UNIVERSITY OF LONDON THESIS

Degree PhD Year 2006 Name of Author PATEL
Dony P

National Institute for Medical Research
Division of Molecular Structure

Functional Studies of Serine/Threonine Protein Kinase Signalling in *Mycobacterium tuberculosis*

A thesis submitted by

Dony P. Patel

in partial fulfilment of the requirements of the

University of London

for the degree of Doctor of Philosophy.

January 2006

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Abstract

Fork-head associated (FHA) domains function in the assembly of signalling complexes through specific interactions with phospho-threonine motifs. FHA domains are found in eukaryotes, and intriguingly in a subset of bacterial species including *Mycobacteria tuberculosis*. In addition, the *M. tuberculosis* genome encodes 11 eukaryotic-like serine/threonine protein kinases (STPKs). In *M. tuberculosis*, the co-expression of FHA domain-containing proteins and STPKs strongly suggest that these bacterial FHA domain-containing proteins engage in phospho-dependent protein-protein interaction, and FHA dependent processes in bacteria are controlled by STPK-dependent phosphorylation. This study describes a body of biophysical and biochemical experiments which characterises the interactions between two STPKs, PknA and PknB, and three FHA domain-containing proteins, Rv0019, Rv0020 and Rv1827. Isothermal titration calorimetry and surface plasmon resonance combined with site-directed mutagenesis are used to investigate interactions of kinase domains with FHA domains. These experiments reveal that the FHA domains interact with specific phospho-threonine residues located within the kinase domain activation loop. Additionally, i*n vitro* kinase assays demonstrated that the interactions also involve phosphorylation of the FHA domains and/or adjacent segments of the protein. These data suggest that STPK-mediated signalling in *M. tuberculosis* involves a complex series of interactions and suggests a 'molecular docking' model for the phosphorylation of FHA domain-containing proteins. Finally, high-throughput assays have been used to identify several inhibitors of PknB which will serve as the basis for development of novel antimicrobial therapies.

Acknowledgements

To the long suffering boss (Dr. Steve Smerdon) and friends in the NIMR, to my parents for their eternal support, to Rikita for all the reasons only she knows.

Contents

List of Figures

Chapter 1 – Introduction

Chapter 2 – Quantitative Techniques in Protein-complex Interactions

Chapter 3 – Materials and Methods

Chapter 4 – Protein Purification and Characterisation

Chapter 5 – Biophysical Characterisation of Kinase-FHA Interactions

Chapter 6 – Kinase Activity of PknA and PknB

Chapter 7 – Inhibitor Studies

List of Tables

1 Introduction

1.1 The Origins of Signal Transduction

The term 'signal transduction' encompasses a multitude of mechanisms by which the transfer of biological information occurs. These mechanisms regulate cellular events such as signal output of cell surface receptors, DNA repair, cell cycle control, cytoskeletal changes and protein degradation. Protein-protein interactions provide the principal framework through which the many signals for these cellular events are initiated. Such protein recognition and interaction events must be tightly controlled, specific and reversible, as they control the route of a signalling pathway. Post-translational modifications are commonly utilised for the regulation of these molecular interactions as well as conformational changes. Post-translational modifications include acetylation, methylation, hydroxylation and ubiquitination, but by far the most commonly employed post-translational modification is phosphorylation. Over 30% of human proteins contain a covalently bound phosphate at some time during their existence. In addition, enzymes that catalyse protein phosphorylation, kinases, have been determined to be one of the largest gene families in eukaryotes, with 518 putative protein genes (Manning *et al.*,2002).

The study of protein phosphorylation began with work carried out in the late 1930s by Carl and Gerty Cori. Their work on the interconversion of phosphorylase *a* and *b* won them the Nobel Prize in Physiology and Medicine in 1947 for "discovering the course of the catalytic conversion of glycogen". While the Coris were unaware of the exact manner of the interconversion of phosphorylase, they demonstrated that an enzyme termed 'prosthetic-group-removing' (or PR) enzyme acted on phosphorylase a to convert it to the b form. This was the first example of what is now known as allosteric activation.

It was twenty years later that the exact chemical nature of the interconversion reaction of phosphorylase *a* and *b* was revealed. Working at the same department of the University of Washington as the Coris, Edwin Krebs and Ed Fischer tried to purify phosphorylase *a* from muscle using the same methods as their predecessors. To clarify

the muscle extract they initially used centrifugation, and were only able to purify the *b* form. If this was replaced by a filtration step, they obtained the *a* form. It was also found that the conversion did not occur if the muscle extract was aged. It was eventually established that ATP was the essential constituent and that this was lost over time. This led to the identification of phosphorylase kinase; an enzyme that catalysed a phosphotransferase reaction in which the terminal phosphoryl group of ATP was transferred to phosphorylase *b*. Further study identified the critical difference between the centrifugation and filtration methodologies. The filter paper contained a source of calcium. Calcium was required for the activity of phosphorylase kinase and other activators of this enzyme.

Despite the ubiquity of phosphorylated proteins in nearly every aspect of cellular life, the identification of phosphorylated amino acids was only discovered in the 1970s. Pioneering work on the *v-Src* protein from Rous sarcoma virus (RSV). *v-Src* was the first oncogene to be discovered (Martin, 1970) and shown to encode an abnormally regulated tyrosine kinase (Collett & Erikson, 1978). During their study of polyomavirus middle T antigen (PyMT), Hunter and Eckhart found that PyMT was phosphorylated (Eckhart *et al.*, 1979). Further analysis showed the phosphorylated amino acid to be a tyrosine. Aberrantly regulated kinase activity is now recognised as a cause or consequence of many human diseases.

1.2 The Transmission of Signals by Protein Phosphorylation

Phosphorylation is the most common regulatory mechanism adopted by eukaryotic cells and to a lesser extent, prokaryotic cells. The coupled reactions of protein phosphorylation and dephosphorylation are catalysed by protein kinases and protein phosphatases, respectively. Working in tandem, these enzymes are able to modify the function of proteins in numerous ways; they can disrupt protein-protein interactions, modulate biological activity, signal subcellular compartmentalisation, translocation, involve signal destruction or alter protein stability.

Protein kinases catalyse the specific transfer of the terminal (γ) phosphoryl group of ATP to either serine or threonine residues in the case of serine/threonine kinases, and to tyrosine in the case of tyrosine kinases. Human genome sequence analysis by Manning *et al* (2002) suggests that humans may possess 388 serine/threonine kinases and 90 tyrosine kinases. A further class of protein kinases are the histidine/aspartate kinases that catalyse the phosphorylation of specific histidine residues (N^{π} or N^{τ} residues on the histidine ring). These enzymes are most prevalent in prokaryotic cells and will be discussed later. Eukaryotes also posses two additional classes of protein kinases, the myosin heavy chain kinase/EF-2 kinase and the mitochondrial protein kinases which including pyruvate dehydrogenase kinase (Krupta and Srinvasan, 2005).

The reversibility, specificity and flexibility of phosphorylation make it ubiquitous as a control mechanism. However, there are many physical consequences of phosphorylation; the addition of a phosphoryl group adds two negative charges to the modified protein. This can lead to profound alterations to electrostatic interactions within the protein that can in turn lead to structural changes. The phosphoryl group is also able to form a maximum of three hydrogen bonds. The free energy of phosphorylation is −12kcal/mol, such a large free energy change ensures that the conformation change of the modified protein is very stable. The cycling between phosphorylation and dephosphorylation can occur very quickly. This means that the effects of phosphorylation can be quickly amplified. The activation of one kinase can lead to it phosphorylating many target proteins in a short space of time, especially other kinases leading to a signalling cascade.

1.3 Protein Phosphorylation in Prokaryotes

Until recently, the prevailing model of prokaryotic signalling largely involved phosphorylation of histidine and aspartic acid residues via two-component regulatory systems. Serine/threonine and tyrosine kinases were thought to be unique to eukaryotes. However, the accumulation of genomic sequence data has revealed that some bacteria contain phosphoester kinases (both Ser/Thr and Tyr kinases) and phosphatases as well as associated adaptor molecules that have eukaryotic homologs.

Extracellular signals such as osmotic stress or sugar concentration are generally transduced into cellular responses by two-component systems (aka. the His-Asp phospho-relay system) in bacteria (Figure 1.1). The sensor component of this system contains an extracelluar ligand-binding domain that, once bound, leads to the autophosphorylation of its intracellular segment on a conserved histidine residue carried in the H-box domain of the histidine kinase. The sensor protein then transfers the phosphoryl group to a conserved aspartate on its associated response regulator protein.

It was not until 1991 that a serine/threonine protein kinase, *pkn1*, from *Myxococcus xanthus* was purified and characterised (Munoz-Dorado *et al.*,1991). It has now been established that bacterial signalling through serine/threonine and tyrosine protein kinase is the rule rather than the exception. This is evident from the presence of these proteins in more than 20 of the 50+ bacterial genomes sequenced to date (Pallen *et al.*, 2004). Analysis of GC content and codon usage suggests that bacterial serine/threonine protein kinases (STPKs) are genuine prokaryotic enzymes rather than genes acquired through horizontal gene transfer from eukaryotes to prokaryotes (Han & Zhang,2001). If horizontal gene transfer of STPKs had occurred from eukaryotes to prokaryotes then this would have occurred early in the evolution of the bacteria that possess the genes (Han and Zhang, 2001). Genome analysis has shown that eukaryotes also utilise two-component systems containing histidine kinases (Wolanin *et al.*,2002).

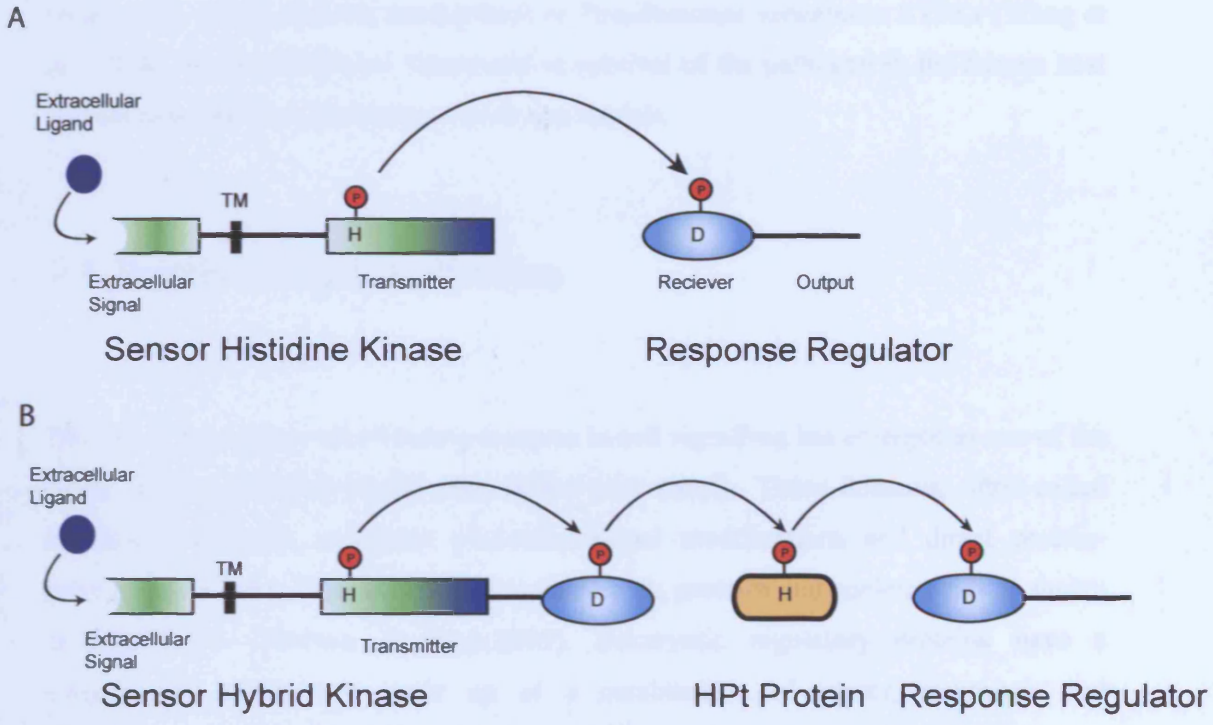


Figure 1.1. Prokaryotic signal transduction systems involving histidyl-aspartyl phosphorelay. Histidine kinase domains are represented as rectangles, receiver domains as ovals and transmembrane domains by black bars. Histidine-containing phospho-transfer protein (HPt) is indicated by rounded rectangles. Sites of phosphorylation upon histidine (H) and aspartic acid (D) residues are indicated. **(A)** A basic two-component regulatory system which employs a histidine kinase that contains an extracellular sensor domain and a response regulator. **(B)** A phospho-relay system that employs a HPt protein intermediate between the sensor hybrid kinase and the response regulator. TM, transmembrane domain.

In prokaryotes, STPKs are generally involved in either growth or pathogenicity. There are a large number of STPKs encoded in the genomes of bacteria that have the ability to differentiate into a new developmental state, including *Anabaena* spp. (Zhang and Libs., 1998), *Streptomyces* (Urabe and Ogawara *et al.*, 1999), and *Myxococcus xanthus* (Inouye *et al.*, 2000). In these species, STPKs are essential for controlling the late stages of development, sporulation and secondary metabolite production. STPKs are also involved in pathogenicity. Examples include *Yersinia pseudotuberculosis* STPK, *yopO* (Barz *et al.*, 2000), and the complement of *Pseudomonas aeruginosa* STPKs (Wang *et al.*, 1998). These STPKs are implicated in survival of the pathogen in the human host and are essential for full virulence in mouse models.

1.4 Protein Interaction domains

The use of modular protein binding domains in cell signalling has emerged as one of the major themes of signal transduction in the past decade. These domains, often called interaction domains, recognise post-translational modifications and direct protein-protein interactions. They can also recognise lipids, proteins and nucleic acids as shown in Figure 1.2 (Pawson & Nash,2003). Eukaryotic regulatory proteins have a multidomain architecture made up of a combination of sensor, enzymatic and interaction domains. The enzymatic domains of regulatory proteins often generate a modified amino acid on a substrate protein, this can then be recognised by an interaction domain. In this way, multi-component signalling complexes may be built up and cytosolic signal transduction cascades generated. Interaction domains control the specificity of signal transduction as well as the kinetics of interactions. They therefore regulate the formation of signalling complexes at appropriate cellular locations, and are essential for the generation of complex cellular behaviour. In contrast, prokaryotic signalling predominantly involves separate genes for enzyme and interaction domains encoded in the same operon.

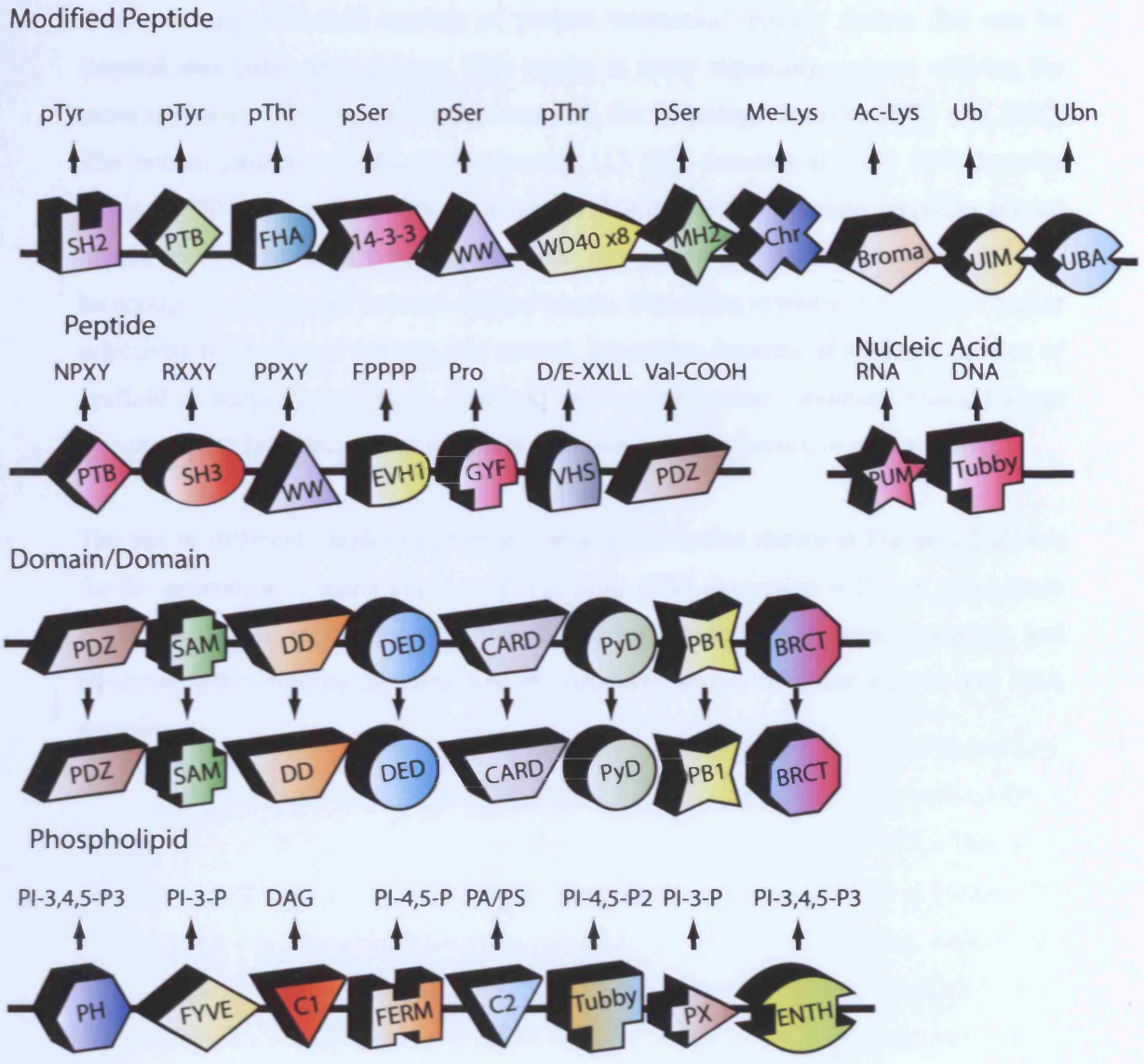


Figure 1.2. Modular interaction domains used in signal transduction. Interaction domains have been identified which associate with proteins, nucleic acids and phospho-lipids. A selection of domains and their binding partners are shown. The full names of domains are given in the SMART (Simple Modular Architecture Research Tool) database (www.smart.embl-heidelberg.de). Adapted from Pawson and Nash, 2003.

A modular system has allowed the tailoring of proteins during evolution. This in turn has facilitated diverse combinations of protein interactions and activities. However, there are only a limited number of protein interaction domain classes that can be inserted into genes and operons. This results in many signalling systems utilising the same interaction domains, as is the case with *Src* homology domains (SH2 and SH3). The human genome encodes approximately 115 SH2 domains and 253 SH3 domains (Pawson, 2004). Specificity can be achieved despite similar domains targeting related motifs. Firstly, the affinity of the interaction domain for a particular peptide motif may be enough to distinguish between related targets. Signalling systems may achieve higher selectivity by the use of not one, but several, interaction domains as well as a number of scaffold or adaptor proteins. In addition, local concentration conditions brought about by subcellular localisation and temporal expression may influence *in vivo* selectivity.

The use of different classes of protein interaction domains shown in Figure 1.2 allows for the generation of many systems of regulation. This discussion will now concentrate on the subset of regulatory processes based on the actions of protein kinases and phosphopeptide-binding domains, and in particular, serine/threonine kinases and FHA domains.

1.5 Phosphopeptide-binding domains

A number of phosphopeptide-binding domains have been studied in detail and several reviews exist on phosphoserine/threonine (pSer/pThr) -binding domains (Yaffe & Smerdon,2004) and phosphotyrosine (pTyr)-binding domains (Pawson, 2004). In general, isolated phosphodependent interaction domains fold in a manner that leaves a ligand-binding surface able to recognise exposed short phosphorylated motifs. For example, growth factor-receptor tyrosine kinases autophosphorylate at specific tyrosine residues on their extended tails in response to an extracellular signal. These phosphotyrosine sites are able to bind effectively with a number of pTyr-binding domains such as SH2 domains and phosphotyrosine binding domains (PTB) (Wybenga-Groot *et al.*,2001).

The versatility of interaction domains is exemplified by the signalling between transforming growth factor-β (TGF-β) receptor kinases and R-SMAD signal transducer proteins. Upon TGF-β binding, the transmembrane Ser/Thr kinase TGF-β receptors autophosphorylates a specific glycine- and serine-rich sequence in the juxtamembrane region at exposed serine residues. The R-SMAD protein contains an N-terminal DNA binding domain (MH1) and a C-terminal protein interaction domain (MH2). This MH2 domain binds to the phosphopeptide motif of the phosphorylated TGF-β kinase. The MH2 domain also associates with the scaffold protein SARA, and once phosphorylated by TGF-β kinase, MH2 can then recognise other phosphorylated MH2 domains. The R-SMAD domains/proteins trimerise through the phosphorylated MH2 C-terminal region, which binds to the same phospho-binding region of the adjacent monomer. In turn, trimers then dissociate from the receptor kinase and translocate to the nucleus (Miyazono *et al.*,2004). In the nucleus the MH1 domain binds to DNA whilst the MH2 domain forms new interactions with various transcriptional components, altering gene expression (Wu *et al.*,2001). Surprisingly, the structures of FHA domains are related to the SMAD MH2 domain described previously despite the lack of sequence homology (Durocher *et al.*, 2000). The FHA domains play a major role in this thesis and will be the topic of the following Section.

1.6 FHA domains

1.6.1 Overview

Forkhead-associated domains, or FHA domains, were originally identified by Hoffman and Bucher (1995). This domain was initially discovered in a subset of Forkhead transcription factor family members, from which their name derives. FHA domains are now known to be present in over 450 proteins, in many organisms ranging from prokaryotes to higher eukaryotes. They are a class of small protein interaction domain found to recognise phosphothreonine peptide motifs (Durocher *et al.*, 2000). The majority of FHA domains studied to date have been shown to selectively bind peptide motifs defined by amino acids from the –4 to the +3 positions relative to a pThr using *in vitro* peptide library screens (Yaffe and Smerdon, 2004). The domains can be 130-140 residues in length with a core homology region of around 50-75 residues. The core homology region can be separated into three conserved regions with five absolutely conserved amino acids (Figure 1.3).

1.6.2 Functional versatility of the FHA domain

One of the best characterised FHA domain-containing proteins with respect to its physiological role and FHA domain structure is Rad53. This is a protein from *Saccharomyces cerevisae* and is involved in the DNA damage cell cycle checkpoint. It has two FHA domains flanking a central serine-threonine protein kinase domain. The FHA domains of Rad53 are involved in the association with phosphorylated Rad9 following DNA damage (Sun *et al.*,1998) where the Rad53/Rad9 interaction leads to a high local concentration of Rad53. Consequently, Rad53 autophosphorylates *in trans* and becomes active, resulting in activation of the Rad53 branch of the checkpoint pathway for DNA damage repair (van den Bosch *et al.*, 2004).

An FHA domain was also found to be important in the CLV1 signal transduction pathway that controls cell proliferation and differentiation in *Arabidopsis thaliana* (Li *et al.*,1999). The interaction between the kinase-associated protein phosphatase (KAPP)

and CLV1 from *A. thaliana* was shown to be dependent on both a kinase interaction (KI) domain present in KAPP, and phosphorylation of CLV1. CLV1 is a serine/threonine protein kinase. A 52 amino-acid stretch of the KI domain is homologous to the conserved region of FHA domains. Mutations in key residues of this FHA homology region abolished binding of KAPP to CLV1.

1.6.3 Structure of FHA domains

Rad53 FHA2 was the first FHA domain structure to be determined using nuclear magnetic resonance (NMR) (Liao *et al.*,1999). This was followed by crystallographic structures of both FHA1 of Rad53, human Chk2 FHA domain and the FHA domain from ring finger containing checkpoint protein, Chfr (Liao *et al.*,2000, Durocher *et al.*,2000, Stavridi *et al.*,2002, ,Li *et al.*,2002). Comparisons of these structures show that the core structure of the FHA domain is an 11 or 12 β-stranded sandwich of around 50-75 amino acids in size (Figure 1.3). Variability in FHA domains arises as a result of insertions in the loops between β-strands and these are often helical. The mechanism of phospho-peptide binding will be discussed based on the structures of Rad53 FHA1 and Chk2 FHA. Later, these examples will be compared to the structure of the first bacterial FHA domain to be solved, that of Rv0020 of *Mycobacterium tuberculosis*.

The mechanism of binding and selectivity is explained by the structures of FHA domains solved in complex with an optimal binding peptide derived from peptide library screening. The FHA domain binds phosphopeptides through direct contact to the phosphopeptide backbone and via hydrogen-bonding to the phosphothreonine residue. These interactions involve a large number of residues, probably forming a binding surface rather than a few localised amino acids (Durocher and Jackson., 2002). This surface is generated by the loops that connect the β-strands β3-β4, β4-β5, β6-β7, and β10-β11 (Figure 1.4). The residues involved in contacting the phospho-peptide backbone or the phosphothreonine itself are best conserved. The specificity-determining residues are less well conserved, allowing FHA domains to impart selectivity which is most prominent for the residue +3 relative to the phosphothreonine (Durocher *et al.*,2000).

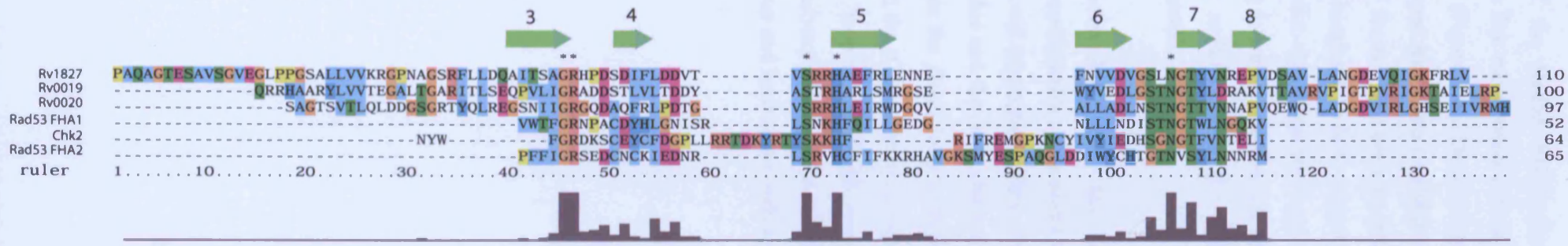


Figure 1.3. Multiple sequence alignment of selected FHA domains. The amino acid sequence and secondary structure elements of human Chk2 FHA domains are shown above the alignment as β-sheets (green arrows). The FHA residues critical for phospho-threonine recognition are indicated by stars. Rv1827, Rv0019 and Rv0020 are from *M. tuberculosis*, Rad53 FHA domains from *S. cerevisae* and Chk2 FHA from humans. Bars below the alignment show are a measure of the conservation at each residue position.

The structure of the Rad53 FHA1:phospho-peptide complex clearly shows the importance of the five most conserved residues of FHA domains (Gly^{69}, Arg^{70}, Ser^{85}, His^{88} and Asn^{107}) (Figure 1.4). These residues are all found around the peptide-binding site. Arg^{70}, Ser^{85} and Asn^{107} are involved in the binding to the peptide backbone whilst in addition, Arg^{70} forms a salt bridging interaction with pThr and Ser^{85} forms a direct contact with the phosphate. In fact, the hydroxyl moiety of Ser^{85} (Ser^{140} in Chk2) forms a hydrogen bond with the phosphate oxygen as well as stabilising the binding pocket by hydrogen bonding to a main chain amide nitrogen on the $\beta6$-$\beta7$ loop. Gly^{69} and His^{88} is in stabilising the architecture of the domain by tethering the loops that bind the phospho-peptide motif.

As mentioned above, Arg^{70}, Ser^{85} and Asn^{107} (Arg^{117}, Ser^{140} and Asn^{166} in Chk2) interact with the peptide main chain atoms that flank the pThr via a number of hydrogen bonds. The conserved asparagine makes contact with the main chain carbonyl oxygen of the pThr + 1 residue and the main chain nitrogen of the pThr + 3 residue. In addition, this residue secures the $\beta6$-$\beta7$ loop to the$\beta3$-$\beta4$ loop. This asparagine residue can also make contact with the pThr γ-methyl group in Chk2 (Li *et al*, 2002). The interaction to secure the $\beta6$-$\beta7$ loop to the $\beta4$-$\beta5$ loop involves the asparagine side chain amino moiety and the carboxyl moiety of Arg^{85} (Thr^{138} of Chk2). Finally, the Arg^{70} contacts the phosphate group and hydrogen bonds to the main chain carbonyl oxygen of the pThr + 2 residue.

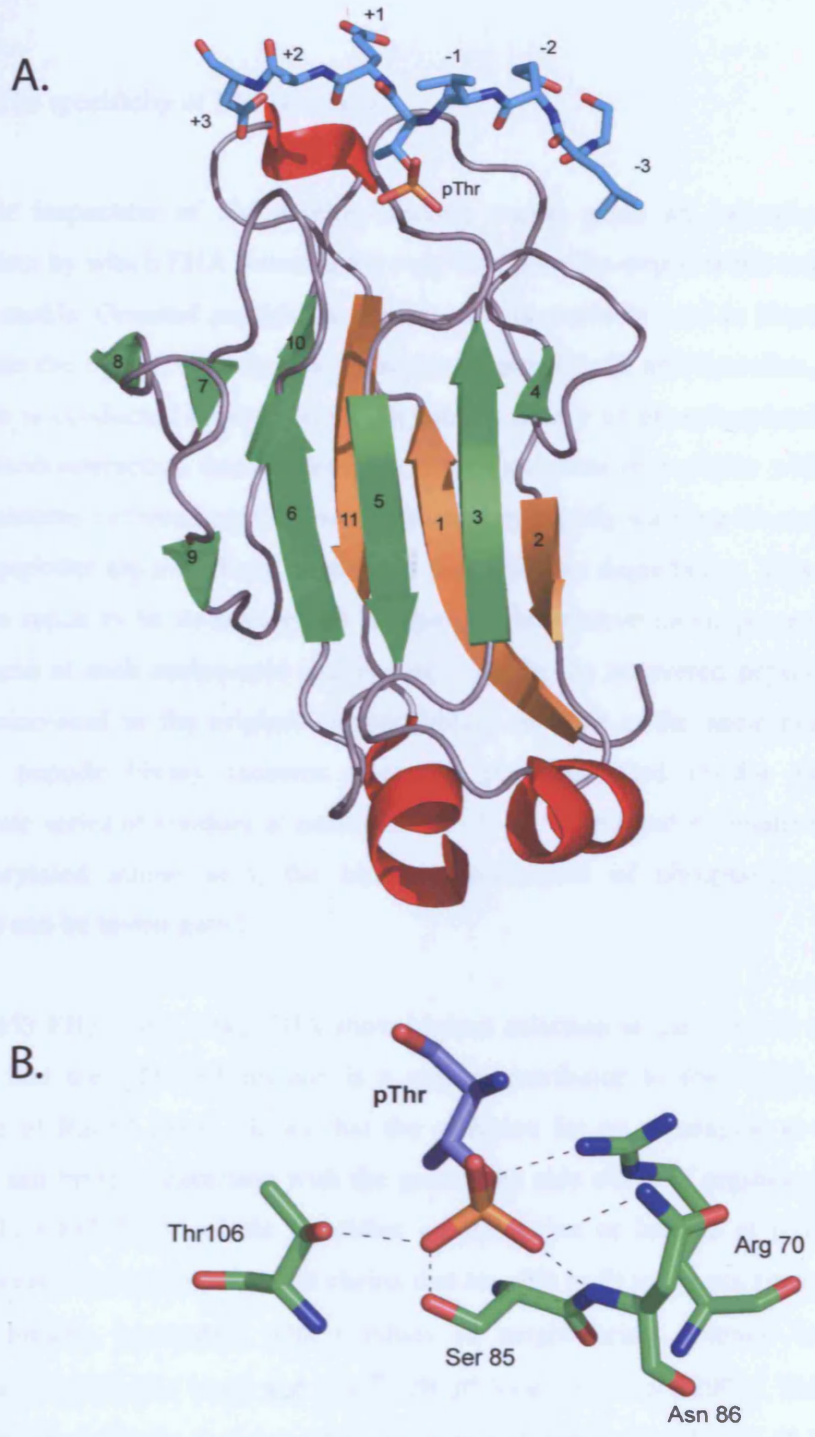


Figure 1.4. The protein fold of the Rad53 FHA1 domain from *Saccharomyces cerevisiae* in complex with its optimal binding peptide. (A) The domain architecture is formed from a total of 11 β-strands that associate to form a compact, β-sandwich fold. Phospho-peptide binding occurs through interaction with residues from the β3-β4, β4-β5, β6-β7, and β10-β11 loops. (B)Key protein-phospho-threonine contacts made by conserved residues of the FHA domain of Rad53 FHA1. The phospho-threonine and key FHA domain residues are shown in stick representation and hydrogen bonds shown as dashed lines.

1.6.4 The specificity of FHA domains

A closer inspection of the peptide binding region gives an indication as to the mechanism by which FHA domains not only bind phospho-peptides but target particular peptide motifs. Oriented peptide library selection is routinely used to identify a peptide that forms the tightest affinity to a particular domain (Yaffe and Smerdon., 2004). This selection is conducted by applying a degenerate library of phosphorylated peptides to resin-bound interaction domain proteins, after incubation of peptides with the protein for 10 minutes, unbound peptides were removed by rapidly washing the resin with PBS. Bound peptides are eluted and sequenced using Edman degradation. This enables raw selection ratios to be determined by comparing the relative molar percentage of each amino-acid at each amino-acid sequencing cycle in the recovered peptides to that of each amino-acid in the original peptide library mixture at the same position. If the original peptide library contains a central phosphorylated residue flanked by a degenerate series of residues at positions –3, -2, -1, +1, +2 and +3 relative to the fixed phosphorylated amino acid, the binding preferences of phospho-peptide binding domains can be investigated.

The Rad53 FHA1 and Chk2 FHA show highest selection at the pThr +3 residue. This implies that the pThr +3 residue is a major contributor to specificity. Indeed, the structure of Rad53 FHA1 shows that the selection for an asparagine at this position forms a salt-bridge interaction with the guanodine side chain of arginine (Durocher *et al*, 2000). Chk2 FHA selects for either an isoleucine or leucine at pThr +3. These amino-acids posses non-polar side chains that are able to fit into a pocket on the domain surface leading to contact with residues of neighbouring β-strand loops namely Ser^{192}/Leu^{193} (β10-β11 loop) and Thr^{148} (β4-β5 loop) (Li *et al.*, 2001). This interaction is brought about by the fact that Chk2 has a helical insertion in the β4-β5 loop, leading to a larger β10-β11 loop than in Rad53, which creates a second binding cleft (Li *et al*, 2001).

There appears to be little structural variation in the FHA domains studied to date. This can be attributed to the restrained nature of the binding of the phospho-peptide in an

extended conformation along the domain surface. This results in a forced selection at the regions of greatest structural variability, namely the surface around the pThr + 3 residue (Li *et al*, 2001).

1.6.5 The similarity between prokaryotic and eukaryotic FHA domains

The structure of a FHA domain from the protein Rv0020 of *Mycobacterium tuberculosis* has recently been solved in complex with its optimal peptide (Westcott S, unpublished data). The structure of Rv0020 FHA is strikingly similar to eukaryotic domains with 10/11 β-strands but has no additional α-helical insertions in the connecting loops (Figure 1.5). Again, the domain shows an absolute requirement for a pThr residue. Its optimal peptide also binds in an extended conformation as described above for Rad53. The FHA binding surface interacts with the phospho-peptide via five hydrogen bonds mediated by side-chain and main-chain atoms of Arg^{459}, Ser^{473}, Asn^{495}, Arg^{474} and Thr^{494}. The only novel interactions of Rv0020 are Arg^{474} and Thr^{494}; all other interactions exist in FHA domains previously discussed. The Arg^{474} was found hydrogen bonded to the phosphate group of the phospho-peptide.

The specificity of Rv0020 for its phospho-peptide is mediated by a number of key contacts; an isoleucine is highly selected for in the pThr +3 position of the peptide in a manner analogous to that of Chk2. Again a specificity 'pocket' is created for a large hydrophobic side chain by the residues Asn^{495} (β6-β7) and His^{519} (β10-β11). Asparagine contributes to the selection of isoleucine by forming one hydrogen bond between main-chain atoms. Taking the structural and specificity data collectively one can class FHA domains into two categories based on their specificity at the pThr + 3 position of their optimal phospho-peptide. The FHA domains of Rv0020 and Chk2 prefer large hydrophobic residues while Rad53 FHA1 and another FHA domain, Cds1p, show preference for an acidic residue (Durocher *et al.*, 2000, Westcott S., unpublished data).

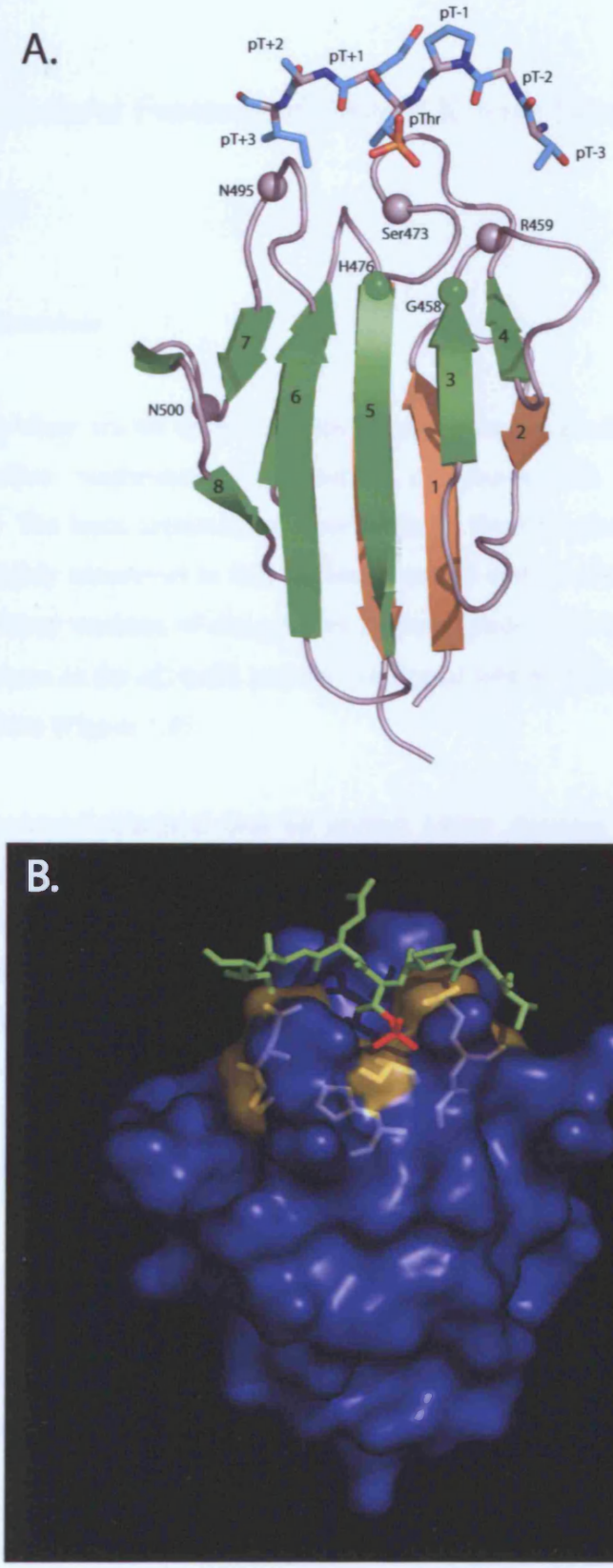


Figure 1.5. (A) The protein fold of the Rv0020 FHA domain from *M. tuberculosis* in complex with its optimal binding peptide showing a characteristic b-stranded sandwich topology. Highly conserved residues (N495, R459, G458, S473 and H476) critical for phospho-threonine recognition are shown as spheres. (B) Surface representation of the FHA domain of *Mycobacterium tuberculosis* protein Rv0020 cocrystallised with its optimal binding peptide. FHA domain surface is coloured blue. Highly conserved residue surfaces coloured yellow and also shown in stick representation (yellow). The pTXXI optimal binding peptide is shown and coloured green. The phosphate group of the phospho-threonine residue is coloured red.

1.7 Structural Features of Protein Kinase Domains

1.7.1 Overview

Currently there are 46 unique protein kinase structures available, which include active and inactive conformations as well as complexes with inhibitors and regulatory subunits. The basic architecture observed in all these structures involves a bilobal fold that is highly conserved in both serine/threonine and tyrosine kinases. The N-terminal lobe (~ 85aa) consists of anti-parallel β-sheets made of five β-strands and a single α-helix, known as the αC-helix and the C-terminal lobe (~ 170aa) is predominately helical (Lew, 2003) (Figure 1.6).

The volume of structural data on protein kinase domains has greatly increased our understanding of how kinases phosphorylate their targets and how this activity is regulated. This Section will attempt to define the structural features of the catalytically active conformation, with reference to the first solved kinase structure, that of PKA, and will outline some key regulatory mechanisms.

1.7.2 The active conformation

ATP is bound within a cleft between the two lobes and is positioned in a precise orientation for phospho-transfer by interactions with ~10 conserved residues. A key interaction coordinates the phosphates of ATP with the backbone of a glycine-rich loop composed of a 'GXGXψG' motif (X represents any amino acid and ψ is usually a tyrosine or phenylalanine). The glycines give flexibility to the loop allowing it to get close to the ATP and the side chain of the aromatic residue caps the site of phosphate transfer.

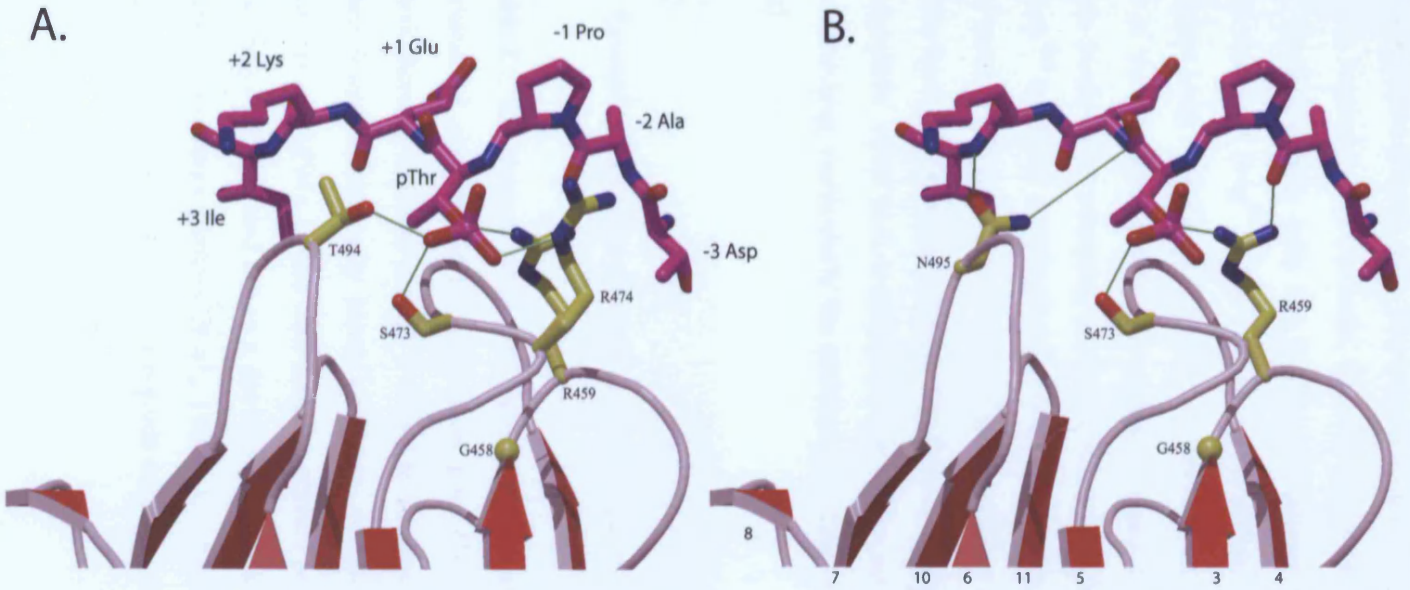


Figure 1.6. Key protein-peptide contacts made by conserved residues of the FHA domain of Rv0020. The peptide and key FHA domain residues are shown in ball-and-stick representation and hydrogen bonds shown as green lines. (A) The phosphate group of the threonine is bound through contacts from S473 and R459, along with contacts from the non-conserved residues T494 and R474. (B) N495 and R459 form hydrogen bonds with the main chain atoms of the peptide.

Further interactions occur to position the phosphates of ATP for catalysis when the conserved Lysine/Glutamate ion pair (Lys^{72}/Glu^{91} in PKA) form contacts with the α- and β-phosphates (Figure 1.7). The conserved Glutamate (Glu^{91} in PKA) of the Glu-Lys ion pair is located in the *Cα-helix*, this is a conserved helix of the N-terminal lobe that makes direct contacts with the N-terminal anchor point of the *activation loop*. A divalent metal ion (Mg^{2+} or Mn^{2+}) binds to the β- and γ-phosphates as well as conserved Asparagine (Asp^{184} in PKA) from an important DFG tripeptide motif at the N-terminal anchor of the important active site segment termed the activation loop. When a protein substrate binds in an extended conformation across the nucleotide binding pocket, an Asp (Asp^{166} in PKA) residing in the *catalytic loop* from the C-terminal lobe acts as a base to remove the proton from the substrate residue hydroxyl group leaving it free to attack the nearby γ-phosphate group. The final conserved feature of the active site is the APE tripeptide motif at C-terminal anchor of the activation loop which is critical for substrate binding, particularly the residue preceding the phospho-acceptor residue (P+1 residue).

1.7.3 Controlling Activity through Activation Loop Conformation

Undoubtedly, the most common mechanism for regulation of kinase activity involves conformational changes of the activation loop (Nolen *et al*, 2004), a multipurpose structural element that has clearly evolved to control activity of the kinase domain. It has been found to sterically block substrate- or co-factor binding in the active site (Hubbard *et al.*, 1994). It can also impede correct positioning of the αC Helix (Engh and Bossemyer, 2001) and act as a docking site for activating or inactivating cofactors or regulatory domains (Sahara *et al.*, 1996). In addition, it provides the P+1 residue-binding pocket to accommodate the peptide substrate (Engh and Bossemyer, 2002).

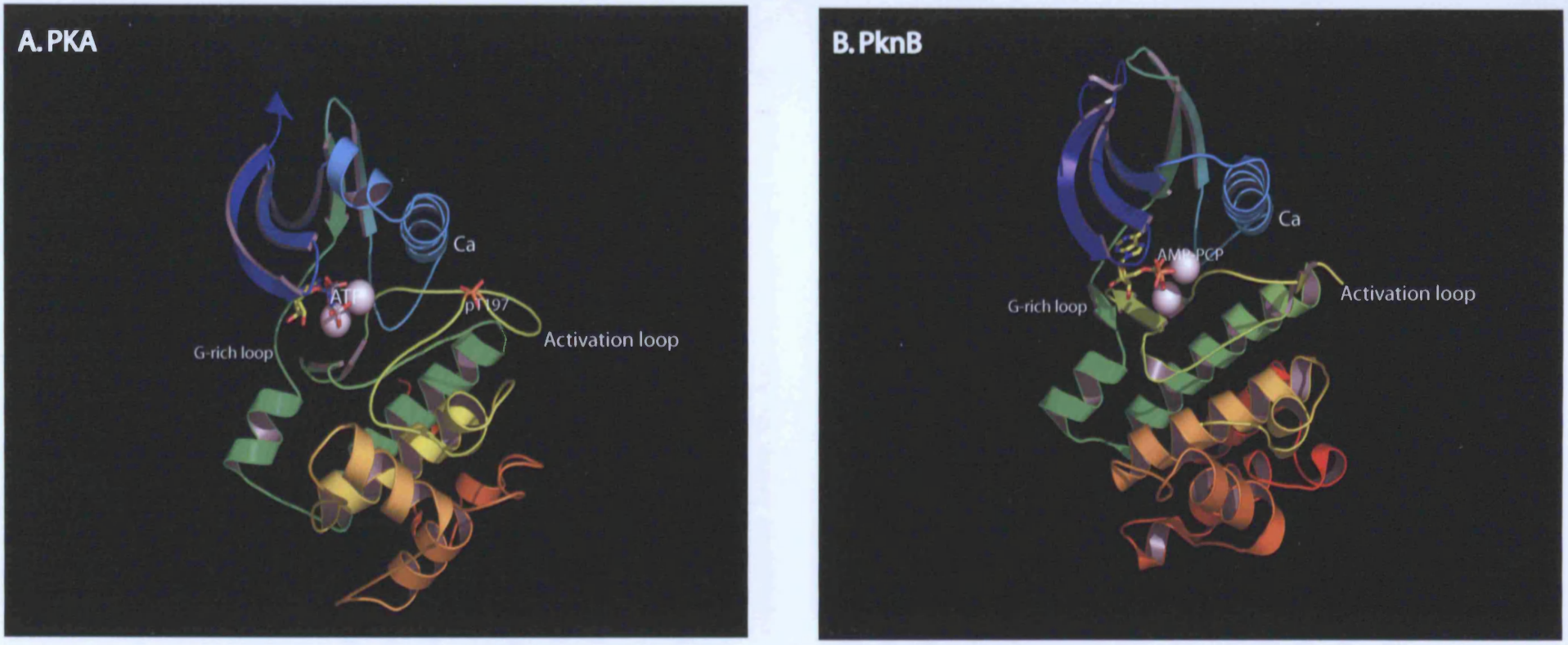


Figure 1.7. The catalytically active conformation of serine/threonine protein kinase domains of Protein kinase A (PKA) and PknB shown in ribbon representation. Key structural elements within the kinase domain including the bound nucleotide are indicated. Panel **(A)** shows the phosphorylated residue critical to activation of PKA, pT197. **(B)** The activation loop is absent from the structure of active PknB due to disorder arising from phosphorylation of four residues within this region.

A number of kinases require the activation loop to be phosphorylated for full activity, Johnson and colleagues noted that all these kinases possess a conserved arginine immediately preceding the conserved aspartate in the catalytic loop (Asp^{166} in PKA) (Johnson *et al.*, 1996). Therefore, this group of kinases, to which PKA and PknB belong, were termed RD kinases. The structure of active PKA reveals a phosphothreonine in the activation loop which allows the formation of a network of key interactions for catalytic activity (Figure 1.8). The importance of phospho-residues in the activation loop is also well illustrated in the active and inactive structures of insulin receptor kinase (IRK). In the inactive unphosphorylated form, the activation loop collapses into the active site and impedes binding of nucleotide and peptide substrate (Hubbard, 1997). Upon phosphorylation of three tyrosine residues, the activation loop moves out of the active site and is involved in both phospho-tyrosine and non-phosphotyrosine interactions resulting in the activation loop being positioned to allow correct orientation of the APE motif for substrate binding (Hubbard *et al*, 1998).

In addition to the activation loop regulatory mechanisms described above, the loop can also be an important site of protein-protein interaction. A recent crystal structure presented evidence of activation loop phosphorylation playing a role in recruitment of downstream signalling proteins. Hubbard and colleagues solved the structure of the *Src* homology domain (SH2) of Adaptor protein, APS (adaptor molecules containing PH and SH2 domains), bound to IRK via two phospho-tyrosines in the activation loop (Hu *et al.*, 2003). Another study (Wolf *et al.*, 2001) demonstrated mutations of activation loop residues within Erk2 kinase that affect its transport to the nucleus by eliminating activation loop interaction with cytosolic anchoring proteins. This newly discovered mechanism for substrate recruitment has implications on the conclusions drawn from data presented here and will be the subject of later discussion.

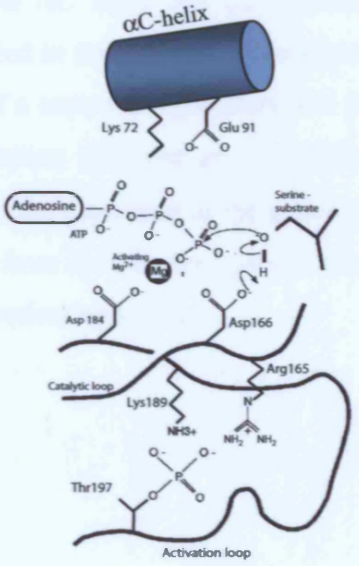


Figure 1.8.Schematic representation of the interactions in the active site of protein kinases and the possible mechanism for the phosphorylation of the serine substrate with the γ-phosphate of ATP. It also illustrates the involvement of the phosphorylated threonine (pThr 197in PKA) in the correct positioning of the key residues for catalysis.
Adapted from Nolen *et al.*, 2004.

1.7.4 Coupling Activity to Cα-Helix conformation

Another common regulatory strategy of kinases involves conformational changes in the αC helix. A number of *Src* kinase domains are inhibited when an N-terminal SH3 domain binds to the surface of the N-terminal lobe, stabilising the αC helix in an inactive conformation. Only when a ligand binds to the SH3 domain does the kinase assume an active conformation (Huse and Kuriyan, 2002). The switch between active and inactive conformation of the αC helix can concurrently cause a switch in the activation loop. This is exemplified in the case of Cyclin-dependent kinase (CDK). The inactive conformation consists of a conserved glutamate out of position for catalysis as a consequence of αC helix rotation (DeBont *et al.*, 1993). The binding of cyclin reorientates the αC helix to place the glutamate in the active site. This also necessitates movement of the activation loop from the inactive position to an active position where it can be phosphorylated by CDK-activating kinases.

Name	ORF	*M.leprea homolog*	Tm[a]	Unique features	Associated FHA domains	Other adjacent/nearby genes	Proposed regulatory function
pknA	rv0015c	+	+	-	Rv0019 and Rv0020	*pknB, pbpA , rodA* and *ppp*	cell elongation/septation
pknB	rv0014c	+	+	Extracellular PASTA domains	Rv0019 and Rv0020	*pknA, pbpA , rodA* and *ppp*	cell elongation/septation
pknD	rv0931c	-	+	PQQ domain, extracellular β-propeller	-	phosphate uptake operon	phosphate uptake
pknE	rv1746	-	+	-	-	ABC-transporter	membrane transport
pknF	rv1743	-	+	-	Rv1747	ABC-transporter	membrane transport
pknG	rv0410c	+	-	TPR motif	-	GlnH	Glutamine uptake, pathogenesis metabolism
pknH	rv1266c	-	+	*AfsK*-like	EmbR	ABC-transporter	Arabinan metabolism
pknI	rv2914c	-	+	Asn in active site	-	*ffh , ftsY*	cell division
pknJ	rv2088	-	+	-	-	transposon	?
pknK	rv3080c	-	-	PDZ and AAA domains	-	*luxA* like	?
pknL	rv2176	+	+	-	-	transcriptional regulator	?

Table 1.1. Features of serine/threonine protein kinases of *Mycobacterium tuberculosis*. A brief survey of the 10 serine/threonine protein kinases and associated functions. Abbreviations: ORF, open reading frame (*M. tuberculosis* H37Rv ORF designation); TM, transmembrane-spanning region; PASTA, penicillin binding domain and ser/thr protein kinase associated domain; PQQ, pyrroloquinoline quinone; TPR, tetratricopeptide. Table adapted from Av-Gay and Everett, 2000.

1.8 FHA and Serine/Threonine Protein Kinase Domains of *Mycobacterium tuberculosis*

1.8.1 Overview

There are 11 STPKs (PknA to L, except C) and 4 protein phosphatases in the *M. tuberculosis* genome and investigation to date implicates many of these STPK in either regulation of developmental processes or pathogenicity (Table 1.1). Also in the *M. tuberculosis* genome, four of the six FHA domain-containing proteins are encoded within gene clusters that also contain STPKs or Ser/Thr protein phosphatases (STPPs) (Figure 1.9) (Pallen *et al.*, 2002). Interestingly, there is a near perfect correlation between the presence of FHA-domain containing proteins with the presence of eukaryotic-like STPK and STPPs in bacterial genomes suggesting synergistic activity of the two protein classes (Ponting *et al*, 1999). Presently, little is known about the physiological role of many FHA domain-containing proteins and even less is known about the function of FHA and STPK proteins in bacteria. The occurrence of FHA domain-containing proteins in other bacterial genomes has been noted, generally in bacteria with more complex life cycles (Pallen *et al.*,2002), perhaps reflecting the requirement for more advanced signalling pathways.

Although the role of these kinases is not fully understood, they are likely to regulate growth, development, metabolic processes and interactions with the host cell (in the case of intracellular parasites). It is hoped that greater understanding of how prokaryotic STPK activities are regulated may identify drug targets to disrupt essential activities in *M. tuberculosis* and other important pathogens. It is important to note that sequence analysis of the 11 STPKs from *M. tuberculosis* shows they are significantly divergent from mammalian kinase domains and should therefore represent tenable targets for therapeutic intervention (Figure 1.10). To date, little is known about the effects of kinase inhibitors on *M. tuberculosis*. Kinase inhibitors erbstatin and staurosporine were shown by

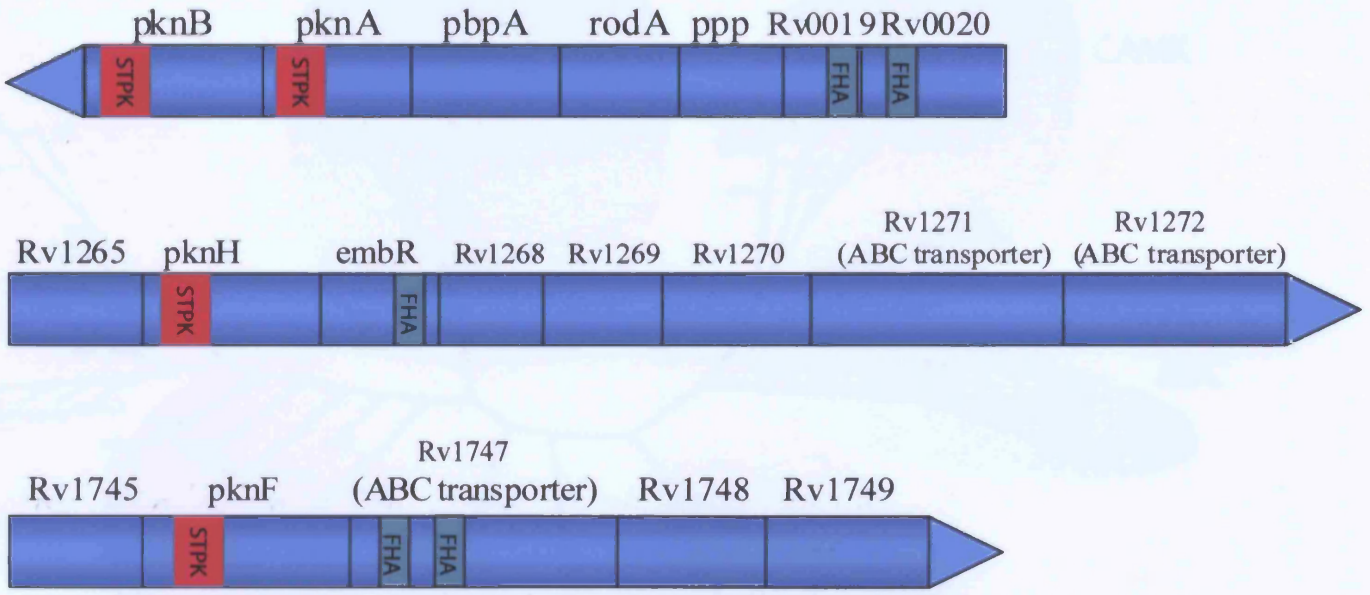


Figure 1.9. Putative operons of *M. tuberculosis* which encode both serine/threonine protein kinases and FHA-domain containing proteins. Serine/threonine protein kinase genes (*pknA, pknB, pknH* and *pknF*) are shown. Genetic locus of forkhead-associated (FHA) domains and serine/threonine protein kinase domains (STPK) are also shown.

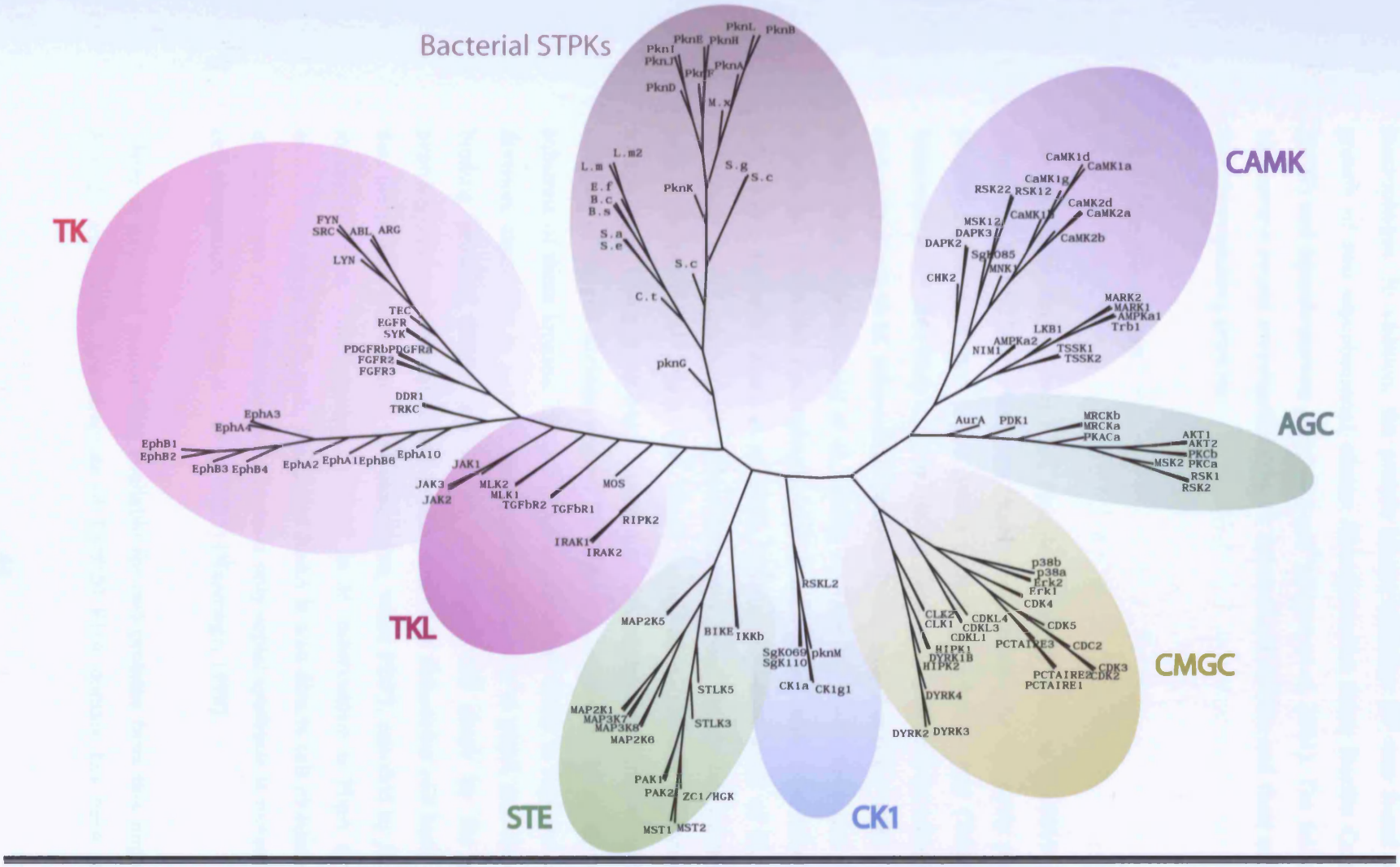


Figure 1.10. Phylogenetic tree depicting the relationship between members of the super family of human protein kinase and selected bacterial protein kinases. The seven major groups of human kinases are labelled and coloured distinctly, bacterial kinases are coloured in grey. Group names: AGC containing PKA, PKG, PKC families; CAMK Calcium/calmodulin-dependent protein kinase; CK1 Casein kinase 1; CMGC Containing CDK, MAPK, GSK3, CLK families; STE homologs of yeast Sterile 7, Sterile 11, Sterile 20 kinases; TK Tyrosine kinase; TKL Tyrosine kinase–like. A selection of human and bacterial kinase domain amino-acid sequences were aligned using ClustalW and a neighbor-joining tree was produced using Phylip software. (S.g, *Streptomyces griseus*; S.c, *Streptomyces coelicolor*; L.m, *Listeria monocytogenes*; E.f, *Enterococcus faecalis*; B.c, *Bacillus cereus*; B.s, *Bacillus subtilis*; S.e, *Staphylococcus epidermidis*; S.a, *Staphylococcus aureus*; C.t, *Clostridium tetani*).

(Prabhakaran *et al.*, 2000) to prevent phagocytosis of *Mycobacterium leprae* by mouse macrophages. In addition, the protein kinase inhibitor H7 was found to inhibit the growth of two mycobacterial strains *Mycobacterium bovis* Bacille Calmette Guerin (BCG) and *Mycobacterium smegmatis* mc^2 (Drews *et al*, 2001). The following Section will review recent investigations on *M. tuberculosis* STPKs and their associated FHA domain-containing proteins.

1.8.2 PknA and PknB

The operon containing two FHA domain-containing proteins (Rv0019 and Rv0020) along with two STPKs (PknA and PknB) also contains the cell cycle protein rodA, a penicillin-binding protein pbpA, and a protein phosphatase ppp (Table 1.1). PknB transcription is stable from early to late log phase with a 10-fold decrease upon entry to stationary phase in *M. tuberculosis* cultures. Presumably this is the same for all proteins in the PknB operon (Kang *et al.*, 2005). Av-Gay and co-workers also detected *pknB* expression in alveolar macrophages taken from a patient with tuberculosis (Av-Gay *et al.*, 1999). Recently, Kang et al, 2005, showed over-expression of PknA and PknB separately in *M. tuberculosis* cells results in a broad, bulgy cell morphology, while partial depletion of PknA and PknB using RNAi caused a narrow elongated morphology. Using a proteomic approach, this study also demonstrated Wag31, an ortholog of the cell division protein DivIVA of gram-positive bacteria, was an *in vivo* substrate of these kinases. These STPKs are therefore likely to regulate cell shape and division, especially in light of their genetic proximity to pbpA and rodA. Penicillin-binding proteins (PBPs) are known to control cell shape by the regulation of peptidoglycan polymerisation and genetic studies in *Esherichia coli* have demonstrated that the protein PBP2 directs cell elongation, while PBP3, encoded by *ftsI*, is involved in cell division. The homolog of PBP2 in *M. tuberculosis* is PbpA (29% identity), encoded by *pbpA*. In E. coli, PBP2 and RodA is also directs cell elongation and RodA mutants grow in a spherical shape because only septal synthesis is increased relative to cell elongation (Popham & Young, 2003) (Nanninga, 1998).

There is structural information available for two proteins from this important operon, Rv0020 and PknB. The structure of Rv0020 FHA domain has been solved by Dr.

Smerdon's lab at the National Institute of Medical Research, Department of Protein Structure (unpublished data) and is discussed in Section 1.6. The structure of PknB kinase domain in complex with a nucleotide analog was published at the start of 2003 by two groups, (Young et al., 2003) (3Å resolution) and (Ortiz-Lambardia *et al.*, 2003) (2.2Å resolution) and was the first known structure of a bacterial STPK. Critically, PknB was found to be autophosphorylated at four residues in its activation loop upon activation and this region was disordered in the crystal structure (Young et al., 2003).

The extracellular domain of PknB contains four penicillin-binding protein and serine/threonine kinase associated (PASTA) domain repeats of approximately 70 residues in length (Yeats et al., 2002). The domain binds to β-lactam antibiotics and their peptidoglycan analogues, and PASTA domains are purported to direct proteins to the site of cell growth. Genetic evidence suggests that the PASTA domain and the transmembrane of a PknB homolog in *Bacillus subtilis* are capable of forming dimers *in vivo* (Madec *et al.*, 2005). Therefore, this enzyme may be localised to areas of cell growth by means of extracellular PASTA domains, leading to oligomerisation and auto-activation. Another regulatory mechanism involves, PPP, the only mycobacterial protein phosphatase which was found to dephosphorylate PknB (Boitel et al., 2003) and PknA (Chopra *et al.*, 2003). It can be hypothesised that a third mechanism of intracellular regulation of these STPK is via the FHA domain containing proteins in this operon.

It is important to note that this operon is conserved throughout the mycobacterial genus (100% identity in *M. bovis* BCG) and in many other Gram-positive actinobacteria, including *Streptomyces coelicolor* and *Corynebacterium glutamicum* (Boitel et al, 2003). The fact that this operon is located near the origin of replication and is continuously transcribed in animal models indicates that it is essential for normal growth of *M. tuberculosis*. Together these data indicate PknA and PknB are involved in a mechanism for the regulation of cell shape, one that may be conserved among gram-positive actinobacteria.

1.8.3 PknD

Currently, little is known about the function of PknD but genomic and expression data suggest that it plays a role in sensing levels of extracellular phosphate (Av-Gay and Everett., 2000). An important insight into the function of extracellular sensor domains of *M. tuberculosis* STPKs was provided by Alber and colleagues with the X-ray crystal structure of the PknD extracellular sensor domain (Good *et al.*, 2004). This domain forms a six-bladed β-sheet fold with a flexible tether to the transmembrane domain. It is similar many other β-propeller motifs with diverse functions and remains the only structure of β-propeller domain linked to a kinase solved to date. The structure contains a 'cup' motif that is the site for ligand/protein-binding surface for many other β-propeller structures (Good *et al.*, 2004). Although the domain seems able to bind extracellular ligands, it is difficult to see how a signal is transmitted to the intracellular kinase domain. Close inspection of the structure suggests there is little conformational change upon ligand binding and the flexibility of the transmembrane linker limits the propagation of any conformational change.

Further work by Grunder and co-workers showed, through the use of *in vitro* kinase assay, PknD was able to phosphorylate one of the Rv1747 FHA domains (Grunder *et al.*, 2005).

1.8.4 PknE and PknF

Initial characterisation of PknE has determined that the protein is located in the membrane fraction (Molle *et al.*,2003). Further work by Grunder and co-workers showed, through the use of *in vitro* kinase assay, PknE and PknF were able to phosphorylate Rv1747 FHA domains and PknF could also phosphorylated Rv0020 (Grunder *et al.*, 2005). Koul *et al* (2001) also determined that PknF was capable of autophosphorylation. More recently, Deol et al (2005) carried out PknF expression studies in *M. tuberculosis* using an anti-sense strategy. A reduction in PknF expression caused faster growth, aberrant septum formation and increased uptake of D-glucose. Over-expression led to a reduced growth rate and swelling of cells. Rv1747 is encoded by a gene that is present in the same operon as PknF and close to PknE (Figure 1.9).

Rv1747 is an ATP-binding cassette transporter protein (ATP) with two N-terminal FHA domains.

Molle *et al.*,2003, have shown that the recruitment and phosphorylation of Rv1747 by PknF is dependent on the two FHA domains of Rv1747. Buxton and colleagues later presented evidence to demonstrate that the growth of an Rv1747 knockout strain of *M. tuberculosis* was reduced *in vivo* but normal *in vitro* (Curry *et al.*, 2005).

1.8.5 PknG

Koul and co-workers found PknG localised to the cytosol of cells (Nugyen *et al.*, 2005). This is consistent with the fact that PknG along with PknK are the only STPK in *M. tuberculosis* that lack a membrane spanning domain. The same group also demonstrated that PknG was secreted within macrophage phagosomes and inhibited phagosome-lysosome fusion thereby mediating intracellular survival of mycobacteria (Walburger *et al.*, 2004). PknG is expressed in *M. bovis* Bacille Calmette-Guérin (BCG) as well as *M. tuberculosis*, but interestingly not in the non-pathogenic mycobacterium, *M. smegmatis*. Macrophages infected with PknG-expressing *M. smegmatis* prevented lysosomal transfer whereas the vast majority of wild-type bacteria were found within the lysosome. Significantly, an inhibitor specific to PknG, tetrahydrobenzothiophene, has been identified and is able to cause a dose-dependent inhibition of *M. tuberculosis* intracellular growth with additional effects on viability outside host cells (Walburger *et al.*, 2004).

Cowley *et al.*, 2004, demonstrated that PknG can autophosphorylate within the kinase domain and its C-terminal tetratricopeptide (TPR) domain. They also demonstrated a direct correlation between PknG expression and glutamate/glutamine levels in *M. tuberculosis*. This concurs with the fact that the PknG gene is encoded in an operon that includes *glnH*, a glutamine binding lipoprotein. However, Koul and colleagues concluded that in *M. bovis* BCG, PknG does not regulate glutamine metabolism (Nugyen *et al.*, 2005). This may suggest differing role for PknG among *M. tuberculosis* and *M. bovis* BCG.

1.8.6 PknH

Molle *et al* (2003) presented evidence for an interaction between the STPK, PknH (Rv1266), and EmbR (Rv1267), an FHA domain containing protein in *M. tuberculosis*. This has further strengthened the case for functional linkage between the STPKs and FHA domain-containing proteins. They also determined that the FHA domain of EmbR is essential for the PknH-EmbR interaction and that abrogation of this interaction by mutation of conserved S85 to alanine in the FHA domain of EmbR prevents phosphorylation by PknH. Furthermore, a study by (Sharma *et al.*,2004), determined that under stress conditions such as low pH and heat shock, PknH was able to autophosphorylate and was downregulated, suggesting a role in adaptation to the host environment. Indeed, PknH knockout *M. tuberculosis* had increased resistance to acidified nitrite stress. The knockout strain was also found to replicate to a higher bacillary load in mouse organs than a wild type strain of *M. tuberculosis*.

EmbR is a transcription factor that is able to modulate the level of arabinosyltransferase activity in *Mycobacterium avium*. This enzyme is involved in the biosynthesis of arabinan. Arabinan is a component of the mycobacterial cell wall as a constituent of arabinogalactan and lipoarabinomannan (LAM). The anti-tuberculosis drug ethambutol (EMB) has been shown to target arabinosyltransferases encoded by the adjacent genes *emb*C, *emb*A and *embB*. Mutations in EmbR have been linked to EMB resistance in human isolates of *M. tuberculosis* (Sreevatsan *et al.*,1997).

1.8.7 Other STPKs in *M. tuberculosis*

No information has been gathered as to the function or properties of any other STPKs in the *M. tuberculosis* except limited characterisation of PknI. Here, standard cloning, expression and characterisation of PknI confirm that it is a functionally active kinase able to autophosphorylate on serine and threonine residues in a manganese-dependent manner (Gopalaswamy *et al.*, 2004).

1.8.8 GarA

Rv1827 is also referred to as GarA (glycogen-accumulation regulator A). This FHA domain-containing protein is almost entirely taken up by its FHA domain. GarA is found in the *M. tuberculosis* cell as both a full-length protein and in a truncated form (residues 31-162) and both were found to induce protective immunity in animal models (Weldingh and Andersen., 1999). The storage of glycogen is necessary for the growth of *Mycobacterium smegmatis* and GarA was determined to be essential for the regulation of this process. Further functional information was provided by a GST pull-down experiment that identified candidate binding partners for GarA, NAD-dependent glutamate dehydrogenase (Rv2476) and ferrodoxin-dependent glutamate synthase (Rv3859)(Westcott S., unpublished data).

The optimal phospho-peptide motif selected for by the GarA FHA domain has a pThrXXY motif (Durocher *et al*, 2000). The pThrXXY phospho-peptide ligand was determined to bind to GarA FHA with a K_D of 0.23μM by isothermal titration calorimetry (ITC). The GarA FHA demonstrates a preference for bulkier hydrophobic/polar residues at the pThr + 3 position of its phospho-peptide ligand in a manner analogous to Rv0020 FHA.

Although there is no obvious associated STPK for GarA, PknG is linked to cellular glutamate/glutamine levels in *M. tuberculosis* and hence may be functionally associated with GarA (Cowley *et al.*,2004). Villarino *et al*, 2005, demonstrate that GarA is an *in vitro* substrate for PknB, D, E and F (Grundner *et al,* 2005).

1.9 *Mycobacterium* and Tuberculosis

The genus *Mycobacterium* consists of rod-shaped Gram-positive bacteria. Most of the 60+ species are harmless although a subset of this genus is able to cause disease in humans and animals. The most severe of these human diseases are tuberculosis and leprosy caused by *M. tuberculosis* and *M. leprae*, respectively. Robert Koch, discovered that *M. tuberculosis* was the causative agent of tuberculosis (TB) in 1882.

Despite extensive investigation into tuberculosis, which has led to a detailed understanding of the disease mechanism, it remains a global epidemic. Mycobacteria are facultative intracellular pathogens that preferentially infect macrophages and multiply within their phagosomes. Infection occurs predominantly in the lungs, but can progress to infection of other anatomical sites. Most people who carry the bacteria suffer no obvious symptoms and cannot pass the disease to others during this acute phase of the infection. Both *M. tuberculosis* and *M. leprae* are essentially non-toxic to the host cells and much of the tissue damage comes about from the persistent activation of macrophages. The chronic nature of the infection leads to sever immunopathology with the formation of granulomas ('tubercles' in TB). Granulomas are small necrotic structures in which bacteria are sequestered in sheaths of phagocytes and lymphocytes. In most cases, acute infection does not occur and the bacteria may remain dormant, localised to the lungs. One-third of the global population are reported to be carriers of TB. A small subset of individuals become symptomatic leading to a mortality rate of 50% if untreated. The mechanism by which the bacteria remain dormant is unknown and the precise reason for reactivation is unclear. Factors predisposing individuals to reactivation may be stress, overcrowding, malnutrition or hormone imbalance.

Throughout history, TB has accounted for more deaths than any other single infectious disease, with the exception of malaria. The World Health Organisation (www.who.int) estimate that of the 6 billion people in the world, approximately 2 billion are infected with *M. tuberculosis* and that each year about 8 million new cases of TB develop. The bacterium causes active disease in between 5 and 10% of those infected. The highest prevalence of TB is in developing countries, which account for 95% of cases

worldwide. In 2002, the disease killed nearly 2 million people, the vast majority of them in low- and middle-income countries (latest WHO data reveals one death from TB every 15 seconds worldwide and in India tuberculosis kills one person every minute). More than half of cases occur between the ages of 15 and 49, leading to significant socio-economic impact. Infections with HIV increase the likelihood that latent TB will develop into active disease, amplifying the effects of TB worldwide. In fact TB kills more AIDS patients than any other infection, and costs the global economy $16 billion each year. Although 95% of the burden falls on developing countries globalisation is fuelling its spread internationally and no country is immune. This has been recognised by the G8 countries, which issued a resolution to halve the morbidity from tuberculosis within 10 years. Recently the WHO sounded a warning over the growing prevalence of drug resistant TB and noted that the worst affected areas were Eastern Europe and Central Asia with as many as 14% of cases in these areas involving multi-drug resistant (MDR) strains of the bacterium. The former Soviet Union has the largest MDR rate of the world and there is a worrying spread of MDR in China

The vaccine for tuberculosis is not very effective. Treating tuberculosis is hampered by diagnostic tests that fail to identify at least half of cases, and by drug treatments that must continue for at least six months to be effective. The difficulty many patients have in completing such long treatment has contributed to new strains of TB that are resistant to many drugs. Since research and development for new TB drugs came to a virtual standstill after the 1960s, TB treatment relies on drugs that are 40-50 years old. This is because the TB market lacks sufficient incentives to support a pipeline of continually improving drugs. Even with the market potentially reaching $700 million by 2010, it is concentrated in poor countries and no single pharmaceutical company has pursued the full development of an anti-TB drug. Standard therapy for pulmonary TB includes isoniazid and rifampicin for 6 months, along with pyrazinamide and ethambutol for the first two months (isoniazid and rifampicin without pyrazinamide and ethambutol may be used for 9 months if necessary). Occasionally used are streptomycin, thiacetazone, rifabutin and rifapentin. The costs of curing MDR-TB are up to 1400 times the cost of regular treatment. Since MDR-TB cannot be adequately treated through the standard short course of therapy, treatment entails up to 2 years of treatment with "second line" drugs including cycloserine, *para*-aminosalicyclic acid (PAS), ethionamide, amikacin, kanamycin and capreomycin, which is much too long to ensure compliance and currently cure rates are limited to 60%. It is hoped that with a new drug and a shortened

treatment regimen of less than two months, then cure rates of 90% could potentially be achieved (WHO).

1.10 The biological perspective

M. tuberculosis is reported to have a large number of eukaryotic-like proteins (Krupta *et al.*, 2005) and therefore has developed one of the most complex signalling systems of all bacteria. This intricate signalling system facilitates the bacterium's ability to lead a complex life cycle and exact such a devastating disease. The primary focus of this project is a putative operon that contains two FHA domain-containing proteins (Rv0019 and Rv0020) along with two STPKs (PknA and PknB) plus a differentially expressed FHA domain-containing protein Rv1827 (GarA) (Figure 1.11). The co-expression of FHA domain-containing proteins and STPKs strongly suggests that these bacterial FHA domain-containing proteins engage in phospho-dependent protein-protein interactions, and FHA dependent processes in bacteria are controlled by STPK-dependent phosphorylation. In addition, the fact that STPK domains and FHA domains are covalently linked in some eukaryotic proteins (i.e. Rad53, see above) implies functional linkage.

Initial yeast two-hybrid screens carried out by the Division of Mycobacterial Research, NIMR, demonstrated an interaction existed between PknA and Rv0020, and PknB and Rv0019. These data provided the opportunity to study the relationship between FHA domains and STPKs in the context of mycobacterial signalling. These interactions may form the basis of a mechanism of phosphorylation dependent control that is a precursor to systems utilised by eukaryotes.

The aim of this study was to further characterise the interactions between PknA and Rv0020, and PknB and Rv0019 along with interaction with Rv1827 through the application of a number of biophysical and biochemical methods. This investigation concentrated on the interrelationship between these two important STPKs and FHA domains. It is hoped that a greater understanding of how prokaryotic STPK activities are regulated may identify drug targets.

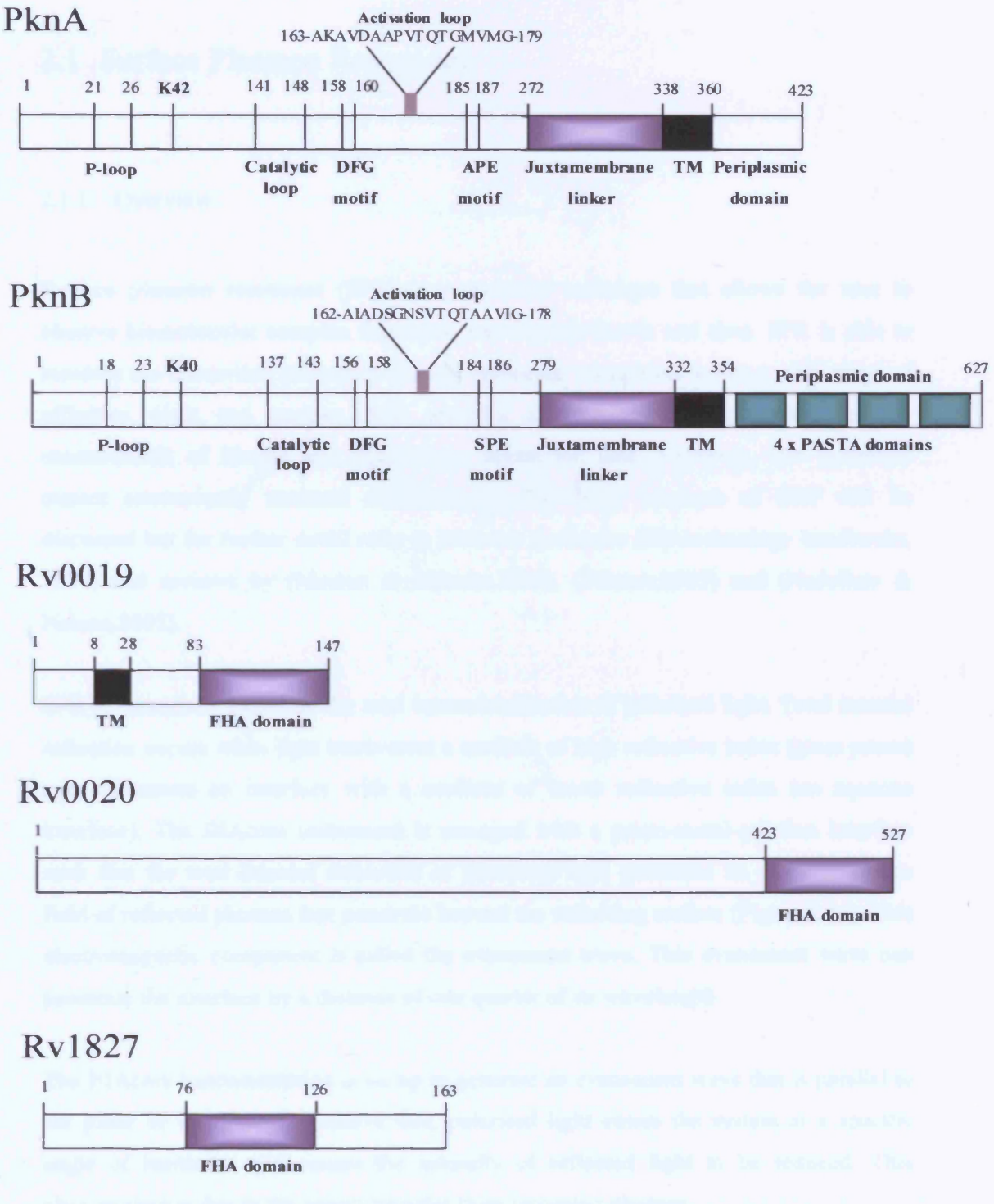


Figure 1.11. Schematic representation of proteins used in this study. *Mycobacterium tuberculosis pknA, pknB and Rv0020* encode predicted transmembrane proteins with a single transmembrane helix (TM) connecting a periplasmic domain to a cytosolic domain. *pknA* and *pknB* encode typical kinases of the STPK family with conserved motifs (shown above). All domains were predicted using the SMART (Simple Modular Architecture Research Tool) database (www.smart.embl-heidelberg.de).

2 Quantitative techniques in protein-complex interactions

2.1 Surface Plasmon Resonance

2.1.1 Overview

Surface plasmon resonance (SPR) is a powerful technique that allows the user to observe biomolecular complex formation and dissociation in real time. SPR is able to measure the interaction properties of multi-molecular complexes across a wide range of affinities, sizes and purities. SPR provides a rapid and robust method for the measurement of kinetic and equilibrium interaction data. However, this technique cannot conveniently measure stoichiometry. The basic concepts of SRP will be discussed but for further detail refer to BIAcore textbooks (BIAtechnology handbooks, 2004) and reviews by (Morton & Myszka,1998), (Winzor,2003) and (Nedelkov & Nelson,2003).

SPR is essentially based on the total internal reflection of polarised light. Total internal reflection occurs when light transverses a medium of high refractive index (glass prism) and encounters an interface with a medium of lower refractive index (an aqueous interface). The BIAcore instrument is arranged with a prism-metal-solution interface such that the total internal reflection of incoming light generates an electromagnetic field of reflected photons that penetrate beyond the reflecting surface (Figure 2.1a). This electromagnetic component is called the evanescent wave. This evanescent wave can penetrate the interface by a distance of one quarter of its wavelength.

The BIAcore instrumentation is set up to generate an evanescent wave that is parallel to the plane of incident. To achieve this, polarised light enters the system at a specific angle of incidence and causes the intensity of reflected light to be reduced. This phenomenon is due to the energy transfer from incoming photons

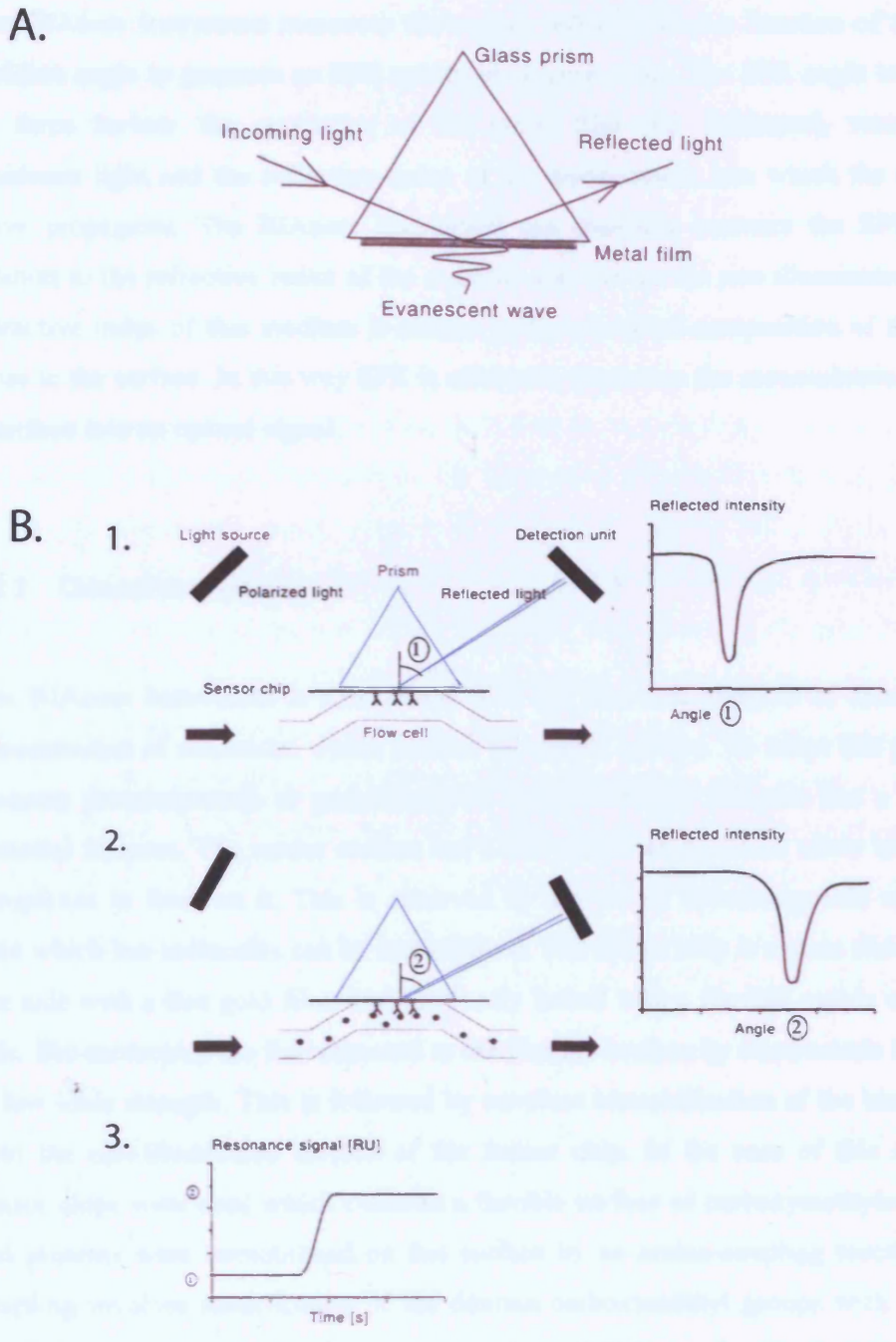


Figure 2.1. Generation and detection of surface plasmon resonance signal. (A) Total internal reflection at the metal coated interface generates an evanescent wave that propagates into the medium of lower refractive index (B) Changes in the refractive index of the surface layer of a solution in contact with the sensor chip are detected by SPR. 1. There is a decrease in reflected intensity at an angle dependant on the refractive index of the medium on the non-illuminated surface. 2. There is a shift in the SPR angle as biomolecules bind to the surface changing the refractive index. 3. A sensorgram is a plot of the SPR angle against time. Adapted from the BIAtechnology handbook.

to surface plasmons (free electrons in the metal film). Thus, this technique is known as surface plasmon resonance.

The BIAcore instrument measures the energy reflectance as a function of the SPR or incident angle to generate an SPR spectrum (Figure 2.1b). The SPR angle is dependent on three factors: the properties of the metal film (i.e. thickness), wavelength of incidence light and the refractive index of the environment into which the evanescent wave propagates. The BIAcore instrument can therefore measure the SPR angle in relation to the refractive index of the aqueous solution on the non-illuminated side. The refractive index of this medium is altered by the chemical composition of the solution close to the surface. In this way SPR is utilised to transduce the accumulation of mass at a surface into an optical signal.

2.2.2 General methodology

The BIAcore instrument is able to use SPR for real-time analysis of changes in the concentration of molecules which contact the sensor surface. To adapt this principle to measure protein/protein or protein/nucleic interactions the BIAcore has a number of essential features. The sensor surface has a non-illuminated surface allow biomolecular complexes to form on it. This is achieved by the use of interchangeable sensor chips onto which bio-molecules can be immobilised. The sensor chip is a glass slide coated on one side with a thin gold film, and covalently bound with a flexible matrix on the other side. Bio-molecules are first attracted to the flexible surface by electrostatic interactions at low ionic strength. This is followed by covalent immobilisation of the bio-molecules onto the non-illuminated surface of the sensor chip. In the case of this study CM5 sensor chips were used which contains a flexible surface of carboxymethylated dextran and proteins were immobilised on this surface by an amine-coupling reaction. Amine coupling involves modification of the dextran carboxymethyl groups with a chemical reaction involving N-hydroxysuccinimide (NHS) and N-ethyl-N'-(dimethyl-aminopropyl)-carbodimide (EDC). This introduces N-hydroxysuccinate esters to the dextran that then react spontaneously with the nucleophilic groups on the ligand, such as amines, to form covalent links.

Once a bio-molecule is immobilised onto the sensor chip the liquid handling system of the BIAcore instrument allows the controlled transport of analytes over the sensor chip. The BIAcore instrument detects the changes in SPR angle, which can be related to any mass increase or decrease on the surface, including specific and non-specifically bound analyte. These changes in resonance signal as a function of time are represented in a sensorgram plot (Figure 2.2).

2.2.3 Real-time biomolecular interaction detection

The association and dissociation of an analyte to the immobilised ligand is monitored to give accurate kinetic information on the interaction (Figure 2.2). In a typical binding study, the association phase consists of an analyte sample being injected over the immobilised ligand at a constant flow rate. As analytes bind to the immobilised ligand there is a concurrent increase in signal, this gives information on the association-rate of the interaction. At the point where there is no further increase in signal, all binding sites have been saturated and an equilibrium phase is obtained. This is a steady-state point defined by the concentration of the analyte in the sample. The signal at equilibrium phase is used to determine the association constants for this interaction. After equilibrium is achieved, the injection of sample is stopped and buffer is passed over the surface, this begins the dissociation phase. All non-covalently bound material is detached from the chip surface giving information on the dissociation-rate of the interaction.

All binding experiments should be accompanied by the use of a reference surface containing no immobilised protein as a control to ensure reliability of BIA data. The contribution of the injected analyte to the total refractive index can be determined from a sensorgram of a reference cell without immobilised ligand. The first step of data analysis should always be to subtract the response of the

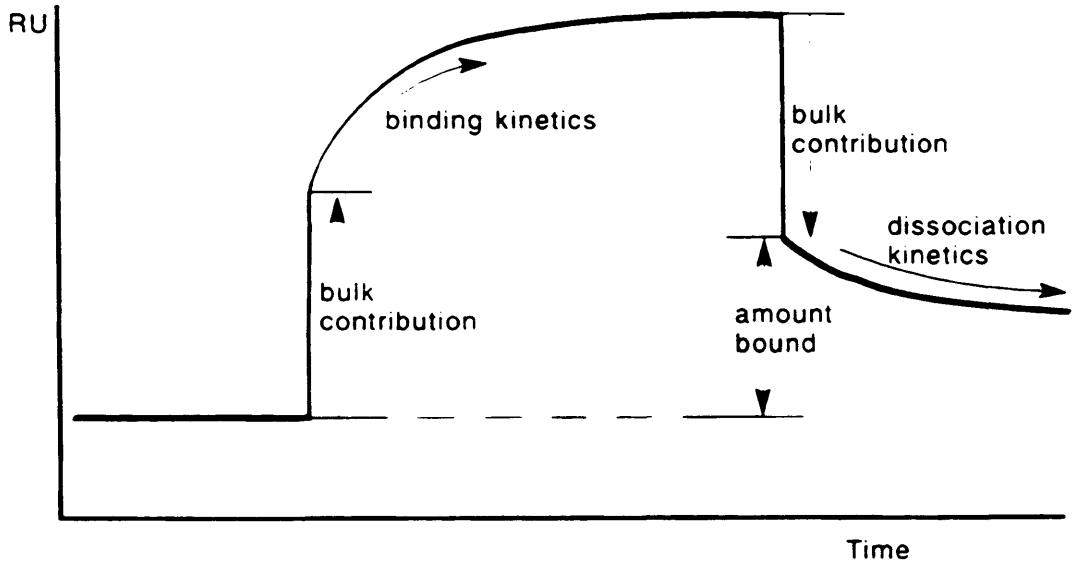


Figure 2.2. Schematic sensorgram from a typical surface plasmon resonance experiment. A sensorgram displays the progress of the interaction at the sensor surface. Adapted from the BIAtechnology handbook.

reference surface from the reaction surface. To obtain the most accurate possible kinetic data for the interaction there are a number of experimental parameters that should be optimised. The experiment should be conducted over a range of analyte concentrations. The concentration of analyte sample should be chosen with the signal to noise limitation in mind. The flow rate should be selected to minimise the effect of the mass-transport limitation (see Section 2.2.6). Faster flow rates can reduce the amount of analyte rebinding during dissociation phase. During dissociation phase, the amount of available binding sites increases leading to rebinding of the analyte. Consequently, the apparent dissociation-rate is lower than the actual rate.

2.2.4 Application areas

SPR techniques have been applied to different areas, such as the studies of interactions involving protein:protein complexes, protein:nucleic acid complexes and protein:compound complexes. Generally the BIAcore instrument is a versatile tool to address a variety of questions:

- The instrument provides a way of screening binding activities from samples without the need for labelling
- The k_{on} and k_{off} rate constants can be determined from the association and dissociation phases of a sensorgram and are often more informative in biological settings than equilibrium constants
- Binding of an analyte to a ligand can be monitored in the presence of a preexisting analyte complex and an immobilised ligand.
- The relative binding affinities and rate constants of a series of analytes can be evaluated

2.2.5 Data analysis

2.2.5.1 Kinetic measurements

Kinetic rate constants are determined with respect to the chosen interaction model for the system. This interaction model should be selected based on background biological knowledge and/or previous biophysical experiments. In general, complex models should only be applied when simple models do not provide a suitable fit to data. The 1:1 binding model, was utilised for this study, and will be discussed here. The model assumes rapid mixing of the analyte with the sensor surface layer in the bulk phase and a single binding step.

A homogenous interaction with a stoichiometry of 1:1 between analyte A and an immobilised ligand B is described below:

$$A + B \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} AB \qquad (1)$$

Where, k_{on} is the association rate constant and k_{off} is the dissociation rate constant. In the association phase, the rate of complex formation is described below:

$$\frac{d[AB]}{dt} = k_{on}[A][B] - k_{off}[AB] \qquad (2)$$

As the concentration of the complex [AB] is measured by the BIAcore instruments as the response, R, at time, t, to express equation **(2)** in terms of SPR signal:

$$\frac{dR}{dt} = k_a[A](R_{max} - R) - k_d R \qquad (3)$$

Where, dR/dt is the rate of change of the SPR signal and R_{max} is the maximum analyte binding capacity. If the expression (k_{on}[A]+k_{off}) is termed k_s, then in accordance with this model the slope of the plot of dR/dt against R will be k_s. A series of experiments are carried out at a range of analyte concentrations will give a range of k_s values. When these k_s values are plotted against the analyte concentrations a straight line should be generated. The slope of the plot is the k_{on} and the intercept of the Y-axis defines the k_{off} value.

During the dissociation phase there is no analyte addition and equation **(2)** is modified to:

$$\frac{dR}{dt} = -k_{off} R \qquad (4)$$

the integrated logarithmic form of equation **(4)** is:

$$\ln \frac{R_0}{R_t} = k_{off}(t - t_0) \qquad (5)$$

where, R_0 is the response at an arbitrary starting time , t_0. If $\ln(R_0/R_t)$ is plotted against $(t\text{-}t_0)$, a straight line will be generated with the slope being k_{off}.

2.2.5.2 Equilibrium measurements

If the interaction is able to reach a steady state, or equilibrium, then equilibrium constants may be determined from BIAcore data using the following scheme.

The equilibrium constants are defined by the following two formulae:

$$K_D = \frac{[A][B]}{[AB]} \qquad (6)$$

$$K_A = \frac{[AB]}{[A][B]} \qquad (7)$$

where K_A is the association constant and K_D is the dissociation constant. Equations **(6)** and **(7)** can be applied to BIAcore data. The concentration of free analyte is equal to the concentration of analyte sample injected. If the surface binding capacity is known, the concentration of free ligand can be derived from the concentration of the complex. The concentration of the complex is the response of equilibrium.

$$K_A = \frac{R_{eq}}{C(R_{max} - R_{eq})} \qquad (8)$$

where, C is the concentration of analyte, R_{max} is the total surface binding capacity and is the steady state binding level. R_{max} and R_{eq} are both expressed in terms of response units (RU).

2.2.6 Complexities and artifacts

Sensorgrams that do not fit a homogeneous interaction with a 1:1 stoichiometry may indicate a more complex mode of interaction or maybe due to artefacts of the measurements. Several frequently observed causes of deviation are discussed in this Section.

The kinetic model in equation **(1)** assumes that the observed rate of binding reflects the interaction kinetics between the analyte and ligand. In practice, the mass transport of analyte to the sensor chip surface can be a limiting factor and to take account of this factor, the reaction should be described as:

$$A_{bulk} \underset{k_m}{\overset{k_m}{\rightleftharpoons}} A_{surface} + B \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} AB \qquad (9)$$

where k_m is the rate constant for mass transport to and from the surface assuming that the rate constant for mass transport is the same in both directions.

During the dissociation phase, if the mass transport is much faster than association of analyte with ligand, the observed binding will be governed by the kinetics between the analyte and the ligand. On the other hand, if mass transport is slower than association, the mass transport process will limit the binding and the apparent on rate will be slower than the true value. During the dissociation phase, analyte molecules removed from the surface by dissociation can either be transported to the bulk solution or rebind to free ligand sites on the surface. The observed dissociation rate is the result of competition between these two processes, and will reflect the dissociation rate constants only when mass transport is fast compared to dissociation.

When mass transport is limiting, the plots of *dR/dt* versus *R* curve downwards instead of being linear. It is important to be aware that this downwards curvature may also be caused by heterogeneity in either the ligand or analyte. If ligands with more than one binding constant are present on the same surface, the observed sensorgram would be the sum of two or more independent sensorgrams, one for each ligand species. At low concentrations, the high affinity sites will dominate the binding, whereas at high concentrations, the low affinity process will become more evident. Ligand heterogeneity may be caused by either the heterogeneity of the starting material or may be created during the immobilisation process.

2.2 Isothermal Titration Calorimetry

2.2.1 Overview

To fully understand the nature of a protein recognition event and the forces that stabilise the interaction, the thermodynamics of association should be characterised. The thermodynamics of association are define by the stoichiometry of the interaction (n), the binding constant (K), the Gibbs free energy of binding (ΔG_b), enthalpy of binding (ΔH_b), entropy of binding (ΔS_b) and heat capacity of binding (ΔC_p). Isothermal titration calorimetry (ITC) is a now a well established technique that allows the calculation of all but one of these parameters in a single experiment. A typical ITC experiment consists of stepwise injections of one reactant into a reaction cell containing the other reactant. The heat evolved or absorbed in a liquid sample as a result of injecting a reactant is measured.

This system provides equilibrium binding constant, the stoichiometry and enthalpy of binding using the Wiseman isotherm theory (Wiseman *et al.*, 1989). Gibbs free energy of binding and the entropy of binding can subsequently be determined once the association constant is established. The most commonly used instrument for conducting ITC is the VP-ITC MicroCalorimeter from MicroCal Inc. (Northampton MA, USA). The VP-ITC MicroCalorimeter is able to directly measure the heat evolved from the association of a ligand to its binding partner. This technique has been widely used to determine the thermodynamic parameters for a range of biomolecular binding events (Leavitt & Freire, 2001). When used in tandem with structural information, important residues within the binding interface can be identified and their energetic contribution to binding can be established. This review of the basic concepts of ITC is based on the VP-ITC user guide, although more detailed guides to using ITC exist (Ladbury & Chowdhry, 1996).

2.2.2 Wiseman isotherm theory

In this study, the Wiseman isotherm theory will be applied to a binding event with a 1:1 stoichiometry involving single-step binding. The independent variables for a single step binding events are the association constant, K_A, the enthalpy, ΔH_b^0, and the number of binding sites. As mentioned above these parameters are directly measured by the VP-ITC MicroCalorimeter, and details of this process will be discussed later. The dependent variables for binding are Gibbs free energy of binding (ΔG_b) and entropy (ΔS_b). ΔG_b and ΔS_b can be calculated from the following equation:

$$\Delta G_b = -RT \ln K_A = \Delta H_b - T\Delta S_b \qquad (10)$$

A binding model to characterise the change in uncomplexed protein after each successive injection is explained below. A binding event with the stoichiometry of 1:1 is described by the binding equilibrium $M + X = MX$ then:

$$K_A = \frac{[MX]}{[M][X]} \qquad (11)$$

and

$$X_{tot} = [X] + [MX] \qquad (12)$$

and

$$M_{tot} = [M] + [MX] = \frac{[MX]}{K_A[X]} + [MX] \qquad (13)$$

where K_A is the equilibrium association constant, X_{tot} is the total ligand concentration and M_{tot} is the total macromolecule concentration in the system. Equation **(12)** can be expressed in terms of the ligand concentration, $[X]$, and substituted into equation **(13)** to give the following quadratic equation:

$$0 = [MX]^2 + [MX](-M_{tot} - X_{tot} - \frac{1}{K_A}) + M_{tot}X_{tot} \quad \textbf{(14)}$$

If

$$b = -M_{tot} - X_{tot} - \frac{1}{K_A}$$

$$c = M_{tot}X_{tot} \quad \textbf{(15)}$$

then the only real root of equation **(14)** is:

$$[MX] = \frac{-b - \sqrt{b^2 - 4c}}{2} \quad \textbf{(16)}$$

substitution of **(15)** into **(16)**:

$$\frac{d[MX]}{dX_{tot}} = \frac{1}{2} + \frac{1 - (1-r)/2 - X_r/2}{\sqrt{X_r^2 - 2X_r(1-r) + (1-r)^2}} \quad \textbf{(17)}$$

where r is equal to $1/(K_A M_{tot})$ and X_r is equal to X_{tot}/M_{tot}. The change in the concentration of MX can be related to the heat change by:

$$dQ = d[MX] \bullet \Delta H_b \bullet V_0 \qquad \textbf{(18)}$$

where ΔH_b is the enthalpy of binding, V_0 is the reaction cell volume and Q is the heat evolved from the reaction. Then, substitution of equation **(17)** into **(18)** yields the final equation:

$$\frac{1}{V_0(dQ/dX_{tot})} = \Delta H_b \left(\frac{1}{2} + \frac{1-(1+r)/2 - X_r/2}{\sqrt{X_r^2 - 2X_r(1-r) + (1+r)^2}} \right) \qquad \textbf{(19)}$$

Equation **(19)** describes a curve, used originally by Wiseman, 1989, which represents binding isotherms for a single-site interaction (Figure 2.3). The microcalorimeter measures the finite differential heat dQ/dX_{tot}. It should be noted that the differential heat measured is dependent on the value of M_{tot} in relation to the values of K_A and to X_{tot}.

A precise values for the equilibrium association constant, K_A, and enthalpy of binding, ΔH_b, can only be established if M_{tot} is in the appropriate range to yield an informative binding isotherm. The shape of a binding isotherm is very sensitive to changes in the $K_A M_{tot}$ value, designated c, as seen in Figure 2.3. In

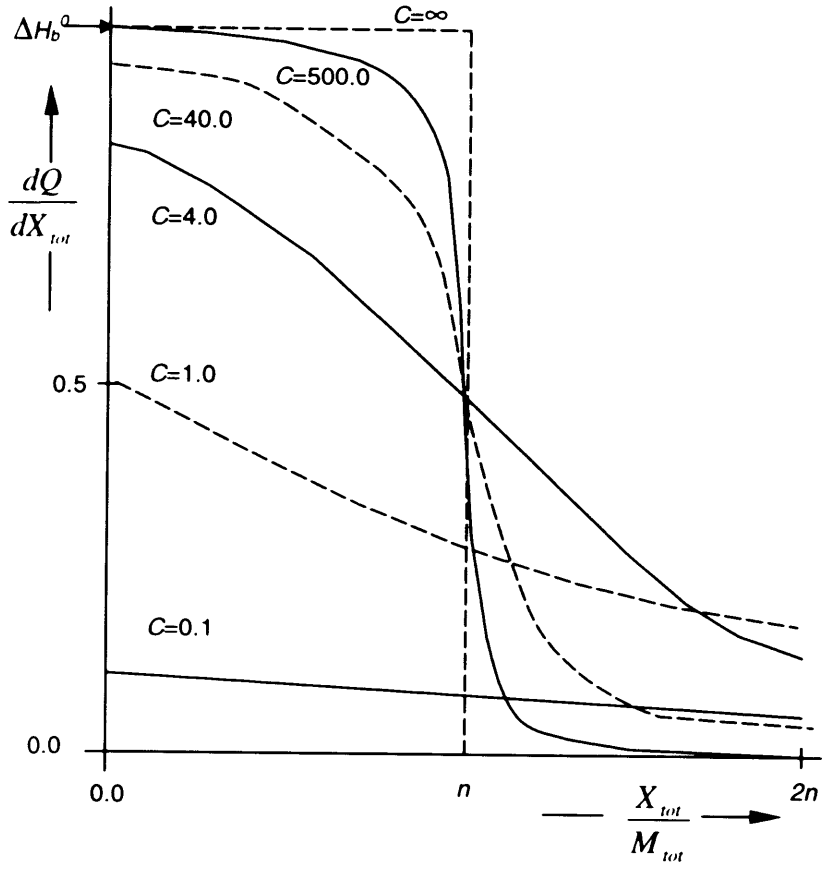


Figure 2.3. Theoretical binding isotherms for a single-site binding interaction. Isotherms for a range of c values (equal to $K_A M_{tot}$) are shown. The plot shows that for c values between 1 and 1000 (experimental K window), small changes in c cause large changes in the shape of the isotherm. For accurate binding constant determination a value of c between 10 and 100 is most desirable. Adapted from BIAtechnology handbook.

other words, reactions with tight K_A must be studied at low macromolecular concentrations and reactions with weak K_A must be studied at high concentrations. Therefore, it is mandatory that the values of c be in the range of 1 to 1000, although c value between 10 and 100 is most desirable. This range is referred to as the 'experimental K window'. A titration with a c value beyond this range causes the isotherm to decrease in curvature and therefore lose information.

2.2.3 General methodology

A schematic of the VP-ITC MicroCalorimeter is shown in Figure 2.4. Samples are introduced to the reaction cell through a narrow access tube using a long-needled syringe. Filtered, degassed water is loaded into the reference cell. The cells must be loaded carefully to avoid introduction of air bubbles and samples should be degassed prior to loading. The second sample is added to the syringe, again in a manner to prevent the acquisition of air bubbles. The injection of sample from the syringe to the reaction cell will follow a pre-set schedule. Specific requirements for schedules in each experiment will be discussed later. All injection schedules have an initial smaller injection to remove any air bubbles and bring the sample in the syringe to the correct volume. In addition, all schedules have a pre-equilibration before the titration commences that lasts 30 minutes and brings both samples to the same temperature.

Both cells have two heaters, and between the cell surfaces, a thermoelectric device is placed to measure the temperature difference ΔT_1. The temperature difference between the inner shield and the outer cell surfaces, ΔT_2, is measured and fed to a heater on the jacket to maintain its temperature close to the temperature of the cells. There is also an outer shield that is maintained at the same temperature in order to minimise fluctuations due to room temperature changes.

The ITC measurement works on a heat compensation principle (Leavitt and Freire, 2001). Throughout a titration, less than a milliwatt of power is dissipated by the heater of the reference cell. This activates the cell feedback circuit to

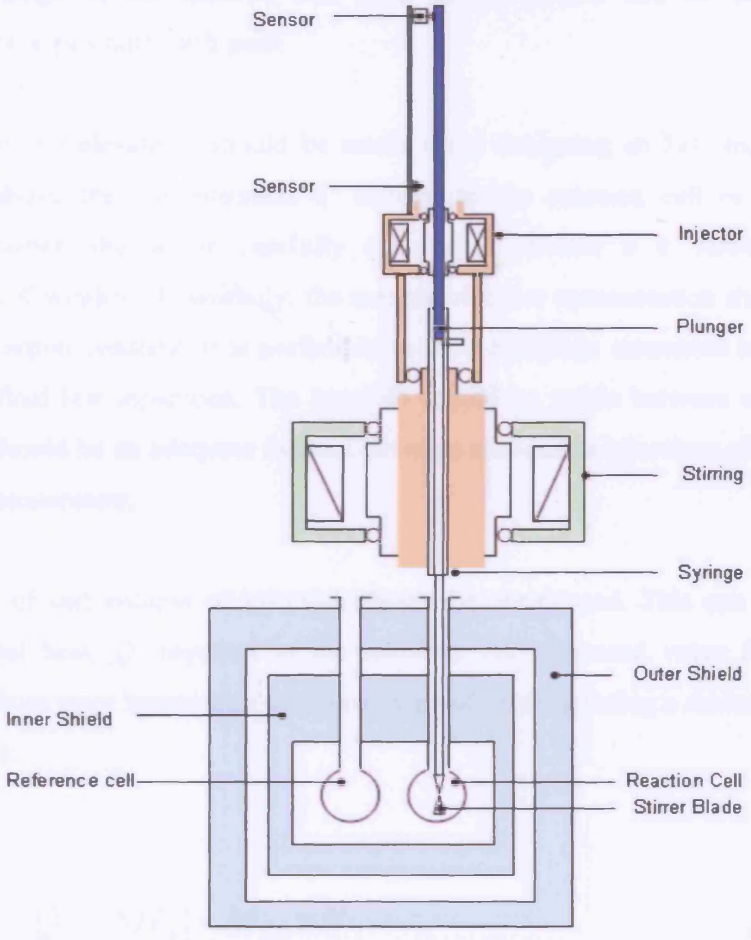


Figure 2.4. Schematic of the VP-ITC instrument. This illustration includes the VP-ITC cells that are kept at a constant temperature by the inner and outer shields (blue). The injector/stirrer assembly (green and brown) allows accurate addition of the ligand through the injection needle, while the stirrer blade ensures complete mixing in the cell throughout the injection period. The reaction cell temperature is maintained at the same temperature as the reference cell by feedback. The change in temperature of the reaction cell during addition of the ligand forms the basis of the ITC measurement.

maintain ΔT_1 at zero. In the absence of an interaction, the feedback power will be a constant that is resting at the baseline value. In the case of an exothermic interaction, there will be a temporary decrease in feedback power causing the reaction cell to return to the baseline value of the reference cell. Conversely, an endothermic reaction will lead to an increase in power to bring the reaction cell to the temperature of the reference cell. The heat change in the reaction cell upon each injection can be determined by integrating the area under each peak.

A number of considerations should be made when designing an ITC experiment. As mentioned above the concentration of sample in the reaction cell of the VP-ITC MicroCalorimeter should be carefully selected to obtain a c value within the experimental K window. Essentially, the macromolecular concentration should be close to the dissociation constant. It is preferable to have complete saturation of the binding sites in the final few injections. The baseline should be stable between each injection and a there should be an adequate interval between successive injections allowing return to baseline temperature.

The number of and volume of injection should be considered. This can be estimated from the total heat, Q, required in the reaction. An estimated value for Q can be determined from prior knowledge of the enthalpy of binding, using a rearranged form of equation **(18)**:

$$Q = \left|\Delta H_b\right| \times M_{tot} \times V \qquad \textbf{(20)}$$

where V is the reaction volume (around 1.3mL for the VP-ITC MicroCalorimeter).

Control titrations can be run to determine the contribution of the heat of dilution of the protein sample to the total heat of the reaction. This is conducted by titrating the a ligand solution into a cell filled with buffer. Conversely, buffer may be titrated into a reaction cell containing the protein sample.

2.2.4 Data analysis

Raw ITC data should initially be corrected by subtracting the heat of dilution. Accurate initial concentrations of samples in the syringe and cell are used to determine the free and bound concentrations of both macromolecules in the reaction cell after each injection. These values are related to the integrated heat signal. This analysis is performed by the VP-ITC data analysis software.

As already described the values for X_{tot} and M_{tot} can be determined using equations **(12)** and **(13)**. When taking equations **(11-13)** into consideration, the complex concentration can be expressed in terms of M_{tot} and K_A:

$$[MX] = [M_{tot}] \times \frac{K_A[X]}{1 + K_A[X]} \qquad \textbf{(21)}$$

As the heat, Q, evolved from the reaction is proportional to the concentration of the complex in an ITC experiment:

$$Q = V\Delta H_b [M_{tot}] \times \frac{K_A[X]}{1 + K_A[X]} \qquad \textbf{(22)}$$

Using floating values of K_A and ΔH_b^0, the raw ITC data can be fitted to Q as a function of $[X]$ using equation **(22)**. Using the approximations of K_A and ΔH_b generated, the ligand concentration is calculated at each point and equation **(22)** is refitted as a function of $[X]$. The fitted curve represents the generation of heat with respect to the ligand concentration. This allows the values of K_A and ΔH_b to be redefined. Further iterations are conducted until constant values for K_A and ΔH_b are achieved.

2.3 Circular Dichroism

2.3.1 Overview

Circular Dichroism (CD) is a widely used technique for the elucidation of protein conformational information. CD is itself an expression of optical activity and all optical activity originates in the absorption bands of dissymmetric compounds. Dissymmetric compounds cannot be superimposed on their mirror image. Proteins are dissymmetric in two ways: Configuration and Conformation.

All amino acids residues with the exception of glycine have a dissymmetric configuration due to the *L* configuration about their α-carbon atom. In addition, threonines and isoleucine residues contain additional dissymmetric centres that increase their optical activity. Conformational dissymmetry is a consequence of the fact that electronic interactions between different residues in a protein also contribute to optical activity and therefore, the secondary structure of a proteins also affects its optical activity. Thus, CD spectroscopy has been widely used to study proteins because it is remarkably sensitive to molecular conformation. This Section reviews the basic methodology of CD spectroscopy and some of its applications, and is based on a number of excellent reviews on the subject (Yang *et al.*, 1986)(Woody, 1996)(Johnson, 1990).

2.3.2 General methodology

In a conventional protein absorption spectrum, several amino acid side chains (notably Tyr, Trp, Phe, His and Met) absorb light strongly in the far-UV region (below 250nm) mainly through the contribution of the peptide bond (amide chromophore). As CD is an absorption phenomenon, the chromophores that contribute to a CD spectrum are the same for a conventional absorption spectrum. However, the parameter measured in a CD experiment is the difference in absorbance from left and right circularly polarised light:

$$DA = A_L - A_R \qquad (23)$$

Where A_L is the measured absorbance for left circularly polarised light and A_R is the absorbance for right circularly polarised light. Electrons participating in the transition to an excited state sense a dissymmetric environment, giving rise to a difference in absorption of light in each direction. These differences in absorption cause the resultant beam (a recombination of both left and right-handed emergent beams) to form an ellipse. CD spectropolarimeters are capable of measuring this difference in absorption and hence the ellipticity.

As CD is an absorbance measurement, it uses a version of the Beer-Lambert Law:

$$\Delta\varepsilon_{mrw} = \frac{\Delta A}{c_r l} \qquad (24)$$

where ε_{mrw} is the mean residue molar extinction coefficient ($M^{-1}cm^{-1}$), c_r is the mean residue molar concentration and l is the path length. The mean residue molar concentration can be determined by multiplying the concentration of protein by the number of amino acids or peptide bonds.

The far-UV CD bands are principally derived from peptide bond absorption and are sensitive to the secondary structure of the protein (α-helix, β-sheet, β-turn and random/unordered main chain conformations). Far-UV spectra are particular informative in the 170nm to 240nm region, where the characteristic contributions of secondary structure are well known (Figure 2.5). In the context of this study, it is important to note that whilst the content of proteins can also be estimated from CD spectra, there is some degree of uncertainty because β structures are less regular than α helices and contribute less to the CD spectrum.

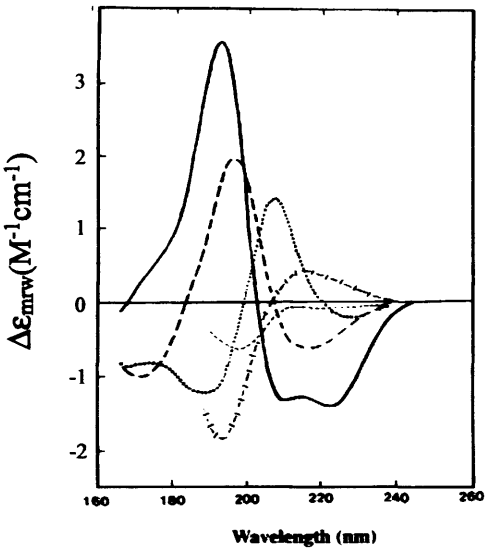


Figure 2.5. Far UV spectra associated with various types of secondary structure. Solid line, α-helix; long dashed line, anti-parallel β-sheet; dotted line, type I β-turn; cross dashed line, extended 3_1-helix or poly (Pro) II helix; short dashed line, irregular structure.

Adapted from Kelly *et al.*, 2005.

2.3.3 CD Applications

Although CD only provides low-resolution structural information, it does have two great strengths: First, it is extremely sensitive to changes in molecular conformation and second, it is applicable in a wide range of solvent conditions and utilises a small amount of protein. In the study of proteins CD spectroscopy is mainly utilised for:

- Determination of protein secondary structure
- Determination of optical purity
- Analysis of the tertiary structure of proteins and conformational changes
- Comparison of the secondary and tertiary structures of wild-type and mutant proteins
- Binding studies involving conformational changes
- Nucleic acid structure and changes upon binding or melting
- Thermodynamic parameters from temperature dependencies of CD signals
- Kinetics of conformational changes on millisecond time scales

Data collection and analysis of CD spectra are outlined in the following Chapter (Section 3.2.5)

3 Materials and Methods

3.1 Materials

3.1.1 Chemicals

Chemicals used in this study were produced from the following companies: Amersham-Pharmacia, Fluka, Novagen, Merck, Pierce, Qiagen, Sigma-Aldrich, and Invitrogen. Contributions from individuals are mentioned specifically.

3.1.2 Enzymes

Benzonase nuclease	Novagen
KOD DNA polymerase	Novagen
Restriction enzymes	New England Biolabs
T4 DNA ligase	New England Biolabs
T4 DNA polymerase	New England Biolabs
PreScission protease	Amersham-Pharmacia

3.1.3 Kits

Qiagen Plasmid Mini Kit	Qiagen
Qiagen Plasmid Midi Kit	Qiagen
Qiagen Gel Extraction	Qiagen
QuikChange Site directed mutagenesis	Stratagene

3.1.4 Chromatography media

DEAE-Sepharose Fast Flow	Amersham-Pharmacia
Glutathione Sepharose 4B	Amersham-Pharmacia
NTA-agarose Superflow	Amersham-Pharmacia
Superdex 75 and 200	Amersham-Pharmacia
Immobilised Metal Affinity Column (IMAC)	Pierce

3.1.5 Microorganisms

E. coli Top10	F- *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) φ80*lac*ZΔM15 Δ*lac*X74 *rec*A1 *ara*D139 Δ(*araleu*) 7697 *gal*U *gal*K *rps*L (StrR) *end*A1 *nup*G (Invitrogen)
E. coli BL21 (DE3)	*B, F⁻, hsdSB (rB⁻, mB⁻), gal, dcm, ompT, λ(DE3* (Novagen)
E. coli Rosetta	*pRARE* containing the tRNA genes *argU, argW, ileX, leuW, glyT, proL, metT, thrT, tyrU,* and *thrU* (Novagen)

3.1.6 Media, special reagents and antibiotics

Luria-Bertani (LB) media	10 g/l Bactotryptone, 10 g/l NaCl, 5 mM NaOH, 5 g/l yeast extract
TAE	40 mM Tris.HCl, pH 8.0; 40 mM Acetic acid; 1 mM EDTA
10X SDS buffer	0.25 M Tris.HCl, 1.92 M Glycine, 1 % SDS
Electroblotting buffer	10x buffer, 22.13g CAPS in 900mL dH_2O, titrate to pH 11 with NaOH and add dH_20 to final volume of 1L.
HBS-EP buffer	10mM HEPES, pH 7.4, 150mM NaCl, 3mM EDTA, 0.005% surfactant P20

Yeast Protein Extraction Reagent (YPER) Pierce

Antibiotics from Sigma were used at the following concentration: 100 μg/mL; Ampicillin and 34 μg/mL; Chloroamphenicol.

Electroblotting buffer 10mM CAPS, pH 11, 10% analytical grade methanol.

3.2 Methods

3.2.1 Recombinant DNA

3.2.1.1 Agarose gel electrophoresis

Agarose gels were prepared and run according to standard procedures (Sambrook, 1989).

3.2.1.2 Isolation of plasmid DNA

DNA isolation was carried out using either QIAprep Spin Miniprep or Midiprep Kits from Qiagen, according to the manufacturer's instructions.

3.2.1.3 Polymerase chain reaction (PCR)

Amplification of DNA fragments was carried out using *KOD* DNA polymerase (Novagen) according to standard procedures (Sambrook, 1989). Fragments were digested according to enzyme specification (NEB Handbook) and purified by 1% agarose gel electrophoresis was followed by extraction using QIAquick Gel Extraction Kit according to manufacturer's protocol.

3.2.1.4 Ligation

Vector and insert DNA were quantified in agarose gels using digested λ-marker (New England Biolabs) as a reference. 100 ng of vector was ligated with a four fold molar excess of insert overnight at 8 °C using T4 ligase (New England Biolabs) according to the manufacturer's protocol.

3.2.1.5 Sequencing of recombinant plasmid constructs

The nucleotide sequence of all synthesised and mutated genes was verified to ensure error-free amplification. Capillary dye terminator sequencing was performed by Cytomyx sequencing services of Cambridge. Sequence analysis was performed using Chromas and DNAstar software.

3.2.1.6 Ligation-Independent Cloning (LIC)

3.2.1.6.1 Overview

The Ligation-Independent Cloning (LIC) method uses a version of the pGEX 6P1 vector (Amersham Biosciences) modified to include a variation of the TEV/LIC linker (Stols et al., 2002) cloned between the BamH1 and EcoR1 restriction endonuclease sites. Using the 3' to 5' exonuclease activity of the phage T4 DNA polymerase, overhanging single-stranded DNA, 15-17 nucleotides in length, is generated in the vector and insert. The exonuclease activity of the DNA polymerase continues until either a cytosine or guanoside nucleotide is reached, provided a complementary dNTP (dC or dG) is present in the reaction mix. Once the base complementary to the unique nucleotide is reached the exonuclease activity ceases as polymerase activity competes with it. LIC-compatible primer and vector sequences are designed to anneal following T4 DNA polymerase digestion. Mixing of the T4 DNA polymerase digested PCR product and vector therefore generates a stable nicked-circular plasmid that can be used directly for transformation without the need for ligation of DNA fragments (Figure 3.1).

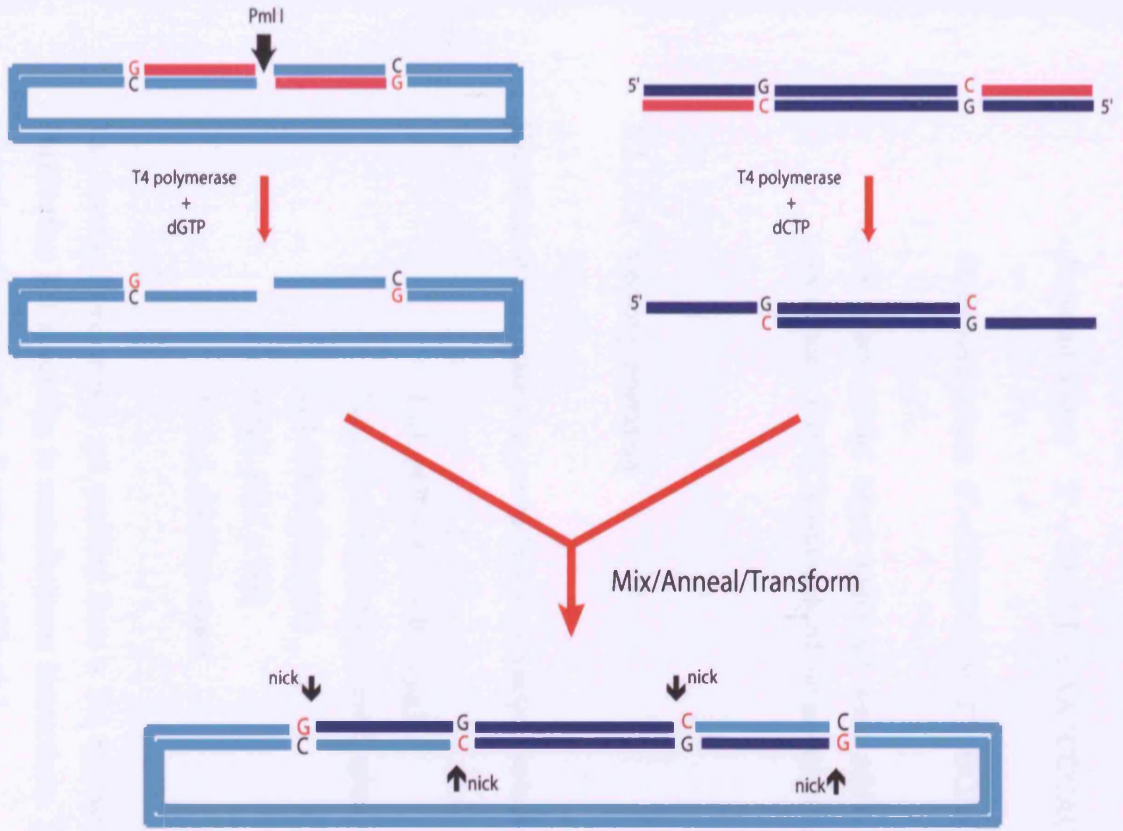


Figure 3.1. Overview of Ligation independent cloning (LIC). 3'-5' exonuclease activity of phage T4 DNA polymerase generates a 15-17 nucleotide single-stranded overhang on both vector and insert. The length of overhang is determined by presence of a single dNTP (dC or dG) in the reaction mixture. This dNTP is added by T4 DNA polymerase to a site opposite the 'unique' G or C nucleotides in the template strand. At this point further exonuclease activity ceases as a 'competition' is set up between the exonuclease and polymerase activities. The sequence of the linker and the 5' and 3' extensions on the PCR product are complementary. Annealing is sufficient to generate a stable nicked-circular plasmid that can be used directly to transform suitable *E. coli* strains with high efficiency.

3.2.1.6.2 Oligonucleotide design

The LIC method requires the following leader sequences to be present in the oligonucleotide primer designs:

Forward primer 5' – TACTTCCAATCCCACGCA..... – 3'

Backward primer 5' – TTATCCACTTCCCACGTTA..... – 3'

18 bases of the target sequence were added to the end of these leader sequences. The PCR was carried out as described previously.

3.2.1.6.3 Vector preparation

Digestion of the vector was carried out in a reaction volume of 50μL containing:

1 μL pGEX 6P1-LIC (1μg)
5μL 10 x PmlI buffer (New England Biolabs)
1μL PmlI (20 units)
0.5μL BSA x 100
42.5μL distilled water

The linerised vector was gel purified from a 1% agarose gel using the QIAGEN Gel Purification Kit according to manufacturers instructions. The linerised vector was then ethanol precipitated and re-dissolved at 300ng/μL.

3.2.1.6.4 T4 DNA polymerase digestion

The linearised vector and the PCR product were digested using T4 DNA polymerase in order to generate the complementary overhangs:

Vector	**PCR product**
0.7pmol PmlI cut vector	0.5pmol PCR product
1μL T4 DNA polymerase (3 units)	0.5μL T4 polymerase
4μL 10x T4 buffer	4μL 10x T4 buffer
1μL 100mM dGTP	1μL 100mM dCTP
Add distilled water to 40μL	Add distilled water to 40μL

The reaction was carried out for one hour at room temperature, followed by 30 minutes at 70°C to heat inactivate the enzyme.

3.2.1.6.5 Annealing

The annealing reaction was carried out in a 9μL volume containing:

1μL vector digest reaction
4μL PCR product digest reaction
5μL distilled water

The annealing reaction was carried out for five minutes at 4°C. 1μL of this reaction was used to transform *E. coli*.

3.2.1.7 Transformation

The heat shock method was used according to the standard protocol (Sambrook, 1989). This involves 30 minute ice incubation of 1 μg DNA and 100-200 μl of chemically competent cells followed by a 45 second heat shock at 42°C and again placed on ice for 15 minutes. 1 ml of LB medium without antibiotic is added to this and incubated at

37°C for 40 minutes with shaking, centrifuged carefully and 100μl cell suspension was plated. Ligated DNA was transformed in *E. coli* Top10 strain, amplified and isolated. The isolated DNA was transformed in *BL21 (DE3)* expression bacteria.

3.2.1.8 Site specific mutagenesis

Site-specific mutagenesis was carried out using the QuickChange kit (Stratagene) according to the manufacturer's protocol.

3.2.1.9 Constructs

Table 3.1. Constructs generated

Construct*	Description	Source
pGBD Rv1827$^{1-162}$	pGBD with a 486bp fragment encoding the entire Rv1827 protein (162 residues)	Provided by J. Currie
pGEX 6P1 Rv1827$^{1-162}$	pGEX with a 486bp BamHI/HindIII fragment encoding the entire Rv1827 protein (162 residues)	*de novo*
pGEX 6P1 Rv1827$^{25-162}$	pGEX with a 411bp BamHI/HindIII fragment encoding the entire Rv1827 protein (137 residues)	Provided by S. Westcott
pGEX 6P1 Rv0019$^{1-156}$	pGEX with a 468bp BamHI/EcoRI fragment encoding the entire Rv019 protein (156 residues)	Provided by J. Currie
pGEX 6P1 LIC Rv0019$^{29-156}$	pGEX with a 381bp LIC fragment encoding the intracellular segment of Rv0019 protein (127 residues)	*de novo*
pGEX 6P1 LIC Rv0019$^{57-156}$	pGEX with a 297bp LIC fragment encoding the FHA domain of Rv0019 protein (99 residues)	*de novo*
pGEX 6P1 Rv0020$^{1-527}$	pGEX with a 2,046bp EcoRI/XhoI fragment encoding the entire Rv0020 protein plus 155 residues prior to the protein start site (682 residues)	Provided by J. Currie
pGEX 6P1 Rv0020$^{155-527}$	pGEX with a 1,581bp EcoRI/XhoI fragment encoding the entire Rv0020 protein (372 residues)	*de novo*
pGEX 6P1 LIC Rv0020$^{423-527}$	pGEX with a 312bp BamHI/XhoI fragment encoding the FHA domain of Rv0020 protein (372 residues)	Provided by S. Westcott
pGEX 6P2 PknB$^{1-626}$	pGEX with a 1,878bp BamHI/XhoI fragment encoding the entire PknB protein (626 residues)	Provided by J. Currie
pGEX 6P1 LIC PknB$^{1-332}$	pGEX with a 996bp LIC fragment encoding the intracellular segment of PknB protein (332 residues)	*de novo*

pGEX 6P1 LIC PknB$^{1-279}$	pGEX with a 837bp LIC fragment encoding the kinase domain of PknB protein (279 residues)	*de novo*
pGEX 6P1 PknA$^{1-431}$	pGEX with a 1,293bp BamHI/XhoI fragment encoding the entire PknA protein (431 residues)	Provided by J. Currie
pET TOPO 151 PknA$^{1-290}$	pET TOPO 151 fragment encoding the entire PknA protein inserted via the TOPO system (290 residues)	*de novo*

*all pGEX vectors encode GST tagged proteins and pET vectors encode six histidine tagged proteins.

3.2.1.10 Point mutants

Table 3.2. Point mutations generated

Construct	Point Mutants
pGEX 6P1 LIC Rv1827$^{1-162}$	S95A (*de novo*)
pGEX 6P1 Rv1827$^{25-162}$	S95A (Provided by S. Westcott)
pGEX 6P1 Rv0020$^{155-527}$	S473A (*de novo*, sense and antisense primers provided by S. Westcott)
pGEX 6P1 Rv0020$^{432-527}$	S473A (*de novo*)
pGEX 6P1 LIC PknB$^{1-279}$	K40A, T171A, T173A, T171A/T173A (*de novo*)
pET TOPO 151 PknA$^{1-290}$	K40A (*de novo*)

3.2.2 Protein expression and purification

3.2.2.1 SDS-PAGE

Separation of proteins of different molecular weight was performed according to (Laemmli, 1970) using denaturing, discontinuous SDS-polyacrylamide gel electrophoresis (SDS-PAGE).

3.2.2.2 Determination of protein concentration

Protein concentration was determined by measuring absorbance of the protein sample at 280nm and dividing the A_{280} by the calculated molar extinction coefficient for each protein and the path length.

3.2.2.3 Electrospray ionisation mass spectrometry (ESI-MS)

Protein molecular weight was determined using a stand-alone syringe pump (Perkin Elmer, Foster City CA) coupled to a Platform electrospray mass spectrometer (Micromass, Manchester, UK). Samples were desalted on-line using a 2mm x 10mm guard column (Upchurch Scientific,Oak Harbor WA) packed with 50 micron Poros RII resin (Perseptive Biosystems, Framingham) inserted in place of the sample loop on a rheodyne 7125 valve. Proteins were injected onto the column in 10% acetonitrile, 0.10% formic acid, washed with the same solvent and then step eluted into the mass spectrometer in 70% acetonitrile, 0.1% formic acid at a flow rate of 10µl/min. The mass spectrometer was calibrated using myoglobin. Experiments were carried out by Dr. Steve Howell, National Institute for Medical Research, Division of Protein Structure.

ESI MS results in proteins gaining multiple positive charges upon ionisation. As each protein molecule differs in the number of added protons, the function of the mass spectrometer is to resolve the mass/charge (*m/z*) and measure the intensities for each set of differently charged protein molecules. The data is plotted in the form of intensity versus *m/z* from which the molecular mass can then be determined using the following equation:

$$m/z = \frac{\left(M + nH^{+}\right)}{n} \qquad (26)$$

where *n* is the integer number of charges on the ions and H is the mass of a proton (1.008Da). As *n* is usually known, simultaneous equations can be used to determine the number of charges by assuming the peaks (*m/z* of differently charged protein molecules) are separated by a single charge. *n* can thus be defined by the equation:

$$n = \frac{m/z(n+1) - H^+}{m/z(n) - m/z(n+1)} \qquad (27)$$

A number average can thus be determined for n by solving for all peaks and the molecular mass of the protein then calculated by substitution of n back into equation **(27)**.

3.2.2.4 Test expression of proteins

To test expression and solubility of proteins expressed from pGEX or pET vectors, plasmids containing the desired insert were transformed into expression bacteria (*BL21-DE3* and *BL21-DE3 Rosetta*). A 50 ml bacteria culture in LB medium was induced at of OD_{600} of ~0.4-0.6 with 500 μM IPTG and grown at 37 °C overnight. This was repeated in triplicate with inductions using 50 μM and 200 μM IPTG and grown at 18 °C overnight.

To test expression, 1 ml of the overnight culture was pelleted, resuspended in water and analysed using SDS-PAGE. To test the solubility, 1 ml bacteria culture was pelleted, resuspended in 100μl Y-PER reagent (Pierce), 5U/mL benzonase, 2mM $MgCl_2$ and lysed for 3 min at room temperature. The solution was centrifuged at 25000 X g for 10 min at 4 °C. Fractions of 5 and 10 μl of the supernatant were analysed by SDS-PAGE.

3.2.2.5 Protein overexpression and preparation of soluble bacteria extract

Proteins were overexpressed using the parameters described in Table 3.3.

Table 3.3. Expression conditions

Plasmid	*E. coli* strain	Antibiotics	Induction at OD_{600}	T / °C	Expression time / h
pGEX 6P1 Rv1827 (all constructs)	BL21DE3	Amp	0.6	25	16
pGEX 6P1 LIC $Rv0019^{29-156}$	BL21DE3	Amp	0.8	18	16
pGEX6P1 LIC $Rv0019^{57-156}$	Bl21DE3	Amp	0.6	25	16
pGEX6P1 $Rv0020^{432-527}$	BL21DE3	Amp	0.6	25	16
pGEX6P1 $Rv0020^{155-527}$	Rosetta	Amp, Cam	0.8	18	16
pGEX 6P1 LIC PknB (all constructs)	Rosetta	Amp, Cam	0.8	18	16
pET TOPO 151 $PknA^{1-290}$	Rosetta	Amp, Cam	0.9	18	16

In all the cases, bacteria were grown at 37 °C until the desired optical density (O.D.) was reached. Upon IPTG induction, temperature was reduced to increase solubility of protein. After expression, bacteria were pelleted at 4000 x *g*, resuspended in their respective buffer (see individual protein purification methodology) plus one EDTA-free protease inhibitor tablet (Roche), 5U/mL units Benzonase, 2mM $MgCl_2$ and 5% YPER. Cells were broken by 30-second bursts of sonification on ice using a Sonifier 450 (Branson Ultrasonics, Danburry, USA) at 40% duty cycle. Insoluble material was

removed by centrifuging at 20,000 x g for 45 min, and the supernatant used for further purification.

3.2.2.6 Purification of Rv1827

Limited proteolysis experiments were carried out to define the FHA domain boundaries in Rv1827 and a stable fragment of 138 residues was identified. For this study both wild-type and truncated form of Rv1827 were purified. All constructs of Rv1827 and Rv1827 S95A were purified according to (Durocher *et al*, 2000). Bacteria were broken in Rv1827 buffer (300mM NaCl and 50mM Tris.HCl pH 8.0) with addition of one EDTA-free protease inhibitor tablet (Roche), 5U/mL Benzonase, 2mM $MgCl_2$ and 5% YPER. The cell supernatant was applied to a GST-column (25 ml, Amersham) equilibrated with Rv1827 optimal buffer. The column was washed extensively with at least 500mL 1M NaCl, 50mM Tris.HCl pH 8.0 to remove bound contaminants and DNA, followed by buffer exchange to Rv1827 optimal buffer plus 1mM DTT. The GST was cleaved by addition of 200 units PreScission protease (Amersham) and overnight incubation at 4°C. Rv1827 was eluted with optimal buffer and the protein was concentrated (5-10mg/mL) using a Vivaspin centrifugal concentrator (5 kDa cutoff) and further purified on a Sephadex 75 gel filtration column using Rv1827 optimal buffer as running buffer. Rv1827 eluted in one peak from gel filtration at an apparent molecular weight of 20kDa. Protein from this peak was collected and concentrated to 10mg/mL, then finally flash frozen in liquid nitrogen for storage at -80°C.

3.2.2.7 Purification of Rv0019

Constructs of Rv0019 were purified in different buffers depending on their predicted isoelectric point ($Rv0019^{29-156}$ pI = 10.38 and $Rv0019^{57-156}$ pI = 8.35). Bacteria were broken in $Rv0019^{29-156}$ buffer (300mM NaCl and 50mM Tris.HCl pH 8.0) or $Rv0019^{59-156}$ buffer (300mM NaCl and 50mM Tris.HCl pH 7.0) with addition of one EDTA-free protease inhibitor tablet (Roche), 5U/mL Benzonase, 2mM $MgCl_2$ and 5% YPER. The cell supernatant was applied to a GST-column (25 ml, Amersham) equilibrated with optimal buffer. The column was washed extensively with at least 500mL 1M NaCl, 50mM Tris.HCl pH 8.0, to remove bound contaminants and DNA; this is especially important as the Rv0019 protein is basic and could be associated with an increase in

bound DNA. The buffer was exchanged for optimal buffer plus 1mM DTT. The GST was cleaved by addition of 200 units PreScission protease (Amersham) and overnight incubation at 4°C. Alternatively, a GST fusion protein was obtained by incubating protein-bound resin with 20mM reduced glutathione and overnight incubation at 4°C. Rv0019 was eluted with optimal buffer and the protein was concentrated (5-10mg/mL) using a Vivaspin centrifugal concentrator (5 kDa cutoff). For anion-exchange, the NaCl concentration of the buffer was adjusted to 50mM by dilution of the protein sample. To remove all remaining bound DNA the protein sample was applied to a Q-Sepharose column equilibrated with 50mM NaCl, 50mM Tris.HCl pH8.0 and flow-through containing protein stripped of DNA was collected. The protein was further purified on a Sephadex 75 gel filtration column using respective protein buffers as running buffers. Both Rv0019 protein fragments eluted with a single peak from gel filtration at an apparent molecular weight of 20kDa. GST-Rv0019$^{29\text{-}156}$ was applied to a S200 column and eluted in one peak from gel filtration at an apparent molecular weight of 50kDa. Protein peaks were collected and concentrated to 10mg/mL, then finally flash frozen in liquid nitrogen for storage at -80°C.

3.2.2.8 Purification of Rv0020

Constructs of Rv0020 were purified differently according to the construct properties. Bacteria were broken in Rv0020 buffer (300mM NaCl and 50mM Tris.HCl pH 8.0) with addition of one EDTA-free protease inhibitor tablet (Roche), 5U/mL Benzonase, 2mM $MgCl_2$ and 5% YPER. The cell supernatant was applied to a GST-column (25 ml, Amersham) equilibrated with optimal buffer. The column was washed extensively with at least 500mL 1M NaCl, 50mM Tris.HCl pH 8.0, to remove bound contaminants and DNA, followed by buffer exchange to the Rv0020 buffer plus 1mM DTT. The GST was cleaved by addition of 200 units PreScission protease (Amersham) and overnight incubation at 4°C. Rv0020$^{423\text{-}527}$ was eluted with optimal buffer and the protein was concentrated (5-10mg/mL) using an Vivaspin centrifugal concentrator (5 kDa cutoff) and further purified on a Sephadex 75 gel filtration column using the respective protein buffer as running buffer. Rv0020 eluted in one peak from gel filtration at an apparent molecular weight of 20kDa. Protein from this peak was collected, concentrated to 10mg/mL and then finally flash frozen in liquid nitrogen for storage at -80°C. Rv0020$^{155\text{-}527}$ was eluted off the GST-column with Rv0020 buffer and the protein was

concentrated to 1mg/mL using a Centricon centrifugal concentrator (10 kDa cutoff). It was previously found that concentration to higher than ~1mg/mL would result in protein aggregation. The protein sample was applied to a Sephadex 200 gel filtration column using respective buffer as running buffer. $Rv0020^{423-527}$ eluted in one peak from gel filtration at an apparent molecular weight of 20kDa, protein from this peak were collected and concentrated to ~1mg/mL, then finally flash frozen in liquid nitrogen for storage at -80°C.

3.2.2.9 Purification of PknB

All constructs of PknB and mutants were purified using the same procedure. Bacteria were broken in PknB buffer (300mM NaCl and 50mM Tris.HCl pH 8.0) with addition of one EDTA-free protease inhibitor tablet (Roche), 5U/mL Benzonase, 2mM $MgCl_2$ and 5% YPER. The cell supernatant was applied to a GST-column (25 ml, Amersham) equilibrated with optimal buffer. The column was washed extensively with at least 500mL 1M NaCl, 50mM Tris.HCl pH 8.0, to remove bound contaminants and DNA. The buffer was exchanged for PknB buffer plus 1mM DTT. The GST was cleaved by addition of 200 units precision protease (Amersham) and overnight incubation at 4°C. PknB was eluted with optimal buffer and the NaCl concentration of the buffer was adjusted to 50mM by dilution of the protein sample in preparation for ion-exchange. The protein sample was applied to a Q-sepharose column equilibrated with 50mM NaCl, Tris.HCl pH8.0 and bound proteins were eluted using a NaCl gradient of 0-300mM, 50mM Tris.HCl pH 8.0. Fractions containing PknB were identified using SDS-PAGE, pooled and the protein was concentrated (2-5mg/mL) using a Vivaspin centrifugal concentrator (10 kDa cutoff). The protein was further purified on a Sephadex 200 gel filtration column using optimal buffer as running buffer. PknB eluted in one peak from gel filtration at an apparent molecular weight of 45kDa. Protein from this peak was collected and concentrated to 5mg/mL, then finally flash frozen for storage at -80°C in liquid nitrogen.

3.2.2.10 Purification of PknA

$PknA^{1-290}$ was purified using the following procedure. Bacteria expressing $PknA^{1-290}$ were broken in PknA buffer (300mM NaCl, 50mM Tris.HCl pH 8.0) with addition of one EDTA-free protease inhibitor tablet (Roche), 5U/mL Benzonase, 2mM $MgCl_2$, 5mM imidazole and 5% YPER. The cell supernatant was applied to an NTA-column (25 ml, Amersham) equilibrated with optimal buffer plus 5mM imidazole. The column was washed extensively with at least 500mL 1M NaCl, 50mM Tris.HCl pH 8.0 and 30mM imidazole, to remove bound contaminants and DNA. Bound protein was eluted off the column using an imidazole gradient from 50-250mM (buffer also contained 50mM NaCl, 50mM Tris.HCl pH 8.0). Fractions containing PknA were identified by SDS-PAGE and pooled. The eluted protein sample was applied to a Q-sepharose column equilibrated with 50mM NaCl, Tris.HCl pH 8.0 and bound proteins were eluted using a NaCl gradient of 0-300mM, 50mM Tris.HCl pH 8.0. Fractions containing PknA were identified using SDS-PAGE, pooled and the protein was concentrated (2-5mg/mL) using a vivaspin centrifugal concentrator (10 kDa cut-off). The protein was further purified on a Sephadex 200 gel filtration column using optimal buffer as running buffer. PknA eluted in one peak from gel filtration at an apparent molecular weight of 45kDa. Protein from this peak was collected and concentrated to 10mg/mL, then finally flash frozen for storage at -80°C in liquid nitrogen.

3.2.3 Limited Proteolysis

3.2.3.1 Limited tryptic digestion

Purified fragment of Rv0019 consisting of residues 29-156 were subjected to limited proteolysis using the protease trypsin and Asp-N (Roche, sequencing grade). Reactions were carried out in a 100µL volume containing protein and trypsin in the digestion buffer (150mM NaCl, 50mM Tris.HCl pH 8.0). The trypsin:protein ratio for this digestion was 1:250 and 1:500. Products of the time course of digestion were 'quenched' by adding 10µL of SDS gel loading buffer and heating the sample to 100°C for five minutes. Samples were resolved using SDS-PAGE on a 12% gel.

3.2.3.2 Protein blotting

Proteins resolved on an SDS-PAGE gel after limited trypsin digestion experiments were transferred to a ProBlott membrane (Applied Biosystems). This was achieved cutting the ProBlott membrane to the size of the gel and wetting in analytical grade methanol. The membrane was then placed in the electroblotting buffer (see Sections 3.1.6). The gel was allowed to equilibrate in electroblotting buffer for five minutes before the blotting 'sandwich' was assembled: pad > filter paper > gel > membrane > filter paper > pad.

The transfer was carried out in an electroblotting tank at a constant voltage of 50V for 45 minutes with cooling using an ice block in the gel tank. After blotting was complete, the membrane was removed and rinsed in distilled water before being immersed in methanol. The membrane was stained using 0.1% Coomassie® Blue R-250 in 40% methanol and 1% acetic acid, and destained using 50% methanol. The bands of interest were then excised for protein sequencing.

3.2.3.3 Protein sequencing

Protein from the western blot was sent to Dr. Mike Bacon of the Babraham Institute, Cambridge, to determine the six N-terminal residues using the Edman degradation technique. The N-terminus of the amino-acid was modified by phenylisothiocyanate (PITC) which was subsequently removed by acid hydrolysis resulting in the release of 2-anilino-5-thiazolinone amino acid derivative. This derivative is then converted to a phenylthiohydantoin amino acid and HPLC analysis is used for amino acid identification. Six amino acid residues were identified in a cyclical sequencing reaction as described above.

3.2.4 Phosphorylation Mapping

2.1.2.2 In-gel digestion of proteins and sample preparation

Phosphorylation mapping was performed by first digesting 10μg of the target protein with 1:10 (w/w) sequencing grade modified trypsin (Promega) in 10mM ammonium bicarbonante. Phosphopeptides were isolated from peptide mixtures using an IMAC Spheres gel with chelated Ga^{3+} (Pierce). The tryptic digest was diluted 1:1 with 20% acetic acid and incubated on the gallium resin for 30 minutes at room temperature. The resin was washed with 10% acetic acid, followed by 10% acetic acid plus 20% acetonitrile. The resin was then equilibrated with MilliQ H_2O followed by elution in 100mM ammonium bicarbonate (pH 9.0) plus 10% acetonitrile.

3.2.4.1 Matrix assisted laser desorption ionization/Time of Flight (MALDI/TOF)

Phosphorylation mapping was performed using a Reflex III MALDI time-of-flight mass spectrometer (Bruker Daltonik, GmbH) by Dr. Steve Howell (NIMR), equipped with a nitrogen laser and a Scout-384 probe, to obtain positive ion mass spectra of digested protein with pulsed ion extraction in reflectron mode. An accelerating voltage of 26kV was used with detector bias gating set to 2kV and a mass cut-off of m/z 650. Thin layer matrix surfaces were prepared using recrystallised α-cyano-4-hydroxycinammic acid and nitrocellulose using the fast evaporation method (Vorm *et al*, 1994). 0.4 μL of acidified digestion supernatant was deposited on the matrix surface and allowed to dry prior to desalting with water.

Specific phosphorylation sites shown to be ambiguous by MALDI-TOF were determined by loading the remaining sample onto a 2mm x 0.8mm C18 microcolumn (LC Packings), washed and step eluted with 60% methanol, 0.1% formic acid directly into an Econo12 nanospray needle (New Objective Inc.). Nanospray mass spectra were

acquired on an LCQ 'classic' quadrupole ion trap mass spectrometer (ThermoQuest Corp.) equipped with a nanospray source (Promega) operated at a spray voltage of 800V and a capillary temperature of 150°C. Tandem (MS^2) mass spectra were acquired at collision energy of 30% and a parent ion isolation width of 3Da. The resulting spectra were analysed for phosphate containing fragments using the MS-Product program in the Protein Prospector Suite (http://prospector.ucsf.edu/). In addition, phospho-peptide mass fingerprints thus obtained were searched against the non-redundant protein database placed in the public domain by the National Centre for Biotechnology Information (NCBI) using the program MASCOT (Perkins et al 1999).

3.2.5 Circular Dichroism

Far-UV spectra were recorded on a Jasco J-715 spectropolarimeter (Jasco) using 1mm fused silica cuvettes. All experiments were carried out by Dr. Steve Martin (NIMR). Protein samples were at ~0.15mg/mL in 50mM Tris.HCl pH 8.0, 150mM NaCl buffer. Multiple scans were averaged and the appropriate buffer blanks were subtracted. In the case of thermal denaturation, the temperature dependence of enthalpy (ΔH) is described by the following equation:

$$\Delta H = \Delta H_0 + \Delta C_p (T - T_0) \qquad \textbf{(26)}$$

where T is the temperature (°K) and ΔH_0 is the values of ΔH at any reference temperature T_0, and ΔC_p is the change in heat capacity at constant pressure (Schellman, 1987). In this model ΔC_p is assumed to be independent of temperature (Privalov and Khechinashvilli, 1947). The experimental data were fitted to equation **(26)** using a standard non-linear least squares fitted procedure written by Dr. S. Martin.

3.2.6 Synthesis of peptides

Peptides were synthesised by solid phase technique with an addition tyrosine residue and two C-terminal lysines to enhance solubility and facilitate concentration determination by optical absorbance. Peptides were purified by reverse phase HPLC following deprotection, and sequence confirmation by MALDI-TOF mass spectroscopy. All peptide synthesis was carried out by Dr. Pete Fletcher (NIMR) or Dr. Will Mawby (University of Bristol).

3.2.7 Characterisation of binding

3.2.7.1 Isothermal titration calorimetry

The enthalpy of binding was determined by titrating 1.3mL of protein solution, at an appropriate concentration, with a solution of either FHA or peptide using a VP-ITC MicroCalorimeter instrument from MicroCal. Both proteins were dialysed overnight in the same buffer containing 50mM Tris.HCl buffer, pH 8.0 and 300mM NaCl and degassed prior to use. The sample cell was maintained at 18°C. The enthalpy of binding was determined from heats of multiple injections of 10μL each. An equilibration time of 300 seconds was set between injections. The heat of dilution of each injection was established as a control by measuring the injection of ligand into buffer. The mean of the enthalpies measured was subtracted from the raw titration data prior to curve fitting.

Data analysis was carried out using the Origin data-analysis software supplied with the VP-ITC. This software uses principles based on Wiseman isotherm theory as discussed in Chapter 2 (Section 2.2).

3.2.7.2 Surface plasmon resonance

All surface plasmon resonance (SPR) experiments were conducted using a Biacore 2000 (Pharmacia) machine. A research grade CM5 sensorchip which contained a dextran sensor surface with four flow cells for protein immobilization. Proteins were immobilised to the dextran using the amine coupling method outlined in Chapter 2 (Section 2.2). For immobilisation, flow cells were activated with a mixture of EDC (0.2 M) and NHS (0.05 M). The PknA and PknB proteins were passed over the surface in a 10mM sodium acetate buffer, pH 4.5-6, at 40-120μg/mL concentrations. The low ionic strength and buffer pH value below the isoelectric point of the protein led to the concentration of protein at the matrix surface by electrostatic attraction allowing spontaneous reaction of lysine residues with the active esters. The remaining active

esters of the chip surface are converted into amides via reaction with ethanolamine. This sequence was performed automatically using a program scripted for the BIAcore 2000 control software and typically lasted 45 minutes. An immobilization level between 500 and 5000 response units and typically achieved by varying parameters such as protein concentration, protein ionic strength, buffer pH and reaction time. Protein was immobilized to the first 3 flow cells and the 4th flow cell was left as a reference 'mock' immobilization surface.

The reference surface makes it possible to subtract the response changes related to non-specific binding, changes in the matrix conformation or instrument derived errors. Kinetic analysis was performed by injecting FHA domain protein in HBS-EP buffer over immobilised protein kinase at a flow-rate of 50μL/min at 24°C. During equilibrium analysis the sensor-gram (*RU* vs time) was recorded at a flow-rate of 5μL/min at 24°C.

Data analysis was performed using a global fitting method on the BIAevaluation 3.0 program. Further details on the models used to fit data are found in Chapter 2. Further analysis of primary data was performed for accurate determination of equilibrium binding constants. After subtracting the sensorgram generated on from the protein surface from the curve generated from the control or 'reference' surface, steady-state response values (RU) were obtained at each FHA concentration. These values were fitted, using non-linear regressional analysis, to a single site saturation curve, where the amount of ligand bound is plotted as a function of the initial FHA concentration.

$$RU = \frac{Capacity.[FHA]}{K_d + [FHA]} \qquad \textbf{(28)}$$

Competition experiments were conducted using the same immobilised proteins and a solution with fixed concentration of FHA domain were passed over them spiked with varying concentrations of synthetic peptide. The steady-state response values (RU) were obtained at each peptide concentration after subtracting the control value. This response value would equate to the maximal binding of the FHA the ligand at a specific concentration of competing peptide. Readings were taken for at least 12 different peptide concentrations and IC_{50} values were calculated by fitting, using non-linear regressional analysis, the response *versus* peptide inhibitor concentration to dose response curve:

$$RU = \frac{Range}{1+\left(\frac{[peptide]}{IC_{50}}\right)^{s}} + \text{Background} \qquad (29)$$

where range is the maximum y range, and s is the slope factor.

Assuming that the interactions investigated are simple 1:1 reversible interactions where the two ligand species are competing for binding to the same site, the IC_{50} calculated for the titration of each peptide in SPR competition experiments can be related to the equilibrium dissociation constant of the complex between the peptide and FHA domain using the following formula:

$$K_D^I = \frac{IC_{50}}{\left(1+\frac{[S]}{K_D}\right)} \qquad (30)$$

where K_D^I is the peptide equilibrium dissociation constant as determined by SPR competition assay, IC_{50} is the concentration of the peptide that inhibits binding of the FHA domain to the surface immobilised kinase by 50%, K_D is the equilibrium dissociation constant for the complex between the FHA and surface immobilised kinase and [S] is the concentration of FHA domain.

3.2.8 Characterisation of phosphorylation activity

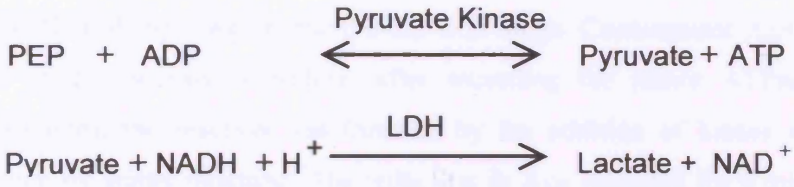
3.2.8.1 *In vitro* radiolabelling kinase assay

The ability of PknA and PknB to autophosphorylate and phosphorylate exogenous substrates was determined by radiolabelling experiments. 2μM PknA or PknB was incubated with 20mM of FHA containing protein or myelin basic protein (Invitrogen). Each assay was conducted in 10μL of buffer containing 150mM NaCl, 50mM Tris.HCl pH 8.0, 5mM $MgCl_2$ and 1mM $MnCl_2$ plus 0.2μCi of [γ-^{32}P]ATP (Amersham). The reaction proceeded for 60 minutes at room temperature and stopped with the addition of SDS-PAGE sample buffer plus EDTA (25mM final concentration). 10μL were subjected to SDS-PAGE on a 4-12% polyacrylamide gel. Gels were fixed and stained in 0.1% Coomassie® Blue R-250, 10% methanol, 10% glacial acetic acid. Gels were destained in 40% methanol, 10% glacial acetic acid and 5% glycerol, then wrapped in Saran wrap for gel drying. Dried gels were exposed to phosphoimage screens (Kodak) overnight. The radiolabelled protein was visualised using the Molecular Dynamics Storm 860 phosphorimaging scanner and quantification analysis was performed using Image Quant software (Molecular Dynamics).

3.2.8.2 *In vitro* coupled kinase assay

Kinetic data for enzymatic activity was determined using the classical coupled pyruvate kinase: lactate dehydrogenase assay followed spectrophotometrically at 340nm (A_{340}) according to the method of (Roskoski, 1983) (Figure 3.2). ATP, phosphoenolpyruvate, NADH, pyruvate kinase and lactate dehydrogenase were purchased from Sigma. Standard reactions were carried out in a buffer containing: 150mM NaCl, 50mM Tris.HCl pH 8.0, 5mM $MgCl_2$, 1mM $MnCl_2$, 20 units/mL pyruvate kinase, 30 units/mL lactate dehydrogenase, 1mM phosphoenolpyruvate and 0.25mM NADH. Optimal concentrations of lactate

A.



B.

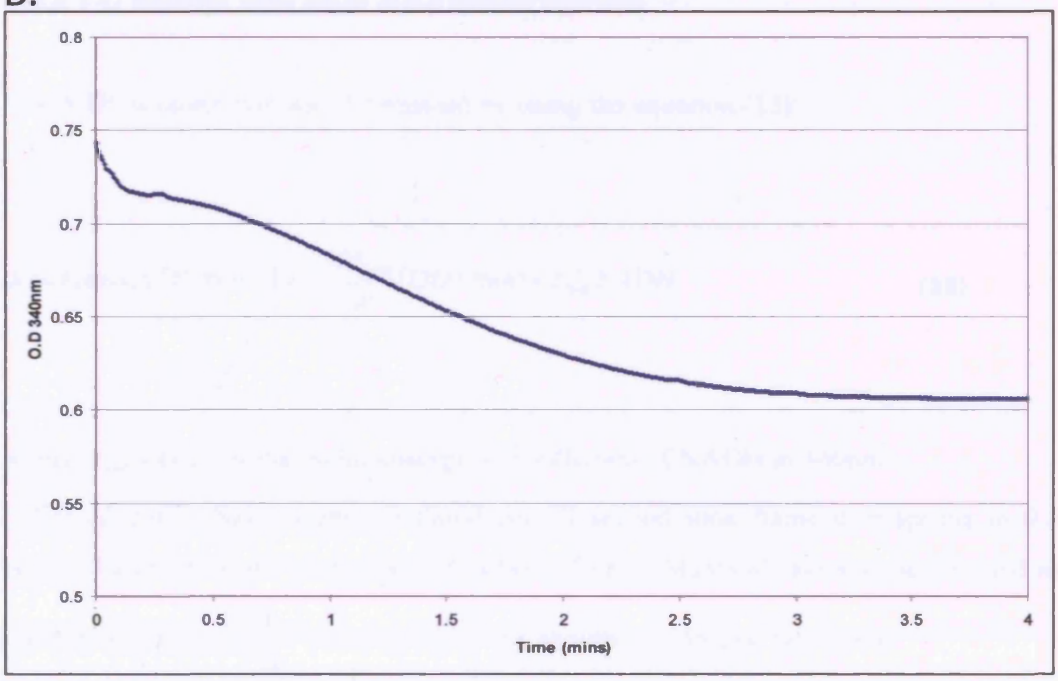


Figure 3.2. *In vitro* coupled kinase activity assay. (A) Enzymatic steps for linking the turnover of ATP to the oxidation of reduced NAD (NADH) by Lactate dehydrogensase (LDH). (B) A typical graph generated by the phosphorylation of $Rv1827^{1-162}$ by PknB. The turnover of ATP is linked to the oxidation of NADH and is followed spectrophotometrically at 340nm. Contaminant ADP is converted into ATP instantaneously; therefore, after recording the reduction in A_{340}, the reaction was initiated by the addition of kinase. The reduction in A_{340} recorded for ~4 minutes.

dehydrogenase and pyruvate kinase were determined in accordance with (Roskoski, 1983). All reactions were carried out in 100μL volume containing 2μM PknA or PknB using $Rv1827^{1-162}$ substrate at concentrations in the range of 1.3μM-500μM and an ATP concentration of 1mM. ATP dose response experiments were also carried out in a 100μL volume containing 2μM PknA or PknB using $Rv1827^{1-162}$ substrate at 40μM and ATP at concentrations in the range of 0.63μM-5mM. As a control the above controls were carried out in the absence of substrate, this would provide a measurement for the auto-phosphorylation activity of PknA and PknB.

Initial NADH input was in the 0.6-1.3 O.D range. Contaminant ADP is converted into ATP instantaneously; therefore, after recording the minor ATPase activity of the components, the reaction was initiated by the addition of kinase and mixing of the solution by gentle pipetting. The reduction in A_{340} recorded for 4 minutes. All enzyme assays were conducted at 25°C, and all experiments were done in triplicate. Absorbance data was collected on a Cary 50 UV-vis spectrophotometer using Enzyme Kinetics mode and readings were made at 0.5 second intervals.

The ATP turnover rate was determined by using the equation **(23)**:

$$Rate[\mu molATP.\min^{-1}] = -\frac{dA_{340}}{dt}[OD/\min] \times \varepsilon_{340}^{-1} NADH \qquad \textbf{(28)}$$

where $\varepsilon_{340}^{-1} NADH$ is the molar absorption coefficient of NADH at 340nm (6220 $M^{-1}cm^{-1}$). Rates were calculated per 10 second time frame overlapping in 0.5 second increments and plotted as a function of time. Maximal rate was determined as the minimal slope ($\frac{dA_{340}}{dt}$) during which the absorbance decline was linear.

Data was fitted to equation **(24)** for normal Michelis-Menten kinetics:

$$v = \frac{V_{max}[S]}{Km + [S]} \qquad \textbf{(29)}$$

where [S] is the substrate, V_{max} is the maximal enzyme velocity and K_m is the Michelis constant. The data was analysed by non-linear regression using GraFit 5.0 Enzyme Kinetics module.

3.2.9 Inhibitor Screening

Inhibitor studies were performed using the Kinase-Glo™ Luminescent Kinase Assay (Promega) according to manufacturers guidelines. All experimented was carried out at Medical Research Council Technology in collaboration with D. Whalley and B. Saxty. Initial optimization of reaction conditions was performed. To determine the optimal concentration of ATP, different concentrations of ATP (0.4-15μM) were made with 1μM PknB and 1mM $MnCl_2$, 150mM NaCl, 50mM Tris.HCl-HCl pH 8.0 in a total reaction volume of 20μL. As a control, a repeat titration was performed in the absence of kinase. The incubation was performed at room temperature for a total of 10 minutes. After 10 minutes the reactions was stopped by the addition of 20μL of Kinase-Glo reagent. Luminescence was read after 10 minutes using a BMG Polarsar reader (luminescence mode, 1 second integration time) and the change in relative luminescence signal (RLU) compared to the zero enzyme control. RLU was plotted against Log [ATP] and the optimal kinase substrate concentration was around the EC_{50}.

To determine the optimal concentration of substrate, $Rv1827^{1-162}$, several concentrations of Rv1827 were made with 1μM kinase and 1.5μM ATP, 1mM $MnCl_2$, 150mM NaCl, 50mM Tris.HCl-HCl pH 8.0 in a total reaction volume of 100μL. As a control, a repeat titration was performed in the absence of kinase. The incubation was performed at room temperature, and at the time points indicated 20μL aliquots were removed and mixed with an equal volume of Kinase-Glo reagent. Luminescence was read after 10 minute incubation using a BMG Polarstar reader (luminescence mode, 1 second integration time) and the change in relative luminescence signal (RLU) compared to the no enzyme control. RLU was plotted against time for each different Rv1827 concentration. The optimal substrate concentration would give a linear decrease in RLU over the time period.

Inhibitor library screening was conducted using the Library of Pharmacologically Active Compounds (LOPAC) (1,240 compounds, Sigma-Aldrich). All stock compounds are purchased dissolved in DMSO. Each compound was tested at a final assay concentration of 10μM in a total reaction volume of 22μL with a final DMSO concentration of 1% and a final ATP concentration of 1.5μM. Aliquots (15μL) of a 1.46x master mix containing 1.46x the final concentration of PknB (0.15μM final concentration), Rv1827 substrate (5μM), NaCl (150mM), Tris.HCl pH 8.0 (50mM), $MnCl_2$ (1mM), Triton X-100 (1%) and DTT (1mM) were added to Matrix 384-well polypropylene plates using the Biomek FX Laboratory automated workstation (Beckman-Coulter). Stock solutions of the inhibitor compounds were made at 108μM and 2μL was added to the master mix using the automated pipettor. The mastermix was then incubated for 30 minutes. The kinase reaction was initiated by the addition of ATP to the reaction buffer using the Wellmate automated pipettor (Matrix). After a reaction time of 20 minutes 22μL of Kinase-Glo reagent was added to each well to stop the reaction and the plate was incubated for a further 10 minutes prior to reading. The plate was read using the BMG Polarstar reader ('Hithunter 1' luminescence mode). As a positive control, compounds were replaced with DMSO in all wells of the last column of the plate. As a negative control no enzyme was added to the wells of the first column of the plate and compound solutions were replaced with DMSO only.

To assess the compounds as inhibitors the percentage inhibition of each compound was calculated as follows:

$$RLU = \frac{1 - [LU\ (\text{data well}) - LU(\text{negative control})]}{LU\ (\text{positive control}) - LU\ (\text{negative control})} \qquad \textbf{(30)}$$

where negative and positive control reactions were carried out in the absence and presence of ATP respectively, and data well reactions were carried out in the presence

of test compounds and ATP. Hits were defined as any compound producing >20% inhibition of PknB activity.

Quality of the assay was evaluated with the *Z'* factor. This is a statistical value which is the ratio of the separation band to the dynamic range of the assay, based on the positive and negative control data and is defined as:

$$Z' = 1 - \frac{(3\sigma_{c+} + 3\sigma_{c-})}{|\mu_{c+} - \mu_{c-}|} \qquad (31)$$

where σ_{c+} and σ_{c-} are the standard deviation of the positive and negative control signals, respectively, and μ_{c+} and μ_{c-} are the averages of the positive and negative control signals, respectively. *Z'* factor values above 0.5 indicate good assay quality and a value of 1 is a perfect assay.

The screening window coefficient (Z-factor) is similar to the previous equation:

$$Z = 1 - \frac{(3\sigma_S + 3\sigma_C)}{|\mu_S - \mu_C|} \qquad (32)$$

where μ_s and μ_c are the means of the population and control (usually background) signals, respectively, and the signal standard deviations are σ_s and σ_c, respectively. As before, *Z* factor values above 0.5 indicate good assay quality and a value of 1 is a perfect assay. This assay informs the user about data variation associated with the sample measurements and reference controls as well as informing about assay signal dynamic range (Zhang *et al.*, 2000).

As a validation and for additional characterisation of 'Hit' compounds IC_{50} values were determined. To each well of a 96-well plate 25µL of a reaction mixture containing 2x the reaction buffer (PknB (0.2µM), Rv1827 substrate (5µM), NaCl (150mM), Tris.HCl-HCl pH 8.0 (50mM), $MnCl_2$ (1mM) and DTT (1mM)) was added except for column 1. In the first column 50µL of inhibitor diluted in the reaction mixture was added. A two-fold serial dilution was made across the plate by transfer pipetting. 25µL of 2x the

concentration of ATP (1.5μM) was added to all the wells to initiate the reaction. The reaction was allowed to continue for 20 minutes at room temperature and readings were taken as previously described. The IC_{50} values were obtained by fitting the data using non-linear regressional analysis to the following equation:

$$RLU = \frac{RLU_{max} IC_{50}}{IC_{50} + [I]} \quad (33)$$

4 Protein purification and characterisation

4.1 Overview

A total of 18 constructs were designed and Table 4.1 lists the properties for each purified protein. Protein expression was carried out in *E. coli* because of the ease in which plasmids can be transferred into the bacterium, the ease of culture, fast replication rate and immediate induction of the expression system.

10 recombinant proteins were made representing three FHA domain-containing proteins Rv0019, Rv0020 and Rv1827 in three forms: 1) as full-length wild type protein (in the case of membrane bound protein Rv0019, a protein representing the full intracellular region was produced) 2) proteins representing the isolated FHA domain were generated (using limited proteolysis to define domain boundaries) 3) mutants proteins where a conserved serine in the FHA domain essential to the phospho-binding function is mutated to an alanine. All proteins were characterised by electrospray mass spectroscopy after purification to confirm expected molecular weight.

Eight recombinant fragments of two STPKs PknA and PknB were also expressed and purified. The full intracellular region of PknB was produced in addition to the isolated kinase domains of PknA and PknB. PknA and PknB were also generated with mutation of conserved kinase domain activation loop threonine residues to alanines and separately with mutation of a conserved catalytic lysine to an alanine. All proteins were characterised by electrospray mass spectroscopy which indicates differing degrees of phosphorylation of kinase constructs displayed as an increase in observed molecular weight due to added phosphate groups. PknA and PknB K40A were also subjected to matrix-assisted laser desorption/ionisation (MALDI)/Time-of-flight (TOF) mass spectrometry in an effort to identify the sites of phosphorylation.

Table 4.1. Recombinant proteins expressed for this study

A. FHA Domain-Containing Proteins

Protein fragment	Residues	Expected Molecular Weight (Da)[a]	Electrospray Molecular Weight (Da)	Purification method[d]
Rv1827$^{1-162}$	162	17630.3	17632.1±1.2	GST, GF75
Rv1827$^{1-162}$ S95A	162	17614.3	17612.3±2.4	GST, GF75
Rv1827$^{25-162}$	137	14691.1	14689.8±1.3	GST, GF75
Rv1827$^{25-162}$ S95A	137	14675.1	14674.8±1.5	GST, GF75
Rv0019$^{28-156}$	156	15282.4	15284.3±1	GST, GF75
Rv0019$^{56-156}$	127	11391.9	11392.8±0.5	GST, GF75
Rv0020$^{155-527}$	527	40615	40609±4.1	GST, GF75
Rv0020$^{155-527}$ S473A	527	40599	40600±2.0	GST, GF75
Rv0020$^{432-527}$	95	12093.3	12105.5±1.5	GST, GF200
Rv0020$^{432-527}$ S473A	95	12077.3	12088.4±2.2	GST, GF200

B. Serine/Threonine Protein Kinases

Protein fragment	Residues	Expected Molecular Weight (Da)[a]	Electrospray Molecular Weight (Da)	Maximum phosphorylation sites[b]	Most abundant phosphorylation state[c]	Purification method[d]
PknA$^{1-290}$	290	34855.9	36374.8±4	17	12	NTA, Q, GF200
PknA$^{1-290}$ K41A	290	34798.2	35038.2	6	4	NTA, Q, GF200
PknB$^{1-279}$	279	31680	31992±2.5	5	3	GST, Q, GF200
PknB$^{1-279}$ T171A	279	31650	31960±4	4	3	GST, Q, GF200
PknB$^{1-279}$ T173A	279	31650	31879±2.7	3	2	GST, Q, GF200
PknB$^{1-279}$ T171A/T173A	279	31620.18	31850.1±3.2	3	2	GST, Q, GF200
PknB$^{1-279}$ K40A	279	31622.3	31777±7.0	2	1	GST, Q, GF200
PknB$^{1-332}$	332	37377.9	37606.3±13.3	5	3	GST, Q, GF200

[a]As calculated from constituent amino-acids using ProtParam tool (www.us.expasy.org)

[b]As identified by the number of different mass/charge peaks differing by 80 Da (molecular weight of a HPO_3). See Figure 4.1 for a typical electrospray mass spectrum.

[c]Determined from the peak with the greatest intensity signal corresponding to the protein phosphorylation state.

[d]Purification methods include GST (glutatione S-transferase affinity purification), NTA (His-tag affinity purification), GF75 (gel filtration on a S75 Superdex column), GF200 (gel filtration on a S200 Superdex column), S (SP sepharose ion-exchange), Q (Q Sepharose anion-exchange).

Proteins were purified using standard methods (Section 3.2.2). The purification of Rv0019 and PknA proved most difficult and the purification of these proteins will be discussed in more detail in this chapter. In addition, the phosphorylation state of PknB is of great relevance to the biophysical and biochemical experiments in this study and will be discussed at the end of this chapter.

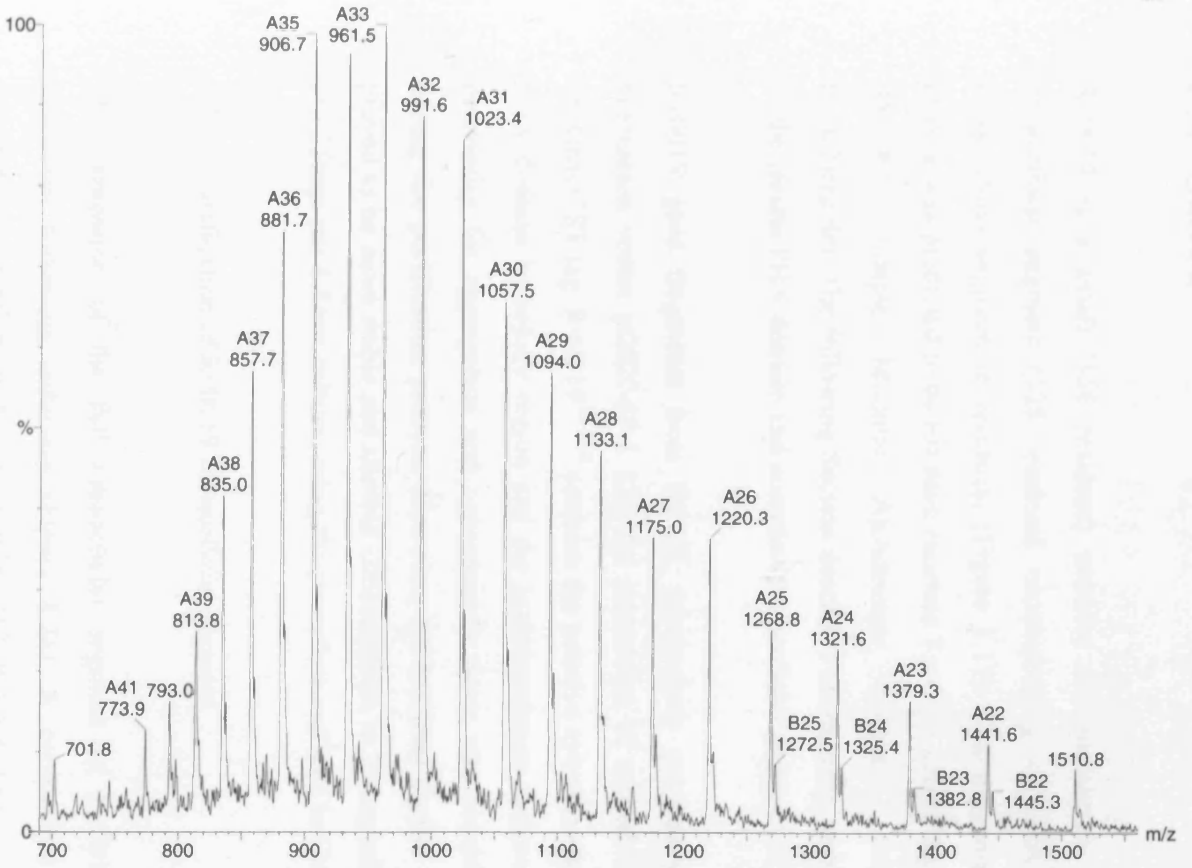


Figure 4.1. Mass spectra of $PknB^{1-279}$ K40A mutant. Mono- and Di-phosphorylated species of $pknB^{1-279}$ K40A mutant protein , A and B, were determined to be, 31697Da and 31777Da, respectively. These mass are equivalent to the calculated mass of $pknB^{1-279}$ K40A (31623Da) plus one phosphate group (A) and two phosphate groups (B).The molecular weight of HPO_3 is 80Da.

4.2 Purification and Characterisation of Rv0019

4.2.1 Overview

Rv0019 is a small (156 residues) putative transmembrane protein containing an intracellular segment (128 residues) encompassing an FHA domain and a short extracellular segment (6 residues) (Figure 1.11). The transmembrane portion of the protein was predicted to be between residues 7 and 28 according to the on-line database SMART (Simple Modular Architecture Research Tool, www.smart.embl-heidelberg.de). The following Section describes the cloning, expression and purification of the isolate FHA domain and complete intracellular segment of Rv0019.

Rv0019 gene fragments from the *M. tuberculosis* genome were inserted into the expression vector pGEX-6P-1 LIC for expression of recombinant protein with a N-terminal GST tag. $Rv0019^{28-156}$ contains the putative cytoplasmic segment including the FHA domain homology region and the juxtamembrane sequence. This protein had a propensity for aggregation and consequently there was considerable loss of protein during the purification process. However, the minimal tryptic fragment $Rv0019^{56-156}$ proved to be more stable and allowed concentration to 10 mg/mL, with an overall yield of ~25mgs per 6 litres culture, using the procedure outlined in Chapter 3.

4.2.2 Purification of Rv0019 intracellular segment

The expression of the full intracellular segment of Rv0019, $Rv0019^{28-156}$, was monitored following induction (Figure 4.2a). A comparison of lanes 1-3 shows detectable expression of fusion protein GST-Rv0019 had occurred two hours after induction. Maximum yield was obtained at three hours post induction. An approximately 90% pure recombinant protein solution was produced after affinity purification and by the removal of the GST tag by means of cleavage of the fusion protein bound glutathione-Sepharose matrix using 3C protease. This resulted in a band of approximately17kDa (Figure 4.2b), which corresponded to the expected size of the recombinant protein.

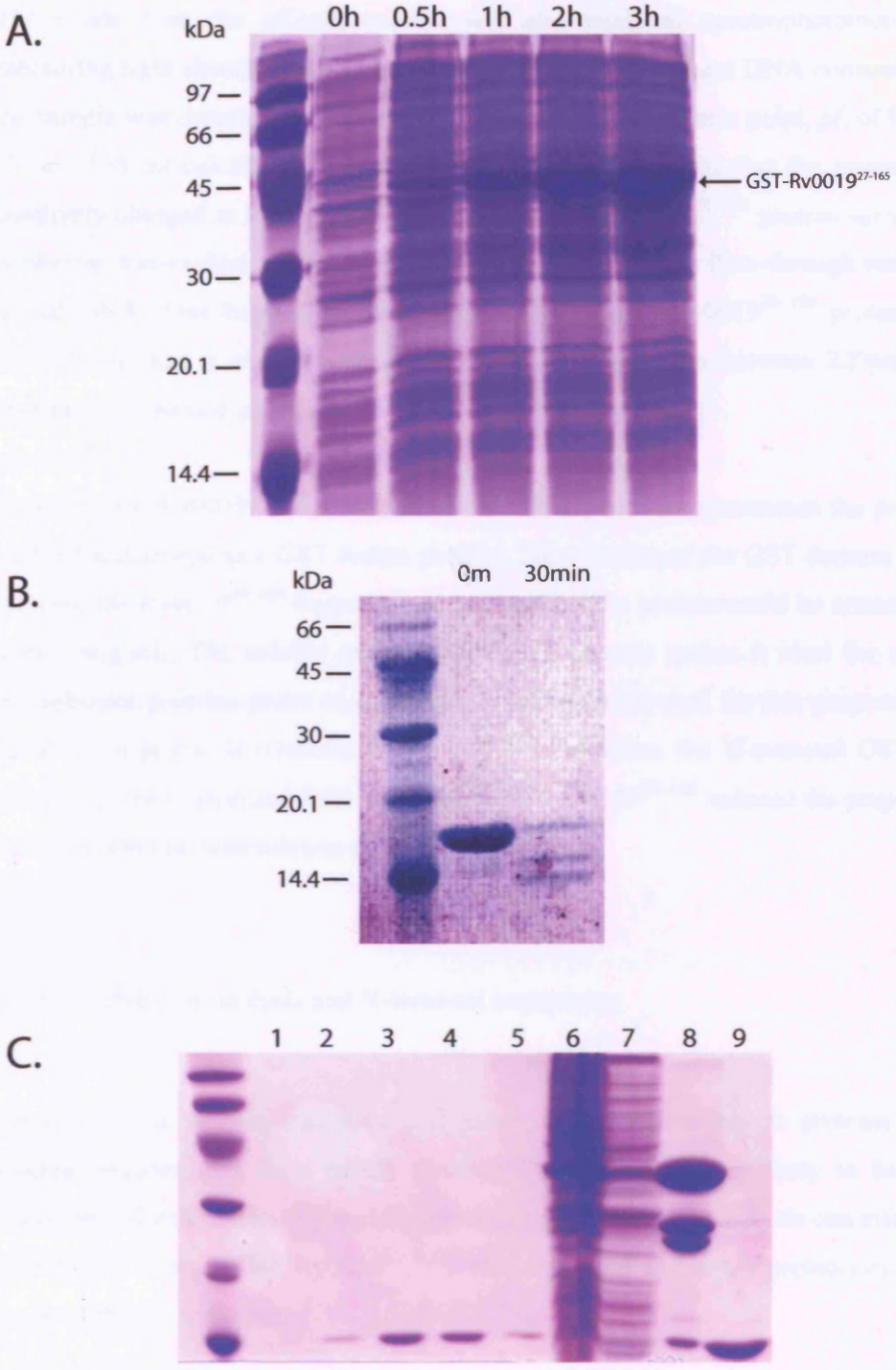


Figure 4.2. SDS-PAGE analysis of FHA-domain containing protein Rv0019. (A) The expression of GST-Rv0019$^{28-156}$ was monitored after induction. Lysates of cells taken 0, 30 min, 1 hr, 2 hr, 3 hr after induction was analysed by SDS-PAGE. Molecular weight marker sizes are shown on the left. These marker are used in all subsequent gels. (B) Tryptic digestion of Rv0019$^{28-156}$ with an enzyme: protein ratio of 1:250 at time points of 0, 30 secs, 1,5,10,15,30 minutes. Purified Rv0019$^{28-156}$ protein produces three bands at molecular weights around 17-14kDa after 30 seconds of digestion with trypsin. The band running slowest corresponds to the undigested Rv0019$_{28-156}$, the other two bands represented the digest products. (c) Purification of the minimal tryptic fragment of Rv0019 (Rv0019$^{59-156}$). Lanes contain samples taken of the protein containing fractions after gel filtration on a S75 column (lanes 1-5), cell lysate before loading on affinity resin (lane 6), and after loading on affinity resin (lane 7), sample taken after 3C protease cleavage of GST-Rv0019$^{59-156}$ bound to affinity resin (lane 8) and 3C cleaved Rv0019$^{59-156}$ prior to gel filtration (lane 9).

The eluate from the affinity column was also analysed spectrophotometrically by measuring light absorbance between 220nm–340nm. Significant DNA contamination of the sample was detected as absorbance at 260nm. The isoelectric point, *pI*, of $Rv0019^{28-156}$ is 10.38, as calculated from the sequence, which suggests that the protein will be positively charged in buffer at pH 8.0. Applying the $Rv0019^{28-156}$ protein sample to a Q Sepharose ion-exchange column at pH 8.0 and collecting the flow-through removed the bound DNA. This bound the DNA whilst allowing the $Rv0019^{28-156}$ protein to pass through and had a significantly improved protein spectrum between 220nm–340nm. However, it retained a propensity to aggregate.

To utilise the $Rv0019^{28-156}$ protein for further biochemical experiments the protein was purified and stored as a GST fusion protein. The addition of the GST domain helped to stabilise the $Rv0019^{28-156}$ fragment and in this form the protein could be concentrated to over 10mg/mL. The soluble nature of the GST domain makes it ideal for stabilising recombinant proteins prone to aggregation and has been used for this purpose in many purification protocols (Grunder *et al.*, 2005). In addition, the N-terminal GST domain along with the C-terminal FHA domain of GST-$Rv0019^{28-156}$ reduced the propensity for the disordered juxtamembrane region to aggregate.

4.2.3 Limited proteolysis and N-terminal sequencing

Regions of the protein that have less structure are susceptible to protease cleavage whereas regions that form tightly structured domains are less likely to be cleaved. Therefore, identification of the cleavage sites using limited proteolysis can infer domain boundaries. The purified $Rv0019^{28-156}$ was subjected to limited proteolysis to define limits of the FHA domain.

A tryptic digestion with an enzyme: protein ratio of 1:250 (w:w) produced three bands at molecular weights between 14-17kDa after 30 minutes (Figure 4.2b). The band with the slowest migration corresponded to the undigested $Rv0019^{28-156}$, whilst the other two bands represented digested products. These bands were transferred to blotting membrane and the smallest of the three bands was excised and subjected to N-terminal

amino acid sequencing. Interestingly, the minimal tryptic fragment of $Rv0019^{28-156}$ transferred to the blotting membrane with greater efficiency than the uncleaved protein suggesting that the fragment has a lower isoelectric point. For the six N-terminal residue positions the observed amino acids were:

1. Q/S **2.** H/P/F/L **3.** A/V **4.** D/F/H **5.** Y **6.** L/V

From this information, the N-terminus was determined to be in the region between residues 54-57 (See Figure 1.3 for Rv0019 amino acid sequence). The ambiguity of the sequence information is due to the presence of a number of tryptic sites in the region between residues 54-57 (three) resulting in a number fragments differing in size by one or two residues. Cloning, expression and purification of a Rv0019 fragment extending from residues 56-156 produced a soluble protein which was mono-dispersed as determined by dynamic light scattering (data not shown). Purified $Rv0019^{56-156}$ protein could be concentrated to 10mg/mL with an overall yield of ~25mgs produce per 6 litre bacterial culture (Figure 4.2c) using the procedure outlined in Chapter 3 (Section 3.2.2).

4.2.4 Circular dichroism analysis of FHA domains

To examine the secondary structure content of FHA domains and to assess secondary structure content, circular dichroism (CD) analysis was performed as described in Materials and Methods (Section 3.2.5). The resulting spectra of FHA domains of known and unknown structure were compared (Figure 4.3) in order to understand why the $Rv0019^{56-156}$ construct, containing the probable isolated FHA domain, was unable to bind phosphor-threonine motifs, as determined by biophysical analysis conducted in Chapter 6.

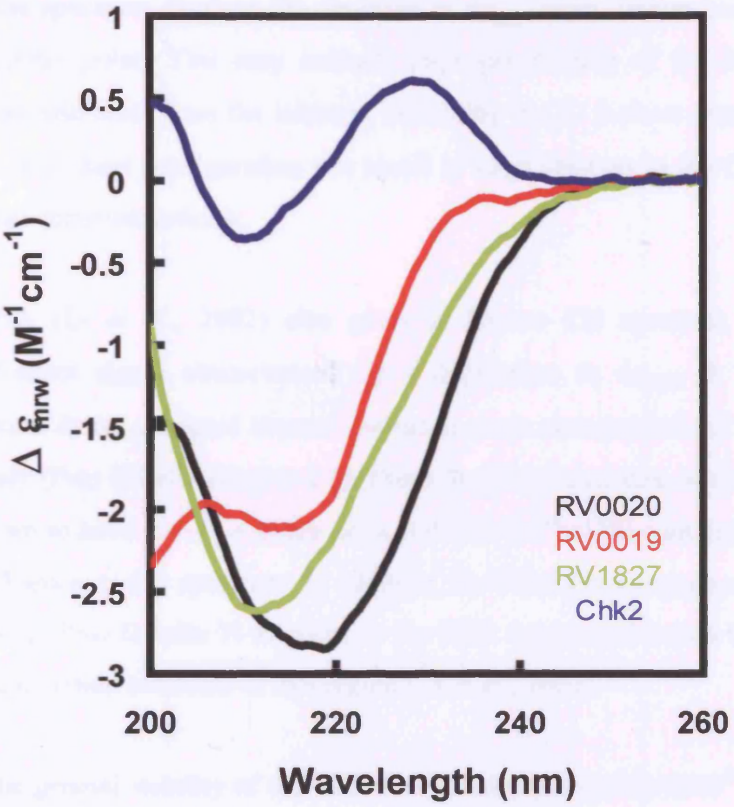


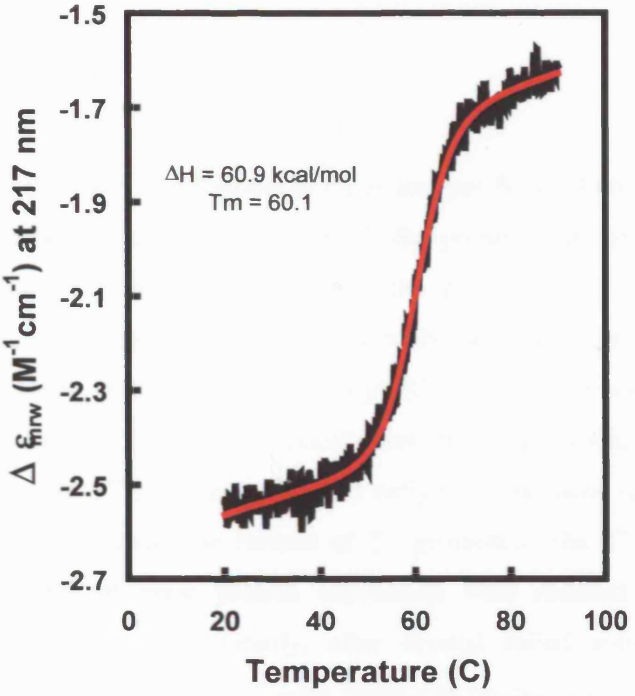
Figure 4.3. Circular dichroism spectra of the FHA domains of Rv1827, Rv0019, Rv1827 of *M. tuberculosis* and human Chk2 FHA domain. Protein samples were at ~0.15mg/mL in 50mM Tris-HCl pH 8.0, 150mM NaCl buffer. Multiple scans were averaged and the appropriate buffer blanks were subtracted.

Comparison of CD spectra immediately shows a distinct variation in spectrum despite being related in structure. The spectra representing Rv1827 and Rv0020 FHA domains seem typical of proteins consisting of mainly β-sheets, with a in mean residue molar extinction coefficient, $\Delta\varepsilon_{mrw}$, with the lowest point at around 220nm. In the case of $Rv0019^{56-156}$ the spectrum displays the decrease in $\Delta\varepsilon_{mrw}$ from 240nm but no increase beyond the 220nm point. This may indicate improper folding of the Rv0019 FHA domain but may also arise from the inherent variability in CD β-sheet measurement, as small changes in β-sheet configuration can result in large changes in the CD signal (S. Martin, personal communication).

The Chk2 FHA (Li *et al.*, 2002) also gives a distinct CD spectrum with a less pronounced β-sheet signal characterised by a depression in $\Delta\varepsilon_{mrw}$ at 210nm. The observed increase in $\Delta\varepsilon_{mrw}$ signal beyond 240nm is more characteristic of an extended 3_{10}-helix or poly (Pro) II helix (Figure 2.5). From the X-ray structure of Chk2 FHA the domain is known to have an ~11-residue helix (Glu127 to Thr138) containing two turns which may influence its CD spectrum. In addition, the Chk2 FHA sequence is predicted to contain a poly (Pro) II helix N-terminal to the FHA domain, although this appeared disordered in the crystal structure of this region (Li *et al.*, 2002)

To ascertain the general stability of the Rv0020 FHA domain and $Rv0019^{56-156}$, thermal unfolding was monitored using far-UV CD. Typical curves are shown in Figure 4.4. The measured T_m and ΔH values are 57.6°C and 59.4kcal/mol for the Rv0020 FHA domains and 60.1°C and 60.9kcal/mol for $Rv0019^{56-156}$. These data suggest that the minimal tryptic fragment of Rv0019, $Rv0019^{56-156}$, is slightly more stable than the Rv0020 FHA domain.

A. Rv0019$^{56-156}$



B. Rv0020$^{432-527}$

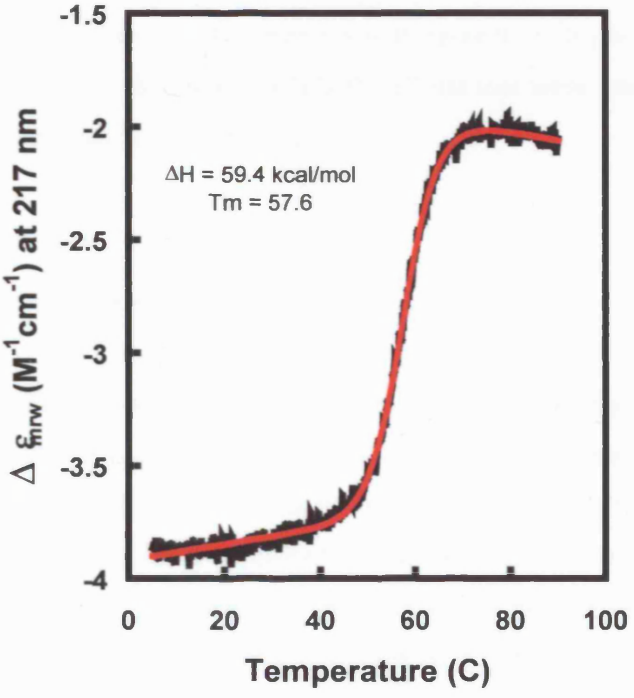


Figure 4.4. Thermal denaturation profile of (A) Rv0019 and (B) Rv0020 FHA domains. Experiments were all conducted in 150mM NaCl, 20mM Tris-HCl pH 8.0. Thermal denaturation was carried out by heating samples at a constant rate of 2°C per min from 5 to 90°C. The black lines represent the recorded CD at 217nm while heating. and red lines are computed derivative plots fitted using a non-linear least-squares analysis.

4.3 Purification and characterisation of PknA

4.3.1 Overview

PknA proved to be the most troublesome to clone and purify of all proteins in this study. Firstly, initial cloning of the protein suggested the protein was toxic to *E.coli* when expressed at basal level that occurs through the *tac* promoter of pGEX vectors. Therefore, only certain fragments, which presumably expressed non-functional protein (no/low kinase activity), could be cloned into pGEX vectors. All constructs were then made using pET vectors to circumvent toxicity problems. pET vectors require the DE3 lysogen which encodes T7 polymerase in host cells for expression of the recombinant protein as expression is under the control of T7 promoter. The T7 promoter system allows greater control of toxic protein expression with reduced basal expression compared to the *tac* promoter. Finally, after several failed solubility trials with constructs of various lengths, $PknA^{1-290}$ gene fragments produced with a N-terminal six-histidine tag gave rise to high levels of soluble expression in *E. coli* although the purified protein was extremely heterogeneous with respect to its phosphorylation state (Table 4.1). The following discussion details the efforts that were made to purify PknA and map the phosphorylation sites of this protein.

4.3.2 Phosphorylation State of PknA

$PknA^{1-290}$ was purified according to the protocol outlined in Section 3 (Materials and Methods). SDS-PAGE analysis of purified PknA showed the apparent molecular weight of PknA differs from the expected molecular weights by around 3kDa (Figure 4.5). The aberrant migration of proteins in SDS-PAGE can be caused by a number of factors including the presence of phosphorylated residues in the protein sample that prevent binding of SDS-molecules to the denature protein polypeptide chain due to their negative charge. Further support for the presence of phosphorylation was found upon

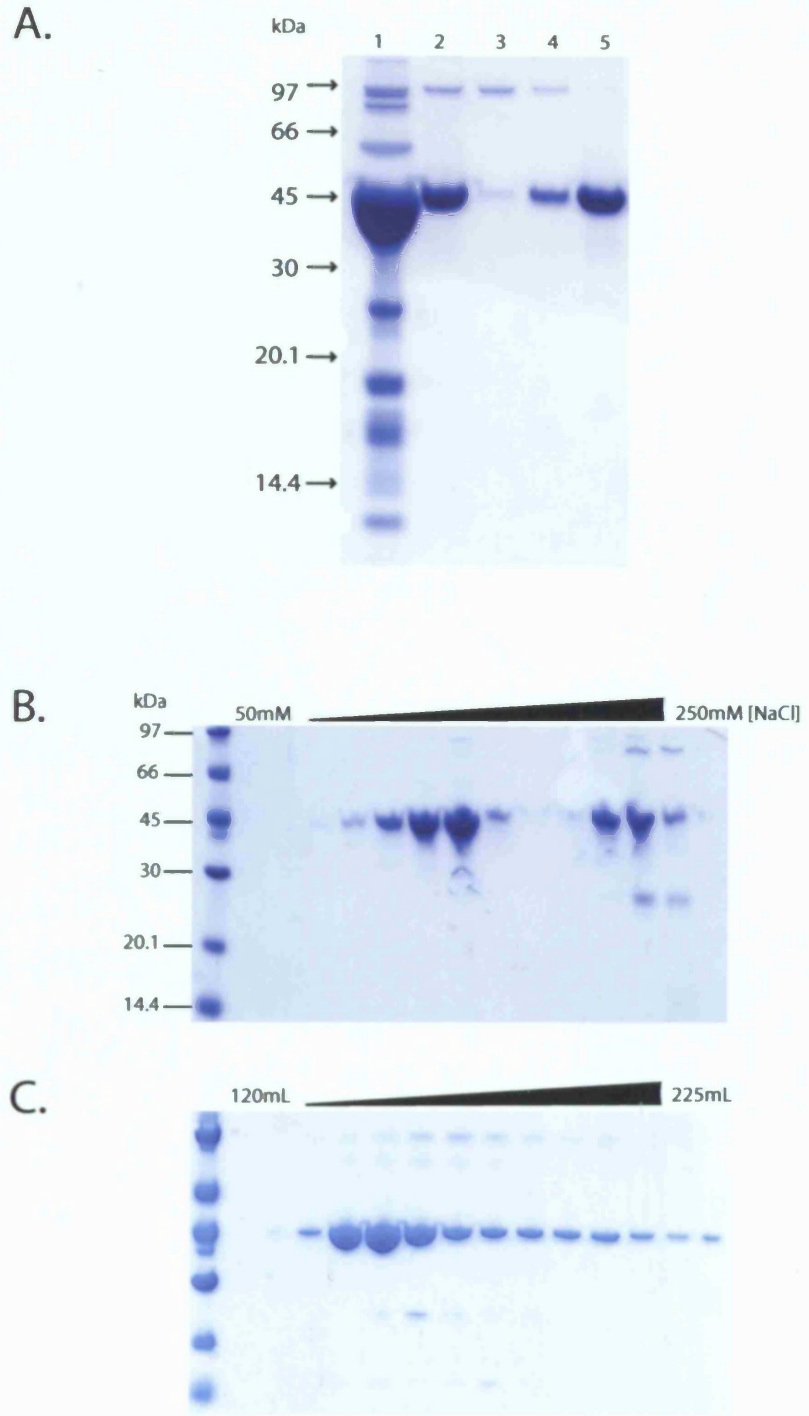


Figure 4.5. SDS-PAGE analysis of recombinant PknA protein. (a) NTA affinity resi n with bound Histidine tagged $PknA^{1-290}$ protein before imidazole elution (lane 1) and after elution (lane2), washing flow-through using 30mM Imidazole (lane3) and samples after protein elution at 100mM and 200mM Imidazole (lanes 4 and 5, respectively). (b) 10µL samples of fractions collected after elution of $PknA^{1-290}$ protein from a Q-sepharose ion-exchange column using 50-250mM NaCl gradient. (c) 10µL samples of protein peak fractions collected after elution of $PknA^{1-290}$ protein from an S200 gel filtration column. The volume at which the protein eluted is shown above the gel based on a column of 300mL.

ion-exchange chromatography. The predicted *pI* of the unphosphorylated form of $PknA^{1-290}$ is 9.42 suggesting it will have a positive in charge at buffer pH 8.0 and bind negatively charged SP-sepharose cation-exchange resin. However, extensive phosphorylation of the protein caused reduction in the pI allowing it to bind to Q Sepharose anion exchange resin (Figure 4.5b). It should be noted that predicted *pI*, is an average of the *pI*s from the constituent amino acids and therefore does not account for basic/acidic region on the protein surface that can influence the binding of proteins to ion-exchange resins. Therefore, $PknA^{1-290}$ may contain a highly acidic region (low *pI*) that facilitates binding to Q-sepharose resin at pH 8.0 even in the absence of phosphorylation.

Subsequent electrospray mass spectroscopy of the protein sample confirmed that PknA was extensively phosphorylated with a minimum of 9 phospho-residues, the most abundant phosphorylation state had 12 phospho-residues, and maximum number observed was 17 (Figure 4.6). Despite reported loss of kinase activity upon mutation of an important catalytic lysine residue in the PknA K42A mutant protein, electrospray data for PknA K42A shows the mutant protein was still phosphorylated on at least three residues. Given this high level of activity, the toxicity associated with $PknA^{1-290}$ may be attributed to non-specific phosphorylation of essential endogenous *E. coli* proteins or cellular ATP shortage.

The phosphorylation of PknA must have occurred within *E. coli* cells prior to purification, due to the absence of ATP post-purification. Many studies have shown that over expression of some eukaryotic kinases in *E. coli* results in autophosphorylation (Hubbard *et al.*, 1998). Hyperphosphorylation is probably due to autophosphorylation through a *trans*-kinase intermolecular reaction as is the case for close homolog PrkC in *Bacillus subtilis*, the only other bacterial kinase where autophosphorylation pattern has been studied (Madec *et al.*, 2002) (Madec *et al.*, 2003).

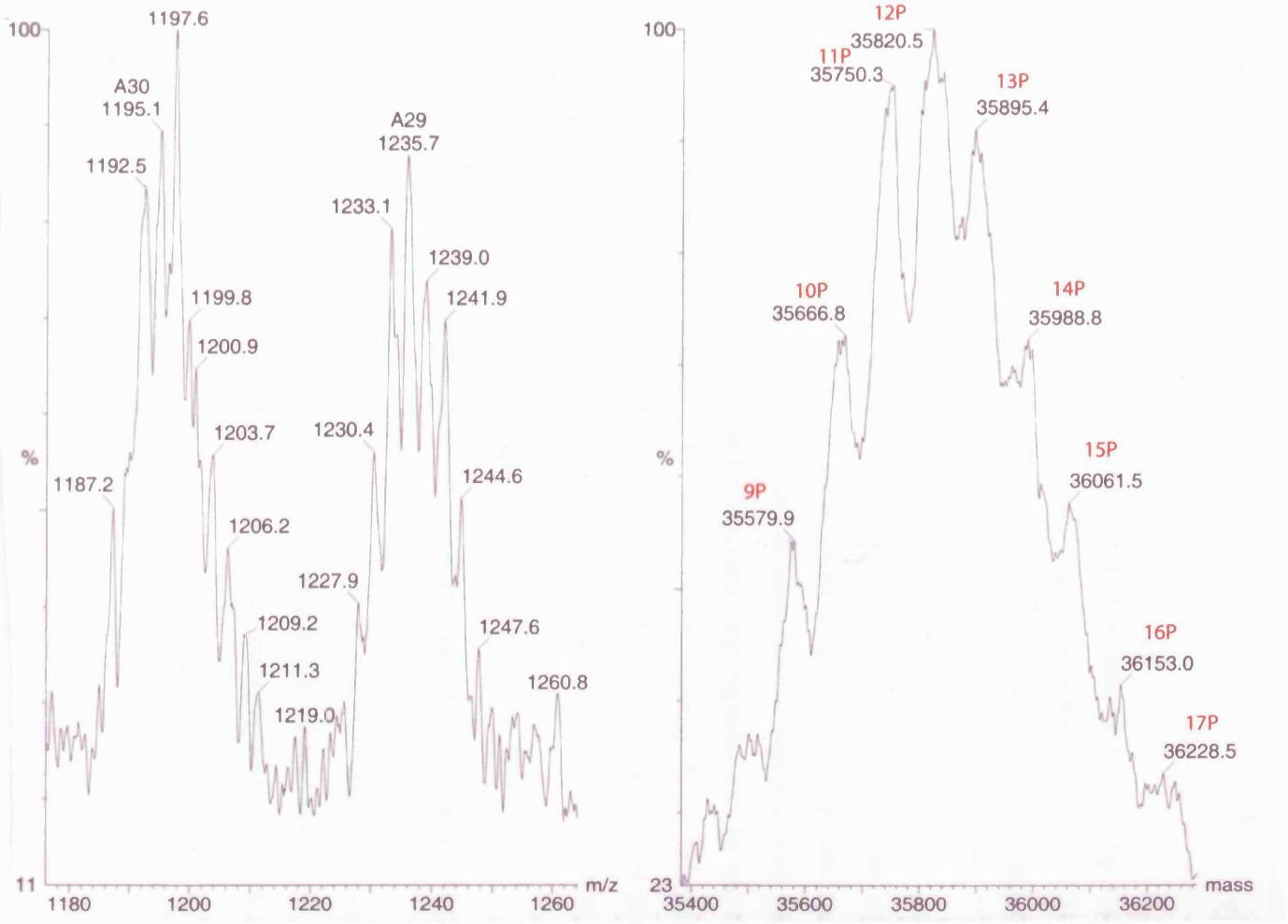


Figure 4.6. Ionisation electrospray mass-spectrum of purified recombinant PknA$^{1\text{-}290}$. The calculated molecular weight was 34855.9Da and the first measured PknA$^{1\text{-}290}$ peak was 35579.9Da, equating to the mass of PknA$^{1\text{-}290}$ plus 9 phosphate groups (80Da each). The number of phosphate groups in each subsequent PknA$^{1\text{-}290}$ protein mass peak is indicated in red above the peaks on the right panel. PknA$^{1\text{-}290}$ protein species with a maximum of 17 phosphate groups and an average of 12 phosphate groups is observed from the spectrum.

4.3.3 Phosphorylation mapping in PknA and PknB

4.3.3.1 Overview

As PknA was observed be multiply phosphorylated at 17 sites by electrospray mass spectroscopy, matrix-assisted laser desorption/ionisation (MALDI)/Time-of-flight (TOF) mass spectrometry was employed to determine the sites of autophosphorylation.

Identification of phosphorylated sites can be achieved by MALDI/TOF measurements of proteolytic digests before and after treatment with a phosphatase. This can be achieved using a generic phosphatase such as alkaline phosphatase, lambda phosphatase or the more specific mycobacterial phosphatase, *ppp* (Boitel *et al.*, 2003). pThr and pSer-containing peptides can be identified from the resulting spectra, indicated by mass values shifting 80Da or multiples thereof (molecular weight of a HPO_3). However, complete detection of all phosphopeptides, particularly those of low stoichiometry, from this complex mixture is difficult for a number of reasons:

- Neutral losses of phosphoric acid and water from phospho-amino acid residues during MALDI/TOF analysis
- Poor desorption/ionisation properties of negatively charged species – especially important in the case of multiply phosphorylated peptides
- Ion suppression effects induced by the presence of dominant unphosphorylated peptides.
- Bias for peptide containing basic residues such as lysine or arginine.

To improve coverage and circumvent these problems, the peptide mixture may be separated by HPLC or enriched for phospho-peptides using affinity methods.

Affinity methods for phosphopeptide enrichment on immobilised gallium ions bound to a chelating gel have been found particularly useful in the identification of phosphorylated peptides for large, multiphosphorlyated proteins (F.Ivins, personal

communication). Therefore, after purification, the phosphorylated $PknA^{1-290}$ was excised from the gel after SDS-PAGE and digested *in situ* with trypsin and Asp-N. The resulting mixture of peptides was applied to an immobilised gallium affinity column and after stringent washing with 10% acetic acid, followed by 10% acetic acid plus 20% acetonitrile, peptides were eluted in a basic buffer (100mM ammonium bicarbonate, pH 9.0) prior to MS analysis.

The principle of this enrichment process is based upon the affinity of the phosphopeptide phosphate groups for the gallium in the range pH 2.0-5.5. Elution of phosphopeptides is performed at a higher pH (>6.0). Acetonitrile was used in elution/wash buffers to reduce non-specific hydrophobic interactions.

The enriched peptide samples were analysed by MALDI-TOF and each candidate peak was subjected to post source decay (PSD). This technique delivers a high burst of energy to the entire peptide sample while a deflector excludes everything outside of a narrow mass range centred on the target peak. If the peptide in this range is phosphorylated, a second peak at 98Da will be seen. This is characteristic of β-elimination of phosphoric acid (H_3PO_4). When the sequence of the protein is known, the mass of the parent peak is usually unique to only one possible tryptic/Asp-N peptide.

If the number of potential sites (Ser/Thr) in a given peptide exceeds the number of phosphorylation sites determined by MALDI-TOF, then the exact position of phosphorylation is still unknown and nanospray mass spectrometry is useful. The entire sample is injected into the nanospray, and the peptides are trapped by a quadrupole ion trap in a resonance determined by their mass to charge (*M/Z*) ratio. The peptide of interest is selected by applying a high energy pulse at all resonances except for that of the target peptide from the ion trap. Then inversely, another pulse is applied at a small window around the resonance for the target peptide. This pulse is not sufficient to cleave the peptide at its weakest bonds. The predominant breakdown product is β-elimination of phosphoric acid as in PSD above. However, in most cases, a subset of the sample breaks down at the weak amide bonds. If this breakage occurs between two ambiguous phosphorylation sites, the fragment containing the phosphorylated residue will have a mass shift of +80Da and the exact phosphorylated residue can be assigned with confidence.

4.3.3.2 Determination of phosphorylation sites in wild-type PknA

Of the 17 phosphorylation sites detected by ionisation-electrospray (17 x 80Da above expected weight), eight were detected by MALDI/TOF with overall sequence coverage of 55%. Using only MALDI/TOF four sites were identified. Additional nanospray fragmentation was required to unambiguously assign the exact location of three sites. One site was narrowed down to a singly phosphorylated peptide with two possible sites (see Figure 4.7 and Table 4.2).

Four phosphorylation sites were found in the N-terminal lobe and four in the C-terminal lobe of PknA with no phospho-residues identified for the vector encoded tag or juxtamembrane linker regions. In addition, the predicted peptide fragment encompassing the activation loop of PknA was not detected by MALDI/TOF. One of the reasons for this maybe the fact that the predicted peptide is relatively large (50 residues) after Asp-N/trypsin digestion and consequently would have difficulty ionising before undergoing fragmentation. Furthermore, the activation loop has a number of potential phosphorylation sites as multiple phosphorylation events are required in this loop for full kinase activity (Boitel *et al.*, 2003). Heavily phosphorylated peptides ionise poorly and therefore are less likely to be detected by MALDI/TOF.

It is important to note that mass spectrometry analysis only gives limited information about the stoichiometry of phosphorylation at a particular site and mainly serves to prove the existence of the phospho-residue in the protein sample. MALDI/TOF also cannot provide reliable detection of phosphorylation at certain residues or a sub-set of residues for reasons listed in Section 4.3.3.1. Whilst, 9 potential sites remained undetermined it was decided further phospho-mapping investigations was not appropriate. Many of the sites currently identified are likely to be non-specific and were found on various sites over the surface of the PknA kinase domain. The biological significance of the bulk of the mapped (and mapped) sites remains to be determined.

4.3.3.3 Determination of phosphorylation sites in PknB K40A

Removal of a conserved catalytic lysine in the kinase domain of PknA and PknB did not completely abrogate autophosphorylation and the reason for this will be discussed in the

following chapter. PknB K40A was found to have two phosphorylation sites. In an effort to determine the phosphorylated residues, a purified PknB K40A protein sample was subjected to MALDI/TOF analysis. The procedure was performed as described above. Unfortunately, neither of the two sites detected by electrospray mass spectroscopy could be located despite 46% sequence coverage, although this did not include the activation loop region.

4.4 The Phosphorylation state of PknB

Sequence analysis of the 626 amino acid protein, PknB, using the SMART database predicts an N-terminal cytoplasmic region and a C-terminal extracellular region separated by a short transmembrane region (Figure 1.11). In this study, two recombinant protein fragments of PknB were expressed and purified from *E. coli* by the methods described in Chapter 3 (Section 3.2.3), $PknB^{1-332}$ and $PknB^{1-279}$.

The protein fragment $PknB^{1-332}$ makes up the entire intracellular region of the transmembrane protein PknB and includes the kinase domain of 279 residues and a juxtamembrane region of 53 residues. If compared to close homolog PrkC in *Bacillus subtilis*, PknB would be expected to form a dimer. In PrkC, autophosphorylation occurs through a *trans*-kinase phosphorylation from one subunit to the opposing monomer following ligand-induced dimerisation of the extracellular PASTA domains (Madec *et al.*, 2002). An X-ray structure of the PknB kinase domain also suggests a dimeric architecture (Young *et al.*, 2003). Using electrospray mass spectrometry, this study found $PknB^{1-332}$ is phosphorylated at total of four sites with a triply phosphorylated species being most abundant. The phosphorylation status of PknB was later determined (Young *et al.*, 2003) and the protein was found to be phosphorylated within the activation loop at residues S162, S166, T171, T173 and within the juxtamembrane linker at T294, S295 and T309 (Figure 4.7). This was confirmed in a more recent study that also mapped phosphorylated residues of PknD, PknE and PknF (Grunder *et al.*, 2005).

Expression and purification of the isolated catalytic core domain, $PknB^{1-279}$, again resulted in a soluble but heterogeneously phosphorylated protein sample. Using high mass accuracy electrospray mass spectrometry we have found that $PknB^{1-279}$ is phosphorylated at a total of five sites with a species phosphorylated at four sites being most abundant. This contradicts previous studies that only found four sites of phosphorylation, as listed above. It is therefore possible that the fifth site of autophosphorylation found here may reside in the LIC vector encoded sequence which adds 13 residues and contains of two serines, although this remains to be confirmed. Mutants of the $PknB_{1-279}$ protein fragment were generated for further analysis and

Table 4.2. Phosphorylated peptides of PknA$^{1-290}$ oberseved by MALDI-TOF

Position on PknA$^{1-290}$	Phosphorylated peptide sequence	-PO_4 groups[a]	Phosphorylated residue
5-12	VGVTLSGR	1[bc]	T8, S10
13-34	YRLQRLIATGGMGQVEVDNR	1	T21
60-89	AEARITAMLNHPGIASVHDYGESQMNGEGR	1,2[c]	T64, T65
117-127	HALDMLEQTGR	1	T125
141-156	DVKPGNILITPTGQVK	1	T152
215-228	RPFAGDGALTVAMK	1	T224
247-260	ELIEITLVKNPAMR	1	T252

[a]Each peptide was found with different degrees of phosphorylation
[b]Phosphorylated residue in mono-phosphorylated peptide was not found
[c]These residues were identified by PSD sequencing

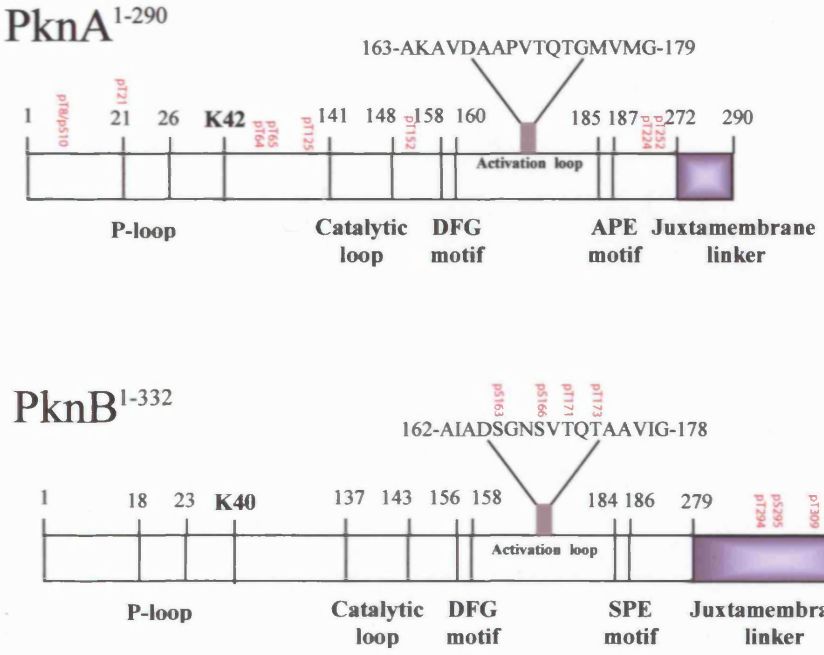


Figure 4.7. Schematic representation of the recombinant protein kinase constructs PknA$^{1-290}$ and PknB$^{1-279}$. Known phosphorylation sites for the two constructs determined by MALDI mass-spectrometry are highlighted in red. All phosphorylated sites determined for PknA$^{1-290}$ were found in this study. Phosphorylation of PknB$^{1-279}$ sites S163, S166, T171, T173, T294 and S295 was first determined by Young *et al.*, 2003 and S309 by Duran *et al.*, 2005.

purified mutant proteins were applied to an electrospray mass spectrometer to confirm the mutation and characterise their phospho-residue content. PknB T173A is maximally phosphorylated on four sites whereas PknB T171A is maximally phosphorylated on three sites. Loss of the T171 therefore has a greater impact on autophosphorylation activity than loss of the T173. Interestingly, PknB K40A was generated as a supposedly 'kinase dead' mutant after removal of an important catalytic lysine that is conserved in all kinases. However, removal of this conserved lysine did not completely abrogate autophosphorylation with two phosphorylation sites identified. All $PknB_{1\text{-}279}$ K40A molecules were phosphorylated on one site with a small proportion phosphorylated on a second site, as determined through electrospray mass-spectrometry. Although this lysine residue is widely conserved and critical for normal kinase function, in the case of PknA and PknB this residue does not prevent partial phosphorylation *in vivo*. Residual activity presumably arises from weak binding of ATP through remaining conserved residues.

4.5 Discussion

In order to deconstruct the signalling pathway of the PknA/PknB operon (Figure 1.9) it was important to produce pure samples of these important proteins as both isolated domains and full cytoplasmic proteins fragments. This task was aided by the fact that the *Mycobacterial* FHA domain-containing Rv0020 and Rv1827 had already been purified as isolated core FHA domains (S. Westcott, unpublished data) (Durocher *et al.*, 2000). In addition, an atomic resolution structure of the core catalytic domain of PknB has been solved (Young *et al.*, 2003) (Ortiz-Lombardia *et al.*, 2003). Challenges remained including the purification of full-length protein of Rv1827, Rv0020, the FHA and cytoplasmic region of Rv0019 and PknA. The cloning and expression of these proteins involved a number of purification problems, as detailed in this chapter. Fortunately, recombinant protein fragments of PknA, Rv0019 (FHA and cytoplasmic region), full-length Rv0020 and full-length Rv1827 were generated in sufficient quantity for subsequent assays.

Although Alber and co-workers expressed the full-length construct of Rv0020 recently (Grunder *et al.*, 2005), little work has been published on PknA or Rv0019 to date. These proteins are important for *M. tuberculosis* viability as they are predicted to play a role in the coordination of cell growth and morphology, possibly in conjunction with other coexpressed proteins including PknB, Rv0020 and Ppp (Figure 1.9) (Section 1.8). Some evidence has suggested that PknA and PknB genes are essential and consequently, they were unable to be removed from the mycobacterial genome (Sassetti *et al.*, 2003). In addition, partial depletion of PknA and PknB in *M. bovis* BCG and *M. smegmatis* resulted in narrow, elongated cell morphology. Over-expression of both proteins in these same strains caused reduction in growth rate and long, broad and in some case branched cell morphology (Yang *et al.*, 2005). Although PknB was found to have a more severe effect on cell morphology than PknA, this illustrates the importance of studying PknA signalling. Until recently, full-length PknA had only been expressed and purified as a fusion with a maltose-binding protein tag (Chaba *et al.*, 2002) and this was used for limited analysis. Here, PknA was purified as a kinase domain with a portion of its juxtamembrane linker (~11 residues) and found to be active.

The cloning, expression and purification of these proteins make it possible to conduct further extensive biochemical and biophysical investigations. It is clear from previous work that FHA domains are involved in the phospho-threonine-dependent signalling directed by STPKs such as PknA and PknB and the genetic linkage of Rv0019 and Rv0020 to these kinases suggests that there is likely to be some interplay between these proteins. The availability of the FHA-domain containing protein Rv1827 means it is now possible to investigate whether PknA and PknB function in a simple linear pathway involving coexpressed proteins or whether they have regulatory roles in conjunction with independently expressed proteins. In the proceeding chapters I have investigated these ideas by using *in vitro* phosphorylation assays and quantitative binding experiments to screen for interactions between these purified proteins.

5 Biophysical Characterisation of Kinase-FHA Interactions

5.1 Introduction

The primary focus of this project is a putative operon in the *M. tuberculosis* genome that contains two FHA domain-containing proteins (Rv0019 and Rv0020) along with two STPKs (PknA and PknB) (Figure 1.9). The co-expression of FHA domain-containing proteins and STPKs, coupled with the fact that STPK domains and FHA domains are often found covalently linked in eukaryotic proteins (i.e. Chk2 and Rad53) strongly suggests that these bacterial FHA domain-containing proteins engage in regulatory phospho-dependent protein-protein interactions. As FHA domains are protein interaction domains which are known to recognise specific phospho-threonine motifs (Durocher *et al.*, 2000) and both PknA and PknB are shown to autophosphorylate at threonine residues (Section 4.3 & 4.4) this indicates FHA domain-containing proteins could form complexes with STPKs.

The potential of kinase-FHA interactions in directing signalling in the *M. tuberculosis* was investigated using a combination of isothermal titration calorimetry (ITC) and surface plasmon resonance (SPR). A detailed explanation of these techniques can be found in Chapter 2. The ITC and SPR experiments described here were carried out in order to establish whether PknA and PknB function in a simple linear pathway where they interact with the co-expressed FHA domain-containing proteins, Rv0019 and Rv0020. Rv1827 is an FHA domain containing protein found encoded at a locus in the *M. tuberculosis* genome far apart from the *pknA* and *pknB* genes and therefore, this protein was investigated as a binding partner for PknA/B to see whether they can play a role in a more complex signal transduction pathway involving FHA domain-containing proteins from other operons. Details on the possible biological role of Rv1827 have been previously discussed (Section 1.8.8).

To characterise kinase-FHA interactions, the kinase domains of PknA and PknB, and FHA domains of Rv0019, Rv0020 and Rv1827 were isolated, then subjected to analysis

by ITC and SPR. In addition phosphopeptides representing the different phosphorylation states of the PknB activation loop were synthesised (Table 5.1), and binding to each FHA domain was measured and compared with the measurements of FHA binding to the intact kinase domains.

5.2 Isothermal Titration Calorimetry Measurements of Kinase-FHA Interactions

5.2.1 Overview

ITC was used in this study to define the equilibrium binding properties and thermodynamic parameters of interactions between STPKs and FHA domains.

Table 5.1. Peptides synthesised for this study

Designation[a]	Sequence
PknB Activation loop	
TQT	A-Y-K-K-N-S-T-Q-T-A-A-V-I
pTQpT	A-Y-K-K-N-S-pT-Q-pT-A-A-V-I
TQpT	A-Y-K-K-N-S-T-Q-pT-A-A-V-I
PTQT	A-Y-K-K-N-S-pT-Q-T-A-A-V-I
Rv1827 N-terminal	
pTTSV	D-E-V-T-V-E-pT-T-S-V-F-R-A-Y-K-K
TpTSV	D-E-V-T-V-E-T-pT-S-V-F-R-A-K-K
Optimal peptide[b]	
Rv1827 optimal peptide	A-L-I-P-pT-Q-L-Y-A-Y-K-K
Rv0020 optimal peptide	D-T-A-P-pT-E-K-I-A-Y-K-K

[a]pT represents phosphorylated threonine residues.

[b]Optimal peptides were determined in a prior study using peptide library screening analysis (Westcott, S., PhD thesis 2003).

In general, when analysing protein-peptide interactions, the peptide was loaded in the syringe at concentrations ranging from 350μM to 1.6mM and the protein in the cell at concentrations between 29μM and 159μM. When analysing protein-protein interactions, the FHA domain was generally loaded in the syringe at concentrations ranging from 217μM to 620μM and the kinase in to the cell at concentrations between 25μM and 59μM. In all titrations, the ratio of concentrations of material in syringe:cell were between 9 and 12 and all were performed at 18°C. Proteins were exchanged from their storage buffer into a buffer containing 150mM NaCl and 50mM Tris-HCl pH 8.0 by overnight dialysis. Phospho-peptides derived from the PknB activation loop amino-acid sequence (168-NSTQTAAVI-176) were dissolved into the identical ITC buffer. In all cases, the resulting data fit to a one-site model for binding.

Table 5.2. Thermodynamic parameters for FHA interactions

Material in syringe	Protein in cell	*n*	K_D (μM)	ΔH_b (kcal/mol)	$T\Delta S_b$ (kcal/mol)	ΔG_b* (kcal/mol)
Protein:Protein						
Rv0020 FHA	PknB Wt	1.3	3.2	-9.2	-1.8	-7.4
	PknB T171A	0.6	2.0	-11.0	-3.3	-7.7
	PknB T173A	0.9	9.4	-15.1	-8.4	-6.7
	PknB T171A/T173A	NDB	-	-	-	-
Rv0020 FHA S473A	PknB Wt	NDB	-	-	-	-
Rv1827 FHA	PknB Wt	0.9	9.3	-7.7	-0.9	-6.8
	PknB T171A	0.7	30.7	-10.1	-4.1	-6.0
	PknB T173A	0.8	22.6	-18.2	-12.0	-6.2
	PknB T171A/T173A	NDB	-	-	-	-
Rv1827 FHA S95A	PknB Wt	NDB	-	-	-	-
FHA:Activation loop peptide[a]						
TQT	Rv0020 FHA	NDB	-	-	-	-
pTQpT		0.7	2.1	-9.8	-2.2	-7.6
TQpT		0.8	3.2	-11.2	-3.8	-7.3
pTQT		0.9	34.0	-11.1	-5.2	-6.0
pTQpT	Rv0020 FHA S473A	NDB	-	-	-	-
TQT	Rv1827 FHA	NDB	-	-	-	-
pTQpT		0.8	11.4	-4.9	1.7	-6.6
TQpT		0.9	14.3	-3.2	3.2	-6.4
pTQT		1.0	23.6	-5.8	0.4	-6.2
pTQpT	Rv1827 FHA S95A	NDB	-	-	-	-

* $\Delta G_b = \Delta H_b - T\Delta S_b$. All ITC experiments were conducted at 291K (18°C).

[a]Full peptide sequences are provided in Table 5.1. (NDB) No detectable binding. Stoichiometry (*n*), equilibrium binding constant (K_D), observed binding enthalpy (ΔH_b), observed binding entropy (ΔS_b), and binding free energy (ΔG_b)

5.2.2 The FHA domains of Rv0020 and Rv1827 interact with the PknB kinase domain

Calorimetric titration of the PknB kinase domain with FHA domains of Rv0020 and Rv1827 generated measurable interactions in the absence of ATP (Figure 5.1). The thermodynamic parameters stoichiometry (n), equilibrium dissociation constant (K_D), observed binding enthalpy (ΔH_b), observed binding entropy (ΔS_b), and binding free energy (ΔG_b) were determined from calorimetric data and are listed in Table 5.2.

It is known that mutation of a conserved serine within the FHA domain of both Rv1827 and Rv0020 FHA domains (S95 and S473, respectively) results in complete loss of phospho-threonine binding (Durocher *et al.*, 2000, Westcott S., PhD thesis 2003). Importantly, no binding was detected upon titration of serine mutants of Rv0020 (Rv0020 S473A) and Rv1827 (Rv1827 S95A) with both phosphopeptides and purified phosphorylated kinase domains (Table 5.2 and Figure 5.2), underlining the fact that FHA-phosphothreonine binding is essential for kinase-FHA complex formation, and that these conserved serine residues are an absolute requirement for phosphothreonine binding.

5.2.3 The site of FHA domains of Rv1827 and Rv0020 interaction is the PknB activation loop

The autophosphorylation of PknB has been mapped to four residues in the activation loop (S166, S169, T171, T173)(Young *et al.*, 2003) and the presence of four phosphorylated residues has been confirmed in wild type PknB protein used here by electrospray mass spectroscopy (Table 4.1 and Figure 4.7). The fact that FHA domains are protein interaction domains that recognise specific phospho-threonine motifs suggests that T171 and/or T173 of the PknB activation loop are the likely binding site(s) for the FHA domains.

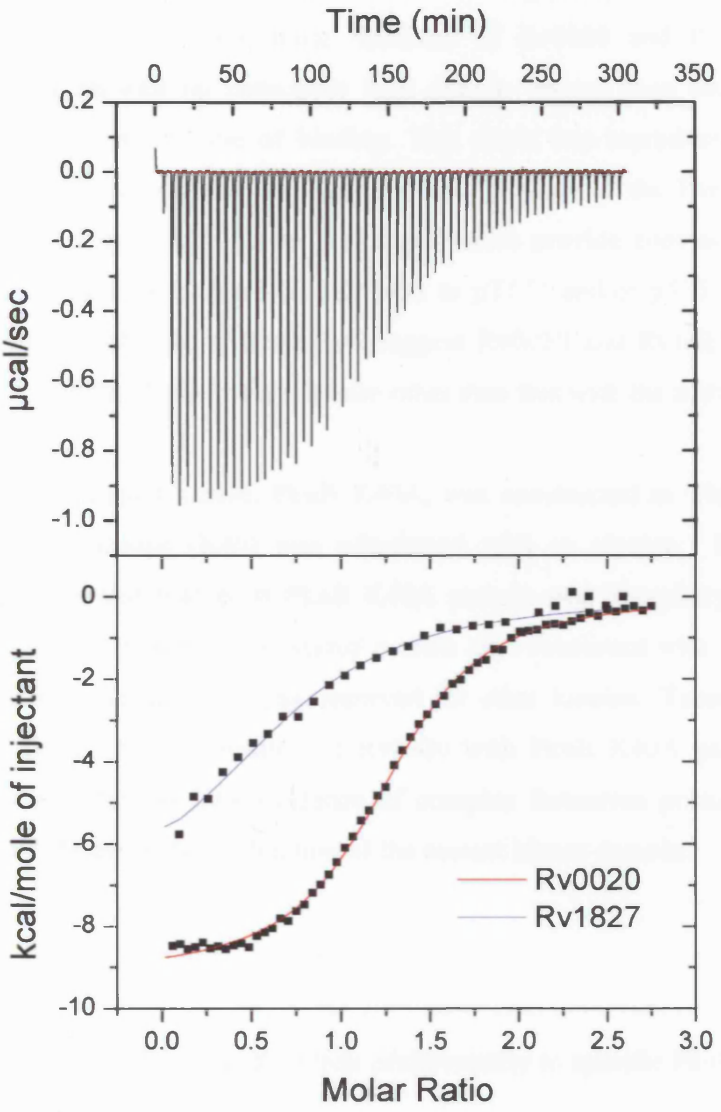


Figure 5.1. Binding isotherms of the FHA domains of Rv1827 and Rv0020 for the phosphorylated kinase domain of PknB, $PknB^{1-279}$, determined by isothermal titration calorimetry. The upper panel shows a representative raw ITC trace for the binding of Rv0020 FHA domains binding to $PknB^{1-279}$. The lower panel shows the overlapping ITC data for interactions between $PknB^{1-279}$ and both Rv0020 and Rv1827 FHA domains fit to a one-site binding model. All of the observed binding stoichiometries were consistent with a 1:1 complex between the FHA and kinase domain. Resultant thermodynamic parameters are shown in Table 11.

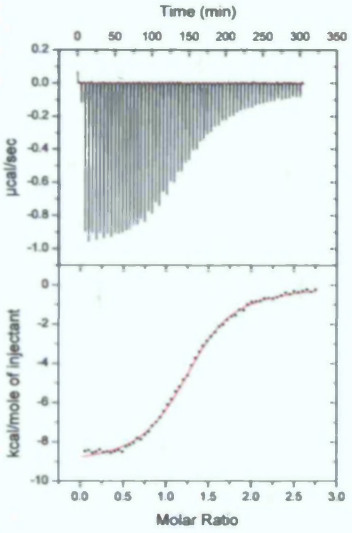
Activation loop residues, T171 and T173 were therefore mutated to alanines with the expectation that this might interfere with PknB-FHA complex formation. After purification of the PknB T171A/T173A kinase domain protein, use of electrospray mass spectrometry confirmed the presence of all remaining phosphorylation sites (S166 and S169)(Table 4.1). The calorimetric titrations of Rv0020 and Rv1827 with PknB T171A/T173A showed no detectable heat change greater than the heat of dilution consistent with complete loss of binding. This result was reproduced when replacing PknB T171A/T173A with an unphosphorylated peptide of the PknB activation loop (Table 5.2 and Figure 5.2). These ITC experiments provide convincing evidence that FHA domains of Rv0020 and Rv1827 bind to pT171 and/or pT173 in the activation loop of PknB. Furthermore, these data suggest Rv0020 and Rv1827 do no form any interactions with the PknB kinase domain other than that with the activation loop.

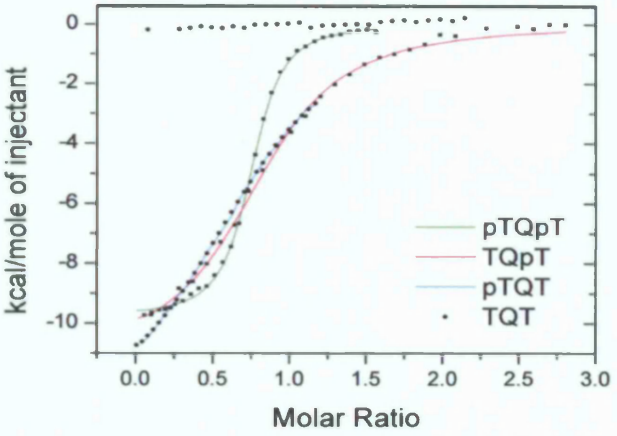
A fourth PknB mutant kinase, PknB K40A, was constructed in which the conserved catalytic lysine residue (K40) was substituted with an alanine. Electrospray mass spectrometry showed that most PknB K40A protein was phosphorylated on one site with a small proportion phosphorylated on two sites consistent with a severe reduction in autophosphorylation activity as observed for other kinases (Table 4.1)(Ozes *et al.*, 1999, Iyer *et al.*, 2005). Titration of Rv0020 with PknB K40A generated a binding isotherm which showed little evidence of complex formation presumably due to the small active (phosphorylated) fraction of the mutant kinase domain.

5.2.4 The Rv0020 FHA domain binds preferentially to specific PknB phosphorylation states

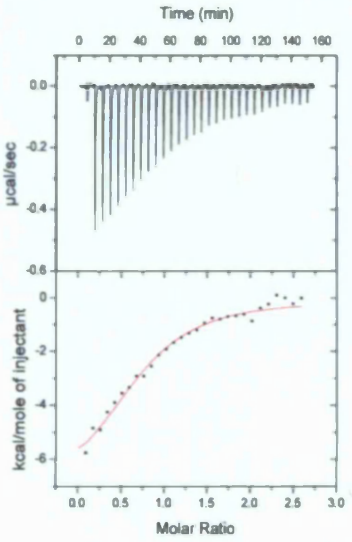
After establishing an interaction between the FHA domain of Rv0020 and the kinase domain of PknB and establishing this involved binding of Rv0020 to phosphothreonines of the activation loop (pT171 and pT173), it was necessary to determine if either phosphorylated threonine was favoured. After generating two PknB mutants (PknB T171A and PknB T173A) and synthesis of activation loop peptides (pTQT, TQpT and pTQpT) a series of ITC experiments were conducted to clarify these issues.

A. Rv0020





B. Rv1827



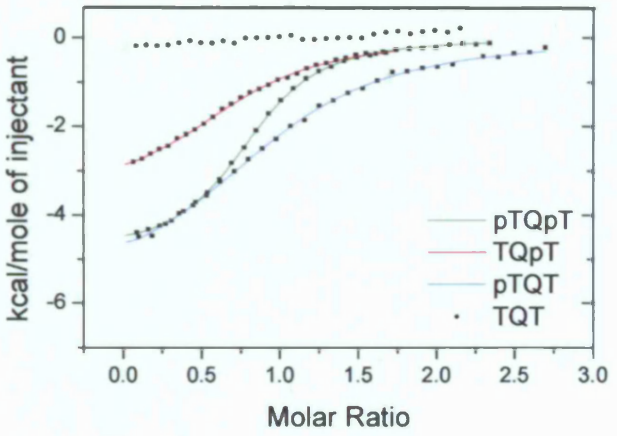


Figure 5.2. Binding isotherms of the FHA domains of Rv1827 and Rv0020 for the phosphorylated kinase domain of PknB, $PknB^{1-279}$, and PknB activation loop peptides determined by isothermal titration calorimetry. The upper left panel shows a representative raw ITC trace for the binding of Rv0020 FHA domains binding to $PknB^{1-279}$, besides which are overlapping ITC data for interactions between Rv0020 FHA and various activation loop peptides fit to a one-site binding model. The lower two panel show the same for the Rv1827 FHA domain. All of the observed binding stoichiometries were consistent with a 1:1 complex between the FHA and kinase domain. Resultant thermodynamic parameters are shown in Table 11.

Rv0020 showed a preference for binding to the PknB activation loop mutants in this order: PknBWt ≈ PknBT173 > PknBT171 only. The equilibrium dissociation constant (K_D) observed for the binding of Rv0020 to PknB (wild-type and T171A) was ~2-3μM (Table 5.2). The interaction of the Rv0020 FHA domain to PknB T173A was weaker, with a K_D of 9.4μM. Similar results were obtained after analysis of protein-peptide interactions. The binding of the Rv0020 FHA domain with peptides pTQpT (representing wild-type PknB) and TQpT (representing PknB T171A) also had a K_D of ~2μM. As might have been predicted, the Rv0020 FHA domain interaction with pTQT (representing PknB T173A) was significantly weaker, with a K_D of 33.5μM, consistent with a small but significant preference of the Rv0020 FHA domain for pT173 over pT171. This observation will be discussed at the end of the Chapter (Section 5.4).

5.2.5 Rv1827 shows no specificity for particular PknB phosphorylation states

In a similar way, the mutant kinase domain and phosphopeptides were used to determine which phosphorylated threonine in the PknB activation loop is favoured by the Rv1827 FHA domain.

In contrast to the Rv0020 FHA interactions described above, Rv1827 showed slight preference for binding to the wild-type kinase domain (K_D of 9μM) over both mutants PknB T171A and T173A (K_D of 31μM and 23μM, respectively) indicating that phosphorylation of both residues contributes equally to overall binding. The binding of the Rv1827 FHA domain with peptides pTQT (representing PknB T173A) and TQpT (representing PknB T171A) gave a K_D of 24μM and 14μM, respectively. The Rv1827 FHA domain interacted with pTQpT (representing wild-type PknB) with a K_D of 11μM. Thus, the results from the protein-protein titrations were recapitulated by the protein-peptide interactions measured although the discrimination was less marked.

5.2.6 Rv0019 is unable to bind phospho-peptides

The minimal tryptic fragment of Rv0019 was determined (Section 4.2.3) encompassing the FHA domain homology region. An ITC experiment was conducted with the cell containing this purified Rv0019 FHA domain (27.8μM) and the syringe containing wild-type PknB kinase domain (175μM). All other conditions were as described above. Surprisingly, the ITC data showed no detectable binding even when repeated under a number of different conditions (i.e. changing protein concentrations, conducting a reciprocal titration by swapping components in the cell and syringe, conducting titrations at different temperatures). In addition, an ITC experiment was conducted with the cell containing purified Rv0019 FHA domain at five times the concentration (132μM) and the syringe an appropriately higher concentration of the TQpT peptide (1.12mM). Again no binding was detected.

Since it seems possible that proteolytic fragment may have contained a structurally compromised FHA domain, ITC experiment was also conducted between the full intracellular domain of Rv0019 as a GST fusion, GST-Rv0019$^{29\text{-}156}$, and the PknB activation loop peptide, pTQpT. Here, the cell contained the GST-Rv0019$^{29\text{-}156}$ protein (27.8μM) and the syringe contained the pTQpT peptide (175μM). No binding signal was observed in either of two separate experiments.

Although the reasons for lack of binding activity of full-intracellular Rv0019 protein or its isolated FHA domain are unclear, there are a number of possible explanations. Firstly, the Rv0019 FHA domain may have a completely different specificity to that of other FHA domains of Rv0020 and Rv1827. Secondly, as mentioned above, Rv0019 may fail to fold correctly when produced in *E.coli*. Nevertheless, the protein construct Rv0019$^{56\text{-}156}$ was stable after purification (i.e. no aggregation of protein samples) and thermal stability similar to that of the Rv0020 FHA domain (Section 4.2.4)(Figure 4.4). Despite displaying good homology to known functional FHA domains (Figure 1.3) the Rv0019 FHA domain does show variation in its far UV CD spectrum, although it nonetheless is consistent with a high proportion of β-sheet structure (Figure 4.3). The sequence of the Rv0019 FHA domain is close enough to the Rv0020 FHA to create a model of the Rv0019 FHA domain based in the Rv0020 FHA domain (Figure 5.3). This served to demonstrate that conserved residues of Rv0019 would be correctly placed for phosphopeptide binding. Sequence analysis using SMART also shows Rv0019 contains

a transmembrane region and therefore it would be membrane-associated along with PknA and PknB. This suggests it could bind to PknA and/or PknB at weak affinities aided by co-localisation with these proteins and the resulting high local concentrations. Nonetheless, the association constants for this interaction in free solution may be below the detection limit of ITC. Finally, it is possible that Rv0019 plays non-phospho-dependent role in TB signalling.

In a related case, the mitotic checkpoint protein Chfr contains an FHA domain that failed to show binding to phosphopeptides in a phosphopeptide library screen (M. Yaffe, unpublished data). The structure of the Chfr FHA was solved with a tungstate, a phosphate analog, bound in the expected phosphate binding site (Stavridi *et al.*, 2002), suggesting that it is able to bind phosphoproteins by a similar mechanism to other known FHA domains.

The phosphorylation of Rv0019 the FHA domain by PknB may represent a mechanism by which STPKs may modulate the FHA domains phospho-peptide binding ability, and this was investigated in the next Chapter. For example phosphorylation may lead to activation of the FHA phosphothreonine binding ability. The important DNA-damage response protein in eukaryotes, Chk2 contains a STPK domain that is also able to phosphorylate its N-terminal FHA domain. In this case, this modification reduces the Chk2 FHA domains affinity for a primary phosphorylation site on the protein (Ahn et al., 2002).

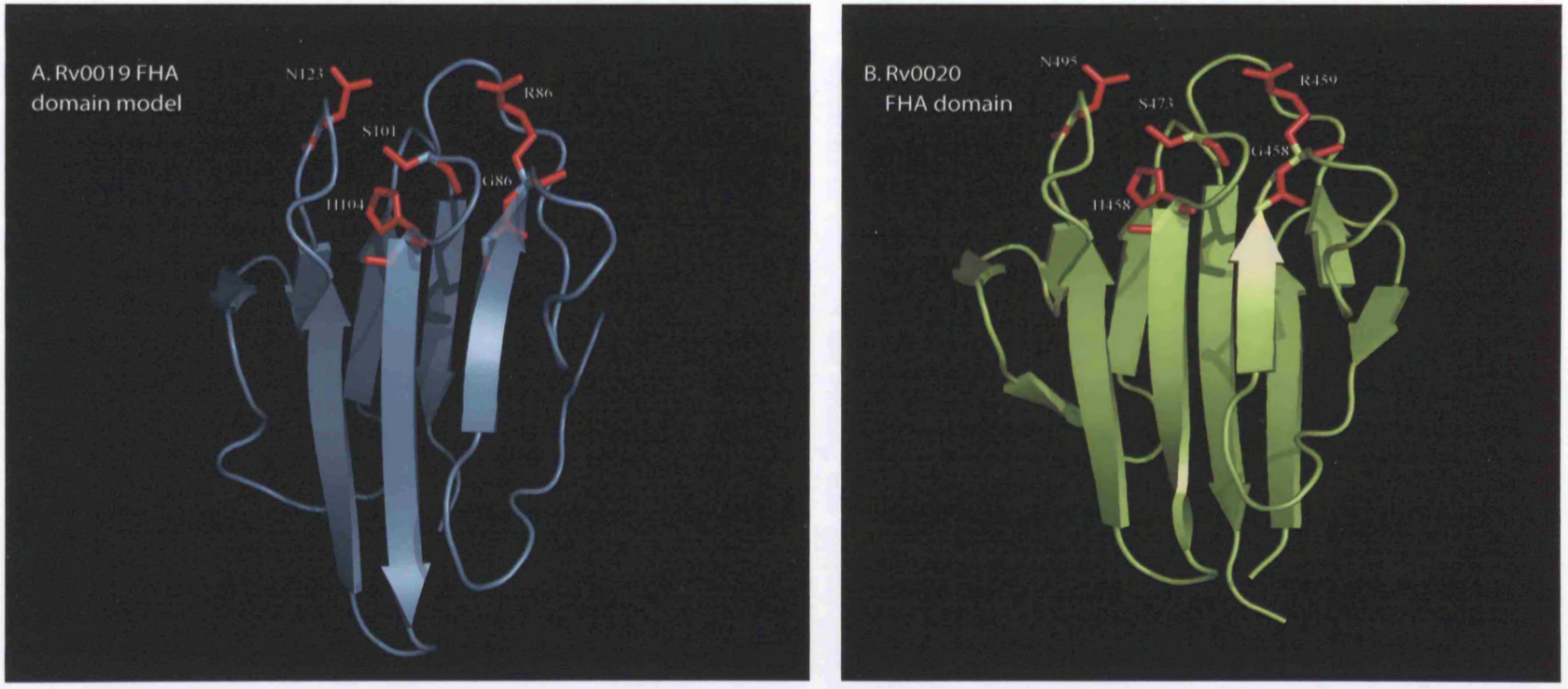


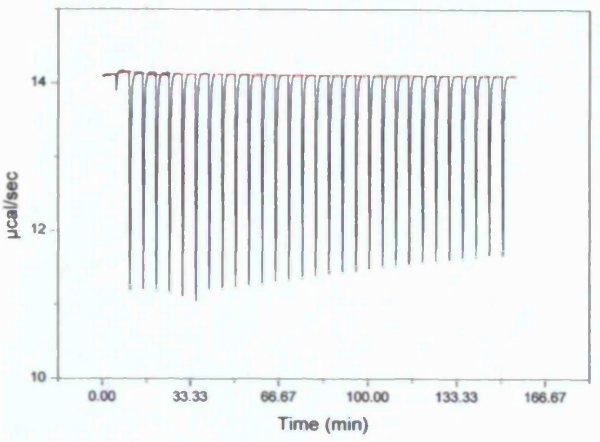
Figure 5.3. A model of the Rv0019 FHA domain based on the known structure of the Rv0020 FHA domain. Both structures are depicted as cartoon representations with conserved residues shown in stick representation in red. The model of the Rv0019 FHA domain was constructed by threading the amino-acid sequence of the Rv0019 FHA homology region onto the structure of the Rv0020 FHA domain (Westcott, S., PhD thesis 2003) using SwissModel first approach mode (Schwede *et al*,. 2003).

5.2.7 The PknA-FHA interaction

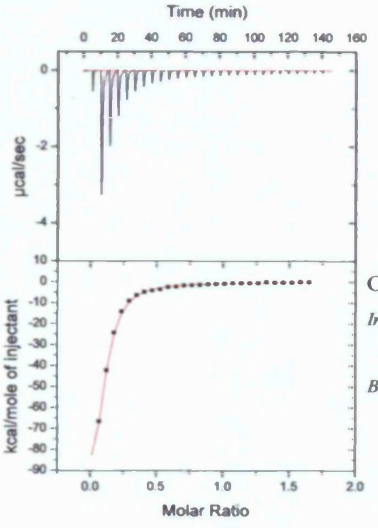
Purification of PknA resulted in a heterogeneous and hyperphosphorylated protein population as determined by electrospray mass spectrometry (Section 4.3). Calorimetric titration, carried out with the Rv0020 FHA domain in the syringe at 608μM and PknA in the cell at 70μM, produced a significant heat change (–3kcal/mol) for all injections without saturation of binding sites (Figure 5.4A). Phosphorylation mapping performed earlier (Section 3.2.4) demonstrated that there are a number of potential binding sites for FHA domains as the protein is heavily autophosphorylated. It is possible that multiple FHA molecules could bind to the phospho-threonines of PknA, as all FHA domains have some non-specific affinity for phospho-threonine residues; this would result explain the slow saturation of binding sites upon titration. To test this theory a reciprocal titration was conducted with the Rv0020 FHA (38μM) in the cell and PknA (300μM) in the syringe. Indeed, this produced a substantial heat change (–70kcal/mol) for early injections of 10μL but the binding was quickly saturated consistent with the presence of multiple sites on the purified PknA kinase domain (Figure 5.4B). This titration displayed an n value of ~0.09 and yielded a K_D of 2μM, which is close to the SPR observed K_D of 1.6μM (Section 5.3.2 and Table 5.3). It should be noted that the apparent thermodynamic and affinity parameters for this interaction would be an amalgamation of many interactions and therefore the autophosphorylation on PknA would have to be reduced to the level of PknB before and meaningful data can be obtained.

Steps were taken to improve the thermodynamic and affinity data of the ITC experiment described above, conducted with the Rv0020 FHA (38μM) in the cell and PknA (300μM) in the syringe (Figure 5.4B), in light of data suggesting the PknA contained multiple FHA binding sites. The raw data generated from the ITC was input into the Origin ITC analysis software, with a higher than actual concentration of PknA also input and data was fitted as before to a one-site model. The concentration of PknA was increased iteratively until the fit provided an n value close to 1. This simulated a decrease in the amount of sites in a single PknA molecule, compensating for the multiple sites present in the PknA molecules and thereby allowing the data to fit better to a one-site model.

A.



B.

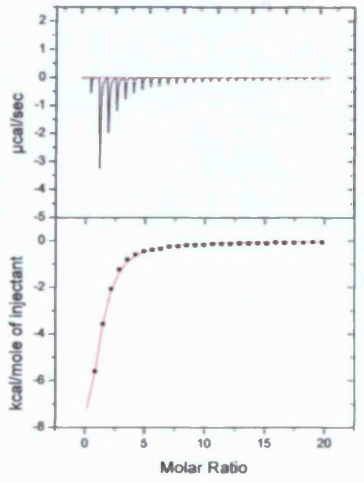


Observed PknA:Rv0020 FHA interaction

Initial concentrations: *Cell=Rv0020 FHA* (38µM)
Syringe=PknA (300µM)

Binding parameters: *n* 0.09, K_D 2µM,
D*H* -131kcal/mol, D*S* –424cal/mol
D*G* –7.6 kcal/mol

C.



Modelled PknA:Rv0020 FHA interaction

concentrations: *Cell=Rv0020 FHA* (38µM)
Syringe=PknA (3600µM)

Binding parameters: *n* 0.97, K_D 29µM,
D*H* -13.6kcal/mol, D*S* –24.4cal/mol
D*G* –6.5 kcal/mol

Figure 5.4. Binding isotherm of the Rv0020 FHA domain for the hyperphosphorylated kinase domain of PknA. (A) a representative raw ITC trace with Rv0020 FHA domain in the syringe at 608µM and PknA in the cell at 70µM. No saturation of binding was evident. (B) the raw ITC trace and binding isotherm generated by a reciprocal ITC experiment with Rv0020 FHA domain in the cell at 13µM and PknA in the cell at 150µM. (C) the binding isotherm generated from the ITC data shown in panel B after increase the computationally increasing the PknA concentration 12-fold. All the available Rv0020 FHA binding sites were quickly saturated. Data fit to a one-site model and the calculated thermodynamic parameters are given to right of the graph.

A value of 0.97 for n was achieved when the true syringe concentration of PknA was increased 12-fold, implying each PknA molecule could bind to 12 FHA domains. The resultant data from this modelled fit (Figure 5.4C) lie within the range of values expected for FHA-phosphothreonine interactions. The enthalpy change of binding, ΔH_b, was –13 kcal/mol, which is similar to that observed for the binding of the Rv0020 FHA domain to PknB (-9.2 to 15.1kcal/mol)(Table5.2). Furthermore, the equilibrium constant, K_D, for the modelled fit was ~30μM, which is within the range expected for FHA-phosphothreonine interactions despite being rather weaker than interaction measured previously involving the Rv0020 FHA domain (Table 5.2). It should be noted that the apparent thermodynamic and affinity parameters for this interaction would be a combination of multiple interactions. Therefore, the relatively weak interaction observed between Rv0020 FHA domain and PknA would be an average of numerous weak non-specific interactions with a small contribution from tight biologically relevant interactions. It is likely that the PknA activation loop is the only true target for binding by the Rv0020 FHA domain, and this interaction may have similar properties to that observed for PknB as activation loop sequences are highly homologous (Figure 1.11). Whilst these data cannot be considered a true reflection of the biological interaction between PknA and Rv0020 they do suggest that PknA may indeed be a bona fide binding partner of Rv0020 *in vivo*.

5.3 Surface Plasmon Resonance

5.3.1 Overview

SPR was performed according to the procedure outlined in Section 3.2.5. In general, when analysing protein-protein interactions it is advisable to immobilise the smaller component of the complex on the surface of the sensor-chip as this provides a greater signal (related to mass change) upon complex formation. In this case, the smaller components (FHA domains) were more difficult to immobilise due to a lack of surface accessible lysine groups that are utilised in the immobilised reaction, and therefore the kinase domains of PknA and PknB were immobilised instead. Immobilisation of the purified kinase domains onto a carboxymethyldextran-modified gold surface was achieved using standard manufacturers guidelines, as described in Section 3.2.7. All proteins were immobilised to a level of around 1000 response units (RU) using a standard buffer containing 10mM sodium acetate, pH 4.5-6, at 40-120μg/mL protein concentrations. Equilibrium binding analysis was performed by injecting FHA domain over the immobilised protein surface and similarly over a control surface at varying concentrations and at a flow-rate of 5μL/min. Kinetic experiments were conducted at a flow-rate of 50μL/min. The sensor chip did not require a regeneration step as binding was relatively weak, meaning all proteins could be removed by running HBS-EP buffer over the surface for 5 minutes to return to a baseline response (Figure 5.5). Readings were taken for at least 12 different FHA concentrations and equilibrium binding constants were calculated by fitting the response *versus* substrate concentration to the single-site saturation curve described in Section 2.1.

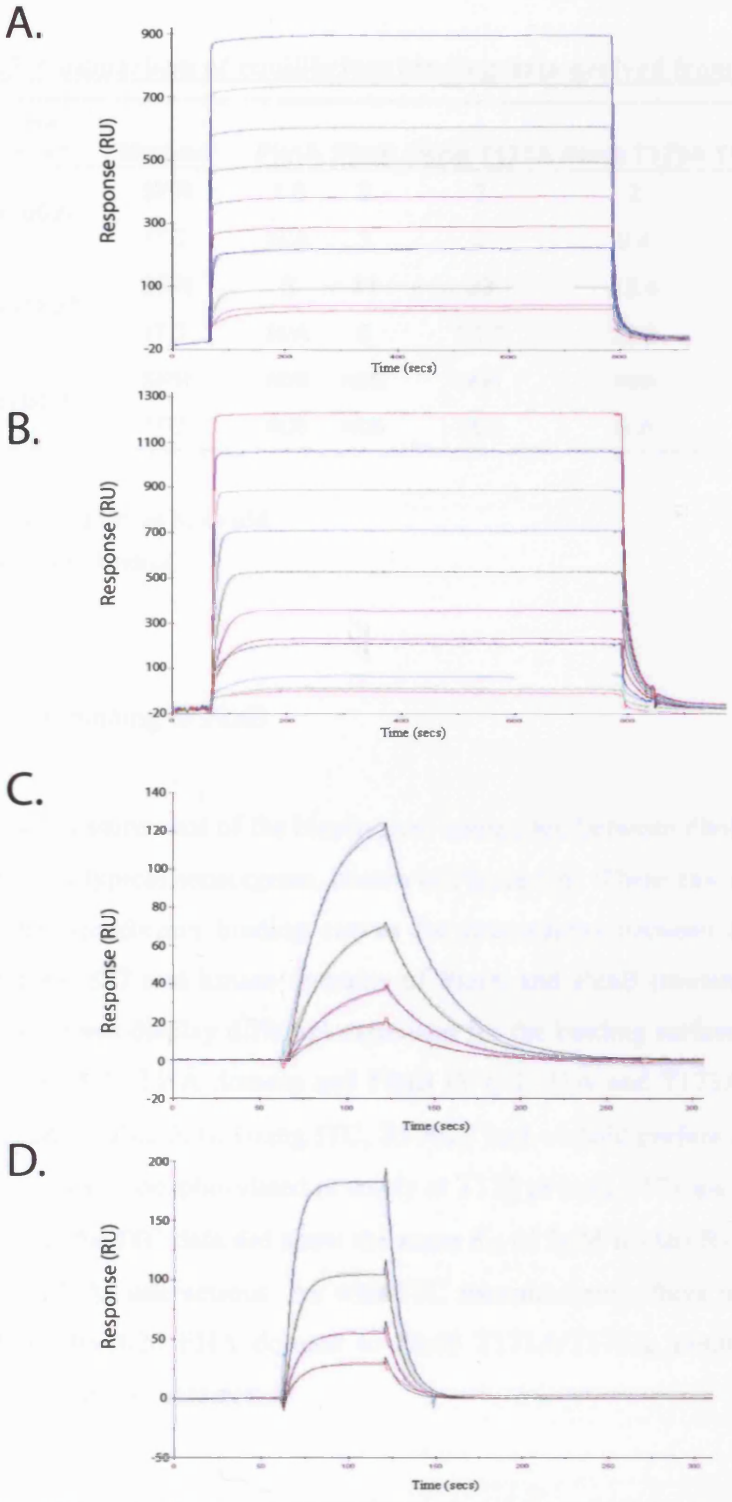


Figure 5.5. Equilibrium and kinetic analysis of Rv0020 binding with immobilised PknA and PknB conducted by Surface Plasmon Resonance. Equilibrium analysis of Rv0020 with immobilised PknA and PknB. (A) A series of concentrations of Rv0020 (0.02, 0.039, 0.08, 0.16, 0.3, 0.63, 1.25, 2.5, 5, 10, 20 μM) were injected over immobilised $PknA^{1-290}$ and (B) $PknB^{1-279}$ until the maximal response was achieved. The flow speed was 5μL/min. The recorded response at equilibrium was corrected for background from a control flow cell. Kinetic analysis of Rv0020 with immobilised PknA and PknB. (C) A series of concentrations of Rv0020 (16, 31, 63, 125nM) were injected over immobilised $PknA^{1-290}$. (D) A series of concentrations of Rv0020 (31, 63, 125, 250nM) were injected over immobilised $PknB^{1-279}$. The flow speed was 50μL/min. The k_{on} and k_{off} were determined from the relevant portions of the sensorgram curve (coloured lines) using BIAevaluation 3.0 software using global fitting to a 1:1 Langmuir binding model. The fitted curves (black lines) are overlaid on the sensorgram traces. All samples were in HBS buffer pH 7.4 at 25°C. Resultant equilibrium binding parameters are shown in Table 12.

Table 5.3. Comparison of equilibrium binding data derived from ITC and SPR

FHA domain	Method*	PknA	PknB	PknB T171A	PknB T173A	PknB T171A/T173A
Rv0020	SPR	1.6	2	2	2	NDB
	ITC	N/A	3	2	9.4	NDB
Rv1827	SPR	5	11	23	13.4	NDB
	ITC	N/A	9	30.7	22.6	NDB
Rv0019	SPR	NDB	NDB	NDB	NDB	NDB
	ITC	NDB	NDB	NDB	NDB	NDB

*All binding data is given as K_D in μM

(NDB) No detectable binding

5.3.2 Rv0020 binding to PknB

The real-time measurement of the biophysical interaction between PknB and Rv0020 by SPR generated a typical sensorgram, shown in Figure 5.6. These raw data were used to determine the equilibrium binding curves for interactions between FHA domains of Rv0020 and Rv1827 and kinase domains of PknA and PknB (mutant and wild-type). Although the curves display different capacities for the binding surface, the interactions between the Rv0020 FHA domain and PknB (Wt, T171A and T173A) resulted in the same K_D (2μM) (Table 5.3). Using ITC, Rv0020 had ~4-fold preference for binding to the activation loops phosphorylated at solely at T173 or both T171 and T173 over T171 only. However, the ITC data did show the same K_D of 2μM for the Rv0020 FHA-PknB (wild-type, T171A) interactions. As with ITC measurements, there was no detectable binding of the Rv0020 FHA domain to PknB T171A/T173A, confirming that these residues are the sites of interaction.

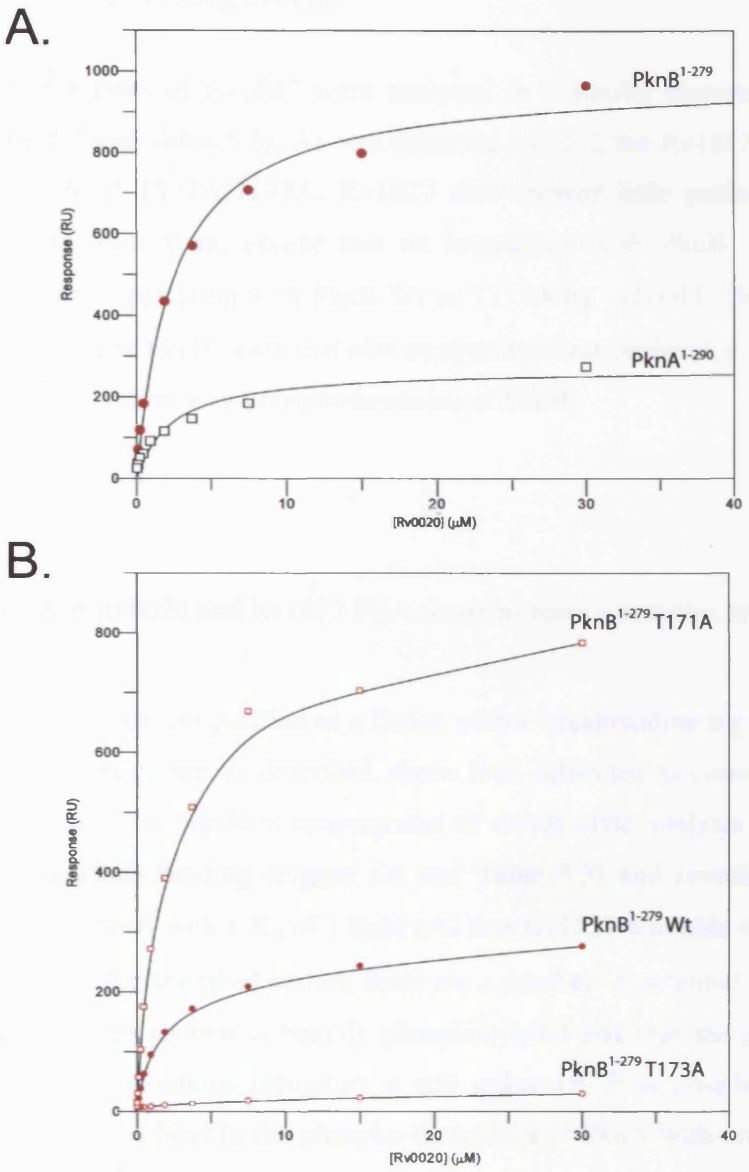


Figure 5.6. Equilibrium analysis of Rv0020 FHA domain with immobilised $PknA^{1-290}$ and $PknB^{1-279}$ conducted using surface plasmon resonance. Data from equilibrium experiments were analysed by fitting to a saturation binding curve using a one-site binding model by non-linear regressional analysis. (A) binding curves for the interaction of Rv0020 FHA domain to wild-type forms of $PknA^{1-290}$ and $PknB^{1-279}$. (B) binding curves for the interaction of Rv0020 FHA domain to wild-type and mutant $PknB^{1-279}$. All samples were injected over SPR surfaces in HBS buffer pH 7.4 at 25°C. Data were fitted to a single-site binding curve using non-linear regressional analysis and all fits were generated from the average of three independent experiments. All values were within 1 standard deviation of the average data point. The measured interaction parameters are shown in table 12.

5.3.3 Rv1827 binding to PknB

The interactions of Rv1827 were analysed in a similar manner to those of Rv0020 (Figure 5.7 and Table 5.3). As was observed by ITC, the Rv1827 FHA domain did not bind to PknB T171A/T173A. Rv1827 also showed little preference for a particular phosphorylation state, except that its interaction with PknB T171A was modestly weaker than interaction with PknB Wt or T173A by ~2-fold. The equilibrium binding data are similar to ITC data that also suggest no clear preference of Rv1827 for binding to either activation loop phosphothreonine of PknB.

5.3.4 The Rv0020 and Rv1827 FHA domains form a complex with PknA

PknA kinase domain purified as a fusion with a hexahistidine tag was immobilised onto the CM5 sensor chip as described above then subjected to conventional SPR analysis (Figure 5.5). The resultant sensorgrams of steady-state analysis were similar to those involving PknB binding (Figure 5.6 and Table 5.3) and revealed that Rv0020 FHA could bind PknA with a K_D of 1.6μM and that Rv1827 was able to bind to PknA with a K_D of 5μM. As described earlier, there are a number of potential binding sites for FHA domains as the protein is heavily phosphorylated and that the precise phospho-motif that the FHA domains recognise is still unknown. It is possible that multiple FHA molecules could bind to the phospho-threonines of PknA with only a subset binding to sites of physiological relevance. However, the stoichiometry of binding cannot be determined as the immobilised species of SPR experiments assumed to be in excess and therefore saturation of the binding sites is unobtainable.

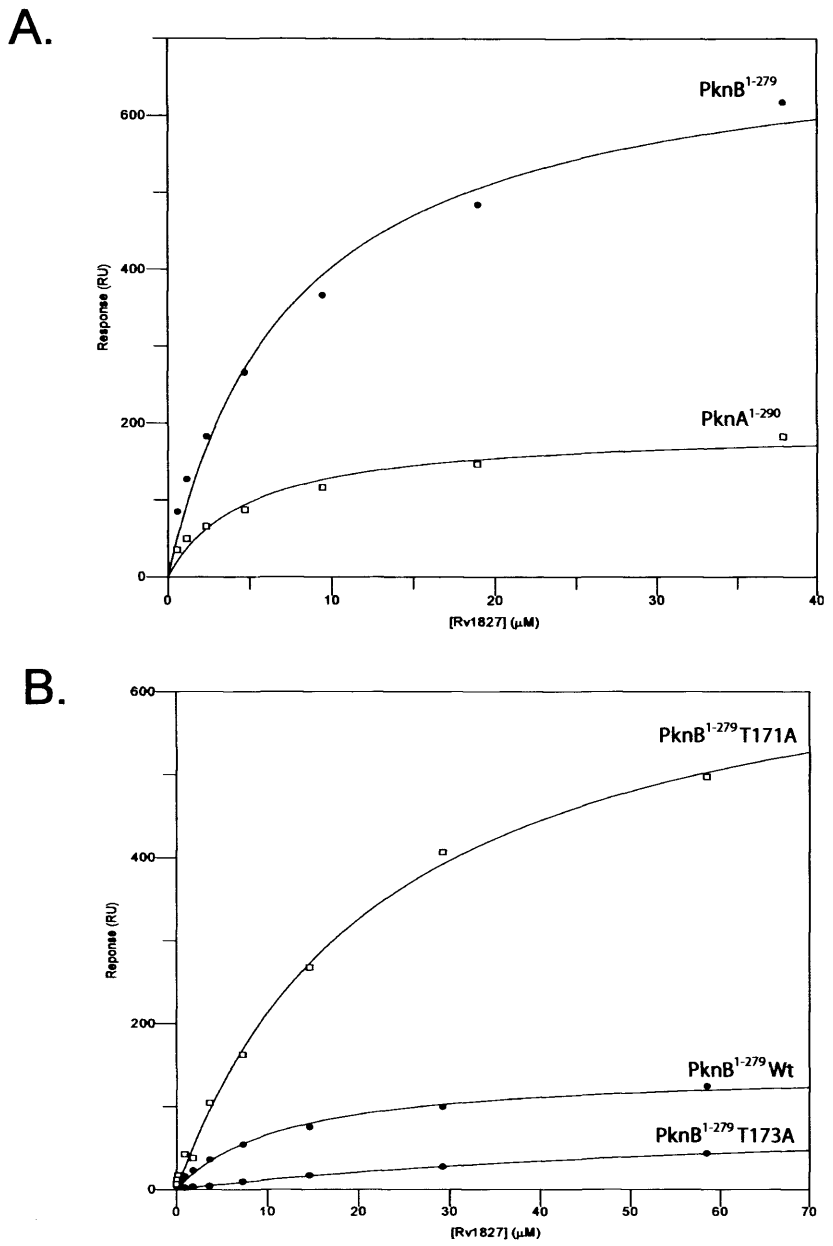


Figure 5.7. Equilibrium analysis of Rv1827 FHA domain with immobilised $PknA^{1-290}$ and $PknB^{1-279}$ conducted using surface plasmon resonance. Data from equilibrium experiments was analysed by fitting to a saturation binding curve using a one-site binding model by non-linear regressional analysis. (A) binding curves for the interaction of Rv1827 FHA domain to wild-type forms of $PknA^{1-290}$ and $PknB^{1-279}$. (B) binding curves for the interaction of Rv0020 FHA domain to wild-type and mutant $PknB^{1-279}$. All samples were injected over SPR surfaces in HBS buffer pH 7.4 at 25°C. Data were fitted to a single-site binding curve using non-linear regressional analysis and all fits were generated from the average of three independent experiments. All values were within 1 standard deviation of the average data point. The measured interaction parameters are shown in table 12.

Although not investigated in detail, it seems likely that the FHA domain containing proteins bind to PknA through the activation loop, which is similar in sequence to that of PknB (Figure 1.11). The binding affinities of the FHA domains from Rv0020 and Rv1827 to wild type PknA are of similar affinity to their interaction with PknB. Importantly, the reduced phosphorylation of full-length Rv0020 and Rv1827 proteins by PknA relative to PknB, as reported in the next Chapter (Section 6.2.3), cannot be attributed to lower interaction affinity of the FHA domains.

5.3.5 Kinetics parameters of FHA-Kinase interactions cannot be determined by SPR

The kinetic rate constants of the binding between PknA and PknB to Rv0020 FHA domain and Rv1827 FHA domain were measured by injecting a series of concentrations of the FHA domains over the immobilised kinase domains at a flow-rate of 50μL/min, significantly faster than for equilibrium analysis. Faster flow rates can reduce the amount of analyte rebinding during dissociation phase making dissociation rate constants easier to evaluate. During dissociation phase, the amount of available binding sites increases leading to rebinding of the analyte. Consequently, the apparent dissociation-rate is lower than the actual rate. Raw data on the response change upon binding (Figure 5.5) were then analysed using the BIAevaluation software. Here, the data acquired from several FHA concentrations were simultaneously fitted to a one-site model by a non-linear least-squares method and analysed according to the integrated rate equations described in Chapter 2. Using this procedure, only a few low FHA domain concentration experiments fitted to an acceptable chi-squared value; a statistical measure for the goodness of fit. From the resultant rate constants, the derived equilibrium constants were in the nanomolar range and did not agree with previous findings from SPR equilibrium experiments and ITC which both show binding affinities in the micromolar range (Table 5.3). By modelling these interactions, it was found that there was an error in fixing an accurate dissociation rate constant through the BIAevaluation global fit process. Therefore, the dissociation phase of the interaction was analysed separately to yield a more precise value for the dissociation rate constant. This was then used to constrain the analysis of the association phase to provide a better

estimate. Although, the data were much improved in terms of chi-squared value, this still failed to fit adequately implying that the rate constants are too fast to be accurately determined by SPR (S. Nicols, personal communication).

5.3.6 Competition studies

Although ITC measurements had suggested a binding preference of Rv0020 for pT173, equilibrium SPR experiments failed to show this difference. In order to eliminate the possible problems associated with the behaviour of differentially phosphorylated PknB proteins on the sensor surface, a series of competition studies were used to determine the binding specificity of Rv0020 and Rv1827 FHA domains to activation loop peptides. This has the important advantage that a single sensor chip surface modified with only wild-type PknB can be used for all the experiments, permitting a more reliable comparison of relative ligand affinities. In brief, the SPR competition experiments were conducted using immobilised PknB and passing a fixed concentration of FHA domain (1μM for Rv0020 and 5μM for Rv1827) over them spiked with varying concentrations of synthetic peptide. As with equilibrium SPR experiments, the steady-state response values (maximum response) were obtained at each peptide concentration after subtracting the control value. This response value would equate to the maximal binding of the FHA domain (Rv0020/Rv1827) to the ligand (PknB Wt) at a specific concentration of competing peptide. Readings were taken from a flow cell containing immobilised wild-type PknB for at least 12 different peptide concentrations and IC_{50} values were calculated by fitting the response *versus* peptide inhibitor concentration to a dose response curve (Section 3.3).

Table 5.4. Equilibrium binding parameters for the interaction of FHA domains of Rv0020 and Rv1827 with PknB activation loop peptides

		*Activation Loop Peptides**			
		pTQT	TQT	TQpT	pTQpT
Rv0020	SPR IC_{50}	71	DNB	15	3.6
	SPR K_D[a]	20	ND	4.2	1.0
	ITC K_D	34	DNB	3.0	2.0
Rv1827	SPR IC_{50}	14	DNB	22	2.6
	SPR K_D[a]	9.6	ND	18	1.9
	ITC K_D	24	DNB	22	11

*All binding data is given in μM units.

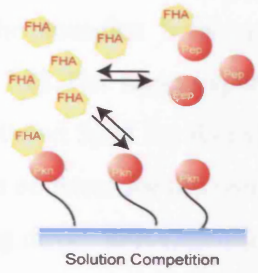
[a]SPR generated K_D values were determined using the following equation:

$$K_D^I = \frac{IC_{50}}{\left(1+\frac{[S]}{K_D}\right)}$$

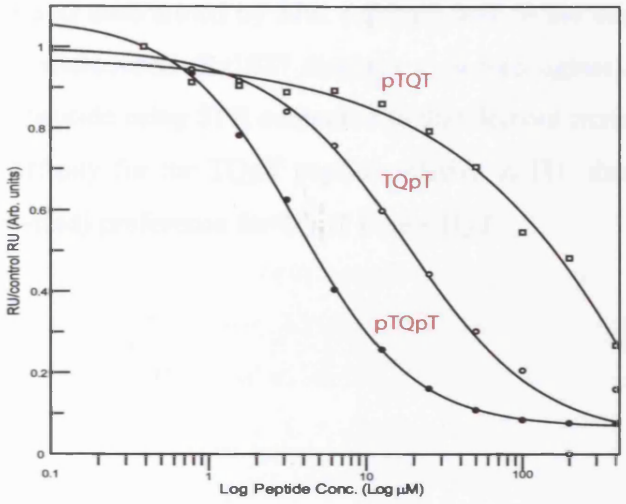
where K_D^I is the peptide equilibrium dissociation constant as determined by SPR competition assay, IC_{50} is the concentration of the peptide that inhibits binding of the FHA domain to the surface immobilised kinase by 50%, K_D is the equilibrium dissociation constant, as measured by SPR (Table 5.3) for the complex between the FHA and surface immobilised kinase and [S] is the concentration of FHA domain.

These experiments demonstrated that the Rv0020 FHA domain binding preference for the various peptide ranks as follows: pTQpT > TQpT > pTQT (Figure 5.8). The unphosphorylated activation loop peptide was not competitive indicating that it cannot bind to either FHA domain. Rv1827 was generally less able to bind the activation loop peptides than Rv0020, and showed no significant specificity for each site, which was consistent with the ITC data. Again, the FHA domain bound more tightly to the doubly phosphorylated peptide than either singly phosphorylated peptide by a factor of 10.

A.



B.



C.

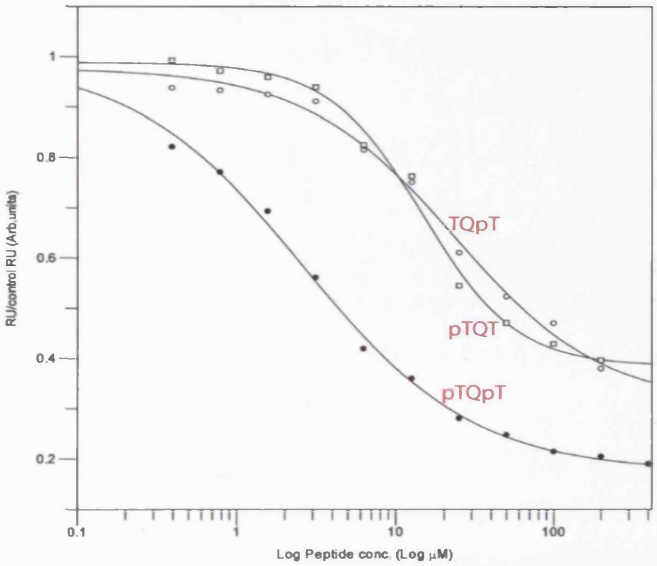


Figure 5.8. Competition binding studies of the interaction of Rv0020 FHA domain with PknB activation loop phospho-peptides. (A) illustration of surface competition interactions. (B) A fixed concentration of Rv0020 FHA domain (1µM) with various concentrations of peptide was injected (5µL/min) over immobilised wild-type PknB. (C) A fixed concentration of Rv1827 FHA domain (5µM) with various concentrations of peptide was injected (5µL/min) over immobilised wild-type PknB. All samples were injected over SPR surfaces in HBS buffer pH7.4 at 25°C. The ratio of RU with and without competing peptide was determined and plotted against the log peptide concentration. Data were fitted to a dose response curve using non-linear regressional analysis and IC_{50} values were determined (Table 13). All fits were generated from the average of three independent experiments and all values were within 1 standard deviation of the average data point.

The IC_{50} calculated for the titration of each peptide can be related to the equilibrium dissociation constant of the complex between the peptide and FHA domain using equation **(30)** (Section 3.2.7.2). For these experiments, the FHA domain concentration was fixed (1μM for Rv0020 and 5μM for Rv1827) and the value of K_D was taken from the equilibrium dissociation constant for the complex between wild-type PknB and FHA previously determined using direct SPR (2μM for Rv0020 and 9μM for Rv1827) (Table 5.3). The calculated equilibrium affinity constants for the binding of FHA domains to the various activation loop peptides are shown in Table 5.4. The peptide equilibrium binding constants as determined by SPR compare well to the direct ITC measurements of FHA-peptide interactions. Rv1827 displays a ~4-fold tighter affinity for the doubly phosphorylated peptide using SPR compared to that derived from ITC. Rv0020 shows a ~4-fold lower affinity for the TQpT peptide relative to ITC data but again displays a significant (~20-fold) preference for TQpT over pTQT.

5.4 Discussion

5.4.1 The specificity of activation loop binding by the Rv0020 and Rv1827 FHA domains

The results presented here provided the evidence for physical interactions between activation loop phospho-threonines of PknB and the FHA domains of Rv0020 and Rv1827. Both ITC and SPR show that mutation of the two threonines in the activation loop of PknB abrogates interaction with these FHA domains.

The FHA domains of Rv0020 and Rv1827 did show differences in their binding specificity towards differentially phosphorylated PknB constructs and activation loop peptides. In the ITC experiments, Rv0020 showed a preference for binding to the activation loops phosphorylated at solely at T173 or both T171 and T173 over T171 (pT173 $\approx$ p171 & p173 > pT171), and this result was supported by SPR. SPR competition studies showed peptides TQpT and pTQpT competed for the binding to the Rv0020 FHA domain with PknB (wild-type) more effectively than the pTQT peptide and therefore bound more tightly to the FHA domain. An earlier series of ITC experiments (Westcott, S., PhD thesis 2003) investigated the binding of the Rv0020 FHA domain to various peptide sequences. The results of these experiments suggested that the Rv0020 FHA domain prefers a medium sized hydrophobic residue (i.e. Leu, Ile or Val) in the +3 position relative to a phospho-threonine. The high-resolution X-ray structure of the Rv0020 FHA domain-optimal peptide complex shows that this hydrophobic residue is accommodated in a hydrophobic pocket at the peptide-binding surface (Westcott, S., PhD thesis 2003, and Section 1.6.5).

These ITC experiments also find the Rv1827 FHA domain demonstrates little preference for a particular activation loop phosphorylation state. Previous peptide-library selection data for the Rv1827 FHA domain that suggests it favoured bulkier hydrophobic residues (i.e. Tyr or Met) in the +3 position (Durocher *et al.*, 2000). These residues are not present in the +3 position to either phospho-threonine in the PknB activation loop and therefore may account for the weaker interaction relative to Rv0020.

5.4.2 A comparison of 'optimal' phospho-peptide interactions with binding of the PknB activation loop

As discussed previously, the 'optimal' or tightest binding peptide sequences for Rv0020 and Rv1827 FHA domains are known from peptide library screening and are shown in Table 5.1. The interactions between both FHA domains and their optimal peptides are significantly stronger than their interaction with their potential *in vivo* partner PknB (35-fold reduction for Rv0020 and 45-fold reduction for Rv1827). This trend is also observed when comparing binding of these FHA domains to their optimal peptides and to the PknB activation loop peptides.

Whether or not Rv0020 or Rv1827 interact with PknA/B is governed by various factors including local concentration conditions brought about by subcellular localisation and temporal expression. Rv1827 may possibly interact with PknB within the *Mycobacterium* cell after Rv1827 was detected as a substrate for PknB phosphorylation in whole-cell protein extract (Villarino *et al.*, 2005). Rv0020 is also likely to interact with PknB given its genetic proximity and co-expression (Figure 1.9). For this reason, data described in this Chapter demonstrates that optimally selected phospho-motifs are not always favoured as binding partners to FHA domains *in vivo* and that the interacting phospho-peptide motif is likely to evolve particular affinities for their targets in order to fit the requirements for the dynamics of its signal transduction cascade. The use of in vitro peptide-library screens to determine 'optimal' binding peptides for a phospho-recognition domain rarely provide a precise biological motif. They do however facilitate studies into the molecular basis for specificity of interaction domains.

5.4.3 Stoichiometry of Kinase-FHA domain interactions

In all cases, ITC data for phospho-peptide interactions could be fitted to a standard one-site binding model and gave stoichiometries that were consistent with a 1:1 molar ratio. This suggests indicating that FHA domains have only one binding site. In contrast, Kinase-FHA domain interactions showed a variable binding stoichiometry between

0.6:1 and 0.9:1. This most likely indicates heterogeneity in the kinase domain protein samples and a resulting variable active fraction. In the case of PknB K40A, the *n* value, which specifies the number of binding sites, was extremely low (0.03) which indicated that only a small proportion of the sample contained phospho-threonine sites.

The fact that the FHA domains of Rv0020 and Rv1827 are able to bind to the wild-type PknB kinase domain with a stiochiometry of one implies that only one FHA domain can only bind to a single PknB activation loop despite the presence of two phosphothreonines. This is possibly because these two phosphothreonine residues are found within the space of three residues and that FHA-phosphopeptide interactions involve close contacts between with the carbon-backbone of the peptide around the phosphothreonine residue. Consequently, the FHA binding surface may perhaps engage two phosphothreonine residues. The SH2 domain of the APS adaptor protein is an example of interaction domain that is able to bind two phosphorylated residues per domain (Hu *et al.*, 2003).

The observed binding parameters observed from an ITC experiments describe the average of the true binding parameters for all the interactions that occur in the titration. As the activation loop of PknB contains two phosphothreonine binding sites and the pT173 site is favoured for binding by Rv0020 ~9-fold over the pT171 site (Table 5.2), the wild-type PknB-Rv0020 FHA domain titration should provide a weaker apparent K_D than that of the PknB T171A-Rv0020 FHA domain titration. As the PknB-Rv0020 FHA domain interaction and PknB T171A-Rv0020 FHA domain interaction show a similar K_D the pT173 site seems to be favoured for binding over the pT171 in the fully phosphorylated PknB activation loop. This may possible be due to the fact that a pT in the –2 position (as would be the case upon pT173 binding) is accommodated by the Rv0020 FHA domains in favour of a pT in the +2 position (as would be the case upon pT171 binding).

5.4.4 The primary/secondary phosphorylation sites of PknB

Figure 5.1 shows the calorimetric titration of Rv0020 and Rv1827 FHA domains with wild-type PknB reveals the number of binding sites (*n*) is close to 1 (0.9), suggesting

that the wild-type PknB preparation is pure and contains a population of kinase domain molecules which are phosphorylated on at least one phospho-threonine. As the data for these titrations shows single site binding, the PknB wild-type sample is likely to contain protein wholly phosphorylated on one or both of the threonines in the activation loop. The values for *n* generated by the titration of Rv0020 and Rv1827 with PknB T173A are 0.9 and 0.8, respectively, which are similar to the wild-type PknB titrations (*n* of 1.3). Thus, the PknB T173A mutant appears to be wholly phosphorylated on T171. The values for *n* generated by the titration of Rv0020 and Rv1827 with PknB T171A are both ~0.65 suggesting that PknB T171A appears to be partially phosphorylated at T173 in the sample. Electrospray data shows PknB T173A is maximally phosphorylated on four sites whereas PknB T171A is maximally phosphorylated on three sites. The T171A mutation therefore seems to have a greater impact on autophosphorylation activity than the T173A mutation.

Taken together these data suggest that T171 is the primary site and T173 is the secondary site for PknB activation loop autophosphorylation. This assumes that PknB has a general preference for threonine over serine as the phosphoacceptor as proposed by evidence from Cantley and colleagues PLSA (Yang *et al.*, 2005). PLSA is a technique used to determine preferred substrate motifs for kinases and is similar to peptide library screen used for phospho-recognition domains. This technique involves generating peptide libraries with a fixed central phospho-acceptor residue (Ser or Thr) and random amino acid sequences flanking this residue. Each degenerate peptide library was phosphorylated in the presence of [γ-^{32}P]ATP by PknA and PknB. The peptides preferred as substrates for the kinase had a highest incorporation of ^{32}P. PLSA provided another important piece of information by demonstrating the importance of a glutamine residue +1 relative to the phosphoacceptor residue for peptide phosphorylation (Yang *et al.*, 2005). This implies the glutamine residue +1 relative to T171 makes T171 the favoured phosphoacceptor in the activation loop and therefore the primary phosphorylation site. It is important to note that there are ~4000 TQ sites in the *M. tuberculosis* proteome suggesting additional factors are involved *in vivo* substrate selectivity.

The conclusion that T171 is the primary phosphorylation site is further supported by electrospray mass spectrometry data and therefore loss of T171 is predicted to have a greater impact on autophosphorylation activity than loss of T173 (Chapter 6).

Interestingly, if Rv0020 binds pT173 of PknB after autophosphorylation of both pT171 and pT173, as indicated by specificity data presented here, this could confer a mechanism to ensure that Rv0020 would only bind to the fully activated kinase domain.

In many RD kinases, the primary phosphorylation site plays an essential role in maintaining the kinase in an active conformation. Phosphorylation of the primary site allows the activation loop to refold optimally for substrate binding. The critical driver in this conformational change is the electrostatic interaction between the primary phosphorylation site and a basic pocket that pins the loop in an active conformation (Nolen *et al.*, 2004). The main conserved residue involved in this process is the arginine that immediately precedes the invariant aspartate in RD kinases, as would be expected this motif is present in the RD kinase PknB. However, the X-ray structure of active PknB implies a disordered activation loop. Others (Boitel *et al.*, 2003) have investigated the surface charge distribution of PknB and highlighted the fact that the area near the active site contains an exposed cluster of basic residues that are favourably positioned for anchoring the phospho-threonines of the activation loop to a fixed position on the domain surface. Data presented here indicates this cannot be the case, as this would preclude the binding of FHA domains to the activation loop.

5.4.5 Conclusion

The observation that FHA domains bind to activation loop residues of the PknB kinase domain may illustrate a mechanism by which the disordered loop becomes more rigid leading to stabilisation of the kinase in an active conformation as proposed in a recent study (Duran *et al.*, 2005). In addition, the presence of the FHA domain may serve to protect the kinase domain for down-regulation by preventing dephosphorylation of the activation by protein phosphatase Ppp. Another potential function of a physical interaction between activation loop phospho-threonines and the FHA domains of Rv0020 and Rv1827 could be to facilitate phosphorylation of FHA domain-containing proteins by PknA and PknB. A series of *in vitro* kinase assays were carried out to shed light on this matter and are described in the following Chapter.

6 Kinase Activity of PknA and PknB

6.1 Introduction

The wild-type kinase domain proteins $PknA^{1-290}$ and $PknB^{1-279}$ were found to be autophosphorylated upon purification (Section 4.3 and 4.4, Table 4.1 and Figure 4.7). The phosphorylation of PknB has been mapped to four residues in the activation loop (S166, S169, T171, T173) (Young *et al.*, 2003) and the presence of four phospho-residues has been confirmed in wild type PknB by electrospray mass spectroscopy (Section 4.4). In the RD class of protein kinases, which includes PknA and PknB, the presence of phosphorylated residues in the activation loop is characteristic of the activated form thus providing an opportunity to identify and study the substrate preferences of these STPKs *in vitro*.

The results presented in Chapter 5 provided strong evidence for physical interactions between the kinase domain of PknB and the FHA domains of Rv0020 and Rv1827. In addition to establishing an interaction between the PknB kinase domain with FHA domains of Rv0020 and Rv1827, the site of interaction was identified as encompassing two phospho-threonines in the activation loop (T171 and T173). This demonstrates that autophosphorylation is not only required for kinase activity in PknB but also that it creates a binding site for FHA domain-containing proteins Rv0020 and Rv1827. I therefore investigated whether the kinase-FHA interaction might be associated with FHA domain-containing protein phosphorylation using two different *in vitro* phosphorylation assays: autoradiographic assays, for initial identification of phosphorylation activity, and where applicable, enzyme-coupled phosphorylation assays for kinetic measurement of the reactions. The results suggest that PknA and PknB do not only function in a simple linear pathway where they phosphorylate the co-expressed FHA-domain containing proteins, Rv0019 and Rv0020, but also play a role in a more complex signal transduction pathway involving FHA-domain containing proteins at other genetic loci.

6.2 PknA and PknB Phosphorylate FHA-domain Containing Proteins

6.2.1 Overview

To identify the substrates of the STPKs, PknA/B, full-length proteins Rv0020 and Rv1827 (from now on called $Rv0020^{155-527}$ and $Rv1827^{1-162}$) and the kinase domains of PknA, PknB were cloned, expressed and purified (Table 4.1). The actual translational start site of Rv0020 differs from the Genbank sequence (NP_214535) and was determined by N-terminal sequencing to be residue 155 (Weldingh *et al.*, 1998). The complete Rv0019 intracellular domain
($Rv0019^{28-156}$) was also purified but as a fusion protein with GST to circumvent poor stability associated with the $Rv0019^{28-156}$ (Section 4.2). The kinase domains of PknA and PknB were incubated at 2μM with 20μM $Rv0020^{1-527}$, $Rv1827^{1-162}$, GST-$Rv0019^{28-156}$ or a non-cognate substrate, MBP in the presence of [γ^{32}-P]ATP. The reaction products were separated using SDS-PAGE and identified by autoradiography. This procedure was also repeated by replacing GST-$Rv0019^{28-156}$, $Rv0020^{155-527}$ and $Rv1827^{1-162}$ with isolated FHA domains or serine-to-alanine mutants of these proteins ($Rv0020^{155-527}$ S473A and $Rv1827^{1-162}$ S85A) that result in complete loss of phosphothreonine-dependent binding (Durocher *et al.*, 2000). A summary of phosphorylation activity for PknA and PknB is provided in Table 6.1.

Table 6.1. PknA and PknB phosphorylate specific FHA domain containing proteins.

	Protein	PknA	PknB
Full-length protein	**Rv1827$^{1-162}$**	++	+++++
	Rv0020$^{155-527}$	+	++
	GST-Rv0019$^{28-156}$	+	+
FHA domain	**Rv1827$^{25-162}$**	-	-
	Rv0020$^{423-527}$	-	-
	Rv0019$^{56-156}$	-	+
Full- length S-A mutant	**Rv1827$^{423-527}$S95A**	-	-
	Rv0020$^{423-527}$S473A	-	-
Controls	**MBP**	++	++
	GST	-	-

(-) represents an experiment with no detectable phosphorylation

(+) represents a qualitative measure of the phosphorylation observed in the experiment with the number of + symbols representing the relative level of apparent phosphorylation determined by band intensity analysis

6.2.2 PknA and PknB phosphorylate FHA domain-containing proteins

PknA and PknB were each able to phosphorylate full-length FHA domain-containing proteins Rv1827$^{1-162}$ and Rv0020$^{155-527}$ (Figure 6.1 and Table 6.1, top panel). An interesting and somewhat surprising observation from these *in vitro* kinase assays is the marked preference of PknA and PknB for Rv1827$^{1-162}$ over Rv0020$^{155-527}$ as a substrate. Comparing band intensities, PknA and B show a ~20-fold and 500-fold higher preference for Rv1827 over Rv0020 respectively. These observations are contrary to the

expectation that co-expressed proteins such as Rv0020 would be a better substrate for PknA and PknB than Rv1827, which is expressed from a separate operon.

6.2.3 PknB is a more active kinase *in vitro* than PknA

There was significant difference in relative phosphorylation activity of PknA and PknB on both Rv0020 and Rv1827 (Figure 6.1a). As expected both STPK kinase domains were able to phosphorylate myelin basic protein (MBP) as a generic kinase substrate, with similar efficiency. Band intensity analysis of the autoradiograph determined using ImageQuant software (Kodak) shows that PknA has a ~10-fold lower activity than PknB on Rv0020 and ~200-fold lower activity on Rv1827. The result was surprising since a much higher autophosphorylation activity was evident for PknA relative to PknB and both are equally effective at phosphorylating the generic kinase substrate MBP.

6.2.4 FHA domains of Rv1827 and Rv0020 are not phosphorylated by PknA or PknB

The kinase activity of PknA and PknB towards isolated FHA domains of Rv0020 and Rv1827 was examined (Figure 6.2 and Table 6.1, middle panel). The N- and C-terminal limits of the FHA domains of $Rv0020^{155-527}$ and $Rv1827^{1-162}$ were isolated by limited proteolysis experiments (Section 4.2.3). The assay conditions remained the same as that of the previous assay (Figure 6.1) with the exception of the kinase concentration, which increased from 2μM to 5μM. Both PknA and PknB auto-phosphorylated but neither exhibited significant activity on FHA domains of Rv0020 or Rv1827. However, PknB did show catalytic activity on the minimal tryptic fragment of Rv0019. These observations strongly suggest that phosphorylation sites in full-length Rv0020 and Rv1827 lie in the amino-acid sequences N-terminal to the FHA domains (Residues 155-431 in Rv0020 and residues 1-24 in Rv1827; Figure 1.11).

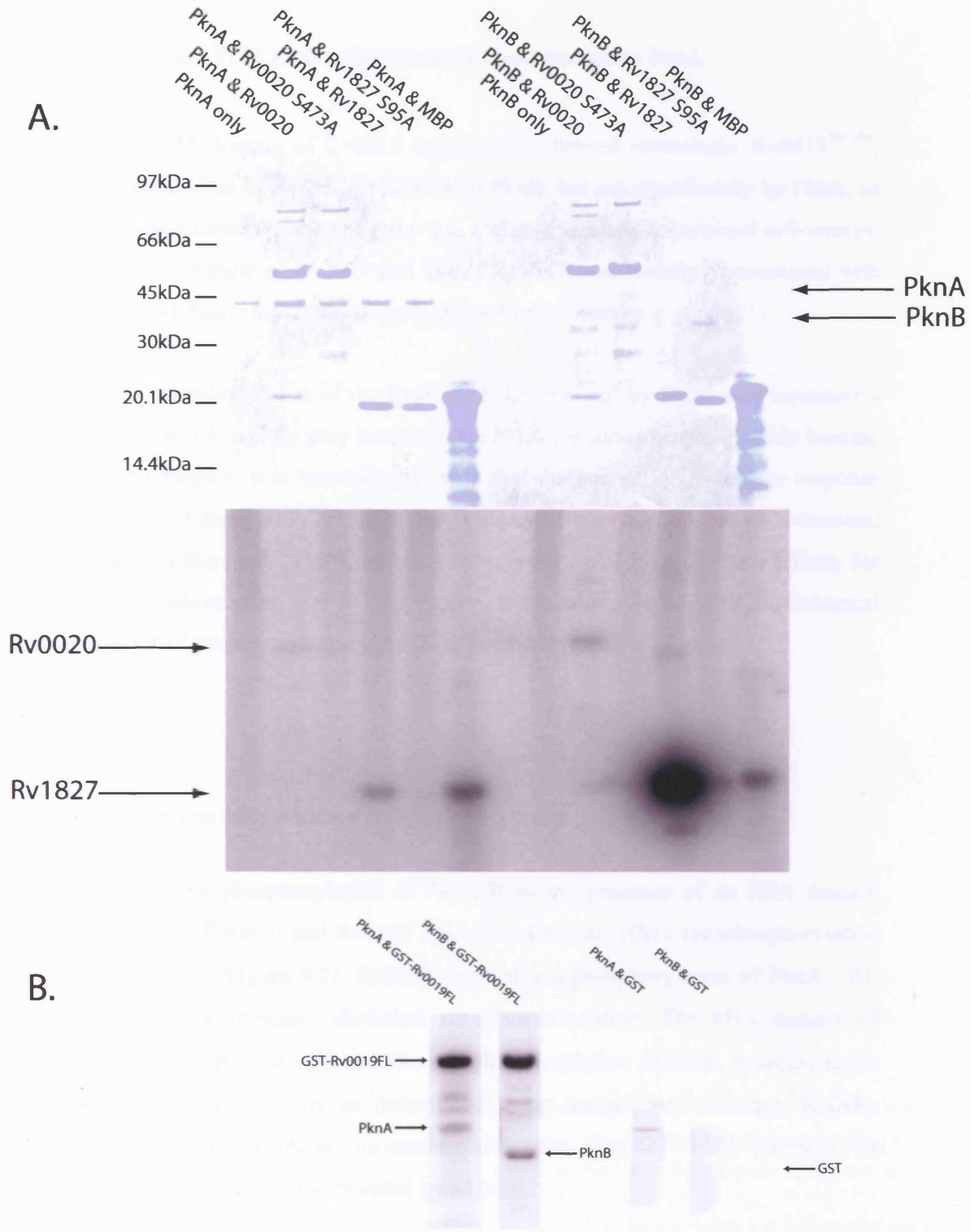


Figure 6.1. Specific phosphorylation activity of *M. tuberculosis* STPKs PknA and PknB on FHA domain-containing proteins. PknA$^{1-290}$ and PknB$^{1-279}$ were assayed in *in vitro* kinase reactions where each kinase was incubated at 2μM alone or with full length protein constructs of Rv1827 (Rv1827$^{1-162}$), Rv0020 (Rv0020$^{1-527}$) and myelin basic protein (panel A) and the intracellular region of Rv0019$^{1-156}$ (panel B), all at 20 μM. Reaction were carried out by incubating the proteins at room temperature for 1 hour in the presence of [γ-^{32}P]ATP. Coomassie staining was performed to verify the amount of protein in each reaction of panel A.

6.2.5 The Rv0019 FHA is phosphorylated by PknB but not by PknA

The minimal FHA domain of Rv0019 suggested by limited proteolysis, $Rv0019^{56-156}$, can be phosphorylated by the kinase domain of PknB, but not significantly by PknA, as observed by the autoradiograph in Figure 6.2, and confirmed by subsequent radioassays. Indeed, the FHA domain of Rv0020 and that of Rv1747 (incidentally co-expressed with PknF), were both found to be phosphorylated by PknF (Grunder *et al.*, 2005).

The efficient phosphorylation of the Rv0019 'FHA domain' by PknB may represent a mechanism by which STPKs may modulate the FHA domains phospho-peptide binding ability. In this respect, it is interesting to note that the human DNA-damage response kinase, Chk2 contains a STPK domain that is also able to phosphorylate its N-terminal FHA domain. In this case, phosphorylation reduces the Chk2 FHA domains affinity for a primary phosphorylation site on the protein (Ahn *et al.*, 2002). The physiological effects of Rv0019 phosphorylation are still to be investigated.

6.2.6 FHA binding affects kinase autophosphorylation

Examination of the phosphorylation of PknA/B in the presence of an FHA domain demonstrates both Rv0020 and Rv1827 FHA domains can affect autophosphorylation of PknA and PknB (Figure 6.2). Rv0020 reduced autophosphorylation of PknA ~20-fold and Rv1827 completely abolished autophosphorylation. The FHA domain of Rv0020 has a more modest effect on the autophosphorylation of PknB, reducing it ~3-fold (based on band intensity as determined using ImageQuant software, Kodak). Surprisingly, when PknB is incubated with the Rv1827 FHA domain its autophosphorylation ability is stimulated by ~10-fold.

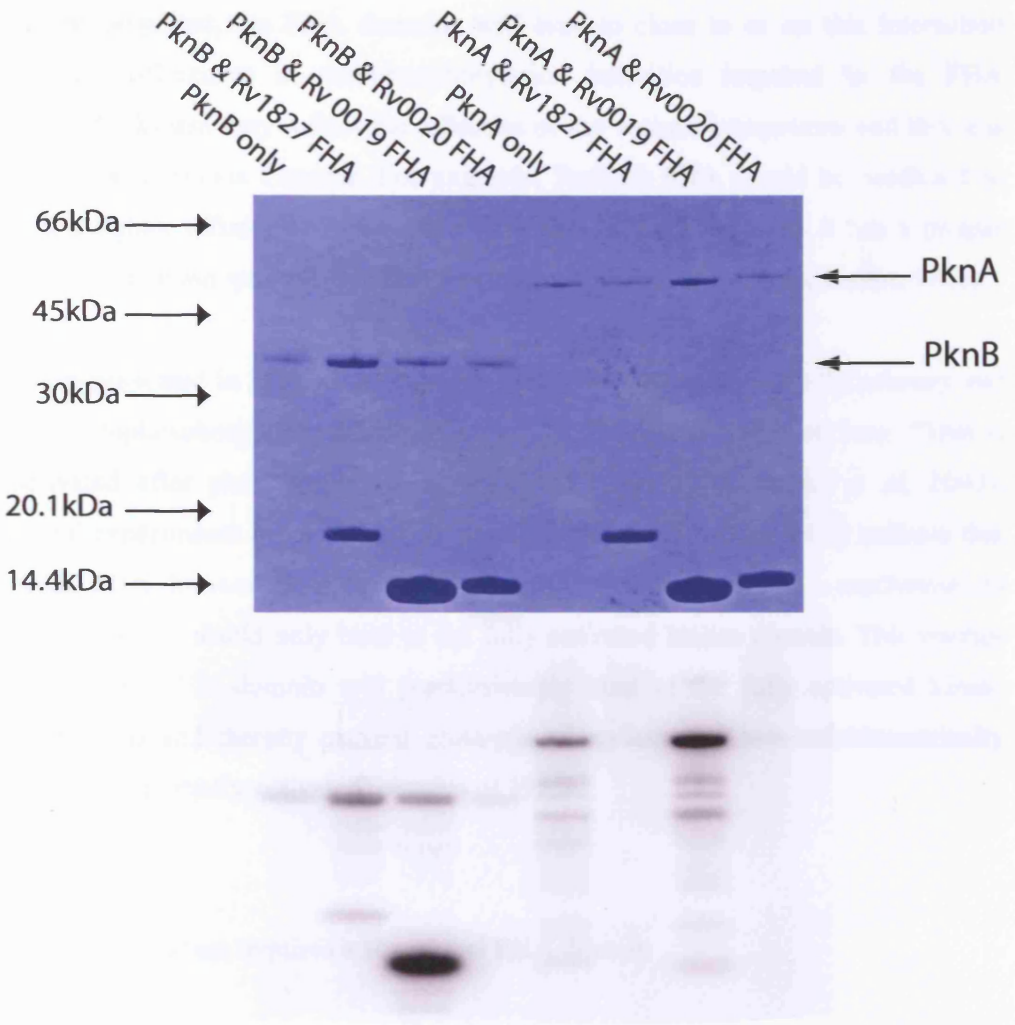


Figure 6.2. Specific phosphorylation activity of *M. tuberculosis* STPKs PknA and PknB on isolated FHA domains. PknA$^{1-290}$ and PknB$^{1-279}$ were assayed in *in vitro* kinase reactions where each kinase was incubated at 5 μM with the purified FHA domains of Rv1827 (Rv1827$^{25-162}$), Rv0020 (Rv0020$^{1-527}$) and Rv0019$^{1-156}$ all at 20 μM. Reaction were carried out by incubating the proteins at room temperature for 1 hour in the presence of [γ-^{32}P]ATP. Coomassie staining was performed to verify the amount of protein in each reaction (upper panel).

The inhibitory effect of FHA domains on autophosphorylation suggests that the domains may bind to the kinase domain in a manner that prevents the intermolecular interaction between kinase domains that is required for autophosphorylation. It can therefore be proposed, the FHA domains will bind to close to or on this interaction surface. The differences in autophosphorylation inhibition imposed by the FHA domains on the kinase may reflect the affinities of the various interactions and this was analysed in the previous Chapter. For example, Rv0020 FHA would be predicted to bind with a tighter affinity to PknA and PknB than Rv1827, because it has a greater inhibitory effect on autophosphorylation and this was shown to be case (Section 5.4.1).

Information presented in Section 5.4.4 shows that pT171 and pT173 are the primary and secondary autophosphorylation sites respectively, in the PknB activation loop. PknB is fully activated after phosphorylation of both T171 and T173 (Biotel *et al*, 2003). Biophysical experiments described in the previous Chapter (Section 5.4.1) indicate that the Rv0020 FHA domain binds preferentially to pT173 which confers a mechanism to ensure that Rv0020 would only bind to the fully activated kinase domain. This implies that the Rv0020 FHA domain will predominately bind to the fully activated kinase domain of PknB and thereby prevent cross-phosphorylation of sub-stoichiometrically phosphorylated (partially activated) species of PknB.

6.2.7 Phosphorylation requires a functional FHA domain

One key finding from the experiment shown in Figure 6.1 was that PknA and PknB were able to phosphorylate $Rv1827^{1-162}$ and $Rv0020^{155-527}$ but no significant phosphorylation of $Rv1827^{1-162}$ S95A or $Rv0020^{155-527}$ S473A mutant proteins was detected for either enzyme (Figure 6.1 and Table 6.1, top panel). Therefore, the mutation of the conserved serines, that are essential to phospho-ligand binding, leads to abrogation of phosphorylation in FHA-domain containing proteins, Rv0020 and Rv1827. Therefore, the fact that these mutant proteins are not phosphorylated may show the importance of phospho-threonine binding for phosphorylation of $Rv0020^{155-527}$ and $Rv1827^{1-162}$.

6.3 Rv1827 is phosphorylated on a conserved Threonine, Thr22

6.3.1 Overview

Phosphorylation mapping studies carried out by Alzari and colleagues, have indentified the site of Rv1827 phosphorylation by PknB to be T22 (Villarino *et al.*, 2005). Multiple sequence alignments of close homologs of Rv1827 (Figure 6.3) show that there is high conservation of not only the phospho-acceptor residue, T22, but all residues in the 20-ETTSVFRAD-28 motif. This observation implies suggests an important role in the physiological function of Rv1827, potentially by serving as a binding site for another FHA-domain containing protein or itself.

6.3.2 The effect of truncation on Rv1827 interactions

Limited proteolysis experiments were carried out to define the FHA domain boundaries in Rv1827. Regions of the protein that have less structure are susceptible to protease cleavage whereas regions that form tightly structured domains are less likely to be cleaved. Therefore, identification of the cleavage sites using limited proteolysis can infer domain boundaries. Limited proteolysis this resulted in a reduction in size from 162 residues to 138 residues (Westcott, S., unpublished data). These previous investigations also demonstrated that the effect of truncation on the Rv1827 FHA domains ability to bind its optimal peptide is negligible; with both protein constructs able bind the peptide with a K_D of ~200nM (Westcott, S., unpublished data).

To determine whether truncation of the protein would affect its ability to interact with the wild-type PknB kinase domain, an ITC experiment was performed at concentrations of 143μM FHA domain and 13.5μM kinase. The resulting thermodynamic and equilibrium data generated were close to that observed for the Rv1827 FHA domain (*n* 1.28, K_D 11.4 μM, ΔH -4.7kcal/mol, $T\Delta S$ -1.9kcal/mol ΔG -6.6 kcal/mol) and show that there was no effect of truncation on the ability of Rv1827 to bind PknB. This confirms that the FHA domain is sufficient for full binding to the kinase domain and isolated phosphopeptides.

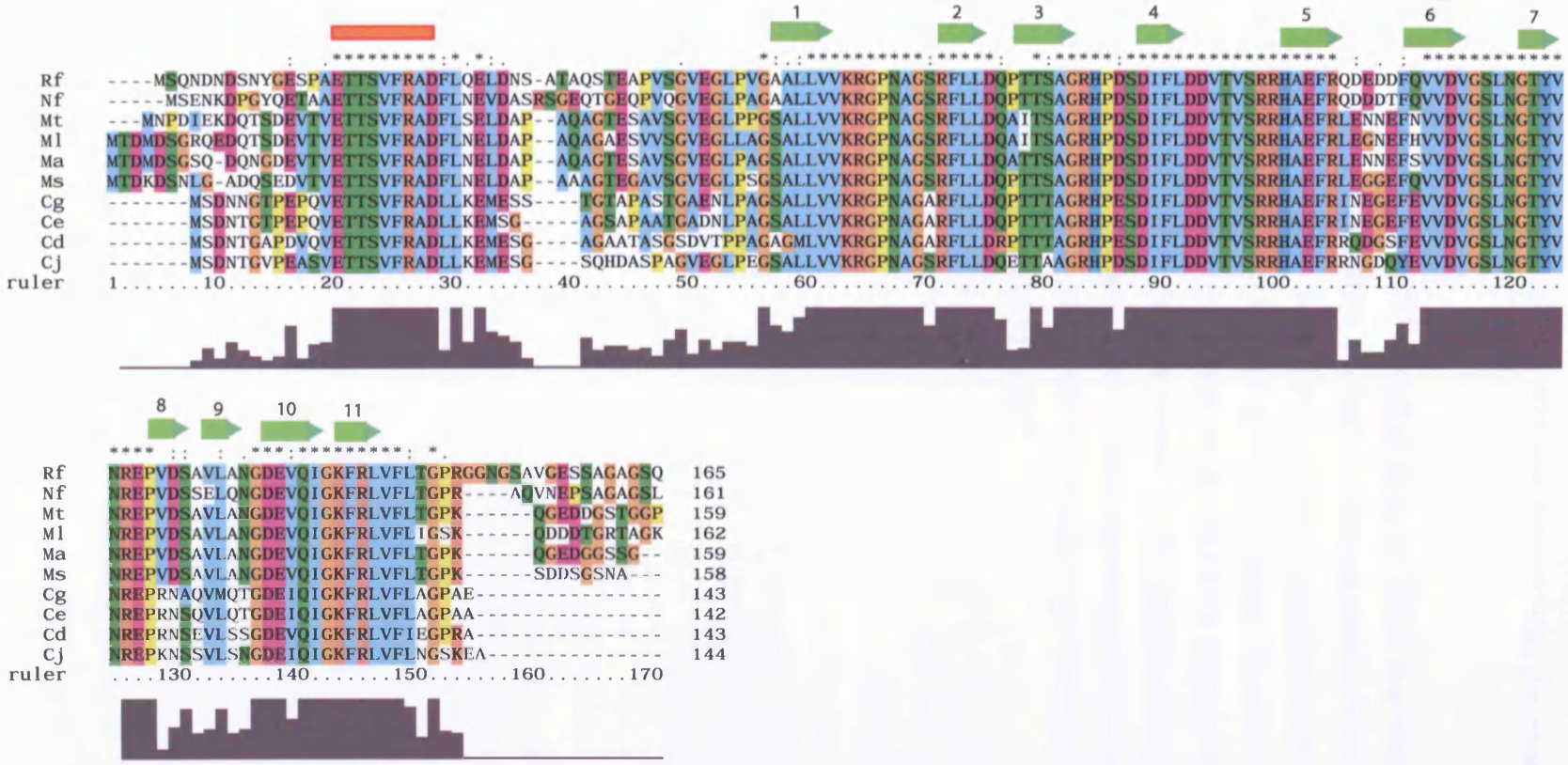


Figure 6.3. Multiple sequence alignment of close homologs of *M. tuberculosis* Rv1827 (GarA). Secondary structure elements (all β-sheets) were predicted from homology to Rv0020 and are shown above the sequence. Conserved residues in the N-terminus outside the FHA domain region containing the putative phosphorylation site, pThr22, of Rv1827 are indicated by a red bar above the sequence. Bars below the alignment show are a measure of the conservation at each residue position. (Mt, *M. tuberculosis*; Ma, *M. avium* sp. paratuberculosis; Ml, *M. leprae*; Ms, *M. smegmatis*; Rf, *Rhodococcus fascians*; Nf, *Nocardia farcinica*; Cg, *Corynebacterium glutamicum*; Ce, *C. efficiens*; Cd, *C. diptheriae,* Cj, C. *jejuni)*

6.3.3 Rv0020 and Rv1827 can interact with the phospho-threonine 22 region of Rv1827

The site of Rv1827 phosphorylation by PknB was mapped to T22 (Villarino *et al.*, 2005) which resides in a highly conserved amino-acid sequence (Figure 6.3). The fact that this N-terminal region is conserved suggests this motif plays an important role in the physiological function of Rv1827. Indeed, this sequence seemed a likely candidate for a dimeric interaction of the Rv1827 FHA domain due to a bulky aromatic group, Phe, at the +3 position relative to the phospho-acceptor residue, T22. In order to investigate this possibility, a phosphorylated peptide representing the phospho-threonine 22 region of Rv1827 with the sequence DEVTVETpTSVFRAYKK was synthesised and binding measured by ITC.

Table 6.2. Thermodynamic parameters for FHA:Rv1827 N-terminal peptide interactions.

Material in syringe[a]	Protein in cell	*n*	K_D (μM)	ΔH_b (kcal/mol)	$T\Delta S_b$ (kcal/mol)	ΔG_b* (kcal/mol)
TTSVF	Rv0020 FHA	NDB	-	-	-	-
TpTSVF		1.0	1.3	-7.5	0.3	-7.8
TTSVF	Rv1827 FHA	NDB	-	-	-	-
TpTSVF		0.9	7.8	-5.6	1.1	-6.7

* $\Delta G_b = \Delta H_b - T\Delta S_b$. All ITC experiments were conducted at 291K (18°C).

[a]Full peptide sequences are provided in Table 5.1. (NDB) No detectable binding. Stoichiometry (*n*), equilibrium binding constant (K_D), observed binding enthalpy (ΔH_b), observed binding entropy (ΔS_b), and binding free energy (ΔG_b)

This phosphopeptide shows detectable binding to the Rv1827 FHA domain with similar binding parameters to Rv1827-pTQpT peptide complex (Table 5.2). This indicates that Rv1827 may be able to form a head-to-tail dimer upon phosphorylation by PknB, assuming that this head-to-tail dimer conformation is sterically feasible. Furthermore, this peptide was found to interact with Rv0020 FHA 6-fold more tightly (K_D of 1.3μM) than the interaction with Rv1827 FHA (K_D of 7.8μM) using ITC.

Therefore, an Rv0020-Rv1827 heterodimer could be expected if T22 of Rv1827 was to be phosphorylated. However, it is important to note that the Rv1827-Rv1827 dimerisation affinity are likely to much greater than the individual affinities of the FHA domain for the TpTSVF site due to cooperativity effects. Similar cooperativity was seen for the Chk2 FHA domain-containing protein homodimer, where the overall dimerisation affinity was substantially greater than the affinity of the Chk2 FHA domain for its phospho-peptide target on the Chk2 protein.

6.4 Kinetic Characterisation of PknA and PknB

6.4.1 Overview

In light of the qualitative information accumulated on phosphorylation of FHA-domain containing proteins Rv1827 and Rv0020 by PknA and PknB, the kinetic parameters for each of these putative substrates was further investigated using an enzyme-coupled spectroscopic assay. The phosphorylation reaction rates were measured by coupling the formation of ADP to the oxidation of NADH using the enzymes pyruvate kinase and lactate dehydrogenase (Figure 3.2a). ATP turnover is stoichiometrically linked to the formation of NAD^+, which results in a decrease in absorbance at 340nm.

The assay described above is one of the most popular assays for measurement of the release of P_i, ADP or H^+. The method is particularly suitable for enzymes sensitive to the accumulation of ADP, as it is based on ADP recycling coupled to the oxidation of NADH. Furthermore, the high molar absorption coefficient of NADH at 340nm wavelength light and the virtual absence of protein or DNA absorbance at this wavelength ensures reliable ATP hydrolysis measurement down to a micromolar concentration scale (Kiianitsa *et al.*, 2003). The assay is valuable in enzyme kinetic studies because it is continuous, therefore, initial velocities are apparent immediately and do not require a laborious series of time points to be sampled as with radioisotopic methods (Roskoski, 1983). A limitation of the technique is that accuracy can drop off when examining enzymes with low turnover rates (i.e. <1 min^{-1}) as the ATPase activity may not be significantly higher than the background rate of NADH decomposition. Therefore, errors can arise from UV-induced NADH decomposition and, spectrophotometer-borne errors accounting for a low signal: noise ratio.

Although $Rv0020^{155-527}$ was not stable enough at high concentrations to measure its kinetic properties as a substrate in phosphorylation reactions, the assay was sensitive enough to measure phosphorylation of full-length Rv1827 ($Rv1827^{1-162}$) by PknA and PknB. The following Section will describe kinetic studies of PknA and PknB with the substrate $Rv1827^{1-162}$. Next, the K_m and V_{max} values were determined for various

activation loop mutations in PknB. In addition, the effects of activation loop threonine autophosphorylation on ATP binding and kinase reactivity were investigated. Two serines in the activation loop of PknB (S162 and S166 are sites of autophosphorylation, in addition to the two threonine residues (Figure 4.6b). However, a mutational analysis of these residues was not carried out since the Rv0020 and Rv1827 FHA domains are only able to bind phospho-threonine residues and S162/166 are known to be unable to activate PknB upon phosphorylation (Boitel *et al.*, 2003).

6.4.2 Kinetics of phosphorylation of Rv1827 by PknA and PknB

Traditional steady-state analyses of the kinase reactions were conducted with the goal of providing values for the *Michaelis constant* (K_m) and maximal reaction velocity (V_{max}) for the phosphorylation of Rv1827 by PknA and PknB.

Kinase assays for Rv1827 with PknA and PknB were carried out in a standard coupled reaction buffer including 2μM kinase, 1mM ATP and an Rv1827 protein concentration ranging from 870nM to 1mM (Section 3.6.2). Purified kinase domains of PknA ($PknA^{1-290}$) and PknB ($PknB^{1-279}$) were used in this experiment as well as purified full-length Rv1827 proteins ($Rv1827^{1-162}$, $Rv1827^{1-162}$ S95A). Reactions were only initiated after a stable A_{340} reading appeared. The initial velocity values were obtained by varying the concentration of Rv1827 under the fixed concentrations of ATP and PknA/PknB. Reading were taken for at least nine different Rv1827 concentrations and kinetic parameters were calculated by fitting the initial velocity *versus* substrate concentration to the Michaelis–Menten equation.

Table 6.3. Kinetic Parameters for the phosphorylation of Rv1827 by PknA and PknB.

Protein	K_m (μM)	V_{max}* (μmolATP.min^{-1})	k_{cat} (min^{-1})	k_{cat}/K_m (μM^{-1}.min^{-1})
PknB wild-type	143±11.4	0.0044±0.0001	22	0.15
PknB T171A	393±15	0.0025±9.6E-5	13	0.032
PknB T173A	352±9	0.0033±0.0002	17	0.047
PknB T171A/T173A	1900±0.07	0.0007±0.0004	3.5	0.0018
PknA	2070±0.9	0.0004±0.0001	2.0	0.00097

*V_{max} is defined in μmol ATP turnover.min^{-1} from 2μM kinase in a 100μL reaction.

These kinetic data support previous radiolabelling experiments which show significantly higher activity exhibited by PknB over PknA (Figure 6.4). The apparent K_m for the reaction catalysed by PknB was 143μM, ~14-fold better than the reaction catalysed by PknA (2mM). However, the activity of PknA was at the close to the limit of detectable phosphorylation using this technique, therefore the kinetic parameters for reaction involving PknA phosphorylation of Rv1827 should be taken with caution. There was also a concurrently high V_{max} displayed by PknB (0.0044μmol.min^{-1}) relative to the V_{max} of PknA (0.0004μmol.min^{-1}). Since the protein was used at subsaturating concentration, particularly in the case for PknA, these apparent kinetic constants should be taken with caution. Higher concentrations of Rv1827 could not be used in these reactions owing to aggregation of the protein at concentrations above 1mM. In the case of PknB the Michaelis–Menten catalysis curve was well defined, particularly in the area around the K_m of the enzyme which indicates the apparent kinetic parameters can be considered close to the true kinetic parameters. Finally, there was no detectable signal above background NADH decomposition for the reaction containing Rv1827$^{1-162}$ S95A.

The possible reasons for the low activity of PknA have been discussed in the previous Section. The steady-state kinetic parameters for the phosphorylation of Rv1827 by PknB was also determined in a more recently study (Villarino *et al*, 2005), the V_{max} found to be 0.00024μmol.min^{-1}, 18-fold lower than the value

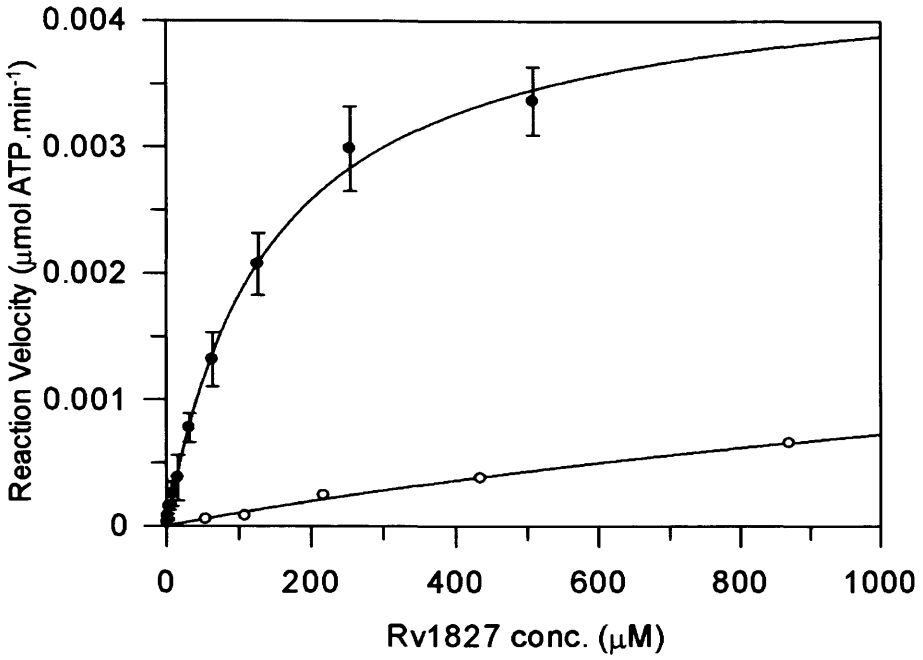


Figure 6.4. Apparent K_m and V_{max} determination for the phosphorylation of $Rv1827^{1-162}$ by $PknA^{1-290}$ and $PknB^{1-279}$ using an *in vitro* enzyme coupled kinase activity assay. The kinetic profiles of $PknA^{1-290}$ (open circles) and $PknB^{1-279}$ (solid circles) at increasing $Rv1827^{1-162}$ concentrations were fit to the *Michealis-Menten* equation using non-linear regressional analysis. Fits were generated from the average of three separate experiments.
Resultant reaction kinetic parameters are shown in Table 8.

determined in this study. However, these results were determined using a less rigorous study based on radiolabelling and using only three different concentrations of Rv1827. Additional disparity between the two results may arise from differences in active site purity of the PknB protein preparations. Catalytic activity may not correlate perfectly with enzyme activity as a consequence of misfolded protein or modified forms of the enzyme. Without conducting an active site titration, good active site homogeneity can be achieved by fractionation of the enzyme preparation on ion-exchange and/or gel filtration assuming misfolded and/or modified forms of the enzyme migrate differently to the active enzyme species (Lew, 2003). The following Section will show that PknB is only fully active upon phosphorylation of two threonines of the activation loop and electrospray mass spectrometry of PknB preparation show that the proportion of this phosphorylation state in the sample is variable. Thus, the PknB protein sample utilised by Villarino *et al* may be substoichiometrically T-loop phosphorylated. Unfortunately, no mass spec data were supplied in the publication to verify this.

6.4.3 Activation loop autophosphorylation and kinase activity

The effect of activation loop autophosphorylation on the activity of PknB phosphorylation of Rv1827 was explored. Previous work has shown that PknB is regulated by phosphorylation and dephosphorylation by the co-expressed phosphatase, *ppp*. This appears to be a general feature of regulatory circuits involving protein kinases genetically linked to protein phosphatases in bacteria such as *Bacillus subtilis* (Obuchowski *et al.*, 2000), *Streptococcus agalactiae* (Rajagopal *et al.*, 2003) and *Streptococcus pneumoniae* (Nováková *et al.*, 2005). Steady-state kinetic analyses of different activation loop mutants of PknB would provide a clearer picture of the effect of alteration in the presumed binding sites for the FHA domains on the FHA domains containing protein phosphorylation. Additionally, due to the potential effects of activation loop mutation on the overall kinase activity and activation, the effects of these mutations on ATP binding was also investigated (Section 6.4.5). Unfortunately, as the wild-type form of PknA displayed low kinase activity, close to the lowest detectable limit ($1 min^{-1}$), activation loop mutants were unlikely to exhibit measurable activity using the enzyme-coupled assay and consequently were not studied.

The phosphorylation of PknB has been mapped to four residues in the activation loop (S166, S169, T171, T173; Figure 4.7b.) and the presence of four phospho-residues has been confirmed in wild type PknB protein preparation by electrospray mass spectroscopy in this study (Table 4.1 and Figure 4.7). Purified activation loop mutants of PknB, T171A, T173A, T171A/T173A and the 'kinase dead' mutant K40A were generated by site-directed mutagenesis. The mutants were expressed and purified according the protocol for wild-type PknB and once the mutations were confirmed by DNA sequence analysis and electrospray mass spectrometry (Section 3.2.2,).

Kinetic analysis revealed that mutation of threonines to alanine in activation loop of PknB had a substantial effect on phosphorylation of Rv1827 (Figure 6.5 and Table 6.3). As expected fully phosphorylated wild-type PknB displayed the best K_m for Rv1827 (143μM), while both triply phosphorylated PknB mutants, PknB T171A and PknB T173A, showed a ~3-fold weaker K_m. The maximal velocity of wild type and mutants forms of PknB was determined and ranks as: wild-type>T173A>T171A>T171A/T173A>K40A. Wild type kinase had an apparent V_{max} of 0.0044μmol.min^{-1}, whereas no significant kinase activity was observed above background for the reaction using PknB T171A/T173A. This suggests a cumulative inhibitory effect of dephosphorylation of the activation loop. Note that the V_{max} values generated in this experiment cannot be considered true V_{max} values until the K_m^{ATP} is determined, to be sure ATP was used in saturating concentrations in the reactions. Reactions carried using a synthetic peptide encompassing the N-terminal region of Rv1827 produced no detectable phosphorylation.

As predicted, these data confirm requirement for phosphorylation on key threonine residues within the activation loop for full activity of PknB. Thr-to-Ala point mutations in this region significantly decrease the kinase activity towards

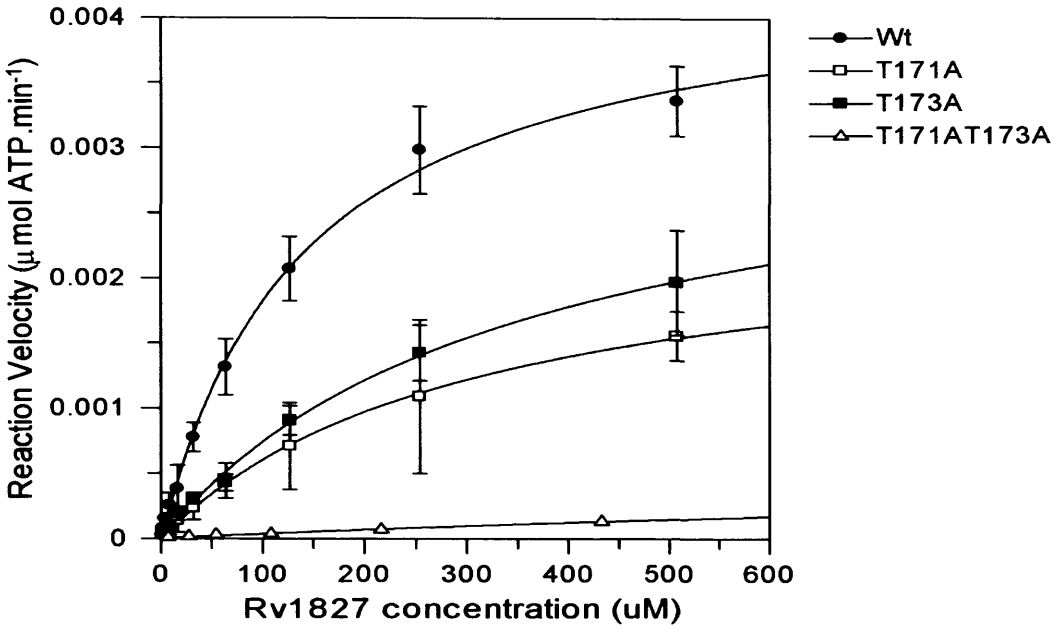
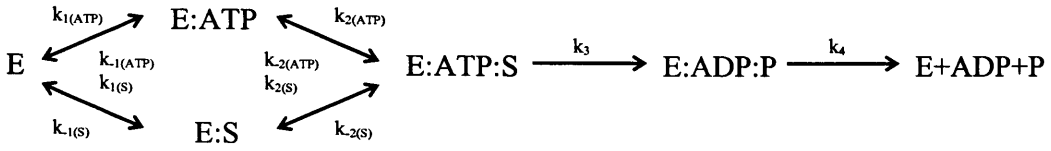


Figure 6.5. Apparent K_m and V_{max} determination for the phosphorylation of $Rv1827^{1-162}$ by (wild-type and activation loop mutant proteins) using an *in vitro* enzyme coupled kinase activity assay. The kinetic profiles of all $PknB^{1-279}$ proteins at increasing $Rv1827^{1-162}$ concentrations were fit to the *Michaelis-Menten* equation using non-linear regressional analysis. Fits were generated from the average of three separate experiments. Resultant reaction kinetic parameters are shown in Table 8.

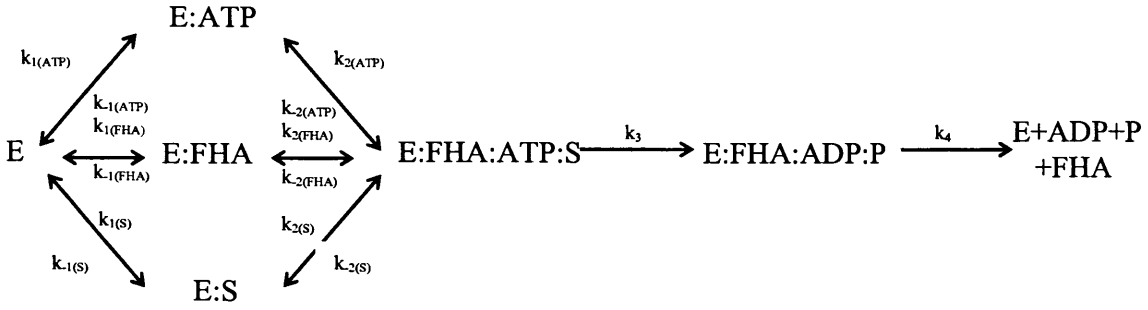
the endogenous substrate Rv1827 and may help to link the atomic resolution structure of PknB with its biochemical function. The two independent crystal structures of PknB in complex with nucleotide reveals disorder in the activation loop even though they depict an enzyme poised for phospho-transfer. Conformational disorder of flexible regions, such as loops or extended coils, which are mobile with respect to the rest of the protein, result in reduced and sometimes lost electron density, as is the case with PknB. Although many of the inactive structures exhibit disorder in these loops (Wybenga-Groot *et al.*, 2001)(Lie *et al.*, 2000)(Sicheri *et al.*, 1997), disorder only correlates with inactivity if it extends to the anchor points of the activation segment. In the case of the PknB structures, disorder seems only confined to the activation loop and P+1 loop allowing the kinase to maintain activity in a similar manner to that seen in the structures of kinases PDK1 (Biondi *et al.*, 2002) and *Src* kinase (Ogawa *et al.*, 2002). The observation that FHA domains bind to activation loop residues may illustrate a mechanism by which the disordered loop become more rigid leading to stabilisation of the kinase in an active conformation.

The interaction of the FHA domains of Rv0020/Rv1827 with the kinase domain of PknA/B and the requirement of kinase-FHA interactions for efficient phosphorylation of full-length Rv0020 and Rv1827 has been demonstrated. The only possible interacting residues are the phospho-threonines on the activation loop, mutations of these residues will not only disrupt the essential phospho-binding interactions, it will also reduce activity of the kinase itself. Scheme 1 shows a basic kinetic scheme for the steady-state phosphorylation of substrates catalysed by a protein kinase under saturating concentrations of a substrate (Lew, 2003). Taken together the information gathered so far from this study suggests a more complex scheme for the for phosphorylation of Rv1827 by PknB (Sheme 2).

Scheme 1:



Scheme 2:



The reactions described in scheme 1 and 2 include the association (k_2) and dissociation (k_{-2}) of both ATP, enzyme (E), product (P) and substrate (S), phosphoryl transfer (k_3), and product release (k_4). The reactions specific to scheme 2 are the association (k_1) and dissociation (k_{-1}) of FHA domain from the kinase: FHA complex and the association (k_1) and dissociation (k_{-1}) of FHA domain from the fully catalytic complex.

6.4.4 A substrate peptide of Rv1827 is not phosphorylated

The site of phosphorylation of Rv1827 by PknB was mapped to T22 residing in the N-terminal non-structured region of the protein (Figure 1.10) by (Villarino *et al.*, 2005). A peptide of the N-terminal sequence containing this phosphorylation site was synthesised with the sequence: 20-ETTSVFRAD-28. To separate the effect of FHA binding from the phosphorylation of Rv1827 by PknB an *in vitro* coupled-enzyme assay was conducted using this N-terminal Rv1827 peptide as the substrate for the reaction. There was no detectable phosphorylation of the peptide at the highest concentration used (7.5mM). This result is in line with previous finding that prove FHA domain of Rv1827 is required for efficient phosphorylation of the N-terminal region.

6.4.5 The effect of ATP concentration on the phosphoryl transfer reaction

To complete the kinetic characterisation of the PknB-Rv1827 phosphorylation reaction the K_m^{ATP} for wild-type and activation loop mutants of PknA and PknB was determined. This would provide more information about the role of autophosphorylation on the activation process of these enzymes. The K_m^{ATP} values are also required to determine whether optimal an ATP concentration were used in previous experiments on the kinetics of Rv1827$^{1-162}$ phosphorylation. Furthermore, the K_m^{ATP} values for wild-type PknA and PknB will be useful to find the optimal ATP concentration in a screen for ATP-competitive inhibitors.

Table 6.4. The effect of ATP concentration on the phosphorylation of PknB by Rv1827.

Protein	K_m^{ATP} (μM)	V_{max}* (μmol ATP.min^{-1})
PknB wild-type	13.6μM ±1.7	0.0011±4.68e-5
PknB T171A	651.4μM±138	0.0005±3.6e-5
PknB T173A	386μM±9	0.0007±5.5e-5
PknB T171A/T173A	1.12mM±0.35	0.0006±7.18e-6

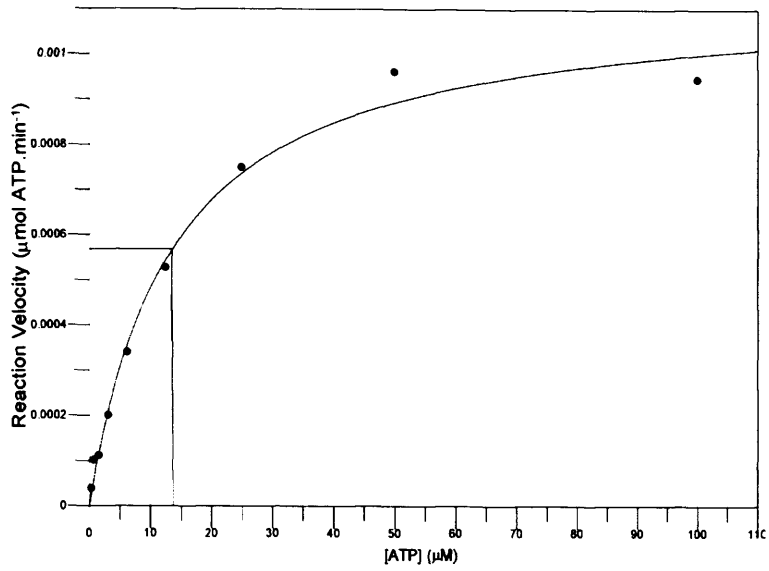
*V_{max} is defined in μmol ATP turnover.min^{-1} from 2μM kinase in a 100μL reaction.

The *in vitro* coupled-enzyme kinase assay was performed using a protocol based on the previous series of assay with the following differences: phosphorylation of Rv1827 was carried out in the standard coupled reaction buffer (Section 3.6.2) plus 2μM kinase, 40μM Rv1827$^{1-162}$ and an ATP concentration ranging from 390nM to 5mM. The initial velocity values were obtained by varying the concentration of ATP under the fixed concentrations of the substrate protein Rv1827. Reading were taken for at least nine different ATP concentrations and kinetic parameters were calculated following the previous protocol.

Kinetic analysis revealed that mutation of activation loop threonines had a substantial effect on catalysis, as expected (Figure 6.6 and Table 6.3). The fully phosphorylated wild-type PknB displayed the best K_m^{ATP} (13.6μM), while PknB T171A and PknB T173A mutant proteins showed 48- and 28-fold weaker K_m^{ATP}, respectively. The double mutant, PknB T171A/T173A, showed weakest K_m^{ATP} (1.12mM). The literature values for the K_m^{ATP} of several important eukaryotic protein kinases are as follows: PKA (12μM), PKC isoforms from tissue (15μM), recombinant PKCβII (70μM) and *Src* kinase (24μM) (Sun *et al.*, 2005). These are all in a similar range to the K_m^{ATP} of PknB. The maximal velocity of wild type PknB 0.0011μmol.min^{-1}, whereas all activation loop mutants display V_{max} values in the range of 0.0007-0.0005μmol.min^{-1}. However, the V_{max} values generated in this experiment cannot be considered true V_{max} as the values were not determined under saturating concentrations of a fixed substrate. As the K_m^{Rv1827} of PknB proteins (native and mutant forms) is known one can see that an Rv1827 concentration of 40μM is subsaturating. This assumes a saturating concentration of substrate would be defined as a 20-fold in excess of the K_m^{Rv1827} but millimolar concentrations of Rv1827 are not possible due to limited protein solubility.

As expected these data confirm that phosphorylation of the activation loop has an effect on PknB kinase activity. As discussed previously (Section 1.7) one of the principle roles of phosphorylation of the activation loop is to correctly configure conserved residues in the active site for catalysis and neutralise a region of high positive charge within the active site (Johnson *et al.*, 1996). Data presented here suggest that the remodelling of the PknB active site upon autophosphorylation is required for ATP association. In the absence of any FHA binding requirement for catalysis K_m is a function of all the steps in the phosphoryl transfer reaction pathway (scheme 2) where

A.



B.

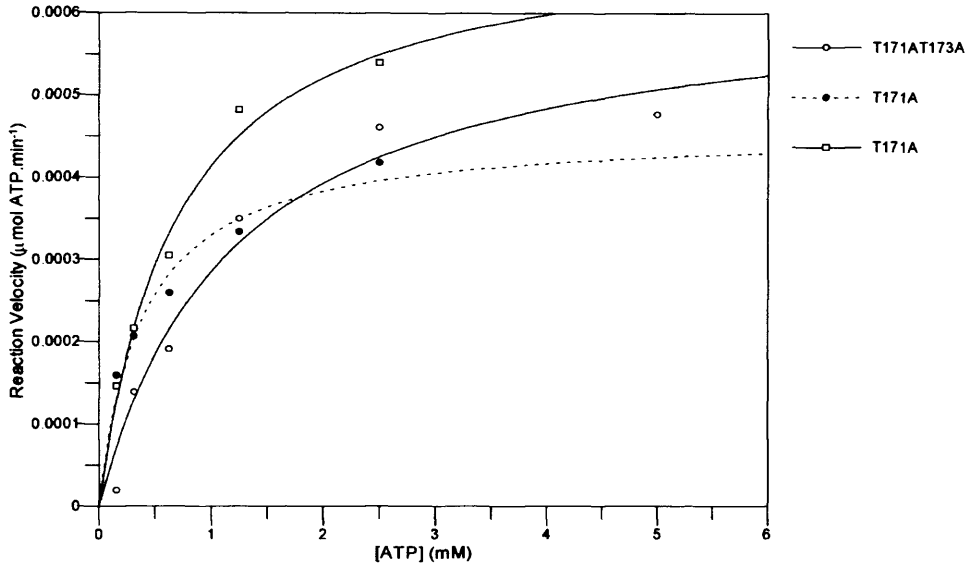


Figure 6.6. Apparent K_m^{ATP} and V_{max} determination for the phosphorylation of Rv1827$^{1-162}$ by PknB (wildtype and activation loop mutant proteins) using an *in vitro* enzyme coupled kinase activity assay. The kinetic profiles of all wild type (panel A) and mutant (panel B) PknB$^{1-279}$ proteins at increasing ATP concentrations were fit to the *Michealis-Menten* equation using non-linear regressional analysis. Fits were generated from the average of three separate experiments, all values were within 2 S.D. of the average data point. Resultant reaction kinetic parameters are shown in Table 9.

$$K_m^{ATP} = \frac{k_{2(ATP)} + k_3}{k_{1(ATP)}} \quad (34)$$

From scheme 2, the K_m^{ATP} is a composite of multiple catalytic reaction steps this parameter does not give direct information on any individual steps. However, if the rate of phosphoryl transfer is slow in relation to that of substrate release (k_{-2}), K_m approximates the ATP binding affinity (Lew *et al.*, 2003). This assumption cannot be made without information on the microscopic constants, particularly k_3. Therefore, the affinity of ATP for the PknB binding site must be directly measured, for example by using fluorescence spectroscopy to determine the binding of a fluorescent ATP analog to the PknB active site.

The K_m^{ATP} was used to determine if the optimal fixed ATP concentration required for subsequent reactions where Rv1827 would be varied. Ideally this concentration would be 20-fold higher than the K_m^{ATP} to ensure that over 90% of enzyme molecules in the reaction are bound to ATP at any one time, this would ensure the ATP association step is not rate-limiting. A concentration of 1mM was chosen as the preferred for assay. As a consequence the only true V_{max} determined from this study would be for the reaction of wild-type PknB with Rv1827, as wild-type PknB K_m^{ATP} was 13.5μM and thus reaction conditions of 1mM ATP would be saturating. Although the majority of the V_{max} values determined in this Chapter cannot be considered the true V_{max} values, they serve to measure relative activity of kinases tested under the same conditions.

6.5 Discussion

6.5.1 Molecular docking is required for FHA-domain containing protein phosphorylation

Purification of full-length protein and isolated FHA domains of Rv0019, Rv0020 and Rv1827 as well as the core catalytic domains of PknA and PknB enabled the design of a series of *in vitro* kinase assays to investigate the interrelationship between these proteins. These experiments were conducted after establishing an interaction between the PknA/B kinase domain and FHA domains of Rv0020 and Rv1827 and identifying the site of interaction on PknB as two phospho-threonines in the activation loop (T171 and T173) (Chapter 5). *In vitro* kinase assays demonstrate PknA and PknB STPK domains are able to phosphorylate the FHA domain-containing proteins. Critically, there was no phosphorylation of the isolated FHA domains of Rv0020, Rv1827 by either PknA or PknB, whereas Rv0019 was phosphorylated within the FHA domain. Hence, the phosphorylation site of Rv1827 and Rv0020 must reside outside the FHA domain. Taken together these data suggest a sequential mechanism for phosphorylation of these FHA domain-containing proteins which consists of initial binding of the FHA domain to the activation loop of the kinase domain via phospho-threonine interactions. This in turn allows the substrate motif to fit into its binding site on the STPK and become phosphorylated (Figure 6.7).

To test this theory, conserved serines of the FHA domains of Rv0020 and Rv1827, which enable the FHA-phospho-threonine interaction, were mutated to alanine residues. PknA/B could not phosphorylate these mutant proteins lending support for a molecular docking mechanism for phosphorylation as explained above. Alber and colleagues reproduced the findings for the interaction between PknB and Rv0020 in a recently published report (Grunder *et al.*, 2005). Indeed, a synthetic peptide of the N-terminal region purported to contain the site of phosphorylation of Rv1827 (Duran *et al.*, 2005) was synthesised but is not phosphorylated by PknB in enzyme-coupled assays, further illustrating the need for an active FHA domain.

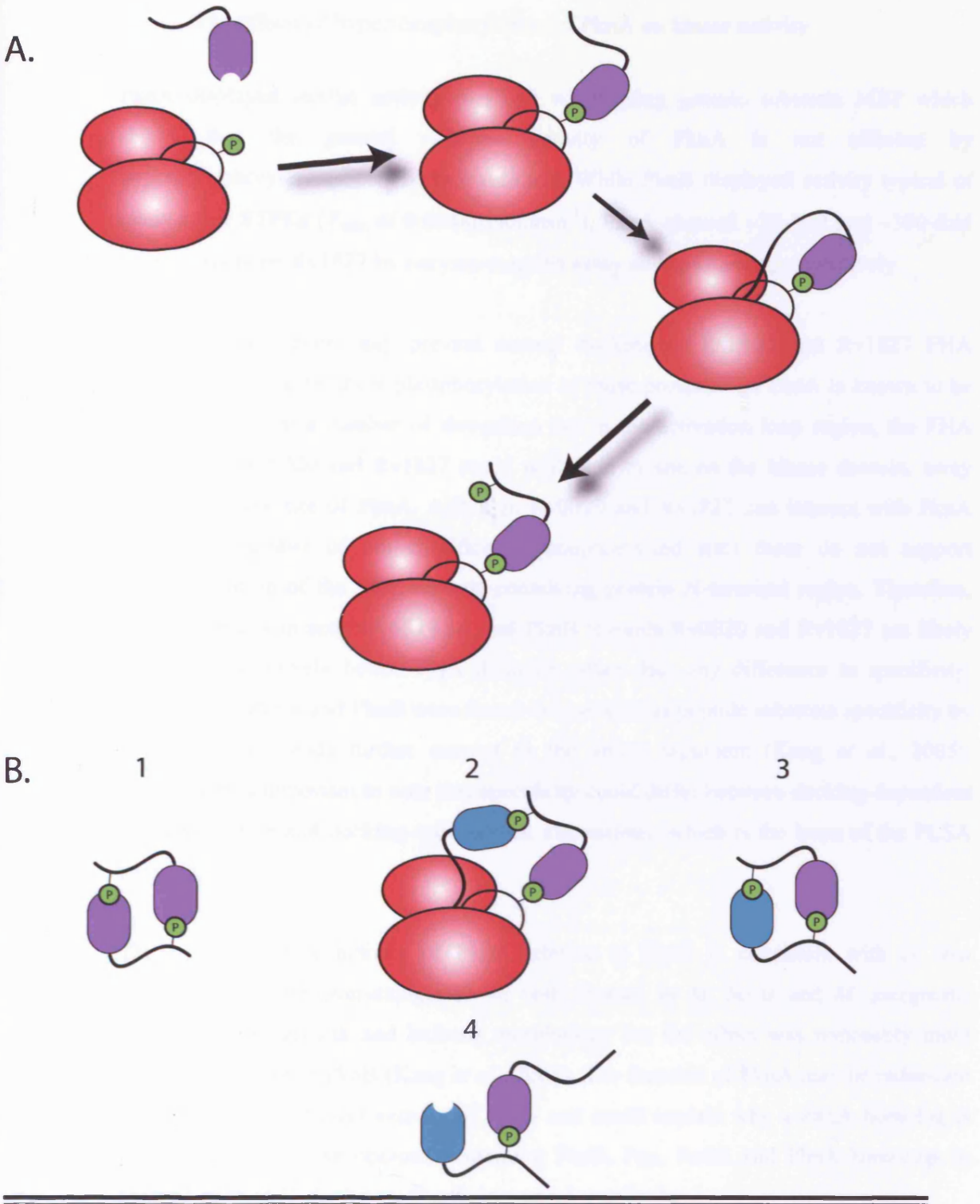


Figure 6.7. A possible mechanism for the interaction of FHA domain-containing proteins with the kinase domains of PknA and PknB. (A) After autophosphorylation of PknA/B at the activation loop, the FHA domain binds to the activation loop of the kinase domain via phospho-threonine interactions. This is followed by movement of the substrate motif (N-terminal portion of the protein) into the activate site of the kinase domain. This leads to phosphorylation of this region. (B) Possible consequences of FHA domain-containing protein phosphorylation. (1) The phosphorylated FHA domain-containing protein may form a homodimer (2) The phosphorylated FHA-domain containing protein may recruit further FHA domain-containing proteins whilst bound to the kinase domain (3) The phosphorylated FHA-domain containing protein may form a heterodimer with other phosphorylated FHA-domain containing proteins. (4) After phosphorylation of many molecules a 'daisy chain' effect may occur.

6.5.2 The effects of hyperphosphorylation of PknA on kinase activity

PknA displayed similar activity to PknB when using generic substrate MBP which shows that the general catalytic activity of PknA is not affected by hyperphosphorylation of its surface residues. While PknB displayed activity typical of many other STPKs (V_{max} of 0.0044μmol.min^{-1}), PknA showed ~20-fold and ~300-fold lower activity on Rv1827 by enzyme-coupled assay and radioassay, respectively.

Hyperphosphorylation may prevent correct docking of Rv0020 and Rv1827 FHA domains resulting in lower phosphorylation of these proteins. As PknA is known to be phosphorylated at a number of threonines not in the activation loop region, the FHA domains of Rv0020 and Rv1827 could well bind to site on the kinase domain, away from the active site of PknA. Although Rv0020 and Rv1827 can interact with PknA through a number of non-specifically phosphorylated sites these do not support phosphorylation of the FHA domain-containing protein N-terminal region. Therefore, any differences in activity of PknA and PknB towards Rv0020 and Rv1827 are likely due non-productively bound FHA domains rather than any difference in specificity. The fact that PknA and PknB were found to have similar peptide substrate specificity by PLSA analysis lends further support to the above argument (Kang *et al.*, 2005). However, it is important to note that specificity could differ between docking-dependent phosphorylation and docking-independent interactions, which is the basis of the PLSA analysis.

The reduced kinase activity of PknA relative to PknB is consistent with *in vivo* experiments where over-expression of both kinases in *M. bovis* and *M. smegmatis* caused reduced growth and bulbous morphology but the effect was noticeably more severe in the case of PknB (Kang *et al.*, 2005). The function of PknA may be redundant due to the more efficient activity of PknB and could explain why a PknA homolog is not present in putative operons containing PknB, Ppp, RodA and PbpA homologs in bacteria such as *C. diptheria*, *C. efficiens* and *S. coelicolor*.

6.5.3 The effects of activation loop dephosphorylation on STPK signalling in TB

Detectable PknB kinase activity on Rv1827 by the enzyme-coupled phosphorylation assay provided the opportunity to investigate the effect of autophosphorylation on kinase activity. Removing threonines from the activation loop by mutation to alanine has a cumulative effect on the reduction of kinase activity when using Rv1827 as a substrate. Previous studies by (Boitel *et al.*, 2003) and more recently (Duran *et al.*, 2005) have demonstrated that this same trend is apparent when using the generic STPK substrate MBP. However, using $Rv1827^{1-162}$ substrate in this study, the conditions inside the mycobacterial cell can be reproduced more accurately. For example the removal of phosphate groups from threonines of the activated PknB activation loop would not only result in generally lower kinase activity by means of a common regulatory strategy used by kinase of the RD class, but also remove the binding site for FHA domains. Consequently, the effect of coexpressed protein phosphatase; Ppp (dephosphorylation of activation loop residues) would be amplified. This could well be a conserved mechanism for regulation of all STPKs in *M. tuberculosis* as indicated by the conservation of these phospho-acceptor threonines in their activation loops (Figure 6.8).

The cross-phosphorylation ability of PknA and PknB was not tested in this set of experiments but has been published recently (Kang *et al.*, 2005). Native forms of PknA and PknB were shown to phosphorylated inactive mutants of PknA and PknB in a set of *in vitro* kinase reactions similar to that described in this Chapter. The existence of this intermolecular interaction implies there may be cross-talk between PknA and PknB during signal transduction involving these proteins.

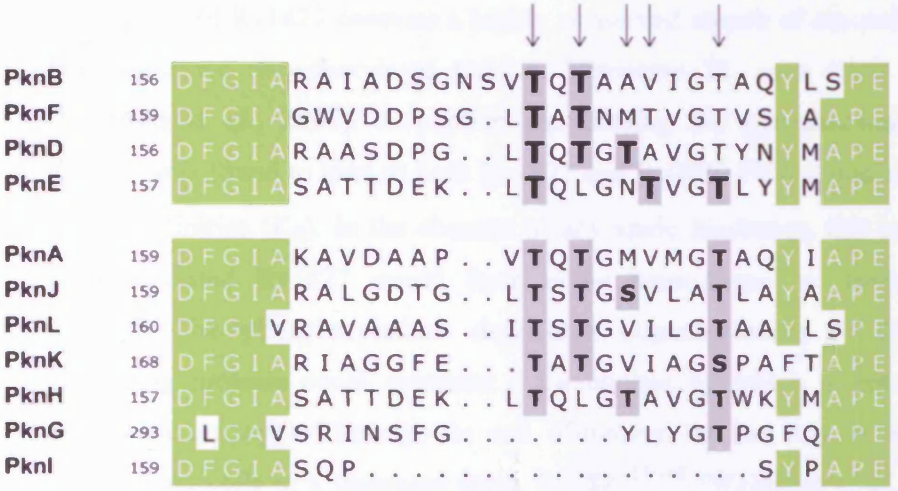


Figure 6.8. Amino-acid sequence alignment of activation loop residues in kinases domains of *M. tuberculosis*. The conserved residues located in the N- and C-terminal anchor regions of the activation loop are highlighted in green. The conservation of residues specific phospho-acceptor residues in the above sequences are indicated by arrows.

Adapted from Duran *et al.*, 2005.

Finally, ITC experiments used to investigate the binding of FHA domains to the phosphorylated motif created by phosphorylation of Rv1827 by PknB provide some tantalising insights into the nature of STPK signalling networks in mycobacteria. The N-terminus of Rv1827 contains a highly conserved stretch of six residues that contains one confirmed phosphorylated residue, Threonine 22, as a result of PknB activity (Villarino *et al.*, 2005). A peptide representing the phosphorylated N-terminal of Rv1827 was found to bind to both Rv1827 and Rv0020 FHA domains, with 7.8μM and 1.3μM affinities (K_D). In the absence of any steric hindrance, this is evidence that the phosphorylated Rv1827 would form either homodimers or hetero-oligomers with Rv0020. The phosphorylation- dependent oligomerisation of these FHA-domain containing proteins could represent a mechanism by which receptor kinase detected signals are propagated through the cell. Moreover, the fact that Rv1827 is found in *M. tuberculosis* cells in a truncated form, $Rv1827^{31-162}$ (Weldingh *et al.*, 1999) where the phosphorylation site has presumably been cleaved off suggests that these FHA domain propagated signals are regulated by separation of these dimers.

7 Inhibitor Studies

7.1 Introduction

Protein kinases have emerged as the second largest potential group of therapeutic targets after G-protein-coupled receptors (Cohen., 2002). Not only are kinases abundant in the human genome (Manning *et al.*, 2002), with 518 currently identified, but are also found in a subset of bacterial genomes, including those that cause disease in humans. The kinase active site can be divided into an Mg^{2+}-ATP complex binding site and a peptide/protein substrate binding site. Compounds that are able to bind either of these two sites with high affinity can be effective inhibitor of this enzyme (Al-obeidi *et al.*, 1998). Despite the large number of proteins in this enzyme class, a sizeable number of small-molecule inhibitors of protein kinases are undergoing human clinical trials, indicating that adequate specificity of inhibitors to targets can be achieved. This is in contrast to detractors of anti-kinase drugs in the early 1990s that argued that there would be difficulties in developing compounds to effectively compete with ATP binding due to overall sequence and structure similarity (Cohen., 2002).

The first inhibitor to be approved for clinical use that seemed to target a single protein kinase, was Rapamycin which inhibited mTOR, a kinase responsible for G1-to-S phase progression (Davies *et al.*, 2000). Rapamycin was important for stimulating interest in the development of other protein-kinase inhibitors. Since then, small molecule inhibitors Iressa (Astrazeneca compound ZD1839) and Tarceva (OSI Pharmaceuticals compound OSI-774) have been shown to specifically inhibit Epidermal Growth Factor Receptor (EGFR) tyrosine kinase, which is over-expressed in many cancers of epithelial origin such as lung and breast cancers, and displayed marked efficacy for a number of other cancers in human trials (Moyer., 2000). However, the most successful kinase inhibitor so far to be developed was Gleevec (Novartis compound STI-571), approved for clinical use in May 2001. It is an ATP competitive inhibitor of Abelson tyrosine kinase (ABL), which has enhanced activity in chronic myelogeneous leukaemia (CML). The structure of the ABL-Gleevec complex showed the drug to be bound in the catalytic

site whilst interacting with the N-terminus of the activation loop. This causes the complexed kinase to revert to an inactive conformation (Schindler., 2000).

In this Chapter, small molecule inhibitor studies on the serine/threonines kinase PknB, are described. It is hoped that a basic knowledge of the catalytic mechanisms of PknB combined with structural and biological information can be translated into a therapeutic strategy to treat tuberculosis. The vaccine for tuberculosis, BCG, is unable to confer significant protection in many parts of the world including India and parts of the USA (www.who.int). The current most effective strategy to treat tuberculosis sufferers involves a cocktail of four drugs taken for six months requiring direct observation by health care workers. Standard therapy for pulmonary TB includes isoniazid and rifampicin for 6 months, along with pyrazinamide and ethambutol for the first two months (isoniazid and rifampicin without pyrazinamide and ethambutol may be used for 9 months, if necessary). Occasionally used are streptomycin, thiacetazone, rifabutin and rifapentin. New drugs are required to shorten this regimen and to combat the spread of new multidrug resistant strains of *Mycobacterium tuberculosis*. The importance of new drugs in treating tuberculosis has been discussed in detail (Section 1.9)

PknB and PknA would appear to be excellent drug targets for the following reasons. Global transposon screens demonstrated PknA and PknB were required for bacterial growth (Sassetti et al., 2003) and attempts to isolate knockout mutants of these kinases have been unsuccessful (K.P. Papavinasasundaram, personal communication). Furthermore, in *M. smegmatis* depletion of both kinases by 50-80% using RNAi results in significant reduction in growth rates *in vivo* as well as producing morphological changes (Yang *et al.*, 2005). Interfering with these targets is likely to attenuate growth of *M. tuberculosis in vitro* and hopefully in *in vivo* infection.

Although the precise role of PknB in mycobacterial signalling is not yet fully understood it is clearly involved in response to its host environment. As shown in this thesis and other studies, PknB interacts with proteins encoded within its operon (Rv0020) and proteins encoded at distant genetic loci (Rv1827). Thus it is likely that there may be more than one final target of PknB, and interfering with its function may have multiple deleterious effects on the bacterium. Sequence analysis shows that PDK- and MARK-family kinases are its closest human homologues (~31% identity within the kinase domain). This contrasts with a significantly greater homology with, for example

M. tuberculosis paralogues (PknL (~46%), PknA and PknJ (44%), PknD (40%) and suggests that inhibitors of pknB may also show cross-specificity for other mycobacterial kinases, many of which have been implicated in growth and/or pathogenicity. Therefore, inhibitors able to simultaneously target a combination of these enzymes would have clear benefits in terms of overcoming problems of drug resistance. PknB inhibitors may also display cross-species specificity as the PknB kinase domains possesses significant homology to other important bacterial kinase domains including those of *Corynebacterium diphtheriae* (60% identity), *Listeria monocytogenes* (45%), *Bacillus anthracis* str. A2012 (42%), *Streptococcus pyogenes* MAS8232 (40%), *Staphylococcus aureus* (37%), *Streptococcus pneumoniae* (39%) and *Clostridium tetani* (43%) potentially making any inhibitor of PknB potentially a broad spectrum agent.

The following Chapter describes efforts to identify small-molecules that specifically inhibit PknB activity. Here, the justification for the preferred screening strategy will be outlined and in particular the nature of the assay and compound libraries used. Initial experiments for optimisation of a high-throughput screen will be discussed, followed by the standard operating procedure for the screen. The subsequent 'hit' compounds will be validated to confirm they are genuine inhibitors.

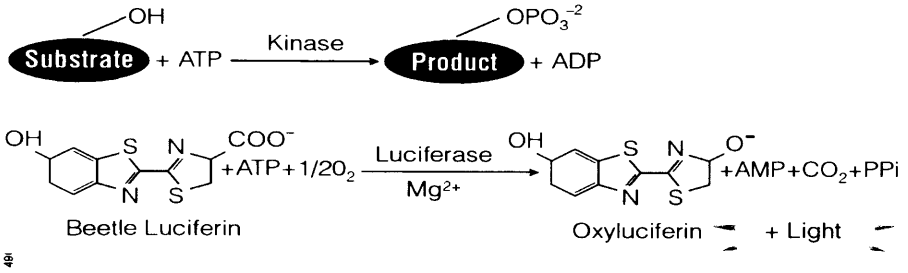
7.2 High-Throughput Inhibitor Screening

7.2.1 Overview

Protein-protein interactions between PknA and B with Rv0020c and Rv1827 (GarA) were found to be phospho-dependent and Rv1827 was found to be a particularly good substrate for PknB *in vitro* (Section 5.2.2). The interaction with Rv1827 has also been found in a global proteomics screen involving separation of PknB phosphorylated cell-free extract by 2D-gel electrophoresis. I have shown that this FHA domain-containing protein interacts with PknB and is highly phosphorylated by this kinase. In contrast generic substrates such as MBP are poorly phosphorylated. Therefore, we have chosen to use Rv1827 as a substrate for high-throughput screening.

In collaboration with the Drug Discovery Group (DDG), part of Medical Research Council Technology (MRCT), a strategy was devised for high-throughput screening of inhibitors to the phosphorylation of Rv1827 by PknB. The protein utilised for all experiments described in this Chapter were purified by me, whilst all experiments were carried out by Dr. Barbara Saxty and David Whalley. The basis of this screening strategy was the Kinase-Glo Luminescent Kinase assay developed by Promega. This method indirectly measures the turnover of ATP and is particularly suitable to high-throughput screening (HTS) for kinase inhibitors. The Kinase-Glo reagent contains a recombinant 'UltraGlow' luciferase enzyme that catalyses the mono-oxygenation of luciferin in the presence of Mg^{2+}, ATP and molecular oxygen to produce one photon of light per ATP turnover (Figure 7.1). The luminescent signal is correlated with the amount of ATP present and is therefore inversely correlated to the kinase activity which is, in turn an indirect measure of ATP consumption.

A.



B.

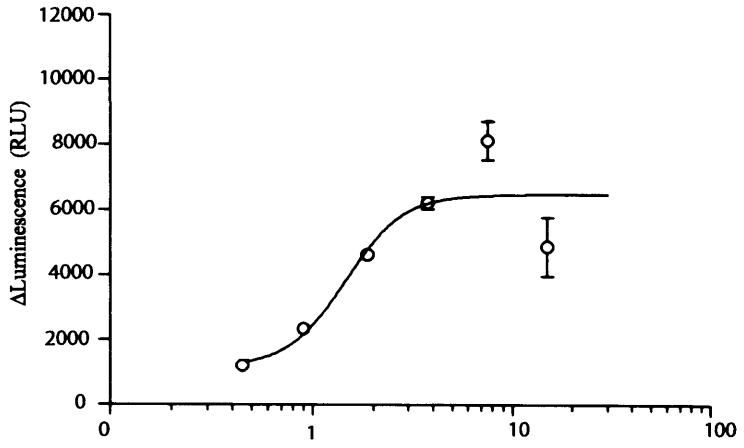


Figure 7.1. The Kinase-Glo kinase activity assay. (a) The mono-oxygenation of luciferin catalysed by UltraGlow luciferase (Promega) in the presence of Mg^{2+}, ATP and molecular oxygen produces one photon of light per turnover. (b) ATP titration curves for $PknA^{1-290}$ and $PknB^{1-279}$. 50μL reactions were made up with ATP concentrations between 0.4μM and 15μM were made with 1μM PknB and 10μM Rv1827, 1mM $MnCl_2$, 150mM NaCl, 50mM Tris-HCl pH 8.0. As a control, a repeat titration was performed in the absence of kinase. The incubation was performed at room temperature for a total of 10 minutes. The reactions were then stopped by the addition of an equal volume of Kinase-Glo reagent. The change in luminescence signal relative the no enzyme control (RLU) was plotted against Log ATP concentration. Fits were generated from the average of three separate experiments and EC_{50} values were determined.

Early kinase high-throughput assays measured the incorporation of radiolabelled ATP into a kinase substrate using a scintillation proximity assay (Amersham). However, more convenient, safe and cost-effective non-radioactive assays are generally preferred. There are a number of other non-radioactive kinase assays that can be utilised for high-throughput screening relying on generation of fluorescent or luminescent signals. These assays are designed to use phospho-peptide binding reagents, to detect the phosphorylated substrate in a quantitative manner, such as antibodies, metal-ion coated particles or fluorescent polymers. The particular drawback of such assays based on chelation of phosphates is that anionic compounds (i.e. ATP, phospholipids and carboxylate compounds) compete for binding to the chelator (Sun *et al.*, 2005). Despite this, antibody-based fluorescent kinase assays have been used to study a number of tyrosine kinases (Braunwalder *et al.*, 1996). Unfortunately, there are no antibodies with sufficient affinity to phospho-serine and phospho-threonine residues.

The Kinase-Glo luminescent kinase assay was chosen as it allows for batch-mode processing of multiple-plates due to the slow degradation of luminescent signal over time, and it has proven validity for HTS (Kupcho *et al.*, 2003). There are, however, problems associated with this assay. It has a narrow ATP dynamic range (<15μM) which is lower that that required for many kinases. In addition, various compounds are found to inhibit the activity of the luciferase enzyme (Sun *et al.*, 2005). Another important drawback is that the Kinase-Glo assay is an end-point assay as it can only measure ATP after the reaction has been stopped rather the continuously measuring ATP consumption. The expertise and necessary automated laboratory facilities of the DDG were important in initiating and performing the screening procedure. Apart from Kinase-Glo reagent kit and multiwell plates used, consumables required for this assay were minimal and the quantity of protein required for each reaction was relatively small owing to the high activity displayed by PknB on Rv1827.

7.2.2 Determination of optimal ATP concentration

The first step towards designing a high-throughput inhibitor screen the ATP concentration was optimisation of initial ATP concentration. To do this, the EC_{50} of the ATP was determined by performing a dose response titration of ATP (Section 3.2.9). In brief, various concentrations of ATP between 0.4 and 15μM were added to 1μM PknB, 1mM $MnCl_2$, 150mM NaCl, 50mM Tris.HCl pH 8.0. The incubation was performed at room temperature for a total of 10 minutes. The reactions were then stopped by the addition of an equal volume of Kinase-Glo reagent. As a control, a repeat titration was performed in the absence of kinase. Luminescence was read after a further 10minute incubation using a BMG Polarstar reader (luminescence mode, 1 second integration time) and the change in relative luminescence signal (RLU) compared to the no enzyme control. RLU was plotted against Log ATP concentration (Figure 7.1) and the preferred kinase substrate concentration was around the EC_{50}.

The EC_{50} for ATP in this assay was 1.5μM, which is significantly lower than the K_m^{ATP} determined from previous assays (13.5μM) (Section 5.3.5). However, this assay measures PknB autophosphorylation rather than Rv1827 phosphorylation, but it was decided all further optimisation work would be carried out at an ATP concentration of 1.5μM corresponding to 30pmol of ATP per reaction. A concentration of ATP below the K_m^{ATP} would be useful in a screen of ATP-competitive inhibitors. This decision was also governed by the fact that this concentration of ATP would give over 50% ATP turnover in a time period between 10-20 minutes, which has been shown previously to provide a good signal:noise ratio and a acceptable *Z'*-factor value. The *Z'*-factor value is describes the quality of the assay and is defined in Chapter 3 (Section 3.2.9).

7.2.3 Determination of optimal kinase and substrate concentration

The optimal concentrations of both kinase and substrate were determined simultaneously using a series of time course experiments (Section 3.2.9). The most convenient time for kinase reactions was around 20 minutes (B. Saxty, personal

communication), and it was important to find a concentration of substrate and kinase that would provide a linear ATP turnover over this time period and an acceptable *Z'* value for the assay. Reactions were performed containing Rv1827 (3, 5 and 7μM) with 0.05μM, 0.1μM and 0.2μM of kinase concentration plus 1.5μM ATP, 1mM $MnCl_2$, 150mM NaCl, 50mM Tris.HCl pH 8.0 in a total reaction volume of 100μL. The incubation was performed at room temperature for 20 minutes, and at the time points indicated 20μL aliquots were removed and mixed with an equal volume of Kinase-Glo reagent. As a control, a repeat titration was performed in the absence of kinase. Luminescence was read after a 10 minute incubation using a BMG Polarstar reader (luminescence mode, 1 second integration time) and the change in luminescence signal relative to the no enzyme control was recorded. The amount of ATP used as determined from the RLU was plotted against time for each concentration of Rv1827 (Figure 7.2).

The preferred substrate concentration should give a linear ATP turnover over the time period of 0-20 minutes. This optimum Rv1827 concentration was determined to be 5μM at a kinase concentration of 0.2μM, as the ATP turnover appeared most linear whilst providing an acceptable *Z'* value (0.67). This indicates the substrate and ATP was not saturating at these concentrations and, therefore, this concentration of protein gave a good indication of reaction velocity after 20 minutes with a high signal:noise ratio. In addition, this condition resulted in 66% turnover of ATP at the end of 20 minutes, which is required for improved *Z'* values (B. Saxty, personal communication).

If the kinase substrate, Rv1827, were to be used at 5μM per 20μL reaction then ~300μL of a stock solution of 500μM would be required for four 384-well plates (discounting void volumes of pipetting apparatus). This is an acceptable protein usage considering 5-8mL of 500μM protein stock solution can be purified from a 6 litre culture volume of protein-expression *E. coli*. With respect to PknB, a final concentration of 2μM per reaction translates as a ~15μL volume of 350μM stock kinase solution which is again acceptable considering 1mL of 350μM stock can be purified from a 40g pellet of protein-expression *E. coli*.

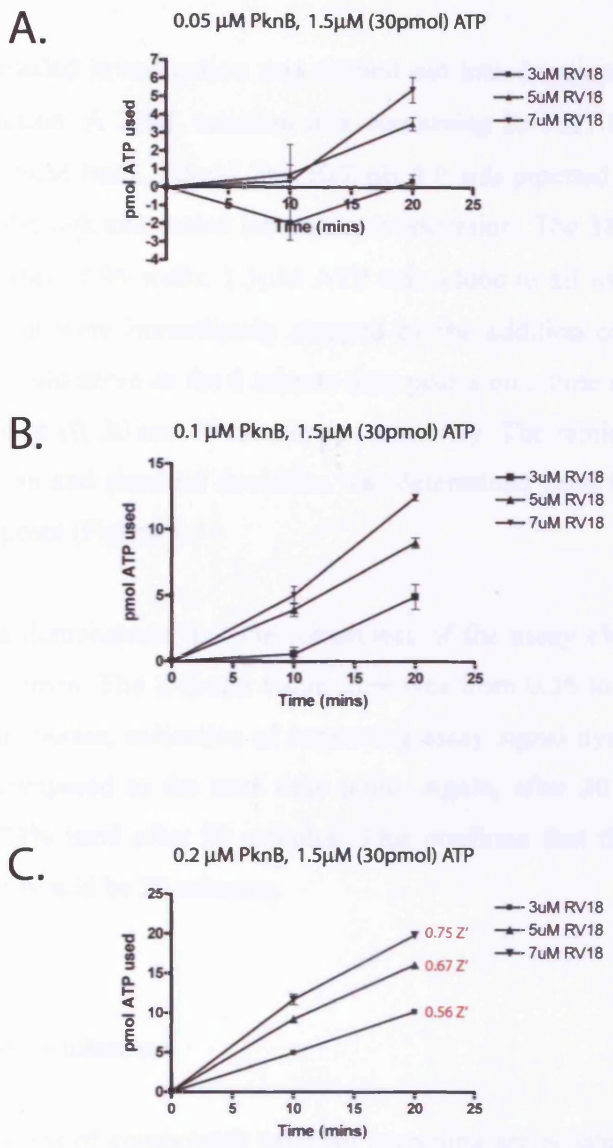


Figure 7.2. Determination of optimal kinase ($PknB^{1-279}$) and substrate ($Rv1827^{1-162}$) concentrations for high-throughput screening. Kinetic profiles of reactions containing a fixed concentration of ATP (1.5µM), three different concentrations of $Rv1827^{1-162}$ (3, 5, 7µM) and three different concentrations of $PknB^{1-279}$ (0.05, 0.1 and 0.2µM, Panels A, B and C, respectively). Reactions shown in panel C demonstrated a linear phase of the reaction over 20 minutes and the *Z'* values for these conditions were determined and are indicated in red.

7.2.4 Determination of optimal reaction time

A more detailed investigation was carried out into duration for which to carry out the kinase reaction. A 20μL reaction mix containing Rv1827 (5μM), PknB (0.2μ), 1mM $MnCl_2$, 150mM NaCl, 50mM Tris.HCl pH 8.0 was pipetted into all wells of a 384-well using the Biomek automated laboratory workstation. The 384-well plate is divided into four quadrants of 96-wells. 1.5μM ATP was added to all wells and the reactions in the first quadrant were immediately stopped by the addition of Kinase-Glo. These halted reactions would serve as the 0 minute time points on a time course. Quadrants 2,3 and 4 were halted at 10, 20 and 30 minutes, respectively. The luminescent readings were taken and the mean and standard deviation was determined from the readings of all wells for each time point (Figure 7.3).

These data demonstrate that the robustness of the assay clearly improves with longer incubation times. The Z-factor value improves from 0.35 to 0.7 between the 10 and 30 minute time points, indicative of improving assay signal dynamic range and lower data variation compared to the zero time point. Again, after 20 minutes 60% of ATP was used and 73% used after 30 minutes. This confirms that the preferred time for assay completion would be 20 minutes.

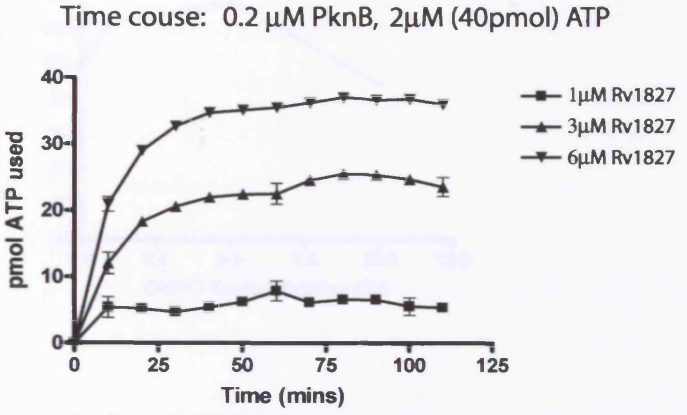
7.2.5 DMSO tolerance

Stock solutions of compounds used for screening are solubilised in dimethyl sulfoxide (DMSO), thus, the effect of DMSO on the phosphorylation of Rv1827 by PknB should be tested. Several of reaction were set up containing various concentrations of DMSO (0.14M-1.4M, 0.1-10% v:v) with 0.2μM of kinase concentration, 5μM Rv1827, 1.5μM ATP, 1mM $MnCl_2$, 150mM NaCl, 50mM Tris.HCl pH 8.0 in a total reaction volume of 100μL. The reaction was incubated for 30 minutes prior to reaction initiation with ATP. ATP turnover was plotted as a percentage of the positive control value for each DMSO concentration (Figure 7.4).

Table 7.1. Summary of 384-well format time couse experiment

Time points (mins)	0	10	20	30
Mean LU	9146.51	5575.75	3723.927	2488.729
S.D.	420.596	342.3418	312.9923	236.3773
***Z'* factor**	NA	0.359012	0.594148	0.703967

A.



B.

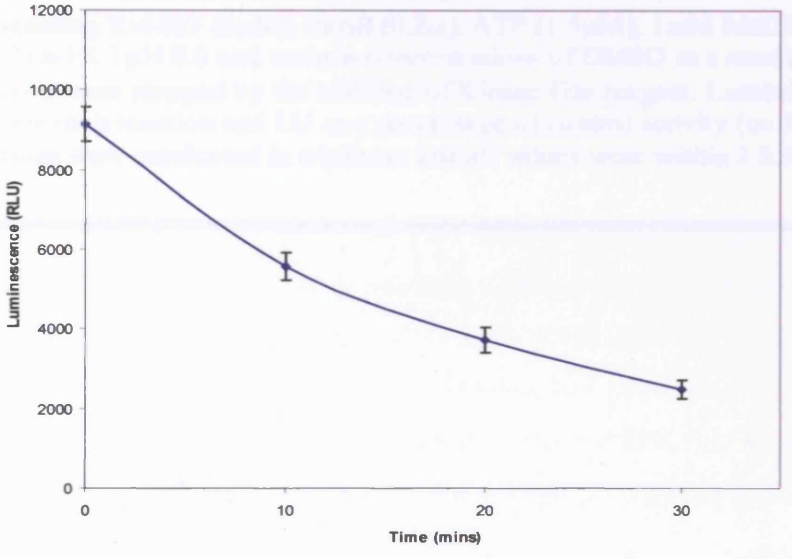


Figure 7.3. Representative IC_{50} data for Staurosporine inhibition of the reaction between kinase ($PknB^{1-279}$) and substrate ($Rv1827^{1-162}$). 4 sets of 96 reactions were carried in a 384-well plate containing Rv1827 (5µM), PknB (0.2µ), ATP (1.5µM), 1mM $MnCl_2$, 150mM NaCl, 50mM Tris-HCl pH 8.0 and various concentrations of Staurosporine in a reaction volume of 20µL. Reactions were stopped by the addition of Kinase-Glo reagent. Luminescence (LU) was recorded for each well and LU as a percentage of control activity (no inhibitor) was plotted against log inhibitor concentration.

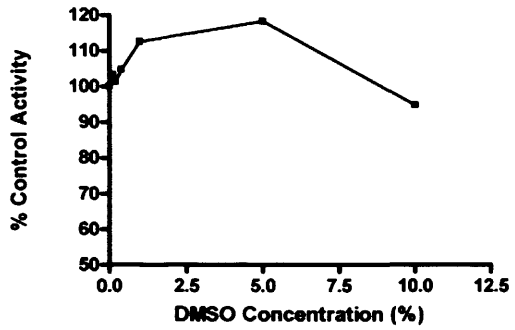


Figure 7.4. DMSO tolerance test for PknB activity with Rv1827. Reactions were carried out in parallel containing Rv1827 (5μM), PknB (0.2μ), ATP (1.5μM), 1mM $MnCl_2$, 150mM NaCl, 50mM Tris-HCl pH 8.0 and various concentrations of DMSO in a reaction volume of 100μL. Reactions were stopped by the addition of Kinase-Glo reagent. Luminescence (LU) was recorded for each reaction and LU as a percentage of control activity (no DMSO) was plotted. Reactions were conducted in triplicate and all values were within 2 S.D. of the average.

Results suggest there is no significant inhibition of PknB activity on Rv1827 by DMSO below ~1M. As high-throughput screening reactions will contain no more than 1% DMSO results will not be affected.

7.2.6 Validation of assay conditions

To establish whether the assay conditions selected above could detect enzyme inhibition, a known inhibitor of STPKs, staurosporine, was titrated across each quadrant of a plate containing the previously determined optimal reaction constituents. This is facilitated on the Biomek automated laboratory workstation due to the fact that all pipetting into 384-well plates is conducted separately for all four quadrants by this machine. This allowed the simultaneous evaluation of data generated from each quadrant to ensure consistency. Therefore, all the following procedural details are given for a 96-well plate format: 20μL of a reaction mixture containing Rv1827 (5μM), PknB (0.2μM), 1mM $MnCl_2$, 150mM NaCl, 50mM Tris.HCl pH 8.0 were added to each well except those in the first column which each contained 40μL of the reaction mix to which 10μM staurosporine had been added. A two-fold serial dilution was made across the plate by transfer pipetting. After preincubation, the reactions were initiated using 1.5μM ATP and stopped after 20 minutes. Readings were taken and the results of this inhibitor concentration titration are plotted (Figure 7.5).

Representative IC_{50} data were obtained and showed good consistency between quadrants with values obtained in each quadrant within 1 standard deviation (S.D) of each other. The average staurosporine IC_{50}, as determined from the four curves, was 126± 25nM. This illustrates that the assay conditions had been correctly adjusted to allow for a wide time window for the linear phase of the reaction. This window of ~20 minutes for the inhibited reaction allowed reproducible determinations of the initial velocity of substrate phosphorylation.

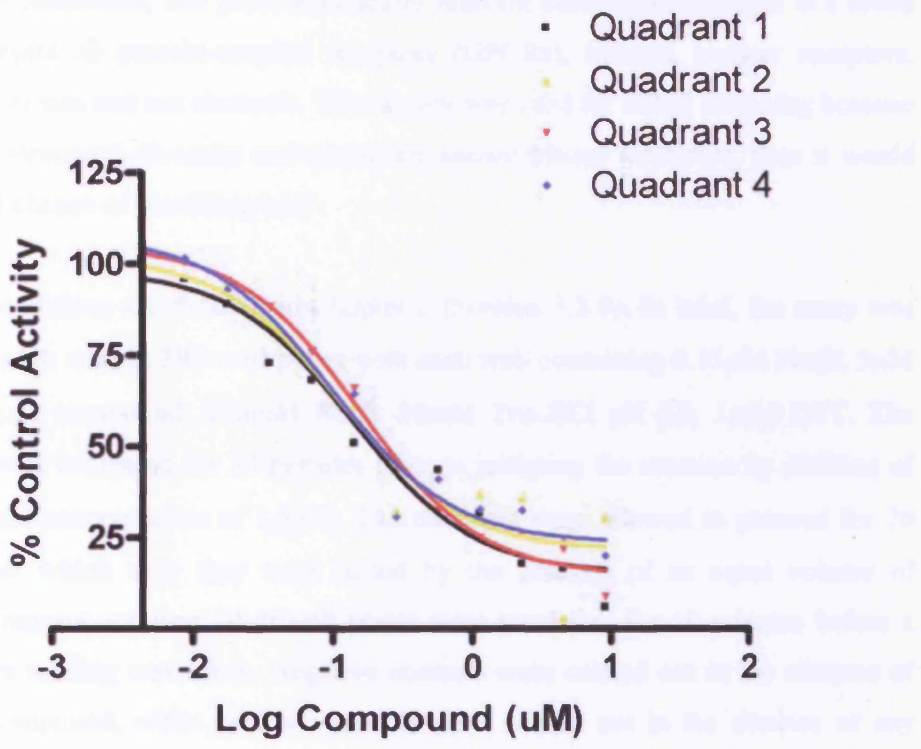


Figure 7.5. Representative IC_{50} data for Staurosporine inhibition of the reaction between kinase ($PknB^{1-279}$) and substrate ($Rv1827^{1-162}$). 4 sets of 96 reactions were carried in a 384-well plate containing Rv1827 (5μM), PknB (0.2μ), ATP (1.5μM), 1mM $MnCl_2$, 150mM NaCl, 50mM Tris-HCl pH 8.0 and various concentrations of Staurosporine in a reaction volume of 20μL. Reactions were stopped by the addition of Kinase-Glo reagent. Luminescence (LU) was recorded for each well and LU as a percentage of control activity (no inhibitor) was plotted against log inhibitor concentration. The staurosporine average IC_{50}, as determined from the four curves, was 126nM.

High-throughput screening of PknB inhibitors

It was decided that the first compound library to be screened would be the commercially available Library of Pharmacologically Active Compounds (LOPAC) from Sigma. This is a collection of 1280 compounds containing sets of approved drugs, failed development candidates, and pharmaceutically relevant compounds directed at a broad range of targets: G protein-coupled receptors (GPCRs), kinases, nuclear receptors, metabolic enzymes and ion channels. This library was used for initial screening because it had good structural diversity and contained known kinase inhibitors, thus it would stand a good chance of providing hits.

The assay conditions are detailed in Chapter 2 (Section 3.2.9). In brief, the assay was carried out batch wise in 384-well plates with each well containing 0.15μM PknB, 5μM Rv1827, 10μM compound, 150mM NaCl, 50mM Tris.HCl pH 8.0, 1mM DTT. The master mix was incubated for 30 minutes prior to initiating the reaction by addition of ATP to a final concentration of 1.5μM. The reactions were allowed to proceed for 20 minutes, after which time they were halted by the addition of an equal volume of Kinase-Glo reagent solution. Multiwell plates were incubated for 10 minutes before a luminescence reading was taken. Negative controls were carried out in the absence of PknB and compound, while positive controls were carried out in the absence of any compounds.

Data were normalised by subtracting the negative control luminescence unit (LU) reading from the data well LU and the percentage activity decrease for each compound was derived from the following formula:

$$RLU = \frac{1 - [LU(\text{data well}) - LU(\text{negative control})]}{[LU(\text{positive control}) - LU(\text{negative control})]}$$

The LOPAC Library was screened in duplicate at a 10μM compound concentration to assess assay performance and hit reproducibility. Compounds detected as 'hits' by the inhibitor screening are those found to inhibit PknB activity by 20% or greater. After the primary screen, 11 hits were obtained. A second assay was conducted with a longer preincubation time (one hour) and a further four hits were identified (Appendix 2). Inter-assay reproducibility could be illustrated using a 'scatter plot' depicting the percentage inhibition by compounds detected over the two assays (Figure 7.6). This showed a general trend for hits to locate near to the diagonal of the plot suggesting good reproducibility, with the exception of the four hits only observed in the second assay. Comparison of statistics for each screen also indicated improved performance for the second assay (Table 7.3). The *Z'*-value greatly improved between the two assays upon longer preincubation (from 0.51 to 0.77), bearing in mind *Z'* factor values above 0.5 indicate good assay quality and a value of 1 is a perfect assay (Zhang *et al.*, 1999).

Table 7.2. Summary of two high-througput screens of the LOPAC chemical library

	Assay 1			Assay 2		
	Average	s.d.	c.v.	Average	s.d.	c.v.
Enzyme control (LU)	2583	105	0.04	2567	104	0.04
No Enzyme control (LU)	9485	1057	0.11	9966	447	0.04
Z-factor	0.51			0.77		
Hit Rate (<80% Activity)	11/1024			15/1024		

(s.d.) standard deviation, (c.v.) coefficient of variance, (LU) luminescent units. The screening window coefficient, Z-factor, is defined in Chapter 3 – Materials and Methods, Section 3.2.9. *Z* factor values above 0.5 indicate good assay quality and a value of 1 is a perfect assay.

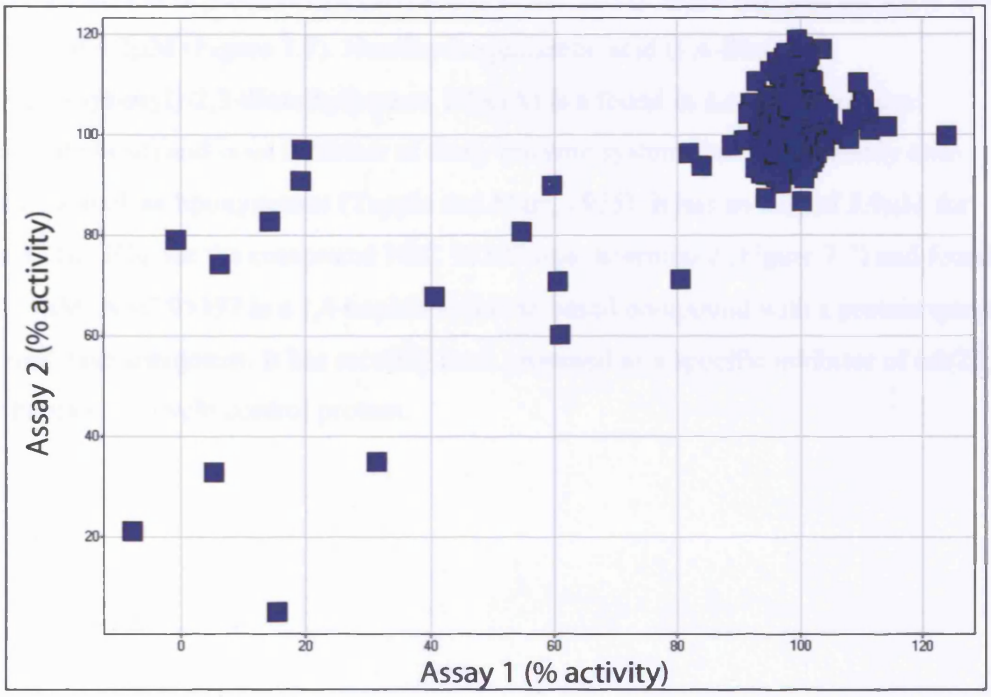


Figure 7.6. Scatter plot indicating interassay reproducibility for the screening of 1024 compounds from the LOPAC library (Sigma). 384-well plate containing Rv1827 (5μM), PknB (0.15μ), ATP (1.5μM), $MnCl_2$ (1mM), NaCl (150mM), Tris-HCl pH 8.0 (50mM) and 10μM compound. Reactions were stopped by the addition of Kinase-Glo reagent. Luminescence (LU) was recorded for each well and LU as a percentage of control activity (no inhibitor) for the two assays was plotted across each axis of the scatter plot. Summary statistics for the both assays are provided in table 15.

7.2.7 Validation of 'hit' compounds

Compounds detected as 'hits' by the inhibitor screening (those found to decrease PknB kinase activity by 20% or greater) were further investigated. A dose-response titration was performed for all hits and the IC_{50} values determined (Materials and Methods, Section 3.2.9). The validation data for three of the hit compounds myricetin, nordihydroguaiaretic acid and NSC 95397 are shown in Figure 7.7.

All three compounds exhibited low micromolar range IC_{50} values (4.2-7.3μM). These data show that myricetin inhibits PknB slightly more than nordihydroguaiaretic acid and NSC 95397. Myricetin (3,5,7-trihydroxy-2-(3,4,5-trihydroxyphenyl)-4H-1-benzopyran-4-one) is a naturally occurring flavonoid that can inhibit a broad range of protein kinases and is ATP competitive (Srivastava *et al.*, 1985). Myricetin was shown to have an IC_{50} of 4.2μM (Figure 7.7). Nordihydroguaiaretic acid (1,4-Bis(3,4-dihydroxyphenyl)-2,3-dimethylbutane, NDGA) is a found in *Larrea divaricata* (Creosote bush) and is an inhibitor of many enzyme systems but most notably anti-oxidants such as lipoxygenase (Tapple and Marr, 1955). It has an IC_{50} of 5.9μM for PknB. The IC_{50} for the compound NSC 95397 was determined (Figure 7.7) and found to be 7.3μM. NSC 95397 is a 1,4-naphthoquinone-based compound with a protein tyrosine phosphatase antagonist. It has recently been proposed as a specific inhibitor of cdc25, the human cell cycle control protein.

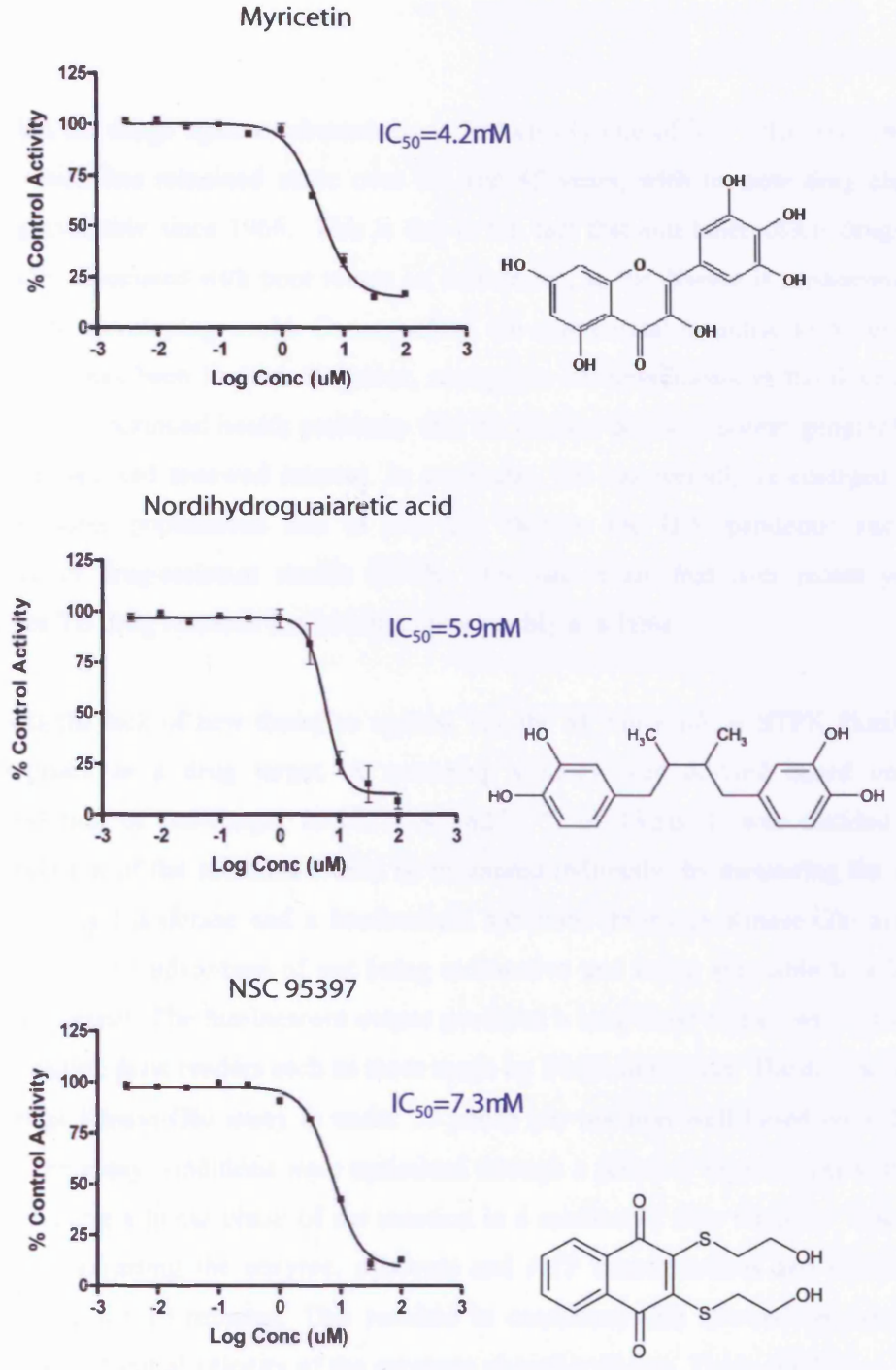


Figure 7.7. Dose response curves of inhibitors of PknB. Inhibition of PknB by myricetin, nordihydroguaiaretic acid and NSC 95397 was assayed at 0.15μM PknB, 5μM Rv1827 and 1.5μM ATP. Inhibitors were initially identified by high-throughput screening of the LOPAC library, Sigma. The IC_{50} values are the average of three individually measured dose response curves.

7.3 Discussion

The market for drugs against tuberculosis is historically one of low value and limited growth, which has remained static over the last 45 years, with no new drug classes becoming available since 1966. This is due to the fact that anti-tuberculosis drugs are traditionally associated with poor return on investment, as the disease is predominantly limited to the developing world. Consequently, the commercial incentive to develop or market drugs has been limited. However, resurgence of tuberculosis in the developed world and the continued health problems that the disease poses in poorer geographical regions has sparked renewed interest. In particular, TB has recently re-emerged as a threat to global populations due to two key factors: the HIV pandemic and the emergence of drug-resistant strains of TB. This has meant that over recent years, funding for TB drug research has become more readily available.

To address the lack of new therapies against TB, the *M. tuberculosis* STPK PknB has been proposed as a drug target. A screening strategy was devised based on the phosphorylation of full-length Rv1827, $Rv1827^{1-162}$, by PknB. It was decided that phosphorylation of the substrate would be measured indirectly, by measuring the ATP depletion using Luciferase and a luminescent substrate (Promega Kinase-Glo assay). This assay has the advantage of not being radioactive and being amenable to a high-throughput format. The luminescent output produces a long-lived signal, which can be read in common plate readers such as those made by BMG and Tecan. The cost of using the Promega Kinase-Glo assay is under 10 pence per reaction well based on a 20μL reaction. The assay conditions were optimised through a series of experiments with the aim of detecting a linear phase of the reaction in a reasonable time window. This was achieved by adjusting the enzyme, substrate and ATP concentrations and measuring ATP turnover for 20 minutes. This resulted in conditions that allowed reproducible determination of initial velocity of the substrate phosphorylation. These conditions were used to accurately determine the IC_{50} value for inhibition of PknB by staurosporine (Section 7.2.5).

The LOPAC library contains 167 known inhibitors of phospho-dependent signalling including those active against protein phosphatases as well as protein kinases. Six of the 15 confirmed hits identified are known protein kinase inhibitors. One compound, identified as a hit compound in both screens, was found to have intrinsic luminescence and interfered with the assay and therefore was discounted. Interestingly, compounds H7 and H8 which have isoquinol-based structures and are STPK broad spectrum inhibitors were represented in the LOPAC library but not identified as hits. H7 was shown to be active against PknB by Av-Gay and co-workers (Av-Gay *et al.*, 1999) and the inhibition of PknB by H7 and H8 were tested by the MRCT and IC_{50} values were determined to be 44 and 140μM, respectively. Therefore, using a compound concentration of 10μM in the previously described high-throughput screen is not expected to identify H7 and H8 as hits.

Staurosporine was not present in the LOPAC library but it is expected to be identified as a hit compound under the high-throughput assay conditions used in this Chapter. Under these conditions the IC_{50} was determined as 26nM. Staurosporine is clearly an effective inhibitor of PknB, although this has been disputed by previous authors (Ortiz-Lombardía *et al.*, 2003). They cite the loss of a hydrogen bond between a staurosporine nitrogen atom and the Asp156, the residue equivalent to conserved Asp102 in PKA. The analysis below shows that staurosporine could in fact interact with enough conserved and non-conserved residues to bind with nanomolar affinity.

A Basic Local Alignment Search Tool (BLAST) search of the human genome with the amino-acid sequence of PknB kinase domain reveals 3-phosphoinositide-dependent protein kinase-1 (PDK1) to be its closest match with ~31% identity. The divergence between the two kinases has been demonstrated previously (Figure 1.10). PDK1 is a member of the AGC (cAMP-dependent, cGMP-dependent, protein kinase C) family of human protein kinases (Figure 1.10), and has a key role in insulin and growth-factor signalling through phosphorylation and subsequent activation of a number of other AGC kinase family members including protein kinase B. The structure of PDK1 in complex with staurosporine has been determined by X-ray crystallography. It is therefore possible to model a potential staurosporine interaction with the kinase domain of PknB (Figure 7.8) by overlapping the structure of PknB with that of the PDK1-staurosporine complex (Komander *et al.*, 2005). These overlapping structures serve to demonstrate that although staurosporine binds via residues common to both PDK1 and

PknB there is sufficient divergence between the two active sites to suggest that inhibitors specific for PknB should be obtainable.

Of the 14 hit compounds identified after screening the LOPAC library, nordihydroguaiaretic acid (1,4-bis(3,4-dihydroxyphenyl)-2,3-dimethylbutane, NDGA) generated most interest. NDGA is found in *Larrea divaricata* (Creosote bush) and extracts of this plant have been used traditionally to treat a number of illnesses in Mexico including tuberculosis (Rivero-Cruz *et al.,* 2005). NDGA is an inhibitor of many enzyme systems but most notably anti-oxidants such as lipoxygenase (Tapple and Marr, 1955). Very recent studies have confirmed NDGA is as also an effective anti-mycobacterial agent with a minimal inhibitory concentration (MIC) value of 50μg/mL (Rivero-Cruz *et al.,* 2005). The MIC of a compound is defined as the lowest concentration of antimicrobial agent that inhibits the growth of the microorganism. This coupled with the fact that NDGA has an IC_{50} of 5.9μM for PknB suggests PknB may be the target for NDGA in mycobacteria, however, there is no information about its mechanism of action or selectivity. Prior work has shown that kinase inhibitor H7 was able to inhibit the growth of *M. smegmatis*, however, currently no work has been published into the effect of kinase inhibitors on the growth of *M. tuberculosis*. This information presented here may provide evidence that inhibition of PknB and/or other kinases in *M. tuberculosis* can prove effective in preventing growth of this bacterium.

In order to identify as many hits as possible which can then act as a starting point for medicinal chemistry, a number of chemical libraries will be screened using the assay developed from this study. These will include libraries that specifically target serine/threonine protein kinases. Initially screening of the 1280 compounds of the LOPAC library served as validation of assay conditions in a high-throughput format as it identified 14 hit compounds after two separate runs, six of which are known inhibitors of kinases. Analysis of results generated by the two runs showed that they were both sufficiently robust (run 1, Z' of 0.51 and run 2, Z' of 0.77). Thus, this assay represents an important discovery route for novel anti-TB chemotherapeutics.

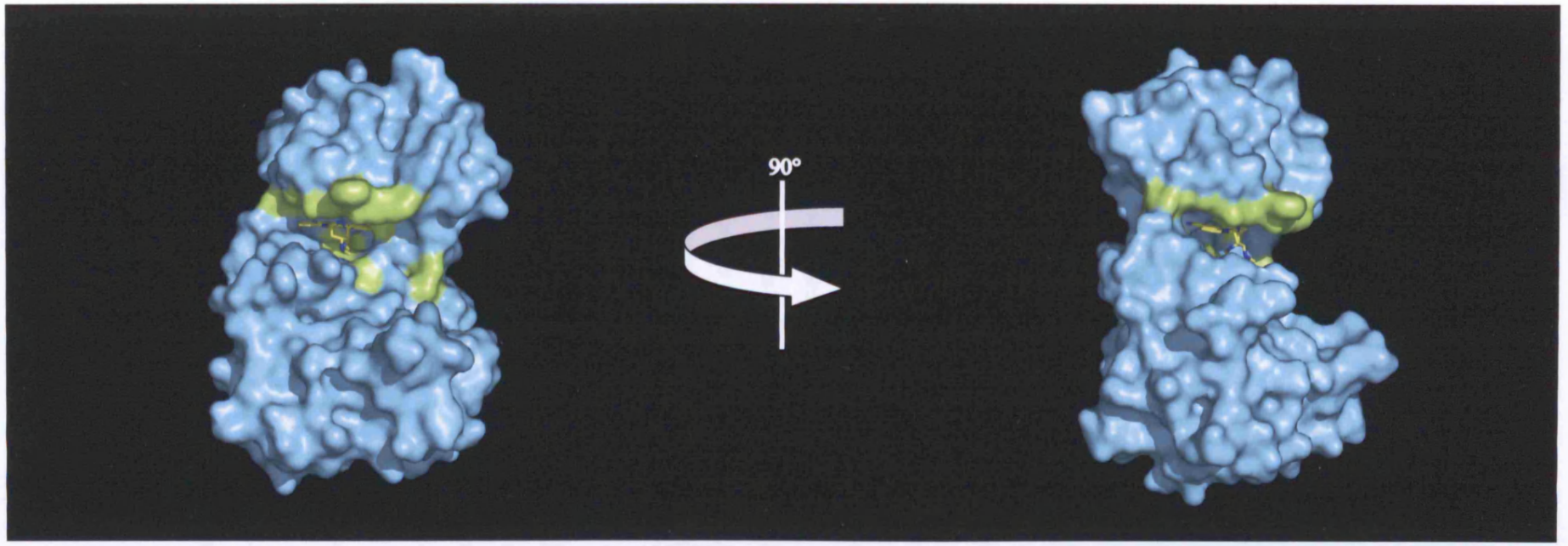


Figure 7.8. Structural basis for the binding of staurosporine into the ATP-binding site of PknB. The structure of PDK1 in complex with staurosporine has been determined by X-ray crystallography (Komander *et al.*, 2005). After overlapping the structure of PknB with that of the PDK1-staurosporine complex, it is possible to model a potential staurosporine interaction with the kinase domain of PknB. The surface representation of the kinase domain of PknB is shown with the bound staurosporine molecule shown in ball-and-stick representation. Active site residue surfaces conserved between PDK1 and PknB are shown in green. Molecules were overlapped by structural similarity using the Magic Fit program of Deepview (www.expasy. org/spdbv)

8 Conclusions

8.1 Introduction

This thesis describes a series of biophysical and biochemical analyses of the interactions between two *M. tuberculosis* STPKs, PknA and PknB, and three FHA domain containing proteins, Rv0019, Rv0020 and Rv1827. After purification of these proteins either as isolated domains or longer fragments, *in vitro* kinase assays were employed to evaluate the interrelationship between each protein. PknA and PknB were found to phosphorylate FHA domain-containing proteins following autophosphorylation. These observed phosphorylation reactions were then dissected through the use of ITC and SPR techniques and biophysical parameters for Kinase-FHA interactions were determined.

These data have added to the growing body of information about the role of STPKs in *M. tuberculosis*, specifically in conjunction with FHA domains. In particular, the first detailed biophysical analysis of interactions between FHA and STPK is presented. A model of the possible mechanism which FHA domain-containing proteins become phosphorylated by STPKs in *M. tuberculosis* has evolved through data generated by this and previous studies (Figure 6.7).

STPKs PknA/B appear to be involved in a signalling pathway concerned with the correct regulation of cell wall synthesis and cell division (Av-Gay, Y., and Everett, M. 2000) (Kang *et al.*, 2005). Perturbation of this pathway, by over- or under-expression, results in gross morphological changes and reduced cell viability (Kang *et al.*, 2005). Consequently PknB represents a potentially important target for anti-TB chemotherapeutics. Therefore, the final part of this study was spent developing a high-throughput kinase assay to screen for inhibitors of PknB. A systematic approach was used to optimise a screening method based on the phosphorylation of full-length Rv1827 by PknB and utilising a luminescent signal for assessing activity.

8.2 Kinase Activity Studies and FHA Interactions

In vitro kinase assays were used to elucidate the molecular reactions between various STPKs and FHA domain-containing proteins of *M. tuberculosis*. Other groups have recently substantiated the work presented here (Grunder *et al.*, 2005, Kang *et al.*, 2005, Villarino *et al.*, 2005). These efforts have now been extended to include preliminary studies of the activities of PknA and the FHA-domain containing protein, Rv0019. The demonstration of Rv1827 phosphorylation by PknB parallels results obtained by (Duran *et al.*, 2005). These authors also show phosphorylation of full-length Rv1827 and map phosphorylation to a single threonine, T22. Furthermore, a portion of Rv0020 N-terminal to the FHA domain was found to be phosphorylated but the site or sites for this remain unknown. Phosphorylation of Rv0020 and Rv1827 was abrogated through the mutation of a key serine residue in the domain which rendered the FHA domains non-functional. The absence of FHA domain phospho-threonine binding ability, as there was no detectable phosphorylation, might suggest a molecular docking model for the interaction between these FHA domain containing proteins and STPKs, as depicted in Figure 6.7.

Two specific phospho-threonine residues on the activation loop of the kinase domain were found to be the binding site for FHA domains Rv0020 and Rv1827. This leads to the conclusion that the absence of these threonines completely eliminates binding. The FHA domain of Rv0020 displayed preferential binding to pT173, the likely secondary phosphorylation site of PknB using ITC and SPR. The FHA domain of Rv1827 showed no such preference. The FHA domain of Rv0020 showed tighter affinity for PknA and PknB than the FHA of Rv1827.

8.3 An Emerging Model for FHA domain-containing protein Phosphorylation

The mechanism by which two FHA-domain containing proteins, Rv1827 and Rv0020 become phosphorylated by STPKs PknB relies on the autophosphorylation of two threonines residing in the activation loop. Functional FHA domains from Rv0020 and Rv1827 efficiently bind to the PknB activation loop consisting of one or two phospho-threonine residues. This molecular docking interaction that allows the N-terminal tails of these proteins to bind in the substrate-binding site of PknB. Thus, the N-terminal portion of Rv0020 and Rv1827 become phosphorylated. This molecular docking mechanism is also likely to occur in reactions involving PknA but further experiments will be required to confirm this or otherwise.

The activation loop of the PknB kinase domain is found to be disordered after autophosphorylation. However, upon FHA binding, the activation loop may be sandwiched between the kinase and FHA domains, consequently becoming ordered. Once the FHA domain binds the activation loop, inactivation of the kinase by a protein phosphatase may be prohibited. This was also suggested for a related interaction discovered by Hubbard and coworkers where the SH2 domain of adaptor protein APS complexes with two phosphotyrosine residues of the insulin receptor kinase (IRK) activation loop (Hu *et al.*, 2003).

The significance of phosphorylation in the region N-terminal to the FHA domain still remains unclear. No domains are predicted in the N-terminal portions of both Rv1827 and Rv0020 from homology searches and they appear not to have any recognisable structure. Results presented in this thesis suggest that the known phosphorylation site of Rv1827 is able to serve as a binding site for both its own FHA domain and for Rv0020. This supports dimer (Rv1827-Rv1827) or hetero-oligomer (Rv0020-Rv1827) formation. An explanation as to the precise purpose of these complexes remains elusive. It is hard to envisage a possible function for these complexes if both components are phosphorylated and one FHA binds to the adjacent FHA domain-containing protein's phosphorylated tail and *vice versa*. However, if Rv1827 were to be phosphorylated and its FHA domain were to remain bound to a kinase activation loop, then binding of any

FHA-domain containing protein onto this phosphorylated tail would result in a multimeric complex which immediately localises the newly bound protein for phosphorylation.

8.4 The Control of *M. tuberculosis* Growth and Morphology with STPKs

Genetic and inhibitor studies carried out mainly in *E. coli* cells have identified that bacteria use two types of peptidoglycan synthesis for cell growth; either cell wall elongation or septal synthesis. The *pknA* and *pknB* genes are part of an operon that includes genes known to be involved in the elongation mode of synthesis; *rodA and pbpA* (Begg *et al.*, 1990) (Figure 1.9). While *rodA* and *pbpA* work in unison to produce a rod morphology, *fts* genes (*ftsA, ftsZ, ftsQ, ftsN, ftsL, ftsK, ftsW* and *zipA*) are required for the septal mode for cell division (Hale *et al.*, 1997). To switch between an elongation mode of synthesis, which involves insertion of peptidoglycan along the cylinder of the cell, to septal synthesis requires insertion of peptidoglycan along the leading edge of the invagination septum (Wang *et al.*, 1998). Septation starts after FtsZ, the bacterial homolog of tubulin, polymerises to form a contractile ring at the future division site. Around 12 additional proteins, many of them *fts* genes, are also recruited to the division site to form the 'divisome' (Goehring *et al.*, 2005).

The evolution of a rod shape in bacteria, such as *M. tuberculosis* and *Bacillus subtilis*, necessitates a mechanism to temporally regulate the switch between elongation and septation. How the cell switches from one process to the other remains unclear (Margolin *et al.*, 2005). *pknA and pknB* have been postulated to be involved in this process because of their genetic proximity to *rodA* and *pbpA*. Recently, Kang et al., (2005), showed that over-expression of PknA and PknB separately in *M. tuberculosis* cells results in a broad, bulgy cell morphology, while partial depletion of PknA and PknB, using an antisense strategy, caused a narrow elongated morphology. These observations are extremely similar to morphology changes associated with mutants of *rodA* and *pbpA* in *E. coli*. This suggests a functional linkage between *PknA, PknB, rodA* and *pbpA*. The nature of the interaction between PknA and PknB with these proteins is

yet to be established. However, PknB does contain extracellular PASTA (penicillin binding protein and STPK associated) domains also present in the extracellular portion of PbpA homolog in *S. pneumoniae*, Pbp2x. Although the PASTA domain is not present in PbpA itself, this does predict an important relationship between the PbpA and PknB: that PbpA and PknB work in concert to exert an effect on cell morphology. The PASTA domain binds to peptidoglycan and their β-lactam antibiotics analogues, and are thought to direct proteins to the site of cell growth (Yeats et al., 2002). Therefore, this enzyme may be localised to areas of cell growth by means of extracellular PASTA domains, leading to oligomerisation and auto-activation. This could produce a signal in response to an extracellular signal that informs the cell that division is required, although, cell division would also be governed by intracellular factors such as the replication status of the genome.

Further intracellular regulation mechanisms of PknA and PknB have also been proposed. Ppp is the only mycobacterial protein phosphatase and was found to be able to dephosphorylate PknB (Boitel et al., 2003) and PknA (Chopra et al., 2003). It can be hypothesised that a third mechanism of regulation of these PknA/B is via the co-expressed FHA domain-containing proteins Rv0020 and Rv0019 and other FHA domain-containing proteins encoded throughout the *M. tuberculosis* genome. Contrary to expectation PknA/B are shown to interact with a protein from disparate operons, namely Rv1827; this suggests possible temporal control of this signalling system. The results presented here suggest PknB is activated upon autophosphorylation, and autophosphorylation of a threonine in the activation loop is essential for docking of FHA domains to the kinase and the subsequent phosphorylation of Rv0020 and Rv1827. This report has uncovered a small part of a larger network of interactions involving the STPKs and FHA domains of *M. tuberculosis*. In many cases, the STPK function has been linked to cell growth including PknF (Deol *et al*, 2005), PknG (Walburger *et al.*, 2004) and PknH (Sharma *et al.*,2004). Combining this work with existing knowledge of STPK-FHA interactions, a complex picture of STPK/FHA signalling emerges where specific interactions are controlled through the phospho-protein binding ability of various FHA domains.

8.5 Towards Novel Anti-tuberculosis Chemotherapeutics

In addition to the biophysical and biochemical characterisation of STPK signalling in *M. tuberculosis*, the latter part of this study focused on the search for a novel intervention strategy to disrupt STPK signalling. Of particular importance, was the goal to discover compounds able to inhibit the phosphorylation of Rv1827 by PknB, an interaction demonstrated during the course of this study. Although, the precise physiological function of PknB remains uncertain, earlier discussion has detailed the importance of PknB and PknA in the *M. tuberculosis* cell and shows them to be a potentially significant drug targets.

Through the development of a robust, high-throughput enzymatic assay, the screening of inhibitors against PknB has been conducted and is detailed in this report. The screening of the LOPAC chemical library has been conducted, which validates the assay protocol, and several larger screens are planned. However, this represents the first step in the longer drug development process.

Once hit compounds are identified, it is important to demonstrate reproducible concentration-dependent activity of the inhibitor and determine the IC_{50} value, as conducted for LOPAC library hits (Section 7.2.7). This is followed by confirmation of hit compounds in alternative *in vitro* kinase assays, such as the enzyme-coupled spectrophotometric assay and radiolabelling assay (Chapter 6). The next phase would be to conduct additional assessment of the physicochemical properties of compounds, including confirmation of structure and purity and tests for chemical reactivity. Finally, it is necessary to confirm the compound in question is not classed as a 'frequent hitter' (i.e. a compound that is a common inhibitor in drug screens).

In addition, selectivity assays will need to be conducted where the compounds are tested for inhibitory properties against a panel of human kinases. At this stage, it would also be important to test any inhibitors of PknB for activity against other *M. tuberculosis* STPKs (i.e. PknA – PknL). Despite the unavailability of such a panel, the purification of many of these kinases in their active forms has been reported. Here PknA has been purified in an active form (Section 4.3). Although in theory it would not be considered

a problem if any potential inhibitor also showed inhibitory activity against any of the other *M. tuberculosis* STPKs, it is important to note that PknH knockout *M. tuberculosis* cells are in fact hypervirulent (Sharma *et al.*,2004) and therefore it may be important to determine kinase specificity. PknH knockout *M. tuberculosis* cells replicate to a higher bacillary load than wild-type upon infection of mice, suggesting PknH may be important in maintaining the latency stage of infection. It may, thus, be advantageous to inhibit PknH and thereby move bacteria out of latency.

Initial evaluation of compounds for inhibition of the *in vitro* replication of Mycobacteria would need to be carried out by following the growth of *M. smegmatis* in culture media. Fortunately, this type of experimentation is relatively straightforward as *M. smegmatis* grows rapidly (approx. 3 days) and growth can be measured by taking optical density (O.D.) readings. Such growth experiments should also be carried out with *M. tuberculosis,* but this experimental system is more complex as growth takes approximately three weeks and such experiments need to be conducted in containment level 3 facilities. In addition to evaluation of activity against wildtype strains of *M. tuberculosis*, it will also be beneficial to investigate activity against variants with resistance to the current antibiotics. If compounds appear to have inhibitory activity in such growth experiments in culture media, then it is necessary to evaluate their effects on the growth of *M. tuberculosis* in macrophages, as described previously (Tascon *et al.*, 2000). When conducting experiments in murine macrophages it will also be important to check in parallel for any cytotoxicity.

Structure-based drug design and/or screening strategies may be utilised in conjunction with traditional screening strategies. Computational methodologies could conceivably be utilised for hit identification as the structure of the target, PknB, and its active site are known. Many drug discovery programmes have involved structure-based virtual screening with successful identification of inhibitors for protein kinases such as CK2 and BCL-ABL tyrosine kinase (Kitchen *et al.*, 2005). In addition to hit identification, other *in silico* techniques can be similarly effective in structure-based lead optimisation. Once a lead compound has been co-crystallised with its target, modern software applications are available which can model compound modifications onto pre-defined core fragments of hits (Ewing *et al.*, 2001). Minor modifications in the chemical structure of a hit can lead to significant changes in the binding affinity. This can facilitate a hit-to-lead transition where the compound potency should be increased by

two to three orders of magnitude (Kitchen *et al.*, 2005). The structure of the PknB kinase domain with a modelled structure of Staurosporine is presented here to illustrate how structural information can help improve the drug discovery process.

9 Future work

A number of important experiments are required to extend this area of research.

It may be possible to reduce the extent of non-specific phosphorylation observed upon purification of PknA (Section 4.3.2). This could be achieved by either co-expressing recombinant PknA with the protein phosphatase, Ppp, using a pET Duet vector or by treatment of PknA with this phosphatase following purification. Ppp has proven to be active on PknA (Boitel *et al.*, 2005) and may aid purification of PknA in a form more akin to its native phosphorylation state. Functional studies with this form of PknA may prove more fruitful than studies conducted using the hyperphosphorylated form. The effects of protein phosphatase co-expression on the expression of Abl and *Src* protein kinase domains have been studied and it was found that this led to increased bacterial expression and reduction in associated toxic effects (Seeliger *et al.*, 2005).

After observation of phosphorylation of the Rv0019 FHA domain it is important to find the site of phosphorylation and determine the effect of phosphorylation on the phospho-threonine binding ability of the domain. This study reports on the inability of the non-phosphorylated form to bind phospho-threonine targets. If Rv0019 were to bind phospho-threonine targets after phosphorylation, this would represent a novel and interesting mechanism for modulation of FHA activity. In addition, the site(s) at which Rv0020 is phosphorylated by PknA and PknB is yet to be determined. Although, phosphorylation is found to be in the region N-terminal region of the Rv0020 FHA domain, identification of the phospho-acceptor residue(s) may aid understanding of Rv0020s physiological role. Although, no function can be assigned to this 373 amino acid N-terminal stretch by homology prediction, phosphorylation may could regulate the function of this region. The fact that Rv1827 is found in *M. tuberculosis* cells in a truncated form, $Rv1827^{31-162}$, where the phosphorylation site may have been cleaved off indicates that FHA domain propagated signals that may be regulated by truncation of the protein. Information about the levels of this truncated form of Rv1827 present in *M. tuberculosis* cells at different growth conditions may help to elucidate the role of Rv1827 and how it is regulated.

The possibility for FHA oligomerisation of Rv0020 and Rv1827 upon phosphorylation of the N-terminal portion of these proteins has been raised. This was postulated after demonstrating the FHA domains of Rv0020 and Rv1827 could bind a peptide representing the N-terminal phosphorylation site of Rv1827 (Section 6.2.9). To investigate this possibility it would be necessary to obtain a purified and phosphorylated form for Rv1827 or Rv0020. Following this, biophysical techniques such as ITC and SPR could again be applied to identify protein interactions and methods such as analytical ultracentrifugation could be helpful in identifying the oligomerisation state.

There remains a lack of information on the physiological role of FHA-domain containing proteins in bacteria, despite their presence in important prokaryotic genomes and their links to many important cellular processes. The requirement for knockout mutant strains lacking FHA-domain containing proteins is therefore crucial. If genes for FHA-domain containing proteins turn out to be essential, disruption of FHA-STPK interaction may prove an effective therapeutic strategy. In the case of Rv0019 and Rv0020, if knockout cells of these proteins display similar morphology to cell where the levels of are reduced, this would imply a direct functional relationship between Rv0019 and Rv0020 and PknA/B.

The opportunity exists to examine the full extent of STPK-FHA interactions in *M. tuberculosis*, as the majority of FHA domains and STPKs have been purified as soluble recombinant proteins. Once soluble forms of these proteins are available one can do a biochemical and biophysical studies similar to those described in chapters 5 and 6, to decipher the complex series of interactions between FHA domains and STPKs present in *M. tuberculosis*. Although a number of FHA-STPK interactions have been described in this study, all of these were observed in the absence of ATP. There may be a difference in binding affinities for these interactions after an ATP is bound within the kinase domain, as this is known to induce a conformational change in the kinase domain and this requires further study. In this regard, co-crystallisation studies of *M. tuberculosis* STPKs with FHA domains will be essential. As yet, the structures of the Rv0020 FHA domain and PknB kinase domain have only been solved individually. A structure of the PknB-Rv0020 FHA complex may clarify the interactions which occur upon complex formation, how the Rv0020 FHA domain can coordinate to a motif containing two phospho-threonines, if the FHA domain interacts with any kinase domain elements other than the activation loop and what conformation the activation

loop of PknB adopts upon FHA domain binding. The question of how the Rv0020 FHA domain can bind to a motif containing two phospho-threonines can also be answered with a structure of the FHA domain in complex with a peptide representing the fully phosphorylated activation loop of PknB. This would also be the first structure of a bacterial FHA domain in complex with a physiological substrate. Additional structures of *M. tuberculosis* FHA domain containing proteins including Rv1827 may add further molecular information to the bacterial branch of known FHA domains.

With respect to the search for inhibitors against PknB, a vast amount of follow up work would be required after initial screening is conducted and hits are identified. Much of this is outlined in the previous Chapter (Sections 8.5). In addition, it may also be advantageous to produce soluble kinase domain protein fragments representing the complete complement of *M. tuberculosis* STPKs for inhibitor screening. Regardless of the potential of such novel inhibitor compounds for therapeutic intervention, the availability of kinase-specific inhibitors will certainly be important tools for further, systematic investigation of these fascinating bacterial signalling pathways.

10 References

Ablooglu, A. J., and Kohanski, R. A. (2001) Activation of the insulin receptor's kinase domain changes the rate-determining step of substrate phosphorylation *Biochemistry* **40**, 504-513

Ahmed, Z., and Pillay, T. S. (2003) Adapter protein with a pleckstrin homology (PH) and an Src homology 2 (SH2) domain (APS) and SH2-B enhance insulin-receptor autophosphorylation, extracellular-signal-regulated kinase and phosphoinositide 3-kinase-dependent signalling *Biochem J* **371**, 405-412

Ahn, J. Y., Li, X. H., Davis, H. L., and Canman, C. E. (2002) Phosphorylation of threonine 68 promotes oligomerization and autophosphorylation of the Chk2 protein kinase via the forkhead-associated domain *Journal of Biological Chemistry* **277**, 19389-19395

Al-Obeidi, F. A., Wu, J. J., and Lam, K. S. (1998) Protein tyrosine kinases: Structure, substrate specificity, and drug discovery *Biopolymers* **47**, 197-223

Av-Gay, Y., and Everett, M. (2000) The eukaryotic-like Ser/Thr protein kinases of Mycobacterium tuberculosis *Trends Microbiol* **8**, 238-244

Av-Gay, Y., Jamil, S., and Drews, S. J. (1999) Expression and characterization of the Mycobacterium tuberculosis serine/threonine protein kinase PknB *Infect Immun* **67**, 5676-5682

Bakal, C. J., and Davies, J. E. (2000) No longer an exclusive club: eukaryotic signalling domains in bacteria *Trends Cell Biol* **10**, 32-38

Barz, C., Abahji, T. N., Trulzsch, K., and Heesemann, J. (2000) The Yersinia Ser/Thr protein kinase YpkA/YopO directly interacts with the small GTPases RhoA and Rac-1 *Febs Letters* **482**, 139-143

Bauer, C. B., Kuhlman, P. A., Bagshaw, C. R., and Rayment, I. (1997) X-ray crystal structure and solution fluorescence characterization of Mg.2'(3')-O-(N-methylanthraniloyl) nucleotides bound to the Dictyostelium discoideum myosin motor domain *J Mol Biol* **274**, 394-407

Begg, K. J., Takasuga, A., Edwards, D. H., Dewar, S. J., Spratt, B. G., Adachi, H., Ohta, T., Matsuzawa, H., and Donachie, W. D. (1990) The Balance between Different Peptidoglycan Precursors Determines Whether Escherichia Coli cells will Elongate or Divide *Journal of Bacteriology* **172**, 6697-6703

Boitel, B., Ortiz-Lombardia, M., Duran, R., Pompeo, F., Cole, S. T., Cervenansky, C., and Alzari, P. M. (2003) PknB kinase activity is regulated by phosphorylation in two Thr residues and dephosphorylation by PstP, the cognate phospho-Ser/Thr phosphatase, in Mycobacterium tuberculosis *Mol Microbiol* **49**, 1493-1508

Bowles, M. R., Hall, D. R., Pond, S. M., and Winzor, D. J. (1997) Studies of protein interactions by biosensor technology: an alternative approach to the analysis of sensorgrams deviating from pseudo-first-order kinetic behavior *Anal Biochem* **244**, 133-143

Braunwalder, A. F., Yarwood, D. R., Hall, T., Missbach, M., Lipson, K. E., and Sills, M. A. (1996) A solid-phase assay for the determination of protein tyrosine kinase activity of c-src using scintillating microtitration plates *Analytical Biochemistry* **234**, 23-26

Chaba, R., Raje, M., and Chakraborti, P. K. (2002) Evidence that a eukaryotic-type serine/threonine protein kinase from Mycobacterium tuberculosis regulates morphological changes associated with cell division *Eur J Biochem* **269**, 1078-1085

Chaba, R., Raje, M., and Chakraborti, P. K. (2002) Evidence that a eukaryotic-type serine/threonine protein kinase from Mycobacterium tuberculosis regulates morphological changes associated with cell division *European Journal of Biochemistry* **269**, 1078-1085

Chopra, P., Singh, B., Singh, R., Vohra, R., Koul, A., Meena, L. S., Koduri, H., Ghildiyal, M., Deol, P., Das, T. K., Tyagi, A. K., and Singh, Y. (2003) Phosphoprotein phosphatase of Mycobacterium tuberculosis dephosphorylates serine-threonine kinases PknA and PknB *Biochem Biophys Res Commun* **311**, 112-120

Cole, S. T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., Gordon, S. V., Eiglmeier, K., Gas, S., Barry, C. E., 3rd, Tekaia, F., Badcock, K., Basham, D., Brown, D., Chillingworth, T., Connor, R., Davies, R., Devlin, K., Feltwell, T., Gentles, S., Hamlin, N., Holroyd, S., Hornsby, T., Jagels, K., Barrell, B. G., and et al. (1998) Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence *Nature* **393**, 537-544

Collett, M. S., and Erikson, R. L. (1978) Protein kinase activity associated with the avian sarcoma virus src gene product *Proc Natl Acad Sci U S A* **75**, 2021-2024

Cowley, S., Ko, M., Pick, N., Chow, R., Downing, K. J., Gordhan, B. G., Betts, J. C., Mizrahi, V., Smith, D. A., Stokes, R. W., and Av-Gay, Y. (2004) The Mycobacterium tuberculosis protein serine/threonine kinase PknG is linked to cellular glutamate/glutamine levels and is important for growth in vivo *Mol Microbiol* **52**, 1691-1702

Curry, J. M., Whalan, R., Hunt, D. M., Gohil, K., Strom, M., Rickman, L., Colston, M. J., Smerdon, S. J., and Buxton, R. S. (2005) An ABC transporter containing a forkhead-associated domain interacts with a serine-threonine protein kinase and is required for growth of Mycobacterium tuberculosis in mice *Infection and Immunity* **73**, 4471-4477

Davies, S. P., Reddy, H., Caivano, M., and Cohen, P. (2000) Specificity and mechanism of action of some commonly used protein kinase inhibitors *Biochemical Journal* **351**, 95-105

Deol, P., Vohra, R., Saini, A. K., Singh, A., Chandra, H., Chopra, P., Das, T. K., Tyagi, A. K., and Singh, Y. (2005) Role of Mycobacterium tuberculosis Ser/Thr kinase PknF: implications in glucose transport and cell division *J Bacteriol* **187**, 3415-3420

Drews, S. J., Hung, F., and Av-Gay, Y. (2001) A protein kinase inhibitor as an antimycobacterial agent *FEMS Microbiol Lett* **205**, 369-374

Duran, R., Villarino, A., Bellinzoni, M., Wehenkel, A., Fernandez, P., Boitel, B., Cole, S. T., Alzari, P. M., and Cervenansky, C. (2005) Conserved autophosphorylation pattern in activation loops and juxtamembrane regions of Mycobacterium tuberculosis Ser/Thr protein kinases *Biochem Biophys Res Commun* **333**, 858-867

Durocher, D. (2003) Bacterial signal transduction: a FHA domainscinating glimpse at the origins of phospho-dependent signal transduction *Trends Microbiol* **11**, 67-68

Durocher, D., Henckel, J., Fersht, A. R., and Jackson, S. P. (1999) The FHA domain is a modular phosphopeptide recognition motif *Mol Cell* **4**, 387-394

Durocher, D., and Jackson, S. P. (2002) The FHA domain *FEBS Lett* **513**, 58-66

Durocher, D., Smerdon, S. J., Yaffe, M. B., and Jackson, S. P. (2000) The FHA domain in DNA repair and checkpoint signaling *Cold Spring Harb Symp Quant Biol* **65**, 423-431

Durocher, D., Taylor, I. A., Sarbassova, D., Haire, L. F., Westcott, S. L., Jackson, S. P., Smerdon, S. J., and Yaffe, M. B. (2000) The molecular basis of FHA

domain:phosphopeptide binding specificity and implications for phospho-dependent signaling mechanisms *Mol Cell* **6**, 1169-1182

Eckhart, W., Hutchinson, M. A., and Hunter, T. (1979) Activity Phosphorylating Tyrosine in Polyoma T-Antigen Immunoprecipitates *Cell* **18**, 925-933

Engh, R. A., and Bossemeyer, D. (2001) The protein kinase activity modulation sites: Mechanisms for cellular regulation - Targets for therapeutic intervention *Advances in Enzyme Regulation, Vol 41* **41**, 121-149

Engh, R. A., and Bossemeyer, D. (2002) Structural aspects of protein kinase control-role of conformational flexibility *Pharmacol Ther* **93**, 99-111

Ewing, T. J. A., Makino, S., Skillman, A. G., and Kuntz, I. D. (2001) DOCK 4.0: Search strategies for automated molecular docking of flexible molecule databases *Journal of Computer-Aided Molecular Design* **15**, 411-428

Frodin, M., Antal, T. L., Dummler, B. A., Jensen, C. J., Deak, M., Gammeltoft, S., and Biondi, R. M. (2002) A phosphoserine/threonine-binding pocket in AGC kinases and PDK1 mediates activation by hydrophobic motif phosphorylation *Embo Journal* **21**, 5396-5407

Goehring, N. W., and Beckwith, J. (2005) Diverse paths to midcell: assembly of the bacterial cell division machinery *Curr Biol* **15**, R514-526

Good, M. C., Greenstein, A. E., Young, T. A., Ng, H. L., and Alber, T. (2004) Sensor domain of the Mycobacterium tuberculosis receptor Ser/Thr protein kinase, PknD, forms a highly symmetric beta propeller *Journal of Molecular Biology* **339**, 459-469

Gopalaswamy, R., Narayanan, P. R., and Narayanan, S. (2004) Cloning, overexpression, and characterization of a serine/threonine protein kinase pknI from Mycobacterium tuberculosis H37Rv *Protein Expr Purif* **36**, 82-89

Grundner, C., Gay, L. M., and Alber, T. (2005) Mycobacterium tuberculosis serine/threonine kinases PknB, PknD, PknE, and PknF phosphorylate multiple FHA domains *Protein Sci* **14**, 1918-1921

Han, G., and Zhang, C. C. (2001) On the origin of Ser/Thr kinases in a prokaryote *FEMS Microbiol Lett* **200**, 79-84

Hu, J., Liu, J., Ghirlando, R., Saltiel, A. R., and Hubbard, S. R. (2003) Structural basis for recruitment of the adaptor protein APS to the activated insulin receptor *Mol Cell* **12**, 1379-1389

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

Hubbard, S. R., Mohammadi, M., and Schlessinger, J. (1998) Autoregulatory mechanisms in protein-tyrosine kinases *Journal of Biological Chemistry* **273**, 11987-11990

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

Hunter, E. W. H. M. (1979) An activity phosphorylating tyrosine in polyoma T antigen immunoprecipitates. *Cell* **18**, 925-933

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

Inouye, S., Jain, R., Ueki, T., Nariya, H., Xu, C. Y., Hsu, M. Y., Fernandez-Luque, B. A., Munoz-Dorado, J., Farez-Vidal, E., and Inouye, M. (2000) A large family of eukaryotic-like protein Ser/Thr kinases of Myxococcus xanthus, a developmental bacterium *Microb Comp Genomics* **5**, 103-120

Iyer, G. H., Moore, M. J., and Taylor, S. S. (2005) Consequences of lysine 72 mutation on the phosphorylation and activation state of cAMP-dependent kinase J Biol Chem 280, 8800-8807

Johnson, L. N., Noble, M. E. M., and Owen, D. J. (1996) Active and inactive protein

kinases: Structural basis for regulation *Cell* **85**, 149-158
Johnson, W. C., Jr. (1990) Protein secondary structure and circular dichroism: a practical guide *Proteins* **7**, 205-214
Kang, C. M., Abbott, D. W., Park, S. T., Dascher, C. C., Cantley, L. C., and Husson, R. N. (2005) The Mycobacterium tuberculosis serine/threonine kinases PknA and PknB: substrate identification and regulation of cell shape *Genes Dev* **19**, 1692-1704
Kang, C. M., Abbott, D. W., Park, S. T., Dascher, C. C., Cantley, L. C., and Husson, R. N. (2005) The Mycobacterium tuberculosis serine/threonine kinases PknA and PknB: substrate identification and regulation of cell shape *Genes Dev* **19**, 1692-1704
Kennelly, P. J., and Potts, M. (1996) Fancy meeting you here! A fresh look at "prokaryotic" protein phosphorylation *J Bacteriol* **178**, 4759-4764
Kiianitsa, K., Solinger, J. A., and Heyer, W. D. (2003) NADH-coupled microplate photometric assay for kinetic studies of ATP-hydrolyzing enzymes with low and high specific activities *Analytical Biochemistry* **321**, 266-271
Kitchen, D. B., Decornez, H., Furr, J. R., and Bajorath, J. (2004) Docking and scoring in virtual screening for drug discovery: methods and applications *Nat Rev Drug Discov* **3**, 935-949
Knighton, D. R., Zheng, J. H., Teneyck, L. F., Xuong, N. H., Taylor, S. S., and Sowadski, J. M. (1991) Structure of a Peptide Inhibitor Bound to the Catalytic Subunit of Cyclic Adenosine-Monophosphate Dependent Protein-Kinase *Science* **253**, 414-420
Komander, D., Kular, G. S., Bain, J., Elliott, M., Alessi, D. R., and van Aalten, D. M. F. (2003) Structural basis for UCN-01 (7-hydroxystaurosporine) specificity and PDK1 (3-phosphoinositide-dependent protein kinase-1) inhibition *Biochemical Journal* **375**, 255-262
Koul, A., Choidas, A., Tyagi, A. K., Drlica, K., Singh, Y., and Ullrich, A. (2001) Serine/threonine protein kinases PknF and PknG of Mycobacterium tuberculosis: characterization and localization *Microbiology* **147**, 2307-2314
Kupcho, K., Somberg, R., Bulleit, B., and Goueli, S. A. (2003) A homogeneous, nonradioactive high-throughput fluorogenic protein kinase assay *Analytical Biochemistry* **317**, 210-217
Ladbury, J. E., and Chowdhry, B. Z. (1996) Sensing the heat: the application of isothermal titration calorimetry to thermodynamic studies of biomolecular interactions *Chem Biol* **3**, 791-801
Leavitt, S., and Freire, E. (2001) Direct measurement of protein binding energetics by isothermal titration calorimetry *Curr Opin Struct Biol* **11**, 560-566
Leonard, C. J., Aravind, L., and Koonin, E. V. (1998) Novel families of putative protein kinases in bacteria and archaea: evolution of the "eukaryotic" protein kinase superfamily *Genome Res* **8**, 1038-1047
Lew, J. (2003) MAP kinases and CDKs: Kinetic basis for catalytic activation *Biochemistry* **42**, 849-856
Li, J., Lee, G. I., Van Doren, S. R., and Walker, J. C. (2000) The FHA domain mediates phosphoprotein interactions *J Cell Sci* **113 Pt 23**, 4143-4149
Li, J., Smith, G. P., and Walker, J. C. (1999) Kinase interaction domain of kinase-associated protein phosphatase, a phosphoprotein-binding domain *Proc Natl Acad Sci U S A* **96**, 7821-7826
Li, J., Williams, B. L., Haire, L. F., Goldberg, M., Wilker, E., Durocher, D., Yaffe, M. B., Jackson, S. P., and Smerdon, S. J. (2002) Structural and functional versatility of the FHA domain in DNA-damage signaling by the tumor suppressor kinase Chk2 *Mol Cell* **9**, 1045-1054

Liao, H., Byeon, I. J., and Tsai, M. D. (1999) Structure and function of a new phosphopeptide-binding domain containing the FHA2 of Rad53 *J Mol Biol* **294**, 1041-1049

Liao, H., Yuan, C., Su, M. I., Yongkiettrakul, S., Qin, D., Li, H., Byeon, I. J., Pei, D., and Tsai, M. D. (2000) Structure of the FHA1 domain of yeast Rad53 and identification of binding sites for both FHA1 and its target protein Rad9 *J Mol Biol* **304**, 941-951

Lin, Q., Buckler, E. S. t., Muse, S. V., and Walker, J. C. (1999) Molecular evolution of type 1 serine/threonine protein phosphatases *Mol Phylogenet Evol* **12**, 57-66

Madec, E., Laszkiewicz, A., Iwanicki, A., Obuchowski, M., and Seror, S. (2002) Characterization of a membrane-linked Ser/Thr protein kinase in Bacillus subtilis, implicated in developmental processes *Mol Microbiol* **46**, 571-586

Madec, E., Stensballe, A., Kjellstrom, S., Cladiere, L., Obluchowski, M., Jensen, O. N., and Serer, S. J. (2003) Mass spectrometry and site-directed mutagenesis identify several autophosphorylated residues required for the activity of PrkC, a Ser/Thr kinase from Bacillus subtilis *Journal of Molecular Biology* **330**, 459-472

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

Margolin, W. (2005) FtsZ and the division of prokaryotic cells and organelles *Nat Rev Mol Cell Biol* **6**, 862-871

Martin, G. S. (1970) Rous sarcoma virus: a function required for the maintenance of the transformed state *Nature* **227**, 1021-1023

Miyazono, K., Maeda, S., and Imamura, T. (2004) Coordinate regulation of cell growth and differentiation by TGF-beta superfamily and Runx proteins *Oncogene* **23**, 4232-4237

Molle, V., Girard-Blanc, C., Kremer, L., Doublet, P., Cozzone, A. J., and Prost, J. F. (2003) Protein PknE, a novel transmembrane eukaryotic-like serine/threonine kinase from Mycobacterium tuberculosis *Biochem Biophys Res Commun* **308**, 820-825

Molle, V., Kremer, L., Girard-Blanc, C., Besra, G. S., Cozzone, A. J., and Prost, J. F. (2003) An FHA Phosphoprotein Recognition Domain Mediates Protein EmbR Phosphorylation by PknH, a Ser/Thr Protein Kinase from Mycobacterium tuberculosis *Biochemistry* **42**, 15300-15309

Morton, T. A., and Myszka, D. G. (1998) Kinetic analysis of macromolecular interactions using surface plasmon resonance biosensors *Methods Enzymol* **295**, 268-294

Morton, T. A., Myszka, D. G., and Chaiken, I. M. (1995) Interpreting complex binding kinetics from optical biosensors: a comparison of analysis by linearization, the integrated rate equation, and numerical integration *Anal Biochem* **227**, 176-185

Moyer, J. D., Barbacci, E. G., Iwata, K. K., Arnold, L., Boman, B., Cunningham, A., DiOrio, C., Doty, J., Morin, M. J., Moyer, M. P., Neveu, M., Pollack, V. A., Pustilnik, L. R., Reynolds, M. M., Sloan, D., Theleman, A., and Miller, P. (1997) Induction of apoptosis and cell cycle arrest by CP-358,774, an inhibitor of epidermal growth factor receptor tyrosine kinase *Cancer Research* **57**, 4838-4848

Munoz-Dorado, J., Inouye, S., and Inouye, M. (1991) A gene encoding a protein serine/threonine kinase is required for normal development of M. xanthus, a gram-negative bacterium *Cell* **67**, 995-1006

Nanninga, N. (1998) Morphogenesis of Escherichia coli *Microbiology and Molecular Biology Reviews* **62**, 110-+

Nedelkov, D., and Nelson, R. W. (2003) Surface plasmon resonance mass spectrometry: recent progress and outlooks *Trends Biotechnol* **21**, 301-305

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

Novakova, L., Saskova, L., Pallova, P., Janecek, J., Novotna, J., Ulrych, A., Echenique, J., Trombe, M. C., and Branny, P. (2005) Characterization of a eukaryotic type serine/threonine protein kinase and protein phosphatase of Streptococcus pneumoniae and identification of kinase substrates *Febs Journal* **272**, 1243-1254

Nguyen, L., Walburger, A., Houben, E., Koul, A., Muller, S., Morbitzer, M., Klebl, B., Ferrari, G., Pieters, J. (2005) Role of protein kinase G in growth and glutamine metabolism of *Mycobacterium bovis* BCG *Journal of Bacteriology* **20**, 7165

Obuchowski, M., Madec, E., Delattre, D., Boel, G., Iwanicki, A., Foulger, D., and Seror, S. J. (2000) Characterization of PrpC from Bacillus subtilis, a member of the PPM phosphatase family *Journal of Bacteriology* **182**, 5634-5638

Ogawa, A., Takayama, Y., Sakai, H., Chong, K. T., Takeuchi, S., Nakagawa, A., Nada, S., Okada, M., and Tsukihara, T. (2002) Structure of the carboxyl-terminal Src kinase, Csk *Journal of Biological Chemistry* **277**, 14351-14354

Ogawara, H., Aoyagi, N., Watanabe, M., and Urabe, H. (1999) Sequences and evolutionary analyses of eukaryotic-type protein kinases from Streptomyces coelicolor A3(2) *Microbiology-Sgm* **145**, 3343-3352

Ortiz-Lombardia, M., Pompeo, F., Boitel, B., and Alzari, P. M. (2003) Crystal structure of the catalytic domain of the PknB serine/threonine kinase from Mycobacterium tuberculosis *J Biol Chem* **278**, 13094-13100

Ozes, O. N., Mayo, L. D., Gustin, J. A., Pfeffer, S. R., Pfeffer, L. M., and Donner, D. B. (1999) NF-kappaB activation by tumour necrosis factor requires the Akt serine-threonine kinase Nature 401, 82-85

Pallen, M., Chaudhuri, R., and Khan, A. (2002) Bacterial FHA domains: neglected players in the phospho-threonine signalling game? *Trends Microbiol* **10**, 556-563

Pawson, T. (2004) Specificity in signal transduction: from phosphotyrosine-SH2 domain interactions to complex cellular systems *Cell* **116**, 191-203

Pawson, T., and Nash, P. (2003) Assembly of cell regulatory systems through protein interaction domains *Science* **300**, 445-452

Peirs, P., Parmentier, B., De Wit, L., and Content, J. (2000) The Mycobacterium bovis homologous protein of the Mycobacterium tuberculosis serine/threonine protein kinase Mbk (PknD) is truncated *FEMS Microbiol Lett* **188**, 135-139

Perkins, D. N., Pappin, D. J. C., Creasy, D. M., and Cottrell, J. S. (1999) Probability-based protein identification by searching sequence databases using mass spectrometry data *Electrophoresis* **20**, 3551-3567

Petrickova, K., and Petricek, M. (2003) Eukaryotic-type protein kinases in Streptomyces coelicolor: variations on a common theme *Microbiology* **149**, 1609-1621

Ponting, C.P., Aravind, L., Schultz, J., Bork, P., Koonin, E.V. (1999) Eukaryotic signalling domains homologues in archaea and bacteria. Anicient ancestry and horizontal gene transfer. *J Mol Biol* **4**, 729-45

Popham, D. L., and Young, K. D. (2003) Role of penicillin-binding proteins in bacterial cell morphogenesis *Curr Opin Microbiol* **6**, 594-599

Prabhakaran, K., Harris, E. B., and Randhawa, B. (2000) Regulation by protein kinase of phagocytosis of Mycobacterium leprae by macrophages *J Med Microbiol* **49**, 339-342

Rajagopal, L., Clancy, A., and Rubens, C. E. (2003) A eukaryotic type serine/threonine kinase and phosphatase in Streptococcus agalactiae reversibly phosphorylate an inorganic pyrophosphatase and affect growth, cell segregation, and virulence *Journal of Biological Chemistry* **278**, 14429-14441

Rivero-Cruz I, A. L., Guerrero JA, Martinez S, Bye R, Pereda-Miranda R, Franzblau S, Timmermann BN, Mata R. (2005) Antimycobacterial agents from selected Mexican medicinal plants. *J Pharm Pharmacol.* **57**, 1117-1126

Roskoski, R., Jr. (1983) Assays of protein kinase *Methods Enzymol* **99**, 3-6

Sahara, S., Sato, K., Kaise, H., Mori, K., Sato, A., Aoto, M., Tokmakov, A. A., and Fukami, Y. (1996) Biochemical evidence for the interaction of regulatory subunit of cAMP-dependent protein kinase with IDA (Inter-DFG-APE) region of catalytic subunit *FEBS Lett* **384**, 138-142

Sambrook, J. F. E. F. M. (1989) *Molecular Cloning. A Laboratory Manual.* Cold Spring Harbor Laboratory Press.

Sassetti, C. M., Boyd, D. H., and Rubin, E. J. (2003) Genes required for mycobacterial growth defined by high density mutagenesis *Molecular Microbiology* **48**, 77-84

Seeliger, M. A., Young, M., Henderson, M. N., Pellicena, P., King, D. S., Falick, A. M., and Kuriyan, J. (2005) High yield bacterial expression of active c-Abl and c-Src tyrosine kinases *Protein Sci*

Sharma, K., Chandra, H., Gupta, P. K., Pathak, M., Narayan, A., Meena, L. S., D'Souza, R. C., Chopra, P., Ramachandran, S., and Singh, Y. (2004) PknH, a transmembrane Hank's type serine/threonine kinase from Mycobacterium tuberculosis is differentially expressed under stress conditions *FEMS Microbiol Lett* **233**, 107-113

Sicheri, F., Moarefi, I., and Kuriyan, J. (1997) Crystal structure of the Src family tyrosine kinase Hck *Nature* **385**, 602-609

Srivastava, A. K. (1985) Inhibition of phosphorylase kinase, and tyrosine protein kinase activities by quercetin *Biochem Biophys Res Commun* **131**, 1-5

Sreevatsan, S., Stockbauer, K. E., Pan, X., Kreiswirth, B. N., Moghazeh, S. L., Jacobs, W. R., Jr., Telenti, A., and Musser, J. M. (1997) Ethambutol resistance in Mycobacterium tuberculosis: critical role of embB mutations *Antimicrob Agents Chemother* **41**, 1677-1681

Stavridi, E. S., Huyen, Y., Loreto, I. R., Scolnick, D. M., Halazonetis, T. D., Pavletich, N. P., and Jeffrey, P. D. (2002) Crystal structure of the FHA domain of the Chfr mitotic checkpoint protein and its complex with tungstate *Structure (Camb)* **10**, 891-899

Stols, L., Gu, M., Dieckman, L., Raffen, R., Collart, F. R., and Donnelly, M. I. (2002) A new vector for high-throughput, ligation-independent cloning encoding a tobacco etch virus protease cleavage site *Protein Expr Purif* **25**, 8-15

Sun, H. Y., Low, K. E., Woo, S., Noble, R. L., Graham, R. J., Connaughton, S. S., Gee, M. A., and Lee, L. G. (2005) Real-time protein kinase assay *Analytical Chemistry* **77**, 2043-2049

Sun, Z., Hsiao, J., Fay, D. S., and Stern, D. F. (1998) Rad53 FHA domain associated with phosphorylated Rad9 in the DNA damage checkpoint *Science* **281**, 272-274

Tascon, R. E., Soares, C. S., Ragno, S., Stavropoulos, E., Hirst, E. M. A., and Colston, M. J. (2000) Mycobacterium tuberculosis-activated dendritic cells induce protective immunityin mice *Immunology* **99**, 473-480

Villarino, A., Duran, R., Wehenkel, A., Fernandez, P., England, P., Brodin, P., Cole, S. T., Zimny-Arndt, U., Jungblut, P. R., Cervenansky, C., and Alzari, P. M. (2005) Proteomic Identification of M.tuberculosis Protein Kinase Substrates: PknB Recruits GarA, a FHA Domain-containing Protein, Through Activation Loop-mediated Interactions *J Mol Biol* **350**, 953-963

Vorm, O., and Mann, M. (1994) Improved Mass Accuracy in Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Mass-Spectrometry of Peptides *Journal of the American Society for Mass Spectrometry* **5**, 955-958

Walburger, A., Koul, A., Ferrari, G., Nguyen, L., Prescianotto-Baschong, C., Huygen,

K., Klebl, B., Thompson, C., Bacher, G., and Pieters, J. (2004) Protein kinase G from pathogenic mycobacteria promotes survival within macrophages *Science* **304**, 1800-1804

Wang, J. Y., Li, C. H., Yang, C. J., Mushegian, A., and Jin, S. G. (1998) A novel serine/threonine protein kinase homologue of Pseudomonas aeruginosa is specifically inducible within the host infection site and is required for full virulence in neutropenic mice *Journal of Bacteriology* **180**, 6764-6768

Wang, P., Byeon, I. J., Liao, H., Beebe, K. D., Yongkiettrakul, S., Pei, D., and Tsai, M. D. (2000) II. Structure and specificity of the interaction between the FHA2 domain of Rad53 and phosphotyrosyl peptides *J Mol Biol* **302**, 927-940

Weldingh, K., and Andersen, P. (1999) Immunological evaluation of novel Mycobacterium tuberculosis culture filtrate proteins *Fems Immunology and Medical Microbiology* **23**, 159-164

Weldingh, K., Rosenkrands, I., Jacobsen, S., Rasmussen, P. B., Elhay, M. J., and Andersen, P. (1998) Two-dimensional electrophoresis for analysis of Mycobacterium tuberculosis culture filtrate and purification and characterization of six novel proteins *Infection and Immunity* **66**, 3492-3500

Winzor, D. J. (2003) Surface plasmon resonance as a probe of protein isomerization *Anal Biochem* **318**, 1-12

Wiseman, T., Williston, S., Brandts, J. F., and Lin, L. N. (1989) Rapid measurement of binding constants and heats of binding using a new titration calorimeter *Anal Biochem* **179**, 131-137

Wolanin, P. M., Thomason, P. A., and Stock, J. B. (2002) Histidine protein kinases: key signal transducers outside the animal kingdom *Genome Biol* **3**, REVIEWS3013

Wolf, I., Rubinfeld, H., Yoon, S., Marmor, G., Hanoch, T., and Seger, R. (2001) Involvement of the activation loop of ERK in the detachment from cytosolic anchoring *Journal of Biological Chemistry* **276**, 24490-24497

Woody, R. W., Sugeta, H., and Kodama, T. S. (1996) [Circular dichroism of proteins: recent developments in analysis and prediction] *Tanpakushitsu Kakusan Koso* **41**, 56-69

Wu, J. W., Hu, M., Chai, J., Seoane, J., Huse, M., Li, C., Rigotti, D. J., Kyin, S., Muir, T. W., Fairman, R., Massague, J., and Shi, Y. (2001) Crystal structure of a phosphorylated Smad2. Recognition of phosphoserine by the MH2 domain and insights on Smad function in TGF-beta signaling *Mol Cell* **8**, 1277-1289

Wybenga-Groot, L. E., Baskin, B., Ong, S. H., Tong, J., Pawson, T., and Sicheri, F. (2001) Structural basis for autoinhibition of the Ephb2 receptor tyrosine kinase by the unphosphorylated juxtamembrane region *Cell* **106**, 745-757

Yaffe, M. B., and Elia, A. E. (2001) Phosphoserine/threonine-binding domains *Curr Opin Cell Biol* **13**, 131-138

Yaffe, M. B., and Smerdon, S. J. (2001) PhosphoSerine/threonine binding domains: you can't pSERious? *Structure (Camb)* **9**, R33-38

Yaffe, M. B., and Smerdon, S. J. (2004) The use of in vitro peptide-library screens in the analysis of phosphoserine/threonine-binding domain structure and function *Annu Rev Biophys Biomol Struct* **33**, 225-244

Yang, J. T., Wu, C. S., and Martinez, H. M. (1986) Calculation of protein conformation from circular dichroism *Methods Enzymol* **130**, 208-269

Yeats, C., Finn, R. D., and Bateman, A. (2002) The PASTA domain: a beta-lactam-binding domain *Trends Biochem Sci* **27**, 438

Young, T. A., Delagoutte, B., Endrizzi, J. A., Falick, A. M., and Alber, T. (2003) Structure of Mycobacterium tuberculosis PknB supports a universal activation mechanism for Ser/Thr protein kinases *Nat Struct Biol* **10**, 168-174

Zhang, C. C., and Libs, L. (1998) Cloning and characterisation of the pknD gene

encoding an eukaryotic-type protein kinase in the cyanobacterium Anabaena sp. PCC7120 *Molecular and General Genetics* **258**, 26-33

Zhang, J. H., Chung, T. D., and Oldenburg, K. R. (1999) A Simple Statistical Parameter for Use in Evaluation and Validation of High Throughput Screening Assays *J Biomol Screen* **4**, 67-73

11 Appendix

Appendix 1

Å	Ångström (0.1 nm)
CAPS	3-[cyclohexylamino]-propanesulphonic acid
DTT	1,4-Dithioerythritol
E. coli	Escherichia coli
EDTA	Ehtylendiamintetraacetate
FHA	Forkhead-associated (domain)
GPCR	G-protein coupled receptors
GSH	Reduced glutathione
GST	Glutathione-S-transferase
GTPase	GTP-hydrolysing Enzyme
HEPES	4-2-hydroxyethyl-1-piperazineethanesulfonic acid
HPLC	High performance liquid chromatography
IPTG	Isopropyl-β-D-1-thiogalactopyranoside
OD_X	Optical density at X nm wavelength of light
Q-sepharose	Quaternary ammonium sepharose
SDS-PAGE	Sodiumdodecylsulfate-Polyacrylamide gelelectrophoresis
SH2	*Src* homology 2 (domain)
SP-sepharose	sulphur propyl sepharose
STPK	Serine/Threonine protein kinase
ITC	Isothermal Titration Calorimetry
SPR	Surface Plasmon Resonance
MALDI/TOF	Matrix-assisted laser desorption/ionisation/time-of-flight
Tris	Tris-(hydroxymethyl)-aminomethan
λ	Wavelength

Appendix 2

Sigma Cat No.	MW	Name	Description	%activity assay 2	%activity assay 1
D 2064	766.6	Dequalinium analog, C-14 linker	Protein kinase C-alpha (PKC-alpha) inhibitor	19.19	90.91
D 8065	168.15	Dephostatin	CD45 protein tyrosine kinase inhibitor	54.68	80.72
G 6416	520.95	GW5074	cRaf1 kinase inhibitor	40.67	67.93
N 1786	310.39	NSC 95397	'Selective, irreversible Cdc25 dual specificity phosphatase inhibitor.	-0.65	79.25
M 6760	318.24	Myricetin	Casein Kinase II inhibitor	31.26	34.91
M 9440	168.15	Me-3,4-dephostatin	Selective inhibitor of protein tyrosine phosphatase 1B and SHPTP-1	14.31	82.79
N 5023	302.37	Nordihydroguaiaretic acid from Larrea divaricata (creosote bush)	Lipoxygenase inhibitor	-7.78	21.15
C 3930	687.71	Calmidazolium chloride	Potent inhibitor of calmodulin activation of phosphodiesterase; strongly inhibits calmodulin-dependent Ca2+-ATPase	59.58	89.93
S 5567	220.23	'SP600125	'Selective c-Jun N-terminal kinase (c-JNK) inhibitor.	60.93	60.23
M 6545	517.41	Mitoxantrone	DNA synthesis inhibitor	60.51	70.93
D 5439	218.21	2,3-Dimethoxy-1,4-naphthoquinone	Redox cycling agent used to study role of ROS (reactive oxygen species)	19.38	97.14
A 1895	422.35	Aurintricarboxylic acid	DNA topoisomerase II inhibitor	15.38	5.10
L 2037	242.28	beta-Lapachone	Induces apoptosis in HL-60 cells; anticancer agent	6.23	74.37
H 2380	213.19	6-Hydroxy-DL-DOPA	Precursor of the catecholaminergic neurotoxin, 6-hydroxydopamine; converted to 6-hydroxydopamine by L-aromatic amino acid decarboxylase	5.35	32.86

Appendix 3

Name	Sequence	Vector	Direction	Protein
PknA T172Ab	caccatgccggtctgggCcacgggcgctgc atc	Topo 151	Reverse	PknA
PknA T172Af	gatgcagcgcccgtgGcccagaccggcatg gtg	Topo 151	Forward	PknA
PknA T174Ab	gcccatcaccatgccggCctgggtcacggg cgc	Topo 151	Reverse	PknA
PknA T174Af	gcgcccgtgacccagGccggcatggtgatg ggc	Topo 151	Forward	PknA
PknA T174Sb	gcccatcaccatgccgCtctgggtcacgggc gc	Topo 151	Reverse	PknA
PknA T174Sf	gcgcccgtgacccagaGcggcatggtgatg ggc	Topo 151	Forward	PknA
PknA290 T268Ab	gcc agg ggg ggg tgC ctg gct ggg ccg cgg	Topo 151	Reverse	PknA
PknA290 T286Af	ccg cgg ccc agc cag Gca ccc ccc cct ggc	Topo 151	Forward	PknA
PknA290b	gccaggggggggtgtctg	pET 22	Reverse	PknA
PknA290b	cgccgaattctca ctcgac gccaggggggggtgtctggct	pGEX 6p2	Reverse	PknA
PknA290b	cgccaagctttca gccaggggggggtgtctggct	pET 42	Reverse	PknA
PknA290b	cgcctca tgcggccgc gccaggggggggtgtctggct	pET 22	Reverse	PknA
PknA290f	tcacccggtgtcctgggcgac	pET 22	Reverse	PknA
PknA290f	ccgggatcc catatg agcccccgagttggcgtgacg	pGEX 6p2	Forward	PknA
PknA290f	cgccaagctttca gccaggggggggtgtctggct	pET 42	Forward	PknA
PknA290f	cgccactagt ctgggagttctgttccaggggccc	pET 42	Forward	PknA
PknA290f	caccgaattc ctgggagttctgttccaggggccc	pET 42	Forward	PknA
PknB T171A173Ab	gccgatcactgctgcgCtctggCtcacgctg tt	pLIC	Reverse	PknB
PknB T171A173Af	aacagcgtgaGccagaGcgcagcagtgatc ggc	pLIC	Forward	PknB
PknB T171Ab	cactgctgcggtctgggCcacgctgttgccg ct	pLIC	Reverse	PknB
PknB T171Ab	cactgctgcggCctgggCcacgctgttgcc gct	pLIC	Reverse	PknB
PknB T171Af	agcggcaacagcgtgGcccagaccgcagc agtg	pLIC	Forward	PknB
PknB T171Af	agcggcaacagcgtgGcccagGccgcagc agtg	pLIC	Forward	PknB
PknB T173Ab	gccgatcactgctgcggCctgggtcacgctg tt	pLIC	Reverse	PknB
PknB T173Af	aacagcgtgacccagGccgcagcagtgatc ggc	pLIC	Forward	PknB
PknB T173Sb	gccgatcactgctgcgCtctgggtcacgctgt t	pLIC	Reverse	PknB
PknB T173Sf	aacagcgtgacccagaGcgcagcagtgatc ggc	pLIC	Forward	PknB

Name	Sequence	Vector	Direction	Protein
Rv0019b	tcacgggcgcaactcgat	Topo 151	Reverse	Rv0019
Rv0019b	cgccgaattctcacgggcgcaactcgat	pGEX 6p2	Reverse	Rv0019
Rv0019f	caccCTGGAAGTTCTGTTCCA GGGGCCCcagcgccggcacgctgca	Topo 151	Forward	Rv0019
Rv0019f	caccggatcccagcgccggcacgctgca	pGEX 6p2	Forward	Rv0019
Rv0020 S473Ab	caa gtg acg gcg TGC cac acc ggt gtc	pGEX 6p1	Reverse	Rv0020
Rv0020 S473Af	gac acc ggt gtg GCA cgc cgt cac ttg	pGEX 6p1	Forward	Rv0020
Rv0020Lf	cacc gaattc atgagtgacaattcgagc	pGEX 6p1	Forward	Rv0020
Rv0020Lr	cgcc ctcgag tcagtgcatgcggacgat	pGEX 6p1	Forward	Rv0020
Rv1827N117Af	gat gtc ggg agt ctc GCc ggc acc tac gtc aac	pGEX 6p1	Forward	Rv1827
Rv1827N117Ar	gtt gac gta ggt gcc gGC gag act ccc gac atc	pGEX 6p1	Reverse	Rv1827
Rv1827R81Af	atc acg tcg gct ggt GCg cat ccc gac agc gac	pGEX 6p1	Forward	Rv1827
Rv1827R81Af	gtc gct gtc ggg atg cGC acc agc cga cgt gat	pGEX 6p1	Reverse	Rv1827
Rv1827S95Af	gac gac gtg acc gtg GCc cgt cgc cat gct gaa	pGEX 6p1	Forward	Rv1827
Rv1827S95Ar	ttc agc atg gcg acg gGC cac ggt cac gtc gtc	pGEX 6p1	Reverse	Rv1827