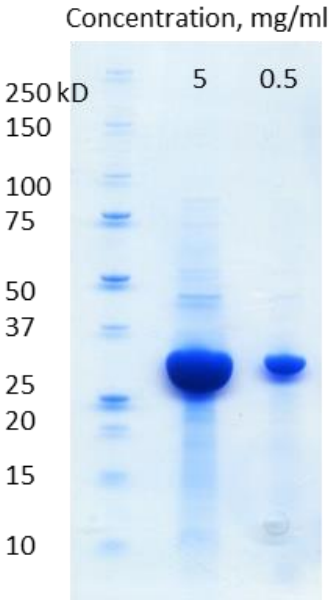
68 50 69 50		
in 3JY0 but not in 3MA3 61 56 73 68 73 69	in 3MA3 but not in 3JY0 None	3MA3 (Dual) partial P-loop, only residues 48-77 considered
in 2BZK but not in 1XR1 78 68	in 1XR1 but not in 2BZK 63 56 73 68 73 69 78 73	Only residues 52-77 considered
in 1XQZ but not in 3JY0 50 45 51 45 52 45 50 46 51 46 73 47 73 48 67 49 73 49 67 50 68 50 73 50 66 54	in 3JY0 but not in 1XQZ 51 43 65 52	1XQZ (apo-enzyme) 3JY0 (Dual)
in 2C3I but not in 3MA3 67 49 68 49 69 49 78 49 68 50 69 50 66 54	in 3MA3 but not in 2C3I 63 56	3MA3 (Dual) partial P-loop, only residues 48-77 considered
in 3JXW but not in 3MA3 73 48 67 49 68 49 69 49 68 50 61 56	in 3MA3 but not in 3JXW 63 56	

Appendix G. Representative examples of purified recombinant proteins

Initial purification of CDK2 by affinity chromatography, SDS-PAGE analysis CDK2 kinase is illustrated as an example.



Purified PIM2 protein used for crystallisation.



Size exclusion chromatography of the purified. Top: chromatogram from ÄKTAxpress with blue curve representing UV trace and red line representing conductivity. Bottom: SDS-PAGE analysis of the eluted fractions.