The Regulation and Inhibition of P-TEFb

Alison Hole

Department of Biochemistry
And
The Queen's College
University of Oxford

A thesis submitted in partial fulfilment for the requirements of the degree of Doctor of Philosophy at the University of Oxford

Trinity 2011

The Regulation and Inhibition of P-TEFb

Alison Hole

Department of Biochemistry and The Queen's College, University of Oxford
submitted in partial fulfilment for the requirements of the degree of Doctor of Philosophy

Trinity 2011

Abstract

Correct regulation of transcription is essential for maintaining a healthy cellular state. During transcription RNA polymerase II (Pol II) proceeds in a regulated manner through several transitions to ensure appropriate control of synthesis and enable correct processing of the pre-RNA. Shortly after initiation Pol II is caused to pause by the binding of factors, DSIF and NELF. To enable transition of Pol II into the elongation phase CDK9/cyclin T phosphorylates the C-terminal domain (CTD) of Pol II, DSIF and NELF. This phosphorylation releases the paused state and provides an alternative set of post-transcriptional modifications on the CTD to generate a binding platform for elongation, histone modifying and termination factors. CDK9/cyclin T is itself regulated within multicomponent complexes. A small activated complex, containing Brd4, recruits CDK9/cyclin T to active sites of transcription, thereby promoting the elongation of transcription. The role of CDK9/cyclin T in the regulation of transcription has resulted in its validation as a drug target against several disease states including cancer, HIV and cardiac hypertrophy.

In this thesis, I present the crystallographic structures of a series of 2-amino-4-heteroaryl-pyrimidine compounds and the roscovitine derivative, (S)-CR8, bound to CDK9/cyclin T and CDK2/cyclin A. In combination with thermal denaturation data and kinetic analysis, these structures have suggested chemical modifications that might be made to increase the CDK9 specificity of these compounds. I have also validated the use of a mutated form of cyclin T for use in the development of CDK9/cyclin T inhibitors.

In addition, I present both structural and kinetic analysis of the Brd4-CDK9/cyclin T interaction. I show that C-terminal fragments of Brd4 enhance the *in vitro* kinase activity of CDK9/cyclin T against the Pol II CTD. Furthermore, I demonstrate that this enhancement may be inhibited by Plk1-mediated phosphorylation of Brd4. Finally, I show that Brd4 binds to a site that spans CDK9 and cyclin T and I propose detailed molecular models of the Brd4-cyclin T interaction.

Acknowledgements

My thanks go to my supervisors, Martin Noble and Jane Endicott. They gave me the freedom to follow my own lines of research and provided enthusiasm and support when my experiments began to become a little too overwhelming.

There are two people whom I cannot thank enough for their expertise and encouragement. Firstly to Sonja Baumli to whom I am grateful for all the advice, support, supervision and time she gave me throughout the 4 years. Secondly, I would like to thank Christina Redfield who helped me set up and understand NMR experiments, but more importantly kept my project and motivation going when all other experiments were struggling to work.

Several people helped me with specific techniques and I am grateful to them all. They include Camilla Oxley, Christiane Riedinger and Jo Claridge who all helped me with NMR. I had many discussions with James Parker about RNA work and would also like to thank him for collecting several crystallography datasets. Many thanks to Paul Emsley for writing Coot and then showing me the tricks it can do! I have tried many more biophysical techniques than mentioned within my thesis, and I would like to thank David Staunton for his patience and his help with all of them. Ed Lowe has kindly spent many hours with me on synchrotron trips and has always been willing to answer my crystallography questions. And finally, thanks to Nick Brown for sharing some of his wealth of knowledge on all things kinase related.

I would like to thank our scientific collaborators. Shudong Wang and Laurent Meijer kindly provided the inhibitors used in this work. Stefan Knapp and Panagis Filippakopoulos, provided constructs of Brd4 and CyclinT2 and Utz Fischer provided Larp7 constructs. I would also like to thank the beamline scientists at the Diamond Light Source. I am grateful to the Wellcome Trust for their funding.

Before the start of my DPhil there were several people who motivated me in science and taught me much about approaching experiments, and life as a scientist. So to Dr Price, Mr Gajadharsingh and Dr McConnaughie thank you. My thanks go to Phil Green and Peter Kilby who let me loose in a “real” lab for the first time. And thank you to Aude Echalier who supervised me in the LMB initially. And by no means least, thank you to Penny, who has guided me with many words of wisdom over the years.

The LMB has been a great place to be during my DPhil. And that is thanks to all the people within it. In no particular order, Maria, Ailsa, Elizabeth, Muhan, Kate, Livia, Eachan, John, Eneida, Max, Burcu, and Matt - thank you for putting up with me and my random questions. And to all the students who went before me thank you for proving that a thesis can be written.

Thank you Steve for the hug whenever I needed one!

And finally my thanks go to my family, Ruth, Philip, David, Patrick, Melanie, Ele, Lawrence and Meredith for their constant support and love.

Abbreviations

A pocket	Adenine binding pocket
ADP	Adenosine diphosphate
AhR	Aryl hydrocarbon receptor
AIC	Akaike's Informative Criteria
AML	Acute myeloid leukemia
AMP-PNP	Adenylyl-imidodiphosphate
ANOVA	Analysis of variance
ATP	Adenosine triphosphate
BD	Bromodomain
BET family	Bromodomain and extraterminal family
Brd	Bromodomain containing protein
BRE	TFIIB recognition element
C-terminal	Carboxy-terminal
CAK	CDK activating kinase
CCP4	Collaborative computational project number 4
CDK	Cyclin dependent kinase
CID	CTD interacting domains
CIITA	Class II transactivator
CKI	CDK inhibitory proteins
CML	Chronic myelogenous leukemia
CTD	C-terminal domain of RBP1 of RNA polymerase II
CTR	C-terminal region of Brd4
Cyc	Cyclin
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DRB	5,6-dichloro-1-b-D-ribofuranosylbenzimidazole
DSF	Differential scanning fluorimetry
DSIF	DRB sensitivity inducing factor
DSS	4,4-dimethyl-4-silapentane-1-sulfonic acid
ΔT_m	Difference in T_m
DTT	Dithiothreitol
E. coli	Escherichia coli
ESI-MS	Electrospray ionization mass spectrometry
ET	Extra-terminal
FL	Full length
GST	Glutathione-s-transferase
H-bonding	Hydrogen bonding
HAT	Histone acetyltransferase
HBS	HEPES buffered saline
HDAC	Histone deacetylase
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
Hexim1	HMBA inducible protein 1
HIV	Human immunodeficiency virus
hLa	Human La protein

HMBA	Hexamethylene bisacetamide
HNSCC	Head and neck squamous cell carcinoma
HPV	Human papilloma virus
HSQC	Heteronuclear Single Quantum Coherence
HSV	Herpes simplex virus
HUNK1	Hormonally up-regulated neu tumor-associated kinase 1
HXMS	Hydrogen/deuterium exchange with mass spectrometry
IC_{50}	Half maximal inhibitory concentration
Inr	Initiator element
IPTG	Isopropyl β-D-1-thiogalactopyranoside
K_d	Dissociation constant
K_i	Dissociation constant for inhibitors
K_m	Michaelis constant
Larp7	La related protein 7
LB	Lysogeny broth
LTR	Long terminal repeat
MAT1	Ménage a trois 1
MBP	Maltose binding protein
MCAP	Mitotic chromosome-associated protein
MePCE	Methylphosphate capping enzyme
MES	2-(N-morpholino)ethanesulfonic acid
MHC	Major histocompatibility complex
miRNA	microRNA
MOPS	3-(N-morpholino)propanesulfonic acid
MPP^+	1-methyl-4-phenylpyridinium
mRNA	Messenger RNA
MTG	Monothioglycerol
MWCO	Molecular weight cut-off
N-terminal	Amino-terminal
NaK-phosphate	Sodium potassium phosphate buffer
NELF	Negative elongation factor
NMR	Nuclear magnetic resonance
NOESY	Nuclear Overhauser enhancement spectroscopy
NUT	Nuclear protein in testis
O-GlcNAc	O-linked N-acetylglucosamine
OD_x	Optical density at x nm
ORF	Open reading frame
P pocket	Phosphate binding pocket
P-TEFb	Positive transcription elongation factor b
PCR	Polymerase chain reaction
PDB	Protein data bank
PEG	Polyethylene glycol
PKA	Protein kinase A
Plk1	Polo-like kinase 1
Pol I	RNA polymerase I
Pol II	RNA polymerase II
Pol III	RNA polymerase III
Pre-mRNA	Precursor mRNA

R pocket	Ribose binding pocket
RNA	Ribonucleic acid
RRM	RNA recognition motif
rRNA	Ribosomal RNA
S. cerevisiae	Saccharomyces cerevisiae
SAR	Structure-activity relationship
SDS	sodium dodecyl sulfate
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SEC	Super elongation complex
Ser2	Serine 2 residues of the CTD
Ser5	Serine 5 residues of the CTD
Ser7	Serine 7 residues of the CTD
Sf9	Spodoptera frugiperda 9
snRNA	Small nuclear RNA
snRNP	Small nuclear ribonucleoproteins
STAT3	Signal transducers and activators of transcription 3
TAR	Trans-activation response element
Tat	Trans-activator of transcription
TB	Terrific broth
TCEP	tris(2-carboxyethyl)phosphine
TEV	Tobacco Etch Virus
TF	Transcription factor
T_m	Melting temperature
TOCSY	Total Correlation Spectroscopy
Tris	2-amino-2-(hydroxymethyl)-1,3-propanediol
tRNA	Transfer RNA
V_{max}	Maximum velocity under substrate concentrations tested

Table of Contents

Table of figures

Table of Tables

1 Introduction

Biology relies on regulated gene expression. This is essential in processes ranging from development and differentiation to apoptosis. The importance of correct gene expression can sadly be seen in cases where either promoter or transcriptional proteins have been mutated, resulting in human diseases such as certain thalassemias, cancer and congenital heart disease (Maston et al. 2006).

There are three different DNA-dependent RNA polymerases that are responsible for the production of nascent RNAs in higher eukaryotes. These are RNA polymerase I (Pol I), RNA polymerase II (Pol II) and RNA polymerase III (Pol III). Each has a separate function transcribing different classes of RNAs in a regulated manner. Pol I is dedicated to synthesizing ribosomal RNA, whilst Pol III transcribes several RNAs including tRNA, 5S rRNA, U6 snRNA and 7SK RNA. Pol II is responsible for transcribing protein-encoding genes, small nuclear RNAs (snRNAs) and microRNAs (miRNAs).

There are 4 main stages in the Pol II transcription cycle: the formation of the preinitiation complex, initiation, elongation and termination. Each stage involves many different proteins to ensure both highly regulated and successful production of fully synthesized and modified RNA.

1.1 Transcription Stages

Preinitiation complex formation and initiation (Fig 1-1, Stages 1 & 2)

The first stage in transcription is the formation of a pre-initiation complex at the promoter. Core promoter sequences, such as the TATA box, initiator element (Inr) and the transcription factor IIB (TFIIB) recognition element (BRE), recruit either TFIID or TFIIB. This induces further recruitment and assembly of RNA polymerase II, in complex with the general transcription factors (TFIIA, TFIIB, TFIID, TFIIE, TFIIF and TFIIH), at the promoter region. This complex is the preinitiation complex (PIC) (reviewed in (Baumann et al. 2010)). Alternatively CpG islands can initiate the formation of the PIC as they often contain several transcription factor binding sites. Further control of this process occurs when combinations of additional cis-acting DNA elements such as enhancers, promoter-proximal elements and silencers are present which recruit various activators or repressors. This enables unique spatial and temporal control of transcription initiation at specific genes (for reviews, see (Maston et al. 2006; Heintzman and Ren 2009)).

On forming the pre-initiation complex, TFIIE and TFIIH melt the DNA duplex, allowing RNA polymerase II access to the single stranded template DNA. This initiates the formation of the first phosphodiester bond and therefore the start of RNA synthesis (Holstege et al. 1996) (Kostrewa et al. 2009).

Elongation (Fig 1-1, Stages 3 & 4)

After 20-50 base pairs have been transcribed, the polymerase is paused by recruitment of two factors: negative elongation factor (NELF) and DRB

sensitivity inducing factor (DSIF) (Wu et al. 2003a). These bind to Pol II and the nascent RNA (Missra and Gilmour 2010). Following phosphorylation by positive transcription elongation factor b (P-TEFb), NELF is released and DSIF becomes a positive elongation factor (Yamada et al. 2006). RNA polymerase continues to track along the gene synthesizing nascent mRNA. Post-transcriptional modifications of the RNA polymerase and associated factors during these transitions from initiation to elongation are discussed in later sections.

Termination and 3' end formation (Fig 1-1, Stage 5)

At the 3' end of the gene, once the RNA polymerase has transcribed the polyadenylation signal, endonucleolytic cleavage and polyadenylation of the nascent RNA occurs. The actual separation of RNA polymerase II from the DNA coding strand in eukaryotes is not fully understood. Currently two models have been proposed: the torpedo model and the allosteric model. The torpedo model involves a 5'-3' exonuclease (Xrn2) that, after nascent RNA cleavage at the poly(A) site, degrades the uncapped RNA, progresses towards the elongating Pol II and subsequently dissociates Pol II from the DNA. The allosteric model proposes that a termination factor is utilized that dissociates the RNA polymerase from the DNA by inducing conformational changes in the elongation complex (Gilmour and Fan 2008; Richard and Manley 2009). It is proposed that transcription termination involves a combination of these models (Banerjee et al. 2009).

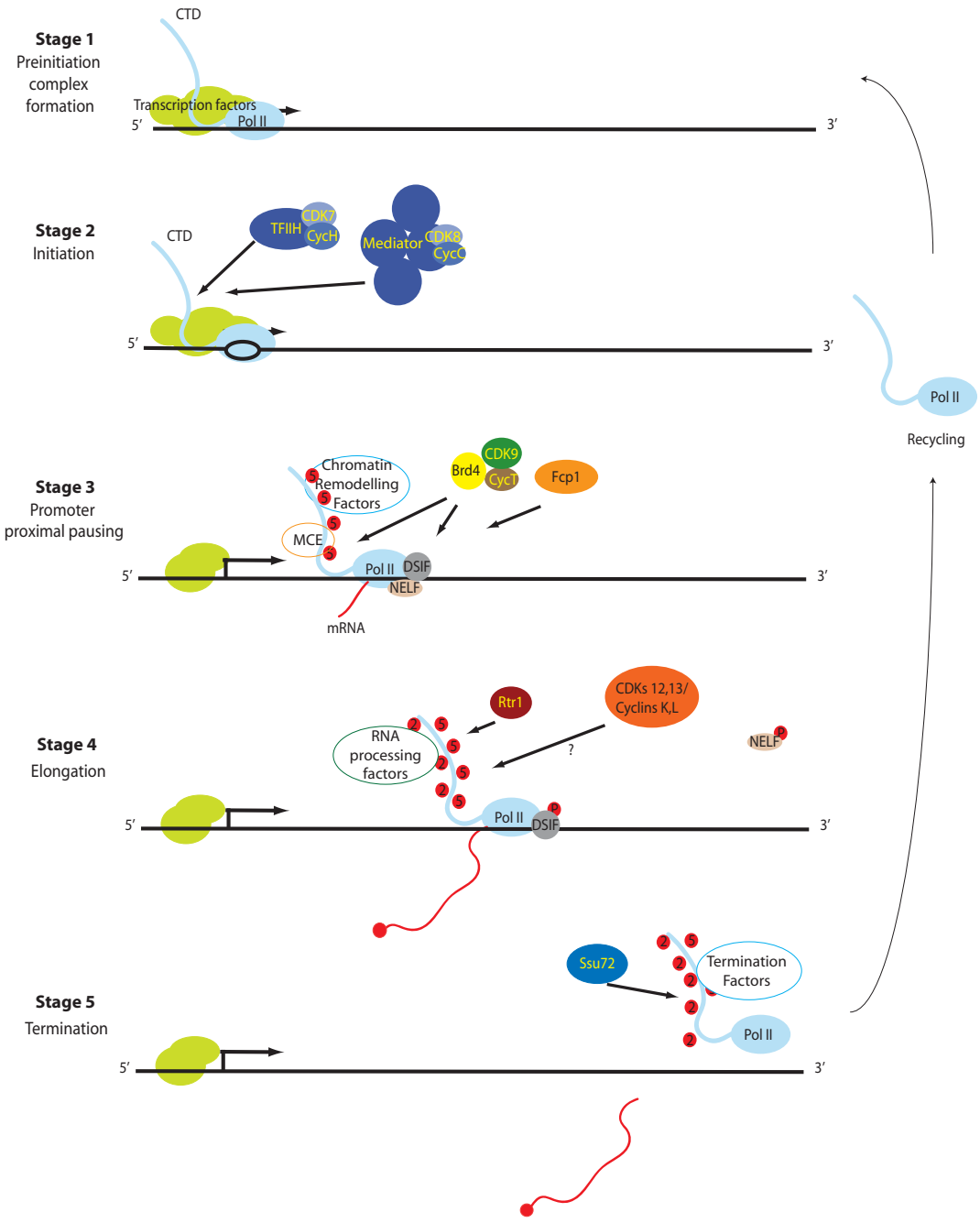


Figure 1-1. The stages of RNA polymerase II transcription in metazoa. The different stages of transcription are depicted with focus on the kinases and phosphatases regulating the phosphorylation state of the C-terminal domain (CTD) of the RPB1 protein subunit of RNA polymerase II Stage1: The preinitiation complex is formed by the recruitment of Pol II by general transcription factors. Stage 2: During initiation the DNA duplex is melted and CDKs 8 and 7, as parts of the mediator and TFIIH complex, respectively, can phosphorylate Ser5 of the heptad repeat of the CTD in Pol II. Pol II is released from the promoter region and the nascent mRNA is synthesised. Stage 3: After around 50 base pairs, Pol II is paused by NELF and DSIF. The capping enzyme, MCE, and chromatin remodelling factors are recruited to the Ser5 phosphorylation site. P-TEFb is recruited, possibly by binding to Brd4, to phosphorylate the negative elongation factors and Ser2 of the CTD. Stage 4: NELF is released and DSIF remains bound to the Pol II where it acts as a positive elongation factor. At this point the phosphatase Fcp1 is thought to be present (Nedea et al. 2003). Pol II is released from the paused state and elongation occurs. Phosphorylation of the Ser2 residues of the CTD recruits RNA processing factors during this stage. CDKs 12 and 13 may phosphorylate the CTD concurrently, while Rtr1 dephosphorylates the CTD. Stage 5: Termination factors bind to the CTD when Ser2 residues are more frequently phosphorylated. Pol II is then released and transcription terminates. Ssu72 and other phosphatases are then proposed to dephosphorylate the CTD of Pol II to enable recycling and reinitiation of the complex for future cycles of transcription.

1.2 Phosphorylation in transcription

A primary method of regulating transcription is through altering the post-translational modifications of proteins involved in transcription. In particular the phosphorylation state of certain proteins has a major influence on transcription. Phosphorylation of proteins is controlled by a combination of kinases (which catalyse the addition of a phosphoryl group) and phosphatases (which catalyse hydrolysis of phosphorylated residues). Phosphorylation during transcription is partially undertaken by cyclin-dependent kinases (CDKs). These are a group of Ser/Thr kinases that function as a heterodimeric complex: the kinase and regulatory cyclin protein.

1.3 The role of RNA polymerase II CTD

RNA polymerase II has a unique C-terminal domain (CTD) on its largest subunit, RBP1, which is essential for cell viability (Meininghaus et al. 2000). The CTD consists of heptad repeats with the consensus sequence Tyr_1-Ser_2-Pro_3-Thr_4-Ser_5-Pro_6-Ser_7 ($Y_1S_2P_3T_4S_5P_6S_7$). Pol II enzymes of different organisms contain a characteristic number of repeats: *S. cerevisiae* have 26 heptad repeats, drosophila contains 45, and the human and mouse CTDs have 52 heptad repeats (Corden 1990).

During the transcription cycle, the residues of the CTD are post-transcriptionally modified. Each stage of transcription is associated with a unique set of modifications. These post-transcriptional modifications provide distinct binding sites for recruiting specific factors appropriate for the various stages in the transcription cycle. mRNA processing enzymes, transcription elongation factors

and chromatin remodelling proteins are recruited by the CTD (reviewed in (Egloff and Murphy 2008)). Thus it is these modifications that provide the "CTD code" (Buratowski 2003). CTD post-translational modifications are tabulated in Table 1-1.

Residue	*Modification*	*Enzyme*	*Function of modification*	*References*
Y1	Phosphorylation	c-Abl, Abl-related gene (Arg)	Transcription activation?	(Baskaran et al. 1993; Baskaran et al. 1996; Baskaran et al. 1997)
S2	Phosphorylation	CDK9, CDK8, CDK12, CDC2, DNA-PK	Transform to elongation complex; splicing; polyadenylation	(Zhang and Corden 1991; Sun et al. 1998; Trigon et al. 1998; Zhou et al. 2000; Bartkowiak et al. 2010)
	Glycosylation?	O-GlcNAc transferase	?	(Kelly et al. 1993; Comer and Hart 2001)
T2	Phosophorylation	CDK1?	?	(Zhang and Corden 1991)
P3	Isomerisation	Pin1	Increased CTD phosphorylation, inactivates Pol II during M phase and Inhibits splicing.	(Xu et al. 2003; Shaw 2007; Xu and Manley 2007)
T4	Glycosylation	O-GlcNAc transferase	?	(Kelly et al. 1993; Comer and Hart 2001)
	Phosphorylation	CDK9	Histone mRNA 3' end formation	(Hsin et al. 2011)
S5	Phosphorylation	TFIIH, CDK8, CDK7, CDC2, ERK1/2	Recruitment of capping enzyme	(Zhang and Corden 1991; Sun et al. 1998; Trigon et al. 1998)
	Glycosylation	O-GlcNAc transferase	?	(Kelly et al. 1993; Comer and Hart 2001)
T5	Phosphorylation	CDK1?	?	(Zhang and Corden 1991)
P6	Isomerisation	Pin1	As above	(Xu et al. 2003; Xu and Manley 2007)
S7	Phosphorylation	DNA-PK CDK7, CDK9	Recruitment of integrator and 3' end processing of snRNAs	(Trigon et al. 1998; Chapman et al. 2007; Egloff et al. 2007; Egloff and Murphy 2008; Akhtar et al. 2009; Glover-Cutter et al. 2009)
	Glycosylation?	O-GlcNAc transferase	?	(Kelly et al. 1993; Comer and Hart 2001)
T7	Glycosylation	O-GlcNAc transferase	?	(Kelly et al. 1993; Comer and Hart 2001)
R1810	Methylation	CARM1	Represses transcription of snoRNA and snRNAs	(Sims et al. 2011)
S13 (of repeat 52)	Phosphorylation	CKII	Stabilises initiation complex, inhibiting transcription	(Cabrejos et al. 2004; Chapman et al. 2004)
K	Ubiquitination	Wwp2	Rpb1 degradation	(Li et al. 2007)

Table 1-1. Table of post-translational modifications on the human CTD and their function.

Initially the CTD is unphosphorylated when recruited to the preinitiation complex at the gene promoter (Laybourn and Dahmus 1989) (See Figure 1-1). During transcription initiation, cyclin dependent kinase 7 (CDK7), as part of the TFIIH complex, phosphorylates the Ser5 residues within the heptad repeat (Trigon et al. 1998). Ser5 phosphorylation releases the mediator complex and allows Pol II to initiate transcription (Max et al. 2007). Additionally Ser5 phosphorylation recruits the capping enzyme to the polymerase and the nascent RNA is then capped by the removal of the terminal phosphate, the addition of an inverted guanosine monophosphate and the subsequent methylation of the guanosine (Ho and Shuman 1999; Fabrega et al. 2003). This is the first modification that occurs to the nascent mRNA. NELF and DSIF pause the polymerase allowing CDK9 to be recruited. CDK9 subsequently phosphorylates these factors and releases the paused Pol II (Nechaev and Adelman 2011). Pol II progresses along the gene and Ser2 is phosphorylated by CDK9 and CDK12. Two termination factors, Pcf11 and Rtt103, bind to Pol II when sequential heptad repeats are phosphorylated at Ser2 and initiate transcription termination (Lunde et al. 2010).

In a similar fashion to the kinases, CTD phosphatases are associated with different stages of the transcription cycle and have different specificities (see Table 1-2). These phosphatases therefore alter the proportion of Ser2 and Ser5 phosphorylation marks on the CTD during transcription. They are also required to enable recycling of the CTD for a new round of transcription (Meinhart et al. 2005).

Phosphatase	Target phosphorylation	Reference
Scp1	Ser5P	(Yeo et al. 2003; Zhang et al. 2006)
Fcp1	Ser2P	(Hausmann and Shuman 2002)
Ssu72	Ser5P	(Krishnamurthy et al. 2004)
Rtr1	Ser5P	(Mosley et al. 2009)

Kinase	Target substrate	Reference
CDK7	Ser5	(Trigon et al. 1998)
CDK8	Ser5	(Sun et al. 1998)
CDK9	Ser2	(Zhou et al. 2000)
CDK12/13	Ser2?	(Bartkowiak et al. 2010)

Table 1-2. The phosphatases and CDK kinases involved in regulating the CTD phosphorylation state.

Chromatin remodelling enzymes are also recruited by the Pol II CTD phosphorylation marks. Set1 (in yeast) binds to Ser5 phosphorylation marks and methylates lysine 4 of histone H3 (H3K4) residues, recruiting histone acetylases, such as NuA3 through Yng1, to modify the chromatin for transcription. Acetylation of histone lysines neutralizes their positive charge, weakening the interaction between the histones and DNA. Ser2 phosphorylated CTD is a binding platform for Set2. Set2 methylates H3K36 such that Rpd3S is recruited and deacetylates the histones behind the transcribing polymerase (Bannister and Kouzarides 2011). This prevents initiation at cryptic sites within the gene. RECQ5 has been identified as being recruited to the CTD when both Ser5 and Ser2 residues are phosphorylated but its function is not fully understood (Bannister and Kouzarides 2011). For further details and references of CTD interacting proteins see Table 1-3.

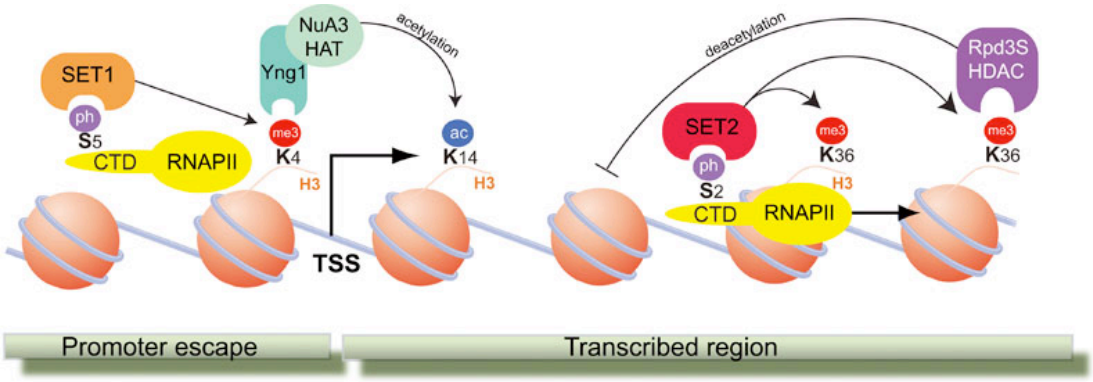


Figure 1-2. Recruitment of histone modifying factors through phosphorylation of Ser5 and Ser2 of the CTD in yeast. Reprinted by permission from Macmillan Publishers Ltd: Cell Research (Bannister and Kouzarides 2011), copyright 2011. **http://www.nature.com/cr/index.html**.

Interacting protein	*Phosphorylation*	*Function*	*Reference*
Mammalian capping enzyme (Mce)	Ser5P	CTD capping	(McCracken et al. 1997; Ho and Shuman 1999; Fabrega et al. 2003)
Integrator	Ser2PSer7P	snRNA transcription	(Egloff et al. 2007; Egloff and Murphy 2008)
Pcf11	Ser2P	Transcription termination	(Lunde et al. 2010)
Rtt103	Ser2P	Transcription termination	(Lunde et al. 2010)
RECQ5	Ser2PSer5P	Genome stability during transcription	(Kanagaraj et al. 2010)
SCAF8	Ser2PSer5P	RNA processing	(Patturajan et al. 1998; Becker et al. 2008)
Pin1	Ser2P or Ser5P	Proline isomerisation and repress transcription	(Yunokuchi et al. 2009)
PCIF1	Ser2P or Ser5P	Inhibit SCP1 phosphatase and transcription repression	(Fan et al. 2003; Hirose et al. 2008; Yunokuchi et al. 2009)
Spt6	Ser2P	mRNA processing and export	(Yoh et al. 2007)
Set1	Ser5P	Histone methyltransferase	(Ng et al. 2003)
Set2	Ser2P	Histone methyltransferase	(Xiao et al. 2003)
Nrd1	Ser5P	snoRNA processing and termination, suppression of cryptic unstable transcripts (CUTs)	(Vasiljeva et al. 2008) reviewed in (Lykke-Andersen and Jensen 2006)

Table 1-3. Proteins that interact with specific phosphorylation states of Pol II CTD. The upper rows correspond to proteins identified in humans. The proteins below the double line have been identified in yeast and are presumed to have a similar function in humans.

Little is understood about the other post-transcriptional modifications. These are included in Table 1-1. These include tyrosine phosphorylation, which has been linked with initiation of transcription and coordination of transcription with cell adhesion, the cell cycle and DNA damage. Furthermore, glycosylation of the CTD has been observed and also suggested to have a role in initiation (Kelly et al. 1993; Comer and Hart 2001). Threonine phosphorylation on residue 2 of the heptad repeat has been reported, but the function of this modification is uncertain (Zhang and Corden 1991), conversely Thr4 phosphorylation is known to be required for histone mRNA 3' end processing (Hsin et al. 2011). Finally, a proline isomerase that alters the cis/trans conformation of prolines within the CTD alters both the phosphorylation status and the rate of transcription (Xu and Manley 2007). Because the structures of many CTD interacting domains (CIDs) are known, including examples of Rtt103, SCAF8, and Pcf11 CTD-binding domains, the significance of proline isomerization can be readily understood: proline residues are often involved in the recognition of the CTD by CID-containing proteins (Meinhart and Cramer 2004; Becker et al. 2008; Lunde et al. 2010). Hence Pin1, a proline isomerase, can alter the recruitment of factors to the CTD, including the kinases and phosphatases (Xu et al. 2003; Xu and Manley 2007).

1.4 Cyclin-Dependent kinases

Transcription regulation is inseparably linked to phosphorylation status. The primary kinases involved in transcription are CDKs. In general, CDK/cyclin

complexes regulate the cell cycle and transcription, although CDK5 and PCTAIRE are examples of CDKs that have other functions in differentiated cells. The activity of cell cycle CDKs is controlled by the presence of the cyclin, cyclin-dependent kinase inhibitory proteins (CKIs), and the phosphorylation state of the kinase. Association of CDKs in large multiprotein complexes provides a further level of regulation of CDK activity, and this phenomenon is particularly apparent for the transcriptional CDKs (Morgan 2007). The transcriptional CDKs will be discussed below and the cell cycle CDKs in Chapter 1.7.

1.4.1 Transcriptional CDKs

There are several CDKs that regulate transcription (Figure 1-3). These differ from the cell-cycle CDKs in their regulation as the cyclins of transcriptional CDKs are not expressed periodically in the same manner that cell cycle cyclins are. Transcriptional CDKs are usually incorporated into larger complexes: CDK8 is part of the mediator complex, CDK7 is complexed within TFIIH, while CDK9 with cyclin T together form P-TEFb. Other CDKs, specifically CDK12 and CDK13, have only recently been found to phosphorylate the CTD and their role in transcription has yet to be established (Bartkowiak et al. 2010).

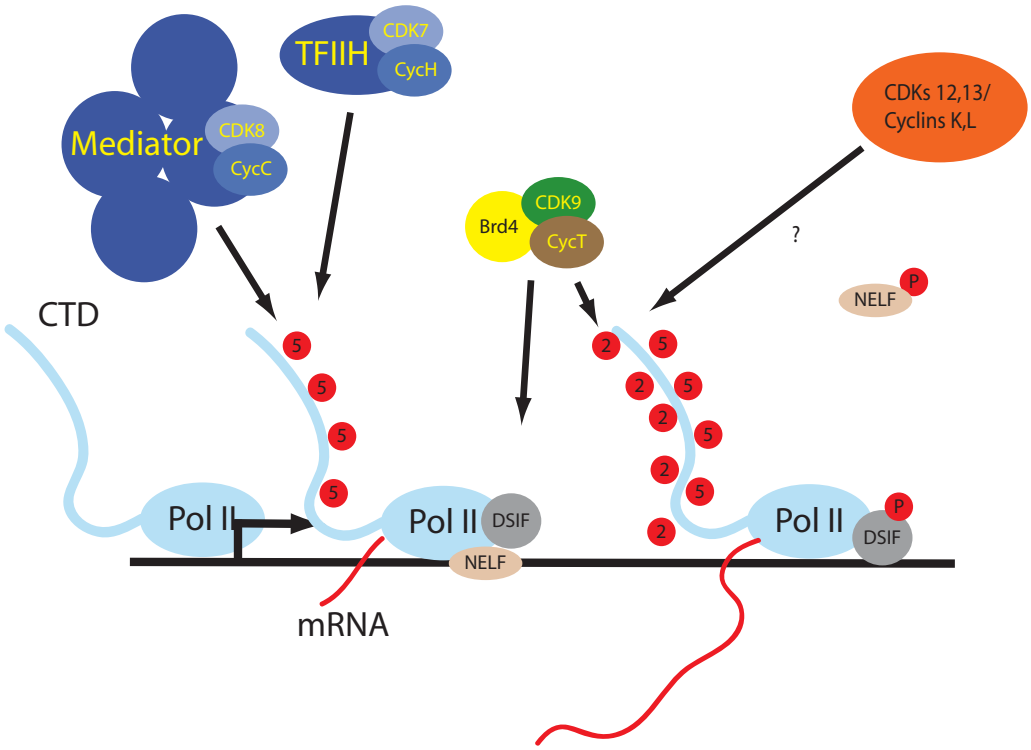


Figure 1-3. The functions of CDKs in transcription. Pol II binds to the promoter region during transcription, whilst the CTD is unphosphorylated. CDK8 and CDK7 as part of the mediator complex and TFIIH respectively enable promoter clearance by phosphorylating the Ser5 of the CTD. The complex is halted by NELF and DSIF. P-TEFb (CDK9/cyclin T) complexed with Brd4 phosphorylates NELF, DSIF and Ser2 of the CTD to enable productive transcription. CDKs 12 and 13 may also phosphorylate Ser2 on the CTD, but more studies are required for improved understanding of their functions.

1.4.1.1 CDK8 (Mediator complex)

The first CDK to be recruited to the gene is CDK8. CDK8/cyclin C is found in certain forms of the Mediator complex. This complex is composed of over 30 proteins in humans and has several different forms (Conaway and Conaway 2011). Its size and complexity have hindered the full understanding of its biochemical functions, about which much remains to be elucidated.

The mediator complex is recruited to the promoter region by different activators and repressors and serves as an interface between regulatory molecules and the core transcription machinery (Malik and Roeder 2010). It

recruits the unphosphorylated Pol II, initiating the formation of the preinitiation complex. Various signaling pathways converge at the mediator complex, and the presence of CDK8 or phosphorylation of the CTD of Pol II releases Pol II for transcription initiation. Pol II is released from the promoter region, but the mediator complex remains bound to the promoter together with a set of general transcription factors in a scaffold complex. The scaffold complex enables additional rounds of Pol II PIC formation and initiation.

Although this model describes the role of the mediator complex in promoting transcription, many studies indicate that the mediator complex and CDK8 regulate transcription negatively as well as positively (see (Galbraith et al. 2010) and examples within). Transcription is either activated or repressed depending on the gene and protein environment.

1.4.1.2 CDK7

CDK7 has fundamental roles in both the cell cycle control and in transcription. It exists in several different complexes with different respective roles. Bound to cyclin H and MAT1, CDK7 acts as the CDK activating kinase (CAK) phosphorylating the activation segments of CDK1, CDK2, CDK4, CDK6 and possibly CDK9. CDK7 phosphorylates CDKs when they are bound to their respective cyclins and hence the phosphorylation by CAK results in the activation of these CDKs and enables progression through the cell cycle. During G_0 CDK7 activity and protein levels are reduced to prevent cell cycle activation.

The transcription factor, TFIIH, consists of CAK (CDK7/cyclin H/Mat1) and 6/7 other subunits. Within this complex the substrate specificity of CDK7 alters to phosphorylate TFIIE, TFIIF, certain nuclear receptors (such as the estrogen receptor α), and residues Ser5 and Ser7 of the heptad repeat of the CTD. This phosphorylation promotes transcription initiation and recruits the mRNA capping enzyme to process the pre-mRNA. For reviews on CDK7 see (Fisher 2005; Lolli and Johnson 2005).

1.4.1.3 CDK 11

Three isoforms of CDK11 exist: $CDK11^{p110}$, $CDK11^{p58}$ and $CDK11^{p46}$. These form CDK/cyclin complexes with one of the 6 cyclin L proteins. Whilst $CDK11^{p58}$ and $CDK11^{p46}$ are associated with mitosis ((Franck et al. 2011) and references within) and apoptosis (Feng et al. 2005) respectively, $CDK11^{p110}$ affects pre-mRNA splicing (Hu et al. 2003). It is found in macromolecular complexes containing transcription elongation factors and mRNA splicing factors such as phosphorylated RNA polymerase, FACT, splicing factor RNPS1 and TFIIF (Trembley et al. 2002; Loyer et al. 2008). However, the substrates of CDK11 that influence splicing and transcription still remain to be determined (Loyer et al. 2011).

1.4.1.4 CDK12 and CDK13

Both CDK12 and CDK13 have only recently been discovered and little has been unambiguously determined about their cellular functions. They are reported to form complexes with cyclins K and L, although this is still a matter of debate

(Chen et al. 2006; Chen et al. 2007; Bartkowiak et al. 2010). They are both able to phosphorylate the CTD and have been implicated in influencing splicing, although again, whether this is directly or indirectly is uncertain (Chen et al. 2006; Chen et al. 2007; Bartkowiak et al. 2010).

The yeast (S. cerevisiae) system has often been utilized to develop models and comparisons for metazoan systems. Yeast differ from metazoa in several aspects, one of which is that they contain two kinases that phosphorylate the Ser2 residues of the CTD. Ctk1 is localized in the central region and towards the 3' UTR of the gene and is the primary CTD kinase (Cho et al. 2001; Mayer et al. 2010). Bur1 associates primarily with the central region of the gene and phosphorylates both SPT5 (DSIF), and the CTD (Murray et al. 2001; Zhou et al. 2009a; Mayer et al. 2010). CDK9 is suggested to be orthologous to Bur1, and CDKs 12/13 to Ctk1 (Bartkowiak et al. 2010). Elucidating the roles of CDK12 and CDK13 will be crucial for developing our understanding of CTD phosphorylation and transcription regulation in metazoa.

1.4.1.5 Role of CDK9 in transcription

CDK9 is required for the synthesis of the majority of Pol II transcripts (Chao and Price 2001). CDK9 interacts with either cyclin T1 or T2 to form P-TEFb (Peng et al. 1998). P-TEFb stimulates transcription elongation by phosphorylating the Pol II CTD as well as the pause-inducing factors, NELF and DSIF (Marshall and Price 1995). P-TEFb is recruited to paused RNA polymerases by general transcription factors (e.g. Brd4), and specific transcription factors (Peterlin and

Price 2006; Takahashi et al. 2011). Here it phosphorylates NELF (Fujinaga et al. 2004), releasing this protein from the halted complex, and DSIF (Ivanov et al. 2000). Upon phosphorylation DSIF becomes a positive transcription factor (Yamada et al. 2006). P-TEFb also phorphorylates Ser2 of the CTD, thereby promoting recruitment of transcription elongation and processing factors as discussed above (Chapter 1.3). Together these phosphorylations enable transcription to continue and the RNA polymerase to enter the elongation stage.

1.5 Regulation of transcription

For a long time transcription was believed to be regulated primarily at the stage of initiation. More recently, genome wide studies have demonstrated the accumulation of Pol II at the 5' end of the gene, around 30 nucleotides downstream of the promoter in yeast (Venters and Pugh 2009), *Drosophila* (Muse et al. 2007; Zeitlinger et al. 2007) and human embryonic stem cells (Guenther et al. 2007). Pol II localized to this site is considered to be in a paused state that represents an additional stage at which transcription of many genes may be regulated. It is suggested that Pol II promoter proximal pausing is likely to be the rate limiting step in transcription for up to 20-30% of *Drosophila* and human genes (Core and Lis 2008).

CDK9, as a subunit of P-TEFb, is the kinase that reverses promoter proximal pausing of RNA polymerase II. The regulation of P-TEFb is, therefore, crucial for the progression of Pol II from paused to elongating. Therefore, understanding P-TEFb regulation is essential to understanding gene expression.

1.6 CDK9/cyclin T interaction proteins

There are two main approaches used by the cell to regulate CDK9/cyclin T. These are by inhibition and activation of its kinase activity and through localizing it to sites of active transcription.

1.6.1 Regulation

The activity of CDK9 is regulated in several ways including phosphorylation (Li et al. 2005), acetylation (Fu et al. 2007; Sabo et al. 2008) and by association with other proteins. It is also contained within several multimeric complexes. For example, it forms a complex with transcription activators and the super elongation complex (SEC) (He et al. 2010; Lin et al. 2010; Sobhian et al. 2010) and has also been associated with a network of other proteins (Bezstarosti et al. 2010). CDK9/cyclin T has been studied primarily in two forms – the large inhibitory complex and the small activatory complex (Figure 1-4). The larger complex is composed of CDK9, cyclin T, Hexim1, 7SK RNA, MePCE and Larp7 (Nguyen et al. 2001; Yang et al. 2001; Michels et al. 2003; Yik et al. 2003; Chen et al. 2004; Jeronimo et al. 2007; Krueger et al. 2008). During high stress conditions and induced differentiation, the smaller complex is formed that contains CDK9, cyclin T and Brd4 (Jang et al. 2005; Yang et al. 2005; He et al. 2006).

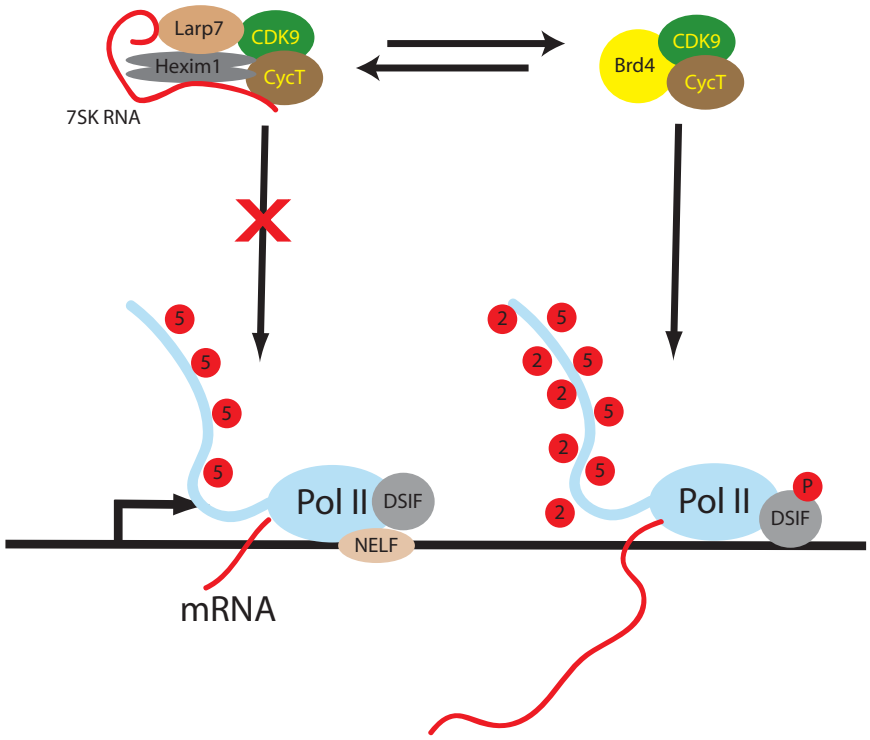


Figure 1-4. Regulation of CDK9 by protein association. The inhibitory complex of P-TEFb contains Larp7, Hexim1, and 7SK RNA. The smaller activating complex consists of CDK9, cyclin T and Brd4**.**

1.6.2 Recruitment of P-TEFb to sites of transcription

The active P-TEFb complex is recruited to promoter regions by both general and specific activators that determine which genes are actively transcribed. Recruitment by specific activators has been implicated in differentiation, inflammatory and immune response, and apoptosis. Peterlin and Price (Peterlin and Price 2006) grouped the activators into four separate classes determined by their mechanisms of interaction with DNA. These are co-activators, which do not bind directly to DNA but require intermediary proteins, DNA bound activators, RNA bound activators and chromatin bound activators (Figure 1-5).

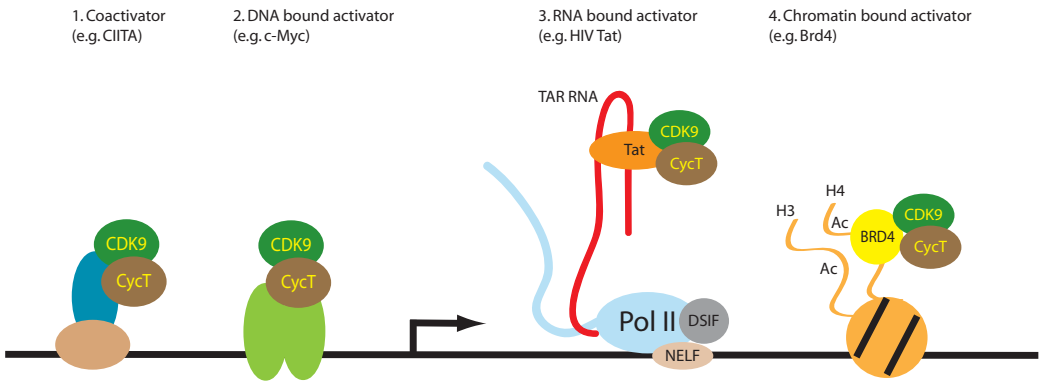


Figure 1-5. Methods of P-TEFb recruitment to promoter regions. CDK9/cyclin T can be recruited to promoter regions via (1) coactivators, (2) DNA bound activators, (3) RNA bound activators, or (4) chromatin bound activators. H3 and H4 refer to the tails of histones 3 and 4 respectively, and Ac indicates acetylated sites. Figure adapted from Molecular Cell 23, Peterlin and Price, Controlling the Elongation Phase of Transcription with P-TEFb 297-305, Copyright 2006, with permission from Elsevier. http://www.cell.com/molecular-cell/home

The most characterized example of recruitment of CDK9/cyclin T to the promoter region is that mediated by the HIV-Tat/TAR complex. The HIV genome encodes for long terminal repeats (LTR) that interact with trans-activator of transcription (Tat), an HIV protein, to recruit host cell P-TEFb (Mancebo et al. 1997; Zhu et al. 1997). Transcription of the HIV LTRs produces a short hairpin loop structure of RNA called the trans-acting response (TAR) RNA element. Tat interacts with TAR and recruits CDK9/cyclin T to the polymerase (Cordingley et al. 1990; Wei et al. 1998). P-TEFb initiates Pol II elongation in the same manner as applies in authentic host transcription (Zhou et al. 2004). This enables up-regulation of transcription from the LTR. The HIV-LTR Tat/TAR system has been crucial in a range of experiments undertaken to develop our understanding of P-TEFb, transcription elongation and HIV virology.

Class II transactivator (CIITA) is an example of a coactivator that recruits CDK9/cyclin T. CIITA is either constitutively expressed in professional antigen presenting cells or can be induced in other cell types. Target genes that are

regulated by CIITA include the major histocompatibility complex (MHC) class II family. CIITA cannot bind to DNA directly but binds via factors that form an enhanseosome at MHC II promoters (reviewed in (Chang et al. 2002)). Here CIITA recruits P-TEFb through cyclin T and stimulates the initiation of MHC II genes (Kanazawa et al. 2000; Drozina et al. 2006).

There are a range of DNA-binding activators that recruit CDK9/cyclin T to promoter regions. These include those that are stimulated on ligand binding and those stimulated by signaling pathways. Included in these families are the aryl hydrocarbon receptor (AhR) (Tian et al. 2003), the PPARγ activator (Iankova et al. 2006), STAT3 (Giraud et al. 2004; Hou et al. 2007) and MyoD (Simone et al. 2002). These are involved in upregulating genes for xenobiotic metabolism, adipocyte differentiation, and myogenic differentiation respectively. Other activators that deserve a slightly longer introduction are c-Myc and NF-κB that are described in more detail below.

c-Myc is a transcription factor activated by mitogenic signals that regulates expression of genes involved in cell proliferation. c-Myc binds to a third of all actively transcribing genes (Rahl et al. 2010). In complex with Max, c-Myc both recognizes a consensus sequence of DNA and interacts with CDK9/cyclin T (Kanazawa et al. 2003; Gargano et al. 2007). It therefore recruits P-TEFb to sites of active transcription, releasing paused Pol II from the promoter-proximal site and enabling transcription to continue.

The NF-κB transcription factor translocates to the nucleus on activation by several events including inflammation (Reviewed in (Oeckinghaus and Ghosh 2009)). NF-κB once in the nucleus is recruited to specific genes involved in either its own regulation or those involved in physiological responses, such as cytokines and chemokines. Depending on the core promoter type P-TEFb is also recruited and activates a subset of these NF-κB-dependent genes (Nowak et al. 2008). NF-κB is also utilized by latent HIV as it initiates transcription from the LTR (Nabel and Baltimore 1987).

The final group of proteins that recruit CDK9/cyclin T are chromatin bound proteins. The CDK9/cyclin T/Brd4 trimeric complex is part of an active CDK9/cyclin T kinase complex. Brd4 is recruited to transcription start sites and enhancer sites by interaction with acetylated histones, the mediator complex and coactivators (Jang et al. 2005; Yang et al. 2005; Zippo et al. 2009). It interacts with CDK9/cyclin T and thereby causes an increased phosphorylation of the CTD and transcription. A total of 10% of genes are proposed to be regulated by Brd4 (Mochizuki et al. 2008).

1.7 CDKs in the cell cycle

As well as specifically phosphorylating the CTD, CDK's are involved in several cellular processes including the cell cycle (CDKs 1, 2, 4, 6 and 7) and other processes such as cytoskeletal development, glucose homeostasis and neurone development (CDK5, reviewed in (Lalioti et al. 2010)). Historically structural studies have focused on the CDKs associated with the cell-cycle. It is from these

studies that a majority of our understanding of CDKs has been derived and hence I have provided a short description of their function below.

Progression through the cell cycle is regulated at several different stages and checkpoints (Morgan 2007). DNA replication occurs during S-phase, and progression through M-phase separates the DNA into two daughter cells. Two growth phases, G_1 and G_2, occur between mitosis and S-phase and between S-phase and mitosis respectively. During the G_1-S phase transition, D-type cyclins are expressed, and associate with CDK2, 4, and 6 to form the corresponding active CDKs 4/6-cyclin D and CDK2/cyclin E complexes (Figure 1-6). These complexes phosphorylate pRb, and thereby release the transcription factor E2F and members of the HDAC (histone deacetylase) family enabling the transcription of proteins required for subsequent cell cycle progression. Furthermore, CDK2/cyclin E phosphorylates centrosomal proteins, releasing them from the centrosome and initiating centrosome duplication. The S/G_2 phase transition is controlled by CDK2/cyclin A. During S phase CDK1/cyclin A phosphorylates proteins involved in DNA replication, such as CDC6 and CDT1. These proteins are then either degraded or exported from the nucleus to prevent multiple replication cycles (reviewed in (Suryadinata et al. 2010)).

Although the roles of specific CDK/cyclin pairs at specific points in the cell cycle are well established, it has been shown by knock-out and knock-down experiments that CDKs 2, 4 and 6 are non-essential for cell cycle progression and similarly D-type and E-type cyclins are not required. Whereas the different CDKs and cyclins appear able to functionally complement each other under such

circumstances, mice lacking certain combinations of these proteins display tissue-specific defects (reviewed in (Malumbres and Barbacid 2009)). CDK1 is, however, essential (Santamaria et al. 2007). CDK3/cyclin C is responsible for both transition from G_0 into G_1 and G_1 to S phase through its phosphorylation of Rb and activating transcription factor 1 (ATF1) (Ren and Rollins 2004).

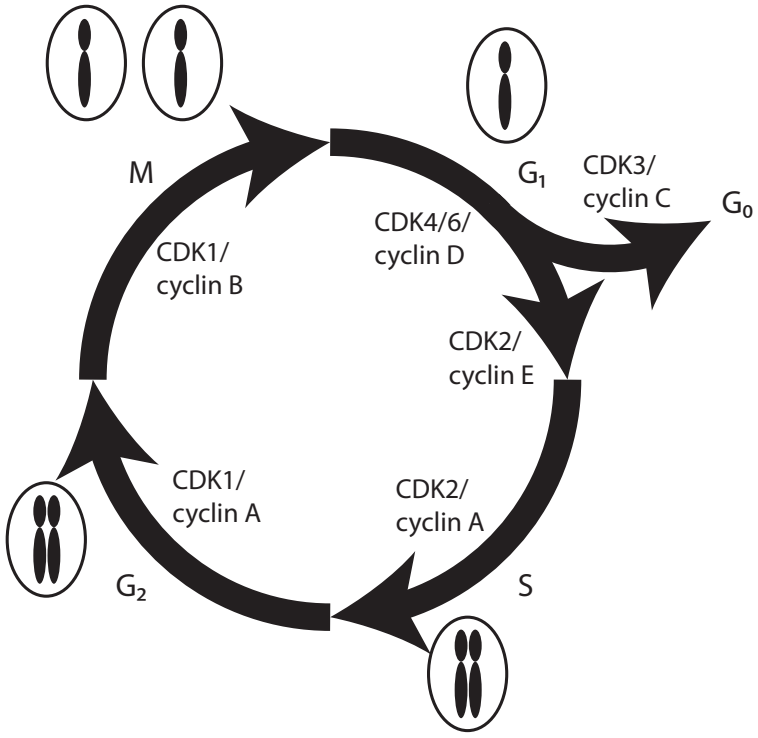


Figure 1-6. Regulating the cell cycle by CDK/cyclins. The relevant CDK/cyclin complexes are shown where their activity is required to progress through the stages of the cell cycle. The cells and chromosome numbers are also depicted within the diagram.

For detailed analysis of CDK substrates see (www.phosphosite.org).

1.8 CDK structural features

As alluded to above, CDKs form dimeric complexes with their respective cyclins to form active kinases. The majority of CDK proteins contain a highly conserved kinase domain (Figure 1-7). These consist of an N- and C- terminal lobe and a hinge region that connects these two lobes. The ATP binding site is located

between the N- and C-terminal lobes. They have mixed secondary structure: 10 β-strands, predominantly in the N-terminal lobe and 11 α-helices predominantly in the C-terminal lobe. The cyclins and peptide substrates interact through mainly the N- and C- terminal lobes of the kinase, respectively.

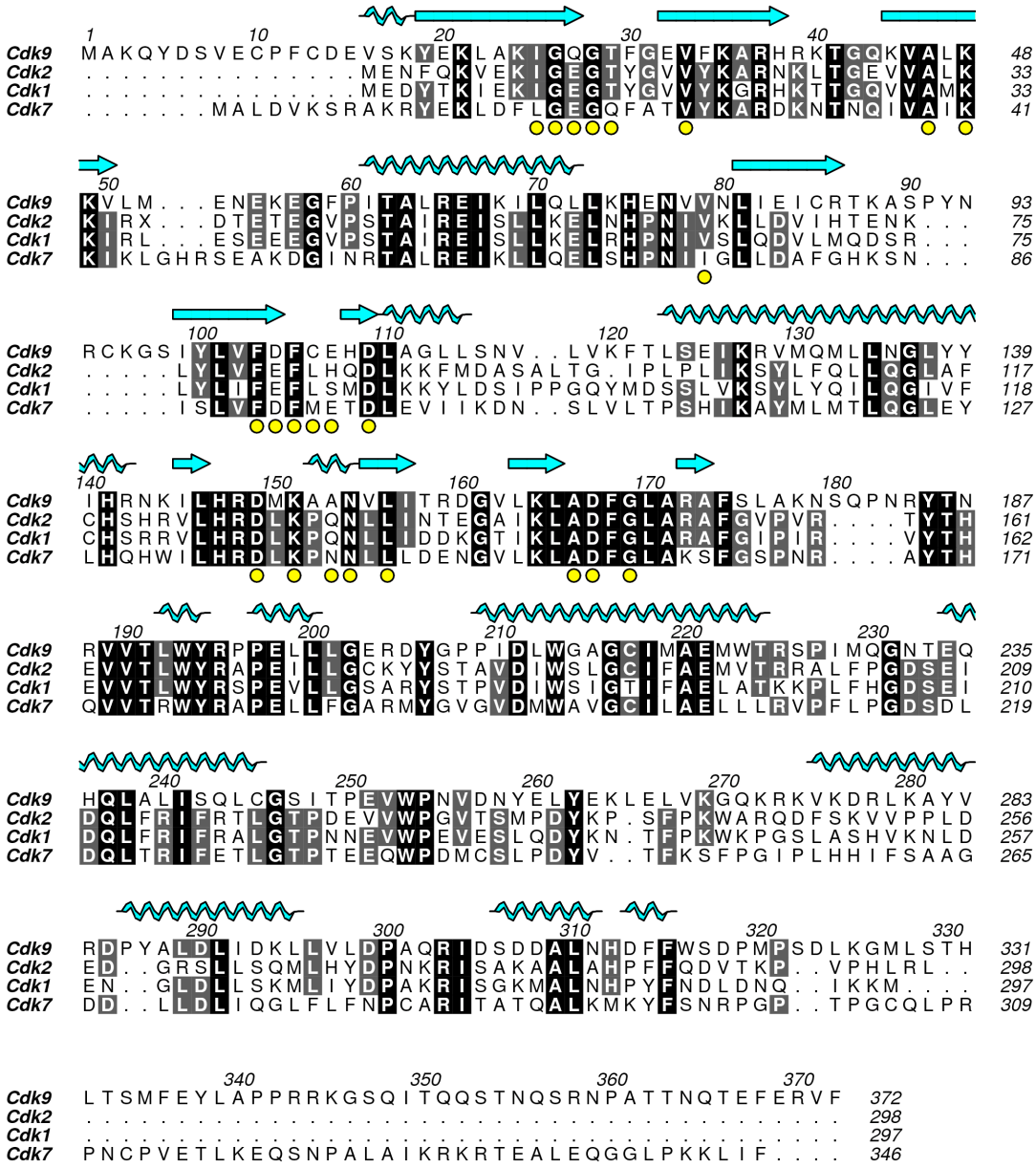


Figure 1-7. Alignment of CDK sequences. CDK9 secondary structure is shown above the sequences and yellow circles indicate residues within CDK9 that are within 5Å of ATP (identified from PDB: 3BLQ). ClustalW2 (Larkin et al. 2007) was used to align the sequences and the image was formed using ALINE (Bond and Schuttelkopf 2009).

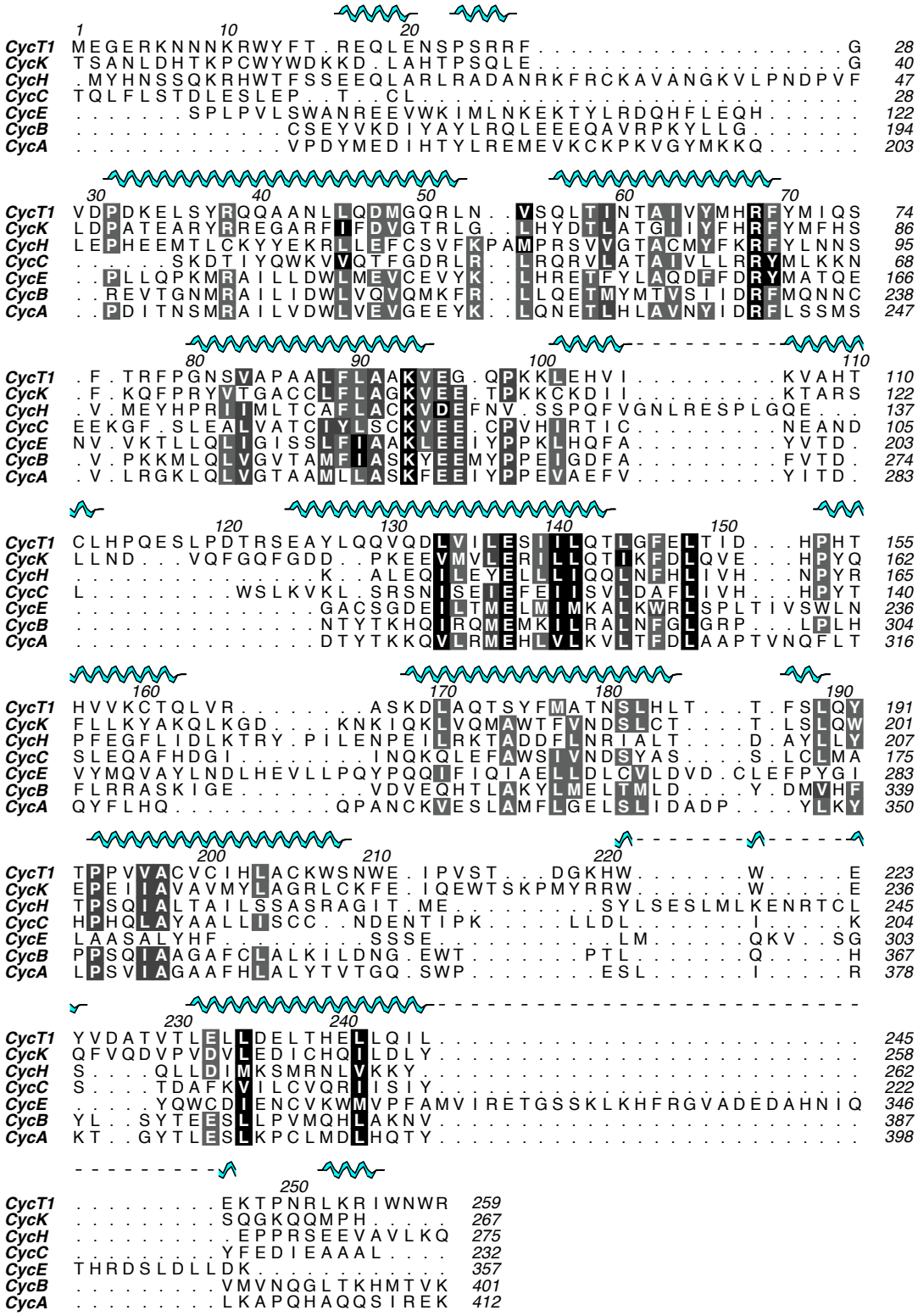


Figure 1-8. Structure-based sequence alignment of the cyclin boxes of cyclins. The secondary structure of cyclin T is show above the sequences.

The cyclins are much more diverse than the CDKs. They are comprised of 2 cyclin boxes, each composed of 5 α-helices (Figure 1-8). These show greater sequence variation than the CDK kinase domains, but maintain the same overall structure of cyclin boxes. In addition, cyclins often have N- and C-terminal extensions that include binding domains for interactions with other proteins.

The textbook description of CDK activation is predominantly based on CDK2 structural analysis. Full activation requires both association of the CDK with its corresponding cyclin and phosphorylation of the threonine on the activation segment (Figure 1-9). The combination of these two events aligns residues within the ATP binding site for ATP binding and phosphotransfer, and opens a binding platform for the protein substrate to bind and receive the γ-phosphate (Brown et al. 1999).

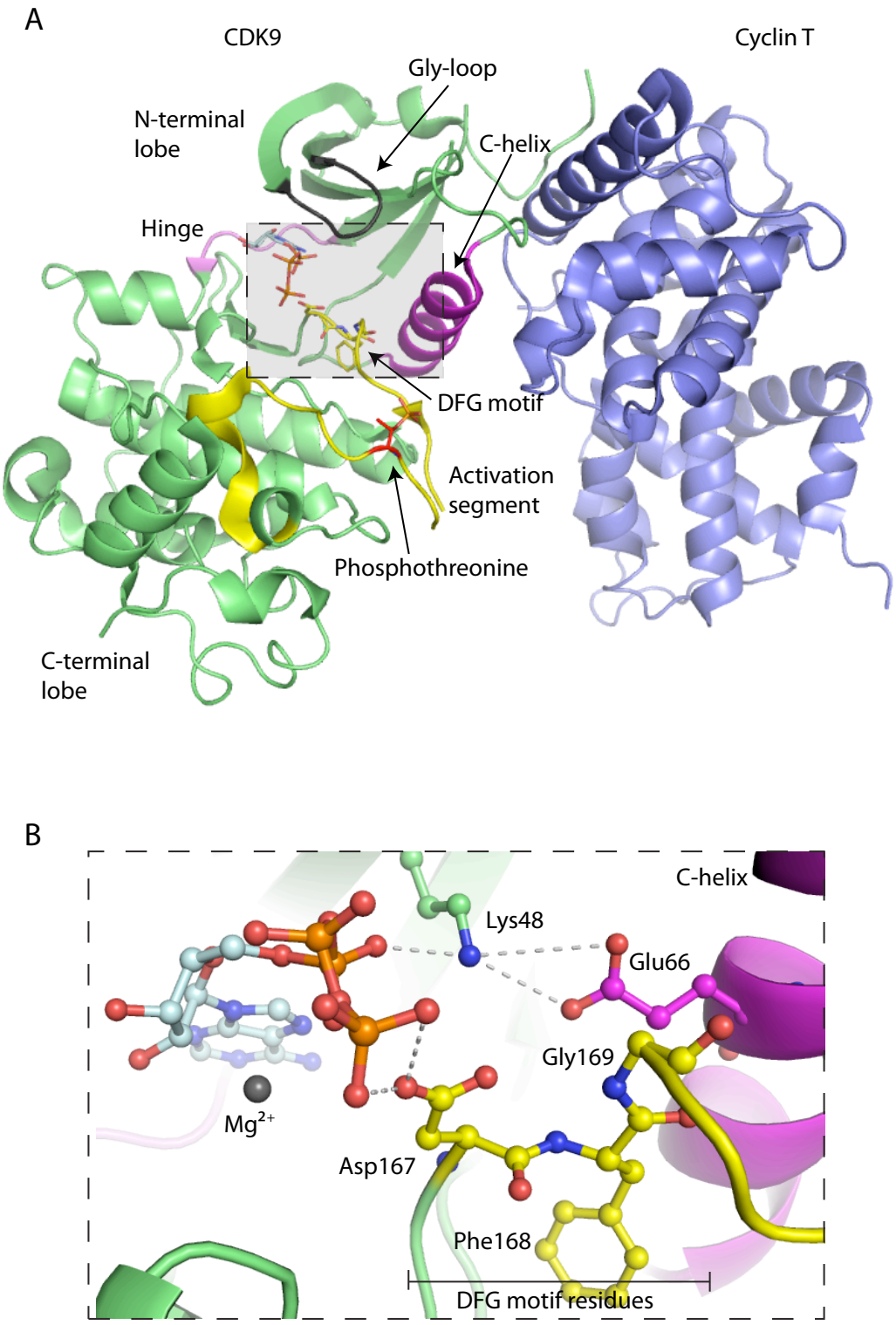


Figure 1-9. Structural features of CDK9/cyclin T (A) and selective residues involved in phosphotransfer (B). The following features are shown in a cartoon representation with the colours they are presented in in parenthesis: cyclin T (slate), CDK9 (green), the C-helix (magenta), the hinge region (violet), the glycine-rich loop (dark grey), and the activation segment (yellow). CDK9 phosphothreonine 186 (red), ATP (pale cyan), CDK9 residues 48, 66, 167, 168 and 169 are shown in ball and stick representation (PDB used:3BLQ). The DFG motif is comprised of CDK9 residues 167-169. The shadowed box in (A) represents the region shown in more detail in (B).

Structural features of particular interest are primarily located on the CDK. These include the C-helix, the hinge region, and the activation segment, which includes the αL12 helix and the DFG motif. The functions of these in catalysis will be described below.

The C-helix, known also as the PSTAIRE helix, is located at the heart of the CDK/cyclin interface. Monomeric CDKs have their C-helices in an inactive "out" conformation. On cyclin binding, the C-helix re-orientates to an "in" conformation, undergoing a rotation and tilting of the helix (Jeffrey et al. 1995). This enables Glu66 of CDK9 (corresponding to Glu51 in CDK2) to form a hydrogen bond network with Lys48 (Lys33 in CDK2) and Asp167 (Asp145 in CDK2) (of the DFG motif). This results in the ATP phosphates and Mg^{2+} ions being aligned correctly for catalysis.

Prior to cyclin binding, residues from the CDK2 αL12 loop, which immediately follows the highly conserved kinase DFG-motif, are helical, and the activation segment is held in a conformation that is incompatible with the binding of peptide substrates (De Bondt et al. 1993). Subsequent to cyclin binding, the segment moves away from the active site enabling the active site to be accessible to ATP. The αL12 residues rearrange into a β-strand. For full activation of the kinase a threonine, near the centre of the activation segment, (Thr160 in CDK2, and Thr186 in CDK9), must become phosphorylated subsequent to cyclin binding. Once phosphorylated, the negative phosphate of the phospho-threonine interacts with several arginines, initiating a network of hydrogen bonds to stabilise the CDK and cyclin, preparing the protein substrate

recognition site for protein binding and catalysis (Russo et al. 1996). It is this location that determines the consensus substrate peptide sequence of CDKs as [S/T*]PX[K/R], where * denotes the phospho-acceptor residue (Brown et al. 1999). Reviews of CDK structural elements and regulation include (Hanks and Hunter 1995; Morgan 2007; Johnson 2009b).

Several crystal structures of CDKs and cyclins have been determined (Lolli 2010). These include CDKs and cyclins crystallised separately, CDK/cyclin complexes in phosphorylated and unphosphorylated states, and in the presence of substrate peptides and inhibitory proteins. Although much is known about CDK2 and its associated complexes, less is known about the transcriptional CDKs. To date, monomeric CDK7 (Lolli et al. 2004), CDK9/cyclin T (Baumli et al. 2008) and very recently CDK8/cyclin C (Schneider et al. 2011) structures have been solved by x-ray crystallography. Comparison of the transcriptional CDK/cyclin complexes to CDK2/cyclin A identified that although the transcriptional CDK/cyclin complexes display the same general overall fold in their CDKs and cyclins, there are some variations between the structures. The most striking difference is in the relative orientation of the cyclin to the CDK within the complexes: in comparison to CDK2/cyclin A, cyclin T is rotated by 26° with respect to the CDK, and cyclin C is rotated by 15° with respect to CDK8 (Baumli et al. 2008; Schneider et al. 2011). These rotations are accompanied by reduced contact between the cyclin and the CDK C-terminal lobe in the transcriptional CDK/cyclin complexes. The respective binding partners of CDK8/cyclin C and CDK9/cyclin T may exploit these variations in structure for binding specificity. Although this structural information is known, much is still

to be discovered about the structure and function of CDKs 11-13, and the complexes and regulation of CDKs 7-9.

1.9 Aims of this thesis

CDK9/cyclin T is a critical protein involved in the regulation of transcription. Since its discovery as a CTD kinase, much progress has been made in the field of transcription towards understanding its role in initiating Pol II elongation. However, there are still many unanswered questions preventing our complete knowledge of its cellular functions. Still little is understood about its interactions within multi-protein complexes. Furthermore, the molecular details by which CDK9/cyclin T is regulated by its interaction partners and recruited to sites of active transcription remains to be discovered. At the same time, its central role in normal and pathogenic processes (described in more detail in the following chapter) makes CDK9/cyclin T an excellent target for the development of inhibitors that might be of benefit in a variety of clinical settings.

In this thesis I focus on the structural and biochemical characterization of complexes formed by CDK9/cyclin T with small molecule chemical inhibitors and with Brd4, the third protein of the P-TEFb small activatory complex. My results provide a fine dissection of the P-TEFb/Brd4 interaction and demonstrate important functional consequences of Brd4-binding to CDK9/cyclin T. They further combine to suggest a small number of molecular models for the structural basis of the Brd4 interaction.

My results also contribute to an emerging picture of how small molecules can achieve selective inhibition of CDK9/cyclin T, as a step on the route to designing drugs and chemical probes that target this enzyme. Taken together, this work lays the foundation for future pharmacological and cell biological approaches to both illuminate the CDK9/cyclin T structure/function relationship, and to exploit that understanding in the rational design of therapies.

2 Chemical inhibitors of CDK9/cyclin T

2.1 Introduction

Several small molecule inhibitors have been synthesised targeting CDKs and there are two primary advantages to developing further small molecules specifically inhibiting CDK9/cyclin T. Firstly, inhibitors that are both specific and potent for CDK9 provide an essential tool to more closely examine the function of P-TEFb within the cell and will enable a more detailed understanding of transcription. Secondly, the role of CDK9 in transcription, and hence protein expression, has implicated CDK9 as a potential major drug target in many disease states and as such it is a candidate for structure-based drug design (Wang and Fischer 2008).

2.1.1 Disease states

Misregulation of P-TEFb is identified in certain pathologies (reviewed in (Romano and Giordano 2008)). In these cases, it is often upregulation of CDK9 that occurs. This results in modifying the fine-balance of cellular protein levels, which would ordinarily maintain the cell in a healthy state. A variety of phenotypes may result from CDK9/cyclin T misregulation that ultimately can lead to several disease states. These include cardiac hypertrophy and certain cancers. Below is a description of the current understanding of certain pathologies and the role that CDK9/cyclin T plays within them.

2.1.1.1 Cardiac hypertrophy

Patients affected by cardiac hypertrophy have enlarged cardiomyocytes and an increased heart size. This can be a result of an adaptive physiological response but in pathological cases, cardiac hypertrophy may result in cardiac arrest or sudden death (Frey and Olson 2003).

At a cellular level, these patients have high levels of protein and RNA in their cardiomyocytes, which also contain increased levels of CDK9 activity (Sano et al. 2004). The importance of this increase in CDK9 activity on cardiac hypertrophy symptoms has been ascertained by studies on cardiomyocytes and mouse models. These have shown that an altered proportion of the active CDK9 complex, either by overexpression of cyclin T1 or decreased Hexim1 expression, increases the propensity for a hypertrophic phenotype (Sano et al. 2002). Conversely, simulating hypertrophic conditions induces a decrease in CDK9/cyclin T associated with the inhibitory Hexim1 and 7SK factors. The signalling pathways involved in this process have not been fully mapped, but both the JAK/STAT and MAPK pathways as well as a micro-RNA, Mir-1 have been implicated (Sayed et al. 2007; Takaya et al. 2009). Of particular pharmacological interest is that downregulation of CDK9, either through knock-out strains or by use of the CDK9-specific inhibitor DRB, can control the growth of cardiomyoctes (Sano et al. 2002). This result supports a clinical application of a CDK9 specific drug in cardiac hypertrophy treatment.

2.1.1.2 Virology

The role of CDK9/cyclin T in RNA synthesis has resulted in several viruses utilising the cellular P-TEFb complex to their own advantage. These include HPV (human papilloma virus), HIV-1 and HIV-2, the Dengue virus, and HSV1 (Herpes Simplex Virus 1). This parasitic relationship enables the viruses to alter the cellular protein expression profile, and both stabilise and replicate their own genomes (Durand et al. 2005; Durand and Roizman 2008; Li et al. 2010).

The hijacking of CDK9/cyclin T by HIV has been studied extensively. As mentioned in the previous chapter, HIV utilises its Tat-TAR protein/RNA complex to dissociate P-TEFb from the inactive complex and recruit active P-TEFb to the HIV-LTR (Figure 2-1) (Contreras et al. 2007). This Tat/TAR/P-TEFb complex remains associated with the HIV transcription site for an extended period of time (compared to other recruitment factors) and is proposed to remain bound to the elongating polymerase (Molle et al. 2007).

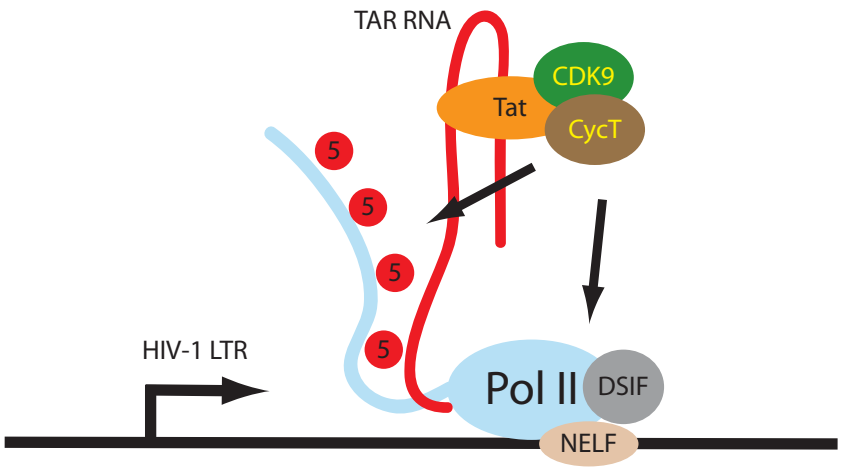


Figure 2-1. Model of recruitment of CDK9/cyclin T1 to the HIV-LTR. Nascent mRNA produces a hairpin structure, TAR. Tat interacts with both CDK9/cyclin T and TAR and therefore recruits P-TEFb to the HIV LTR promoter.

Several studies have shown HIV-driven transcription to be more sensitive to active P-TEFb levels than are cellular host processes (Chiu et al. 2004; Biglione et al. 2007). For this reason, and because the targeted protein will not be susceptible to resistance-conferring mutations, P-TEFb could be a possible protein drug target against HIV and potentially other viruses.

2.1.1.3 Cancer

CDK9 is a proposed drug target for several cancers, as indicated by diverse experiments ranging from cell studies to clinical trials.

On the cellular level, inhibitors of CDK9 reduce the transcription of proteins with short-lived mRNAs and thus subsequently reduce the concentration of these proteins within the cell. A number of these proteins have been identified and include cell cycle regulators (e.g. cyclin D and PLK), anti-apoptotic factors (e.g. Bcl-2, Mcl-1), and HDM2 (a regulator of p53, a protein involved in inducing apoptosis) (Konig et al. 1997; Lam et al. 2001; Alonso et al. 2003). Conversely, increased levels of active CDK9 would result in increased expression of these proteins, deregulating the cell cycle and preventing apoptosis.

Studies on several cell lines using a CDK9 specific inhibitor, flavopiridol, support the importance of CDK9 as a drug target. Cell lines tested include lymphoma cells, prostate carcinoma epithelial cells, breast carcinoma cells, head and neck squamous cell carcinomas (HNSCC), and melanoma cells. In all of these cases the effect of flavopiridol administration is either apoptosis or cell-cycle arrest (Drees

et al. 1997; Parker et al. 1998; Patel et al. 1998; Li et al. 2000; Lu et al. 2000). Furthermore, sensitivity to flavopiridol was shown to be independent of p53 status, and was apparent in cell lines that are insensitive to other antiapoptotic stimuli (Demidenko and Blagosklonny 2004).

Further support comes from in vivo experiments on Nu/nu mice where overexpression of cyclin T1 in nude mice resulted in tumor growth (Moiola et al. 2010). Moreover, CDK9 inhibition has been found effective in treatment of xenograft cancer models, generated by transfecting nude mice with carcinomas such as human small cell lung carcinoma, breast carcinoma, human non small cell lung cancer or leukemia. Results show either tumor growth delay, reduced tumor growth or regression when the mice were treated with flavopiridol (Arguello et al. 1998; Patel et al. 1998; Lu et al. 2000).

2.1.1.4 Inflammation

Furthermore, P-TEFb has been suggested as a target for the treatment of inflammation. The molecular basis for this is linked to the association between CDK9/cyclin T and NF-κB. NF-κB is a critical transcription factor activated by extracellular signals including those of inflammation (Oeckinghaus and Ghosh 2009). The interactions between RelA (a member of the NF-κB family) and both P-TEFb and Brd4 have indicated that P-TEFb regulates a subset of NF-κB controlled genes involved in inflammation (Nowak et al. 2008; Huang et al. 2009). This is further supported by the observations that several signalling cascades controlling inflammation ultimately require CDK9/cyclin T for

transcriptional regulation of specific gene sets and that inhibition of CDK9 reduces inflammation (Hou et al. 2007; Smallie et al. 2010; Schmerwitz et al. 2011). Therefore it has been suggested that P-TEFb is a possible target for treatment of inflammatory diseases (Schmerwitz et al. 2011).

2.1.2 Successful development of kinase inhibitors

The involvement of kinases in a number of diseases has made this class of proteins a target for drug discovery and development. Several drugs against kinases have successfully been approved for clinical use and many more are currently in clinical trials (reviewed in (Grant 2009; Johnson 2009a)). These include the monoclonal antibody, Herceptin, that targets receptor tyrosine kinases ErbB2 and HER2 in breast cancer. There are also several small molecule inhibitors including Gleevec, a BCR-Abl tyrosine kinase inhibitor. Gleevec has been particularly successful in treatment for chronic myelogenous leukemia (CML). Structure-based analysis has informed us of the mechanism of inhibition of several kinase targeting drugs (Noble et al. 2004), improved our understanding of drug-resistant mutations and is enabling the development of second generation and more potent inhibitors (reviewed in (Williams et al. 2005; Schenone et al. 2011)). This combination of results supports structure based inhibitor design against kinases as a valid approach for developing CDK9-specific compounds.

2.1.3 Present availability of inhibitors

There is a range of inhibitors currently available that have been identified, isolated and developed against the CDK/cyclin family of kinases (Table 2-1). They exhibit a variety of chemical scaffolds (Figure 2-2) and a corresponding breadth in potency and specificity. Several are pan-CDK inhibitors such as roscovitine but a few show selectivity towards a specific CDK/cyclin. Currently, DRB and flavopiridol are described as relatively CDK9-specific inhibitors and are often used to probe the cellular function of CDK9/cyclin T.

	CDK IC_{50} or K_i (μM)				
Compound	1	2	4	7	9
DRB	17	>10	>10	>10	0.34
Flavopiridol	0.07	0.19	0.34	0.11	0.003
Roscovitine	1.3	0.25	>10	0.25	0.79
SNS-032	0.28	0.05	0.93	0.06	0.005
R547	0.0001	0.0003	0.001	0.17	0.005
CAN508	44	69	14	26	0.35
Indirubin 3′-monoxime	0.18	0.44	3.3	> 4	0.05
Meriolin 3	0.17	0.011	> 0.1	> 0.1	0.006
(S)-CR8	0.15	0.08	N/A	N/A	0.11
S3-35	0.012	0.004	N/A	0.114	0.006

Table 2-1. Selectivity of various CDK inhibitors adapted from Wang and Fischer (Wang and Fischer 2008) with additional compounds from (Bettayeb et al. 2008; Liu et al. 2011b). Either the half maximal inhibitory concentration (IC_{50}) or the dissociation constant (K_i) are stated.

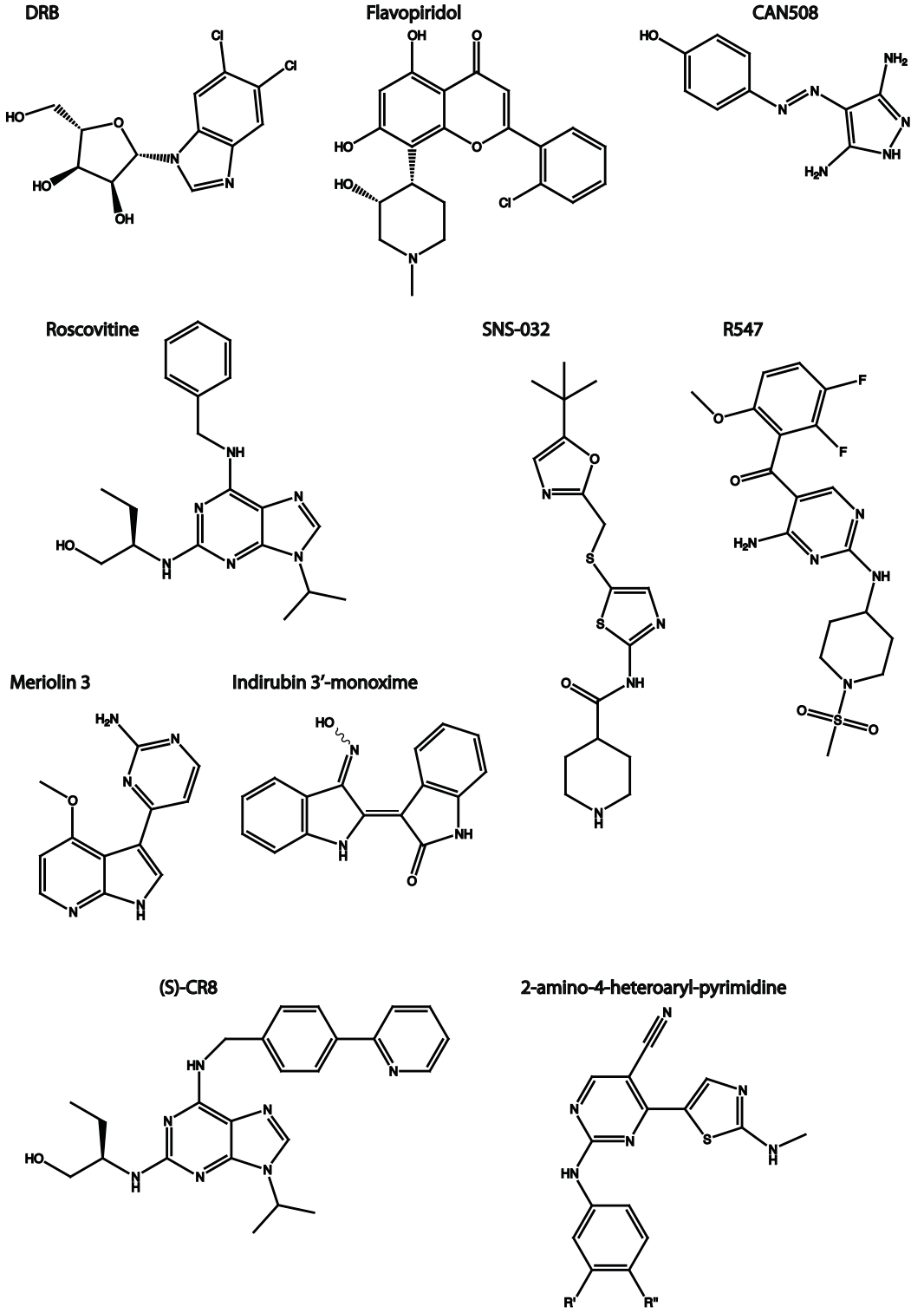


Figure 2-2. Chemical structures of a selection of CDK inhibitors.

Several of these CDK inhibitors including those targeting CDK9 have already entered the clinical trial stages of drug development. There are nearly 60 clinical trials that are either completed or in progress administering flavopiridol solely or in combination with other chemotherapeutic drugs. These are directed at patients with cancers that include leukemias, lymphomas, breast, prostate, ovarian, kidney, liver and gastric cancers. In addition there are several other CDK inhibitors, such as P276-00, SNS-032, R547 and roscovitine, that have also entered clinical trials (clinicaltrials.gov).

Unfortunately these drugs are not without their problems. Administration of flavopiridol during clinical trials has required unexpectedly high dosages, and in addition, this compound has poor solubility and induces DNA double strand breaks in the cell (Liu et al. 2011a). The pharmacokinetics of roscovitine require improvement as this compound is rapidly metabolised in vivo to a less active compound, and it also has poor specificity and relatively low potency (Nutley et al. 2005; Vita et al. 2005; McClue and Stuart 2008). Finally, SNS-032 has been shown to have poor bioavailability in certain animal models as it appears to be a P-glycoprotein substrate and can therefore be efficiently exported from cells (Kamath et al. 2005).

One approach to produce drugs with a higher therapeutic index is to improve selectivity for CDK9 over other kinases, thus reducing the side effects and improving the success of clinical trials. Hence development of second generation inhibitors is under way.

2.1.4 Development of CDK9-specific inhibitors

Many inhibitors have been designed and synthesized by our collaborators (L. Meijer (CNRS, Roscoff, France) and S. Wang (University of Nottingham, UK)) to target CDK9. Using a combination of kinetic studies and structural information we intend to develop inhibitors with improved CDK9 selectivity and potency. The first compound described is (S)-CR8, a roscovitine derivitive, and the second is a class of 2-amino-4-heteroaryl-pyrimidine compounds.

(S)-CR8 was synthesized as a second-generation analogue of (R)-roscovitine (Oumata et al. 2008). Although roscovitine is in clinical trials, its pharmacokinetic properties and poor potency towards cell lines provides opportunity for improvement. (S)-CR8 is related to (R)-roscovitine and MPP^+ (1-methyl-4-phenylpyridinium). MPP^+ is absorbed into cells through the dopamine transporter, and thus (S)-CR8 was designed to be more potent to cells. This approach was successful, with (S)-CR8 having an increased potency of *circa* 50 fold towards a range of cell lines in comparison to roscovitine (Bettayeb et al. 2008). Its presence decreases Pol II phosphorylation by *circa* 50 fold. And in neuroblastoma cells this results in decreased expression of Mcl-1, an anti-apoptotic factor, and increased levels of free Noxa, a pro-apoptotic factor, resulting in cell death (Bettayeb et al. 2010).

The 2-amino-4-heteroaryl-pyrimidine series were developed from initial hits from CDK-targeting virtual screens. Subsequent structure-based design in combination with biochemical studies led to further development of these inhibitors with improved activity (Wu et al. 2003b). Many chemical variants of

these have been synthesized and studied for their *in vitro* activity against CDKs and in some cases for their cellular activity. Some structure-activity relationships (SAR) have allowed the development of models to explain difference in potency and selectivity towards the CDKs. Extensive cellular studies of S3-35 have been carried out on both normal and cancerous cell lines (Liu et al. 2011b). These showed that the activity of S3-35 both specifically targets CDK9 in cells, and induces apoptosis in cancerous cells more readily than in normal cells. These studies also showed that, unlike flavopiridol, S3-35 does not induce DNA double-strand breaks (Liu et al. 2011b).

In vitro kinetic studies were undertaken by our collaborators in order to determine K_i values for the various inhibitors studied (Table 2-1, Table 2-2). These experiments were performed to assess inhibitory activity against CDK1, CDK2, CDK7, as well as CDK9 providing an indication of the selectivity of the inhibitors within the CDK family. The potency against CDK9/cyclin T of the majority of these inhibitors is extremely promising.

	Kinase inhibition K_i (μM)			
Inhibitor	CDK1/A	CDK2/E	CDK7/H	CDK9/T1
S2-106	0.217	0.147	1.871	0.043
S3-35	0.012	0.004	0.114	0.006
S2-104	0.007	0.012	0.173	0.011
S2-83	0.072	0.123	0.405	0.022
S-77	0.094	0.568	0.045	0.007
S-68	0.091	0.139	0.208	0.007
S2-122	N/A	N/A	N/A	>15
S2-124	N/A	N/A	N/A	6.185

Table 2-2. K_i values for inhibitors within the 2-amino-4-heteroaryl-pyrimidine series.

2.1.5 Structural features of CDK9 for exploitation by ATP competitive inhibitors

To date there are only three published structures of CDK9/cyclin T in the presence of an inhibitor. These are flavopiridol (Baumli et al. 2008), (S)-CR8 (described in the results section) (Bettayeb et al. 2010) and DRB (Baumli et al. 2010). Contrastingly, there are many structures of CDK2/inhibitor and CDK2/cyclin A/inhibitor complexes deposited in the protein data bank (PDB).

This is as a result of early models of eukaryotic cell-cycle regulation providing a strong rationale for exploiting CDK2 as an anti-cancer drug target. This reasoning drove several successful inhibitor discovery programs that developed robust protocols for preparing CDK2-containing crystals. This opened up further possibilities for using CDK2 as a surrogate drug-receptor in structural studies of kinase inhibitor-binding. The extensive number of structures of CDK2/cyclin A/inhibitors deposited has resulted in the ATP-binding site of CDK2 being described in terms of potential drug-targeting pockets for CDKs in general.

Liao (Liao 2007) described the ATP binding site in the active kinase as containing two different clefts, a front and back cleft. The back cleft is protected by the gatekeeper residue, which in both CDK2 and CDK9 is a bulky phenylalanine (residues Phe80 (CDK2) and Phe103 (CDK9)). Thus this back pocket is inaccessible to small molecular probes in CDK2 and CDK9 alike. The front cleft has three main pockets that are defined by the positions of the ATP groups when bound: the A pocket (adenine binding pocket), the R pockets (ribose binding pocket) and the P pockets (where the phosphates are located) (Figure 2-3 (B)). All pockets have contributions from residues within both the N-

and C- terminal lobes of CDKs. The adenine pocket has additional interactions with the hinge residues, which provide hydrogen donors and acceptors to this hydrophobic pocket. These hinge residues are often exploited by kinase inhibitors (Ghose et al. 2008). The glycine-rich loop is a highly flexible loop forming the roof of the ATP-binding site and its residues line both the ribose and phosphate pockets. However, because of its flexibility it is able to adopt a range of conformations, which makes it challenging to infer its effect on inhibitor binding. Because of this, and because it is solvent-exposed, the phosphate pocket is suggested to be a less important pocket to target in comparison with the A and R pockets. As well as the pockets that are identified by ATP binding, two additional solvent exposed pockets are described, E_0 (or hydrophobic pocket I) and E_1, and an additional pocket, K, has been located at the back of the ribose pocket (Figure 2-3) (Zhang et al. 2009). Few residues differ between CDK2 and CDK9 within the ATP binding site and these are shown in Figure 2-3.

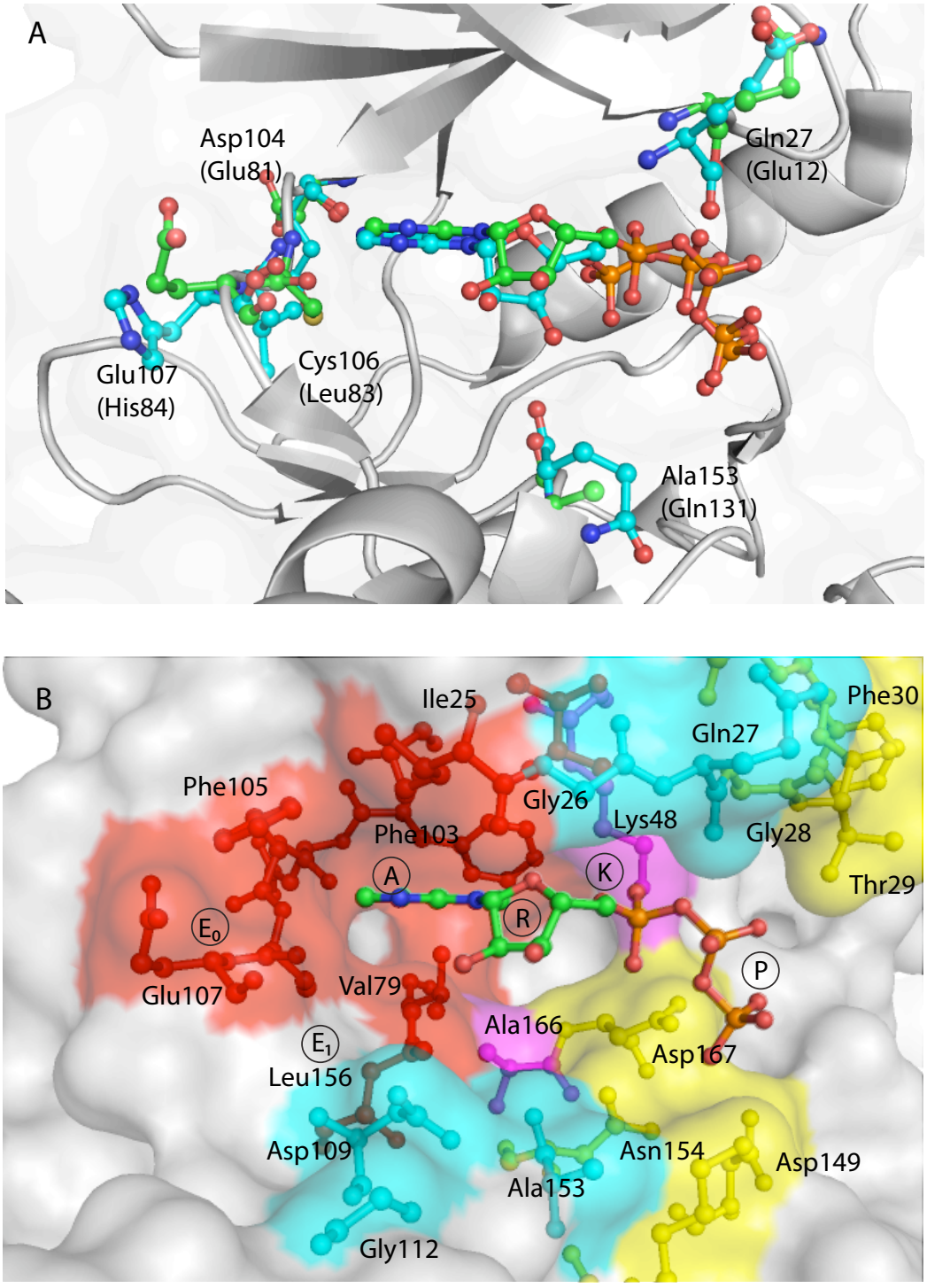


Figure 2-3. The CDK9 ATP-binding site. (A) Residues that are located within 5Å of ATP in the CDK9 ATP-binding site and differ between CDK9 (green) and CDK2 (blue) are depicted in ball and stick mode. CDK9 residues are labelled above the equivalent CDK2 residues, in parenthesis. (B) CDK9 residues that contribute to the ATP binding site pockets. The surfaces of the pockets are shown in red – the A pocket; cyan – the R pocket; yellow – the P pocket; and magenta – the K pocket. The locations of the A, R, P, K, E_0 and E_1 pockets are additionally shown on the structure by their respective letters.

As well as serving as an experimental surrogate, CDK2-inhibitor structures can indicate how inhibitors of other kinases, including CDK9, might bind to a related off-target enzyme. Such information might be useful in understanding and enhancing inhibitor specificity. Therefore the structural studies presented in this chapter include studies on CDK2 as well as CDK9.

2.1.6 Aims

CDK9 is a validated drug target for several medical conditions. Comparative structural characterization of ATP-competitive inhibitors bound to CDK9 and CDK2 can provide information about the chemical character of the inhibitors studied. It can also demonstrate the unique steric and chemical properties that each CDK possesses, and that might be exploited in specific inhibitor design.

Combining structural, binding and kinetic inhibition data on inhibitors developed against CDK9/cyclin T will enable better understanding of how CDK9-specific inhibition may be achieved, thereby providing valuable knowledge for further inhibitor developments. From these studies proposals can be suggested for improving both potency and selectivity of inhibitors.

The aims of the work presented in this chapter are:

- To validate the use of a mutated form of cyclin T1, possessing superior crystallogenesis properties, as an appropriate tool for use in inhibitor binding studies.

- To assess the binding to CDK9/cyclin T of a range of inhibitors, including the roscovitine derivative, (S)-CR8, and compounds from the 2-amino-4-heteroaryl-pyrimidine series.
- To conduct crystallographic analysis of these inhibitors bound to both CDK9/cyclin T and CDK2/cyclin A in order to support structure-based inhibitor design.

2.2 Methods

For detailed general laboratory methods please refer to Appendix A.

2.2.1 CDK9/cyclin T protein constructs

Several CDK9/cyclin T protein constructs were used throughout this thesis. These are summarized in Table 2-3.

Protein Name	Protein residues within construct (mutations)
$CDK9_{FL}$	1-372
$CDK9_{330}$	1-330
$CDK9_{330\ T186A}$	1-330, (T186A)
cyclin T_{WT}	1-259
cyclin T_{F241L}	1-259, (F241L)
cyclin $T_{Q77R,\ E96G,\ F241L}$	1-259, (Q77R, E96G, F241L)
cyclin T_{288}	1-288

Table 2-3. CDK9 and cyclin T constructs. Throughout the text, constructs are identified by the protein name provided here.

These constructs were selected based on their properties. The complex of $CDK9_{330}$/cyclin $T_{Q77R,\ E96G,\ F241L}$ can be crystallised for structural analysis of P-TEFb (Baumli et al. 2008). Kinetic analysis was primarily undertaken with the full-length CDK9 ($CDK9_{FL}$)/cyclin T (residues 1-259) (cyclin T_{WT}). This complex contains the cyclin boxes of cyclin T and the full-length CDK9 that includes the C-terminal tail required for the ordered kinetic reaction mechanism of phosphorylation. The mutation of T186A in CDK9 prevents the phosphorylation of this residue within the activation segment and thus inactivates the kinase. Cyclin T_{F241L} was produced to determine the influence of the E96G mutation introduced for crystallographic purposes. Finally cyclin T_{288} was used for H/D

mass spectrometry providing a longer cyclin T construct to probe the Brd4-cyclin T interaction.

Phe241Leu was introduced into the cyclin T_{WT} gene within pVL1393 using Quikchange II Site-Directed Mutagenesis Kit (Stratagene). After sequence conformation, the plasmid was transfected into *Spodoptera frugiperda 9* (Sf9) cells following instructions from the BD BaculoGold transfection manual (BD Bioscience). The virus was amplified 3 times before use in expressions.

2.2.2 Inhibitor Synthesis

Inhibitors were provided as lyophilized solids. (S)-CR8 was kindly provided by L. Meijer (CNRS, Roscoff, France) and the 2-amino-4-heteroaryl-pyrimidine series of compounds were provided by S. Wang (University of Nottingham). Synthesis methodology of several of these inhibitors have been published (Galons et al. 2009; Wang et al. 2010). When required, inhibitors were resuspended in DMSO at concentrations between 10mM and 50mM and subsequently stored at -20°C.

2.2.3 CDK2/cyclin A expression and purification

For interaction and structural studies, CDK2 (corresponding to the full-length human protein) and cyclin A (residues 174-432 of the human protein cyclin A2) were expressed individually and purified as the CDK/cyclin complex. GST-CDK2/GST-CIV1 and cyclin A were expressed separately in *E. coli* strain B834. Cells were grown in lysogeny broth (LB) to an OD_{600} <1. Once sufficient cell density had been reached, CDK2 expression was induced for over 20 hours at

18°C with 80μM isopropyl β-D-1-thiogalactopyranoside (IPTG). Cyclin A expression was induced overnight at 18°C with 100μM IPTG. Cells expressing CDK2 were harvested by centrifugation and resuspended in modified HBS (0.2M NaCl, 40mM Hepes, pH 7.0, 0.01% MTG containing protease inhibitors (Roche)) whilst cyclin A-expressing cells were resuspended in RB (0.3M NaCl, 50mM Tris, pH 7.0, 10mM $MgCl_2$, 0.01% MTG, containing protease inhibitors). All pellets were stored at -20°C until use.

For purification of CDK2/cyclin A, cell suspensions were thawed and lysed by the addition of lysozyme and sonication. Cell lysates were clarified by centrifugation. Clarified GST-CDK2-containing lysate was loaded onto a gravity flow column containing glutathione Sepharose 4B beads (GE Healthcare) and washed with modified HBS. In a second step, clarified cyclin A lysate was loaded onto the GST-CDK2 bound glutathione beads. These beads, bound with GST-CDK2/cyclin A, were washed with modified HBS and the protein complex was eluted in 20mM glutathione in modified HBS buffer. The elution was incubated with GST-3C-protease overnight. The solution containing the cleaved CDK2/cyclin A was injected on a Superdex 200 (26/60) size-exclusion chromatography column (GE Healthcare), pre-equilibrated in modified HBS. The fractions containing CDK2/cyclin A were combined and the remaining GST and GST-3C protease was removed on a gravity flow column containing glutathione Sepharose 4B beads. The final purified protein was concentrated to 11.9mg/ml, flash frozen and stored at -80°C.

2.2.4 CDK9/cyclin T expression and purification

All GST-CDK9 and GST-cyclin T proteins were co-expressed in *Sf9* cells at 1million cells/ml. Cyclin T baculovirus was infected at a multiplicity of infection (MOI) >0.3 and CDK9 at MOI >0.5. Cells were incubated at 27°C for 72 hrs before harvesting.

Insect cells were collected by centrifugation (250g for 20 minutes at 4°C) and resuspended in buffer containing 500mM NaCl, 20mM Tris pH 8.0, 5mM DTT, 10% glycerol and protease inhibitors. Cells were lysed on ice by sonication and lysate was clarified by centrifugation. GST-fusion proteins were purified from the clarified lysate by glutathione Sepharose affinity chromatography using a gravity flow column containing glutathione Sepharose 4B beads. Beads were washed with buffer containing 1M NaCl, 20mM Tris pH 8.0, 5mM DTT, 10% glycerol and subsequently with buffer containing 200mM sodium sulfate, 20mM Tris pH 8.0, 5mM DTT, 10% glycerol. After 3C cleavage overnight, $CDK9_{330}$/cyclin T protein constructs were autophosphorylated for 1hr at 30°C in 500mM NaCl, 20mM Tris pH 8.0, 5mM DTT, 10% glycerol supplemented with 0.1mM ATP and 10μM $MgCl_2$. CDK9/cyclin T proteins were purified on a Superdex 200 (26/60) size exclusion chromatography column, preequilibrated in buffer containing 500mM NaCl, 20mM Tris pH 8.0, 5mM DTT and 10% glycerol. CDK9/cyclin T containing fractions were pooled and the remaining GST was subsequently removed on glutathione Sepharose 4B beads. CDK9/cyclin T was concentrated to 3.9mg/ml for crystallisation.

2.2.5 CDK2/cyclin A crystallisation

Purified CDK2/cyclin A was incubated on ice for 20 minutes in the presence of the inhibitors at 0.2mM concentration. The solution was filtered through a microfiltration unit (Amicon ultra-free-MC) before setting hanging drops of 1μl (protein solution) and 1μl (reservoir solution) and storing at 4°C. Crystallisation reservoir conditions were screened between 1-1.25M ammonium sulfate, 0.5-0.9M KCl, in the presence of 100mM Hepes, pH 7.0, and 5mM DTT. All reservoir solutions were 1ml in volume. Crystals were fished and cryo-protected in 7M sodium formate before flashcooling in liquid nitrogen.

2.2.6 CDK9/cyclin T crystallisation

$CDK9_{330}$/cyclin $T_{Q77R,\ E96G,\ F241L}$ was crystallised at 4°C by sitting drop methods. Drops were 2μl in volume with a 1:1 protein:reservoir solution ratio. The reservoirs were a total volume of 1ml and consisted of 10-16% PEG 1000, 100mM NaK-phosphate pH 6.2, 500mM NaCl, and 4mM TCEP. Crystals were soaked in the reservoir solution containing 1mM inhibitor before buffer exchange into a cryo-protectant solution of 15% glycerol, 1mM inhibitor in the reservoir solution. CDK9/cyclin T/inhibitor crystals were fished and flash frozen in liquid nitrogen.

$CDK9_{330}$/cyclin T_{F241L} was crystallised at 3.8mg/ml in the same manner as $CDK9_{330}$/cyclin $T_{Q77R,\ E96G,\ F241L}$. However, the reservoir solution required to crystallise this complex was 8% PEG 1000, 100mM NaK-phosphate pH 6.2, 600mM NaCl, 4mM TCEP.

2.2.7 Data collection and processing

Data was collected at Diamond light source, on beamlines I-02, I-03, I-04 or I24. Mosflm (1994) or XDS (Kabsch 2010) was used for indexing and integrating data. Data was scaled using Scala. Initial models were produced by molecular replacement using phaser or rigid body refinement in phenix.refine (Adams et al. 2010). Models were iteratively improved by alternate rounds of refinement using either Refmac or phenix.refine and rebuilding in Coot (Emsley et al. 2010). Restraints were produced in Prosmart and applied to CDK9/cyclin T structures refined in Refmac. Inhibitor restraints were produced using phenix.elbow builder and confirmed by visual inspection of the resulting PDB and cif files.

2.2.8 Thermal denaturation data

0.95μg protein was incubated with 20μM inhibitor, 5x sypro orange (Molecular Probes) in 20mM Hepes pH 7.0, 500mM NaCl, 10% glycerol, 3mM DTT. The total reaction volume was 15μl. The temperature was raised in 1°C steps initating from 25°C and terminating at 80°C. The fluorescence emitted at 570nm was measured using a Stratagene mx3005P RT-PCR 305 machine.

2.3 Results

2.3.1 Stabilisation of CDKs in the presence of inhibitors

K_i or IC_{50} values have been determined for a range of inhibitors against CDK9 and other CDKs (Table 2-1 and Table 2-2). These inhibitors display a range of potency and specificity for the CDKs tested. To augment our knowledge of a selected range of these inhibitors and of the CDK9/cyclin T ATP binding site, differential scanning fluorimetry (DSF) was undertaken on CDK9/cyclin T complexes in the presence and absence of a subset of inhibitors.

DSF assesses the inhibitor/protein interaction by measuring the difference in a protein's melting temperature (T_m) in the presence and absence of the inhibitor. A large positive difference in melting temperature (ΔT_m) suggests that the protein is stabilized in the presence of the inhibitor and hence interacts favourably. Small or no changes in melting temperature imply little stability gained from the interaction, and hence weak binding of the inhibitor. Correlations between ΔT_m and both binding constants and inhibition have previously been observed for inhibitors of the PIM-1 kinase (Bullock et al. 2005). This approach is marketed under the name Thermofluor™ by Johnson & Johnson (Cummings et al. 2006). The Structural Genomics Consortium in Oxford have demonstrated that a similar approach can be extremely effective in identifying and characterizing ligands of protein kinases (Fedorov et al. 2007), and it has consequently been adopted in this laboratory for that purpose.

Thermofluor has several advantages as a technique in studying inhibitor-protein interactions such as those for CDK9/cyclin T inhibitors. Firstly, it complements the kinetic analysis by providing a method to distinguish between kinetic inhibition and binding. Secondly, it has the benefit of allowing alternative protein constructs to be compared independently of their original activity. And thirdly, it has the potential to be a standard simple high-throughput laboratory technique.

In order to validate this technique for the 2-amino-4-heteroaryl-pyrimidine series of inhibitors the stabilization of $CDK9_{FL}$/cyclin T_{WT} induced by the presence of a selection of both standard CDK inhibitors and members of this series (S2-106, S3-35, S2-104, S2-83, S-77, S-68, S2-122 and S2-124) were analysed (Figure 2-4).

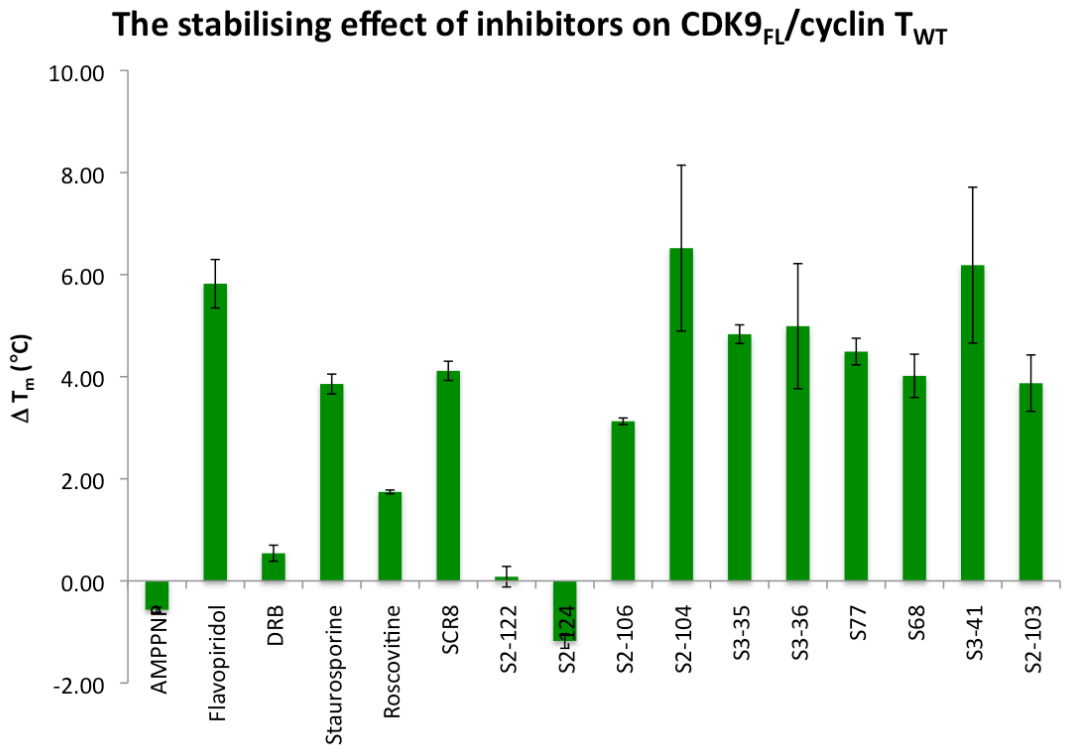
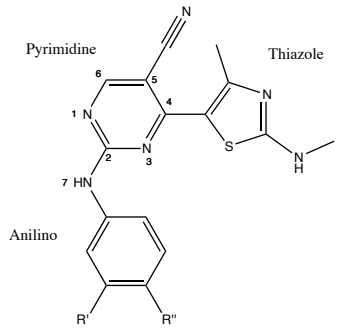


Figure 2-4. The stabilizing effect of inhibitors on CDK9/cyclin T as measured by Thermofluor. Experiments were repeated 3 times and error bars represent standard deviations.

Initial analysis of two well-characterised CDK9 inhibitors, flavopiridol and DRB, indicated encouraging results for this technique. Flavopiridol is one of the most potent CDK9 inhibitors published to date (K_i : 0.003μM) and correspondingly generates a ΔT_m of 5.8 ± 0.5°C indicating flavopiridol exerts a large stabilizing effect on $CDK9_{FL}$/cyclin T_{WT}. DRB, although selective, is much less potent (K_i : 0.34μM) than flavopiridol. This is reflected in the DSF results in which it stabilizes CDK9 by 0.5 ± 0.2°C. The two compounds, roscovitine and (S)-CR8, which share the same chemical scaffold, also provide support for this technique. $CDK9_{FL}$/$CylinT_{WT}$ in the presence of (S)-CR8 remains stable at a higher temperature (ΔT_m : 4.1 ± 0.2°C) then in the presence of roscovitine (ΔT_m : 1.7 ± 0.04°C). Therefore implying that the second generation compound of roscovitine, (S)-CR8, provides more stability than its predecessor.

The 2-amino-4-heteroaryl-pyrimidine compounds can be separated into two groups according to the DSF data. The first group contains inhibitors S2-122 and S2-124 and the second group contains the remaining 8 compounds analysed. The first set of compounds (S2-122 and S2-124) show a destabilizing effect on CDK9/cyclin T in the DSF assay and have K_i values greater than 6μM. These two compounds are distinct from the other compounds tested within this series as they contain an additional bulky benzene ring on the thiazole ring. This benzene is likely to be poorly accommodated within the CDK9 ATP binding site and therefore neither inhibit nor stabilize CDK9/cyclin T to the same extent as the other compounds (Figure 2-5).

A



B

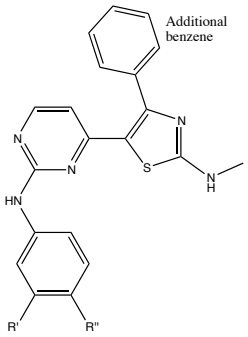


Figure 2-5. The chemical scaffold of the 2-amino-4-heteroaryl-pyrimidine series (A). The scaffold of inhibitors S2-124 and S2-122 (B).

Conversely the remaining 8 compounds have K_i values of less than 0.05μM. These inhibition values are reflected in the positive ΔT_ms of greater than 3°C measured for these compounds. Furthermore, S3-41 has both a comparable ΔT_m (6.2 ± 1.5°C) and K_i (0.006μM) to flavopiridol. The correlation observed between K_i and ΔT_m values serves to validate its use in subsequent studies of inhibitor binding to CDK9.

Thermal denaturation can also be used to compare inhibitor-mediated stabilization of different kinases, and hence provide a measure of inhibitor selectivity. For this purpose, DSF data was collected for a panel of CDK9 inhibitors binding to CDK2/cyclin A. A comparison of thermal denaturation profiles for $CDK9_{FL}$/cyclin T_{WT} and CDK2/cyclin A in the presence of certain inhibitors (Figure 2-6) is discussed below.

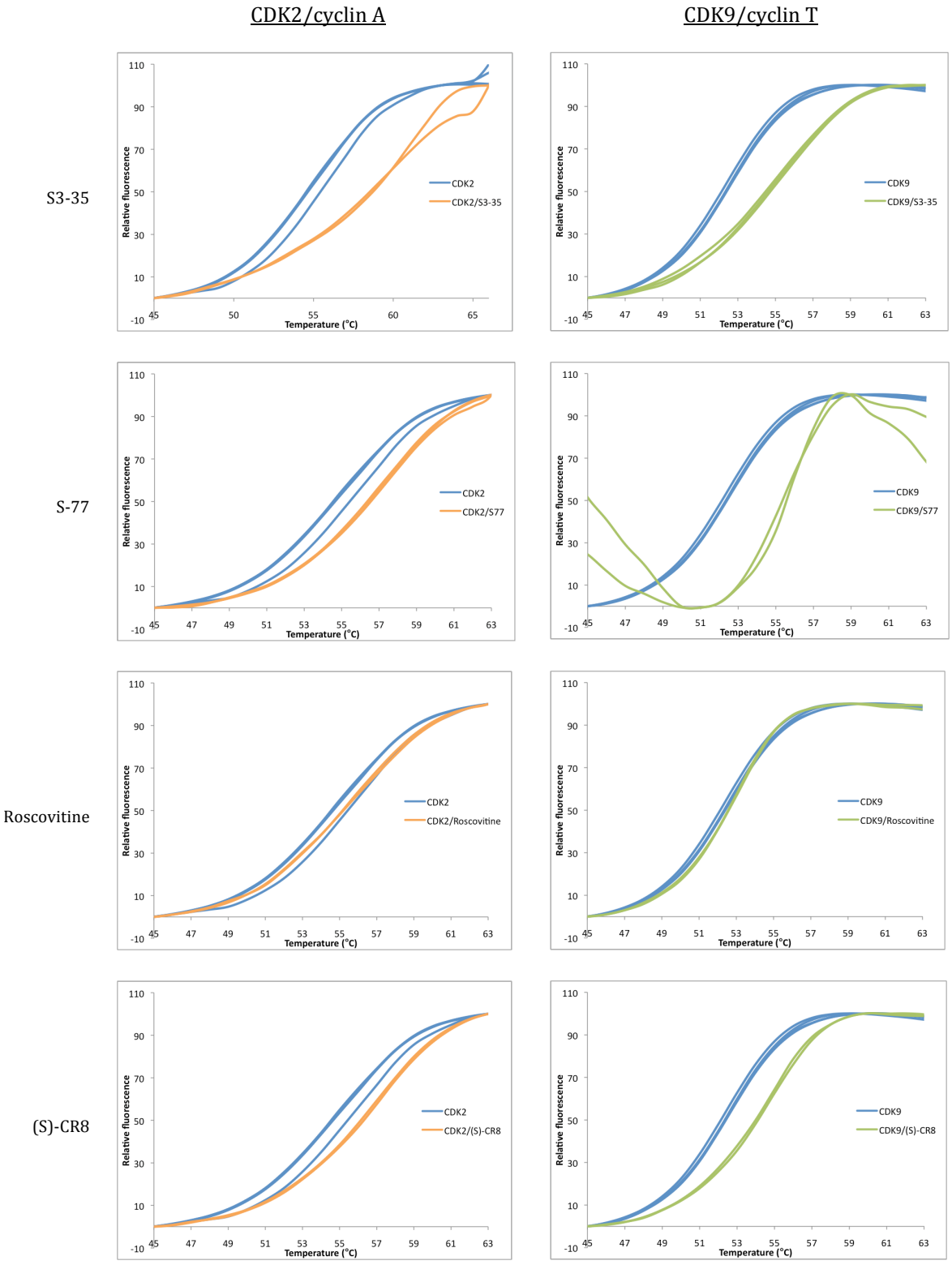


Figure 2-6. Thermal denaturation data for CDK2/cyclin A and CDK9/Cyclin T in the presence of a selection of inhibitors. A minimum of two representative curves is shown for each experiment.

S3-35 is a compound with similarly high K_i values for both CDK9 (K_i : 0.006μM) and CDK2 (K_i : 0.004μM) and correspondingly the DSF curves show significant stabilization of both enzymes. Conversely, S77 is selective for CDK9 (K_i : 0.007μM) over CDK2 (K_i : 0.568μM) and as such induces a much larger T_m shift for CDK9 than for CDK2. Finally, the presence of roscovitine seems to stabilize both CDKs to only a minor extent, whereas (S)-CR8 stabilised CDK9 much more substantially than CDK2. These results correlate well with the IC_{50} results (IC_{50}s - Roscovitine CDK2 : 0.25μM, CDK9 : 0.79μM; (S)-CR8 CDK2 : 0.08μM, CDK9 : 0.11μM).

Thirdly this technique was used to study the effect that inhibitors have on different constructs of CDK9. During catalysis, the C-terminus of CDK9 folds across the ATP binding site and is necessary for ordering the enzymatic pathway of CDK9 (Baumli et al, manuscript in preparation) and in some instances the CDK9 C-terminus influences inhibitor binding (Baumli et al. 2010). To determine if the C-terminal tail influences the binding of (S)-CR8 and the 2-amino-4-heteroaryl-pyrimidine we tested the stabilization of a truncated version of $CDK9_{330}$/cyclin T_{WT} and compared the results to those for $CDK9_{FL}$/cyclin T_{WT} (Figure 2-7).

The inhibitors tested, with the exception of S2-103, show little difference in the stabilizing effect on $CDK9_{FL}$/cyclin T_{WT} and $CDK9_{330}$/cyclin T_{WT} suggesting that, in general, the C-terminal tail neither forms any additional interactions with the inhibitors nor displaces the inhibitors.

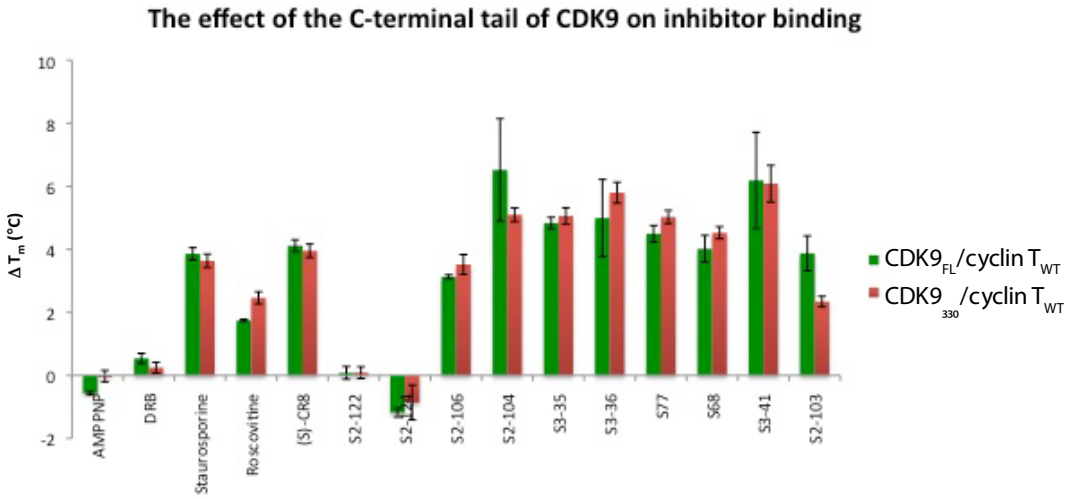


Figure 2-7. The stabilization of CDK9 length variants in the presence of different inhibitors.

2.3.2 CDK9/cyclin T crystal optimization

To date, several crystal structures of CDK9/cyclin T/inhibitor complexes have been published with the same protein constructs (Baumli et al. 2008; Baumli et al. 2010; Bettayeb et al. 2010). These structures have shown that this CDK9/cyclin T construct and crystal form can be reproducibly formed, soaked with inhibitors and diffract reliably to a resolution of *circa* 3Å. These published crystals have the potential to be modified which thus may result in an improvement in the quality of the data obtained. Firstly, an increased resolution would benefit the interpretation of the data tremendously. And secondly, the cyclin T construct includes 3 mutations within its sequence: Gln77Arg, Glu96Gly, Phe241Leu (Q77R, E96G, F241L) and the effect of these mutations on the CDK/cyclin structure is unknown (see Figure 2-8 for the location of these mutations on Cyclin T).

An alternative CDK9 construct was designed, starting from residue 8, with the ambition to improve the resolution of the crystal diffraction. The first 7 residues are not apparent in the electron density, so their removal would remove a flexible region from within the crystals and potentially improve their order. Unfortunately, the truncated CDK9 protein was not expressed in insect cells even after trials with varied viral MOIs.

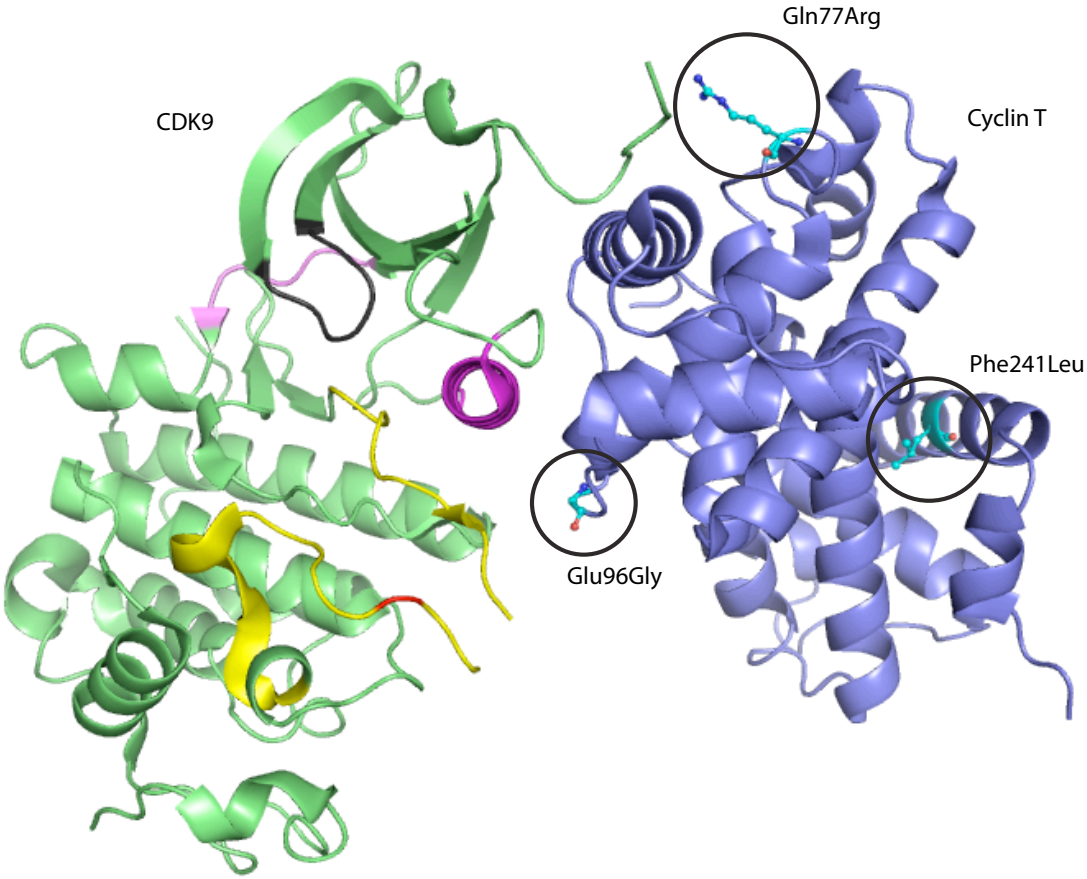


Figure 2-8. The location of mutations Gln77Arg, Glu96Gly and Phe241Leu on Cyclin T. The following features are shown in cartoon with the colours they are presented in in parenthesis: cyclin T (slate), CDK9 (green), the C-helix (magenta), the hinge region (violet), the glycine-rich loop (dark grey), the activation segment (yellow) and phosphothreonine 186 of CDK9 (red). Residues 77, 96, 241 of cyclin T are highlighted and presented in cyan using a ball and stick representation.

The effect of the three mutations within cyclin T on the structure of CDK9/cyclin T is unknown. Residues 77 and 241 are located on the top and far side of the cyclin respectively (Figure 2-8) and are therefore remote and not

expected to influence the enzymatic activity or ATP-binding site of the kinase. Residue 96 of cyclin T is located in the interface between the C-terminal lobe of CDK9 and the N-terminal cyclin box of cyclin T. It is also adjacent to the C-helix of CDK9 and close to the activation segment phosphothreonine residue. Therefore, this mutation has the potential to alter catalytic and inhibitor-binding properties of the enzyme.

To address this issue, I coexpressed cyclin T_{WT} and GST-$CDK9_{330}$. Subsequent purification produced a homogenous sample of the $CDK9_{330}$/cyclin T_{WT} complex. Sparse matrix screens of crystallisation conditions were set up with the protein. Crystals formed in many conditions but their primary morphologies were very fine needles or needle clusters. Optimisation screens based on several of these conditions were set up. Crystals from both the sparse matrix screen and optimisation screens were collected and sent to microfocus beamlines. Unfortunately no protein diffraction was visible.

The C-terminal helix of cyclin T is involved in crystal contacts. The mutation Phe241Leu (F241L) indirectly aids crystal contact formation through stabilising a helical conformation of this region within the cyclin. To aid crystallisation, I reintroduced the mutation into cyclin T with the hypothesis that this CDK9/cyclin T complex would crystallise containing a single mutation distant from the ATP binding site of CDK9. Crystals of the $CDK9_{330}$/cyclin T_{F241L} complex could reproducibly be produced under similar crystallisation conditions to $CDK9_{330}$/cyclin $T_{Q77R, E96G, F241L}$. These crystals diffracted to *circa* 3.5Å on the I04-

1 beamline at the Diamond Light Source. A dataset from a crystal was collected on beamline I24 that diffracted to 3.3Å.

To prevent model bias when rebuilding the structure, initial maps were phased using a model derived from rigid body refinement of PDB:3BLH, excluding the CDK9 C-helix (residues 55-75) and a section of the activation segment (residues 170-176). External structure restraints for Refmac (1994) were also produced from this truncated PDB file. After a round of TLS (Translation, Liberation and Screw tensors) and restrained refinement, clear difference density was visible for the C-helix and activation segment. Overlay of 3BLH showed no alteration of these structural elements and they were subsequently rebuilt. Cyclin T residue 96 and CDK9 residue 65 were refined as glycine and alanine respectively until a more complete model was built.

The overall structure of both the CDK and the cyclin are very similar to both the apo (3BLH) and AMPPNP bound structures (3BLQ) of CDK9/cyclin T (which both contain the triple mutant) (Figure 2-9). The glycine rich loop is unstructured in the structure presented here, and the β3-αC loop has a slightly altered conformation. However the C-helix remains in the same orientation and position as the other structures.

Clear electron difference density was visible for both reverted residues, Gly96 and Ala65. Rebuilding the correct residues showed hydrogen bonds between Arg65 (CDK9) and Gly96 (cyclin T). The electron density for Arg65 does not extend sufficiently to support interactions towards the phosphothreonine,

although the weak electron density around both the arginine and phosphothreonine residues suggests flexibility that might allow for transient interactions.

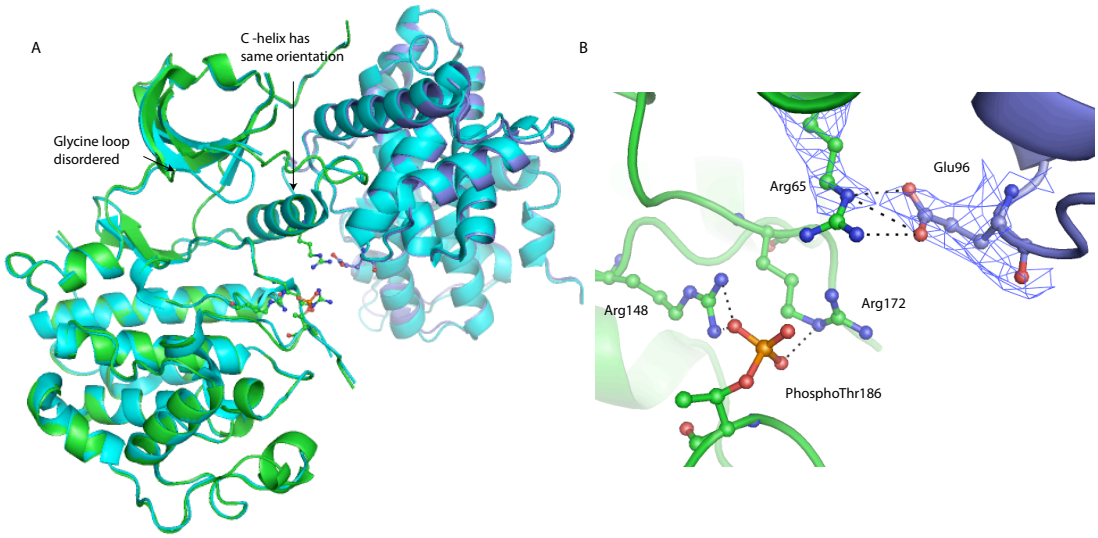


Figure 2-9. The effect of cyclin T mutation, Glu96Gly, on the CDK9/cyclin T structure. Structures are overlaid (A) and a close up figure of the Glu96 and interacting residues in $CDK9_{330}$/cyclin T_{F241L} are shown in (B). $CDK9_{330}$/cyclin $T_{Q77R, E96G, F241L}$ (PDB:3BLH) is shown in cyan, $CDK9_{330}$/cyclin T_{F241L} is shown in green (CDK9) and slate blue (cyclin T). Final $2F_o$-F_c electron density is shown at 1σ around residue Arg65 (CDK9) and Glu96 (cyclin T).

Due to the poorer diffraction of crystals of $CDK9_{330}$/cyclin T_{F241L} compared to crystals of the triple mutant, $CDK9_{330}$/cyclin $T_{Q77R, E96G, F241L}$, and because there is minimal difference between the structure of $CDK9_{330}$/cyclin T_{F241L} and $CDK9_{330}$/cyclin $T_{Q77R, E96G, F241L}$, the triple mutated construct was used for all of the following structural studies.

2.3.3 Crystallisation of CDK9/cyclin T/inhibitor and CDK2/cyclin A/inhibitor complexes

To obtain a basis for developing potent and specific inhibitors of CDK9/cyclin T over other CDK/cyclin complexes, structural analysis of the series of 2-anilino-4-

thiazole-pyrimidine compounds were conducted in both CDK9/cyclin T and CDK2/cyclin A.

CDK2/cyclin A/inhibitor cocrystals grew successfully under standard CDK2/cyclin A/inhibitor crystallisation conditions (Echalier et al. 2008). For CDK9/cyclin T/inhibitor complexes the inhibitors were soaked into apo crystals of $CDK9_{330}$/cyclin $T_{Q77R,\ E96G,\ F241L}$. Diffraction images were obtained and processed and the structures were determined.

2.3.4 Structure determination

Rigid body refinement or molecular replacement was used to produce initial models for subsequent rounds of refinement and rebuilding. Additional restraints from CDK9/cyclin T (PDB:3BLH) and/or a chimera including CDK9/cyclin T/DRB (PDB:3MY1) loops were imposed during refinement of CDK9 structures. The majority of the CDK and cyclin components could successfully be rebuilt. However, some maps produced poor electron density for more flexible regions of the CDK/cyclin structures. In these cases either loops were deleted or, if possible, modelled from other deposited structures.

Clear electron density corresponding to an inhibitor located within the ATP binding site was visible in all of the difference density maps (Figure 2-10). During rebuilding and refinement, ambiguity in the inhibitor binding pose was resolved by refining the major alternative conformations suggested in the electron density, and comparing the refined occupancy and/or b-factors of the

resulting models. All of the inhibitors contained fragments that showed flexibility and the model for the final inhibitor positions and conformations that are described here are inferred from both the crystallographic data, and their capacity to form favourable interactions with the CDK. Crystallographic parameters can be seen in Appendix B.

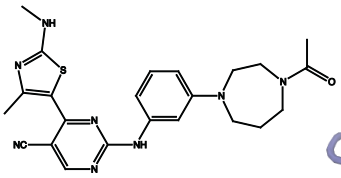
Figure 2-10. Inhibitor interaction modes with CDK2 and CDK9. Chemical structures of inhibitors are shown on the left. Their position within the ATP binding site of CDK2 (centre) and CDK9 (right) are shown. The $2F_o$-F_c electron density is contoured at 1σ. The figure is duplicated to provide both enlarged and more easily comparable images.

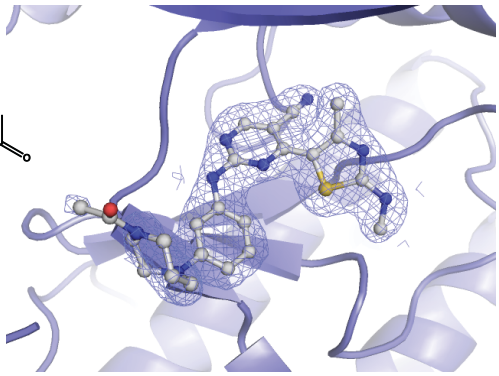
Inhibitor structure

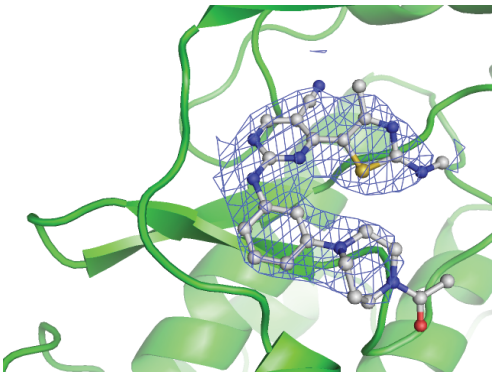
CDK2 ATP binding pocket

CDK9 ATP binding pocket

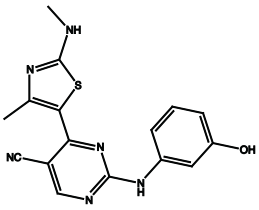
S68

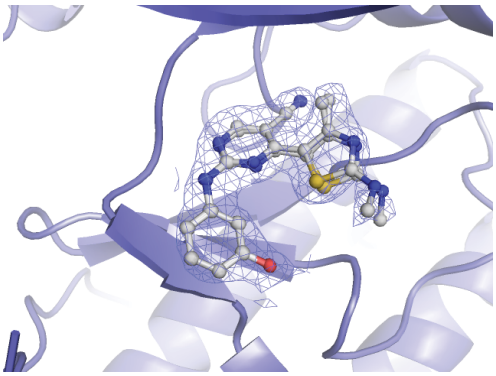


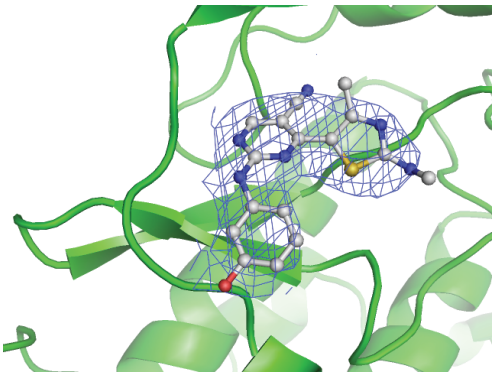




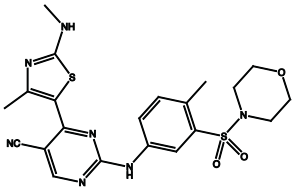
S2-104

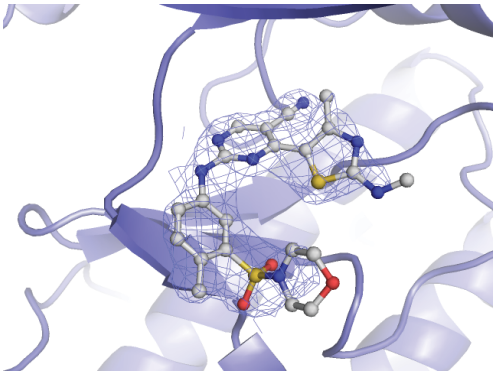


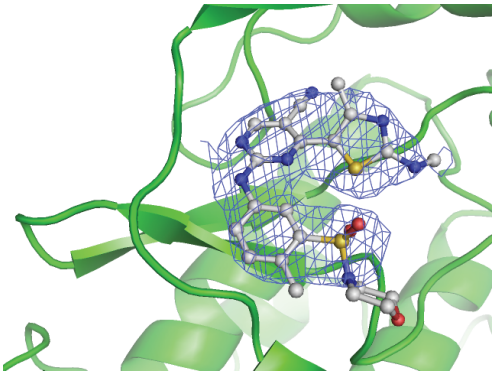




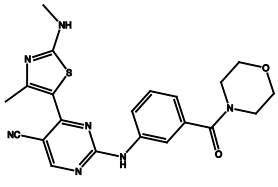
S2-83

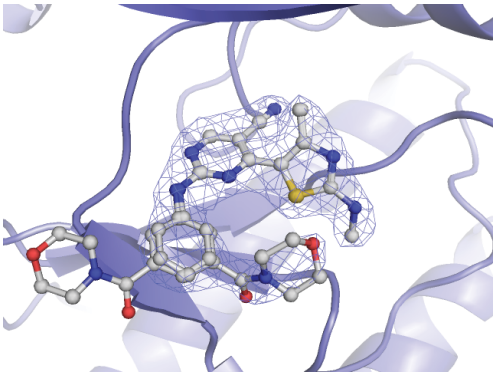




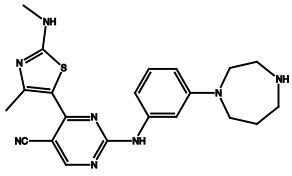
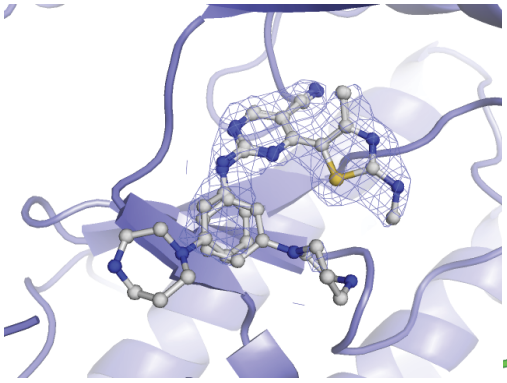
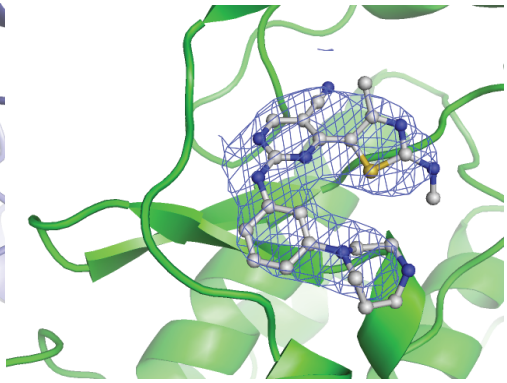
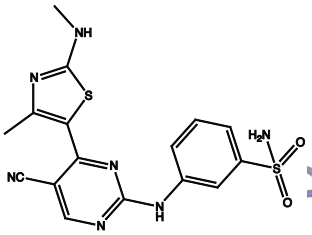
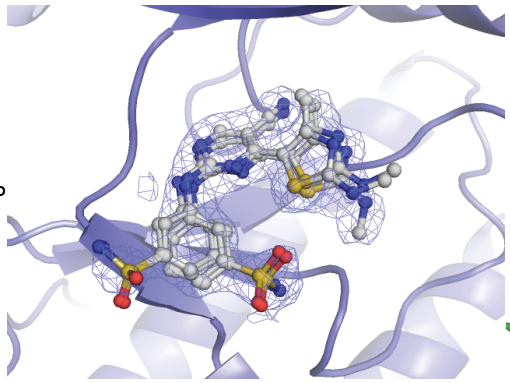
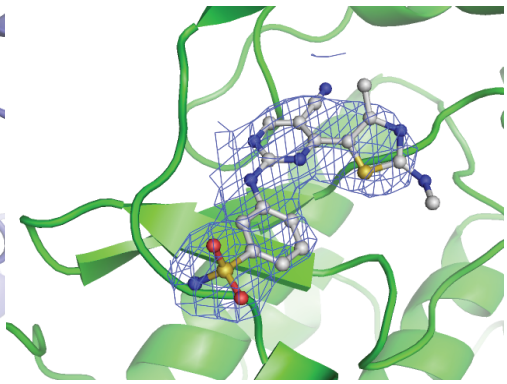
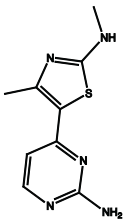
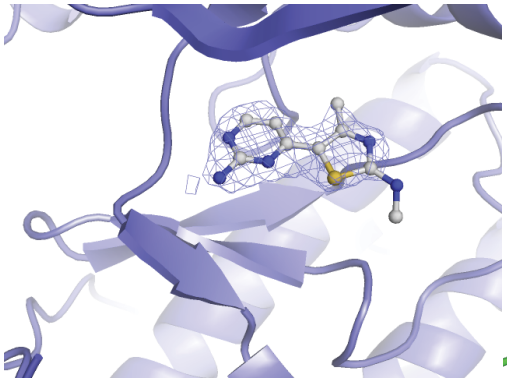
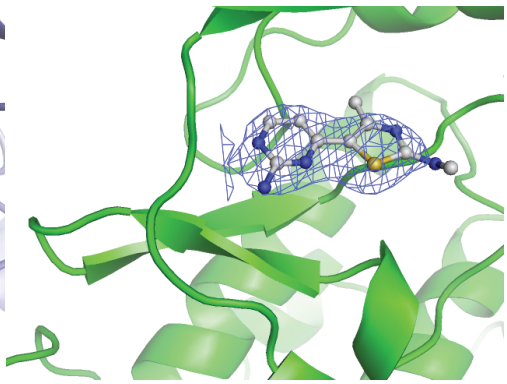
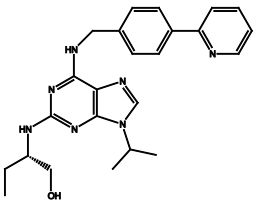
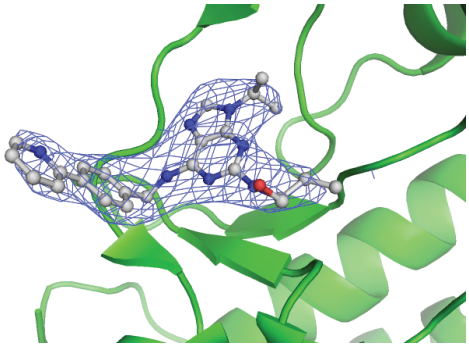


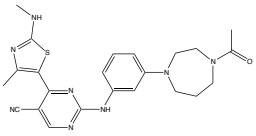
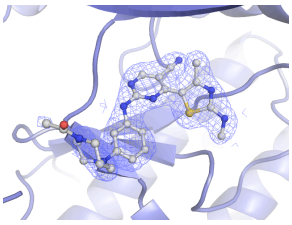
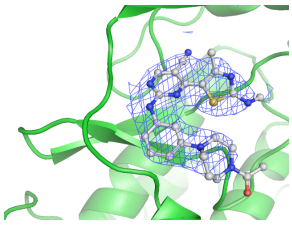
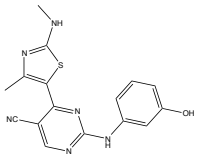
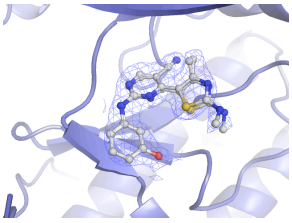
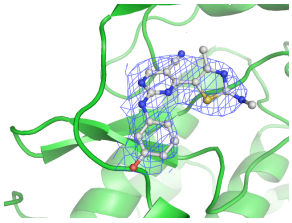
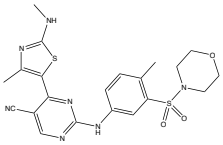
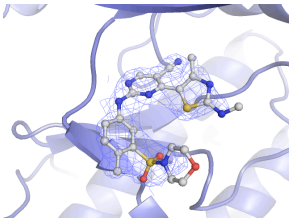
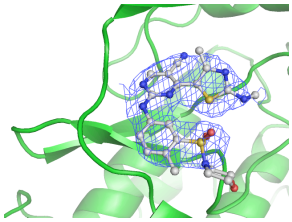
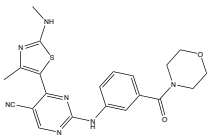
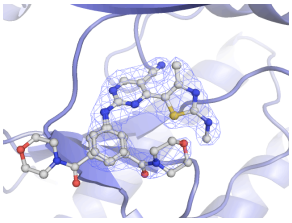
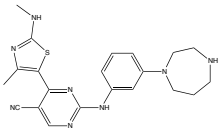
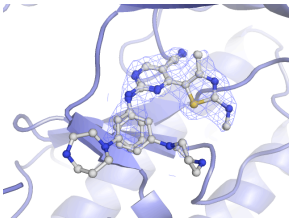
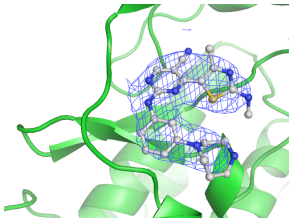
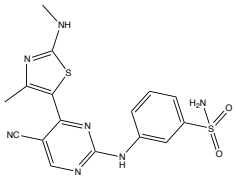
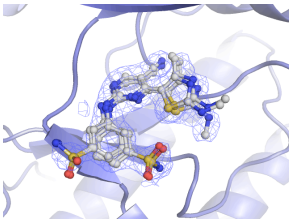
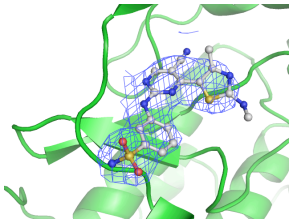
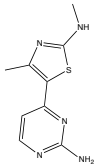
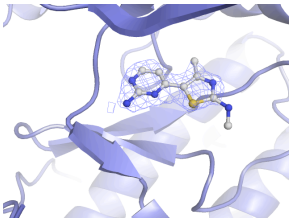
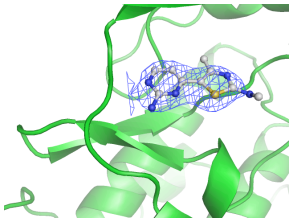
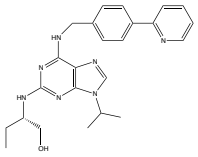
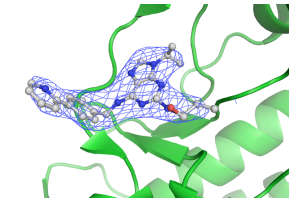
S2-106





Crystals diffracted to low resolution only

Inhibitor structure	CDK2 ATP binding pocket	CDK9 ATP binding pocket
S77 		
S3-35 		
E29 		
SCR8 	R-CR8 structure previously determined	

Inhibitor structure	CDK2 ATP binding pocket	CDK9 ATP binding pocket
S68 		
S2-104 		
S2-83 		
S2-106 		Crystals diffracted to low resolution only
S77 		
S3-35 		
E29 		
SCR8 	R-CR8 structure previously determined	

2.3.5 Binding of 2-anilino-4-thiazole-pyrimidine compounds within the ATP binding site of CDK9/cyclin T and CDK2/cyclin A.

Structures of 2-anilino-4-thiazole-pyrimidine inhibitors have previously been published bound to either the active CDK2/cyclin A complex or the inactive CDK2 protein (Wu et al. 2003b; Wang et al. 2004a; Wang et al. 2004b; Wang et al. 2010). To further understand the structure-activity relationship (SAR) for this class of inhibitors we can compare their binding modes within the ATP binding sites of both CDK9/cyclin T and CDK2/cyclin A.

This class of inhibitors interacts with CDK9 and CDK2 in a similar binding mode. The three main ring groups, the thiazole, pyrimidine and aniline, are located within the ATP binding site in similar positions irrespective of the additional R-groups of the inhibitor. Therefore I will initially describe the interactions that are consistent between the inhibitors, and subsequently the specific interactions and observations that are distinct between the compounds (Chapter 2.3.6). Descriptions of inhibitor interactions with both CDK9 and CDK2 are included for comparison between these two kinases. Parentheses are used to denote CDK2 residues involved in the interactions with the inhibitors.

As predicted, the inhibitors exploit the hinge region of both CDK9 and CDK2 forming two common hydrogen-bonding interactions (Figure 2-11). N1 of the pyrimidine ring accepts a hydrogen bond from the peptide nitrogen of Cys106 (Leu83) while N7 donates a hydrogen bond to the peptide carbonyl of Cys106 (Leu83). The pyrimidine ring is sandwiched between the hydrophobic sidechains of Ala46 (Ala31) and Leu156 (Leu134), with which it forms extensive

van der Waal's interactions. At the back of the ATP binding site the nitrile fits into the pocket close to the gatekeeper residue Phe103 (Phe80), forming a weak, but probably favourable, lone pair-π interaction. The CDK2/cyclin A structures, with higher resolution, show at least one water molecule trapped in a pocket behind the nitrile. These waters are able to form a hydrogen-bond network with the backbone of residues (Asp145) and (Phe146) and with the sidechain of (Glu51). The equivalent residues in CDK9 are conserved and adopt a similar conformation and would therefore support the presence of a comparable hydrogen-bonded network.

A B

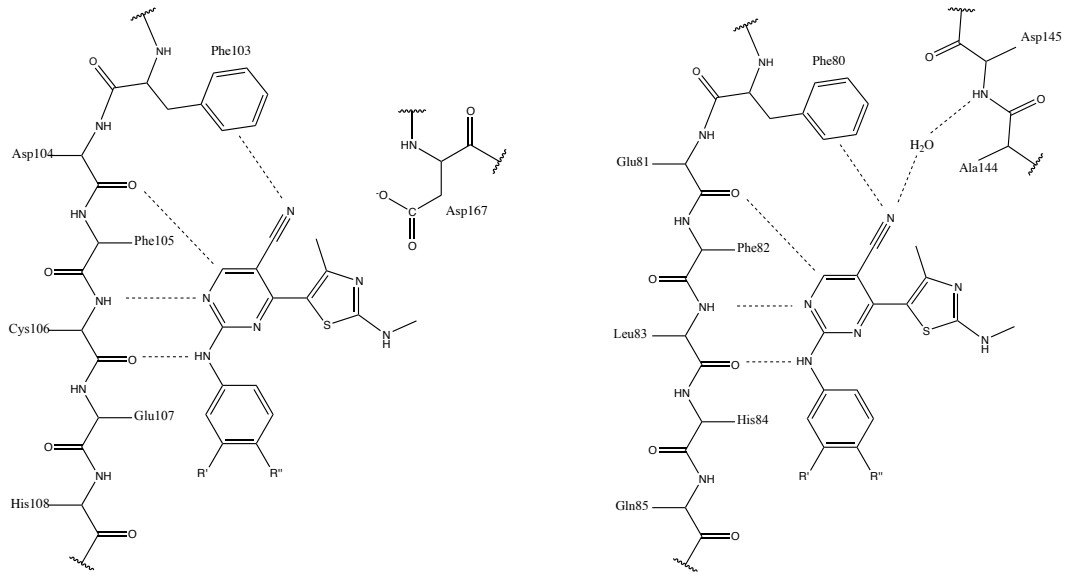


Figure 2-11. Binding mode of inhibitor scaffolds to CDK9 (A) and CDK2 (B) residues. Favourable interactions are shown as a dashed line.

The thiazole has less well-defined electron density in both CDK9 and CDK2 indicating its ability to adopt alternative conformations. In CDK9 the location of the thiazole ring within the ATP-binding site forces Lys48 into a more defined position (in comparison to its unresolved position in the apo CDK9/cyclin T (3BLH) structure). In the major apparent conformation, the amino group

hydrogen bonds with Asp167. In CDK2 the electron density indicates two main conformations for the thiazole moiety. In both, the methyl group is able to interact favourably with the phenyl ring of (Phe80). The first conformation maintains a more planar conjugated ring system with the pyrimidine, whilst in the second conformation the methyl-amino group is able to form favourable interactions with Asp145. Towards the front of the ATP binding pocket, the aniline ring is packed from above by Ile25 (Ile10), with which it forms favourable van der Waal's contacts.

When the C-terminal lobes of the CDK9 structures are superimposed, the pyrimidine rings of the inhibitors are brought into close superposition, consistent with the pyrimidine rings sharing a well-defined disposition with respect to the C-terminal kinase lobe (Figure 2-12). By contrast, the anilino groups show small variations, apparent in particular in the S3-35 and S2-83 complexes. The different behavior of the pyrimido and analino groups is accommodated by a variation in the torsion angles around the two rotatable bonds that connect them. For CDK2 there is a similar trend, with the pyrimidine located in a relatively fixed position with respect to the C-terminal lobe, and the anilino group varying somewhat in position.

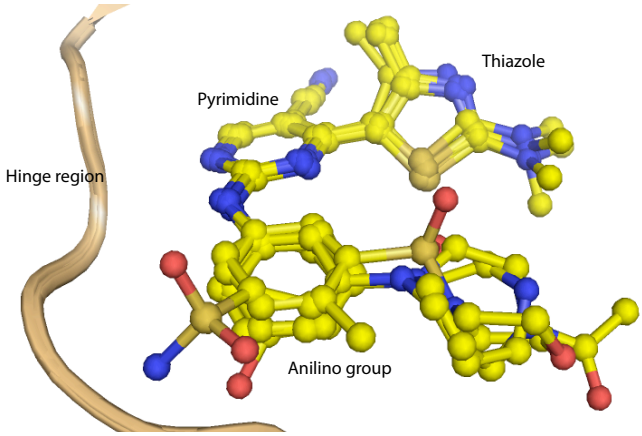


Figure 2-12. Overlay of 2-anilino-4-thiazole-pyrimidine inhibitors within the CDK9 ATP binding site.

2.3.6 Inhibitor-specific interactions in the ATP binding site

Variation of chemical groups within the compounds dissects specific volumes within the ATP binding site. Moieties that interact favourably with CDK9 and unfavourably with CDK2 can be used in future inhibitor design to produce more selective inhibitors against CDK9. Conversely groups that show no apparent preferential interactions for CDK9 over CDK2 are redundant in aiding specificity, although may contribute to the inhibitor potency. Below is a description of the interactions between CDK2 and CDK9 and the alternative R groups of the 2-anilino-4-thiazole-pyrimidine inhibitors.

1) S2-106

Within the CDK2/cyclin A crystallographic lattice there are two CDK/cyclin dimers, and hence also two inhibitors. The crystal structure of CDK2/cyclin A/S2-106 shows variation between the inhibitors bound illustrating two alternate binding modes. In both cases the electron density of the morpholine ring is weak, reflecting the flexibility of this ring. In one

conformation the electron density supports an “inward” conformation of binding and in the other the “outward” conformation (Figure 2-13). This ring in the “inward” conformation can pack against Ile10 of the glycine rich loop, whilst in the “outward” conformation the morpholine ring exits into the solvent channel with minimal contact with the protein. Incidentally Lys89 adopts different conformations in these two inhibitor-binding modes, in the “outward” conformation Lys89 approaches the inhibitor, and in the “inward” conformation Lys89 moves away. This inhibitor was not successfully crystallised with CDK9/cyclin T.

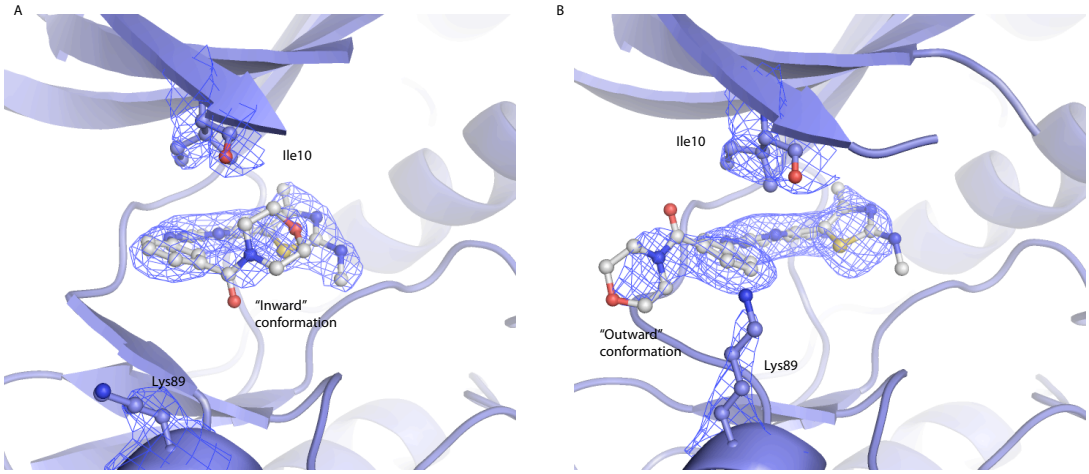


Figure 2-13. S2-106 bound to CDK2/cyclin T in the “inward” (A) and “outward” (B) conformations. In addition to the inhibitor, residues 10 and 89 of CDK2 are shown in a ball and stick representation. $2F_o$-F_c electron density is shown at 1σ.

2) S3-35

The sulphonamide group is oriented differently in CDK9 and CDK2. The electron density shows that the group adopts an “outward” conformation when bound to CDK9 interacting with carboxyl backbone of residue Glu-107 (Figure 2-14). This conformation causes the backbone of the hinge region to move (in comparison to

the apo structure of CDK9/cyclin T (PDB: 3BLH)) and twists the relative orientation of the pyrimidine and anilino rings out of the favourable planar conjugated system. In CDK2 the sulphonamide primarily adopts an "inward" conformation within the ATP binding site, forming an interaction with Asp86 and maintaining a more planar ring system. This inhibitor structure most clearly shows the possibility of at least two orientations for the thiazole group within CDK2, highlighting the flexibility in its binding modes.

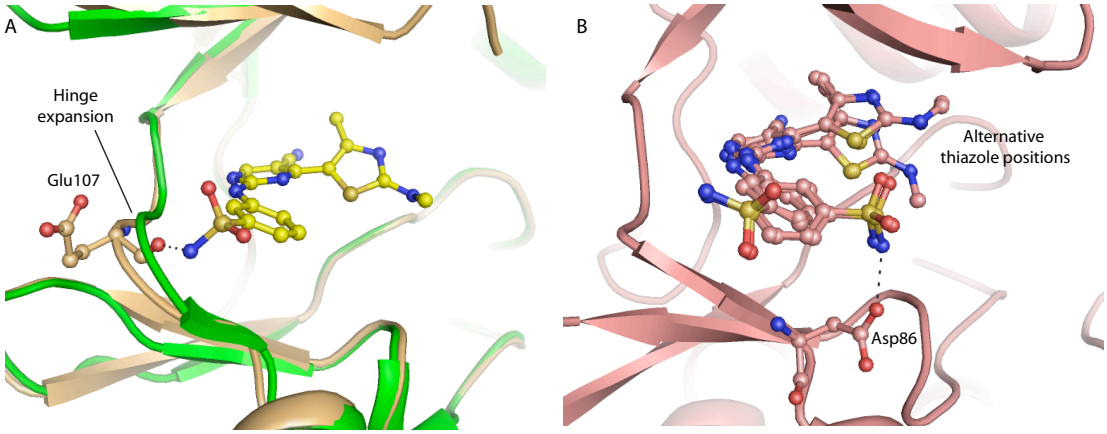


Figure 2-14. S3-35 within the ATP binding sites of CDK9 (A) and CDK2 (B). CDK9/cyclin T in the absence of inhibitor (green), and in the presence (sandy) are shown. S3-35 is shown within the CDK2 ATP binding site (pink) indicating the alternative thiazole positions that this group could accommodate according to the electron density. Although the "inward" conformation is favoured with stronger electron density, all 4 conformations are shown as there is additional weak electron density in the "outward" confromation.

3) S2-104

In CDK2 the hydroxyl moiety of S2-104 adopts an "inward" position forming a hydrogen bond with Asp86. Conversely in CDK9 the hydroxyl group has preference for an "outward" conformation with the potential to form a long-range hydrogen bond with the carbonyl of Glu107. In a similar manner to CDK2, reorientation of the hydroxy-anilino group to adopt an "inward" conformation would enable a hydrogen bond with the corresponding Asp 109.

4) S2-83

In CDK9 there is clear electron density for an inward conformation of S2-83. However the morpholine group is poorly defined and appears to adopt several conformations within the ATP binding site.

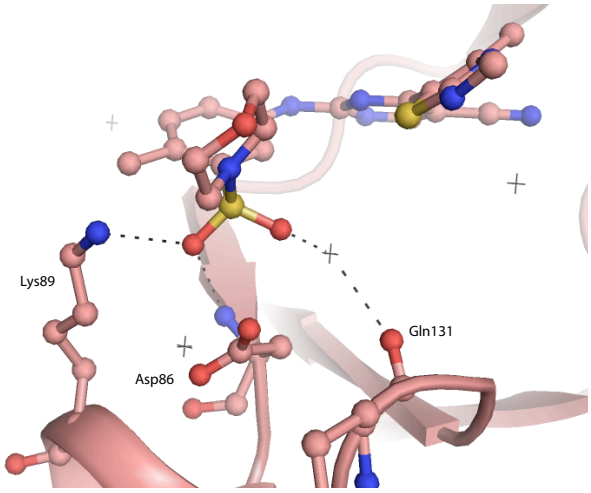


Figure 2-15. S2-83 within the ATP binding site of CDK2. Water molecules are displayed as crosses and polar interactions by dashed lines.

S2-83 also binds in an "inward" conformation to CDK2 (Figure 2-15). The sulphonyl group forms favourable long range interactions to the sidechains of Lys89 and peptide NH of Asp86, but also is close to the unfavourable Asp86 sidechain. A water molecule is found between the carbonyl of Gln131 and the sulphonyl. The morpholine ring packs up against Ile10 and Gly11 of the glycine-rich loop. It is represented by weaker electron density representing the flexibility of this saturated ring.

5) S-77

The electron density of the saturated 1,4-diazepine ring attached to the meta position of the benzene, demonstrates its flexibility when bound to CDK2. No electron density corresponding to this ring is present at a level of 1σ. Two

conformations of S-77 have been built into the structure as possible alternative binding modes. Neither of these modes suggests either favourable or unfavourable protein-inhibitor interactions, correlating to the absence of electron density for this 7-membered ring. The inhibitor bound within the second CDK2/cyclin A dimer in the asymmetric unit shows some preference for "outward" binding. However, interpretation of this should be limited due to the proximity of a crystallographic symmetrical cyclin molecule.

In CDK9 the 1,4-diazepine ring clearly adopts the "inward" position which may be due to reduced solvent exposure of the hydrophobic ring in this conformation.

6) S-68

The addition of an acetyl group to the saturated 1,4-diazepine ring forming an amide, favours an "outward" conformation in CDK2. Continuous electron density is present from the benzene ring away from the centre of the ATP-binding site of CDK2. There is some unaccounted for electron density in the alternative conformation which is likely to reflect some flexibility within these groups. The structural explanation for this preferential orientation is unclear. A combination of a favourable interaction between Lys20 and the amide in the outward conformation and steric hindrance in the inward conformation may contribute. In CDK9 S-68 binds in an "inward" conformation in a similar manner to S-77. Again there appear to be no striking structural reason for this orientation.

A summary of the binding mode and interactions made by these inhibitors to CDK9/cyclin T or CDK2/cyclin A is shown in Table 2-4.

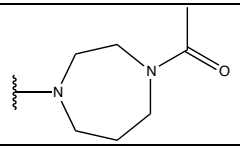
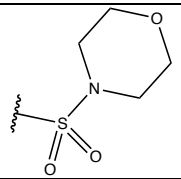
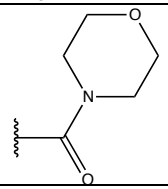
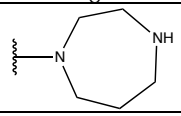
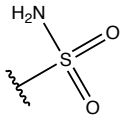
			CDK9				CDK2			
Inhibitor	R' group (meta-)	R'' group (para-)	K_i (μM)	Conformation "Inward"	Conformation "Outward"	Specific interactions	K_i (μM)	Conformation "Inward"	Conformation "Outward"	Specific interactions
S68			0.007	✓	x	None	0.139	x	✓	Glu8
S2-104			0.011	x	✓	Glu107	0.012	✓	x	Asp86
S2-83			0.022	✓	x	None	0.123	✓	x	Lys89 Asp86
S2-106			0.043	No data			0.147	✓	✓	Asp86
S77			0.007	✓	x	None	0.568	✓	✓	None
S3-35			0.006	x	✓	Glu107	0.004	✓	✓	His84 Asp86

Table 2-4. Binding modes of the 2-amino-4-heteroaryl-pyrimidine class of inhibitors within CDK9 and CDK2. The general chemical structure of these inhibitors is shown in Figure 2-5.

2.3.7 The structure of a CDK9/cyclin T/(S)-CR8 complex

To compare the binding modes of CR8 inhibitors in CDK9/cyclin T and in CDK2/cyclin A the structure of a CDK9/cyclin T/(S)-CR8 complex was solved. (S)-CR8 is bound within the ATP-binding site of the kinase (Figure 2-16).

(S)-CR8 adopts a similar binding mode within the CDK9 ATP binding site to that of (R)-CR8 bound to CDK2 (PDB ID: 3DDP, (Bettayeb et al. 2008)). N7 of the purine ring and the 6-amino group form hydrogen bonds to the peptide bond nitrogen and the carbonyl respectively of Cys106 of the hinge region. The ATP N1 interaction with the hinge backbone is mimicked by N7 of (S)-CR8 but the interaction made by the 6-amino group of the inhibitor has no equivalent in the ATP-bound structure. Residues Ile25 and Leu156 sandwich the (S)-CR8 purine ring from above and below respectively. The phenyl-pyridine ring of the inhibitor is located outside of the ATP binding site where it can form favourable interactions with the side chains of Ile25, Phe105 and Ala23 that are reminiscent of the interactions between (R)-CR8 and CDK2 within the CDK2/cyclin A/(R)-CR8 complex structure (Bettayeb et al. 2008). The weaker electron density for the (S)-CR8 phenyl-pyridine moiety, however, suggests that there is some flexibility in the phenyl-pyridine ring location. Similarly, the electron densities for the hydroxy-methyl and ethyl groups are weaker indicating additional conformational variability of these groups within the crystal.

A B

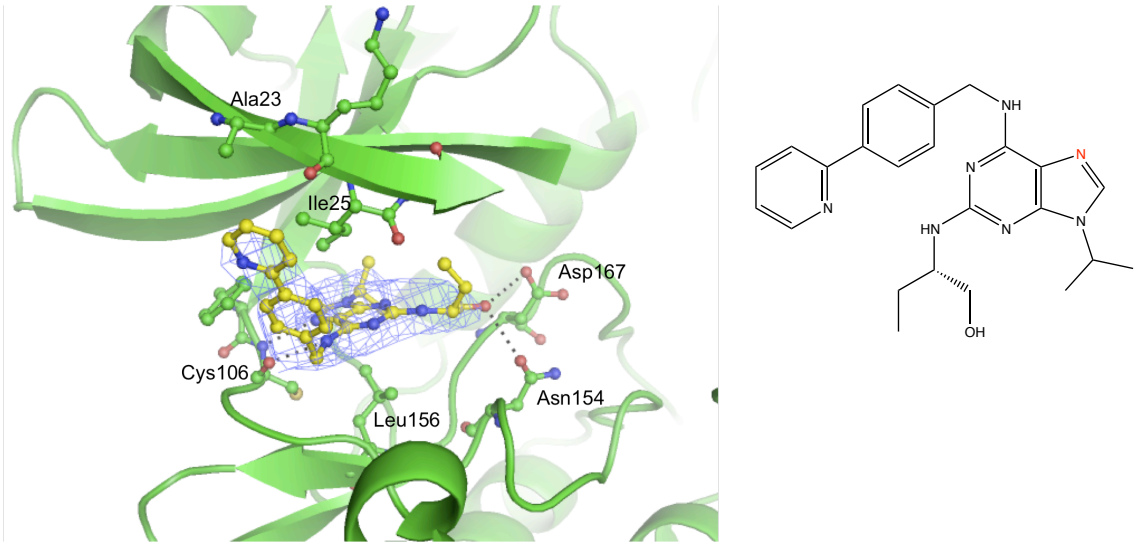


Figure 2-16. Close view of (S)-CR8 within the ATP binding pocket of CDK9/cyclin T. Hydrogen bonds are illustrated with dashed lines and the final $2F_0$-F_c map contoured at 1σ is displayed as a blue mesh. (B) The chemical structure of (S)-CR8. N7 is displayed in red. Reprinted from SAGE Publications Ltd. Genes & Cancer, 1 (4), Bettayeb et al., 2010, copyright 2010, http://online.sagepub.com.

2.3.8 Tertiary structural changes on inhibitor binding

Previously published descriptions of inhibitors binding to CDK9/cyclin T identify apparently unique tertiary structural changes of CDK9 on inhibitor binding (Baumli et al. 2008; Baumli et al. 2010). Confirming this observation, all inhibitors studied here induce structural changes in CDK9 (Figure 2-17). By contrast, the majority of CDK2/cyclin A inhibitor structures show no significant structural reorientations in the presence of an inhibitor.

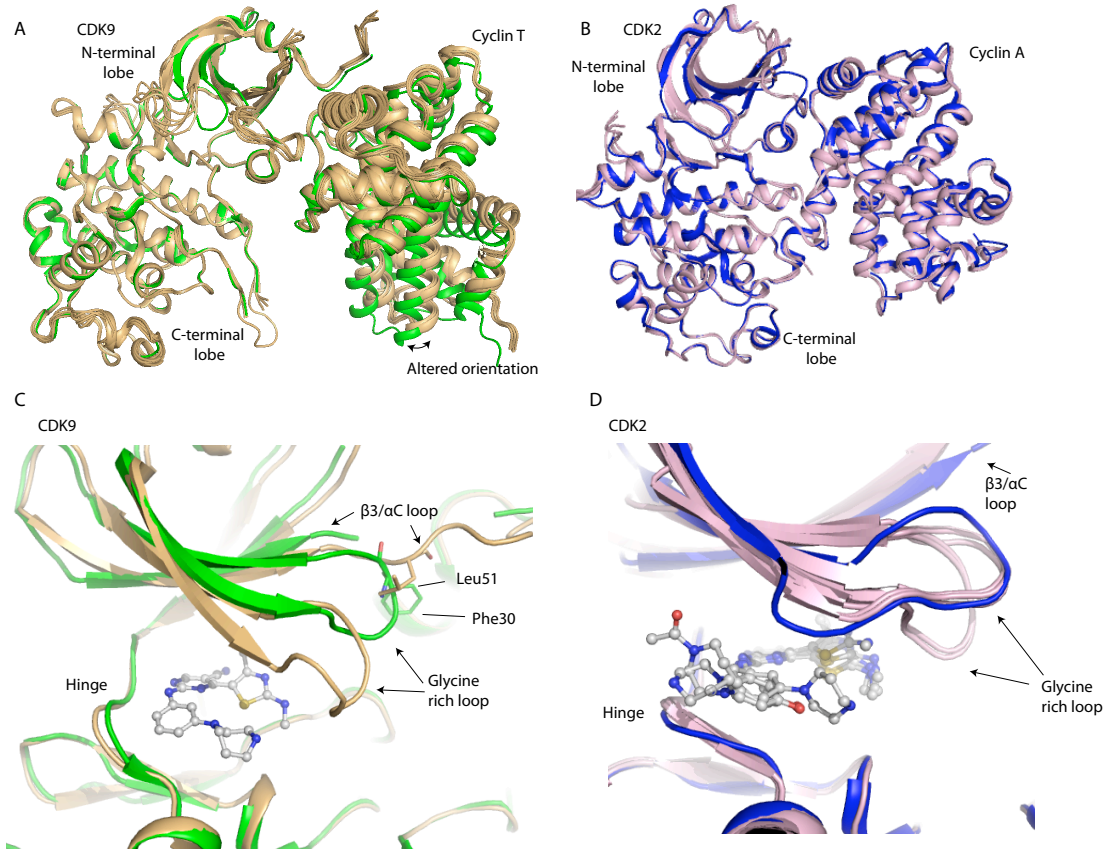


Figure 2-17. Inhibitor-induced conformational changes. (A) Comparison of CDK9/cyclin T with (orange) and without (green) inhibitor. (B) Comparison of CDK2, with (pink) and without (blue) inhibitor. Movements of the glycine rich loop, the β3/αC loop and the hinge region (C) & (D). Kinase complexes were superimposed on their kinase C-terminal lobes (residues 101-286 (CDK2) and 123-316 (CDK9)).

There are three main differences observed between the apo and inhibitor bound states of CDK9. These, and the corresponding regions in CDK2, are described below.

i) The glycine rich loop

As observed with all previous CDK9/cyclin T/inhibitor structures deposited to date, the glycine rich loop folds down over the inhibitors. Weak electron density and high B-factors indicate that it occupies more than one such closed conformation in the crystal. Concurrently the β3/αC loop lowers partially to occupy the site of the apo glycine loop, with Leu51 structurally replacing Phe30.

The glycine-rich loop remains in a more open conformation in the majority of CDK2/cyclin A/inhibitor structures presented here. Although in the presence of S77 and S68 the loop lowers and contacts Asp127 in the C-terminal lobe.

ii) Relative orientation of N- and C- terminal lobes

Superposition of the C-terminal CDK9 domains reveals that the N- and C-terminal lobes of CDK9 have different orientations with respect to each other in the presence and absence of inhibitor. This is particularly apparent on inspection of the cyclin subunit. This rotation is also seen in the CDK9/cyclin T/HIV-Tat (PDB: 3MIA) structure and in other CDK9/cyclin T/inhibitor structures (Baumli et al. 2008; Tahirov et al. 2010).

Conversely in all but two CDK2/cyclin A/inhibitor structures there is no significant rotation of cyclin A with respect to CDK2: the cyclins remain in the same orientation with respect to CDK2 as is seen in the apo structure. CDK2/cyclin A in the presence of either S2-106 or S3-35 does, however, induce conformational movement of both the C-helix and the cyclin with respect to the C-terminus of CDK2 (Figure 2-18). This conformation of CDK2/cyclin A appears unique to the complexes with S2-106 or S3-35 – it is not seen in 50 other CDK2/cyclin A/inhibitor structures available in the PDB. This new conformation has the C-helix in neither the active “in”, partially active or inactive “out” conformation and is further characterised by a displacement of many of the cyclin helices. It is not clear from the structure what property or interactions of the inhibitor has induced this change in conformation.

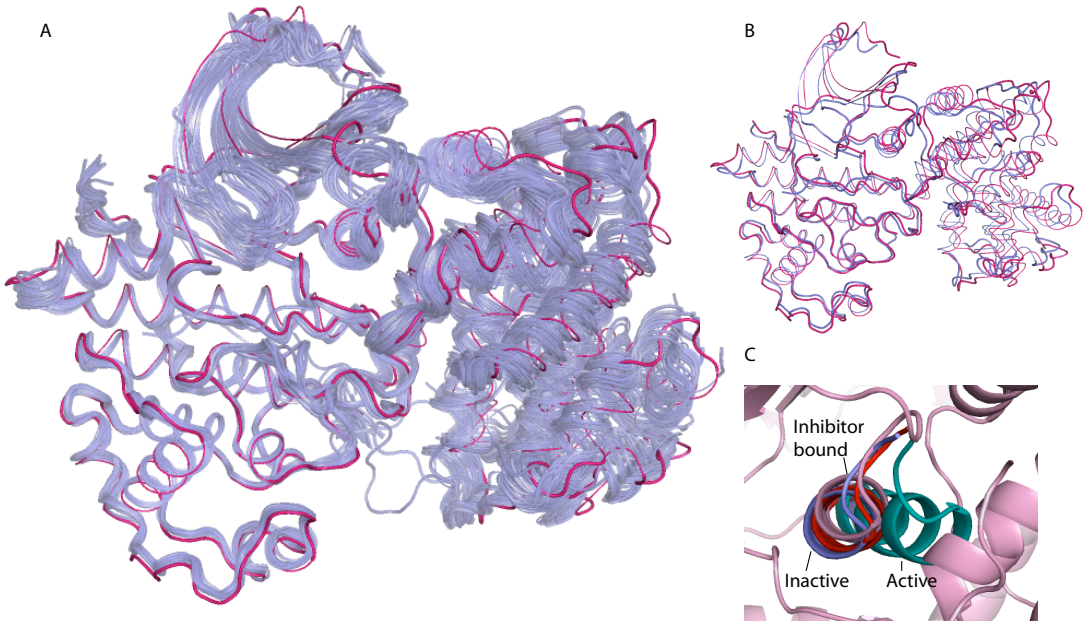


Figure 2-18. CDK2/cyclin A structures superposed on the C-terminal domain of the CDK. (A) A comparison of many deposited PDB structures of CDK2/cyclin (blue structures) with CDK2/cyclin A/S2-106 (pink). (B) Sturcture of CDK2/cyclin A with a peptide bound (PDB: 1QMZ) (blue), in comparison with CDK2/cyclin A/S2-106. (C) Overlay of C-helices located in various positions due to the activation state of CDK2. CDK2 in a partially active unphosphorylated state (PDB: 1FIN (red)), fully active phosphorylated state (PDB: 1QMZ (blue)) and inactive CDK2 (PDB:1HCK (teal)) and S2-106-bound CDK2/cyclin A (pink).

iii) Hinge expansion

Residues 106-107 of the CDK9 hinge region show some flexibility (the Cα of residue 106 has an altered position of *circa* 1Å) to move away from the centre of the ATP-binding site. These movements may reduce steric clashes between the protein and inhibitor and permit the formation of hydrogen bonds with more favourable geometry.

The conformation of the hinge region of CDK2 varies less in the different inhibitor complexes than does that of CDK9. The extended hinge region, residues 107-109, constricts the lower region of the ATP binding site by enclosing in on the ATP binding site in comparison with CDK9. This may be a result of several

factors. The first and most striking variation between CDK2 (and CDK1) and CDK9 in this region is the non-conserved residues, Ala111 (Lys88 in CDK2) and Gly112 (Lys89 in CDK2) (see Figure 1-7). These sequence differences change both the volume available and the electrostatic environment at the front of the ATP binding pocket. Furthermore, although (Lys89) can be flexible, it often forms a favourable hydrogen bond with (Asp86) resulting in it lying close to the ATP binding pocket.

Secondly there are two additional non-conserved residues, Glu107 (His84 in CDK2) and His108 (Gln85 in CDK2), in the hinge region that may prevent lowering of the extended hinge backbone to accommodate the inhibitors.

2.4 Discussion

2.4.1 Suitable crystallisation constructs

To date all the published crystal structures of CDK9/cyclin T/inhibitor complexes have used a CDK9 construct that is truncated at its C-terminus, and a cyclin T construct that contains 3 mutations (Baumli et al. 2008; Baumli et al. 2010; Bettayeb et al. 2010). To confirm that this is a suitable complex for use in inhibitor binding studies, I have used crystallography and differential scanning fluorimetry to assess the consequences of these mutations on the structure and inhibitor binding properties of CDK9/cyclin T.

Firstly, I have shown that the C-terminal tail of CDK9 has no effect on the thermal stability provided to the CDK9/cyclin T complex when measured in the presence of (S)-CR8 or a series of 2-anilino-4-thiazolepyrimidine compounds. This finding presents a contrast to the positive effect that the C-terminal tail has on the binding of the inhibitor DRB (Baumli et al. 2010). Taken together, these results suggest that, unlike the case with DRB, the C-terminal tail of CDK9 does not stabilise additional contacts between CDK9 and the inhibitors described here.

Secondly, I have shown that the crystal structure of $CDK9_{330}$/cyclin T_{F241L} and $CDK9_{330}$/cyclin $T_{Q77R,\ E96G,\ F241L}$ are very similar. The main differences are in the glycine-rich loop and the β3-αC loop, both of which are inherently flexible regions. The mutation Glu96→Gly does, however, prevent the formation of hydrogen bonds between cyclin T and Arg65, from the C-helix of CDK9. Analysis by thermal denaturation showed that the E96→Gly mutation significantly

destabilises the protein but has little effect on the profile of further-stabilisation induced by the binding of a range of inhibitors (Analysed by Dr. S. Baumli).

$CDK9_{330}$/cyclin $T_{Q77R, E96G, F241L}$ reproducibly crystallises to higher resolution than other complexes tested. The demonstration that the inhibitor binding properties of this complex closely resemble those of $CDK9_{FL}$/cyclin T_{WT} validates its use in crystallographic inhibitor binding studies.

2.4.2 CDK9/cyclin T/HIV-Tat inferences

The structure of CDK9/cyclin T/HIV-Tat/AMPPNP (PDB: 3MIA) was a breakthrough in the field of transcription. The authors compared the structure to that of CDK9/cyclin T/AMPPNP (PDB: 3BLQ) and suggested that the presence of Tat confers several structural changes on the kinase (Tahirov et al. 2010), including the formation of additional hydrogen bonds between Arg65 (on the C-helix) and phosphothreonine 186. They speculated that this interaction might cause the further conformational changes that were seen in the glycine rich loop and the β3-αC loop.

The structure of CDK9/cyclin T presented here shows that Glu96 (cyclin T) reorientates Arg65 (CDK9) through the formation of hydrogen bonds. The weak electron density of this residue implies that any interaction that Arg65 makes with phosphothreonine 186 is transient in nature. The interaction does not appear to be as stable as it is in the crystal complex in the presence of Tat. This

finding suggests that Tat may further stabilise the interaction between Arg65 and phosphothreonine 186.

Both the glycine rich loop and the β3-αC loop in the structure of CDK9/cyclin T presented here adopt distinct conformations to those seen in either 3BLQ or 3MIA. It is possible, therefore, that variation seen in the structures of these loops derives from their inherent flexibility, rather than being a direct consequence of Tat binding.

2.4.3 CDK9 movements on inhibitor binding

In all of the structures of inhibitors bound to CDK9/cyclin T published to date and presented here, there are several features that distinguish CDK9 from other CDKs. Unlike other CDKs, CDK9 always undergoes large conformational changes upon inhibitor binding. These movements result in both a re-orientation of its N- and C-terminal lobes and a downward movement of the glycine rich loop. The glycine rich loop is further stabilised over the ATP-binding site by the lowered β3-αC loop (Figure 2-19). In several of the structures described here Phe30 enters the ATP binding site in a similar way to that seen in the CDK9/cyclin T/flavopiridol complex. Overall, these results suggest that CDK9 may be more structurally malleable than other CDKs that have been structurally characterised. The ability of CDK9 to adopt additional conformations in comparison with other CDKs may contribute to the selectivity of flavopiridol and other CDK9-selective inhibitors.

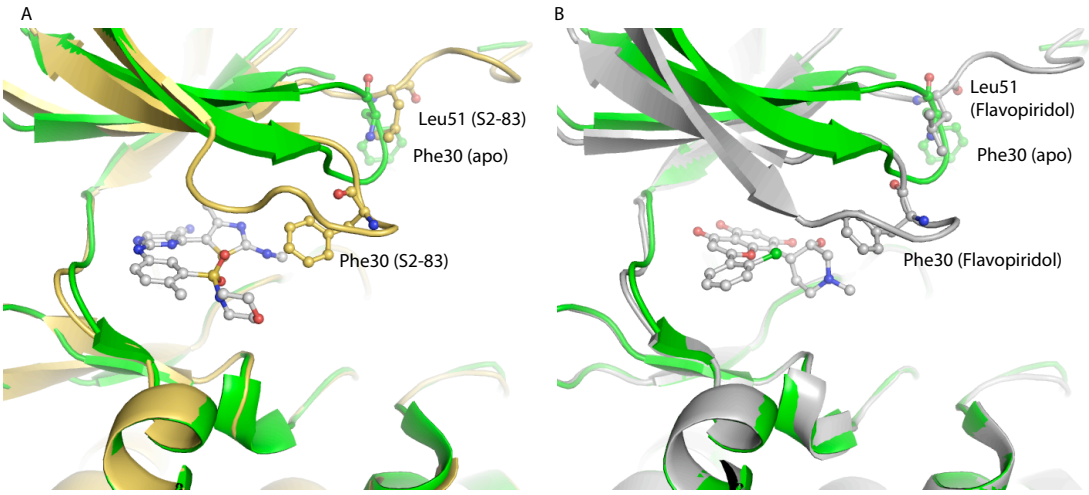


Figure 2-19. Lowering of the glycine rich loop and Leu51 (of the β3-αC loop) of CDK9 in the presence of inhibitors. Phe30 of the glycine rich loop of CDK9 lowers in the presence of S2-83 (A) and flavopiridol (B) in comparison to in the absence of an inhibitor. CDK9/cyclin T is illustrated in cartoon representation in the apo state (green), in the presence of S2-83 (pale orange) and in the presence of flavopiridol (grey). Residues 30 and 51 of CDK9 and the inhibitors are represented as balls and sticks.

2.4.4 Novel CDK2/cyclin A conformation

In the presence of S2-106 or S3-35, CDK2/cyclin A adopts a conformation that appears to lie outside the normal distribution of conformations seen in other CDK2/cyclin A complexes. This conformation has altered locations for the N-terminal CDK β-sheets and cyclin α-helices. The C-helix is relocated to a position different from that of the active (PDB: 1QMZ), partially active (PDB: 1FIN) and inactive conformations (PDB: 1HCK). This novel conformation is likely to represent a state transiently sampled by CDK2. Bao et al (2011) suggested that the variation in structures deposited to the PDB of CDK2 and their molecular dynamic simulations reflect the flexibility of CDK2 and the range of space it samples (Bao et al. 2011). By adopting distinct, less active conformations prior to activation and substrate binding, CDK2 reduces the frequency of the transition state being occupied and therefore reduces futile reactions of ATP hydrolysis until all substrates are correctly aligned and the CDK becomes more stable (Bao

et al. 2011). Therefore this state is likely to be an additional inactive state sampled to prevent futile ATP hydrolysis.

This state has not been seen in the presence of any other inhibitors bound to CDK2 and is therefore likely to be associated with an energetic penalty that is compensated by the interactions that the inhibitor forms with the restructured protein. The explanation for why these specific inhibitors produce this altered state of CDK2/cyclin A is, however, unclear from structural analysis.

2.4.5 Selectivity of inhibitors between CDK2 and CDK9

Suitable inhibitors for clinical trials are both potent, to compete with high cellular ATP concentrations, and selective, to minimise side effects of targeting other cellular proteins (Johnson 2009a). To date, with the exception arguably of flavopiridol, there are no inhibitors of CDK9 available with both of these properties.

CDKs have highly conserved ATP binding sites, making it difficult to synthesise inhibitors specific to a single CDK (see Figure 1-7). Structural biology can help to overcome these challenges by defining the interactions that stabilise inhibitor binding to target and off-target proteins. Using structures to target selectivity between a target protein and a closely related off-target protein is a reasonable way to try to enhance potency and selectivity in the course of a drug discovery program. I have described an extensive body of work, comparing the structural and biophysical characteristics of the binding of a series of 2-anilino-4-thiazole-

pyrimidine compounds to CDK2 and CDK9. CDK2 was chosen as the off-target kinase in this study due to both its experimental tractability and its close sequence homology to CDK9.

As described above, both kinases interact with this class of inhibitors in much the same way: the pyrimidine and anilino groups pack against the hinge region for optimal hydrogen bonding, whilst the thiazole group shows substantial flexibility in its location. There are, however, some significant differences in how the inhibitors bind to CDK9 or CDK2, and these differences may aid in developing selectivity further.

CDK9 shows a greater level of induced fit than does CDK2 when binding to this class of inhibitors. In particular, larger less flexible groups appear to be accommodated better within the ATP binding site of CDK9 than in CDK2. The extended hinge region of CDK9 enables it to accept larger inhibitors in the binding region. The relatively small amino acids Ala111 and Gly112 of CDK9, which are equivalent to residues Lys88 and Lys89 of CDK2, generate a more open active site in CDK9, with additional space available to accommodate larger inhibitors.

Of interest is the different preference of orientation of the inhibitor R groups when bound to the different kinases. In CDK9/cyclin T the larger groups are orientated in an “inward” position, whilst in CDK2/cyclin A there is often little preference for either conformation. The smaller R groups on the other hand have a preference for an “outward” conformation in CDK9 and an “inward” or no

preference for CDK2. Further stabilization of the different binding modes adopted by this scaffold when in complex with CDK2 or CDK9 would be an approach to enhance the selectivity of their inhibition.

Using all of the available CDK9/cyclin T/inhibitor structures a general overview of the ATP-binding site space utilized by these inhibitors can be obtained. The A, R and K pockets are all exploited, but the E_0, E_1 and P pockets less so. Figure 2-18 shows that the P pocket is varied in size, due to the alternative conformations that the glycine-rich loop adopts, and is therefore more difficult to target. Although some of the inhibitors are located in the E pockets these could potentially be exploited further for improved inhibitor characteristics.

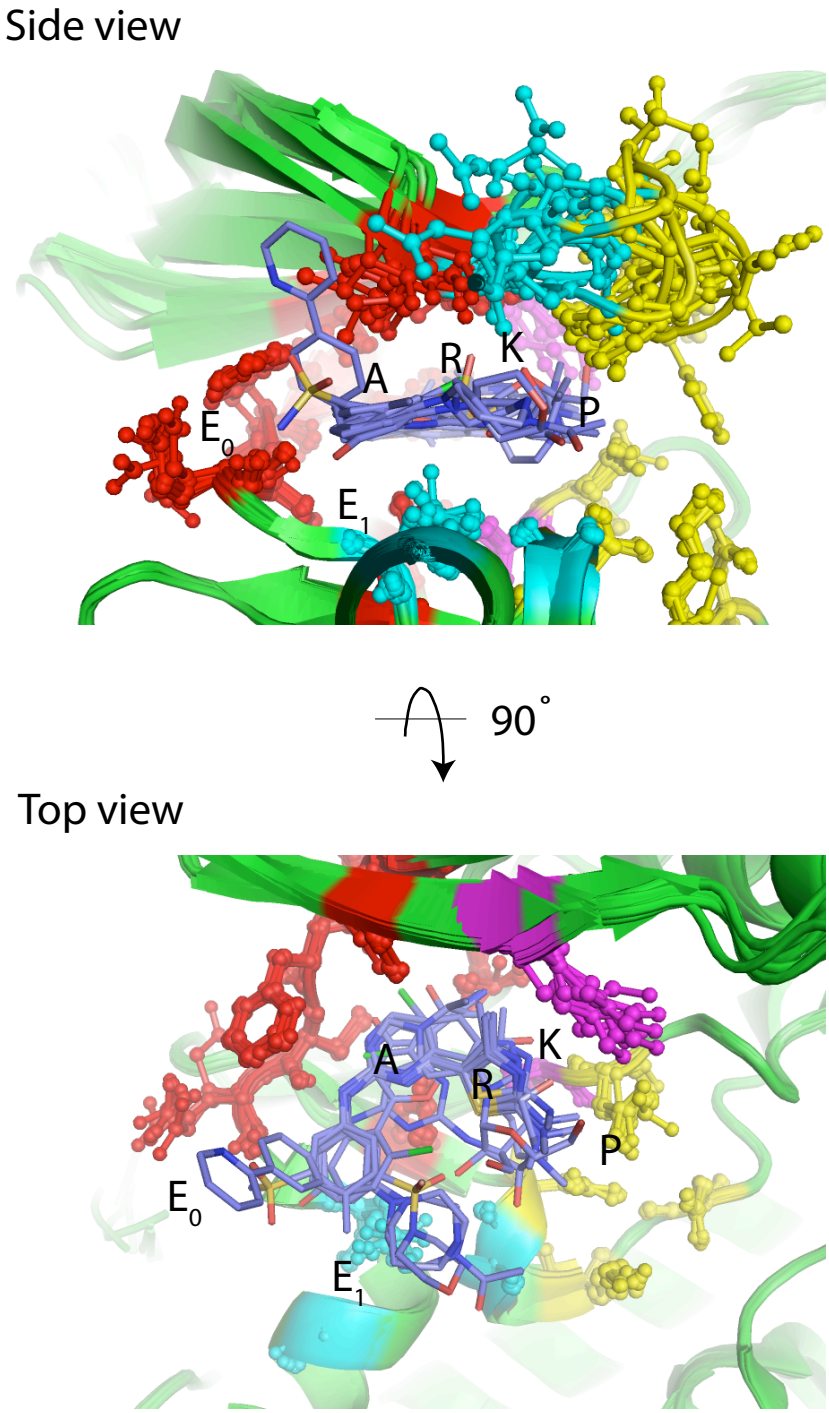


Figure 2-20. Overlay of inhibitors bound to the ATP binding pocket of CDK9. Residues lining the A pocket (red), R pocket (cyan), P pocket (yellow), and K pocket (magenta) are shown as sticks.

2.4.6 Suggested inhibitor improvements

Designing inhibitors to selectively inhibit specific kinases is a challenging aim. The work above leads to several suggestions for structure-based inhibitor improvements. The K_i values of this range of inhibitors imply that they are sufficiently potent against CDK9, but it may be that they could be improved by obtaining a greater degree of selectivity.

Several related inhibitors to the 2-anilino-4-thiazole-pyrimidine compounds presented above have been studied previously. A series of inhibitors based on 2,4-bisanilino pyrimidine and 4,6-bisanilino pyrimidine scaffolds have been studied in the context of CDK2 and CDK4 inhibition (Anderson et al. 2003; Beattie et al. 2003; Breault et al. 2003). And similarly, 2-anilino-4-thiazole-pyrimidine compounds have previously been crystallised using inactive monomeric CDK2, and have been tested against a range of CDKs (Wu et al. 2003b; Wang et al. 2004a). Using the published K_i values and the data described here several suggestions can be put forward to improve the inhibitors' selectivity without compromising their potency for CDK9.

Firstly, it is clear from the data that the 2-(phenylamino)pyrimidine-5-carbonitrile scaffold provides much of the potency of these inhibitors in both CDK2 and CDK9 and should therefore not be altered. The various R groups (see Table 2-4) often show quite distinct conformations whether bound to CDK2 or CDK9. This provides an opportunity to exploit the differences between these two ATP-binding sites. Large flexible rings on the anilino ring are accommodated by

CDK9 but not in CDK2, whilst smaller moieties provide little specificity between the two kinases.

Introducing a combination of R groups may provide specificity between CDK9 and other CDKs. For example, introducing a combination of S2-83 and S2-104, by producing a compound with R groups, 3-SO_2morpholine, 4-methyl, 5-SO_2NH_2, may provide more optimal bonds towards CDK9, but less favourable for CDK2. Another possibility would be to test the effect of introducing two rings to the anilino group as we have shown CDK9 to have more flexibility to accept such groups within its ATP-binding site than CDK2 . A final suggestion for these groups is to reduce the polarity of the substituent rings as they are often exposed more to solvent in CDK2 than in CDK9.

The thiazole group of this class of inhibitors appears flexible within the ATP binding sites. Introducing a bridge between itself and the pyrimidine to prevent rotation increases potency towards CDK1 and CDK7 but potencies towards CDKs 2, 4 and 9 remain similar (McIntyre et al. 2010; Wang et al. 2010). There are few obvious suggestions to improve the potency and selectivity provided by this group. A comparison of the flavopiridol molecule with this class of inhibitors does not suggest any simple hybrids of the two molecules to be synthesised. It is, however, possible that the location of the thiazole ring within these compounds is in a suitable position for chemical modifications to improve the pharmacokinetic and other drug-like properties of the inhibitors.

Although the structural analysis does not explain the conformational changes induced in CDK2/cyclin A by both S3-35 and S2-106 it suggests that these inhibitors provide unique characteristics that induce this novel conformation. Therefore slight modifications to these inhibitors may provide a significant decrease in CDK2 affinity but retain CDK9 affinity and thus be an approach to improve inhibitor selectivity for CDK9.

Finally, CDK9 contains a cysteine within the hinge region that could potentially be exploited by a complementary inhibitor that is able to form a covalent bond. Such a compound would irreversibly inhibit the kinase and would not require the specificity towards CDK9 that non-covalent inhibitors would. Only a handful of other kinases (BUBR1, CDK5, CDK10, CDKL-1,4, LKB1, MYT1, PLK4, SgK-110, 269) have a cysteine at this position and would therefore be susceptible to covalent modification by a covalent inhibitor targeting this residue (Zhang et al. 2009).

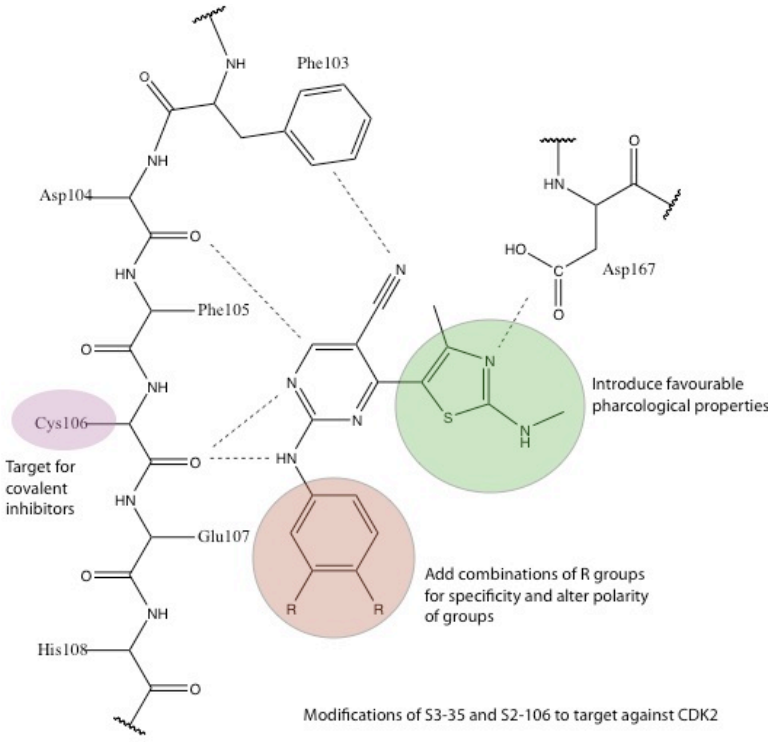


Figure 2-21. Proposed areas for chemical modification of 2-anilino-4-thiazole-pyrimidine.

2.5 Conclusion

Previously all CDK9/cyclin T structures containing bound inhibitors have been determined using $CDK9_{330}$/cyclin $T_{Q77R,\ E96G,\ F241L}$. Structural validation with $CDK9_{330}$/cyclin T_{F241L} and thermal denaturation data has shown that this combination of CDK9 and triply mutated cyclin T is a valid surrogate for CDK9/cyclin T structural studies of inhibitor binding modes. Furthermore, the improved quality of diffraction from crystals of $CDK9_{330}$/cyclin $T_{Q77R,\ E96G,\ F241L}$ mean that it is the preferred construct to use for crystallographic purposes.

An extensive array of CDK9/cyclin T/inhibitor structures presented in this chapter has greatly strengthened the hypothesis that CDK9/cyclin T has an inherently flexible ATP binding site that can adapt readily to accommodate a range of inhibitors. This is highlighted by the extensive flexibility in the glycine rich loop and the alternative relative orientations of the CDK lobes observed.

In addition to structural analysis of the 2-anilino-4-thiazole-pyrimidine compounds within the CDK9 binding site, their binding modes within CDK2 have been presented. In studying these inhibitors within CDK2 a novel conformation of CDK2 was identified. Binding of this class of inhibitors to the two kinases provides us with several observations to aid inhibitor design for improved CDK9 potency. From these observations a range of suggestions for second generation inhibitors have been proposed to improve specificity for CDK9 with respect to CDK2 and potentially other cyclin-dependent kinases.

3 Structure and function of the CDK9/cyclin T/Brd4 ternary complex

3.1 Introduction

CDK9/cyclin T is recruited to the sites of active transcription by transcription factors (as discussed in Chapter 1.6.2). Here it phosphorylates the Pol II pausing factors and the CTD of Pol II. These phosphorylations enable Pol II to transition from a paused state into the elongation phase, allowing transcription to continue along these genes.

The first general factor that was identified as recruiting CDK9/cyclin T to promoters was Brd4. The interaction between Brd4 and CDK9/cyclin T and the functional implications of this interaction are the focus of this chapter.

Brd4 (Bromodomain-containing protein 4) is also known as HUNK1 (hormonally up-regulated neu tumor-associated kinase 1) and MCAP (mitotic chromosome-associated protein). It was initially identified as a protein that regulates G2 to M transition, is able to remain bound to chromosomes during mitosis and the expression of which is related to stimuli for cell growth and inhibition (Dey et al. 2000). It was further identified as a protein involved in regulating cell-cycle progression (Maruyama et al. 2002; Dey et al. 2003; Mochizuki et al. 2008) controlling cell growth (Maruyama et al. 2002) and cell viability (Houzelstein et al. 2002) and, as mentioned previously, recruiting CDK9/cyclin T to transcription start sites (Mochizuki et al. 2008).

After the initial identification of Brd4's involvement in cell cycle transitions it was identified as a gene that is subject to chromosomal rearrangement found in aggressive forms of carcinoma (French et al. 2001). It has since been suggested as a valid target for both nuclear protein in testis (NUT) midline carcinoma and acute myeloid leukemia (AML). NUT midline carcinomas are caused by rearrangement of chromosomes 19 and 15, resulting in a fusion between NUT and Brd4 (NUT-Brd4) (French et al. 2003). As well as this, Brd4 has been implicated in contributing to breast cancer survival (Crawford et al. 2008). In support of Brd4 being a valid target for anticancer drug discovery JQ1, a drug that binds to the bromodomains of Brd4, caused significant growth arrest and apoptosis of NUT midline carcinoma cells and AML cells (Filippakopoulos et al. 2010; Zuber et al. 2011).

3.1.1 Brd4 in transcription

Brd4 is recruited to promoter regions by interactions with several different proteins and by a variety of signaling events (Figure 3-1).

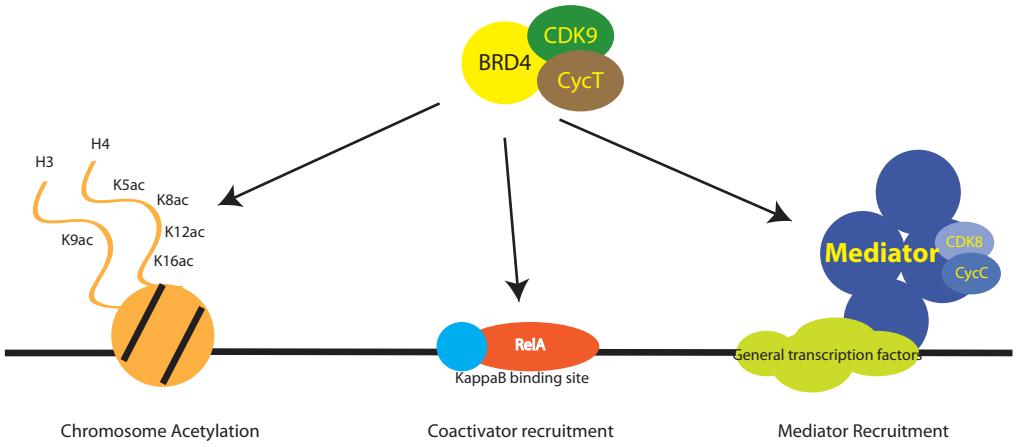


Figure 3-1. Mechanisms of recruitment of Brd4 to promoter regions. Brd4 can be recruited to promoter regions either in association with acetylated histones, RelA or the Mediator complex.

Prior to stimulation of primary response genes, both HDACs and specific histone methylation marks are present at the promoter region maintaining an inactive promoter region. On stimulation, specific kinases and HATs are recruited to promoters. These alter the histone code by phosphorylating and acetylating the histones. Subsequently Brd4, is recruited by its bromodomains to these histone acetylation marks recruiting P-TEFb in the process (Dey et al. 2003; Hargreaves et al. 2009; Zippo et al. 2009; Kapoor-Vazirani et al. 2011).

As well as interactions with acetylated histones, Brd4 can interact with additional proteins at the sites of transcription. Components of the mediator complex have been identified as interaction partners of Brd4 and therefore can stabilize the presence of Brd4 at promoter regions (Jang et al. 2005; Yang et al. 2005).

Also, Brd4 is recruited to the promoters of a subset of genes by interactions between its bromodomains and acetylated RelA (an element of the NF-κB signaling pathway). Here it acts as a coactivator of transcription and therefore is involved in increasing transcription of a subset of NF-κB target genes (Huang et al. 2009).

Of particular interest to transcription of certain genes is the fact that Brd4 remains bound to chromosomes during mitosis (Dey et al. 2000; Dey et al. 2003). Here it acts as an epigenetic marker for genes that are transcribed at the end of mitosis or early in G_1. Association of Brd4 with these genes survives the process

of mitosis and thereby enables their rapid transcription immediately thereafter (Dey et al. 2009).

3.1.2 The use of Brd4 by the Human Papilloma virus

Another protein that is capable of interacting with Brd4 and which may therefore bring CDK9/cyclin T to sites of transcription is the human papilloma virus (HPV) E2 protein. There are over 100 different strains of HPV, which are generally categorized into 'high risk' and 'low risk' depending on their resulting oncogenic phenotypes. The high risk strains such as HPV-16 and HPV-18 are associated with cervical cancer, whereas low risk forms, such as HPV-6 and HPV-11, are linked to benign lesions.

HPV has a small genome containing 8 open reading frames (ORFs), one of which encodes the HPV protein E2 that interacts with Brd4 (reviewed in (Hamid et al. 2009)). There is some variation in the function of E2 depending on the HPV strain, but in general it is involved in HPV replication, viral genome partitioning into daughter cells subsequent to mitosis, and transcription. The Brd4-E2 interaction is implicated in some of these functions and makes additional contributions to the lifecycle of HPV. Firstly, E2 interacts with both the HPV genome and Brd4. As Brd4 binds to host chromosomes during mitosis, several HPV strains have been suggested to utilize this system in order to enable compartmentalization of its genome into the daughter cell nuclei subsequent to mitosis (You et al. 2004; Abbate et al. 2006). Secondly, Brd4 stabilises E2, by preventing its degradation by the proteasome (Gagnon et al. 2009; Zheng et al.

2009). And finally, Brd4 has been shown to influence E2's transcriptional activator and repressor function (McPhillips et al. 2006; Schweiger et al. 2006; Wu et al. 2006; Smith et al. 2010) although this process is still poorly understood (Hamid et al. 2009).

E2 contains three domains: an N-terminal transcription regulation domain, a central "hinge" region, and a C-terminal β-barrel DNA-binding domain. It is the transcription regulation domain that interacts with the C-terminal region (CTR) of Brd4 and the structure of this complex that has been determined by X-ray crystallography (Abbate et al. 2006).

3.1.3 Overview of Brd4 domain structure

Brd4 is a member of the BET family of proteins. This family of proteins contains a protein-protein interacting extra-terminal (ET) domain (that interacts with several transcription proteins including the histone methyltransferase NSD3 that is associated with activating chromatin) (Rahman et al. 2011), a Ser/Glu/Asp-rich SEED region and a tandem array of acetyl-lysine binding bromodomains (see Figure 3-2 for domain organization). BET proteins are present in yeast, plants and animals. Yeast have two BET proteins, BDF1 and BDF2, whilst in humans there are Brds 2-4 and Brdt (reviewed in (Wu and Chiang 2007)).

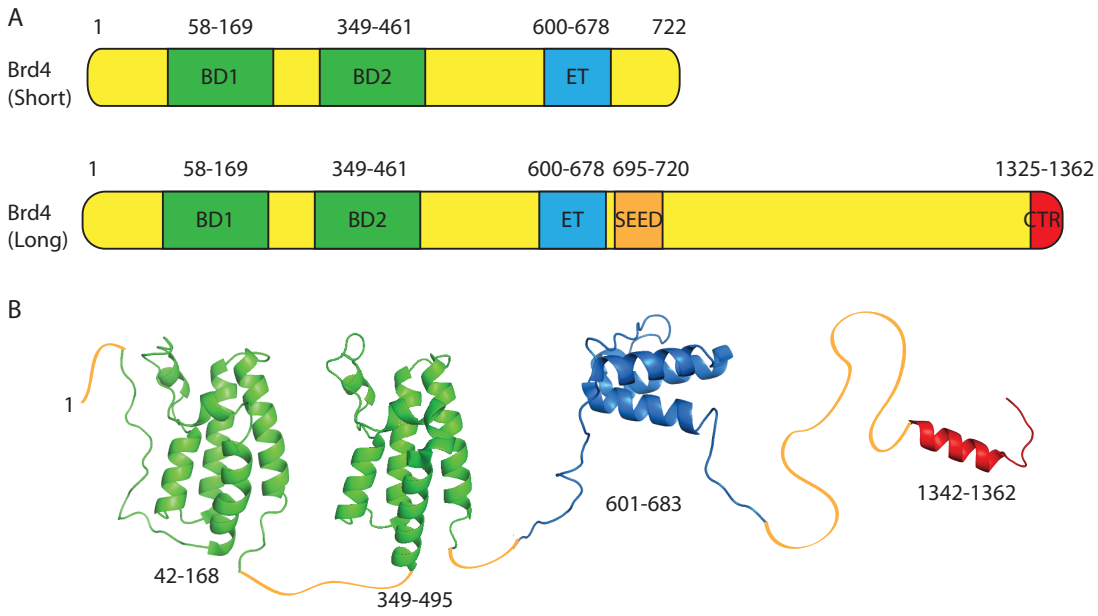


Figure 3-2. Domain structures of Brd4. A: Domains of the long and short isoforms of Brd4. B: Schematic of the long isoform of Brd4 including the known structures of the bromodomains (BD), extra-terminal (ET) and C-terminal region (CTR). The SEED motif is a Ser/Glu/Asp- rich region. (PDBs: BD1 - 2OSS; BD2 - 2OUO; ET - 2JNS; CTR - 2NNU)

Two isoforms of Brd4 exist. The short isoform consists of 947 amino acids whilst the longer isoform has an extended C-terminus and contains 1362 residues. These two isoforms are generated by alternative splicing and 3' end formation. The longer form is most abundant in humans (Mochizuki et al. 2008).

The structures of the two bromodomains (Liu et al. 2008; Vollmuth et al. 2009; Filippakopoulos et al. 2010), the ET domain (Lin et al. 2008) and the C-terminal region complexed with the human papilloma virus protein, E2, have been determined (Abbate et al. 2006). Residues between the ET domain and the CTR are predicted, by sequence analysis, to be unstructured.

3.1.4 Structural information of P-TEFb-protein interactions

Many factors involved in transcription have been suggested to bind to CDK9/cyclin T (Kanazawa et al. 2000; Lee et al. 2001; Tian et al. 2003; Giraud et al. 2004; Iankova et al. 2006; Hou et al. 2007; Nojima et al. 2008). However, there is still little known about how and where they bind to CDK9/cyclin T. Even within the inhibitory 7SK snRNP, few details are known about the interactions with CDK9/cyclin T and the other components within it. Whilst Larp7 interacts with CDK9 (Markert et al. 2008) (see Appendix E for work in progress on this and the Larp7-7SK interactions), 7SK interacts with residues 255-333 of cyclin T (Chen et al. 2004) and residues 254-272 of cyclin T are believed to interact with Hexim1 as Hexim1 competes with Tat for interaction with cyclin T (Garber et al. 1998; Michels et al. 2003; Schulte et al. 2005). The knowledge that has been garnered from *in vitro* and yeast two-hybrid binding studies show much diversity in the structural interaction sites of transcription factors with CDK9/cyclin T. For example STAT3 binds solely to CDK9 (Giraud et al. 2004), PPARγ binds both to CDK9 and cyclin T1/T2 independently (Iankova et al. 2006) and AhR binds to the C-terminus of cyclin T1 (Tian et al. 2003). It is not known whether the binding of CDK9 and/or cyclin T to these different putative partners is mutually permissive or mutually exclusive. This ambiguity obscures the likely composition of CDK9/cyclin T complexes at sites of transcription. The one exception where more information is known about the interaction site between CDK9/cyclin T and an interaction partner is for the interaction with HIV-Tat.

As mentioned in previous chapters, hijacking of P-TEFb by HIV-Tat results in the up-regulation of transcription from the HIV-LTR (Mancebo et al. 1997). HIV-Tat

is capable of competing with Brd4 for binding to CDK9/cyclin T (Yang et al. 2005; Bisgrove et al. 2007). Although the crystal structure of HIV-Tat bound to CDK9/cyclin T was published recently (Tahirov et al. 2010), little is known about the nature of the Brd4/CDK9/cyclin T complex. As such, it is possible that the competition between Brd4 and HIV-Tat for binding to CDK9/cyclin T results from shared or overlapping binding sites, or from an allosteric effect that communicates between two discrete binding sites. It is expected that a better picture of the Brd4-CDK9/cyclin T interaction will contribute to our understanding of the role of P-TEFb in transcription of both cellular and viral genes.

3.1.5 Aims

The aims of the experiments presented within the chapter include the following:

- To define the elements of Brd4 and P-TEFb that stabilize the Brd4/CDK9/cyclin T ternary complex.
- To establish the effect of Brd4-engagement on the activity of P-TEFb *in vitro*
- To develop a molecular model of the Brd4/CDK9/cyclin T complex.

3.2 Methods

3.2.1 Cloning

A codon-optimised gene encoding the Brd4-CTR (residues 1279-1362) was synthesized by Top Gene Technologies (www.topgenetech.com). PCR fragments of Brd4 were amplified from this plasmid using Phusion polymerase, digested with restriction enzymes SalI and BamHI, and ligated into the vector, pET28b (Table 3-1). The *E. coli* constructs to express the Brd4 tandem bromodomains (residues 44-460) and cyclin T2 (residues 7-263, N208D) (Baumli et al. 2008) were a gift from the Structural Genomics Consortium (SGC), Oxford.

Protein	Construct
Brd4	1303-1362
Brd4	1319-1362
Brd4	1322-1362
Brd4	1343-1362

Table 3-1. Constructs of Brd4 cloned for experiments. Throughout the text, Brd4 constructs are identified by including the residue numbers bracketed in subscript after the protein name.

3.2.2 Expressions and purification

3.2.2.1 Brd4

$Brd4_{(1279\text{-}1362)}$ was expressed in *E. coli* arctic express cells in LB. Cells were induced at cell density OD_{600} ~0.7, with 1mM isopropyl β-D-1-thiogalactopyranoside (IPTG) at 18°C overnight. Cell pellets were resuspended in buffer containing 500mM NaCl, 20mM Tris pH 8.0, 5mM DTT, 10% glycerol and cOmplete EDTA-free protease inhibitor cocktail tablets (Protease inhibitors,

Roche). Cells were lysed on ice by sonication, and lysozyme and benzonase were added. The suspension was centrifuged at 48,000g for 1hr at 4°C. Clarified lysate was diluted 1:1 in 500mM NaCl, 20mM Tris pH 8.0, 10% glycerol buffer and loaded onto a Ni-NTA agarose gravity flow column. The beads were washed in buffer containing 20mM imidazole, 500mM NaCl, 20mM Tris pH 8.0, 10% glycerol. Proteins were eluted in 500mM imidazole, 500mM NaCl, 20mM Tris pH 8.0, 10% glycerol. The His-tag was cleaved overnight by the addition of 25mM $CaCl_2$, 5mM DTT and 24U thrombin. The solution was subsequently dialysed for 2 hours into 20mM imidazole pH 6.0, 50mM NaCl, 5mM DTT. The proteins were separated by cation exchange (Mono S column), eluting with a linear buffer gradient from 50mM NaCl to 2M NaCl in 20mM MES pH 6.0, 5mM DTT. Fractions were collected, supplemented with 25mM $CaCl_2$, and cleaved further with thrombin for 2 hours at room temperature. This solution was passed over a gravity flow Ni-NTA column and the flow-through collected. Finally, the protein was purified by gel filtration (HiLoad Superdex75 26/60) equilibrated in 150mM NaCl, 20mM Tris pH 8.0, 5mM DTT. Fractions were pooled, concentrated to 2.2 mg/ml, aliquoted into eppendorf tubes, and flash frozen by immersion into liquid nitrogen, prior to storage at -80°C.

All other Brd4 proteins were expressed in *E. coli* arctic express cells in minimal media. Cells were induced at a cell density OD_{600} ~0.6 with 0.1mM IPTG at 18°C overnight. Cells were harvested and resuspended in buffer containing 500mM NaCl, 20mM Tris pH 8.0, 5mM DTT, 10% glycerol and supplemented with protease inhibitors, and stored at -20°C. After thawing, the cells were sonicated on ice and centrifuged for 1hr at 48,000g at 4°C. The clarified lysate was diluted

in 500mM NaCl, 20mM Tris pH 8.0, 10% glycerol to lower the DTT concentration for Ni affinity purification. This solution was loaded onto Ni-NTA agarose gravity flow column. After loading the beads were subsequently washed with 20mM imidazole, 1M NaCl, 20mM Tris pH 8.0, 10% glycerol. The beads were re-equilibrated into 500mM NaCl, 20mM Tris pH 8.0, 5mM DTT, 10% glycerol. 3C protease was incubated with the beads overnight to cleave the His-tag from Brd4. After elution, Brd4 was purified by gel filtration (HiLoad 16/60 Superdex30) in 150mM NaCl, 20mM Tris pH 8.0, 5mM DTT, 10% glycerol. Because of the lack of Trp, Tyr or Cys in these protein constructs the proteins were traced by detection of their absorbance at 254 nm. Purified Brd4 was concentrated to 1.7 mg/ml, aliquoted, flash frozen and stored at -80°C. Protein concentrations were estimated using Coomassie protein assay reagent (Pierce) calibrated with BSA solution.

^{15}N-labelled Brd4$_{(1322\text{-}1362)}$ was expressed in cells under the same conditions as unlabelled Brd4 in the minimal media with $^{15}NH_4Cl$ substituted for NH_4Cl. Lysis, centrifugation, affinity purification and 3C cleavage were undertaken as described above. The cleaved protein was purified by size exclusion gel filtration (HiLoad 16/60 Superdex30) in 50mM $NaPO_4$ pH 6.0, 2mM DTT. Brd4 containing fractions were pooled, concentrated, flash frozen in liquid nitrogen and stored at -80°C.

Brd4 peptides containing residues 1322-1345 and 1343-1362 were synthesized by Severn Biotech Ltd (www.severnbiotech.com). They were resuspended in

150mM NaCl, 20mM Tris pH 8.0, 5mM DTT to a protein concentration of 5mg/ml. The pH was additionally adjusted to pH 8 when necessary.

3.2.2.2 CDK9, cyclin T1 and CDK9/cyclin T1

The CDK9/cyclin T1 complexes were expressed and purified as in Chapter 2.2.4. CDK9 and cyclin T1 were expressed individually in *Sf9* insect cells and purified by the same method as the complex. The buffer used during size exclusion chromatography varied depending on the subsequent experiment in which the proteins were to be used. Proteins for kinase assays were purified in 150mM NaCl, 20mM Tris pH 8.0, 5mM DTT. CDK9/cyclin T for NMR was purified in 50mM $NaPO_4$ pH 8.0, 100mM NaCl, 10mM DTT.

3.2.2.3 Cyclin T2

For NMR, cyclin T2 was expressed in *E. coli* in LB. Cells were induced at a cell density corresponding to OD_{600}: 0.6, with 0.1mM IPTG at 18°C for 5 hours. Cell pellets were resuspended in 50mM HEPES pH 7.5, 500mM NaCl, 5% glycerol, 5mM DTT and protease inhibitors and stored at -20°C. Cells were lysed on ice by sonication, and centrifuged at 48,000g for 1 hour at 4°C. The soluble lysate fraction was diluted 1:1 in 500mM NaCl, 20mM Tris pH 8.0, 5% glycerol and loaded onto a gravity flow Ni-NTA agarose column. The beads were washed with 20mM imidazole, 20mM Tris pH 8.0, 500mM NaCl, 5% glycerol. The protein was eluted in 500mM imidazole, 20mM Tris pH 8.0, 10% glycerol. The His tag was cleaved off overnight at 4°C with TEV protease whilst dialysed into 50mM Tris pH 8.0, 1mM DTT, 100mM NaCl. The cleaved cyclin T2 was purified further by gel

filtration (HiLoad 26/60 Superdex 75) in 50mM $NaPO_4$ pH 6.5, 10mM DTT, 100mM NaCl. Fractions containing cyclin T2 were concentrated, aliquoted, flash frozen in liquid nitrogen, and stored at -80°C.

For kinase assays cyclin T2 was purified similarly. However, the size exclusion chromatography column was run in 150mM NaCl, 20mM Tris pH 8.0, 5mM DTT.

3.2.2.4 GST-CTD

The GST-CTD, a 52 heptad repeat of the CTD fused to GST, was expressed in *E. coli* in terrific broth (TB). Cells grew to a density of OD_{600}: 0.55 before being induced at 16°C with 1mM IPTG overnight. Cells were pelleted, and resuspended in 500mM NaCl, 20mM Tris pH 8.0, 5mM DTT and protease inhibitors. Cells were lysed on ice by sonication and lysozyme in the presence of Dnase. After centrifugation, at 48,000g for 1 hour at 4°C, the clarified lysate was loaded onto glutathione Sepharose 4B beads in a gravity flow column. The beads were subsequently washed with 1M NaCl, 20mM Tris pH 8.0, 5mM DTT. GST-CTD was eluted in 20mM glutathione, 500mM NaCl, 20mM Tris pH 8.0, 5mM DTT and flash frozen. Aliquots were stored at -80°C.

3.2.2.5 Plk1

pACYC-Duet-PKA-MBP-KD38312 was expressed in *E. coli*. The vector contains both murine PKA and human Plk1 kinase domain residues 38-312 for co-expression (Cheng 2006). Cells were grown to OD_{600} *circa* 0.7 in LB and induced with 1mM IPTG at 16°C overnight. Cells were resuspended in Plk1 buffer

(300mM NaCl, 50mM Tris pH 8.0, 0.01% MTG, 5mM imidazole, 10% glycerol and protease inhibitors). After sonication and centrifugation, the lysate was loaded onto Ni-NTA beads. The beads were washed with Plk1 buffer, and proteins were eluted in Plk1 buffer containing 200mM imidazole. The protein was incubated with 3C protease overnight. The elution was purified by size exclusion chromatography (HiLoad 26/60 Superdex 200) with 50mM Tris pH 8.0, 300mM NaCl, and 0.1% MTG. Fractions corresponding to those containing Plk1 were pooled, further purified with Ni-NTA beads to remove contaminating PKA and MBP, and concentrated. Protein was flash frozen and stored at -80°C.

3.2.3 Pull-down analysis

Brd4 constructs were expressed in ~50 ml LB. After sonication and centrifugation, His-Brd4 lysates were incubated with Ni-NTA agarose for 1hr at 4°C. The beads were washed with ice-cold 10mM β-mercaptoethanol, 0.1% NP-40 in 20mM imidazole, 500mM NaCl, 20mM Tris pH 8.0, 10% glycerol. 70µg $CDK9_{330}$/cyclin $T_{Q77R,\ E96G,\ F241L}$ was added to the beads and incubated for 2 hours at 4°C with rotation, followed by 4 washes. Proteins were separated by SDS-PAGE and detected by coomassie stain.

3.2.4 Analytical gel filtration

An analytical Superdex 200 gel filtration column was equilibrated in 150mM NaCl, 20mM Tris pH 8.0, 5mM DTT, 10% glycerol. 0.5mg His-$Brd4_{(1279\text{-}1362)}$ and 0.8mg $CDK9_{330\ T186A}$/cyclin $T_{Q77R,\ E96G,\ F241L}$ were incubated in a total volume of

200µl on ice for 20 mins before injection. For controls, 0.5mg His-Brd4$_{(1279-1362)}$ (200µl) and 0.2mg CDK9/cyclin T (50µl) were loaded individually.

3.2.5 Kinetic analysis of Brd4 as a modulator of CDK9/cyclin T

Standard kinetic analysis was carried out with 10ng CDK9/cyclin T complexes, 100µM ATP, 15ng GST-CTD in 10mM $MgCl_2$, 50mM Tris pH 8.0, 1mM DTT and γ-^{32}P-ATP (Perkin Elmer) in a total volume of 10µl. Brd4, GST-CTD and ATP concentrations were varied depending on the experiment. Samples were incubated for 5 minutes at 30°C and the reaction stopped by the addition of SDS-PAGE loading buffer. Proteins were separated by SDS-PAGE gel. After staining, destaining and drying the gel, the ^{32}P incorporated into the GST-CTD was measured using a phosphor imager and analysed with QuantityOne (Version 4.6.3, Bio-Rad Laboratories) and GraphPad Prism (Version 5.0d for Mac OS X, GraphPad Software, San Diego California USA, www.graphpad.com).

Cyclin T isoform analysis was studied under the following experimental conditions. 28ng CDK9$_{330}$ and 25ng of cyclin T1$_{Q77R,\ E96G,\ F241L}$ or cyclin T2 were incubated with 15µg GST-CTD, 100µM ATP, 10mM $MgCl_2$, 50mM Tris pH 8, 1mM DTT and 0.5µCi γ-^{32}P-ATP in a total volume of 10µl. Experiments were done in the presence or absence of 600ng Brd4$_{(1322-1362)}$. Samples were incubated for 5 minutes at 30°C and the reaction stopped by the addition of SDS-PAGE loading buffer. Proteins were separated by SDS-PAGE gel and analysed as above.

3.2.6 NMR

3.2.6.1 Assignment

^{15}N Brd4$_{(1322\text{-}1362)}$ assignment in 6M urea

Urea was added to a final concentration of 6M to ^{15}N-labeled Brd4$_{(1322\text{-}1362)}$. The protein solution was concentrated to *circa* 2.3 mg/ml (440μM). D_2O was added to a final concentration of 5%. All NOESY and TOCSY experiments were measured on a 600MHz NMR spectrometer at 15°C. 3D NOESY-HSQC and TOCSY-HSQC data were collected with 400ms and 70ms mixing times respectively. All NMR data were processed using nmrPIPE (Delaglio et al. 1995) and viewed using CcpNmr Analysis v2.1.5 (Fogh et al. 2006).

Transfer of assignments to 0M urea, pH 6.5, 20°C

A series of HSQC spectra were obtained of ^{15}N-labeled Brd4 in sequential dilutions of urea. Initial spectra were taken at 50mM $NaPO_4$ pH 6.0, 2mM DTT, 6M urea, 5% D_2O and at 15°C. Subsequent spectra were obtained at 5M, 4.2M, 3.5M, 2.9M, 2M, 1M urea by dilutions in 50mM $NaPO_4$ pH 6.5, 2mM DTT, 5% D_2O. Brd4 in 0M urea, 50mM $NaPO_4$ pH 6.5, 2mM DTT, 5% D_2O was used for final spectra taken at 15°C and 20°C. All HSQC spectra were obtained on a 600 MHz NMR spectrometer.

3.2.6.2 NMR Interaction analysis

HSQC spectra of 50μM Brd4 both in complex with and in the absence of 20μM CDK9/cyclin T were collected using a 750 MHz spectrometer. The samples were

in 5% D_2O, 0.4mM DSS (4,4-dimethyl-4-silapentane-1-sulfonic acid), 50mM sodium phosphate, pH 6.5, 100mM NaCl, 10mM DTT and data acquired at 20°C.

15μM and 30μM cyclin T2 were incubated with 65μM ^{15}N $Brd4_{(1322\text{-}1362)}$ at pH 6.5 in the presence of 5% D_2O. HSQC spectra were obtained at 15°C on a 600 MHz spectrometer. Combined chemical shifts were calculated as follows

$$Combined\ chemical\ shift\ change = \sqrt{(\Delta\delta_{1H})^2 + \left(\frac{\Delta\delta_{15N}}{5}\right)^2}$$

3.2.7 H/D exchange mass spectrometry

$Brd4_{(1322\text{-}1362)}$ was purified as described above. $CDK9_{FL}$/cyclin T_{288} was purified from cell lysates by affinity purification as described earlier (Chapter 2.2.4). Subsequent to 3C protease cleavage and elution from the affinity column, CDK9/cyclin T was aliquoted. CDK9/cyclin T complex was purified by gel filtration (Superdex 200 10/300 GL) at 4°C. For the purification of CDK9/cyclin T/Brd4, an excess of Brd4 was incubated with CDK9/cyclin T prior to gel filtration. For both complexes the gel filtration was run in 150mM NaCl, 20mM Tris pH 8.0, 10% glycerol, 3mM DTT. CDK9/cyclin T was concentrated to 0.85 mg/ml and CDK9/cyclin T/Brd4 was concentrated to 2.8 mg/ml. Samples were analysed by hydrogen/deuterium-exchange mass spectrometry in collaboration with ExSAR (New Jersey, USA, www.exsar.com).

3.2.8 Brd4 phosphorylation

Initial Radioactive analysis

2.2μg $Brd4_{(1279-1362)}$ was incubated with 100μM ATP, 0.15μCi γ-^{32}P-ATP, in the presence and absence of 2μg Plk1 in 10mM $MgCl_2$, 50mM Tris pH 8.0, 1mM DTT for 1 hour at 30°C. The reaction volume was 10μl. Phosphorylation was terminated by the addition of SDS-PAGE running buffer. Samples were separated by SDS-PAGE gel and analysed as above.

Purification of phosphorylated Brd4

250μg $Brd4_{(1322-1362)}$ was incubated with 70μg Plk1, in the presence of 1mM ATP and in 1x kinase buffer (10mM $MgCl_2$, 50mM Tris pH 8.0, 1mM DTT). The proteins were incubated at 30°C for 120 minutes.

Phosphorylated Brd4 was separated from ATP, ADP and Plk1 by gel filtration (HiLoad 16/60 Superdex 30) in 150mM NaCl, 20mM Tris, 5mM DTT. Phosphorylation was confirmed by ESI-MS.

Kinase assays with phosphorylated Brd4

10ng $CDK9_{FL}$/cyclin T_{WT} was incubated with increasing concentration of phosphorylated or unphosphorylated Brd4. The reaction was in 100 mM ATP, 15 μg GST-CTD, 1x kinase buffer (10 mM $MgCl_2$, 50 mM Tris pH 8.0, 1 mM DTT), 0.2 μCi γ-^{32}P-ATP. The reaction volumes were 10μl. Samples were incubated at 30°C for 5 minutes. SDS-PAGE loading buffer was added to terminate the reaction and the incorporation of ^{32}P was measured as above.

3.2.9 Model of Brd4/cyclin T

Superpose (Ccp4) was used to superimpose residues of Brd4 upon putative equivalent amino acids of Tat. Tat coordinates were drawn from the structure of CDK9/cyclin T/Tat (PDB: 3MIA), while Brd4 residues were drawn from the structure of E2/Brd4 (PDB: 2NNU). Brd4 residues 1346-1353 or 1350-1357 were superposed on Tat residues 35-42.

3.3 Results

3.3.1 Domain interactions

Several interaction sites between Brd4 and CDK9/cyclin T have been mapped previously. Brd4-CDK9/cyclin T interactions have been localised to both the bromodomains and the C-terminal region of Brd4 (Jang et al. 2005; Bisgrove et al. 2007; Vollmuth et al. 2009). In addition Brd4-binding domains have been identified on CDK9, cyclin T within residues 426-516 (Cho et al. 2007) and implicated at the Tat-binding domain on cyclin T (Jang et al. 2005; Bisgrove et al. 2007).

To determine the fragments of Brd4 that could interact with the catalytic domains of CDK9 and cyclin T *in vitro* pull-down analysis was undertaken (Figure 3-3). These experiments were used to identify the interaction sites between Brd4 and CDK9/cyclin T. Two His-tagged constructs of Brd4 were expressed: the tandem bromodomain construct ($Brd4_{(44-460)}$), and the C-terminal region of Brd4 ($Brd4_{(1279-1362)}$). Whereas His-tagged $Brd4_{(1279-1362)}$ could mediate the pulldown of $CDK9_{330}$/cyclin $T_{Q77R, E96G, F241L}$, His-tagged $Brd4_{(44-460)}$ could not.

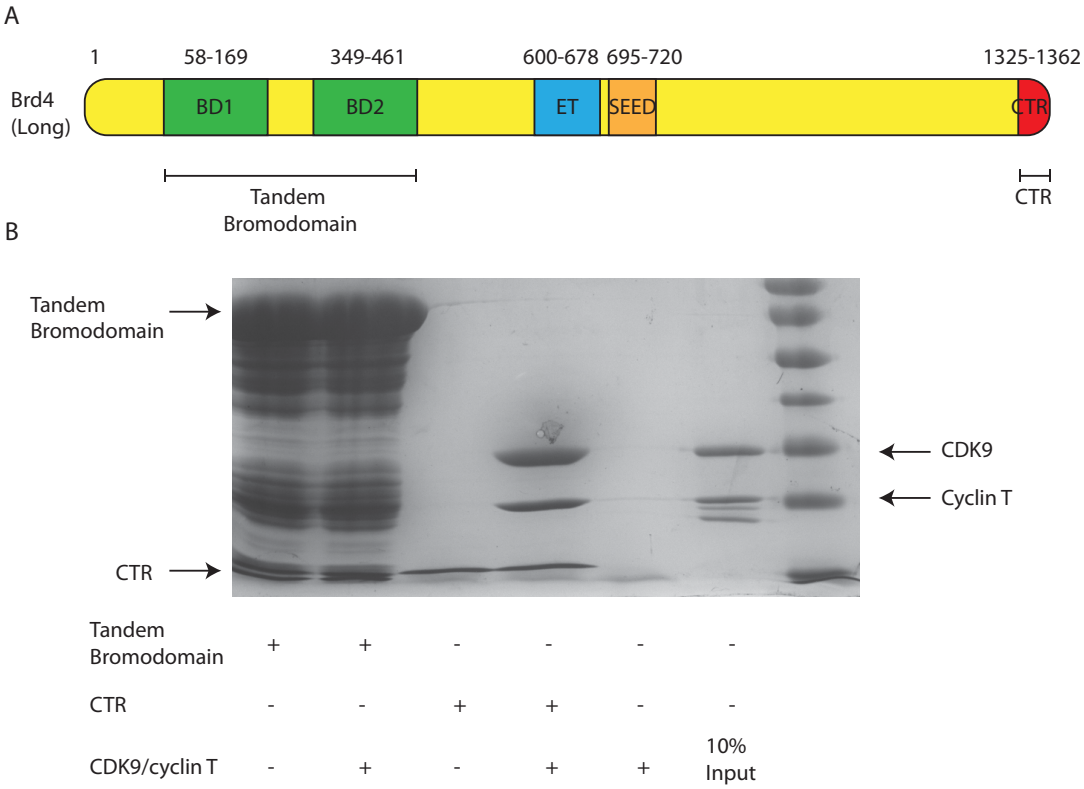


Figure 3-3: The C-terminal region (CTR) of Brd4 interacts with the minimum kinase domains of CDK9/cyclin T. A: Domain structure of Brd4 showing the tandem bromodomain construct and C-terminal region construct. B: Pulldown of $CDK9_{330}$/cyclin $T_{Q77R,\ E96G,\ F241L}$ using His-Brd4 constructs. $Brd4_{(1279\text{-}1362)}$ that encodes the C-terminal region, was readily purified on Ni-NTA beads. However, the tandem bromodomains proved to be unstable resulting in multiple C-terminally truncated products. The protein bands in this gel (and all gels presented in this thesis) are stained with Coomassie blue. Molecular weight markers (Right hand lane) correspond to proteins of masses 17, 26, 34, 43, 55, 72, 95 kDa.

3.3.2 Analytical gel filtration analysis

Gel filtration was used to both confirm the interaction detected by pull-down analysis and to assess the stability of the CDK9/cyclin T/Brd4 complex. The His-$Brd4_{(1279\text{-}1362)}$/$CDK9_{(330\ T186A)}$/cyclin $T_{(Q77R,\ E96G,\ F241L)}$ complex and the individual components were run on an analytical gel filtration column. No significant change in elution volume was observed between the CDK9/cyclin T peaks in the presence and absence of Brd4. However, SDS-PAGE analysis of selected fractions revealed that Brd4 co-elutes with CDK9/cyclin T when all three proteins are present (Figure 3-4). These results confirm an interaction between the CTR of

Brd4 and CDK9/cyclin T. From previous experiences within the lab, this behaviour is observed for complexes with the dissociation constant (K_d) ≤1μM.

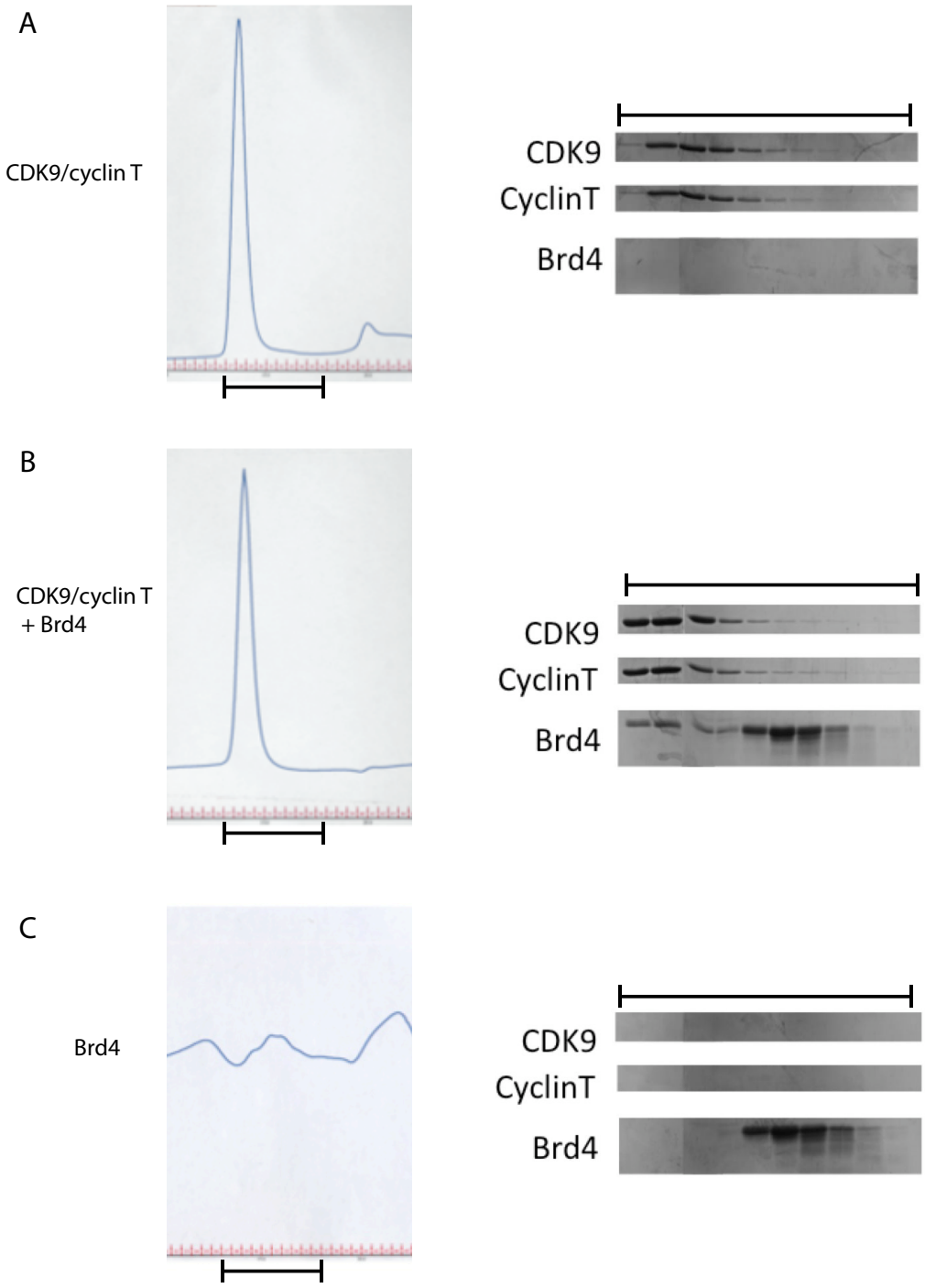


Figure 3-4: Analysis of the interaction between CDK9/cyclin T and the Brd4-CTD (His-Brd4$_{(1279\text{-}1362)}$). In each panel, the LHS figure shows the gel filtration chromatogram (measured at 280 nm) and the RHS an analysis of selected fractions by SDS-PAGE. Fractions shown for each experiment correspond to identical elution times. A bar under each gel filtration profile highlights those fractions run. Brd4 does not absorb at 280nm and hence the chromatogram for Brd4 shows no discernible peak. (A) CDK9/cyclin T, (B) CDK9/cyclin T/Brd4, (C) Brd4.

3.3.3 Identification of the minimum Brd4 fragment that interacts with CDK9/cyclin T

Based on secondary structural analysis, constructs of the Brd4-CTR varying in length were cloned to identify a minimum sequence that can interact with CDK9/cyclin T (Figure 3-5). These constructs were used for initial pull-downs with $CDK9_{330}$/cyclin $T_{Q77R,E96G,F241L}$. All constructs expressed well and bound to CDK9/cyclin T (Figure 3-6). The very C-terminal 40 residues of Brd4 are therefore sufficient to interact with $CDK9_{330}$/cyclin $T_{Q77R,E96G,F241L}$.

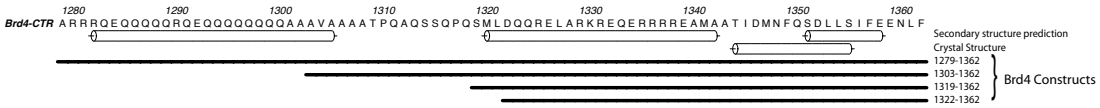


Figure 3-5: Domain structure of Brd4, sequence analysis and cloned constructs. Helices, predicted by Jpred3 (Cole et al. 2008) or observed in the crystal structure of PDB file 2NNU, are depicted as cylinders. Three further constructs identified by the three lowest lines were cloned from the full-CTR construct.

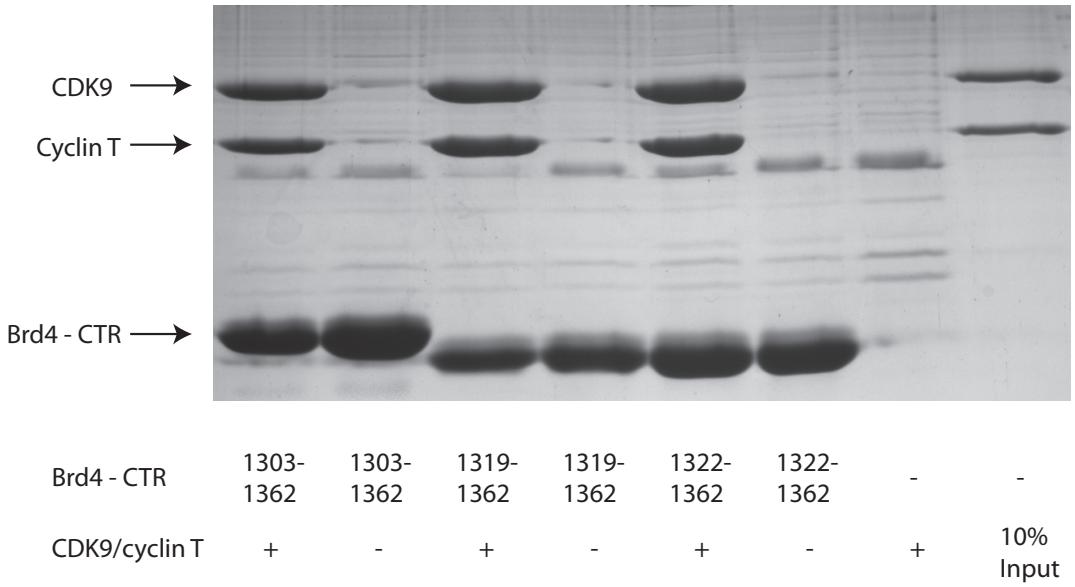


Figure 3-6: Residues 1322-1362 of Brd4 pull-down CDK9/cyclin T. His-Brd4 constructs were used as bait to capture $CDK9_{330}$/cyclin $T_{Q77R,E96G,F241L}$ on Ni-NTA agarose beads.

3.3.4 C-terminal constructs of Brd4 activate CDK9$_{FL}$/cyclin T1$_{WT}$

Having identified that the C-terminal 40 residues of Brd4 are sufficient for binding CDK9/cyclin T, the influence of Brd4 on CDK9 kinase activity was investigated. For these experiments, the phosphorylation of GST-CTD by CDK9/cyclin T was assayed in the presence of increasing amounts of Brd4. Assays were undertaken using standard kinase conditions (see Chapter 3.2.5) and using a GST-tagged version of the 52 heptad repeat CTD. Phosphotransfer was detected by measuring the incorporation of radioactive phosphate ^{32}P from ^{32}P-γ-ATP into the CTD substrate. All kinase assays were incubated at 30°C for 5 minutes to ensure the experiments were performed in the initial rate phase (see Appendix C). An increase in ^{32}P incorporation therefore relates directly to an increase in kinase activity.

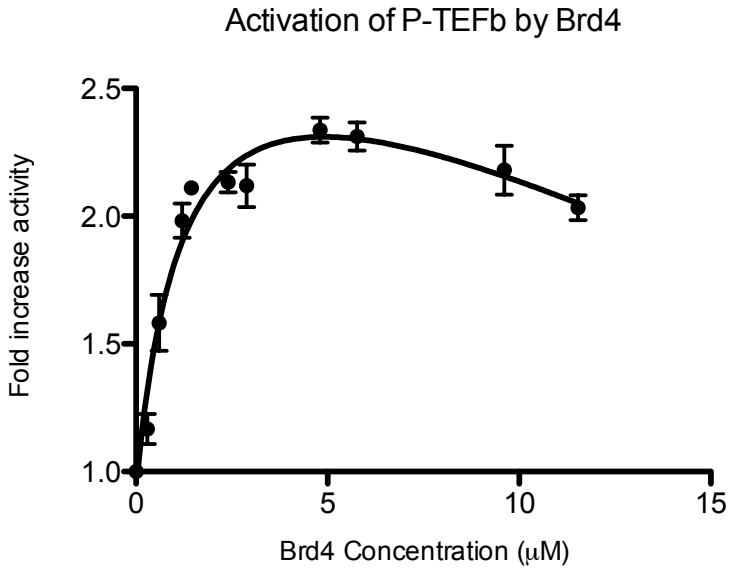


Figure 3-7. Brd4$_{(1322-1362)}$ stimulates activation of CDK9$_{FL}$/cyclin T$_{WT}$. Activity of CDK9/cyclin T was measured through the incorportation of ^{32}P.

Increasing the concentration of Brd4$_{(1322-1362)}$ increased the kinetic activity of CDK9$_{FL}$/cyclin T$_{WT}$ by a maximum of 2.2 fold (Figure 3-7). Subsequent addition

of Brd4 reduces the effect on CDK9 and the CDK9 activity drops. The reason for this is unclear but is likely to be an artifact of the assay system. A similar effect has been observed in the lab when CDK9, in the presence of random peptides, had a lower kinase assay than in the absence of the peptides.

Using a one site binding model, and assuming that activity is directly related to interaction, the predicted K_d is 1.3μM +/- 0.35μM. This is in agreement with the complex being stable on gel filtration (see Figure 3-4) and is within the range of a moderate/weak protein-protein interaction (Kastritis et al. 2011).

3.3.5 The C-terminal 40 residues possesses the full Brd4 CTR activator effect

The length of the Brd4 CTR may be a more extended section of the C-terminus of Brd4 than the 40 residues tested above. In order to determine if longer constructs were capable of greater enhancement of CDK9 activity or if shorter constructs remained sufficient for activating CDK9/cyclin T a variety of Brd4 constructs were tested for their activity (Figure 3-8 and Figure 3-9).

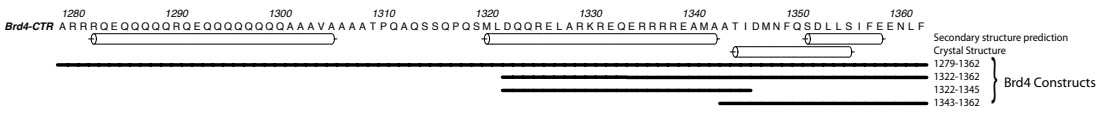


Figure 3-8. Domain structure of Brd4 and constructs tested for kinase activity. In addition to $Brd4_{(1279-1362)}$ and $Brd4_{(1322-1362)}$, $Brd4_{(1322-1342)}$ and $Brd4_{(1342-1362)}$, shown above by lines, were used to investigate the residues of Brd4 required to influence the kinase activity of CDK9/cyclin T. As before, the predicted secondary helical structure and that observed in the crystal structure of PDB file 2NNU are depicted as cylinders.

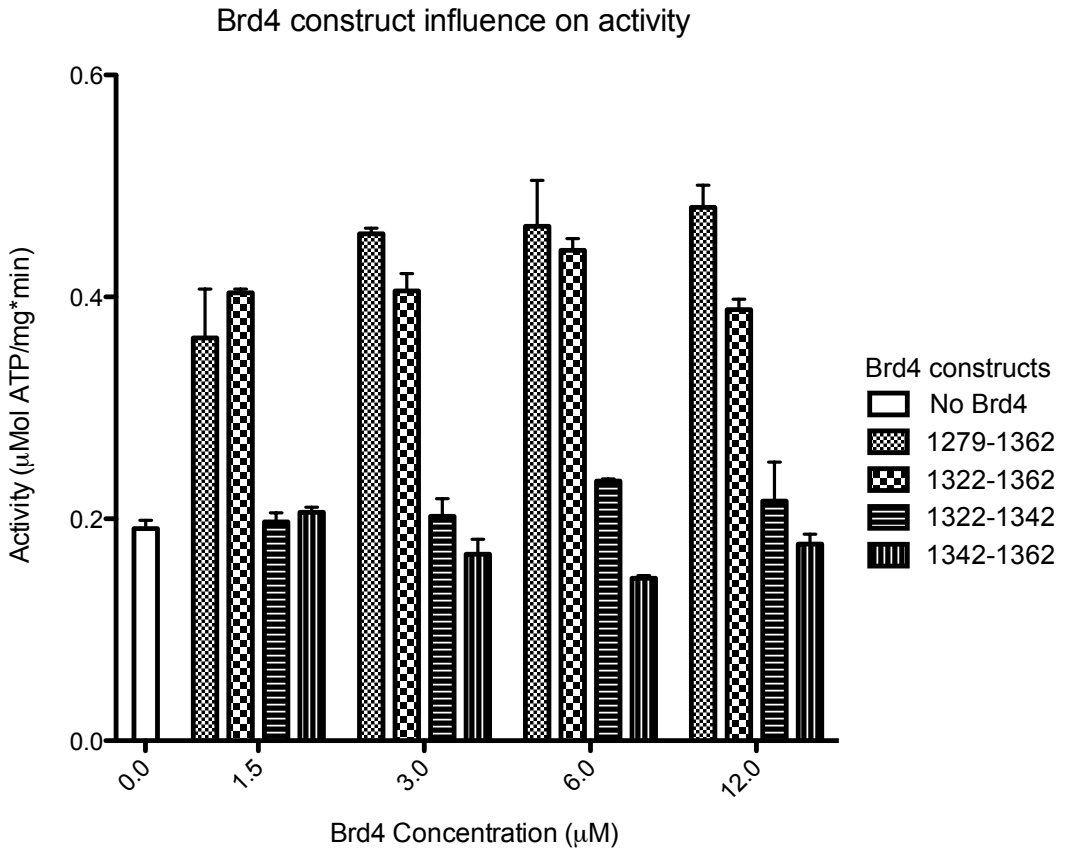


Figure 3-9. Relative increase in activity of $CDK9_{FL}$/cyclin $T1_{WT}$ in the presence of a range of Brd4 CTR constructs at increasing concentrations.

Whilst increased concentrations of $Brd4_{(1279\text{-}1362)}$ and $Brd4_{(1322\text{-}1362)}$ increase the activity of $CDK9_{FL}$/cyclin T_{WT} to similar extents, the shorter constructs of $Brd4_{(1322\text{-}1342)}$ and $Brd4_{(1342\text{-}1362)}$ have little effect on CDK9 activity. Taken together, these results provide strong evidence that residues 1322-1362 of Brd4 are a minimal functioning unit that activates CDK9.

3.3.6 P-TEFb activation is brought about by increasing V_{max} while leaving K_m largely unaltered

To determine the mechanism by which Brd4 increases the kinetic activity of CDK9, a more detailed kinetic analysis was undertaken. Varying the

concentration of substrate enables the K_m (the concentration of substrate at half maximum enzymatic activity) and V_{max} (the enzymes maximum velocity under the conditions tested) of CDK9/cyclin T to be determined with respect to the titrated substrate. Undertaking these experiments in the presence of increasing concentration of Brd4 determines the effect that Brd4 has on these two kinetic parameters. Such experiments provide more detailed information on whether Brd4 enhances substrate binding or increases the maximum turnover rate that the kinase is able to achieve (Dixon and Webb 1979; Leskovac 2003; Cornish-Bowden 2004).

Evaluating the effect of titrating Brd4 constructs into assays that contained varying concentration of either substrate, ATP or GST-CTD, showed that Brd4 activates CDK9/cyclin T under a range of substrate concentrations (Figure 3-10 and Figure 3-11). Activities measured while varying ATP concentration reached a plateau at high ATP concentrations, enabling more accurate determination of V_{max} and K_m. The same levels of saturation could not be achieved for the CTD due to its limited solubility.

The data was analysed based on the Michaelis-Menten model and equation,

$$v = \frac{V_{max}[S]}{K_m + [S]}$$

where, v is the enzyme velocity, V_{max} is the enzymes maximum velocity, [S] refers to the substrate concentration, and K_m is the substrate concentration at which the enzyme yields half-maximal velocity.

Rates measured while varying only the concentrations of ATP and Brd4 show that in the presence of higher concentrations of Brd4 the kinase activity of CDK9 saturates at a higher rate (Table 3-2). However, the half-concentration of ATP at which this occurs remains unaltered independent of the Brd4 concentration. This shows that Brd4 does not affect the ATP K_m but affects the V_{max} with respect to ATP. This indicates that the affinity of CDK9/cyclin T for ATP is unaltered in the presence of Brd4. In effect, at this CTD concentration the Brd4 is acting in a manner analogous to a non-competitive inhibitor. The Brd4 will bind to a location that is separate from the ATP binding site, but may increase the overall rate of the enzyme-catalysed reaction.

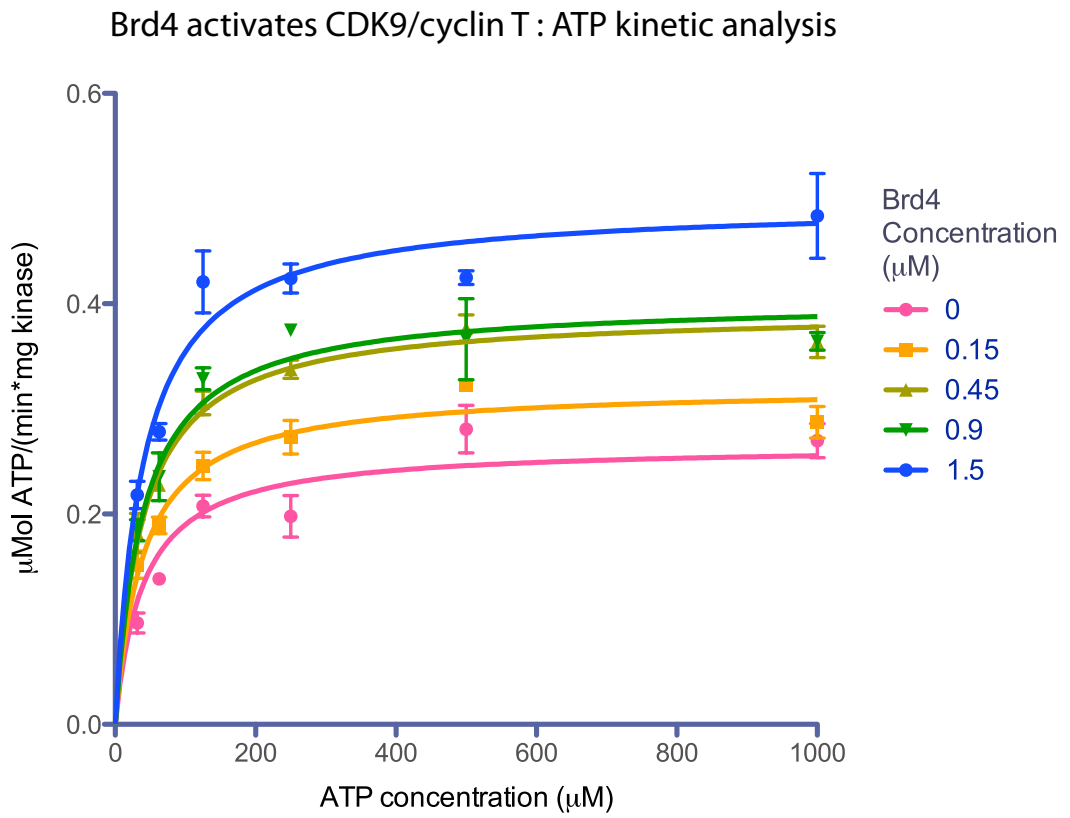


Figure 3-10. Analysis of ATP-related kinetic parameters of CDK9$_{FL}$/cyclin T$_{WT}$ in the presence of Brd4. Data is presented in a non-linear format. The Michaelis-Menten equation was used for analysis and statistical analysis was undertaken using Akaike's Informative Criteria (AICc) (Akaike 1974; Hurvich and Tsai 1989).

Brd4 (µM)	0	0.15	0.45	0.9	1.5	Overall
V_{max} (µMol ATP/mg)	0.27	0.32	0.39	0.40	0.50	-
V_{max} std error (µMol ATP/mg)	0.01	0.01	0.01	0.01	0.01	-
K_m (µM)	39.81	39.81	39.81	39.81	39.81	39.81
K_m std error (µM)	3.09	3.09	3.09	3.09	3.09	3.09
R-factor	0.78	0.85	0.92	0.82	0.86	0.90

Table 3-2. Analysis of CDK9/cyclin T kinetic values with respect to ATP. Analysis of Graph A, Figure 3-9 determined the above values and errors of the data. The values were determined assuming a constant K_m across datasets (subsequent to AICc statistical analysis).

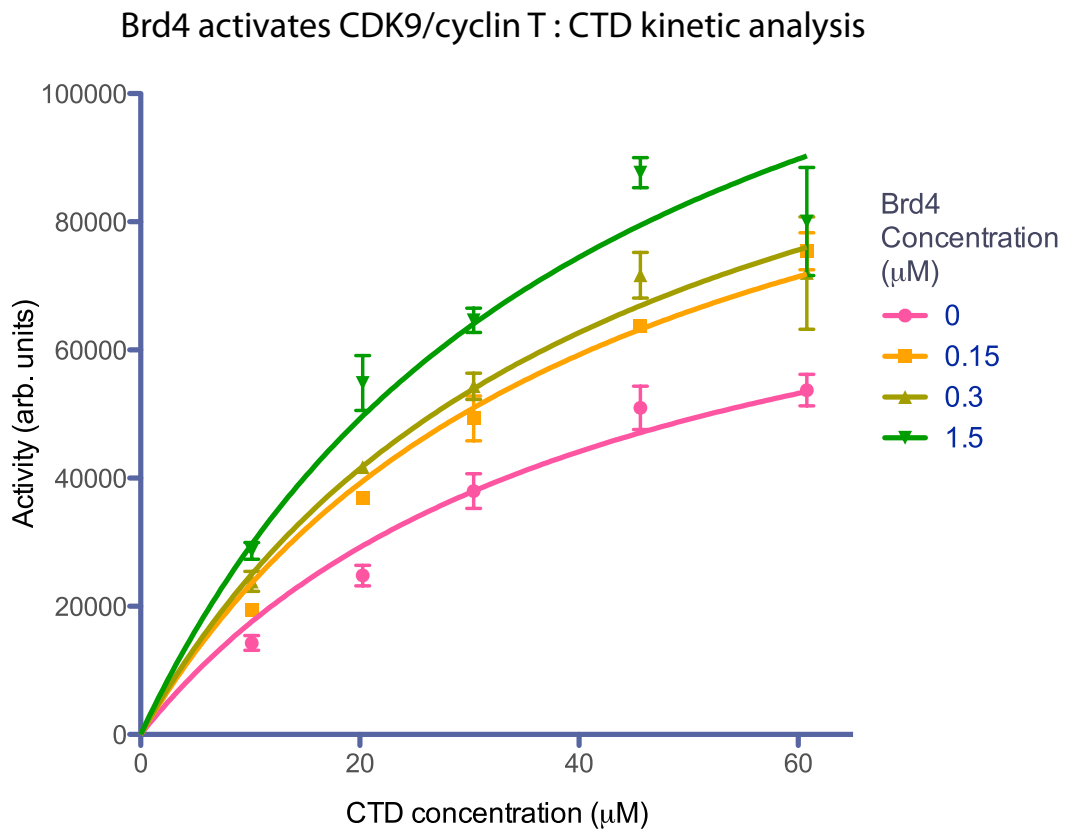


Figure 3-11. Analysis of CTD-related kinetic parameters of CDK9$_{FL}$/cyclin T$_{WT}$ in the presence of Brd4. Data is presented in a non-linear format. The Michaelis-Menten equation was used for analysis and statistical analysis was undertaken using Akaike's Informative Criteria (AICc).

Brd4 (μM)	0	0.15	0.3	1.5	Global
V_{max} (Arb units)	90453	121438	128517	152630	-
V_{max} std error	8040	10178	10681	12342	-
K_m (μM)	41.99	41.99	41.99	41.99	41.99
K_m std error (μM)	6.380	6.380	6.380	6.380	6.380
R-factor	0.9121	0.9583	0.8761	0.7930	0.91

Table 3-3. Analysis of CDK9/cyclin T kinetic values with respect to CTD. Analysis of Graph A, Figure 3-10 determined the above values and errors of the data. The values were determined assuming a constant K_m across datasets (subsequent to AICc statistical analysis).

The CTD analysis revealed the same results as seen for ATP – the K_m remains constant and the V_{max} increases with respect to the CTD (Table 3-3 and Figure 3-11). This therefore further implies that Brd4 binds to neither substrate binding sites but inherently increases the activity of the kinase.

3.3.7 Both full-length and truncated CDK9/cyclin T are activated by Brd4

CDK9 with the truncated C-terminal tail is the construct that we have mostly used for crystallographic analysis due to its superior crystallogenesis. Recently it has been found that the C-terminal tail influences the kinetic mechanism of substrate phosphorylation by enforcing an ordered kinetic reaction (Baumli et al, Manuscript in preparation). In this case ATP binds first, followed by the GST-CTD. After the phosphotransfer has occurred the phosphorylated CTD is released before finally ADP is released. In the absence of the C-terminal tail the kinetic reaction is a random event reaction.

To determine if the C-terminal tail of CDK9 influences the activation of CDK9/cyclin T by Brd4, the increase in kinase activity of $CDK9_{330}$/cyclin T_{WT} was measured in the presence of an increasing concentration of Brd4. The results (Figure 3-12) show an increase in activity due to Brd4 similar to that observed for full length CDK9. Thus the enhancement of kinase activity effected by Brd4 appears to be independent of the precise kinetic pathway followed by the CDK9 constructs considered. Its activation effect is therefore independent of the C-terminal tail and the ordered kinetic mechanism imposed by the C-terminus.

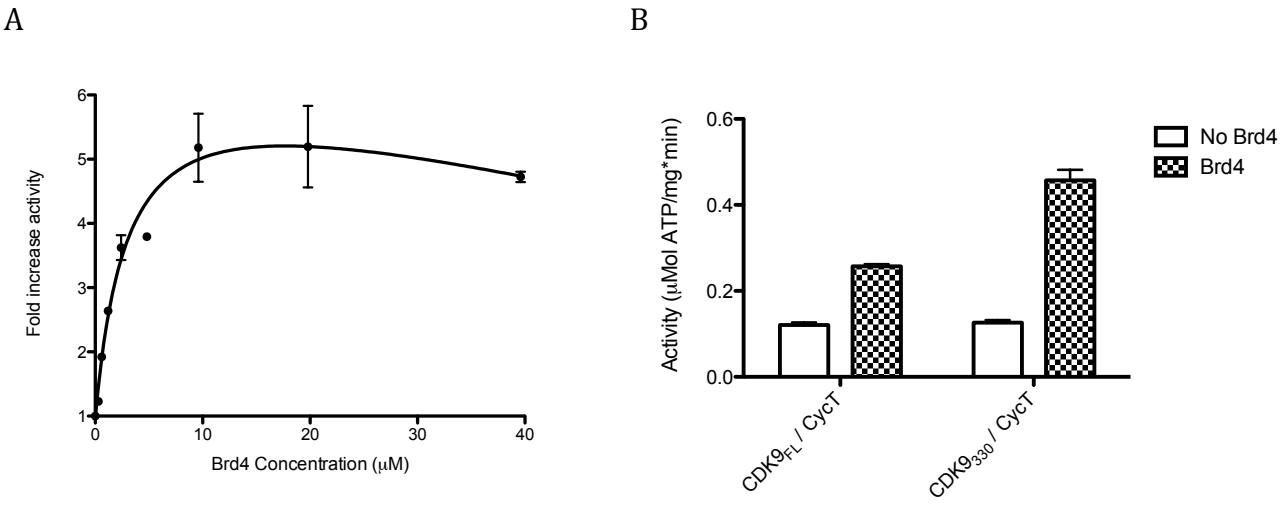


Figure 3-12. Activation of CDK9$_{330}$/cyclin T$_{WT}$ by Brd4. A: Increase in activity of CDK9$_{330}$/cyclin T$_{WT}$ is concentration dependent. B: CDK9$_{330}$/cyclin T$_{WT}$ is activated more by 25pmol Brd4$_{(1322-1362)}$ than is CDK9$_{FL}$/cyclin T$_{WT}$

This assay also provides information on the maximum fold increase that Brd4 can provide to CDK9 and can be used to provide an estimate of the dissociation constant between Brd4 and CDK9/cyclin T. CDK9$_{330}$ increases in activity by 5.8 ± 0.82 fold in the presence of Brd4$_{(1322-1362)}$, whereas CDK9$_{FL}$ increases in activity by 2.2 ± 0.26 fold. Brd4 therefore has a greater effect on activity of the shorter construct than that of the full-length kinase under standard assay conditions. The data also provides information on the K_d of Brd4 for CDK9$_{330}$ (**Table 3-4**). For the full length construct the K_d between CDK9/cyclin T and Brd4 is determined to be 1.3μM and for the CDK9 construct lacking the C-terminal tail the K_d is similar (3μM). The binding of Brd4$_{(1322-1362)}$ to P-TEFb is therefore unlikely to be through the C-terminal tail of CDK9 or to be influenced by the tail.

	K_d (μM)
CDK9$_{FL}$/cyclin T$_{WT}$	1.3 +/- 0.35
CDK9$_{330}$/cyclin T$_{WT}$	3.0 +/- 1.1

Table 3-4. Dissociation constants for Brd4 interactin with CDK9/cyclin T variants. K_d values were calculated from kinase assays.

3.3.8 The activation of P-TEFb by Brd4 is independent of the cyclin T isoform

Knowing that the CDK9 tail does not influence the binding of Brd4 to CDK9/cyclin T a similar approach was taken with cyclin T isoforms to determine if Brd4 has the same activation effect on cyclin T1 and cyclin T2. Although cyclin T1 and cyclin T2 expression levels are similar in most cell types, their expression is induced differently within macrophage cells (De Luca et al. 2001a; De Luca et al. 2001b; Liou et al. 2006). They are both also responsible for regulating a unique subset of genes (Ramakrishnan et al. 2011). It could therefore be hypothesised that the interaction of Brd4 with P-TEFb might depend on the isoform of cyclin T involved, so as to allow differential control of the transcription of the genes that characterize cell type and/or developmental stage.

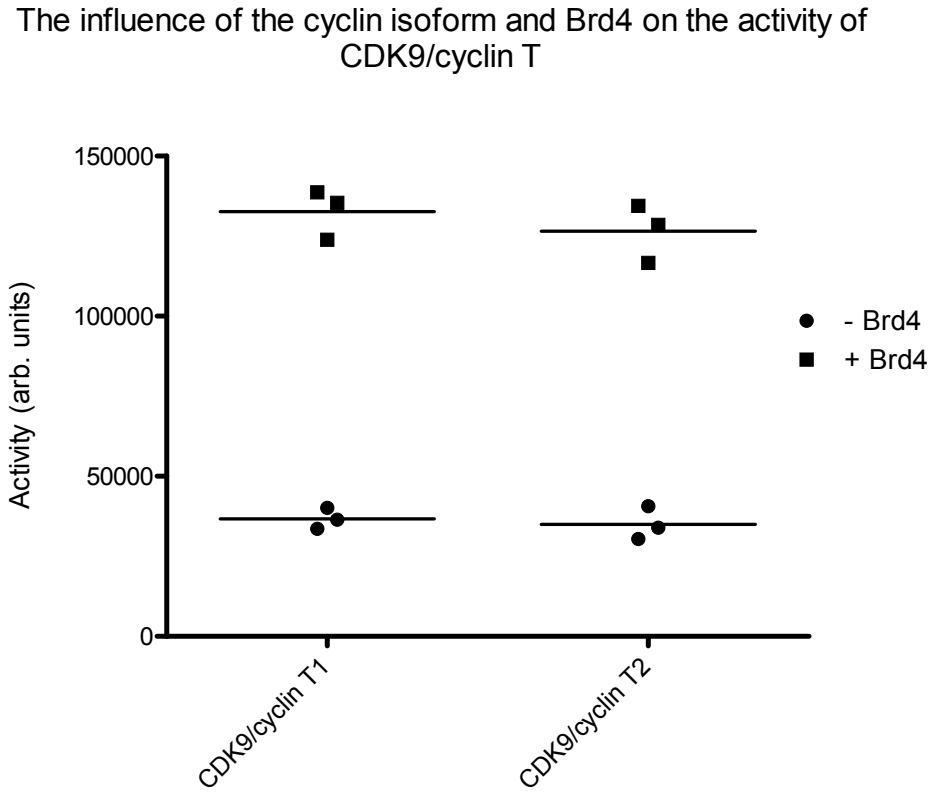


Figure 3-13. The influence of cyclin T variant on CDK9 kinase activity. Cyclin T isoforms T1 and T2 complexed to CDK9 have the same activity and are additionally activated to the same extent by Brd4$_{(1322-1362)}$,

To determine whether cyclin T1 and cyclin T2 differentially affect the activity of CDK9 on the GST-CTD substrate, the activity of $CDK9_{330}$ in the presence of each cyclin isoform was analysed. The results show little difference in CDK9 activity in the presence of the two cyclin isoforms (Figure 3-13). Furthermore, analysis of the data using a 2-way ANOVA test suggests that the activity of CDK9 is independent of the cyclin isoform present. This data shows that CDK9 activity does not depend on the specific cyclin associated with CDK9.

Secondly, Brd4 was introduced into this assay in order to ascertain whether Brd4 has a differential effect on the activity of CDK9 when complexed with different cyclin T isoforms. In this assay, the presence of Brd4 caused the expected increase in activity of CDK9/cyclin T1, and, moreover, caused an enhancement of the activity of CDK9/cyclin T2. There was no statistically significant difference in the magnitude of Brd4 enhancement of CDK9/cyclin T1 and CDK9/cyclin T2 suggesting that the activation of CDK9/cyclin T by Brd4 is independent of the isoform of cyclin T that is present. This result suggests that the effect of Brd4 on transcription is not restricted to genes that are dependent on cyclin T1, but rather applies equally to all genes activated by P-TEFb.

As Brd4 has the same effect on activation of CDK9 irrespective of the cyclin T isoform present, it follows that any interaction that might exist between Brd4 and cyclin T must be primarily located on a patch conserved between the two isoforms. The two different cyclin T constructs used share 80% identity and 88% similarity, as reported by NCBI BLAST. A combination of structural and sequence

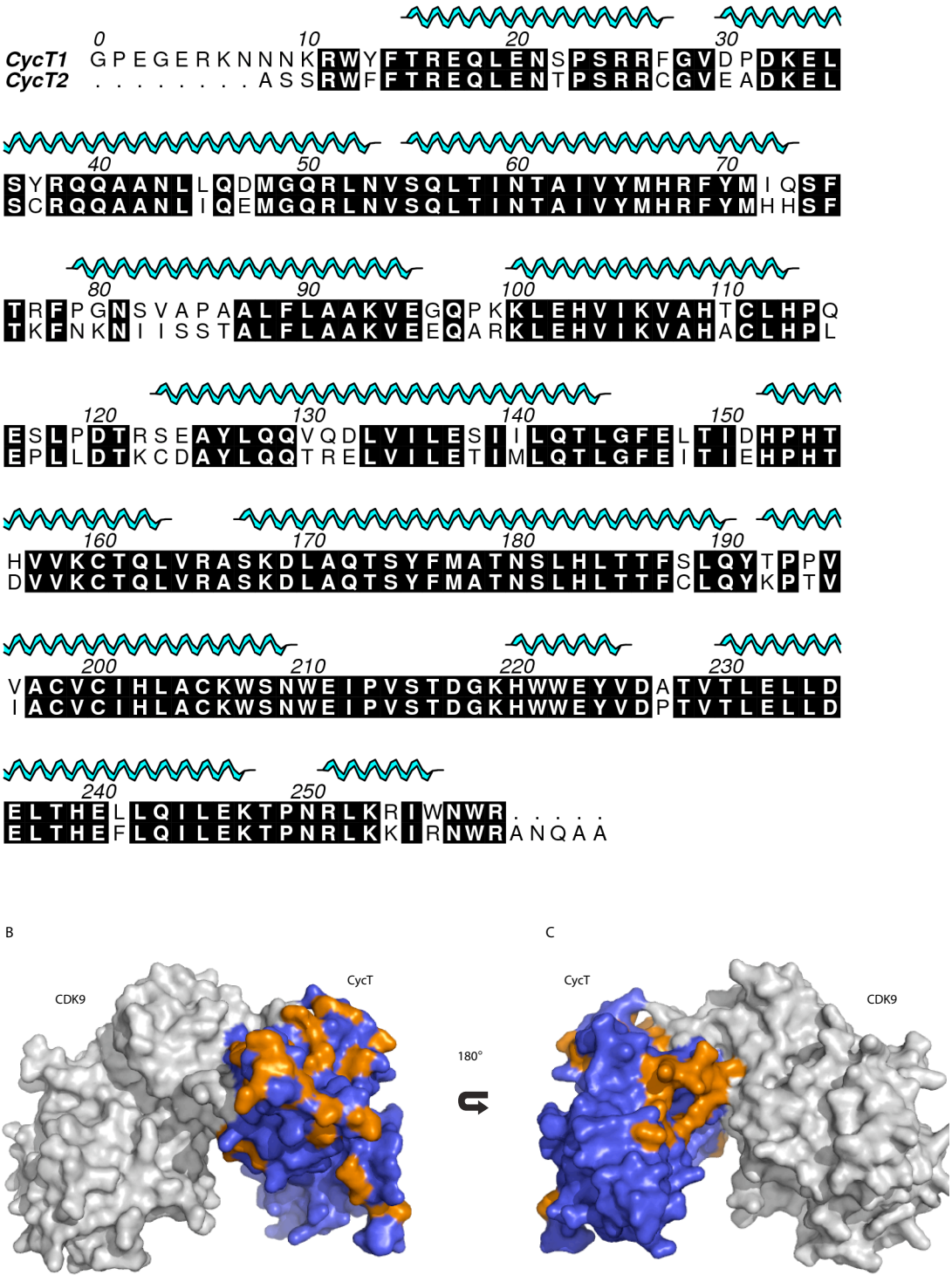


Figure 3-14. Sequence alignment (A) and structural conservation (B) of the two cyclin T constructs. (B) The surface image of CDK9/cyclin T (PDB: 3BLH) is shown. The kinase is coloured grey. Conserved cyclin residues are blue and non-conserved residues are shown in orange. (C) is rotated 180° with respect to (B).

analysis identifies that the majority of non-conserved residues are located in the N-terminus of the cyclin at the top of the CDK9/cyclin T complex (Figure 3-14). Residues that are conserved between the isoforms are located on the C-terminal cyclin box fold. Brd4 could thus be interacting in the cleft formed between the C-terminal kinase lobe and C-terminal cyclin box.

3.3.9 CDK9/cyclin T interacts with most residues across the C-terminal 40 residues of Brd4

The pull-down and kinetic analyses so far described have shown that the C-terminal 40 amino acids of Brd4 are sufficient to activate CDK9/cyclin T, and that this activation derives mostly from increasing the apparent V_{max} of the reaction. This latter result suggests that the Brd4 interaction site on CDK9/cyclin T is likely to be on the CDK, but it remains possible that some part of the interaction is with the cyclin subunit. In order to more precisely define the interaction site, NMR experiments were undertaken.

NMR provides a method that enables analysis of the structure of proteins in solution. It is also a powerful technique to study protein-protein interactions and identify interaction sites in the solution state. Furthermore, once peaks in the ^{15}N HSQC spectrum are assigned to their corresponding residues in the protein, this technique provides the resolution to distinguish individual residues affected by the interaction. More information on this technique is provided in Appendix D.

^{15}N-labelled Brd4$_{(1322-1362)}$ was expressed in *E. coli* and subsequently purified by affinity chromatography and gel filtration. The incorporation of ^{15}N was confirmed by ESI-MS. HSQC spectra of this sample produced a single set of well-dispersed peaks. The distribution of the chemical shifts (Figure 3-16) indicate that the protein has no high order structure and is therefore likely to be unstructured in solution.

Assignment of peaks to residues requires data collection over at least 48 hours under standard NMR conditions. The Brd4 protein sample degraded and aggregated over this time period. Both thermal denaturation and aggregation assays were used to identify buffer conditions that stabilize the protein and that are also compatible with NMR. Unfortunately no condition was found that sufficiently stabilized the protein for the period of time required for 3D NMR experiments. Therefore urea was added to the protein sample, fully denaturing the protein and allowing a higher concentration of stable protein to be achieved. HSQC spectra recorded in the presence and absence of urea showed similar distribution of peaks, supporting the conclusion that the protein is primarily unstructured in solution.

3.3.9.1 Assignment

Several 3D spectra were collected on ^{15}N Brd4$_{(1322-1362)}$ in the presence of 6M urea. Residues were assigned to peaks from the HSQC using data obtained from a combination of 3D TOCSY-HSQC, NOESY-HSQC and HSQC-NOESY-HSQC experiments (Figure 3-15). 39 of the 43 residues could successfully be assigned.

The two N-terminal residues, both cloning artefacts, did not contribute a peak to the HSQC and therefore were not assigned. Residues Gln1333 and Glu1334 could not unambiguously be assigned.

A B

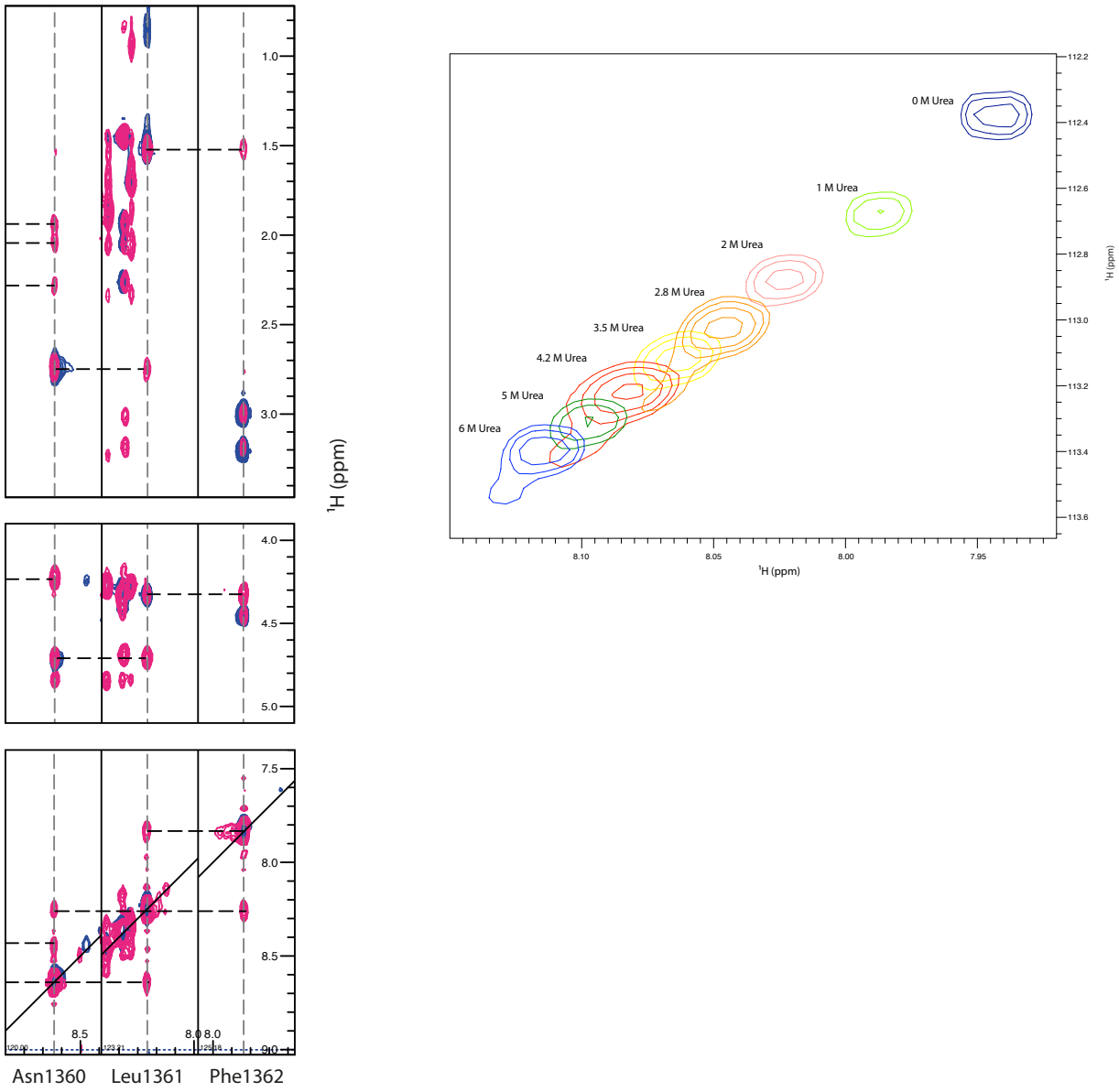


Figure 3-15. Backbone assignment of ^{15}N Brd4$_{(1322-1362)}$. (A) Strip plots for residues 1360-1362 are displayed. Peaks from the TOCSY-HSQC (pink) and NOESY-HSQC (blue) are shown. (B) Example of chemical shift changes induced by the presence of a urea gradient from 6M to 0M urea. (Residue Thr1344 is shown).

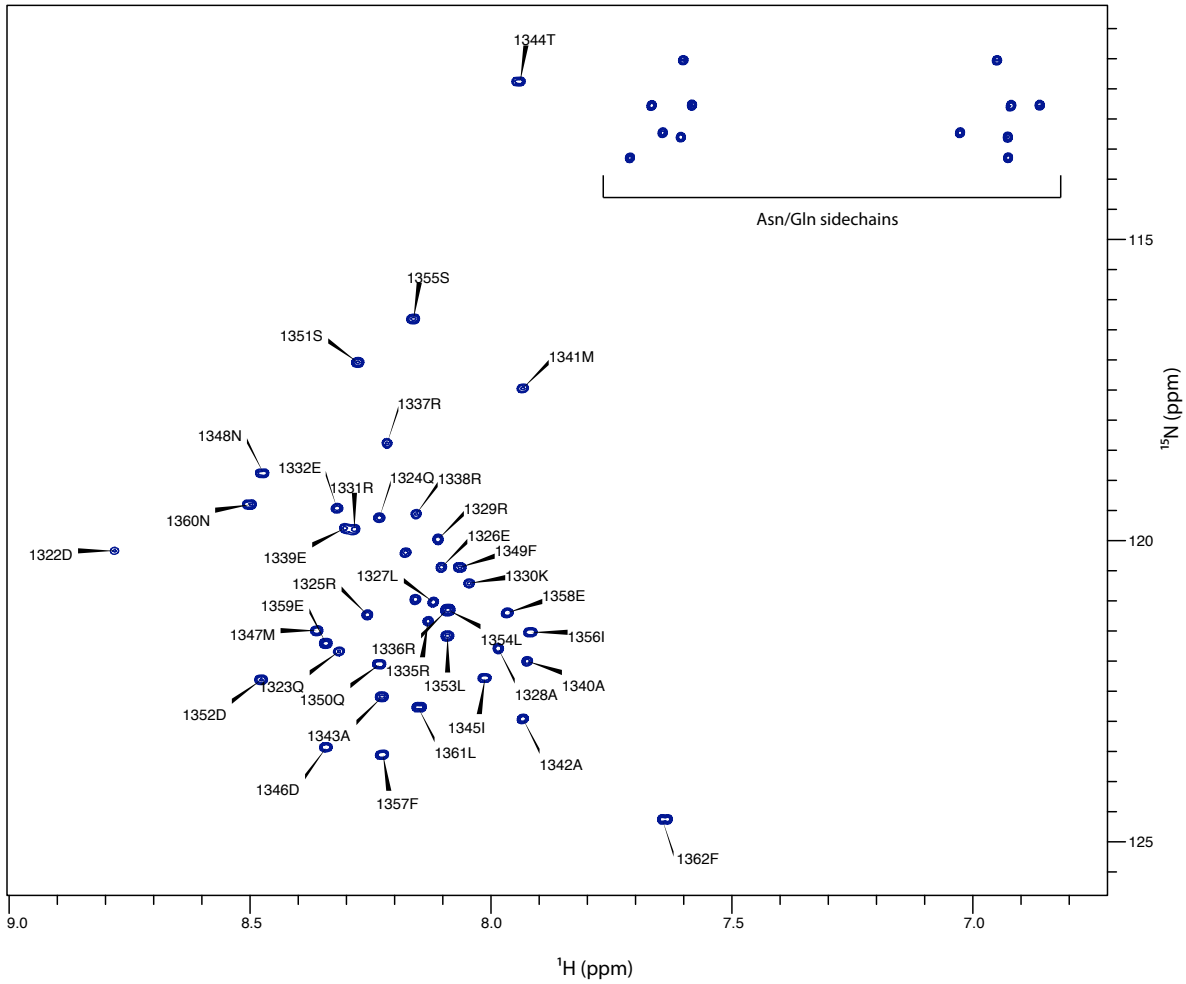


Figure 3-16. Assignment of Brd4 (1322-1362) peaks under conditions of 0M urea, pH 6.5, and at 15 °C

HSQCs were collected at a series of urea concentrations to enable the assignments to be transferred from 6M urea to 0M urea conditions (Figure 3-15 & Figure 3-16). Spectra were also collected at 20°C for analysis of particular protein-protein interactions. All peaks, with the exception of two, were unambiguously assigned under these conditions.

3.3.9.2 Brd4-CDK9/cyclin T Interaction

Subsequent to assignment, HSQCs were set up with ^{15}N Brd4$_{(1322\text{-}1362)}$ in the presence of CDK9$_{330}$/cyclin T$_{Q77R,\ E96G,\ F241L}$ (and in the absence of urea).

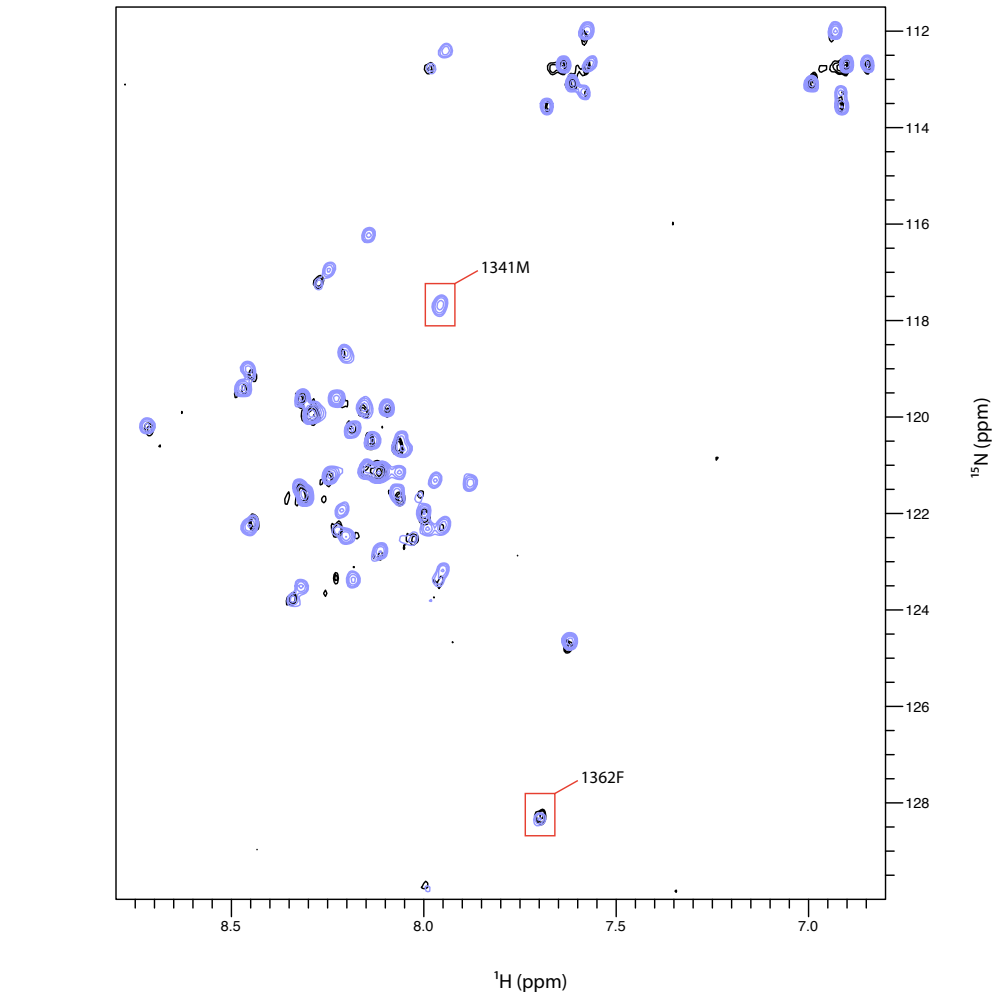
Interactions between ^{15}N labeled proteins and binding proteins show three primary alternative characteristics on binding. As binding partners are titrated in peaks observed in a protein or peptide HSQC spectra may change either in position, in intensity with the appearance of a new set of peaks, or in intrinsic width. Because peak broadening is accompanied by a decrease in height, extremely broad peaks may become undetectable.

The change in the spectral peaks is due to an altered chemical environment at the site of the backbone NH of that particular residue and is determined by the interaction kinetics. Peak shifts or broadening are thus an indication of either an altered conformation of the labeled protein at that residue or an interaction with the binding protein.

Introduction of CDK9$_{330}$/cyclin T$_{Q77R,\ E96G,\ F241L}$ into the NMR experiment drastically decreased the number of peaks visible in the spectrum (Figure 3-16). In complex, the molecular weight of CDK9/cyclin T is *circa* 70kDa, and interaction of Brd4 with proteins of this mass results in peak broadening in the NMR spectra. Certain peaks remained, although at a much decreased intensity,

allowing some distinction between more or less affected residues. Therefore regions of Brd4 that are influenced to a greater or lesser extent by CDK9/cyclin T can be identified. The N- and C- terminal residues of Brd4 retain a higher intensity than others residues, indicating that they possess more flexibility than other residues within the sequence, irrespective of whether Brd4 has bound CDK9/cyclin T.

In general, residues across the entire 40mer protein appear to be significantly involved in forming interactions with CDK9/cyclin T. This result is consistent with the results of the pull-down analysis described in Chapter 3.3.3.



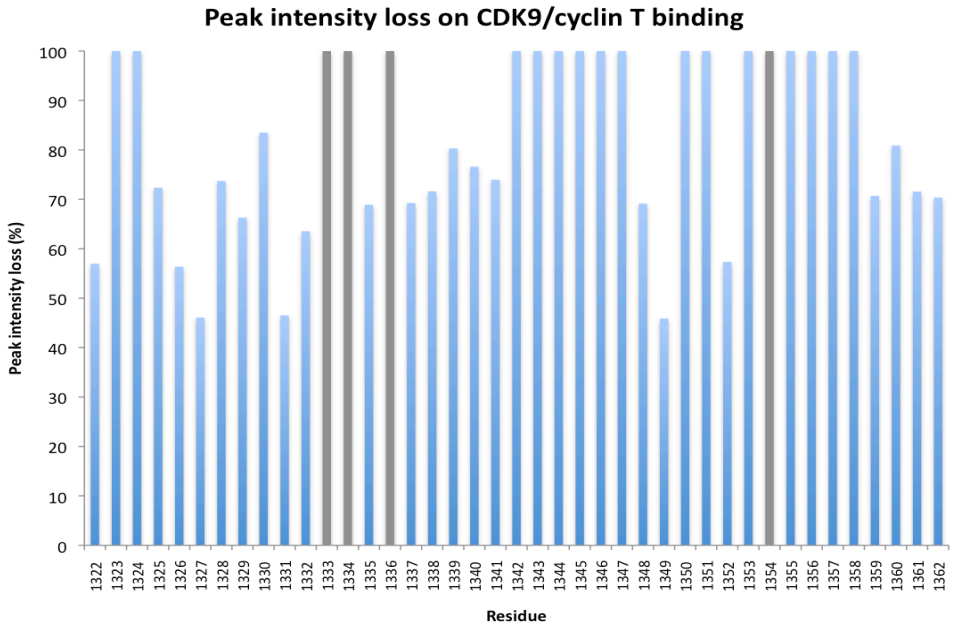


Figure 3-17. CDK9/cyclin T induces loss of intensity to Brd4 peaks. The HSQC show spectra collected with (black) and without (blue) CDK9/cyclin T. Highlighted are peaks assigned to residues 1341 (peak loses intensity) and residue 1362 (peak remains with a similar intensity). This data is shown in a graph below the spectra. Black lines indicate residues that are unassigned or that become ambiguous in the presence of CDK9/cyclin T. The x-axis represents residue numbers.

3.3.10 Cyclin T alone brings about chemical shift changes in ^{15}N labeled Brd4

The kinetic analysis indicated that Brd4 is highly likely to interact with CDK9 in order to increase the activity of the kinase. However, to confirm this, and to determine the contribution that cyclin T may make to the interaction, $CDK9_{330}$ and cyclin $T1_{WT}$ were titrated into ^{15}N-labelled Brd4. Both of these proteins, however, had limited solubility, and therefore the titrations showed little effect on the ^{15}N Brd4 spectra obtained. Cyclin T2, however, proved to have more amenable properties for the NMR conditions required. Spectra of cyclin T2 titrations were collected and used to ascertain if Brd4 binds to CDK9 and/or cyclin T.

Titrations of cyclin T2 into labeled Brd4 produced a pattern of shifts in several of the peaks in the HSQC of Brd4, whereas other peaks remained unaltered in position (Figure 3-18). Peaks that shift correspond to residues of Brd4 that undergo fast exchange in the presence of cyclin T2, i.e. Brd4 residues that, in the presence of cyclin T, experience an altered chemical environment. This environmental change may result either from proximity to cyclin T2 or from an induced conformational change. These peaks map to residues 1342-1362 at the very C-terminus of Brd4. In particular residues Ala1343, Leu1354 and Phe1357 appear to be intimately involved in the interaction of Brd4 with cyclin T2. Cyclin T2 therefore interacts with the C-terminal 22 residues of Brd4.

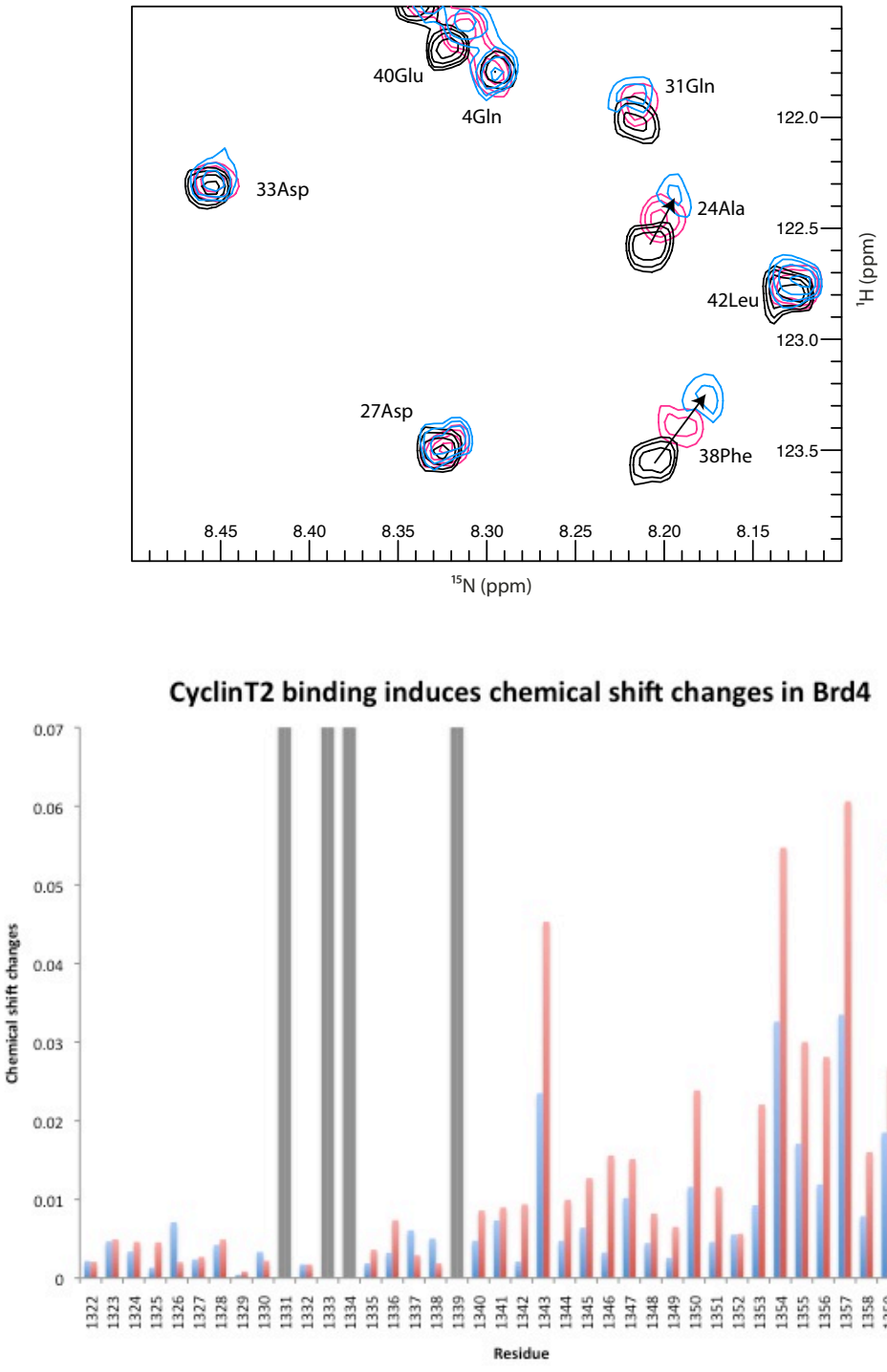


Figure 3-18. Cyclin T2 induces chemical shift changes of specific residues of Brd4. (A) HSQC showing shifts of certain peaks in the absence of cyclin T2 (black) and at different concentrations of cyclin T2 (15μM (pink) and 30μM (blue)). (B) Chemical shift changes of Brd4 residues induced by cyclin T2. Residues that are not assigned or cannot be traced through the titration are shown as grey bars.

3.3.11 H/D exchange experiments suggest Brd4 interacts with cyclin T and CDK9

Several proteins that interact with P-TEFb have been shown to interact with either CDK9 and/or cyclin T and so far I have shown that Brd4 interacts with both the CDK and cyclin. Despite many attempts, more detailed analysis of the interaction between Brd4 and CDK9/cyclin T did not prove successful using traditional structural techniques such as NMR and crystallography. I therefore analysed the interaction between Brd4 and CDK9 and cyclin T by H/D exchange mass spectrometry (HXMS). These experiments were undertaken in collaboration with ExSAR.

This technique monitors the rate of hydrogen/deuterium exchange of protein backbone NH groups. After a period of exchange the protein is transferred to conditions that prevent further exchange and is proteolytically digested into peptides. The peptides are identified by mass spectrometry and the relative abundance of protons and deuterons that are incorporated is determined. Regions of exchange can subsequently be mapped onto the structure of CDK9/cyclin T.

The exposure of the protein to deuterium is due to a combination of factors. As well as solution conditions, the solvent accessibility and the presence of hydrogen bonding affect the rate of transfer. This technique therefore indicates both protein dynamics and, in the presence of ligands, likely binding sites. For reviews of this technique see (Garcia et al. 2004; Busenlehner and Armstrong 2005). Varying the time allowed for H/D exchange and the presence and absence of Brd4 therefore allows the identification of areas within CDK9 and cyclin T that

experience a change in the rate of H/D exchange when CDK9/cyclin T binds to Brd4.

Subsequent to H/D exchange and proteolysis, mass spectrometry successfully detected high signal quality peptides covering 99% of $CDK9_{FL}$/cyclin T_{288}. Close to full coverage of both CDK9 and cyclin T resulted in mapping the H/D exchange on close to all residues (Figure 3-19). The first two residues of the peptides chosen for tracking cannot be included during the analysis due to experimental processes. The minimal peptide fragments that can be identified limit the resolution of this technique. The minimum number of residues resolved for this analysis was 4. On average, the location of binding sites can be defined within 12 residues for CDK9 and 14 residues for cyclin T. This might correspond to 3 or 4 turns of an α-helix, respectively.

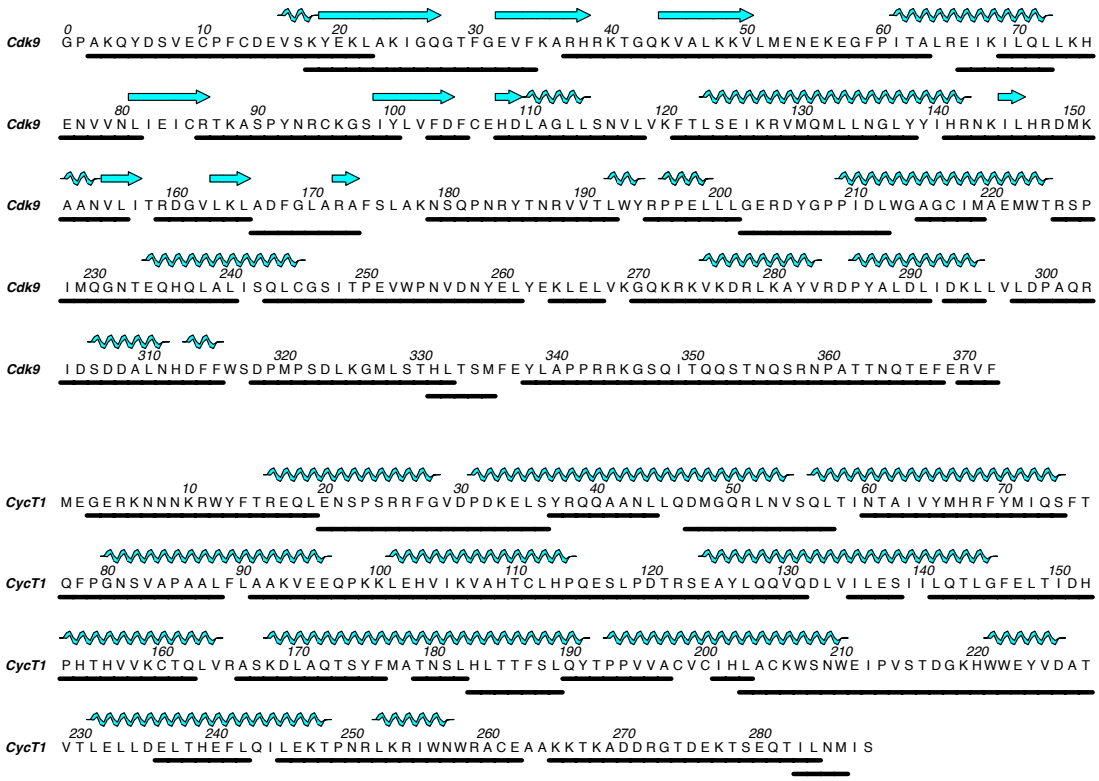


Figure 3-19. Illustration of the peptide coverage of CDK9/cyclin T for HXMS. The peptides used for analysis in H/D exchange mass spectrometry experiments are shown as black bars below the sequence.

3.3.11.1 CDK9/cyclin T plasticity

After five time points (30, 100, 300, 1000, and 3000 seconds) levels of deuteration for $CDK9_{FL}$/cyclin T_{288} were analysed (Figure 3-20). These results highlight both the inherently flexible regions of CDK9/cyclin T and regions that are relatively exposed to the solvent and are therefore able to undergo H/D exchange.

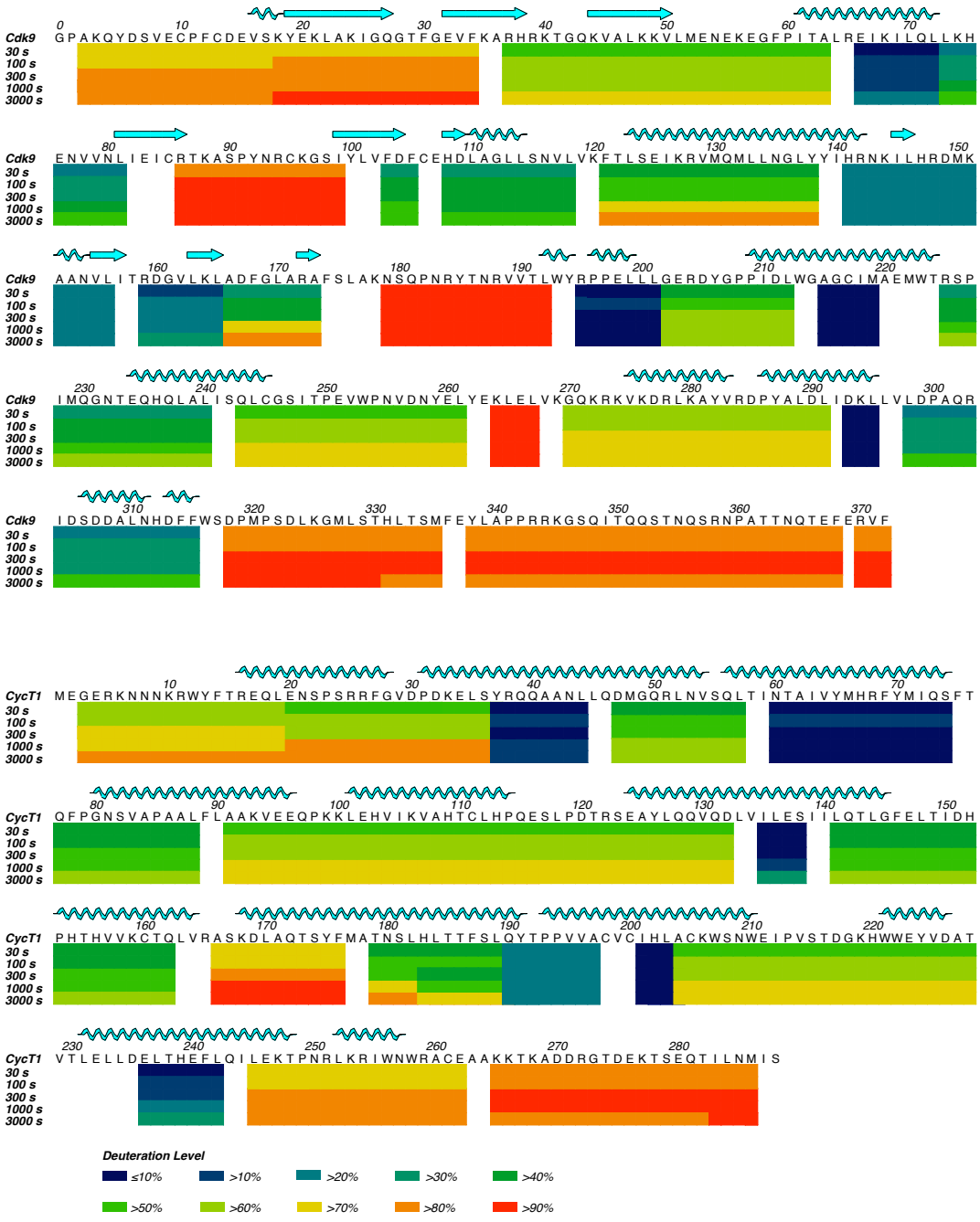


Figure 3-20. The plasticity of CDK9/cyclin T as determined by the incorporation of deuterium over time periods of 30, 100, 300, 1000, 3000s. Each block represents a peptide analysed and each row is a different time point. Dark blue indicates little incorporation of deuterium whereas red bars indicate full deuteration of peptide.

Analysis of the deuterium exchange of CDK9/cyclin T in the absence of any additional proteins indicates a general increase in deuterium incorporation over time as would be expected. This is represented in Figure 3-20 and Figure 3-21, by a transition of segment colour from dark blue, through green, to red during the time periods 30, 100, 300, 1000, 3000 seconds. The regions that are most susceptible to exchange in the first instance are primarily in CDK9 and include the N and C-termini, the activation segment, the β4-β5 loop, and residues 264-277. The majority of these sections are not seen within the crystal structure, presumably due to structural disorder. The activation segment has no secondary structure and therefore its peptide backbone forms few hydrogen bonds. It is therefore both flexible and exposed to solvent and, correspondingly, becomes rapidly deuterated.

Throughout the deuterium exchange process additional hydrogens become exposed due to conformational breathing of the protein. After 3000 seconds, the majority of the protein complex is over 50% deuterated resulting from this conformational plasticity of the complex. There are, however, certain regions, mainly α-helical elements that are located in the core of the proteins, which do not exchange hydrogens within this time period and under the conditions tested. The exchanging protons in these regions are both less accessible to solvent and often involved in a hydrogen bond network, requiring correspondingly more time to exchange.

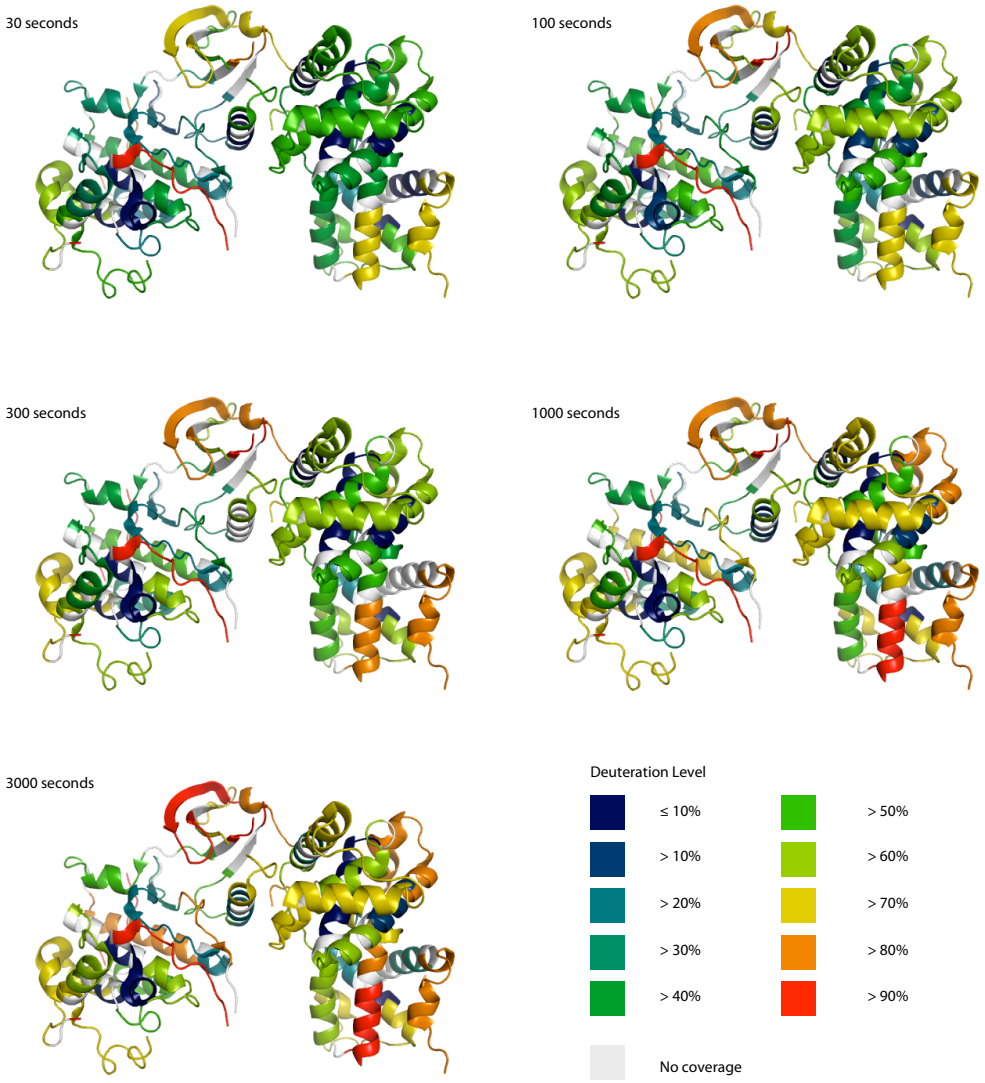


Figure 3-21. Structural analysis of the plasticity of CDK9/cyclin T. Deuterium exchange was mapped onto the CDK9/cyclin T structure (PDB: 3BLH) when measured after 30, 100, 300, 1000, 3000 seconds.

3.3.11.2 Brd4 Interaction sites on CDK9/cyclin T

H/D exchange was measured in the same manner as above in the presence of $Brd4_{(1322\text{-}1362)}$ to define the Brd4 interaction sites on CDK9/cyclin T. The rate of incorporation of deuterium in certain stretches of CDK9/cyclin T was found to vary, depending on whether CDK9/cyclin T was complexed with Brd4 (Figure

3-22). Mapping this variation onto the CDK9/cyclin T structure provides key information on the interaction between Brd4 and CDK9/cyclin T (Figure 3-23). Differences in the extent of peptide deuteration were observed at 30, 100, 300, 1000 and 3000 seconds. As expected, the differences observed at each of these timepoints are correlated. In order to highlight changes in the rate of H/D exchange caused by Brd4 binding, therefore, two values were extracted for each peptide fragment and used for analysis: firstly the fractional difference in H/D exchange apparent at 30 seconds, and secondly the maximum observed fractional difference for each peptide over the experiment. Restriction of H/D exchange observed on CDK9/cyclin T due to Brd4 is proposed to locate the Brd4 interaction sites on CDK9/cyclin T. This analysis suggests that the binding of Brd4 by CDK9/cyclin T is mediated by two distinct subsites.

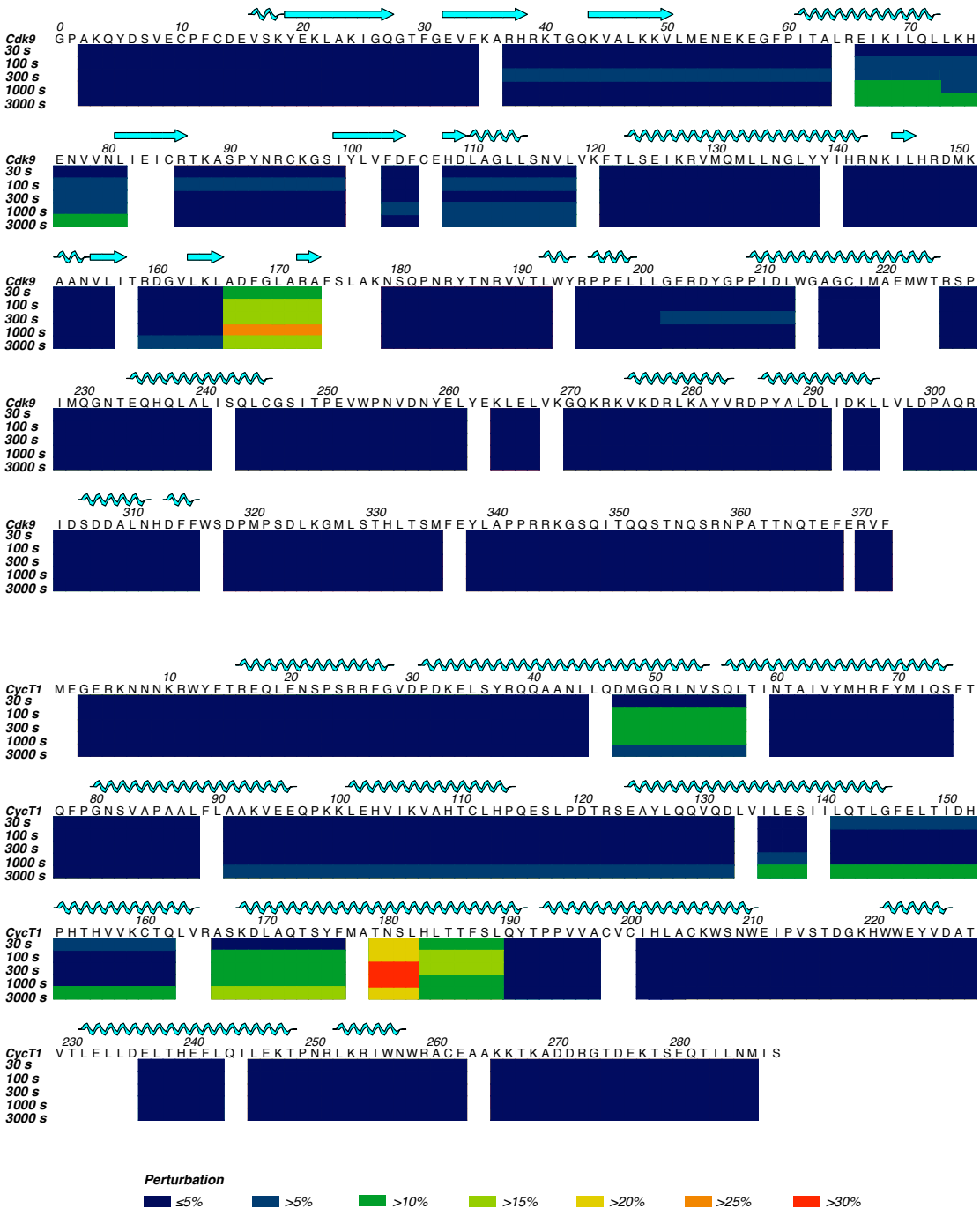


Figure 3-22. Difference in deuteration levels of CDK9/cyclin T upon binding of Brd4. The representation is as in Figure 3-19.

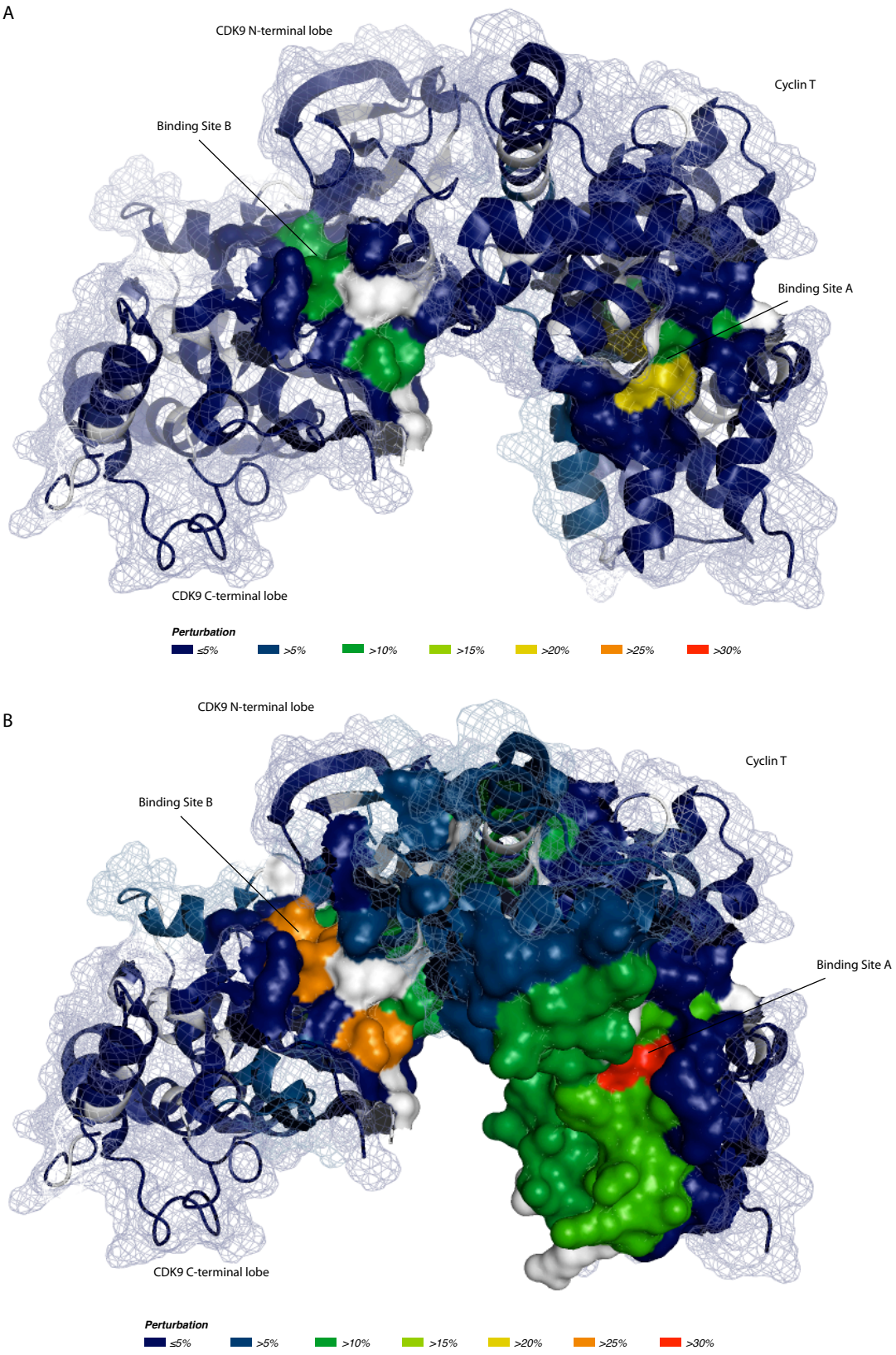


Figure 3-23. The difference in deuterium levels upon Brd4 binding mapped onto the CDK9/cyclin T structure. The difference is mapped after 30s (A) and 3000s (B). CDK9/cyclin T is represented in cartoon form. The surface is represented as a mesh in all regions unaltered by Brd4 binding. The surface representation is shown for residues effected by Brd4 and those within 5Å of effected residues.

The first is on the front of the cyclin between residues 179-189. These residues, which comprise two of the peptide fragments identified in the raw MS data (179-182 and 183-189) are on the H2' α-helix and form a distinct binding pocket (Figure 3-24). This pocket is lined by residues Arg251, Arg254, Phe176, Asn180, Thr179, Asn60, Gln56, Gln46, Asn43, His183 and Tyr175. These residues contribute hydrogen bonding capability to a site often occupied by Tris in CDK9/cyclin T crystal structures. Overall the pocket is relatively polar and has a positively charged patch provided by the two arginine sidechains at the base.

In addition this pocket is shown to interact with HIV-Tat (Tahirov et al. 2010). In the CDK9/cyclin T/Tat crystal structure Tat residues 34-41 adopt an α-helix at this pocket. In particular residues Gln35 (Tat) and Asn180 (cyclin T) form a favourable hydrogen bond. The overlap between Tat and this Brd4 binding pocket suggest a direct competition between Tat and Brd4 on cyclin T consistent with previous reports (Jang et al. 2005; Yang et al. 2005; Bisgrove et al. 2007).

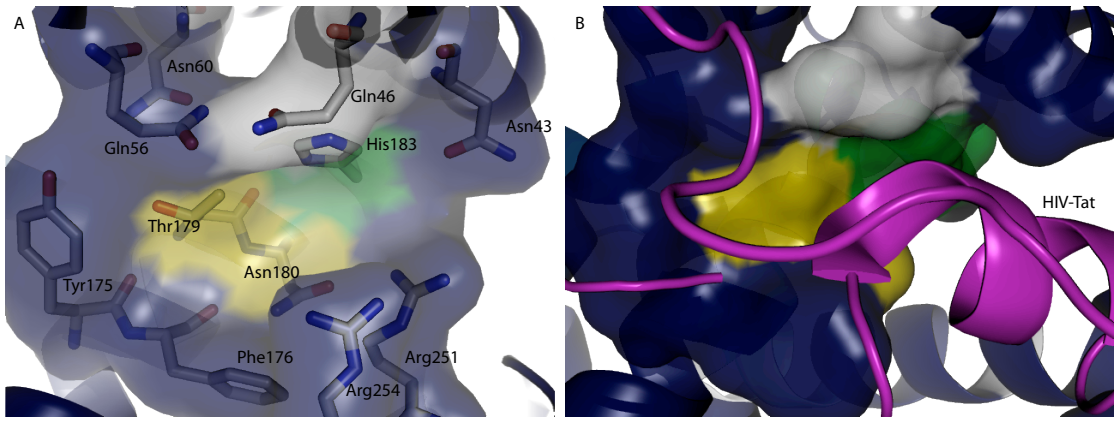


Figure 3-24. Brd4-binding pocket on cyclin T. Residues lining the pocket are shown in stick representation in (A). The interaction of HIV-Tat (magenta) with cyclin T at the same pocket is shown in (B). The surface of cyclin T is represented as for Figure 3-22.

The second interaction site is at residues 167-174 of CDK9. These residues include the highly conserved kinase DFG motif at the start of the activation segment and are located in the ATP-binding site, close to the C-helix. Correct orientation of the DFG residues in the phosphorylated, cyclin-bound active conformation of the kinase aligns the ATP for phosphotransfer to the protein substrate.

More extended regions affected by Brd4 are located around the two interaction sites identified above and the groove between the CDK9 C-terminal lobe and cyclin T. These regions extend through the groove to the back of the CDK and cyclin. Helices H1, H5, H2' and H3' in cyclin T were protected by Brd4 from H/D exchange. Furthermore there was reduced H/D exchange in the C-helix and at the back of the ATP binding pocket in CDK9. These results are consistent with the kinetic and NMR data showing that Brd4 interacts with both CDK9 and cyclin T, and that, moreover, the binding of Brd4 has an impact on the residues that influence the rate of the kinase reaction.

3.3.12 Proposed model of the Brd4/cyclin T complex

The combination of information determined within this thesis, and knowledge from previously determined crystal structures allow us to suggest a model for the interaction of Brd4 with cyclin T. The NMR and H/D exchange data presented above identify two Brd4 interactions subsites within CDK9/cyclin T and, moreover, define the orientation in which Brd4 binds between these two sites.

The Brd4-binding site on cyclin T overlaps with the site at which an α-helical region of the HIV Tat protein binds to cyclin T in the CDK9/cyclin T/Tat complex. As the C-terminus of Brd4 adopts an α-helical conformation on binding to another of its binding partners, namely HPV-16 E2, a model is suggested in which this same C-terminal helix interacts with cyclin T at a similar position to that at which the Tat helix binds. By this model, the very C-terminal residues of Brd4 are proposed to interact with the cyclin T, following a path similar to that of Tat, towards CDK9, where the more N-terminal residues of $Brd4_{(1322\text{-}1362)}$ are proposed to interact at the second interaction site.

Sequence analysis of the helical residues of Tat and Brd4 suggest two different helical registers in which Brd4 may be aligned to Tat in order to generate an atomic model of the complex: these correspond to aligning Brd4 residues 1350-1357 (Model 1) or Brd4 residues 1346-1353 (Model 2) on residues 35-42 of Tat (Figure 3-25). Applying these alignments to the atomic models of Brd4 (from PDB 2NNU) and Tat (as observed in PDB file 3MIA) produce two different atomic models for the Brd4/cyclin T interaction. No remodelling of the side chains was done in this analysis. Although some sidechain rearrangement would be required to minimize steric clashes generated by this relatively simple superposition, the data available was not considered sufficient to make possible such a refinement of the simple models described here. Both models show several strengths that support the cyclin T pocket as a possible interaction site of Brd4.

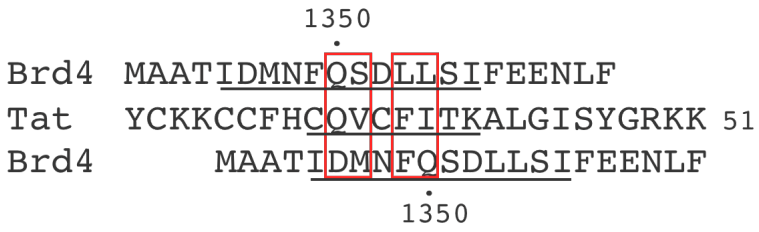


Figure 3-25. Sequence-based alignment of Brd4 and HIV-Tat. The two alignments provide the basis for the models of Brd4 bound to cyclin T.

The two models of Brd4 overlay well with the helical turns of HIV-Tat (Figure 3-26). As the Brd4 helix is longer than the Tat helix, clashes arise in both models of Brd4. In Model 1 this results in a clash at the N-terminus of Brd4: while this may indicate an error in the model, it is also possible that the Brd4 helix, which we believe to form only upon binding to E2 or cyclin T, is able to melt to prevent this. In Model 2 it is the C-terminus of Brd4 that clashes with the C-terminal helix of the cyclin T construct. This cyclin T helix has been seen to adopt different conformations in different crystal structures of CDK9/cyclin T. Accordingly, either the C-terminus of the Brd4 helix may adopt an α-helical conformation or the C-terminal cyclin T helix may adopt a different conformation if this model accurately reflects the binding mode of Brd4 on CDK9/cyclin T.

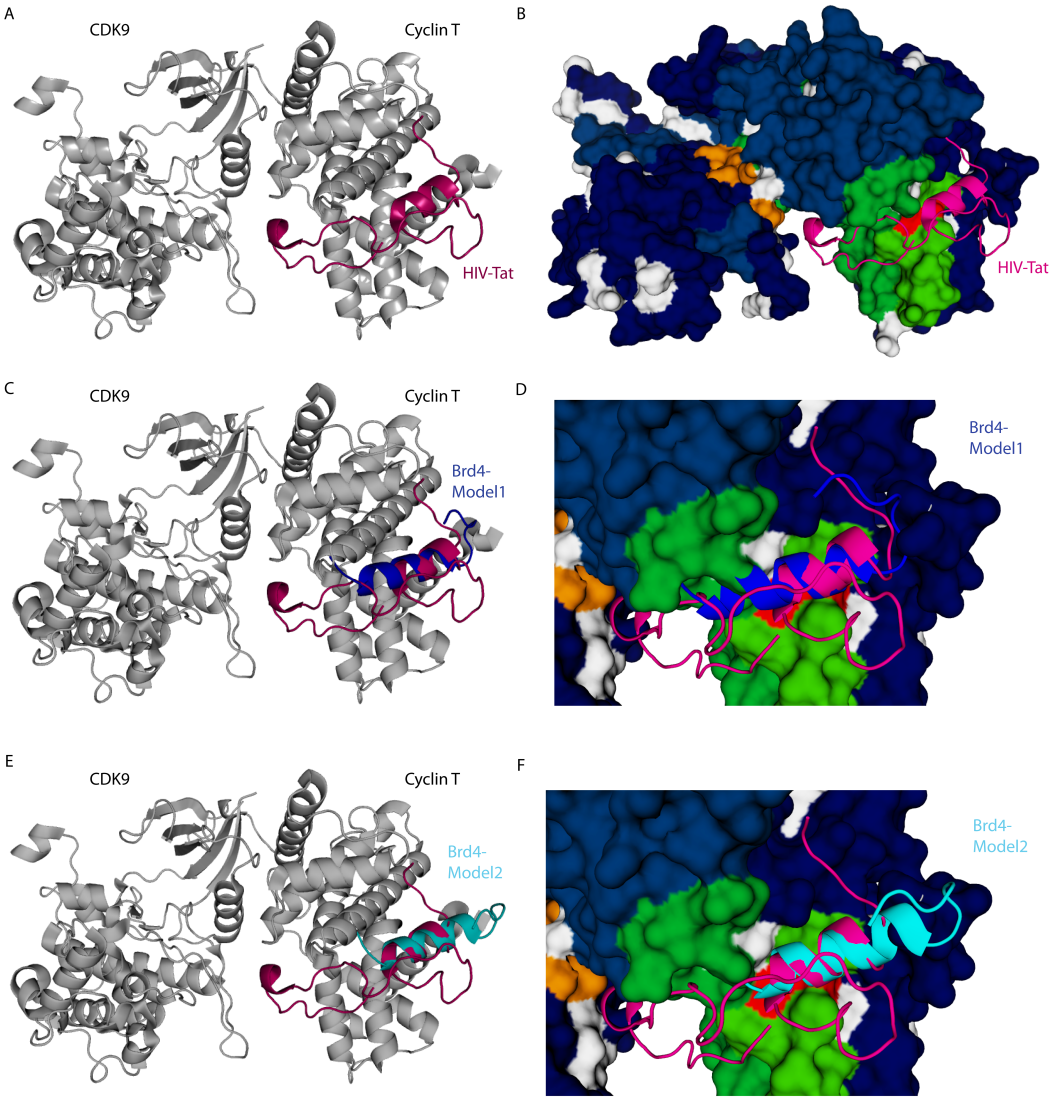


Figure 3-26. Interaction of Tat with CDK9/cyclin T and proposed models of Brd4 C-terminus interaction with cyclin T. (A) and (B) illustrate the interaction of HIV-Tat (magenta) with CDK9/cyclin T. (C) and (D) illustrate the Model 1 of Brd4 interacting with cyclin T whilst (E) and (F) illustrate Model 2. Model 1 is shown in blue and Model 2 is shown in cyan. In (A), (C) and (E) CDK9/cyclin T is shown in cartoon format. (B), (D) and (F) show the surface of CDK9/cyclin T with the H/D exchange analysis depicted with the same colour scheme as previously. The PDB used for these figures was 3MIA.

In the alignment used to generate Model 1, Gln1350 (Brd4) is positioned at the site of Gln35 (Tat) (Figure 3-27). A hydrogen-bonding interaction observed between Gln35 (Tat) and Asp180 (cyclin T) would therefore be maintained by Gln1350 of Brd4. The remaining residues on the hydrophobic face of the Tat helix that interact with cyclin T are replaced by hydrophobic residues in Brd4.

These are however less bulky in Brd4 and would likely result in reduced surface contacts.

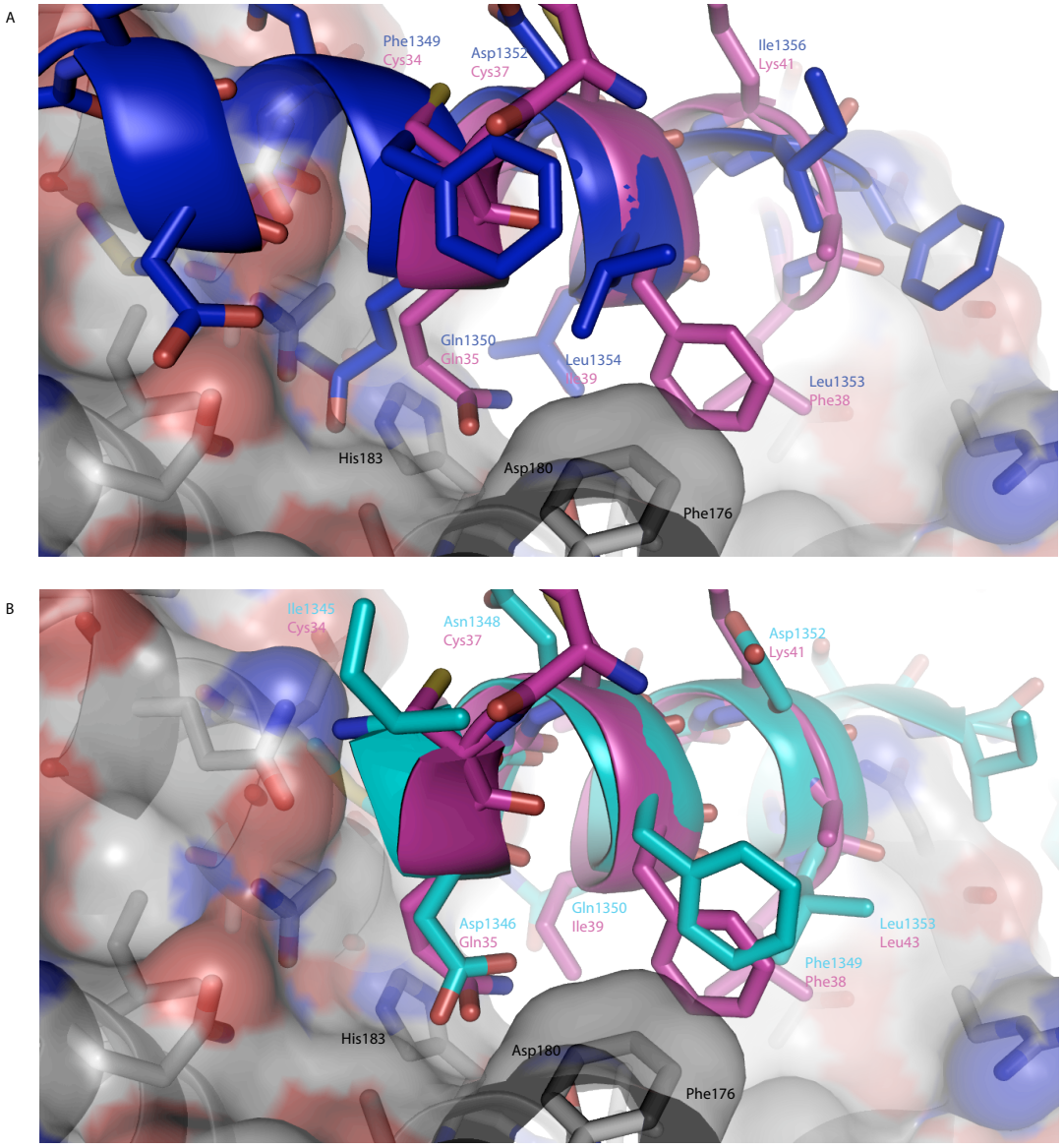


Figure 3-27. The interaction sites of cyclin T and Brd4 for the two models presented. (A) Model 1 of Brd4 (blue) and (B) model 2 of Brd4 (cyan) are shown with HIV-Tat (pink) for comparison. The surface of cyclin T is shown with certain residues depicted as sticks. PDB:3MIA.

The second model shows Gln1350 of Brd4 in close proximity to His183 of cyclin T. This histidine is located deep within the binding pocket and the interaction between this Gln and His would favour this α-helical alignment. The rotameric conformation that Met1347 adopts in the E2 structure sterically

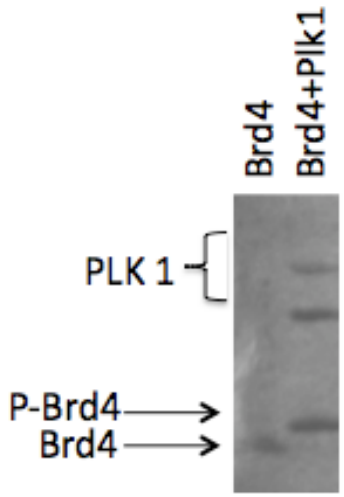
clashes with cyclin T residues if simply superimposed. This may be avoided either by a melting of the helix before this residue or by a rearrangement of a combination of local side chains. The phenylalanine position is maintained between Brd4 and Tat in this arrangement and would therefore likely bind, providing a similar contribution to the strength of the interaction.

3.3.13 Phosphorylation of Brd4 in the C-terminal region inhibits its capacity to activate P-TEFb

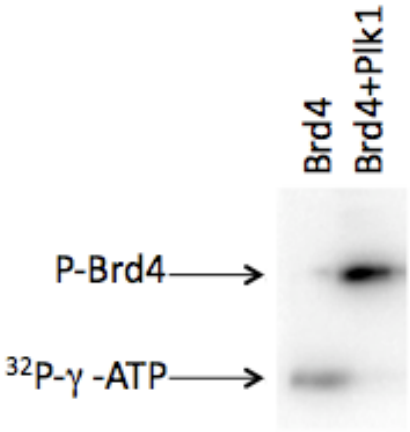
The C-terminal region of Brd4 (residues 1224-1362) can be phosphorylated by Plk1 *in vitro* (Wang et al. 2009). Knowing that the 40mer construct that interacts with CDK9/cyclin T contains one threonine and two serine residues, I hypothesized that phosphorylation may occur at these sites and provide an additional layer of regulation to transcription of Brd4-dependent genes.

In vitro, phosphorylation by Plk1 on the $Brd4_{(1279\text{-}1362)}$ could be detected by a band shift and by detection of ^{32}P incorporation from ^{32}P-γ-ATP (with the same approach as used for the CDK9 kinase assays) (Figure 3-28). Furthermore, ESI-MS identified a change in mass between unphosphorylated and phosphorylated $Brd4_{(1322\text{-}1362)}$ equivalent to that which would be expected by addition of 3 phosphoryl groups. It can therefore be concluded that *in vitro* all serine and threonine residues within the C-terminal 40 residues of Brd4 can be phosphorylated by Plk1.

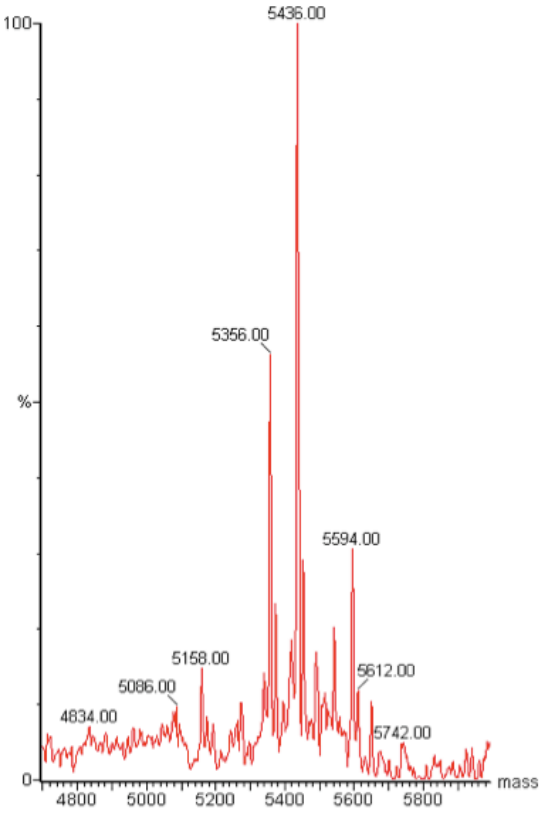
A



B



C



D

1322 1362

dqqrelark reqerrrrea maaTidmnfq SdllSifeen lf

Figure 3-28. In vitro phosphroylation of Brd4 CTR. Incorporation of phosphate was detected by band shifts (A) and ^{32}P incorporation (B) into Brd4$_{(1277-1362)}$. (C) Mass spec data of phosphorylated Brd4$_{(1322-1362)}$ shows an increase in mass consistent with incorporation of 3 phosphates. (D) The sequence of Brd4$_{(1322-1362)}$ with residues that are phosphorylated shown in red text.

In order to determine if the phosphorylations affected the ability of Brd4 to activate CDK9/cyclin T, kinase assays were carried out, after purifying the Brd4 away from contaminating Plk1 and ATP. Kinase assays of $CDK9_{FL}$/cyclin T_{WT} in the presence of phosphorylated $Brd4_{(1322\text{-}1362)}$ present showed a significant difference in activation of the kinase activity to that of unphosphorylated Brd4 (Figure 3-29). This may result either from the phosphorylated Brd4 being unable to bind to CDK9/cyclin T or from the phosphorylated Brd4 binding to CDK9/cyclin T but failing to activate the kinase in the same manner as unphosphorylated Brd4. At higher concentrations the kinase activity of CDK9/cyclin T increased resembling more closely that of unphosphorylated Brd4 activation. This is likely due to some unphosphorylated Brd4 being present within the sample, and therefore contributing to the increased activity at higher Brd4 concentrations.

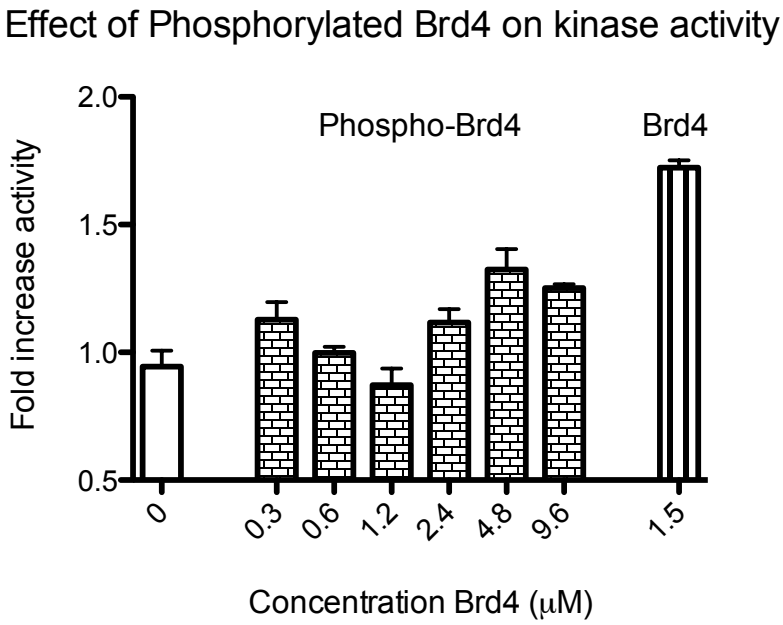


Figure 3-29. Phosphorylated $Brd4_{(1322\text{-}1362)}$ does not activate CDK9/cyclin T to the same extent as unphosphorylated Brd4. Analysis of fold increase in activity of $CDK9_{FL}$/cyclin T_{WT} in the presence of increasing concentration of Brd4 proteins. Experiments were repeated 3 times and the error bars represent standard errors.

3.4 Discussion

3.4.1 Minimal CDK9/cyclin T interacting fragment of Brd4

To date several interaction sites between CDK9/cyclin T and Brd4 have been mapped. The bromodomains of Brd4 were initially identified as interacting with cyclin T residues (426-516) (Jang et al. 2005). Subsequently deletion analysis identified independent interactions between the CTR (specifically the C-terminal 34 residues) of Brd4 and both CDK9 and cyclin T (Bisgrove et al. 2007). Since then further studies have confirmed the interaction of Brd4 with cyclin T residues 426-516 (Cho et al. 2007) and identified that bromodomain 2 can bind a peptide of cyclin T acetylated at residue 390 (Vollmuth et al. 2009).

In agreement with these previously published results, the bromodomains were unable to interact with $CDK9_{330}$ and the cyclin boxes of cyclin T. For the first time, I have shown that the final 40 amino acids of Brd4 are sufficient to interact with a minimal active form of CDK9/cyclin T. This was confirmed by results from pull-down analysis, analytical gel filtration and NMR. Furthermore NMR was able to distinguish that the very C-terminal 1342-1362 residues of Brd4 interact with cyclin T and, by inference, residues from the region 1322-1341 interact with the CDK.

3.4.2 Brd4 Activates CDK9/cyclin T

Brd4 is proposed to activate the transcription of over 10% of genes transcribed in G1 (Mochizuki et al. 2008) and yet Brd4 has been suggested to reduce the activity of CDK9/cyclin T on GST-CTD (Zhou et al. 2009b). Here I show that, in addition to its role in P-TEFb recruitment to the transcription start sites, the C-terminal residues of Brd4 can increase the kinetic activity of CDK9 towards the Pol II CTD substrate.

My results show that Brd4 acts as an activator or, more accurately, a moderator of CDK9/cyclin T. The general kinetic scheme for a single-substrate reaction with an activator is shown in Figure 3-30 (where R is the reactants, E is the enzyme, A is the activator or inhibitor, and P represents the products). A reaction mechanism that includes a moderator has two kinetic constants describing irreversible product formation. Studies presented above provide information of the activity of Brd4 only under standard kinase activity conditions. Full kinetic analysis was not undertaken due to the complexity of such an analysis, particularly for two-substrate reactions, and the difficulty of preparing the protein substrate, GST-CTD, and Brd4 at sufficiently high concentrations.

A: Inhibitor scheme

$$\begin{array}{ccccc} R+E+A & \rightleftharpoons & RE+A & \xrightarrow{K_1} & E+A+P \\ \downarrow\uparrow & & \downarrow\uparrow & & \\ R+EA & \rightleftharpoons & REA & & \end{array}$$

B: Modifier/Activator scheme

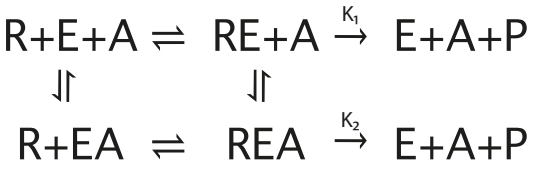


Figure 3-30. Kinetic mechanisms for reactions in the presence of inhibitors or modifiers.

The effect of Brd4 on the kinetic constants of CDK9/cyclin T indicated that Brd4 influences the phosphotransfer reaction by CDK9/cyclin T. Additional experiments demonstrated that Brd4 is also able to activate $CDK9_{330}$/cyclin T, a form of the enzyme in which substrates are bound in a random order. The activation by Brd4 is therefore shown to be independent of the sequential ordered kinetic mechanism imposed by the CDK C-terminal tail (Baumli et al, Manuscript in preparation).

The identification of a Brd4 binding subsite comprising the activation segment of CDK9 supports one mechanism by which Brd4 might activate CDK9. The binding of Brd4 decreases H/D exchange at the C-helix, either by interacting with it or by reducing its flexibility. Both the activation segment and the C-helix are involved in catalysing phosphotransfer. The residues affected in the activation segment include the DFG motif. Within this motif the aspartate is known to chelate the magnesium ions in the active site. This contributes to the correct orientation of the phosphate groups of ATP for successful phosphotransfer of the γ-phosphate to the substrate peptide (Hanks and Hunter 1995), and stabilizes the build up of negative charge that accompanies formation of a phosphoryl group. Glu66 is located within the C-helix. This residue, when in position for catalysis, hydrogen

bonds with Lys48, which in turn positions the α- and β- phosphates of ATP. Even in the absence of Brd4 the residues involved in phosphotransfer are correctly aligned. However, Brd4 was shown by kinase experiments to increase that V_{max} with respect to both ATP and the CTD. Brd4 may achieve this by stabilizing the residues involved in phosphotransfer, increasing the probability of productive transit along the reaction coordinate, whilst still allowing sufficient flexibility in the kinase to enable substrate recruitment, phosphotransfer and product release.

Several other studies involving CDK9 and other CDKs have shown an increase in activity in the presence of a third binding partner. However, no study has analysed the kinetics to the same extent and observed the same phenomenon as seen here. The first structurally described moderator of CDK9/cyclin T, HIV-Tat, binds to CDK9/cyclin T and alters its substrate specificity to phosphorylate Ser5 of the RNA polymerase CTD (Garber et al. 2000). This does not appear to be the case for CDK9/cyclin T in the presence of Brd4. A change in the specificity to allow phosphorylation of a wider range of potential sites in the GST-CTD model substrate would be expected to result in an apparent increase in substrate concentration and therefore a lowered apparent K_m towards the CTD in the presence of Brd4. The data presented shows only an increase in V_{max} towards the CTD, with an unmodified apparent K_m. Hence, unlike Tat, Brd4 probably does not appear to alter the substrate specificity of CDK9/cyclin T.

CDK7/cyclin H activity is increased by several of its binding partners. CIITA, a co-activator of MHCII gene expression, was shown to both recruit CDK7 to specific gene promoters and to enhance Ser5 phosphorylation by CDK7 (Spilianakis et al.

2003), closely paralleling the effects of Brd4 on CDK9/cyclin T. CDK7/cyclin H is also stimulated in the presence of MAT1, a partner that increases CDK7 activity against both CDK2, its cell-cycle target, and a construct containing 4 CTD consensus heptad repeats (Yankulov and Bentley 1997; Busso et al. 2000), a model peptide related to its transcription target. The magnitude of this activation, 2-3 fold, is of a similar order of magnitude to that observed for CDK9 in the presence of Brd4. No kinetic studies have been undertaken to determine the details of activation by MAT1 or CIITA and it is unclear from the published results how these increase kinase activity. In the case of MAT1, it is also unclear whether the kinase specificity is altered. This modulation of the activity of transcriptional kinases, achieved by incorporating different interaction partners into multiprotein CDK complexes, may prove to be a general mechanism by which transcription is controlled.

Whether an increase in activity of CDK9/cyclin T, effected by Brd4, has an influence on transcription within the cell is a crucial question that remains to be answered. There are, however, several mechanisms by which activation of CDK9/cyclin T could be expected to influence transcription, a few of which are suggested below.

The activation of CDK9/cyclin T by Brd4 was only studied here using a CTD model substrate. However, the increase in activity was shown to be a V_{max} effect, rather than a K_m effect, so that it might be predicted that this effect will be reproduced with NELF and DSIF as targets. Increasing the activity of CDK9/cyclin T towards these substrates would enable a more rapid transition

from paused to elongating Pol II in response to signaling cascades that recruit Brd4 to promoter regions.

Brd4 is both located at the promoter and coding regions of a gene (Cho et al. 2010). This variable distribution of Brd4 localisation has the potential to alter the Ser2 phosphorylation distribution on the CTD at different locations within a gene-coding region. Different Ser2 phosphorylation patterns on the CTD provide alternative combinations of binding sites for transcription factors. These alternatives may provide an additional mechanism to achieve the timely recruitment of specific factors to the CTD.

3.4.3 Cyclin T variants

CDK9 is primarily found complexed to cyclin T1 (80%), whilst cyclin T2a (10%) and cyclin T2b (10%) bind the remaining CDK9 (De Luca et al. 2003). Cyclin T1 and cyclin T2 share similar expression profiles, both varying expression in different cell types (De Luca et al. 2001a; De Luca et al. 2001b). However, they have certain differing properties. Cyclin T2a and cyclin T2b are unable to interact with Tat and TAR whereas human cyclin T1 interacts and promotes Tat transactivation (Wimmer et al. 1999). Cyclin T2a and cyclin T2b are splice variants and only differ at their very C-termini. The analysis undertaken here is of the cyclin boxes of cyclin T2, which are identical between cyclin T2a and cyclin T2b. Cyclin T1 and cyclin T2 also are each individually responsible for the upregulation of a subset of genes (Ramakrishnan et al. 2011) .

Brd4 increases the activity of CDK9/cyclin T1 and CDK9/cyclin T2 equally. Its activation of P-TEFb is therefore independent of the cyclin isoform which is complexed to CDK9. Correspondingly, during transcription of genes, Brd4 is predicted to show the same effect on transcription irrespective of the cyclin T isoform within the P-TEFb compex. It is therefore suggested not to be a factor in determining the alternative gene expression between cyclin T isoforms.

3.4.4 Brd4 binding sites on CDK9/cyclin T

Analysis of the interaction by H/D exchange mass spectrometry has identified two primary sites of interaction of Brd4 on CDK9/cyclin T, one on the cyclin and the other on the CDK. NMR data shows that the extreme C-terminal part of Brd4 (residues 1342-1362) binds to the cyclin and therefore, by inference, that the N-terminus of the 40mer binds to the CDK. This develops on the pull-down analysis of Bisgrove *et al* that showed that Brd4 interacts at independent sites on both CDK9 and cyclin T (Bisgrove et al. 2007).

The residues of the binding site on CDK9 include the activation segment. Previously, mutations of CDK9 residue Ser175 were identified to disrupt P-TEFb-Brd4 interactions (Yang et al. 2005). Although this residue was unfortunately not covered by the mass spectrometry analysis an adjacent segment of residues (residues 166-173) was. Notably, that segment was highlighted as being better protected from H/D exchange by the presence of Brd4 in agreement with previously published results. Ser175 has been identified as a residue that controls CDK9 activity depending on its phosphorylation state. However,

conflicting reports show both upregulation and downregulation of the kinase by phosphomimetic mutation (Chen et al. 2004; Ammosova et al. 2011). The H/D exchange experiments described here suggest that this residue may be important in the interaction of Brd4 and could replicate the activating function of the phosphorylation.

Mutations of residue Thr186, that is phoshorylated for full activation of the kinase, has previously been shown not to affect Brd4 binding (Yang et al. 2005), although *in vivo* Brd4-containing P-TEFb complexes have been shown to be unphosphorylated on Thr186 (Chen et al. 2008). Thr186 is so rapidly exchanged that analysis by H/D mass spectrometry provided no further information on the importance of Thr186 phosphorylation status for Brd4 binding.

3.4.5 Competing CDK9/cyclin T-containing complexes

It is well established that HIV-Tat, Hexim1 and Brd4 are capable of competing for CDK9/cyclin T to ensure transcription of the HIV-LTR, inhibition of CDK9/cyclin T or release of P-TEFb from the large inactive complex respectively (Michels et al. 2003; Jang et al. 2005; Schulte et al. 2005; Yang et al. 2005; Barboric et al. 2007; Bisgrove et al. 2007; Krueger et al. 2010). Developing our structural understanding of these mutually exclusive complexes of P-TEFb will help to understand the regulation of CDK9/cyclin T both in "normal" and HIV-infected cells.

A crystal structure of CDK9/cyclin T/Tat has been solved showing the interaction sites of Tat on both CDK9 and cyclin T (Tahirov et al. 2010). These were described as including a region located on the activation segment of CDK9, and a groove that extends between the cyclin repeats and the H5' helix of cyclin T. The Hexim1 interaction has been suggested to interact with residues 254-272 of cyclin T (Garber et al. 1998; Michels et al. 2003; Schulte et al. 2005).

The two interaction sites identified here between CDK9/cyclin T and Brd4 parallel those of Tat for CDK9/cyclin T. Overlay of the Brd4/cyclin T interaction site on CDK9/cyclin T/Tat shows that this groove is likely to be exploited by both Brd4 and Tat. Furthermore, the H/D exchange experiments suggest that Brd4 contacts residues on CDK9 adjacent to the ATP binding site, overlapping with those that interact with Tat. It appears from our results that Brd4 has a larger footprint on the CDK9 structure than does Tat. Therefore Brd4 appears to be able to compete with Tat both on the CDK and cyclin subunit.

The Tat sequence and Brd4 sequence show little, if any, homology. Tat contains a proline rich region and includes several zinc fingers. Contrastingly, Brd4 contains no prolines within the C-terminal 40mer, and has no cysteines or histidines. Both Brd4 and Tat contain an arginine rich stretch, but whereas this is N-terminal in Brd4, it is C-terminal in Tat (and not present in the crystallised structure). It is therefore somewhat surprising that they appear able to compete for the same interaction surfaces.

In addition to Tat, Brd4 and Hexim, CIITA and Tax compete for the binding site on cyclin T. CIITA competes with both Hexim1 and Tat (Kanazawa et al. 2000) (Kohoutek et al. 2006). Tax is a transcriptional activator encoded by the human T-lymphotropic virus type 1 (HTLV-1), which causes adult T-cell leukemia. It has been shown that Tax competes with Brd4 for binding of CDK9/cyclin T but that this competition involves the Brd4 bromodomain/cyclin T acetyl-residue site (Cho et al. 2010). Additional similarity exist between Tax and Brd4 as both induce CDK9 autoinhibition by phosphorylation on CDK9 Thr29 (Zhou et al. 2006; Zhou et al. 2009b). It has been suggested that Brd4 is recruited to the transcription site where it initiates phosphorylation on CDK9 Thr29 (Zhou et al. 2009b). Release of this phosphorylation will then enable CDK9 activity and control transcriptional elongation. This method of control may be specific to the viral promoters as it has not been observed in cellular gene regions.

The specificity of Tat for the different cyclin T variants is likely to derive from the interface between the cyclin T1 and the C-terminus of Tat and TAR. This interface, located distally to the CDK-binding interface, is where the isoforms differ most in sequence, and is remote from the region I propose to mediate Brd4 binding (Figure 3-31).

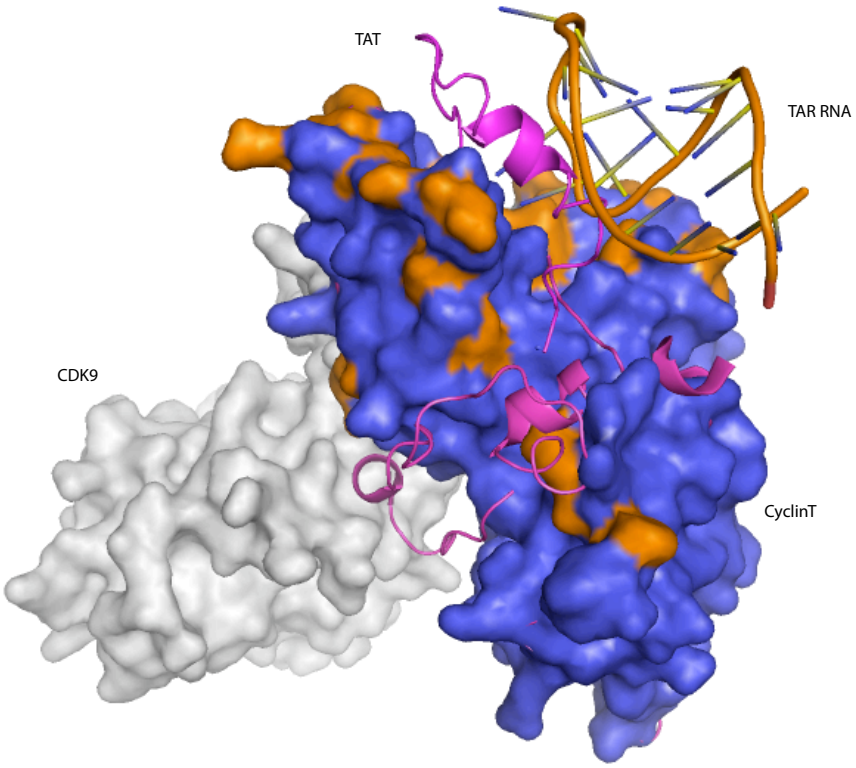


Figure 3-31. Overlay of CDK9/cyclin T/Tat structures. Surface representations of CDK9 (grey) and cyclin T (blue and orange) are shown. Blue residues are conserved between cyclin T1 and cyclin T2 whereas orange are not conserved. Tat (pink) and TAR (orange backbone) are shown as cartoon structures. PDB ids: 3MIA and 2W2H.

3.4.6 Model of the cyclin T/Brd4 interaction

From the data presented in this thesis and previously published structures I propose a structure of Brd4 bound to cyclin T. I suggest that, in a similar interaction to Tat, Brd4 interacts with all cyclin T isoforms through an induced α-helical segment, and then traverses the cleft between CDK and cyclin to interact with the activation segment of CDK9. These models are consistent with mutagenesis and interaction analysis of residues within the C-terminus of Brd4 undertaken by Bisgrove et al (Bisgrove et al. 2007).

Several predictions arise from the models suggested. Firstly that the C-terminal 20 residues of Brd4 interact with cyclin T. Accordingly, this stretch of residues would be unable to interact with CDK9 and therefore would be capable of

binding, but not activating CDK9/cyclin T. CDK9/cyclin T kinetic competition assays with $Brd4_{(1322-1362)}$ and $Brd4_{(1341-1362)}$ would test this hypothesis. Similarly, $Brd4_{(1341-1362)}$ should be capable of competing with Tat for this interaction site. Furthermore, selected mutations of residues on either Brd4 or cyclin T could distinguish between the two possible Brd4 models proposed here.

3.4.7 The same residues of Brd4 bind E2 and CDK9/cyclin T

The regulation of transcription by the HPV E2 protein is still not fully understood, while the influence that Brd4 has on the viral transcription is also not clear (Hamid et al. 2009). The ability of E2 to repress transcription has been suggested to be both dependent and independent of Brd4 (Wu et al. 2006; Schweiger et al. 2007). The uncertainty of whether E2/Brd4/CDK9/cyclin T can form a quaternary complex hinders the understanding of this complex system (Yan et al. 2010).

HPV E2 interacts with Brd4 through the C-terminal residues of Brd4 (Abbate et al. 2006). Upon binding, these C-terminal 20 residues of Brd4 are induced into a helical structure to interact with the transcription regulatory domain of E2. While the involvement of an identical part of Brd4 in the interactions with E2 or cyclin T might suggest that the two interactions are mutually exclusive, the possibility of a quaternary complex cannot be excluded. In such a complex, E2 might bind to one face of the Brd4 C-terminal helix and CDK9/cyclin T to the opposing face. An alternative hypothesis is that an extended section of Brd4

interacts with E2 and the C-terminal 20 residues of Brd4 are not required to maintain association with E2.

There are three reasons that prevent the exclusion of a quaternary complex forming. Firstly, the helix of Brd4 in the E2-Brd4 structure binds not only to a primary E2-molecule, but also to a crystallographically related E2 molecule. Moreover, the crystal contact involves a similar surface area as the primary Brd4-E2 interaction, indicating the potential for protein-protein interactions on both faces of the helix (Figure 3-32). Secondly, the Brd4 CTD (although a construct that is longer that the 20mer helix) has been shown to simultaneously interact with Plk1 and the HPV-5 E2 protein (Wang et al. 2009). And thirdly, the Brd4-E2 and Brd4-CDK9/cyclin T interactions show different dependencies on salt concentrations (Yang et al. 2005; Abbate et al. 2006). The Brd4-E2 interaction is stable under high salt concentrations, suggesting a predominantly hydrophobic interaction surface, in agreement with the interface observed in the crystal structure. Brd4, on the other hand, dissociates from CDK9/cyclin T under relatively low salt concentrations, implying that the binding surface is mediated by polar contacts. These observations suggest that Brd4 may be capable of binding to both E2 and CDK9/cyclin T on different Brd4 surfaces simultaneously.

Two of the three residues whose peaks shift the most in the NMR cyclin T titration data are located on the "solvent-exposed" side of the Brd4-E2 interaction. Whilst Phe1357 is buried in a charged pocket of E2, Leu1354 occupies a reasonably hydrophobic pocket of the crystallography related E2 monomer and Ala1343 (the N-terminal residue in the crystal structure) is

partially solvent exposed. These residues may be capable of interacting with CDK9/cyclin T whilst maintaining the α-helix seen in the E2-Brd4 crystal structure. From the results presented it therefore remains inconclusive as to if a quaternary complex composed of CDK9, cyclin T, E2, Brd4 can exist.

Further knowledge of the interaction between Brd4 and CDK9/cyclin T will help dissecting the influence of the different proteins, the signaling pathways and the various situations involved in the complexity of E2 transcription regulation.

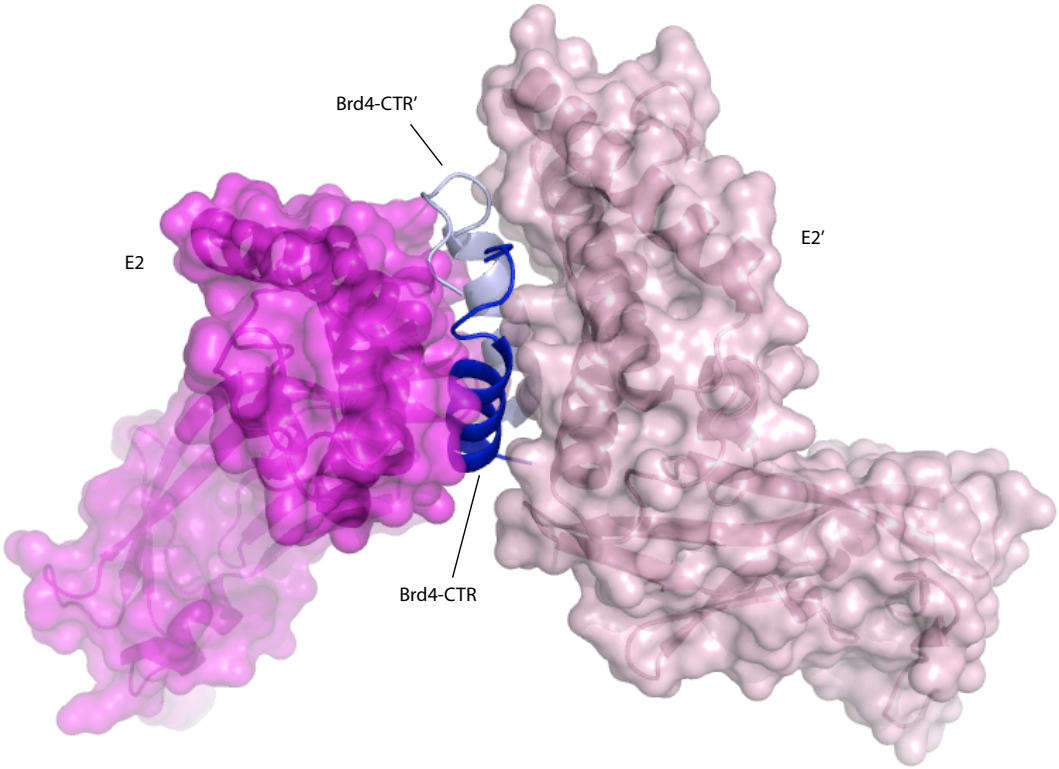


Figure 3-32. Interaction of Brd4 with E2 in the crystal structure. E2 (pink) and Brd4-CTR (blue) represent monomers in a single unit cell, and E2' (pale pink) and Brd4' (pale blue) are monomers in an adjacent unit cell.

3.4.8 Additional regulation by phosphorylation of Brd4

Phosphorylation is a primary method of regulating transcription. At all stages the phosphorylation state of key proteins determines whether transcription can

proceed or is halted. Even within the P-TEFb and its immediate interacting proteins, phosphorylation is a method of regulation. The incorporation of CDK9/cyclin T into its associating complexes also depends on the phosphorylation of the proteins, such as Hexim1, with which it interacts (Contreras et al. 2007).

Phosphorylation of Brd4 provides a potential mechanism by which the cell might control CDK9 recruitment to a subset of genes. Here we showed that the C-terminal region of Brd4 can be phosphorylated *in vitro*. This phosphorylation appears to directly alter the activatory effect that Brd4 exerts on CDK9/cyclin T, and is therefore likely to influence gene transcription. It will be of interest to determine if this phosphorylation prevents the binding of Brd4 to CDK9/cyclin T, or only the subsequent activation of its kinase activity. Furthermore, CDK9/cyclin T phosphorylates many of its binding partners, and so may also be capable of phosphorylating Brd4 in a feedback mechanism. Such an additional phosphorylation would be expected to provide an additional layer to the regulation of transcription in cells.

3.4.9 Concluding model of Brd4 in transcription

I propose therefore that the CTR of Brd4, consisting of the C-terminal 40 residues, binds CDK9/cyclin T kinase and cyclin domains respectively and that the bromodomains of Brd4 interact with an extended section of cyclin T between residues 390-516, including the site of acetylation on cyclin T. In binding to CDK9/cyclin T, Brd4 competes with Hexim1, a component of the 7SK snRNP

inhibitory complex, and thereby localizes P-TEFb at sites of active transcription identified by acetylated histones. In addition, Brd4 increases the kinetic activity of CDK9 increasing the rate of phosphorylation on Pol II CTD and, presumably, DSIF and NELF. This enables a rapid transition of Pol II into the elongation phase of transcription and a consequent increase in mRNA production of this subset of activated genes.

Brd4 also has the potential to regulate transcription by other mechanisms. Firstly Brd4 induces inhibitory autophosphorylation of CDK9 (Zhou et al. 2009b). Secondly, we have shown that phosphorylation of the C-terminus of Brd4 appears to prevent the activation of CDK9. These regulatory phosphorylations, induced by or within Brd4, add an additional layer of regulation to an already complex process that involves the coordination and collaboration of many signaling pathways to regulate transcription by Pol II.

4 Conclusion and future directions

Correct transcription regulation is crucial to the maintenance of a healthy cellular state. CDK9/cyclin T regulates a key transition in transcription in which RNA polymerase II enters the elongation phase from the paused state. This transition has been identified as a major stage of regulation for genes. This thesis has described efforts to develop the understanding of CDK9/cyclin T at a molecular level and hence develop our knowledge of this transition of Pol II to transcription elongation.

In this thesis I have described work on the inhibition (by small chemical probes) and activation of CDK9/cyclin T.

Development of specific inhibitors of high potency towards CDK9 has a two-fold advantage. Firstly more specific inhibitors will enable cell biologists to undertake experiments to probe CDK9 function with minimal cross-reactivity against other cellular kinases. Secondly, as CDK9 activity is altered in several disease states, studies of inhibitors enable an improved understanding of the CDK9 ATP binding site and hence enable structure-aided drug development. I have described the interactions of a class of CDK9 inhibitors in complex with CDK9/cyclin T and CDK2/cyclin A. These have enabled a clearer understanding of the CDK9 ATP-binding site, and, based on favourable and unfavourable protein-inhibitor interactions identified, provided suggestions for the further development of these inhibitors. Progress towards more specific inhibitors will now require the

synthesis of compounds based on these suggestions and subsequent kinase inhibition analysis against an array of kinases including CDK12 and CDK13.

The second focus of this thesis is on the interaction between Brd4 and CDK9/cyclin T, which has been probed here by both kinetic and structural approaches. Brd4 interacts with both the mediator complex and acetylated histones at sites of active transcription. Additionally it interacts with CDK9/cyclin T and recruits this complex to the promoter region to enable phosphorylation of NELF, DSIF and the Pol II CTD. Pol II can then progress into the elongation phase.

From studies I have described within this thesis the minimal CDK9/cyclin T-interaction region of Brd4 has been mapped. Furthermore this fragment has been shown to increase the activity of CDK9/cyclin T by influencing V_{max}, and that may therefore be independent of the identity of the substrate. Additionally, from a range of interaction studies, I have proposed a model of Brd4 binding to CDK9/cyclin T that would be mutually exclusive with the binding of the HIV-Tat protein. Finally I have shown that it is possible that phosphorylation on Brd4 modulates its interaction with P-TEFb, and may therefore be an additional point at which P-TEFb recruitment and hence transcription elongation could be regulated.

Confirming these *in vitro* observations of the interactions between Brd4 and CDK9/cyclin T within a cellular context is crucial to the development of our

understanding of the function of Brd4 with respect to CDK9/cyclin T and its importance in the regulation of transcription.

A. General laboratory methods

A.1 Molecular cloning

A.1.1 Competent cell preparation and transformation

E. coli strains XL1-gold and Arctic express were prepared for subsequent transformations by the following protocol. A single colony was grown overnight at 37°C in 10ml LB. 200ml LB was inoculated with the cell culture and grown at 37°C. At a cell density measuring OD_{600} 0.4-0.6 cells were pelleted at 3800g for 10 minutes at 4°C. Pellets were resuspended in a total volume of 40ml ice-cold sterile TFB1 buffer (30mM potassium acetate, 100mM rubidium chloride, 10mM calcium chloride, 50mM manganese chloride, 15% glycerol, pH 5.8). Suspensions were incubated on ice for 30 minutes. Cells were again pelleted under the same conditions as above. The cell pellet was resuspended in 8ml ice-cold sterile TFB2 (10mM MOPS, 75mM calcium chloride, 10mM rubidium chloride, 15% glycerol, pH 6.8). After incubation on ice for a subsequent 15 minutes, cell suspensions were aliquoted in 200 μl aliquots, flash frozen in liquid nitrogen and stored at -80°C.

For transformation, 50-100μl competent cells were incubated on ice for 5 minutes with *circa* 100ng plasmid DNA. Cells were heatshocked at 42°C for 30 seconds and placed back on ice. 0.7ml sterile LB was added to the cells and they were subsequently incubated at 37°C for 30 minutes. Cells were pelleted by centrifugation for 4 minutes at 6000g and resuspended in 100μl LB. Cell cultures were plated onto LB-agar supplemented with the appropriate antibiotics.

A.1.2 Plasmid DNA preparation

DNA plasmids were transformed into *E.coli* XL1-gold cells. A single colony was grown overnight at 37°C in 3-10 ml LB with the appropriate antibiotic. Cells were pelleted and plasmids were purified using the Qiagen Miniprep Kit. Typically DNA was eluted in 30 μl water.

A.1.3 Production of Novel Clones

DNA fragments were amplified using either Phusion polymerase (New England Biolabs) or Herculase polymerase (Agilent Technologies). The PCR reactions were set up according to the standard protocols for these polymerases. Primers were synthesized by Integrated DNA Technologies. The PCR products were purified using the PCR Purification Kit (Qiagen). Restriction digests were carried out with enzymes obtained from New England Biolabs. The reactions were carried out according to the manufacturer's double digest protocols at 37°C for at least 1 hour. Digested DNA produces were separated by 1% agarose gel electrophoresis and purified using the Gel Extraction Kit (Qiagen). Digested PCR products and vectors were ligated using the Rapid DNA ligation kit (Roche). Ligation solutions were transformed into *E. coli* XL1-gold cells. Plasmid DNA was prepared from cell cultures grown from single colonies. Sequences were confirmed by DNA sequencing undertaken by SourceBioscience, Department of Biochemistry, Oxford.

A.1.4 Mutagenesis

Primer design and site-directed mutagenesis was undertaken according to and using the Quikchange II site directed mutagenesis manual (Stratagene). DNA plasmids were prepared and the introduced mutations were confirmed as described above.

A.2 *Sf9* insect cell maintenance

A.2.1 Maintenance

Sf9 insect cells were maintained in Sf-900 II SFM (1x) media (Gibco). Every 2-3 days they were diluted to a density of 0.5-1 million cells per ml to maintain a density of <5 million cells/ml. Cells were incubated at 26°C, with gentle shaking (120 rpm).

A.2.2 Baculoviral plasmid transfection and virus amplification

Plasmid DNA was transfected into *Sf9* insect cells using the BD BaculoGold Transfection Kit (BD Biosciences). Higher titre stock solutions were produced from subsequent rounds of amplification. Viruses were reamplified by infection of cells (at 1 million/ml) with P2 virus stock. After 3-4 days cell solutions were centrifuged at room temperature, and the virus-containing supernatant was stored at 4°C.

A.3 Protein expression

Specific details for expression and purification of each protein are presented in the methods section within the main text. Recipes for media production are provided in Table A – 1.

Lysogeny broth (LB)	10g tryptone 5g yeast extract 10g NaCl
LB agar	10g tryptone 5g yeast extract 10g NaCl 16g Agar
(^{15}N) minimal medium	200ml 5xM9 medium 1ml 0.1M $CaCl_2$ 2ml 1M $MgSO_4$ 20ml 20% glycerol 20ml 8.5% yeast nitrogen base (without amino acids or ammonium sulfate) 15ml 10% $^{(15)}NH_4Cl$
5x M9 medium	64g $Na_2HPO_4.7H_2O$ 15g KH_2PO_4 2.5g NaCl
Terrific broth (TB)	12g tryptone 24g yeast extract 4ml 87% glycerol 100ml 10x salts
10x salts (for TB)	23.13g KH_2PO_4 123.41g K_2HPO_4

Table A - 1 Recipes for bacterial culture media (preparation for one litre)

A.3.1 E. coli Protein expression

For all proteins, with the exception of CDK2/cyclin A, that were expressed in *E. coli*, plasmids were initially transformed into *E. coli* arctic express cells. These were grown at 37°C overnight on LB agar plates supplemented with the appropriate antibody (100μg/ml ampicillin or 30μg/ml kanamycin). Colonies were picked and grown in *circa* 5ml LB containing antibiotic at 37°C either overnight or for *circa* 5 hours. A fraction of this culture was transferred to flasks

containing the appropriate media and antibiotic for large-scale expression. After induction with IPTG the cells were harvested from large scale cultures by centrifugation at 4°C, for 20 minutes at 4000g.

A.3.2 *Sf9* Protein expression

For protein expression, cells (at 1million/ml) were infected with fresh virus. A range of MOI (multiplicity of infection) was used for the infections. Infected cell solutions were incubated for 72 hours before harvesting cell pellets at 4°C, for 20 minutes at 500g.

A.4 Protein purification and characterisation

A.4.1 Affinity purification of proteins

Gravity flow affinity chromatography was conducted using columns with a bed volume of ~2ml and was carried out at 4°C. For GST-tagged proteins, glutathione Sepharose 4B (GE Healthcare) gravity flow columns were used. His-tagged proteins were purified on Ni-NTA agarose (Qiagen) gravity flow columns. His-tagged affinity purifications were all undertaken in buffers containing <2mM DTT to prevent reduction of the nickel. After loading with clarified lysate, the beads were washed with at least 20mls of the appropriate washing buffers.

A.4.2 General FPLC (Fast protein liquid chromatography) methods

Ion-exchange and gel filtration columns were run on Äkta FPLC Chromatographic systems. Columns were preequilibrated in a minimum of 1

column volume of buffer prior to protein injection. All the proteins were detected by UV absorbance at 280nm, with the exception of Brd4 proteins that were detected at 254nm.

A.4.3 Ultrafiltration and protein concentration determination

Proteins were concentrated in Vivaspin sample concentrators (GE Healthcare) at 4°C. $Brd4_{1322\text{-}1362}$ was always concentrated in a 2ml 3000-MWCO concentrator. Protein concentrations were determined by UV absorbance at 280 nm for all proteins with the exception of Brd4. Extinction coefficients were calculated based on the protein amino acid content using Protparam (http://web.expasy.org/protparam/). Brd4 concentration was determined using the Coomassie protein assay reagent (Pierce) and UV absorbance at 595 nm. BSA solutions of known concentrations were used for calibration for this technique.

A.4.4 Polyacrylamide gel electrophoresis

Tris-tricine and Tris-glycine gels were used to separate and identify proteins. Brd4 samples were primarily run on 16% Tris-tricine gels, whilst all other samples were run on Tris-glycine gels ranging from 8-15% acrylamide content (see Table A - 2 for gel and buffer solutions). Samples were mixed with 5x SDS loading buffer (250mM Tris pH 6.8, 500mM DTT, 10% SDS, 0.5% bromophenol blue, 50% glycerol) prior to being loaded into the wells of the gel. Electrophoresis of Tris-tricine gels was carried out at 100V, and Tris-glycine gels were run at 150-200V. Gels were heated in the microwave for 1 minute in water,

followed by 1 minute heating in stain (0.01% coomassie G, 0.1% HCl, 10% ethanol). All gels were made using the Bio-Rad Mini Protean 3 system.

	Tris-glycine	Tris-tricine
Stacking gel	4.1ml H_2O 1ml 30% acrylamide mix[1] 0.75ml 1M Tris (pH 6.8) 0.06ml 10% SDS 0.06ml 10% ammonium persulfate 0.006ml TEMED	8ml H_2O 1ml AB-3[2] 3ml Gel buffer (3x) 0.09ml 10% ammonium persulfate 0.009ml TEMED
Resolving gel	3.4-6.9ml H_2O 4.0-7.5ml 30% acrylamide mix[1] 3.8ml 1.5M Tris (pH 8.8) 0.15ml 10% SDS 0.15ml 10% ammonium persulfate 0.009ml TEMED (Total volume : 15ml)	7ml H_2O 10ml AB-3[2] 10ml Gel buffer (3x) 3ml glycerol 0.1ml ammonium persulfate 0.01ml TEMED
Cathode buffer (1x)	25mM Tris 250mM glycine 0.1% SDS	0.1M Tris pH 8.25 0.1M Tricine 0.1M SDS
Anode buffer (1x)	Same as cathode	0.1M Tris pH 8.9 0.0225M HCl
Gel buffer (3x)	N/A	3M Tris pH 8.45 1M HCl 0.3% SDS

Table A - 2. Recipes for the production of Tris-glycine and Tris-tricine gels for SDS-PAGE analysis. [1]30% (w/v) acrylamide:0.8% (w/v) bisacrylamide. Stock solution 37.5% (Protogel). [2]Acrylamide:Bisacrylamide (32:1). Stock solution 48% (Severn Biotech).

A.4.5 Mass spectrometry

Proteins were confirmed by electrospray ionization mass spectrometry (ESI-MS). Samples were desalted prior to ESI-MS using Ziptips containing C_4 resin (Millipore).

B. Crystallography Tables

	CDK2/CycA/S3-35	CDK2/CycA/S77	CDK2/CycA/S68	CDK2/CycA/S2-106	CDK2/CycA/S2-104
Data collection					
Beam line	Diamond I-04	Diamond I-03	Diamond I-03	Diamond I-04	Diamond I-02
Space group & unit cell (Å)	$P2_12_12_1$ a=77.20; b=140.37; c=155.17 $\alpha=\beta=\gamma=90°$	$P2_12_12_1$ a=73.81; b=134.55; c=149.17 $\alpha=\beta=\gamma=90°$	$P2_12_12_1$ a=73.90; b=133.81; c=148.38 $\alpha=\beta=\gamma=90°$	$P2_12_12_1$ a=77.08; b=141.16; c=155.52 $\alpha=\beta=\gamma=90°$	$P2_12_12_1$ a=74.05; b=135.08; c=148.19 $\alpha=\beta=\gamma=90°$
Resolution (highest resolution shell) (Å)	29.86-2.05 (2.16-2.05)	29.83-2.26 (2.38-2.26)	52.36-2.05 (2.16-2.05)	52.26-2.40 (2.53-2.40)	49.91-2.10 (2.21-2.10)
Total observations	381071 (37285)	308028 (43211)	296524 (43881)	209450 (25810)	350905 (51105)
Unique	101828 (12695)	69656 (9807)	92057 (13331)	65528 (8880)	87330 (12605)
R_{merge}	0.044 (0.544)	0.071 (0.0496)	0.080 (0.537)	0.127 (0.464)	0.085 (0.549)
Multiplicity	3.7 (2.9)	4.4 (4.4)	3.2 (3.3)	3.2 (2.9)	4.0 (4.1)
Mean I/s_I	17.9 (2.0)	13.4 (2.8)	8.6 (2.0)	5.1 (1.7)	11.9 (2.5)
Completeness	96.5% (83.7%)	99.2% (96.6%)	99.2% (99.4%)	97.9% (93.0%)	99.9% (99.9%)
Refinement Statistics					
(highest resolution shell) (Å)	(2.08-2.05)	(2.29-2.26)	(2.07-2.05)	(2.43-2.40)	(2.12-2.10)
Total number of atoms	9667	9233	9356	9333	9678
Number of waters	628	363	519	331	687
R	19.98 (32.9)	18.10 (23.09)	18.87 (24.92)	20.20 (31.71)	18.38 (25.19)
R_{free} (highest resolution shell)	23.98 (37.13)	21.71 (28.66)	22.60 (28.48)	24.86 (38.73)	21.91 (28.26)
Rms bonds	0.019	0.003	0.005	0.016	0.004
Rms angles	1.586	0.702	0.902	1.499	0.774

	CDK2/CycA/S2-83	CDK2/CycA/E29	CDK9/CycT/ S3-35	CDK9/CycT/S2-104	CDK9/CycT/(S)-CR8
Data collection					
Beam line	Diamond I-02	Diamond I-03	Diamond I-02	ESRF ID14-EH4	Diamond I-02
Space group & unit cell (Å)	$P2_12_12_1$ a=74.07; b=135.41; c=148.63 $\alpha=\beta=\gamma=90°$	$P2_12_12_1$ a=74.30; b=134.80; c=147.90 $\alpha=\beta=\gamma=90°$	H3 a=b=174.11; c=99.261 $\alpha=\beta=90°$; $\gamma=120°$	H3 a=b=174.59; c=99.42 $\alpha=\beta=90°$; $\gamma=120°$	H3 a=b=173.25; c=99.24 $\alpha=\beta=90°$; $\gamma=120°$
Resolution (highest resolution shell) (Å)	59.54-2.45 (2.58-2.45)	74.30-2.11 (2.23-2.11)	60.04-3.10 (3.27-3.10)	50.40-3.16 (3.33-3.16)	50.01-3.00 (3.14-3.00)
Total observations	191491 (28077)	297822 (42616)	73375 (10822)	69409 (9897)	76157 (11267)
Unique	55186 (7979)	85111 (12263)	20202 (2943)	19246 (2800)	21867 (3222)
R_{merge}	0.107 (0.534)	0.055 (0.483)	0.064 (0.593)	0.104 (0.548)	0.074 (0.542)
Multiplicity	3.5 (3.5)	3.5 (3.5)	3.6 (3.7)	3.6 (3.5)	3.5 (3.5)
Mean I/s_I	7.3 (2.1)	16.2(2.6)	10.8 (2.0)	8.2 (2.1)	9.7 (2.1)
Completeness	99.3% (99.7%)	99.2 (99.1)	99.3% (99.5%)	99.6% (99.0%)	98.4% (99.3%)
Refinement Statistics					
(highest resolution shell) (Å)	(2.51-2.45)	(2.14-2.11)	(3.18-3.10)	(3.24-3.16)	(3.14-3.00)
Total number of atoms	9111	9418	4597	4598	4589
Number of waters	287	428	8	6	9
R	19.59 (32.07)	18.30 (24.82)	17.35 (32.5)	16.50 (29.8)	16.23 (26.99)
R_{free} (highest resolution shell)	25.8 (37.14)	21.85 (29.73)	21.99 (40.7)	20.89 (34.9)	21.47 (33.57)
Rms bonds	0.006	0.008	0.006	0.007	0.009
Rms angles	0.964	1.052	0.882	1.041	1.153

	CDK9/CycT/S68	CDK9/CycT/E29	CDK9/CycT/S77	CDK9/CycT/S2-83	CDK9/cyclin T_{F241L}
Data collection					
Beam line	Diamond I-03	Diamond I-03	Diamond I-03	Diamond I-03	Diamond I24
Space group & unit cell (Å)	H3 a=b=172.32; c=98.82 α=β=90°; γ=120°	H3 a=b=173.51; c=98.14 α=β=90°; γ=120°	H3 a=b=172.80; c=98.88 α=β=90°; γ=120°	H3 a=b=174.18; c=98.30 α=β=90°; γ=120°	H3 a=b=173.61; c=97.47 α=β=90°; γ=120°
Resolution (highest resolution shell) (Å)	82.39-3.01 (3.17-3.01)	86.75-2.66 (2.81-2.66)	86.40-3.08 (3.25-3.08)	87.09-2.96 (3.12-2.96)	81.79-3.23 (3.41-3.23)
Total observations	61068 (8985)	151585 (13692)	70331 (10036)	121625 (18367)	58769 (8830)
Unique	20836 (3057)	31058 (4184)	20079 (2956)	23220 (3407)	17300 (2557)
R_{merge}	0.049 (0.386)	0.037 (0.434)	0.058 (0.511)	0.068 (0.581)	0.070 (0.489)
Multiplicity	2.9 (2.9)	4.9 (3.3)	3.5 (3.4)	5.2 (5.4)	3.4 (3.5)
Mean I/s_I	13.5 (2.4)	23.4 (2.4)	15.4 (2.4)	16.3 (2.8)	9.9 (2.4)
Completeness	96.2 (96.6)	98.6 (91.1)	99.0 (99.6)	99.9 (99.9)	98.8 (99.7)
Refinement Statistics					
(highest resolution shell (Å)	(3.01-3.09)	(2.66-2.73)	(3.09-3.17)	(2.96-3.04)	(3.23-3.32)
Total number of atoms	4595	4564	4645	4618	4506
Number of waters	12	33	12	14	0
R	18.62(27.3)	19.39 (32.6)	18.79(31.4)	18.68 (30.5)	16.67 (27.2)
R_{free} (highest resolution shell)	22.05 (29.7)	22.00 (34.7)	21.27(34.6)	21.93 (32.0)	21.79 (36.1)
Rms bonds	0.0102	0.006	0.008	0.009	0.018
Rms angles	1.325	0.942	1.121	1.251	1.81

Table B - 1. Crystallographic parameters

C. Determination of Initial Rate conditions

C.1 Method

Standard conditions

For comparison with the kinase experiments carried out in Chapter 3, the following description is for reaction volumes of 10μl. 10ng CDK9/cyclin T, 15μg GST-CTD in 1x kinase buffer (10mM $MgCl_2$, 50mM Tris pH 8.0, 1mM DTT) in the presence or absence of $Brd4_{(1322-1362)}$ were incubated together. 150ng $Brd4_{(1322-1362)}$ was added per reaction in $CDK9_{FL}$/cyclin T_{WT} experiments, 500ng Brd4 was added per reaction in $CDK9_{330}$/cyclin T_{WT} experiments. Kinase activity was initiated with 100μM ATP (final concentration) and 0.5μCi ^{32}P-γ-ATP. Total reaction volume was 140μl. The reaction was incubated at 30°C. After 1, 2, 3, 4, 5, 6, 7, 8, 10, 15, 20, 40, 60, 120 minutes 10μl of the reaction was terminated by addition to SDS-PAGE loading buffer. Subsequent SDS-PAGE gel separation and phosphor-incorporation analysis was carried out as described before.

1mM ATP analysis

The same conditions were used as for the Standard conditions method above with the following exceptions and clarifications. Only the $CDK9_{FL}$/cyclin T_{WT} complex was assessed in the absence or presence of 150ng $Brd4_{(1322-1362)}$. Reactions were initiated with 1mM ATP (final reaction concentration) with 0.25μCi ^{32}P-γ-ATP present per 10μl reaction. Total reaction volumes were 150μl from which 10μl was removed at each time point.

C.2 Results

A series of experiments were undertaken to confirm that CDK9/cyclin T kinase assays, undertaken in conditions used in this thesis, were measured under initial rate conditions for appropriate enzymatic analysis. Analysis of phosphorylation incorporation into the CTD by CDK9/cyclin T over a time period indicates the region at which initial rate kinetic models can be applied. This was confirmed for standard experimental conditions for both constructs of CDK9 and in the presence and absence of Brd4 (Figure C -1).

The graph shows a primarily linear tendency for the first 10 minutes of the experiment. Therefore experiments undertaken in less than 10 minutes are within a suitable range to be analysed according to standard Michaelis Menton kinetics. All CDK9/cyclin T kinase experiments in this thesis were carried out for 5 minutes to ensure this condition.

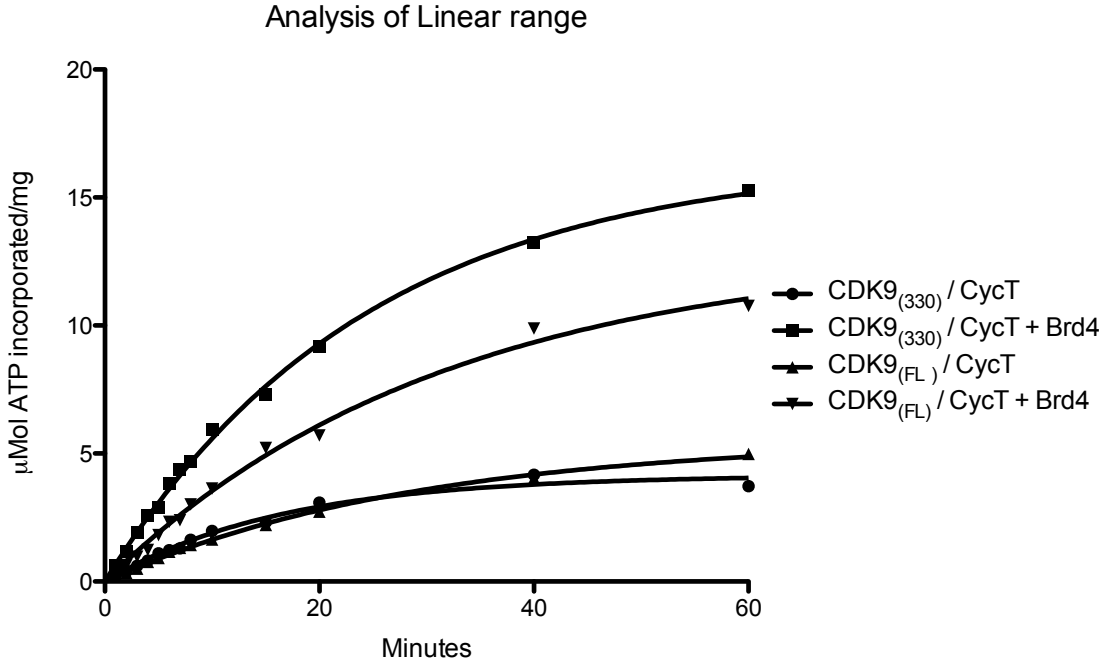


Figure C - 1. Analysis of CDK9/cyclin T kinase activity, under standard conditions, defines that the assay is within the linear range kinetics for at least 5 minutes. Curves were fitted to an exponential association and all had R-square goodness of fit >0.98. These results are from a single experiment.

To ensure experiments in which the ATP concentration was varied were within the linear range, a similar approach was used. Analysis of the reaction in conditions of 1mM ATP was also confirmed to be linear within a 5 minute incubation time (Figure C – 2). This ensured no significant product inhibition contribution to the reaction.

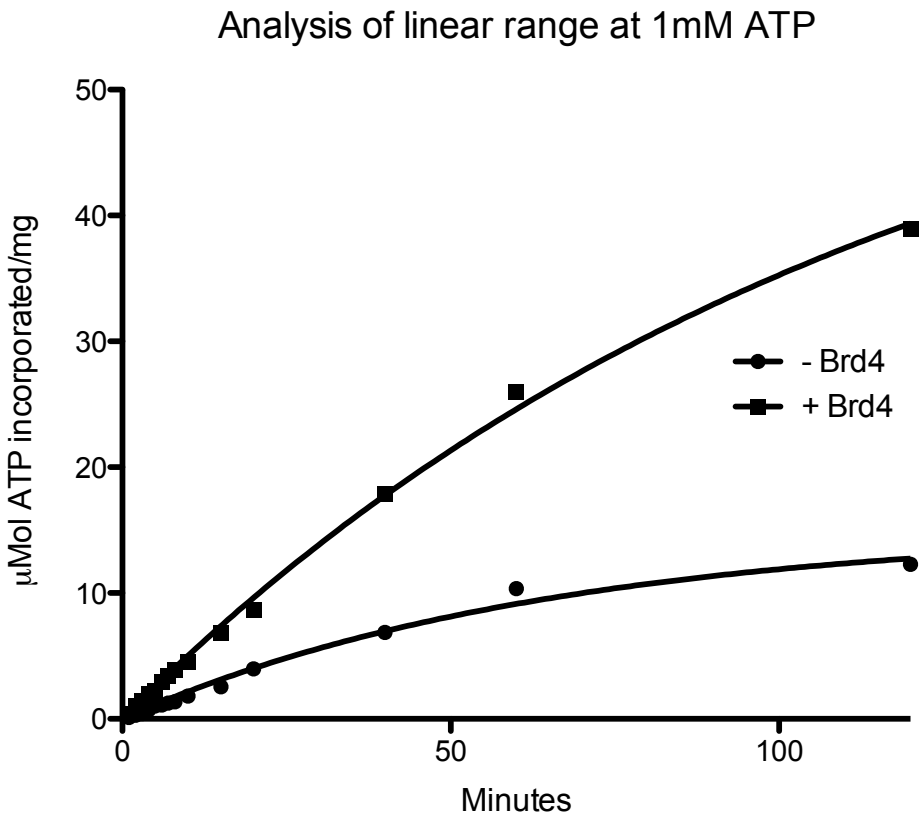


Figure C – 2. Analysis of CDK9$_{FL}$/cyclin T kinase activity at 1mM ATP concentrations defines that the assay under these conditions is within linear range kinetics for at least 5 minutes. Curves were fitted to an exponential association and all had R-square goodness of fit >0.98. These results were from a single experiment.

D. NMR

This appendix by no means delves into the depths of the theory of NMR. Nor does it cover the extent to which NMR can provide information on proteins and nucleic acids. It is meant as a brief summary to aid the understanding of the techniques used within this thesis. For more detailed information on the theoretical and practical aspects of NMR the following books can be recommended (Rattle 1995; Cavanagh 2007).

Basic principles

NMR spectroscopy detects the transition between nuclear spin states within specific atoms, including ^{1}H, ^{15}N and ^{13}C. The spin states of these atoms are affected by magnetic fields and applying a magnetic field produces a split in energy of the different nuclear spin states. Within this magnetic field the nuclear spins precess with a magnetic moment at the Larmor frequency, ω,

$$\omega = \gamma B_0$$

This frequency (ω) is dependent on the applied magnetic field (B_0) and the gyromagnetic ratio of the atom (γ).

Applying a second electromagnetic field (B_1) to the atoms perturbs the nuclear spins from their equilibrium state. Removal of this field enables the nuclear spins to relax back to the original state. This relaxation is measured during the experiments and is

known as the free induction decay (FID). Fourier transformation of the FID produces the typical NMR frequency spectra.

Each nucleus within the magnetic field is affected by its local environment. The surrounding electron clouds are also effected by the applied magnetic field and therefore provide a shielding effect (σ) on the nucleus. This results in nuclei of the same atom type possessing different resonance frequencies. To account for the different magnetic fields applied to the nuclei by different spectrometers, the chemical shift (δ) is used as a measure of resonance position, v_{obs}.

$$v_{obs} = \frac{\gamma B_0 (1 - \sigma)}{2\pi}$$

and,

$$\delta = \frac{v_{ref} - v_{obs}}{v_{ref}} x10^6$$

where v_{ref} is the reference frequency at 0ppm.

The properties of the nuclear spin depend on both the atom and the chemical environment that it is found in. Therefore the chemical shift can provide information on both of these factors.

2D/3D experiments

More complex electromagnetic pulse sequences, with differing time variables, are used to correlate a range of frequencies. These frequencies can be based on the properties of different nuclei (e.g. ^{1}H and ^{15}N). This information can then be used to ascertain information about correlated atoms. Higher dimensional spectra also provide improved resolution.

HSQC (Heteronuclear Single Quantum Correlation Spectroscopy)

^{1}H-^{15}N HSQCs are common NMR spectra obtained for proteins. These display a peak for each correlated ^{15}N and ^{1}H within the protein. Each peak therefore corresponds to one residue within the protein. However, prolines are not represented in these spectra as they do not contain a N-H within their peptide bonds. In addition, side chains of residues containing N-H's will produce additional peaks.

In the absence of full chemical shift assignments HSQCs provide limited information. The dispersion of peaks provides qualitative information on the folded state of the protein. Information about the global effects of different buffer conditions and ligands on the labeled protein can be obtained by HSQCs. However, without full residue assignments, little detailed information can be determined.

Assignment of spectra

For structure determination and binding studies there are various different approaches that can be used to assign peaks to specific residues. These primarily depend on the isotopic enrichment of atoms within the protein.

Three different experiments were used in this thesis for assignment of the ^{15}N labeled protein. These were NOESY- HSQC, TOCSY-HSQC and HSQC-NOESY-HSQC.

In practical terms, NOESY-HSQC and TOCSY-HSQC spectra enable atoms, related through space and through covalent bonds, respectively, to be correlated. For ^{15}N labeled proteins TOCSY-HSQC provides information on the type of amino acid that each peak in the HSQC corresponds to, whilst NOESY-HSQCs allow sequential residues to be identified. A HSQC-NOESY-HSQC spectrum displays cross-peaks corresponding to NOEs between NH protons only. Combining this information, in theory, provides sufficient information to unambiguously assign all peaks within a HSQC spectrum to a residue. However, the behaviour of the protein sample does not always enable ideal data to be collected.

Additional labeling of a protein with ^{13}C enables a much more extensive range of experiments to be undertaken. This allows peaks to be more confidently assigned and additional residue characteristics can be determined.

Protein-protein interaction studies

NMR is a valuable technique for studying protein-protein interactions. It can provide information on the location of interaction sites, binding affinity, stoichiometry and interaction kinetics. As mentioned above, the resonance frequencies and relaxation times of the nuclear spin depends on both the atom and the environment that it is within. Change in environment on binding a second protein could be due to an altered conformation of the labeled protein or direct interaction with the binding protein.

There are three primary effects that binding interactions can have on the HSQC spectra of a protein. These depend on the timescale of the dissociation constant of the protein interaction being measured (Figure D - 1).

During slow exchange two chemical shifts are observed. One shift corresponds to the bound state and the second corresponds to the apo protein. In this scenario an increase in binding protein concentration results in a shift in intensity ratio between the two peaks from the unbound state to the bound state. In this scenario, the line widths provide information on the exchange rates of the interaction.

A fast dissociation rate, relative to the NMR timescale, results in a single frequency that is a weighted average of the bound and unbound states. This results in the classical NMR shifts, seen as peaks migrating across the HSQC spectrum. The dissociation constant can be determined by titrating increasing concentrations of binding protein and measuring the chemical shifts.

In the third scenario, the interaction occurs within an intermediate time frame (relative to the NMR timescale). In this case the peaks decrease in intensity due to line broadening and may disappear. This is therefore the least favourable kinetic state to observe interactions by NMR. However, altering the conditions that the spectra are collected in may change the exchange regime of the system to either fast or slow exchange.

It is also worth noting that it is possible to see a combination of these effects in the HSQC frequency of the protein. This may be a result of chemical shift differences between the free and bound states for different residues within the protein. Another comment to make is that broadening of peaks may occur on binding due to a decreased tumbling rate when a larger complex is formed.

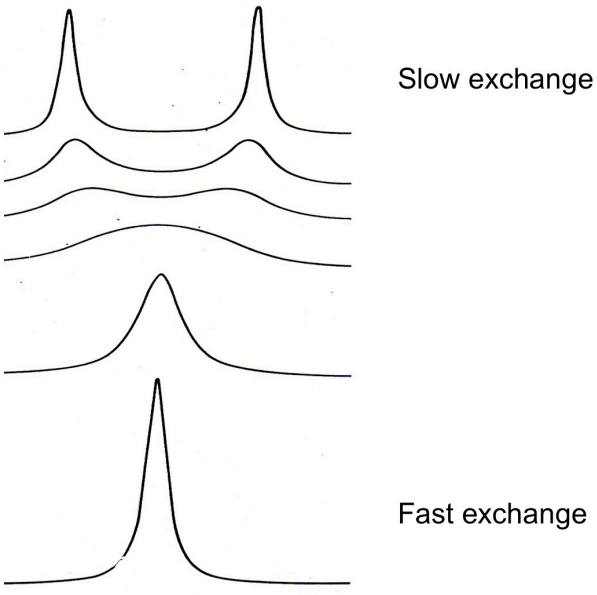


Figure D - 1. Peak distribution under fast and slow exchange rates.

E. Progress towards Structural studies of Larp7

E.1 Introduction

The CDK9/cyclin T inhibitory particle is composed of proteins Hexim1, MePCE, Larp7 (La-related protein 7) and RNA 7SK. Larp7 was initially identified as being significantly mutated in gastric cancers (Mori et al. 2002). It has subsequently been shown to influence expression of tumor-promoter genes through its regulation of P-TEFb (He et al. 2008). Only recently has Larp7 been recognized as binding and protecting all 7SK within the cell and a member of the P-TEFb inhibitory complex (He et al. 2008; Krueger et al. 2008; Barboric et al. 2009). The Larp family of proteins also contains Larps 1 and 4 (Bousquet-Antonelli and Deragon 2009). Much less is known about the function of these Larps but it has been suggested that they influence mRNA metabolism and stability and are involved in mitosis, cell migration, and muscle development (Bayfield et al. 2010; Burrows et al. 2010).

The human La (hLa) protein, which shares the domain structure of Larp7 (Figure E - 1), has been studied in some detail. In a similar manner to Larp7, but with considerably less specificity, hLa protects pre-processed RNAs transcribed from Pol III and Pol II. These include pre-tRNAs, snRNAs and nascent 7SK RNA (reviewed in (Bayfield et al. 2010)).

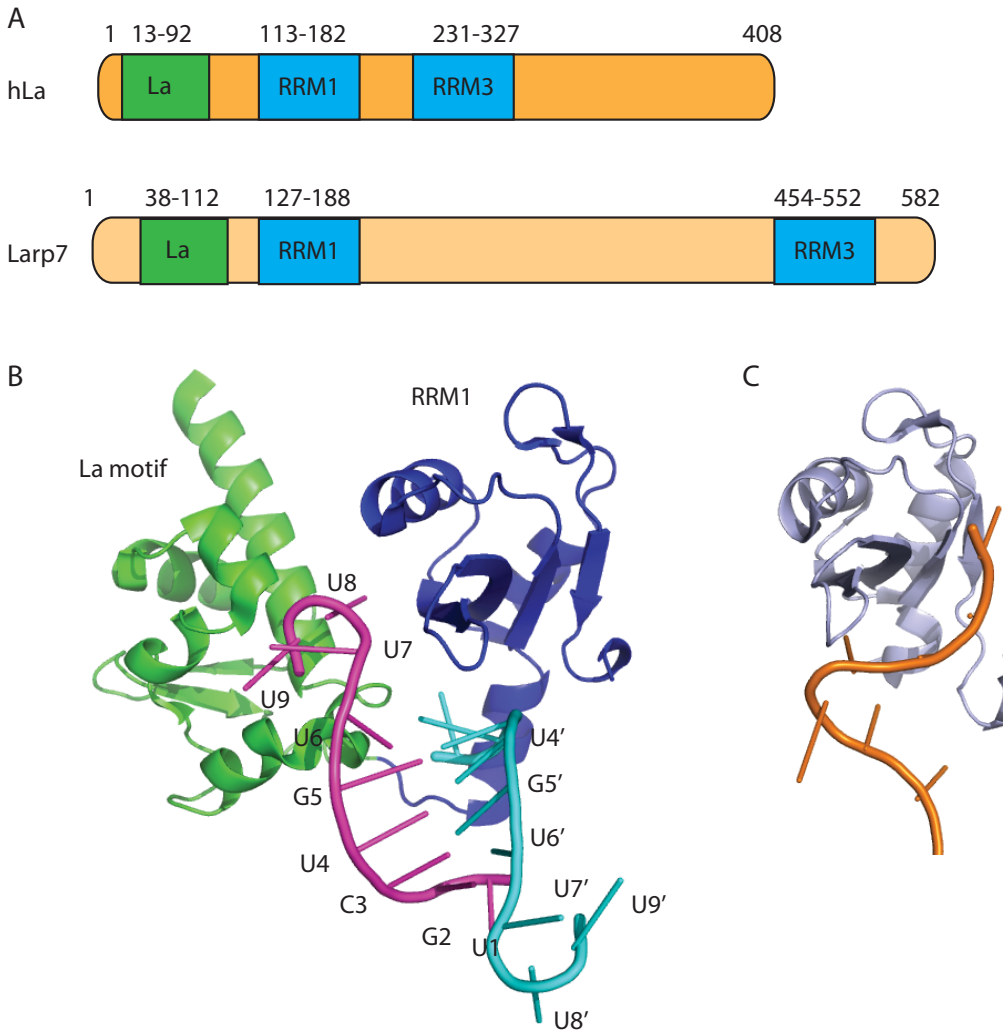


Figure E - 1: Larp7 and hLa comparison. (A) Domain structure of hLa and Larp7. (B) Structure of hLa bound to RNA oligomer, $UGCUGUUUU_{OH}$. The La motif (green), RRM1 (dark blue), and RNA strands (magenta and cyan) are all drawn as cartoons (PDB: 1ZH5). (C) A more typical structure of a RRM domain bound to nucleic acid oligomer. The structure shown is that of hnRNP A1 RRM in complex with a polydeoxyribonucleotide (PDB: 1U1Q). The RRM is shown in the same orientation as that of hLa.

hLa and Larps contain an N-terminal La domain, that consists of a La motif adjacent to an RNA recognition motif (RRM). Both hLa and Larp7 contain an additional atypical C-terminal RRM. The residues involved in RNA interactions within the La domains of Larp7 and hLa are highly conserved suggesting a similar interaction mechanism with RNA. The similarity between the two domains is such that in the absence of hLa the Larp7 La domain can protect RNA 3' ends from degradation (He et al. 2008). The atypical RRM3 of Larp7 is

required for incorporation of Larp7 into the 7SK snRNP complex and therefore is likely to provide the specificity of Larp7 for 7SK RNA binding (Barboric et al. 2009).

Larp7 interacts with several of the components of the 7SK snRNP. The La motif and RRM1 domain of Larp7 and the 3' hairpin loop of 7SK, particularly the 3'UUUU-OH nucleotides, are required for the interaction between 7SK and Larp7 (He et al. 2008; Markert et al. 2008). Furthermore, Larp7 has been shown to interact with MePCE (Xue et al. 2010), CDK9 (Markert et al. 2008) and Hexim1 (He et al. 2008; Markert et al. 2008).

To date there are no structures of the Larp proteins, La or RRM domains, either apo or in association with RNA. There are, however, structures of the hLa La motif and RRM1 (Alfano et al. 2004; Dong et al. 2004; Teplova et al. 2006). The La domain of hLa consists of an atypical winged helix-turn-helix encoded by the La motif and a typical RRM domain. The crystal structure of the N-terminus of hLa bound to RNA shows the interaction surface primarily between the La motif and the RNA (Teplova et al. 2006). Unexpectedly the RNA binds to the RRM surface in an alternative interaction region to the common RNA binding face of RRMs (Figure E - 1). RRMs usually interact with RNA through their β-strand face and loops (Clery et al. 2008). This face in hLa remains available for additional protein-protein or protein-RNA interactions. It is therefore of interest to determine if this is a common feature of Larp family of proteins and if the RRM1 domain has additional functions.

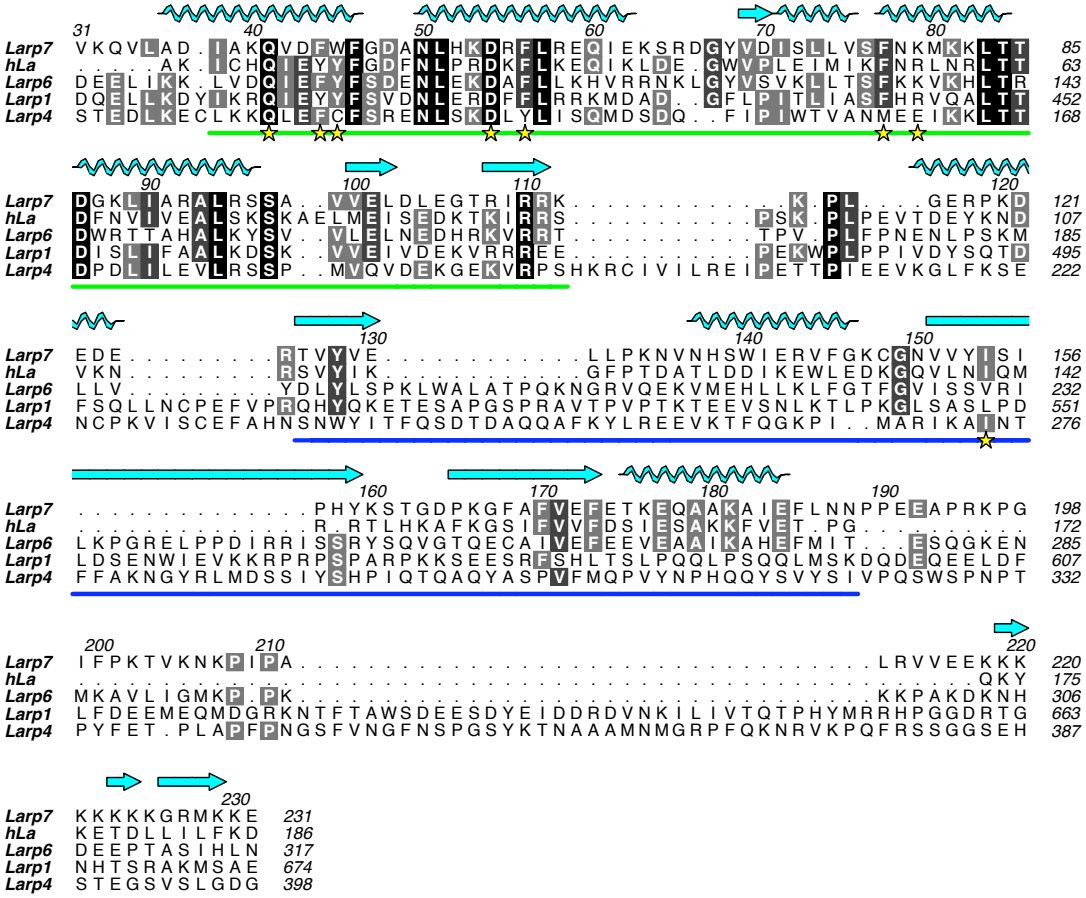


Figure E - 2. Sequence alignment of La-related proteins. Illustrated below the sequences are the La motif (green line), RRM1 (blue line) and residues that interact with the 3 3' RNA nucleotides (stars). The hLa secondary structure is illustrated above the sequences.

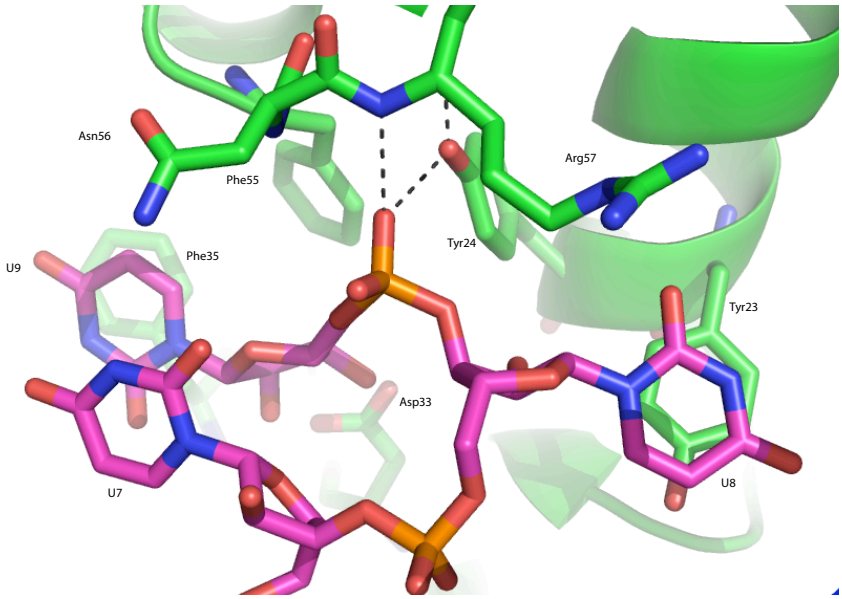


Figure E - 3. Interactions between the 3 3' nucleotides and hLa (PDB:1ZH5). Dashed lines show the hydrogen bonds between the phosphodiester oxygen and residues not conserved between hLa and Larp7. RNA is shown in magenta and hLa in green.

The majority of the residues in the La motif that form the interaction with the 3' end of the RNA are conserved between hLa and Larp7 (Figure E - 2). On close inspection however there are a few anomalies: in particular, residue Y24 and R57 are not conserved. The sidechain of Y24 and the backbone peptide of R57 hydrogen bond with a non-bridging phosphodiester oxygen between the final two bases of the RNA (Figure E - 3). Although an alternative residue at position 57 would be predicted to be accommodating, alternative residues at position 24 have shown a loss in RNA binding to hLa (Teplova et al. 2006). In particular the Y24F mutation prevents RNA binding. Surprisingly the corresponding residue in Larp7 is a phenylalanine. By solving the structure of Larp7 it may be possible to discover if the RNA binding site at this position is maintained in the same location on the La motif and if so how this mutation is accommodated to interact with 7SK.

E.1.1 Objectives

The main objective of the work presented in this chapter was to develop purification protocols of a range of constructs for protein crystallisation with an aim to structure determination. These constructs also provide a basis for future dissection of Larp7 protein-protein and RNA-protein interactions by *in vitro* analysis.

E.2 Methods

E.2.1 Cloning

E. coli constructs to express Larp7 were a gift from Professor Fischer (Department of Biochemistry, Theodor-Boveri-Institute, University of Wurzburg, Germany). These included Larp7 (fl), Larp7 La motif (1-123), Larp 7 LaRRM1 (1-189) and Larp7 RRM3 (375-589) (Markert et al. 2008). Further Larp7 constructs (Table E - 1) were produced by PCR fragments amplified using Phusion polymerase, digested with SalI and BamHI, and ligated into pGEX vectors.

Protein	Construct	Vector
Larp7	16-187	pGEX
Larp7	22-187	pGEX
Larp7	29-187	pGEX
Larp7	12-199	pGEX
Larp7	16-199	pGEX
Larp7	22-199	pGEX
Larp7	29-199	pGEX

Table E - 1 Constructs of Larp7 cloned for experiments. Throughout the text, constructs are identified by including the residue numbers bracketed in subscript after the protein name.

E.2.2 Expression and purification

GST-Larp7 constructs were expressed at 20˚C overnight in *E. coli* strain Arctic express induced with 0.2mM IPTG at OD_{600} ~0.6. The cells were harvested by centrifugation at 4˚C, for 20 minutes at 4000g. Cells were resuspended in 500mM NaCl, 20mM Tris pH 8.0, 5mM DTT and 10% glycerol, flash frozen and stored at -20˚C.

Cells were thawed, benzonase was added to the suspension, and the cells were lysed by sonication on ice. The lysed cells were then centrifuged at 48,000g for 1 hour at 4˚C. The lysate was diluted to <150mM NaCl before purification on a heparin column. The column was washed in 150mM NaCl, 20mM Tris pH 8.0, 5mM DTT and eluted in 1M NaCl, 20mM Tris pH 8.0, 5mM DTT, 10% glycerol. The eluate was then bound to glutathione 4B Sepharose and washed sequentially

with buffer consisting of 1M NaCl, 20mM Tris pH 8.0, 5mM DTT, 10% glycerol, followed by buffer containing 200mM sodium sulfate, 20mM Tris pH 8.0, 5mM DTT and 10% glycerol, and finally in buffer containing 150mM NaCl, 20mM Tris pH 8.0, 5mM DTT. The protein was eluted in 150mM NaCl, 20mM Tris pH 8.0, 5mM DTT supplemented with 20mM glutathione and cleaved with 3C protease, loaded onto a heparin column and eluted over a linear concentration gradient of NaCl. Finally the protein was purified by gel filtration in 200-500mM NaCl, 20mM Tris pH 8.0, 5mM DTT, 10% glycerol. Larp7-containing fractions were concentrated, flash frozen and stored at -80˚C.

E.2.3 Limited proteolysis and N-terminal sequencing

20µg protein was digested with 1.2µg protease. Trypsin digests were incubated at 37˚C and chymotrypsin at 25˚C. Samples were incubated for 1, 3, 10, 30, 60 minutes before being terminated by addition of SDS-loading buffer. N-terminal sequencing was performed by Tony Willis (Protein Characterisation Facility, Department of Biochemistry, Oxford)

E.2.4 Crystallisation trials

Crystallisation trials were set up in a sitting drop format in 96 well plates. Reservoir volumes were 80µl and the drops were composed of 100nl protein solution and 100nl reservoir solution. Plates were incubated at 4˚C and 20˚C.

E.3 Results and Discussion

E.3.1 Larp7 Expression

For structure determination, constructs of the La domain, La-RRM1 domains, RRM3 domain, and the full-length protein were expressed. All these domains were purified and the La domain, RRM3 domain and full-length protein were put into crystallisation trials (Figure E - 4).

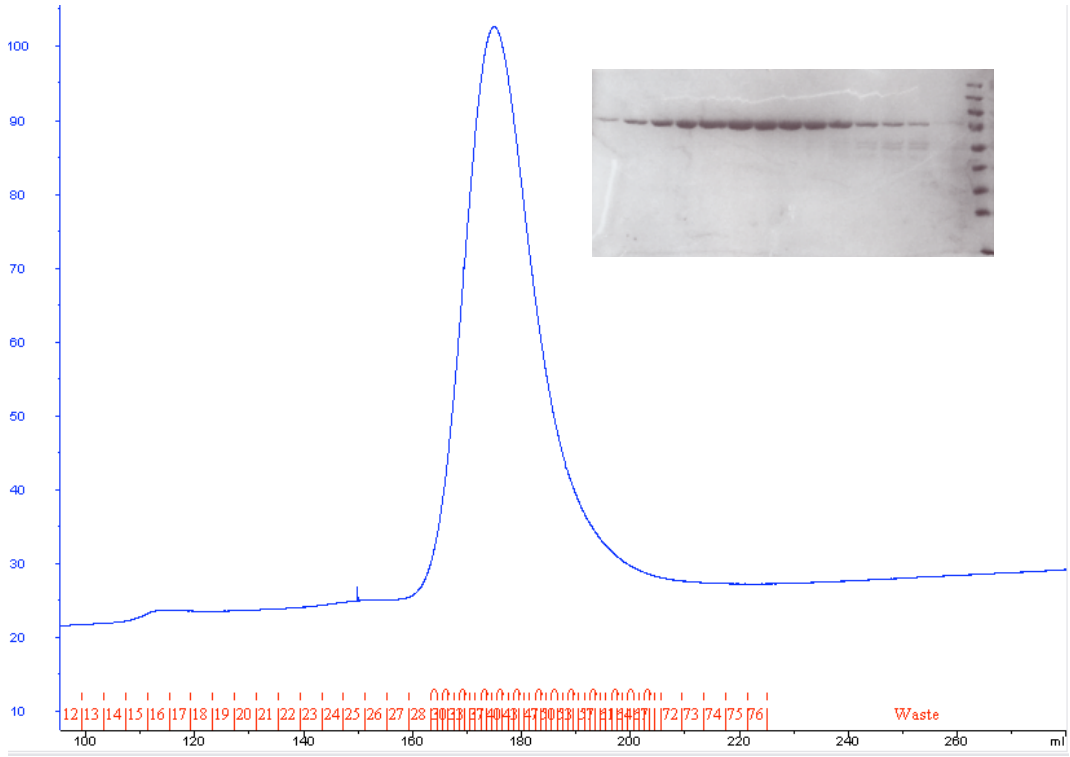


Figure E - 4: Gel filtration chromatogram of full-length Larp7 and selected fractions analysed by SDS-PAGE.

E.3.2 Larp7 limited proteolysis

To establish optimal Larp7 domain boundaries for further crystallisation trials, full-length Larp7 and La-RRM1 were digested with chymotrypsin and trypsin (Figure E - 5). Several bands were identified as more stable constructs resistant to degradation. These samples were sent for N-terminal sequencing. Results both from these experiments, and a stable fragment protein that co-purified with the La-RRM1 construct, allowed the identification of a stable La-RRM1 fragment within residues 12-185. Based on these results, secondary structure predictions

and the hLa structure, several additional constructs were designed in order to optimize the protein for crystallisation trials.

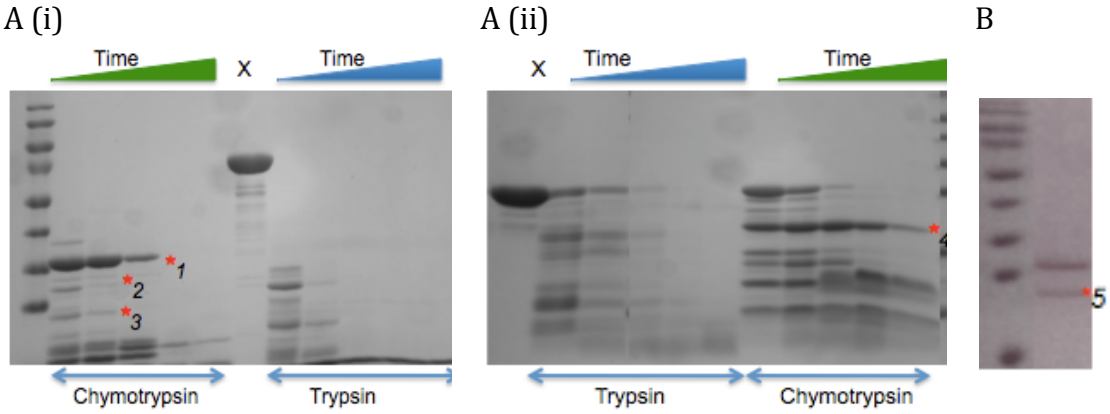


Figure E - 5: A: Time courses of protease digestion. (i): Digest of full-length Larp7. (ii): Digest of La-RRM1 domain of Larp7. B: "Contaminant" protein that copurifies with La-RRM1 domain. X - undigested sample. * - Fragments sent for N-terminal sequencing.

E.3.3 Expression tests and purification

Although expression tests of the new constructs indicated soluble expression of the GST-fusion proteins, larger scale purifications proved problematic (Table E - 2). Both Larp7$_{(16\text{-}199)}$ and Larp7$_{(22\text{-}187)}$ constructs precipitated heavily during overnight 3C cleavage. Although increasing the reaction temperature and incubating for shorter time periods reduced this problem it became apparent that the removal of GST generated protein that was prone to both N-terminal degradation and to aggregation.

Construct	Cloned	Expression tested
16-187	✓	
22-187	✓	✓
29-187	✓	
12-199	✓	✓
16-199	✓	✓
22-199	✓	
29-199	✓	

Table E - 2. Table of Larp7 constructs designed, successfully cloned and then used for large-scale expression tests.

From this analysis, construct Larp7$_{(12\text{-}199)}$ was generated and found to be stable following removal of the GST tag. This construct was purified and concentrated to 11.5mg/ml.

E.3.4 Crystallisation trials

A range of crystallisation screens was set up at different temperatures with various Larp7 constructs (Table E - 3). Unfortunately, to date these have not yielded any diffracting protein crystals.

Temperature (°C)	Construct	Protein Concentration (mg/ml)	Screen
4	RRM2	9.8	Index
			Wizard
			JCSG
			Pact
			MD1/2
	FL	8.0	MD1/2*
			Pact*
		7.0	JCSG
			Pact
			MD1/2
			Wizard
	La	9.0	Pact*
			MD1/2*
	Larp7 (12-199)	11.5	Stura/Natrix
			Wizard
			Index
			JCSG
20	RRM2	9.5	Index
			Pact
			JCSG Plus
			Wizard
			MD1+2
	La	9.0	Wizard
			JCSG
			Pact
			MD1/2

Table E - 3. Crystallisation trials of Larp7 constructs. * - indicates screens set up with a 2:1 ratio protein:reservoir solution.

E.4 Conclusion and future directions

Having tested many crystallisation conditions for several different constructs, Larp7 appears not to be amenable to crystallography. Larp7/7SK complexes may provide a more suitable crystallography target. It will be interesting to determine minimal double or single stranded RNA fragments that interact with the Larp7 protein constructs produced. This will not only develop our understanding of the Larp7-7SK interactions but also provide suggestions for Larp7-7SK co-crystallisation trials.

F. Bibliography

www.clinicaltrials.gov. In.

1994. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr* **50**(Pt 5): 760-763.

Abbate, E.A., Voitenleitner, C., and Botchan, M.R. 2006. Structure of the papillomavirus DNA-tethering complex E2:Brd4 and a peptide that ablates HPV chromosomal association. *Mol Cell* **24**(6): 877-889.

Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W. et al. 2010. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**(Pt 2): 213-221.

Akaike, H. 1974. New Look at Statistical-Model Identification. *Ieee Transactions on Automatic Control* **Ac19**(6): 716-723.

Akhtar, M.S., Heidemann, M., Tietjen, J.R., Zhang, D.W., Chapman, R.D., Eick, D., and Ansari, A.Z. 2009. TFIIH kinase places bivalent marks on the carboxy-terminal domain of RNA polymerase II. *Mol Cell* **34**(3): 387-393.

Alfano, C., Sanfelice, D., Babon, J., Kelly, G., Jacks, A., Curry, S., and Conte, M.R. 2004. Structural analysis of cooperative RNA binding by the La motif and central RRM domain of human La protein. *Nat Struct Mol Biol* **11**(4): 323-329.

Alonso, M., Tamasdan, C., Miller, D.C., and Newcomb, E.W. 2003. Flavopiridol induces apoptosis in glioma cell lines independent of retinoblastoma and p53 tumor suppressor pathway alterations by a caspase-independent pathway. *Mol Cancer Ther* **2**(2): 139-150.

Ammosova, T., Obukhov, Y., Kotelkin, A., Breuer, D., Beullens, M., Gordeuk, V.R., Bollen, M., and Nekhai, S. 2011. Protein phosphatase-1 activates CDK9 by dephosphorylating Ser175. *PLoS One* **6**(4): e18985.

Anderson, M., Beattie, J.F., Breault, G.A., Breed, J., Byth, K.F., Culshaw, J.D., Ellston, R.P., Green, S., Minshull, C.A., Norman, R.A. et al. 2003. Imidazo[1,2-a]pyridines: a potent and selective class of cyclin-dependent kinase inhibitors identified through structure-based hybridisation. *Bioorg Med Chem Lett* **13**(18): 3021-3026.

Arguello, F., Alexander, M., Sterry, J.A., Tudor, G., Smith, E.M., Kalavar, N.T., Greene, J.F., Jr., Koss, W., Morgan, C.D., Stinson, S.F. et al. 1998. Flavopiridol induces apoptosis of normal lymphoid cells, causes immunosuppression, and has potent antitumor activity In vivo against human leukemia and lymphoma xenografts. *Blood* **91**(7): 2482-2490.

Banerjee, A., Sammarco, M.C., Ditch, S., Wang, J., and Grabczyk, E. 2009. A novel tandem reporter quantifies RNA polymerase II termination in mammalian cells. *PLoS One* **4**(7): e6193.

Bannister, A.J. and Kouzarides, T. 2011. Regulation of chromatin by histone modifications. *Cell Res* **21**(3): 381-395.

Bao, Z.Q., Jacobsen, D.M., and Young, M.A. 2011. Briefly bound to activate: transient binding of a second catalytic magnesium activates the structure and dynamics of CDK2 kinase for catalysis. *Structure* **19**(5): 675-690.

Barboric, M., Lenasi, T., Chen, H., Johansen, E.B., Guo, S., and Peterlin, B.M. 2009. 7SK snRNP/P-TEFb couples transcription elongation with alternative splicing and is essential for vertebrate development. *Proc Natl Acad Sci U S A* **106**(19): 7798-7803.

Barboric, M., Yik, J.H., Czudnochowski, N., Yang, Z., Chen, R., Contreras, X., Geyer, M., Matija Peterlin, B., and Zhou, Q. 2007. Tat competes with HEXIM1 to increase the active pool of P-TEFb for HIV-1 transcription. *Nucleic Acids Res* **35**(6): 2003-2012.

Bartkowiak, B., Liu, P., Phatnani, H.P., Fuda, N.J., Cooper, J.J., Price, D.H., Adelman, K., Lis, J.T., and Greenleaf, A.L. 2010. CDK12 is a transcription elongation-associated CTD kinase, the metazoan ortholog of yeast Ctk1. *Genes Dev* **24**(20): 2303-2316.

Baskaran, R., Chiang, G.G., Mysliwiec, T., Kruh, G.D., and Wang, J.Y. 1997. Tyrosine phosphorylation of RNA polymerase II carboxyl-terminal domain by the Abl-related gene product. *J Biol Chem* **272**(30): 18905-18909.

Baskaran, R., Chiang, G.G., and Wang, J.Y. 1996. Identification of a binding site in c-Ab1 tyrosine kinase for the C-terminal repeated domain of RNA polymerase II. *Mol Cell Biol* **16**(7): 3361-3369.

Baskaran, R., Dahmus, M.E., and Wang, J.Y. 1993. Tyrosine phosphorylation of mammalian RNA polymerase II carboxyl-terminal domain. *Proc Natl Acad Sci U S A* **90**(23): 11167-11171.

Baumann, M., Pontiller, J., and Ernst, W. 2010. Structure and basal transcription complex of RNA polymerase II core promoters in the mammalian genome: an overview. *Mol Biotechnol* **45**(3): 241-247.

Baumli, S., Endicott, J.A., and Johnson, L.N. 2010. Halogen bonds form the basis for selective P-TEFb inhibition by DRB. *Chem Biol* **17**(9): 931-936.

Baumli, S., Lolli, G., Lowe, E.D., Troiani, S., Rusconi, L., Bullock, A.N., Debreczeni, J.E., Knapp, S., and Johnson, L.N. 2008. The structure of P-TEFb

(CDK9/cyclin T1), its complex with flavopiridol and regulation by phosphorylation. *EMBO J* **27**(13): 1907-1918.

Bayfield, M.A., Yang, R., and Maraia, R.J. 2010. Conserved and divergent features of the structure and function of La and La-related proteins (LARPs). *Biochim Biophys Acta* **1799**(5-6): 365-378.

Beattie, J.F., Breault, G.A., Ellston, R.P., Green, S., Jewsbury, P.J., Midgley, C.J., Naven, R.T., Minshull, C.A., Pauptit, R.A., Tucker, J.A. et al. 2003. Cyclin-dependent kinase 4 inhibitors as a treatment for cancer. Part 1: identification and optimisation of substituted 4,6-bis anilino pyrimidines. *Bioorg Med Chem Lett* **13**(18): 2955-2960.

Becker, R., Loll, B., and Meinhart, A. 2008. Snapshots of the RNA processing factor SCAF8 bound to different phosphorylated forms of the carboxyl-terminal domain of RNA polymerase II. *J Biol Chem* **283**(33): 22659-22669.

Bettayeb, K., Baunbæk, D., Delehouze, C., Loaëc, N., Hole, A.J., Baumli, S., Endicott, J.A., Douc-Rasy, S., Bénard, J., Oumata, N. et al. 2010. CDK Inhibitors Roscovitine and CR8 Trigger Mcl-1 Down-Regulation and Apoptotic Cell Death in Neuroblastoma Cells. *Genes & Cancer* **1**: 369-380.

Bettayeb, K., Oumata, N., Echalier, A., Ferandin, Y., Endicott, J.A., Galons, H., and Meijer, L. 2008. CR8, a potent and selective, roscovitine-derived inhibitor of cyclin-dependent kinases. *Oncogene* **27**(44): 5797-5807.

Bezstarosti, K., Ghamari, A., Grosveld, F.G., and Demmers, J.A. 2010. Differential proteomics based on 18O labeling to determine the cyclin dependent kinase 9 interactome. *J Proteome Res* **9**(9): 4464-4475.

Biglione, S., Byers, S.A., Price, J.P., Nguyen, V.T., Bensaude, O., Price, D.H., and Maury, W. 2007. Inhibition of HIV-1 replication by P-TEFb inhibitors DRB,

seliciclib and flavopiridol correlates with release of free P-TEFb from the large, inactive form of the complex. *Retrovirology* **4**: 47.

Bisgrove, D.A., Mahmoudi, T., Henklein, P., and Verdin, E. 2007. Conserved P-TEFb-interacting domain of BRD4 inhibits HIV transcription. *Proc Natl Acad Sci U S A* **104**(34): 13690-13695.

Bond, C.S. and Schuttelkopf, A.W. 2009. ALINE: a WYSIWYG protein-sequence alignment editor for publication-quality alignments. *Acta Crystallogr D Biol Crystallogr* **65**(Pt 5): 510-512.

Bousquet-Antonelli, C. and Deragon, J.M. 2009. A comprehensive analysis of the La-motif protein superfamily. *RNA* **15**(5): 750-764.

Breault, G.A., Ellston, R.P., Green, S., James, S.R., Jewsbury, P.J., Midgley, C.J., Pauptit, R.A., Minshull, C.A., Tucker, J.A., and Pease, J.E. 2003. Cyclin-dependent kinase 4 inhibitors as a treatment for cancer. Part 2: identification and optimisation of substituted 2,4-bis anilino pyrimidines. *Bioorg Med Chem Lett* **13**(18): 2961-2966.

Brown, N.R., Noble, M.E., Endicott, J.A., and Johnson, L.N. 1999. The structural basis for specificity of substrate and recruitment peptides for cyclin-dependent kinases. *Nat Cell Biol* **1**(7): 438-443.

Bullock, A.N., Debreczeni, J.E., Fedorov, O.Y., Nelson, A., Marsden, B.D., and Knapp, S. 2005. Structural basis of inhibitor specificity of the human protooncogene proviral insertion site in moloney murine leukemia virus (PIM-1) kinase. *J Med Chem* **48**(24): 7604-7614.

Buratowski, S. 2003. The CTD code. *Nat Struct Biol* **10**(9): 679-680.

Burrows, C., Abd Latip, N., Lam, S.J., Carpenter, L., Sawicka, K., Tzolovsky, G., Gabra, H., Bushell, M., Glover, D.M., Willis, A.E. et al. 2010. The RNA

binding protein Larp1 regulates cell division, apoptosis and cell migration. *Nucleic Acids Res* **38**(16): 5542-5553.

Busenlehner, L.S. and Armstrong, R.N. 2005. Insights into enzyme structure and dynamics elucidated by amide H/D exchange mass spectrometry. *Arch Biochem Biophys* **433**(1): 34-46.

Busso, D., Keriel, A., Sandrock, B., Poterszman, A., Gileadi, O., and Egly, J.M. 2000. Distinct regions of MAT1 regulate cdk7 kinase and TFIIH transcription activities. *J Biol Chem* **275**(30): 22815-22823.

Cabrejos, M.E., Allende, C.C., and Maldonado, E. 2004. Effects of phosphorylation by protein kinase CK2 on the human basal components of the RNA polymerase II transcription machinery. *J Cell Biochem* **93**(1): 2-10.

Cavanagh, J. 2007. *Protein NMR spectroscopy : principles and practice*. Academic Press, Burlington, Mass. ; London.

Chang, C.H., Gourley, T.S., and Sisk, T.J. 2002. Function and regulation of class II transactivator in the immune system. *Immunol Res* **25**(2): 131-142.

Chao, S.H. and Price, D.H. 2001. Flavopiridol inactivates P-TEFb and blocks most RNA polymerase II transcription in vivo. *J Biol Chem* **276**(34): 31793-31799.

Chapman, R.D., Heidemann, M., Albert, T.K., Mailhammer, R., Flatley, A., Meisterernst, M., Kremmer, E., and Eick, D. 2007. Transcribing RNA polymerase II is phosphorylated at CTD residue serine-7. *Science* **318**(5857): 1780-1782.

Chapman, R.D., Palancade, B., Lang, A., Bensaude, O., and Eick, D. 2004. The last CTD repeat of the mammalian RNA polymerase II large subunit is important for its stability. *Nucleic Acids Res* **32**(1): 35-44.

Chen, H.H., Wang, Y.C., and Fann, M.J. 2006. Identification and characterization of the CDK12/cyclin L1 complex involved in alternative splicing regulation. *Mol Cell Biol* **26**(7): 2736-2745.

Chen, H.H., Wong, Y.H., Geneviere, A.M., and Fann, M.J. 2007. CDK13/CDC2L5 interacts with L-type cyclins and regulates alternative splicing. *Biochem Biophys Res Commun* **354**(3): 735-740.

Chen, R., Liu, M., Li, H., Xue, Y., Ramey, W.N., He, N., Ai, N., Luo, H., Zhu, Y., Zhou, N. et al. 2008. PP2B and PP1alpha cooperatively disrupt 7SK snRNP to release P-TEFb for transcription in response to Ca2+ signaling. *Genes Dev* **22**(10): 1356-1368.

Chen, R., Yang, Z., and Zhou, Q. 2004. Phosphorylated positive transcription elongation factor b (P-TEFb) is tagged for inhibition through association with 7SK snRNA. *J Biol Chem* **279**(6): 4153-4160.

Cheng, R. 2006. DPhil Thesis.

Chiu, Y.L., Cao, H., Jacque, J.M., Stevenson, M., and Rana, T.M. 2004. Inhibition of human immunodeficiency virus type 1 replication by RNA interference directed against human transcription elongation factor P-TEFb (CDK9/CyclinT1). *J Virol* **78**(5): 2517-2529.

Cho, E.J., Kobor, M.S., Kim, M., Greenblatt, J., and Buratowski, S. 2001. Opposing effects of Ctk1 kinase and Fcp1 phosphatase at Ser 2 of the RNA polymerase II C-terminal domain. *Genes Dev* **15**(24): 3319-3329.

Cho, W.K., Jang, M.K., Huang, K., Pise-Masison, C.A., and Brady, J.N. 2010. Human T-lymphotropic virus type 1 Tax protein complexes with P-TEFb and competes for Brd4 and 7SK snRNP/HEXIM1 binding. *J Virol* **84**(24): 12801-12809.

Cho, W.K., Zhou, M., Jang, M.K., Huang, K., Jeong, S.J., Ozato, K., and Brady, J.N. 2007. Modulation of the Brd4/P-TEFb interaction by the human T-lymphotropic virus type 1 tax protein. *J Virol* **81**(20): 11179-11186.

Clery, A., Blatter, M., and Allain, F.H. 2008. RNA recognition motifs: boring? Not quite. *Curr Opin Struct Biol* **18**(3): 290-298.

Cole, C., Barber, J.D., and Barton, G.J. 2008. The Jpred 3 secondary structure prediction server. *Nucleic Acids Res* **36**(Web Server issue): W197-201.

Comer, F.I. and Hart, G.W. 2001. Reciprocity between O-GlcNAc and O-phosphate on the carboxyl terminal domain of RNA polymerase II. *Biochemistry* **40**(26): 7845-7852.

Conaway, R.C. and Conaway, J.W. 2011. Function and regulation of the Mediator complex. *Curr Opin Genet Dev* **21**(2): 225-230.

Contreras, X., Barboric, M., Lenasi, T., and Peterlin, B.M. 2007. HMBA releases P-TEFb from HEXIM1 and 7SK snRNA via PI3K/Akt and activates HIV transcription. *PLoS Pathog* **3**(10): 1459-1469.

Corden, J.L. 1990. Tails of RNA polymerase II. *Trends Biochem Sci* **15**(10): 383-387.

Cordingley, M.G., LaFemina, R.L., Callahan, P.L., Condra, J.H., Sardana, V.V., Graham, D.J., Nguyen, T.M., LeGrow, K., Gotlib, L., Schlabach, A.J. et al. 1990. Sequence-specific interaction of Tat protein and Tat peptides with the transactivation-responsive sequence element of human immunodeficiency virus type 1 in vitro. *Proc Natl Acad Sci U S A* **87**(22): 8985-8989.

Core, L.J. and Lis, J.T. 2008. Transcription regulation through promoter-proximal pausing of RNA polymerase II. *Science* **319**(5871): 1791-1792.

Cornish-Bowden, A. 2004. *Fundamentals of enzyme kinetics*. Portland, London.

Crawford, N.P., Alsarraj, J., Lukes, L., Walker, R.C., Officewala, J.S., Yang, H.H., Lee, M.P., Ozato, K., and Hunter, K.W. 2008. Bromodomain 4 activation predicts breast cancer survival. *Proc Natl Acad Sci U S A* **105**(17): 6380-6385.

Cummings, M.D., Farnum, M.A., and Nelen, M.I. 2006. Universal screening methods and applications of ThermoFluor. *J Biomol Screen* **11**(7): 854-863.

De Bondt, H.L., Rosenblatt, J., Jancarik, J., Jones, H.D., Morgan, D.O., and Kim, S.H. 1993. Crystal structure of cyclin-dependent kinase 2. *Nature* **363**(6430): 595-602.

De Luca, A., De Falco, M., Baldi, A., and Paggi, M.G. 2003. Cyclin T: three forms for different roles in physiological and pathological functions. *J Cell Physiol* **194**(2): 101-107.

De Luca, A., Russo, P., Severino, A., Baldi, A., Battista, T., Cavallotti, I., De Luca, L., Baldi, F., Giordano, A., and Paggi, M.G. 2001a. Pattern of expression of cyclin T1 in human tissues. *J Histochem Cytochem* **49**(6): 685-692.

De Luca, A., Tosolini, A., Russo, P., Severino, A., Baldi, A., De Luca, L., Cavallotti, I., Baldi, F., Giordano, A., Testa, J.R. et al. 2001b. Cyclin T2a gene maps on human chromosome 2q21. *J Histochem Cytochem* **49**(6): 693-698.

Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J., and Bax, A. 1995. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* **6**(3): 277-293.

Demidenko, Z.N. and Blagosklonny, M.V. 2004. Flavopiridol induces p53 via initial inhibition of Mdm2 and p21 and, independently of p53, sensitizes

apoptosis-reluctant cells to tumor necrosis factor. *Cancer Res* **64**(10): 3653-3660.

Dey, A., Chitsaz, F., Abbasi, A., Misteli, T., and Ozato, K. 2003. The double bromodomain protein Brd4 binds to acetylated chromatin during interphase and mitosis. *Proc Natl Acad Sci U S A* **100**(15): 8758-8763.

Dey, A., Ellenberg, J., Farina, A., Coleman, A.E., Maruyama, T., Sciortino, S., Lippincott-Schwartz, J., and Ozato, K. 2000. A bromodomain protein, MCAP, associates with mitotic chromosomes and affects G(2)-to-M transition. *Mol Cell Biol* **20**(17): 6537-6549.

Dey, A., Nishiyama, A., Karpova, T., McNally, J., and Ozato, K. 2009. Brd4 marks select genes on mitotic chromatin and directs postmitotic transcription. *Mol Biol Cell* **20**(23): 4899-4909.

Dixon, M. and Webb, E.C. 1979. *Enzymes*. Longman, London.

Dong, G., Chakshusmathi, G., Wolin, S.L., and Reinisch, K.M. 2004. Structure of the La motif: a winged helix domain mediates RNA binding via a conserved aromatic patch. *EMBO J* **23**(5): 1000-1007.

Drees, M., Dengler, W.A., Roth, T., Labonte, H., Mayo, J., Malspeis, L., Grever, M., Sausville, E.A., and Fiebig, H.H. 1997. Flavopiridol (L86-8275): selective antitumor activity in vitro and activity in vivo for prostate carcinoma cells. *Clin Cancer Res* **3**(2): 273-279.

Drozina, G., Kohoutek, J., Nishiya, T., and Peterlin, B.M. 2006. Sequential modifications in class II transactivator isoform 1 induced by lipopolysaccharide stimulate major histocompatibility complex class II transcription in macrophages. *J Biol Chem* **281**(52): 39963-39970.

Durand, L.O., Advani, S.J., Poon, A.P., and Roizman, B. 2005. The carboxyl-terminal domain of RNA polymerase II is phosphorylated by a complex containing cdk9 and infected-cell protein 22 of herpes simplex virus 1. *J Virol* **79**(11): 6757-6762.

Durand, L.O. and Roizman, B. 2008. Role of cdk9 in the optimization of expression of the genes regulated by ICP22 of herpes simplex virus 1. *J Virol* **82**(21): 10591-10599.

Echalier, A., Bettayeb, K., Ferandin, Y., Lozach, O., Clement, M., Valette, A., Liger, F., Marquet, B., Morris, J.C., Endicott, J.A. et al. 2008. Meriolins (3-(pyrimidin-4-yl)-7-azaindoles): synthesis, kinase inhibitory activity, cellular effects, and structure of a CDK2/cyclin A/meriolin complex. *J Med Chem* **51**(4): 737-751.

Egloff, S. and Murphy, S. 2008. Cracking the RNA polymerase II CTD code. *Trends Genet* **24**(6): 280-288.

Egloff, S., O'Reilly, D., Chapman, R.D., Taylor, A., Tanzhaus, K., Pitts, L., Eick, D., and Murphy, S. 2007. Serine-7 of the RNA polymerase II CTD is specifically required for snRNA gene expression. *Science* **318**(5857): 1777-1779.

Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. 2010. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**(Pt 4): 486-501.

Fabrega, C., Shen, V., Shuman, S., and Lima, C.D. 2003. Structure of an mRNA capping enzyme bound to the phosphorylated carboxy-terminal domain of RNA polymerase II. *Mol Cell* **11**(6): 1549-1561.

Fan, H., Sakuraba, K., Komuro, A., Kato, S., Harada, F., and Hirose, Y. 2003. PCIF1, a novel human WW domain-containing protein, interacts with the

phosphorylated RNA polymerase II. *Biochem Biophys Res Commun* **301**(2): 378-385.

Fedorov, O., Marsden, B., Pogacic, V., Rellos, P., Muller, S., Bullock, A.N., Schwaller, J., Sundstrom, M., and Knapp, S. 2007. A systematic interaction map of validated kinase inhibitors with Ser/Thr kinases. *Proc Natl Acad Sci U S A* **104**(51): 20523-20528.

Feng, Y., Ariza, M.E., Goulet, A.C., Shi, J., and Nelson, M.A. 2005. Death-signal-induced relocalization of cyclin-dependent kinase 11 to mitochondria. *Biochem J* **392**(Pt 1): 65-73.

Filippakopoulos, P., Qi, J., Picaud, S., Shen, Y., Smith, W.B., Fedorov, O., Morse, E.M., Keates, T., Hickman, T.T., Felletar, I. et al. 2010. Selective inhibition of BET bromodomains. *Nature* **468**(7327): 1067-1073.

Fisher, R.P. 2005. Secrets of a double agent: CDK7 in cell-cycle control and transcription. *J Cell Sci* **118**(Pt 22): 5171-5180.

Fogh, R.H., Vranken, W.F., Boucher, W., Stevens, T.J., and Laue, E.D. 2006. A nomenclature and data model to describe NMR experiments. *J Biomol NMR* **36**(3): 147-155.

Franck, N., Montembault, E., Rome, P., Pascal, A., Cremet, J.Y., and Giet, R. 2011. CDK11(p58) is required for centriole duplication and Plk4 recruitment to mitotic centrosomes. *PLoS One* **6**(1): e14600.

French, C.A., Miyoshi, I., Aster, J.C., Kubonishi, I., Kroll, T.G., Dal Cin, P., Vargas, S.O., Perez-Atayde, A.R., and Fletcher, J.A. 2001. BRD4 bromodomain gene rearrangement in aggressive carcinoma with translocation t(15;19). *Am J Pathol* **159**(6): 1987-1992.

French, C.A., Miyoshi, I., Kubonishi, I., Grier, H.E., Perez-Atayde, A.R., and Fletcher, J.A. 2003. BRD4-NUT fusion oncogene: a novel mechanism in aggressive carcinoma. *Cancer Res* **63**(2): 304-307.

Frey, N. and Olson, E.N. 2003. Cardiac hypertrophy: the good, the bad, and the ugly. *Annu Rev Physiol* **65**: 45-79.

Fu, J., Yoon, H.G., Qin, J., and Wong, J. 2007. Regulation of P-TEFb elongation complex activity by CDK9 acetylation. *Mol Cell Biol* **27**(13): 4641-4651.

Fujinaga, K., Irwin, D., Huang, Y., Taube, R., Kurosu, T., and Peterlin, B.M. 2004. Dynamics of human immunodeficiency virus transcription: P-TEFb phosphorylates RD and dissociates negative effectors from the transactivation response element. *Mol Cell Biol* **24**(2): 787-795.

Gagnon, D., Joubert, S., Senechal, H., Fradet-Turcotte, A., Torre, S., and Archambault, J. 2009. Proteasomal degradation of the papillomavirus E2 protein is inhibited by overexpression of bromodomain-containing protein 4. *J Virol* **83**(9): 4127-4139.

Galbraith, M.D., Donner, A.J., and Espinosa, J.M. 2010. CDK8: A positive regulator of transcription. *Transcr* **1**(1): 4-12.

Galons, H., Oumata, N., Ferandin, Y., and Meijer, L. 2009. Practical Synthesis of Roscovitine and CR8. *Organic Process Research & Development* **13**(3): 641-644.

Garber, M.E., Mayall, T.P., Suess, E.M., Meisenhelder, J., Thompson, N.E., and Jones, K.A. 2000. CDK9 autophosphorylation regulates high-affinity binding of the human immunodeficiency virus type 1 tat-P-TEFb complex to TAR RNA. *Mol Cell Biol* **20**(18): 6958-6969.

Garber, M.E., Wei, P., KewalRamani, V.N., Mayall, T.P., Herrmann, C.H., Rice, A.P., Littman, D.R., and Jones, K.A. 1998. The interaction between HIV-1 Tat and human cyclin T1 requires zinc and a critical cysteine residue that is not conserved in the murine CycT1 protein. *Genes Dev* **12**(22): 3512-3527.

Garcia, R.A., Pantazatos, D., and Villarreal, F.J. 2004. Hydrogen/deuterium exchange mass spectrometry for investigating protein-ligand interactions. *Assay Drug Dev Technol* **2**(1): 81-91.

Gargano, B., Amente, S., Majello, B., and Lania, L. 2007. P-TEFb is a crucial co-factor for Myc transactivation. *Cell Cycle* **6**(16): 2031-2037.

Ghose, A.K., Herbertz, T., Pippin, D.A., Salvino, J.M., and Mallamo, J.P. 2008. Knowledge based prediction of ligand binding modes and rational inhibitor design for kinase drug discovery. *J Med Chem* **51**(17): 5149-5171.

Gilmour, D.S. and Fan, R. 2008. Derailing the locomotive: transcription termination. *J Biol Chem* **283**(2): 661-664.

Giraud, S., Hurlstone, A., Avril, S., and Coqueret, O. 2004. Implication of BRG1 and cdk9 in the STAT3-mediated activation of the p21waf1 gene. *Oncogene* **23**(44): 7391-7398.

Glover-Cutter, K., Larochelle, S., Erickson, B., Zhang, C., Shokat, K., Fisher, R.P., and Bentley, D.L. 2009. TFIIH-associated Cdk7 kinase functions in phosphorylation of C-terminal domain Ser7 residues, promoter-proximal pausing, and termination by RNA polymerase II. *Mol Cell Biol* **29**(20): 5455-5464.

Grant, S.K. 2009. Therapeutic protein kinase inhibitors. *Cell Mol Life Sci* **66**(7): 1163-1177.

Guenther, M.G., Levine, S.S., Boyer, L.A., Jaenisch, R., and Young, R.A. 2007. A chromatin landmark and transcription initiation at most promoters in human cells. *Cell* **130**(1): 77-88.

Hamid, N.A., Brown, C., and Gaston, K. 2009. The regulation of cell proliferation by the papillomavirus early proteins. *Cell Mol Life Sci* **66**(10): 1700-1717.

Hanks, S.K. and Hunter, T. 1995. Protein kinases 6. The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification. *FASEB J* **9**(8): 576-596.

Hargreaves, D.C., Horng, T., and Medzhitov, R. 2009. Control of inducible gene expression by signal-dependent transcriptional elongation. *Cell* **138**(1): 129-145.

Hausmann, S. and Shuman, S. 2002. Characterization of the CTD phosphatase Fcp1 from fission yeast. Preferential dephosphorylation of serine 2 versus serine 5. *J Biol Chem* **277**(24): 21213-21220.

He, N., Jahchan, N.S., Hong, E., Li, Q., Bayfield, M.A., Maraia, R.J., Luo, K., and Zhou, Q. 2008. A La-related protein modulates 7SK snRNP integrity to suppress P-TEFb-dependent transcriptional elongation and tumorigenesis. *Mol Cell* **29**(5): 588-599.

He, N., Liu, M., Hsu, J., Xue, Y., Chou, S., Burlingame, A., Krogan, N.J., Alber, T., and Zhou, Q. 2010. HIV-1 Tat and host AFF4 recruit two transcription elongation factors into a bifunctional complex for coordinated activation of HIV-1 transcription. *Mol Cell* **38**(3): 428-438.

He, N., Pezda, A.C., and Zhou, Q. 2006. Modulation of a P-TEFb functional equilibrium for the global control of cell growth and differentiation. *Mol Cell Biol* **26**(19): 7068-7076.

Heintzman, N.D. and Ren, B. 2009. Finding distal regulatory elements in the human genome. *Curr Opin Genet Dev* **19**(6): 541-549.

Hirose, Y., Iwamoto, Y., Sakuraba, K., Yunokuchi, I., Harada, F., and Ohkuma, Y. 2008. Human phosphorylated CTD-interacting protein, PCIF1, negatively modulates gene expression by RNA polymerase II. *Biochem Biophys Res Commun* **369**(2): 449-455.

Ho, C.K. and Shuman, S. 1999. Distinct roles for CTD Ser-2 and Ser-5 phosphorylation in the recruitment and allosteric activation of mammalian mRNA capping enzyme. *Mol Cell* **3**(3): 405-411.

Holstege, F.C., van der Vliet, P.C., and Timmers, H.T. 1996. Opening of an RNA polymerase II promoter occurs in two distinct steps and requires the basal transcription factors IIE and IIH. *EMBO J* **15**(7): 1666-1677.

Hou, T., Ray, S., and Brasier, A.R. 2007. The functional role of an interleukin 6-inducible CDK9.STAT3 complex in human gamma-fibrinogen gene expression. *J Biol Chem* **282**(51): 37091-37102.

Houzelstein, D., Bullock, S.L., Lynch, D.E., Grigorieva, E.F., Wilson, V.A., and Beddington, R.S. 2002. Growth and early postimplantation defects in mice deficient for the bromodomain-containing protein Brd4. *Mol Cell Biol* **22**(11): 3794-3802.

Hsin, J.P., Sheth, A., and Manley, J.L. 2011. RNAP II CTD phosphorylated on threonine-4 is required for histone mRNA 3' end processing. *Science* **334**(6056): 683-686.

Hu, D., Mayeda, A., Trembley, J.H., Lahti, J.M., and Kidd, V.J. 2003. CDK11 complexes promote pre-mRNA splicing. *J Biol Chem* **278**(10): 8623-8629.

Huang, B., Yang, X.D., Zhou, M.M., Ozato, K., and Chen, L.F. 2009. Brd4 coactivates transcriptional activation of NF-kappaB via specific binding to acetylated RelA. *Mol Cell Biol* **29**(5): 1375-1387.

Hurvich, C.M. and Tsai, C.L. 1989. Regression and Time-Series Model Selection in Small Samples. *Biometrika* **76**(2): 297-307.

Iankova, I., Petersen, R.K., Annicotte, J.S., Chavey, C., Hansen, J.B., Kratchmarova, I., Sarruf, D., Benkirane, M., Kristiansen, K., and Fajas, L. 2006. Peroxisome proliferator-activated receptor gamma recruits the positive transcription elongation factor b complex to activate transcription and promote adipogenesis. *Mol Endocrinol* **20**(7): 1494-1505.

Ivanov, D., Kwak, Y.T., Guo, J., and Gaynor, R.B. 2000. Domains in the SPT5 protein that modulate its transcriptional regulatory properties. *Mol Cell Biol* **20**(9): 2970-2983.

Jang, M.K., Mochizuki, K., Zhou, M., Jeong, H.S., Brady, J.N., and Ozato, K. 2005. The bromodomain protein Brd4 is a positive regulatory component of P-TEFb and stimulates RNA polymerase II-dependent transcription. *Mol Cell* **19**(4): 523-534.

Jeffrey, P.D., Russo, A.A., Polyak, K., Gibbs, E., Hurwitz, J., Massague, J., and Pavletich, N.P. 1995. Mechanism of CDK activation revealed by the structure of a cyclinA-CDK2 complex. *Nature* **376**(6538): 313-320.

Jeronimo, C., Forget, D., Bouchard, A., Li, Q., Chua, G., Poitras, C., Therien, C., Bergeron, D., Bourassa, S., Greenblatt, J. et al. 2007. Systematic analysis of

the protein interaction network for the human transcription machinery reveals the identity of the 7SK capping enzyme. *Mol Cell* **27**(2): 262-274.

Johnson, L.N. 2009a. Protein kinase inhibitors: contributions from structure to clinical compounds. *Q Rev Biophys* **42**(1): 1-40.

-. 2009b. The regulation of protein phosphorylation. *Biochem Soc Trans* **37**(Pt 4): 627-641.

Kabsch, W. 2010. Xds. *Acta Crystallogr D Biol Crystallogr* **66**(Pt 2): 125-132.

Kamath, A.V., Chong, S., Chang, M., and Marathe, P.H. 2005. P-glycoprotein plays a role in the oral absorption of BMS-387032, a potent cyclin-dependent kinase 2 inhibitor, in rats. *Cancer Chemother Pharmacol* **55**(2): 110-116.

Kanagaraj, R., Huehn, D., MacKellar, A., Menigatti, M., Zheng, L., Urban, V., Shevelev, I., Greenleaf, A.L., and Janscak, P. 2010. RECQ5 helicase associates with the C-terminal repeat domain of RNA polymerase II during productive elongation phase of transcription. *Nucleic Acids Res* **38**(22): 8131-8140.

Kanazawa, S., Okamoto, T., and Peterlin, B.M. 2000. Tat competes with CIITA for the binding to P-TEFb and blocks the expression of MHC class II genes in HIV infection. *Immunity* **12**(1): 61-70.

Kanazawa, S., Soucek, L., Evan, G., Okamoto, T., and Peterlin, B.M. 2003. c-Myc recruits P-TEFb for transcription, cellular proliferation and apoptosis. *Oncogene* **22**(36): 5707-5711.

Kapoor-Vazirani, P., Kagey, J.D., and Vertino, P.M. 2011. SUV420H2-mediated H4K20 trimethylation enforces RNA polymerase II promoter-proximal pausing by blocking hMOF-dependent H4K16 acetylation. *Mol Cell Biol* **31**(8): 1594-1609.

Kastritis, P.L., Moal, I.H., Hwang, H., Weng, Z., Bates, P.A., Bonvin, A.M., and Janin, J. 2011. A structure-based benchmark for protein-protein binding affinity. *Protein Sci* **20**(3): 482-491.

Kelly, W.G., Dahmus, M.E., and Hart, G.W. 1993. RNA polymerase II is a glycoprotein. Modification of the COOH-terminal domain by O-GlcNAc. *J Biol Chem* **268**(14): 10416-10424.

Kohoutek, J., Blazek, D., and Peterlin, B.M. 2006. Hexim1 sequesters positive transcription elongation factor b from the class II transactivator on MHC class II promoters. *Proc Natl Acad Sci U S A* **103**(46): 17349-17354.

Konig, A., Schwartz, G.K., Mohammad, R.M., Al-Katib, A., and Gabrilove, J.L. 1997. The novel cyclin-dependent kinase inhibitor flavopiridol downregulates Bcl-2 and induces growth arrest and apoptosis in chronic B-cell leukemia lines. *Blood* **90**(11): 4307-4312.

Kostrewa, D., Zeller, M.E., Armache, K.J., Seizl, M., Leike, K., Thomm, M., and Cramer, P. 2009. RNA polymerase II-TFIIB structure and mechanism of transcription initiation. *Nature* **462**(7271): 323-330.

Krishnamurthy, S., He, X., Reyes-Reyes, M., Moore, C., and Hampsey, M. 2004. Ssu72 Is an RNA polymerase II CTD phosphatase. *Mol Cell* **14**(3): 387-394.

Krueger, B.J., Jeronimo, C., Roy, B.B., Bouchard, A., Barrandon, C., Byers, S.A., Searcey, C.E., Cooper, J.J., Bensaude, O., Cohen, E.A. et al. 2008. LARP7 is a stable component of the 7SK snRNP while P-TEFb, HEXIM1 and hnRNP A1 are reversibly associated. *Nucleic Acids Res* **36**(7): 2219-2229.

Krueger, B.J., Varzavand, K., Cooper, J.J., and Price, D.H. 2010. The mechanism of release of P-TEFb and HEXIM1 from the 7SK snRNP by viral and cellular

activators includes a conformational change in 7SK. *PLoS One* **5**(8): e12335.

Lalioti, V., Pulido, D., and Sandoval, I.V. 2010. Cdk5, the multifunctional surveyor. *Cell Cycle* **9**(2): 284-311.

Lam, L.T., Pickeral, O.K., Peng, A.C., Rosenwald, A., Hurt, E.M., Giltnane, J.M., Averett, L.M., Zhao, H., Davis, R.E., Sathyamoorthy, M. et al. 2001. Genomic-scale measurement of mRNA turnover and the mechanisms of action of the anti-cancer drug flavopiridol. *Genome Biol* **2**(10): RESEARCH0041.

Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R. et al. 2007. Clustal W and Clustal X version 2.0. *Bioinformatics* **23**(21): 2947-2948.

Laybourn, P.J. and Dahmus, M.E. 1989. Transcription-dependent structural changes in the C-terminal domain of mammalian RNA polymerase subunit IIa/o. *J Biol Chem* **264**(12): 6693-6698.

Lee, D.K., Duan, H.O., and Chang, C. 2001. Androgen receptor interacts with the positive elongation factor P-TEFb and enhances the efficiency of transcriptional elongation. *J Biol Chem* **276**(13): 9978-9984.

Leskovac, V. 2003. *Comprehensive enzyme kinetics*. Kluwer Academic/Plenum, New York ; London.

Li, H., Zhang, Z., Wang, B., Zhang, J., Zhao, Y., and Jin, Y. 2007. Wwp2-mediated ubiquitination of the RNA polymerase II large subunit in mouse embryonic pluripotent stem cells. *Mol Cell Biol* **27**(15): 5296-5305.

Li, L.L., Hu, S.T., Wang, S.H., Lee, H.H., Wang, Y.T., and Ping, Y.H. 2010. Positive transcription elongation factor b (P-TEFb) contributes to dengue virus-stimulated induction of interleukin-8 (IL-8). *Cell Microbiol*.

Li, Q., Price, J.P., Byers, S.A., Cheng, D., Peng, J., and Price, D.H. 2005. Analysis of the large inactive P-TEFb complex indicates that it contains one 7SK molecule, a dimer of HEXIM1 or HEXIM2, and two P-TEFb molecules containing Cdk9 phosphorylated at threonine 186. *J Biol Chem* **280**(31): 28819-28826.

Li, Y., Bhuiyan, M., Alhasan, S., Senderowicz, A.M., and Sarkar, F.H. 2000. Induction of apoptosis and inhibition of c-erbB-2 in breast cancer cells by flavopiridol. *Clin Cancer Res* **6**(1): 223-229.

Liao, J.J. 2007. Molecular recognition of protein kinase binding pockets for design of potent and selective kinase inhibitors. *J Med Chem* **50**(3): 409-424.

Lin, C., Smith, E.R., Takahashi, H., Lai, K.C., Martin-Brown, S., Florens, L., Washburn, M.P., Conaway, J.W., Conaway, R.C., and Shilatifard, A. 2010. AFF4, a component of the ELL/P-TEFb elongation complex and a shared subunit of MLL chimeras, can link transcription elongation to leukemia. *Mol Cell* **37**(3): 429-437.

Lin, Y.J., Umehara, T., Inoue, M., Saito, K., Kigawa, T., Jang, M.K., Ozato, K., Yokoyama, S., Padmanabhan, B., and Guntert, P. 2008. Solution structure of the extraterminal domain of the bromodomain-containing protein BRD4. *Protein Sci* **17**(12): 2174-2179.

Liou, L.Y., Haaland, R.E., Herrmann, C.H., and Rice, A.P. 2006. Cyclin T1 but not cyclin T2a is induced by a post-transcriptional mechanism in PAMP-activated monocyte-derived macrophages. *J Leukoc Biol* **79**(2): 388-396.

Liu, X., Shi, S., Lam, F., Pepper, C., Fischer, P.M., and Wang, S. 2011a. CDKI-71, a novel CDK9 inhibitor, is preferentially cytotoxic to cancer cells compared to flavopiridol. *Int J Cancer*.

-. 2011b. CDKI-71, a novel CDK9 inhibitor, is preferentially cytotoxic to cancer cells when compared with flavopiridol. *Int J Cancer*.

Liu, Y., Wang, X., Zhang, J., Huang, H., Ding, B., Wu, J., and Shi, Y. 2008. Structural basis and binding properties of the second bromodomain of Brd4 with acetylated histone tails. *Biochemistry* **47**(24): 6403-6417.

Lolli, G. 2010. Structural dissection of cyclin dependent kinases regulation and protein recognition properties. *Cell Cycle* **9**(8): 1551-1561.

Lolli, G. and Johnson, L.N. 2005. CAK-Cyclin-dependent Activating Kinase: a key kinase in cell cycle control and a target for drugs? *Cell Cycle* **4**(4): 572-577.

Lolli, G., Lowe, E.D., Brown, N.R., and Johnson, L.N. 2004. The crystal structure of human CDK7 and its protein recognition properties. *Structure* **12**(11): 2067-2079.

Loyer, P., Busson, A., Trembley, J.H., Hyle, J., Grenet, J., Zhao, W., Ribault, C., Montier, T., Kidd, V.J., and Lahti, J.M. 2011. The RNA binding motif protein 15B (RBM15B/OTT3) is a functional competitor of serine-arginine (SR) proteins and antagonizes the positive effect of the CDK11p110-cyclin L2alpha complex on splicing. *J Biol Chem* **286**(1): 147-159.

Loyer, P., Trembley, J.H., Grenet, J.A., Busson, A., Corlu, A., Zhao, W., Kocak, M., Kidd, V.J., and Lahti, J.M. 2008. Characterization of cyclin L1 and L2 interactions with CDK11 and splicing factors: influence of cyclin L isoforms on splice site selection. *J Biol Chem* **283**(12): 7721-7732.

Lu, K., Shih, C., and Teicher, B.A. 2000. Expression of pRB, cyclin/cyclin-dependent kinases and E2F1/DP-1 in human tumor lines in cell culture and in xenograft tissues and response to cell cycle agents. *Cancer Chemother Pharmacol* **46**(4): 293-304.

Lunde, B.M., Reichow, S.L., Kim, M., Suh, H., Leeper, T.C., Yang, F., Mutschler, H., Buratowski, S., Meinhart, A., and Varani, G. 2010. Cooperative interaction of transcription termination factors with the RNA polymerase II C-terminal domain. *Nat Struct Mol Biol* **17**(10): 1195-1201.

Lykke-Andersen, S. and Jensen, T.H. 2006. CUT it out: silencing of noise in the transcriptome. *Nat Struct Mol Biol* **13**(10): 860-861.

Malik, S. and Roeder, R.G. 2010. The metazoan Mediator co-activator complex as an integrative hub for transcriptional regulation. *Nat Rev Genet* **11**(11): 761-772.

Malumbres, M. and Barbacid, M. 2009. Cell cycle, CDKs and cancer: a changing paradigm. *Nat Rev Cancer* **9**(3): 153-166.

Mancebo, H.S., Lee, G., Flygare, J., Tomassini, J., Luu, P., Zhu, Y., Peng, J., Blau, C., Hazuda, D., Price, D. et al. 1997. P-TEFb kinase is required for HIV Tat transcriptional activation in vivo and in vitro. *Genes Dev* **11**(20): 2633-2644.

Markert, A., Grimm, M., Martinez, J., Wiesner, J., Meyerhans, A., Meyuhas, O., Sickmann, A., and Fischer, U. 2008. The La-related protein LARP7 is a component of the 7SK ribonucleoprotein and affects transcription of cellular and viral polymerase II genes. *EMBO Rep* **9**(6): 569-575.

Marshall, N.F. and Price, D.H. 1995. Purification of P-TEFb, a transcription factor required for the transition into productive elongation. *J Biol Chem* **270**(21): 12335-12338.

Maruyama, T., Farina, A., Dey, A., Cheong, J., Bermudez, V.P., Tamura, T., Sciortino, S., Shuman, J., Hurwitz, J., and Ozato, K. 2002. A Mammalian bromodomain protein, brd4, interacts with replication factor C and inhibits progression to S phase. *Mol Cell Biol* **22**(18): 6509-6520.

Maston, G.A., Evans, S.K., and Green, M.R. 2006. Transcriptional regulatory elements in the human genome. *Annu Rev Genomics Hum Genet* **7**: 29-59.

Max, T., Sogaard, M., and Svejstrup, J.Q. 2007. Hyperphosphorylation of the C-terminal repeat domain of RNA polymerase II facilitates dissociation of its complex with mediator. *J Biol Chem* **282**(19): 14113-14120.

Mayer, A., Lidschreiber, M., Siebert, M., Leike, K., Soding, J., and Cramer, P. 2010. Uniform transitions of the general RNA polymerase II transcription complex. *Nat Struct Mol Biol* **17**(10): 1272-1278.

McClue, S.J. and Stuart, I. 2008. Metabolism of the trisubstituted purine cyclin-dependent kinase inhibitor seliciclib (R-roscovitine) in vitro and in vivo. *Drug Metab Dispos* **36**(3): 561-570.

McCracken, S., Fong, N., Rosonina, E., Yankulov, K., Brothers, G., Siderovski, D., Hessel, A., Foster, S., Shuman, S., and Bentley, D.L. 1997. 5'-Capping enzymes are targeted to pre-mRNA by binding to the phosphorylated carboxy-terminal domain of RNA polymerase II. *Genes Dev* **11**(24): 3306-3318.

McIntyre, N.A., McInnes, C., Griffiths, G., Barnett, A.L., Kontopidis, G., Slawin, A.M., Jackson, W., Thomas, M., Zheleva, D.I., Wang, S. et al. 2010. Design,

synthesis, and evaluation of 2-methyl- and 2-amino-N-aryl-4,5-dihydrothiazolo[4,5-h]quinazolin-8-amines as ring-constrained 2-anilino-4-(thiazol-5-yl)pyrimidine cyclin-dependent kinase inhibitors. *J Med Chem* **53**(5): 2136-2145.

McPhillips, M.G., Oliveira, J.G., Spindler, J.E., Mitra, R., and McBride, A.A. 2006. Brd4 is required for e2-mediated transcriptional activation but not genome partitioning of all papillomaviruses. *J Virol* **80**(19): 9530-9543.

Meinhart, A. and Cramer, P. 2004. Recognition of RNA polymerase II carboxy-terminal domain by 3'-RNA-processing factors. *Nature* **430**(6996): 223-226.

Meinhart, A., Kamenski, T., Hoeppner, S., Baumli, S., and Cramer, P. 2005. A structural perspective of CTD function. *Genes Dev* **19**(12): 1401-1415.

Meininghaus, M., Chapman, R.D., Horndasch, M., and Eick, D. 2000. Conditional expression of RNA polymerase II in mammalian cells. Deletion of the carboxyl-terminal domain of the large subunit affects early steps in transcription. *J Biol Chem* **275**(32): 24375-24382.

Michels, A.A., Nguyen, V.T., Fraldi, A., Labas, V., Edwards, M., Bonnet, F., Lania, L., and Bensaude, O. 2003. MAQ1 and 7SK RNA interact with CDK9/cyclin T complexes in a transcription-dependent manner. *Mol Cell Biol* **23**(14): 4859-4869.

Missra, A. and Gilmour, D.S. 2010. Interactions between DSIF (DRB sensitivity inducing factor), NELF (negative elongation factor), and the Drosophila RNA polymerase II transcription elongation complex. *Proc Natl Acad Sci U S A* **107**(25): 11301-11306.

Mochizuki, K., Nishiyama, A., Jang, M.K., Dey, A., Ghosh, A., Tamura, T., Natsume, H., Yao, H., and Ozato, K. 2008. The bromodomain protein Brd4 stimulates G1 gene transcription and promotes progression to S phase. *J Biol Chem* **283**(14): 9040-9048.

Moiola, C., De Luca, P., Gardner, K., Vazquez, E., and De Siervi, A. 2010. Cyclin T1 overexpression induces malignant transformation and tumor growth. *Cell Cycle* **9**(15): 3119-3126.

Molle, D., Maiuri, P., Boireau, S., Bertrand, E., Knezevich, A., Marcello, A., and Basyuk, E. 2007. A real-time view of the TAR:Tat:P-TEFb complex at HIV-1 transcription sites. *Retrovirology* **4**: 36.

Morgan, D.O. 2007. *The cell cycle : principles of control*. New Science Press Ltd in association with Oxford University Press, London.

Mori, Y., Sato, F., Selaru, F.M., Olaru, A., Perry, K., Kimos, M.C., Tamura, G., Matsubara, N., Wang, S., Xu, Y. et al. 2002. Instabilotyping reveals unique mutational spectra in microsatellite-unstable gastric cancers. *Cancer Res* **62**(13): 3641-3645.

Mosley, A.L., Pattenden, S.G., Carey, M., Venkatesh, S., Gilmore, J.M., Florens, L., Workman, J.L., and Washburn, M.P. 2009. Rtr1 is a CTD phosphatase that regulates RNA polymerase II during the transition from serine 5 to serine 2 phosphorylation. *Mol Cell* **34**(2): 168-178.

Murray, S., Udupa, R., Yao, S., Hartzog, G., and Prelich, G. 2001. Phosphorylation of the RNA polymerase II carboxy-terminal domain by the Bur1 cyclin-dependent kinase. *Mol Cell Biol* **21**(13): 4089-4096.

Muse, G.W., Gilchrist, D.A., Nechaev, S., Shah, R., Parker, J.S., Grissom, S.F., Zeitlinger, J., and Adelman, K. 2007. RNA polymerase is poised for activation across the genome. *Nat Genet* **39**(12): 1507-1511.

Nabel, G. and Baltimore, D. 1987. An inducible transcription factor activates expression of human immunodeficiency virus in T cells. *Nature* **326**(6114): 711-713.

Nechaev, S. and Adelman, K. 2011. Pol II waiting in the starting gates: Regulating the transition from transcription initiation into productive elongation. *Biochim Biophys Acta* **1809**(1): 34-45.

Nedea, E., He, X., Kim, M., Pootoolal, J., Zhong, G., Canadien, V., Hughes, T., Buratowski, S., Moore, C.L., and Greenblatt, J. 2003. Organization and function of APT, a subcomplex of the yeast cleavage and polyadenylation factor involved in the formation of mRNA and small nucleolar RNA 3'-ends. *J Biol Chem* **278**(35): 33000-33010.

Ng, H.H., Robert, F., Young, R.A., and Struhl, K. 2003. Targeted recruitment of Set1 histone methylase by elongating Pol II provides a localized mark and memory of recent transcriptional activity. *Mol Cell* **11**(3): 709-719.

Nguyen, V.T., Kiss, T., Michels, A.A., and Bensaude, O. 2001. 7SK small nuclear RNA binds to and inhibits the activity of CDK9/cyclin T complexes. *Nature* **414**(6861): 322-325.

Noble, M.E., Endicott, J.A., and Johnson, L.N. 2004. Protein kinase inhibitors: insights into drug design from structure. *Science* **303**(5665): 1800-1805.

Nojima, M., Huang, Y., Tyagi, M., Kao, H.Y., and Fujinaga, K. 2008. The positive transcription elongation factor b is an essential cofactor for the activation of transcription by myocyte enhancer factor 2. *J Mol Biol* **382**(2): 275-287.

Nowak, D.E., Tian, B., Jamaluddin, M., Boldogh, I., Vergara, L.A., Choudhary, S., and Brasier, A.R. 2008. RelA Ser276 phosphorylation is required for activation of a subset of NF-kappaB-dependent genes by recruiting cyclin-dependent kinase 9/cyclin T1 complexes. *Mol Cell Biol* **28**(11): 3623-3638.

Nutley, B.P., Raynaud, F.I., Wilson, S.C., Fischer, P.M., Hayes, A., Goddard, P.M., McClue, S.J., Jarman, M., Lane, D.P., and Workman, P. 2005. Metabolism and pharmacokinetics of the cyclin-dependent kinase inhibitor R-roscovitine in the mouse. *Mol Cancer Ther* **4**(1): 125-139.

Oeckinghaus, A. and Ghosh, S. 2009. The NF-kappaB family of transcription factors and its regulation. *Cold Spring Harb Perspect Biol* **1**(4): a000034.

Oumata, N., Bettayeb, K., Ferandin, Y., Demange, L., Lopez-Giral, A., Goddard, M.L., Myrianthopoulos, V., Mikros, E., Flajolet, M., Greengard, P. et al. 2008. Roscovitine-derived, dual-specificity inhibitors of cyclin-dependent kinases and casein kinases 1. *J Med Chem* **51**(17): 5229-5242.

Parker, B.W., Kaur, G., Nieves-Neira, W., Taimi, M., Kohlhagen, G., Shimizu, T., Losiewicz, M.D., Pommier, Y., Sausville, E.A., and Senderowicz, A.M. 1998. Early induction of apoptosis in hematopoietic cell lines after exposure to flavopiridol. *Blood* **91**(2): 458-465.

Patel, V., Senderowicz, A.M., Pinto, D., Jr., Igishi, T., Raffeld, M., Quintanilla-Martinez, L., Ensley, J.F., Sausville, E.A., and Gutkind, J.S. 1998. Flavopiridol, a novel cyclin-dependent kinase inhibitor, suppresses the growth of head and neck squamous cell carcinomas by inducing apoptosis. *J Clin Invest* **102**(9): 1674-1681.

Patturajan, M., Wei, X., Berezney, R., and Corden, J.L. 1998. A nuclear matrix protein interacts with the phosphorylated C-terminal domain of RNA polymerase II. *Mol Cell Biol* **18**(4): 2406-2415.

Peng, J., Zhu, Y., Milton, J.T., and Price, D.H. 1998. Identification of multiple cyclin subunits of human P-TEFb. *Genes Dev* **12**(5): 755-762.

Peterlin, B.M. and Price, D.H. 2006. Controlling the elongation phase of transcription with P-TEFb. *Mol Cell* **23**(3): 297-305.

Rahl, P.B., Lin, C.Y., Seila, A.C., Flynn, R.A., McCuine, S., Burge, C.B., Sharp, P.A., and Young, R.A. 2010. c-Myc regulates transcriptional pause release. *Cell* **141**(3): 432-445.

Rahman, S., Sowa, M.E., Ottinger, M., Smith, J.A., Shi, Y., Harper, J.W., and Howley, P.M. 2011. The Brd4 extraterminal domain confers transcription activation independent of pTEFb by recruiting multiple proteins, including NSD3. *Mol Cell Biol* **31**(13): 2641-2652.

Ramakrishnan, R., Yu, W., and Rice, A.P. 2011. Limited redundancy in genes regulated by Cyclin T2 and Cyclin T1. *BMC Res Notes* **4**: 260.

Rattle, H.W.E. 1995. *An NMR primer for life scientists*. Partnership, Fareham.

Ren, S. and Rollins, B.J. 2004. Cyclin C/cdk3 promotes Rb-dependent G0 exit. *Cell* **117**(2): 239-251.

Richard, P. and Manley, J.L. 2009. Transcription termination by nuclear RNA polymerases. *Genes Dev* **23**(11): 1247-1269.

Romano, G. and Giordano, A. 2008. Role of the cyclin-dependent kinase 9-related pathway in mammalian gene expression and human diseases. *Cell Cycle* **7**(23): 3664-3668.

Russo, A.A., Jeffrey, P.D., and Pavletich, N.P. 1996. Structural basis of cyclin-dependent kinase activation by phosphorylation. *Nat Struct Biol* **3**(8): 696-700.

Sabo, A., Lusic, M., Cereseto, A., and Giacca, M. 2008. Acetylation of conserved lysines in the catalytic core of cyclin-dependent kinase 9 inhibits kinase activity and regulates transcription. *Mol Cell Biol* **28**(7): 2201-2212.

Sano, M., Abdellatif, M., Oh, H., Xie, M., Bagella, L., Giordano, A., Michael, L.H., DeMayo, F.J., and Schneider, M.D. 2002. Activation and function of cyclin T-Cdk9 (positive transcription elongation factor-b) in cardiac muscle-cell hypertrophy. *Nat Med* **8**(11): 1310-1317.

Sano, M., Wang, S.C., Shirai, M., Scaglia, F., Xie, M., Sakai, S., Tanaka, T., Kulkarni, P.A., Barger, P.M., Youker, K.A. et al. 2004. Activation of cardiac Cdk9 represses PGC-1 and confers a predisposition to heart failure. *EMBO J* **23**(17): 3559-3569.

Santamaria, D., Barriere, C., Cerqueira, A., Hunt, S., Tardy, C., Newton, K., Caceres, J.F., Dubus, P., Malumbres, M., and Barbacid, M. 2007. Cdk1 is sufficient to drive the mammalian cell cycle. *Nature* **448**(7155): 811-815.

Sayed, D., Hong, C., Chen, I.Y., Lypowy, J., and Abdellatif, M. 2007. MicroRNAs play an essential role in the development of cardiac hypertrophy. *Circ Res* **100**(3): 416-424.

Schenone, S., Bruno, O., Radi, M., and Botta, M. 2011. New insights into small-molecule inhibitors of Bcr-Abl. *Med Res Rev* **31**(1): 1-41.

Schmerwitz, U.K., Sass, G., Khandoga, A.G., Joore, J., Mayer, B.A., Berberich, N., Totzke, F., Krombach, F., Tiegs, G., Zahler, S. et al. 2011. Flavopiridol protects against inflammation by attenuating leukocyte-endothelial

interaction via inhibition of cyclin-dependent kinase 9. *Arterioscler Thromb Vasc Biol* **31**(2): 280-288.

Schneider, E.V., Bottcher, J., Blaesse, M., Neumann, L., Huber, R., and Maskos, K. 2011. The Structure of CDK8/CycC Implicates Specificity in the CDK/Cyclin Family and Reveals Interaction with a Deep Pocket Binder. *J Mol Biol* **412**(2): 251-266.

Schulte, A., Czudnochowski, N., Barboric, M., Schonichen, A., Blazek, D., Peterlin, B.M., and Geyer, M. 2005. Identification of a cyclin T-binding domain in Hexim1 and biochemical analysis of its binding competition with HIV-1 Tat. *J Biol Chem* **280**(26): 24968-24977.

Schweiger, M.R., Ottinger, M., You, J., and Howley, P.M. 2007. Brd4-independent transcriptional repression function of the papillomavirus e2 proteins. *J Virol* **81**(18): 9612-9622.

Schweiger, M.R., You, J., and Howley, P.M. 2006. Bromodomain protein 4 mediates the papillomavirus E2 transcriptional activation function. *J Virol* **80**(9): 4276-4285.

Shaw, P.E. 2007. Peptidyl-prolyl cis/trans isomerases and transcription: is there a twist in the tail? *EMBO Rep* **8**(1): 40-45.

Simone, C., Stiegler, P., Bagella, L., Pucci, B., Bellan, C., De Falco, G., De Luca, A., Guanti, G., Puri, P.L., and Giordano, A. 2002. Activation of MyoD-dependent transcription by cdk9/cyclin T2. *Oncogene* **21**(26): 4137-4148.

Sims, R.J., 3rd, Rojas, L.A., Beck, D., Bonasio, R., Schuller, R., Drury, W.J., 3rd, Eick, D., and Reinberg, D. 2011. The C-terminal domain of RNA polymerase II is modified by site-specific methylation. *Science* **332**(6025): 99-103.

Smallie, T., Ricchetti, G., Horwood, N.J., Feldmann, M., Clark, A.R., and Williams, L.M. 2010. IL-10 inhibits transcription elongation of the human TNF gene in primary macrophages. *J Exp Med* **207**(10): 2081-2088.

Smith, J.A., White, E.A., Sowa, M.E., Powell, M.L., Ottinger, M., Harper, J.W., and Howley, P.M. 2010. Genome-wide siRNA screen identifies SMCX, EP400, and Brd4 as E2-dependent regulators of human papillomavirus oncogene expression. *Proc Natl Acad Sci U S A* **107**(8): 3752-3757.

Sobhian, B., Laguette, N., Yatim, A., Nakamura, M., Levy, Y., Kiernan, R., and Benkirane, M. 2010. HIV-1 Tat assembles a multifunctional transcription elongation complex and stably associates with the 7SK snRNP. *Mol Cell* **38**(3): 439-451.

Spilianakis, C., Kretsovali, A., Agalioti, T., Makatounakis, T., Thanos, D., and Papamatheakis, J. 2003. CIITA regulates transcription onset viaSer5-phosphorylation of RNA Pol II. *EMBO J* **22**(19): 5125-5136.

Sun, X., Zhang, Y., Cho, H., Rickert, P., Lees, E., Lane, W., and Reinberg, D. 1998. NAT, a human complex containing Srb polypeptides that functions as a negative regulator of activated transcription. *Mol Cell* **2**(2): 213-222.

Suryadinata, R., Sadowski, M., and Sarcevic, B. 2010. Control of cell cycle progression by phosphorylation of cyclin-dependent kinase (CDK) substrates. *Biosci Rep* **30**(4): 243-255.

Tahirov, T.H., Babayeva, N.D., Varzavand, K., Cooper, J.J., Sedore, S.C., and Price, D.H. 2010. Crystal structure of HIV-1 Tat complexed with human P-TEFb. *Nature* **465**(7299): 747-751.

Takahashi, H., Parmely, T.J., Sato, S., Tomomori-Sato, C., Banks, C.A., Kong, S.E., Szutorisz, H., Swanson, S.K., Martin-Brown, S., Washburn, M.P. et al. 2011.

Human mediator subunit MED26 functions as a docking site for transcription elongation factors. *Cell* **146**(1): 92-104.

Takaya, T., Ono, K., Kawamura, T., Takanabe, R., Kaichi, S., Morimoto, T., Wada, H., Kita, T., Shimatsu, A., and Hasegawa, K. 2009. MicroRNA-1 and MicroRNA-133 in spontaneous myocardial differentiation of mouse embryonic stem cells. *Circ J* **73**(8): 1492-1497.

Teplova, M., Yuan, Y.R., Phan, A.T., Malinina, L., Ilin, S., Teplov, A., and Patel, D.J. 2006. Structural basis for recognition and sequestration of UUU(OH) 3' temini of nascent RNA polymerase III transcripts by La, a rheumatic disease autoantigen. *Mol Cell* **21**(1): 75-85.

Tian, Y., Ke, S., Chen, M., and Sheng, T. 2003. Interactions between the aryl hydrocarbon receptor and P-TEFb. Sequential recruitment of transcription factors and differential phosphorylation of C-terminal domain of RNA polymerase II at cyp1a1 promoter. *J Biol Chem* **278**(45): 44041-44048.

Trembley, J.H., Hu, D., Hsu, L.C., Yeung, C.Y., Slaughter, C., Lahti, J.M., and Kidd, V.J. 2002. PITSLRE p110 protein kinases associate with transcription complexes and affect their activity. *J Biol Chem* **277**(4): 2589-2596.

Trigon, S., Serizawa, H., Conaway, J.W., Conaway, R.C., Jackson, S.P., and Morange, M. 1998. Characterization of the residues phosphorylated in vitro by different C-terminal domain kinases. *J Biol Chem* **273**(12): 6769-6775.

Vasiljeva, L., Kim, M., Mutschler, H., Buratowski, S., and Meinhart, A. 2008. The Nrd1-Nab3-Sen1 termination complex interacts with the Ser5-phosphorylated RNA polymerase II C-terminal domain. *Nat Struct Mol Biol* **15**(8): 795-804.

Venters, B.J. and Pugh, B.F. 2009. A canonical promoter organization of the transcription machinery and its regulators in the Saccharomyces genome. *Genome Res* **19**(3): 360-371.

Vita, M., Abdel-Rehim, M., Olofsson, S., Hassan, Z., Meurling, L., Siden, A., Siden, M., Pettersson, T., and Hassan, M. 2005. Tissue distribution, pharmacokinetics and identification of roscovitine metabolites in rat. *Eur J Pharm Sci* **25**(1): 91-103.

Vollmuth, F., Blankenfeldt, W., and Geyer, M. 2009. Structures of the dual bromodomains of the P-TEFb-activating protein Brd4 at atomic resolution. *J Biol Chem* **284**(52): 36547-36556.

Wang, S. and Fischer, P.M. 2008. Cyclin-dependent kinase 9: a key transcriptional regulator and potential drug target in oncology, virology and cardiology. *Trends Pharmacol Sci* **29**(6): 302-313.

Wang, S., Griffiths, G., Midgley, C.A., Barnett, A.L., Cooper, M., Grabarek, J., Ingram, L., Jackson, W., Kontopidis, G., McClue, S.J. et al. 2010. Discovery and characterization of 2-anilino-4- (thiazol-5-yl)pyrimidine transcriptional CDK inhibitors as anticancer agents. *Chem Biol* **17**(10): 1111-1121.

Wang, S., Meades, C., Wood, G., Osnowski, A., Anderson, S., Yuill, R., Thomas, M., Mezna, M., Jackson, W., Midgley, C. et al. 2004a. 2-Anilino-4-(thiazol-5-yl)pyrimidine CDK inhibitors: synthesis, SAR analysis, X-ray crystallography, and biological activity. *J Med Chem* **47**(7): 1662-1675.

Wang, S., Wood, G., Meades, C., Griffiths, G., Midgley, C., McNae, I., McInnes, C., Anderson, S., Jackson, W., Mezna, M. et al. 2004b. Synthesis and biological activity of 2-anilino-4-(1H-pyrrol-3-yl) pyrimidine CDK inhibitors. *Bioorg Med Chem Lett* **14**(16): 4237-4240.

Wang, W.S., Lee, M.S., Tseng, C.E., Liao, I.H., Huang, S.P., Lin, R.I., and Li, C. 2009. Interaction between human papillomavirus type 5 E2 and polo-like kinase 1. *J Med Virol* **81**(3): 536-544.

Wei, P., Garber, M.E., Fang, S.M., Fischer, W.H., and Jones, K.A. 1998. A novel CDK9-associated C-type cyclin interacts directly with HIV-1 Tat and mediates its high-affinity, loop-specific binding to TAR RNA. *Cell* **92**(4): 451-462.

Williams, S.P., Kuyper, L.F., and Pearce, K.H. 2005. Recent applications of protein crystallography and structure-guided drug design. *Curr Opin Chem Biol* **9**(4): 371-380.

Wimmer, J., Fujinaga, K., Taube, R., Cujec, T.P., Zhu, Y., Peng, J., Price, D.H., and Peterlin, B.M. 1999. Interactions between Tat and TAR and human immunodeficiency virus replication are facilitated by human cyclin T1 but not cyclins T2a or T2b. *Virology* **255**(1): 182-189.

Wu, C.H., Yamaguchi, Y., Benjamin, L.R., Horvat-Gordon, M., Washinsky, J., Enerly, E., Larsson, J., Lambertsson, A., Handa, H., and Gilmour, D. 2003a. NELF and DSIF cause promoter proximal pausing on the hsp70 promoter in Drosophila. *Genes Dev* **17**(11): 1402-1414.

Wu, S.Y. and Chiang, C.M. 2007. The double bromodomain-containing chromatin adaptor Brd4 and transcriptional regulation. *J Biol Chem* **282**(18): 13141-13145.

Wu, S.Y., Lee, A.Y., Hou, S.Y., Kemper, J.K., Erdjument-Bromage, H., Tempst, P., and Chiang, C.M. 2006. Brd4 links chromatin targeting to HPV transcriptional silencing. *Genes Dev* **20**(17): 2383-2396.

Wu, S.Y., McNae, I., Kontopidis, G., McClue, S.J., McInnes, C., Stewart, K.J., Wang, S., Zheleva, D.I., Marriage, H., Lane, D.P. et al. 2003b. Discovery of a novel family of CDK inhibitors with the program LIDAEUS: structural basis for ligand-induced disordering of the activation loop. *Structure* **11**(4): 399-410.

Xiao, T., Hall, H., Kizer, K.O., Shibata, Y., Hall, M.C., Borchers, C.H., and Strahl, B.D. 2003. Phosphorylation of RNA polymerase II CTD regulates H3 methylation in yeast. *Genes Dev* **17**(5): 654-663.

Xu, Y.X., Hirose, Y., Zhou, X.Z., Lu, K.P., and Manley, J.L. 2003. Pin1 modulates the structure and function of human RNA polymerase II. *Genes Dev* **17**(22): 2765-2776.

Xu, Y.X. and Manley, J.L. 2007. Pin1 modulates RNA polymerase II activity during the transcription cycle. *Genes Dev* **21**(22): 2950-2962.

Xue, Y., Yang, Z., Chen, R., and Zhou, Q. 2010. A capping-independent function of MePCE in stabilizing 7SK snRNA and facilitating the assembly of 7SK snRNP. *Nucleic Acids Res* **38**(2): 360-369.

Yamada, T., Yamaguchi, Y., Inukai, N., Okamoto, S., Mura, T., and Handa, H. 2006. P-TEFb-mediated phosphorylation of hSpt5 C-terminal repeats is critical for processive transcription elongation. *Mol Cell* **21**(2): 227-237.

Yan, J., Li, Q., Lievens, S., Tavernier, J., and You, J. 2010. Abrogation of the Brd4-positive transcription elongation factor B complex by papillomavirus E2 protein contributes to viral oncogene repression. *J Virol* **84**(1): 76-87.

Yang, Z., Yik, J.H., Chen, R., He, N., Jang, M.K., Ozato, K., and Zhou, Q. 2005. Recruitment of P-TEFb for stimulation of transcriptional elongation by the bromodomain protein Brd4. *Mol Cell* **19**(4): 535-545.

Yang, Z., Zhu, Q., Luo, K., and Zhou, Q. 2001. The 7SK small nuclear RNA inhibits the CDK9/cyclin T1 kinase to control transcription. *Nature* **414**(6861): 317-322.

Yankulov, K.Y. and Bentley, D.L. 1997. Regulation of CDK7 substrate specificity by MAT1 and TFIIH. *EMBO J* **16**(7): 1638-1646.

Yeo, M., Lin, P.S., Dahmus, M.E., and Gill, G.N. 2003. A novel RNA polymerase II C-terminal domain phosphatase that preferentially dephosphorylates serine 5. *J Biol Chem* **278**(28): 26078-26085.

Yik, J.H., Chen, R., Nishimura, R., Jennings, J.L., Link, A.J., and Zhou, Q. 2003. Inhibition of P-TEFb (CDK9/Cyclin T) kinase and RNA polymerase II transcription by the coordinated actions of HEXIM1 and 7SK snRNA. *Mol Cell* **12**(4): 971-982.

Yoh, S.M., Cho, H., Pickle, L., Evans, R.M., and Jones, K.A. 2007. The Spt6 SH2 domain binds Ser2-P RNAPII to direct Iws1-dependent mRNA splicing and export. *Genes Dev* **21**(2): 160-174.

You, J., Croyle, J.L., Nishimura, A., Ozato, K., and Howley, P.M. 2004. Interaction of the bovine papillomavirus E2 protein with Brd4 tethers the viral DNA to host mitotic chromosomes. *Cell* **117**(3): 349-360.

Yunokuchi, I., Fan, H., Iwamoto, Y., Araki, C., Yuda, M., Umemura, H., Harada, F., Ohkuma, Y., and Hirose, Y. 2009. Prolyl isomerase Pin1 shares functional similarity with phosphorylated CTD interacting factor PCIF1 in vertebrate cells. *Genes Cells* **14**(9): 1105-1118.

Zeitlinger, J., Stark, A., Kellis, M., Hong, J.W., Nechaev, S., Adelman, K., Levine, M., and Young, R.A. 2007. RNA polymerase stalling at developmental control

genes in the Drosophila melanogaster embryo. *Nat Genet* **39**(12): 1512-1516.

Zhang, J. and Corden, J.L. 1991. Identification of phosphorylation sites in the repetitive carboxyl-terminal domain of the mouse RNA polymerase II largest subunit. *J Biol Chem* **266**(4): 2290-2296.

Zhang, J., Yang, P.L., and Gray, N.S. 2009. Targeting cancer with small molecule kinase inhibitors. *Nat Rev Cancer* **9**(1): 28-39.

Zhang, Y., Kim, Y., Genoud, N., Gao, J., Kelly, J.W., Pfaff, S.L., Gill, G.N., Dixon, J.E., and Noel, J.P. 2006. Determinants for dephosphorylation of the RNA polymerase II C-terminal domain by Scp1. *Mol Cell* **24**(5): 759-770.

Zheng, G., Schweiger, M.R., Martinez-Noel, G., Zheng, L., Smith, J.A., Harper, J.W., and Howley, P.M. 2009. Brd4 regulation of papillomavirus protein E2 stability. *J Virol* **83**(17): 8683-8692.

Zhou, K., Kuo, W.H., Fillingham, J., and Greenblatt, J.F. 2009a. Control of transcriptional elongation and cotranscriptional histone modification by the yeast BUR kinase substrate Spt5. *Proc Natl Acad Sci U S A* **106**(17): 6956-6961.

Zhou, M., Deng, L., Lacoste, V., Park, H.U., Pumfery, A., Kashanchi, F., Brady, J.N., and Kumar, A. 2004. Coordination of transcription factor phosphorylation and histone methylation by the P-TEFb kinase during human immunodeficiency virus type 1 transcription. *J Virol* **78**(24): 13522-13533.

Zhou, M., Halanski, M.A., Radonovich, M.F., Kashanchi, F., Peng, J., Price, D.H., and Brady, J.N. 2000. Tat modifies the activity of CDK9 to phosphorylate serine 5 of the RNA polymerase II carboxyl-terminal domain during

human immunodeficiency virus type 1 transcription. *Mol Cell Biol* **20**(14): 5077-5086.

Zhou, M., Huang, K., Jung, K.J., Cho, W.K., Klase, Z., Kashanchi, F., Pise-Masison, C.A., and Brady, J.N. 2009b. Bromodomain protein Brd4 regulates human immunodeficiency virus transcription through phosphorylation of CDK9 at threonine 29. *J Virol* **83**(2): 1036-1044.

Zhou, M., Lu, H., Park, H., Wilson-Chiru, J., Linton, R., and Brady, J.N. 2006. Tax interacts with P-TEFb in a novel manner to stimulate human T-lymphotropic virus type 1 transcription. *J Virol* **80**(10): 4781-4791.

Zhu, Y., Pe'ery, T., Peng, J., Ramanathan, Y., Marshall, N., Marshall, T., Amendt, B., Mathews, M.B., and Price, D.H. 1997. Transcription elongation factor P-TEFb is required for HIV-1 tat transactivation in vitro. *Genes Dev* **11**(20): 2622-2632.

Zippo, A., Serafini, R., Rocchigiani, M., Pennacchini, S., Krepelova, A., and Oliviero, S. 2009. Histone crosstalk between H3S10ph and H4K16ac generates a histone code that mediates transcription elongation. *Cell* **138**(6): 1122-1136.

Zuber, J., Shi, J., Wang, E., Rappaport, A.R., Herrmann, H., Sison, E.A., Magoon, D., Qi, J., Blatt, K., Wunderlich, M. et al. 2011. RNAi screen identifies Brd4 as a therapeutic target in acute myeloid leukaemia. *Nature*.