Structural and functional characterisation of cyclin dependent kinase 1 containing complexes

Svitlana Mykolaivna Korolchuk

Thesis submitted for the degree of Doctor of Philosophy

April 2018

Northern Institute for Cancer Research
Faculty of Medical Sciences
Newcastle University
I. Abstract

CDK1 belongs to the family of serine/threonine kinases that require cyclin binding and phosphorylation for activation. CDK1 is partnered by cyclin A or cyclin B and is the only essential CDK that is required to drive the eukaryotic cell cycle. Because of their roles in driving cell growth and division, the activities of some members of the CDK family are upregulated in different cancer types making them attractive targets for therapy. The structure of monomeric CDK1 bound to CKS1 and in a complex with cyclin B are presented. These structures confirm the conserved inactive monomeric CDK fold and show how it can be remodelled by cyclin binding. These structures have revealed important insights for our understanding of CDK function and regulation as well as for the development of new more potent and specific drugs. A range of biophysical and biochemical techniques have been used to confirm and develop hypotheses as to the regulation and activity of CDK1 proposed by the structures. The melting temperature of CDK1 was estimated and compared to that of the cyclin B bound complex, CDK2 and CDK2-cyclin A to determine if the smaller interface between CDK1 and cyclin B may have implications for protein stability. The CDK1–cyclin B structures also reveal potential novel protein interaction sites that might regulate CDK1 activity. These findings inspired further investigations to explore binding sites by structural biology methods. CDK1 can phosphorylate the largest number of substrates in comparison to other CDKs. This relaxed substrate specificity was explored with reference to the insights provided by the CDK1-cyclin B-CKS2 structure. The similarities in sequence and secondary structure exhibited by the CDK family pose challenges for inhibitor design. The binding modes of a diverse set of inhibitors were characterized by a range of biophysical methods and their structures were determined by X-ray crystallography. The results of these experiments highlight the importance of conformational plasticity of the whole molecule in shaping the ATP binding site.
II. Acknowledgements

This thesis would not have been possible without valuable involvement of the number of people.

First of all I would like to express my deepest gratitude to my main mentor and supervisor Professor Jane Endicott who has given me a job and allowed to take it further to a next level of development and do PhD project. She has been a very knowledgeable and helpful supervisor through the years teaching me different aspects of cell cycle, molecular biology and biochemistry in general. Her endless support, encouragement and patience were vital for my progression and process of learning.

I would also like to thank my second supervisor Martin Noble for his useful advice in a field of structural biology and for his optimistic approach to science and life which makes the difference. I am very grateful for both of them for reading this thesis and providing helpful feedback.

I am grateful to my advisory panel Prof. Herbie Newell and Dr. Paula Salgado for their advice, expertise and guidance. The very special thanks has to be said to Nicholas Brown for introducing me to the CDK1 crystallography project and generated some constructs and initial data. This work has benefited enormously from involvement of Mat Martin and later on Daniel Wood to whom I am very grateful. Big thank you is going to Arnaud Basle for his help with crystal data collection, structure solution and his guidance. Also I want to thank my ex-colleagues Dr. Will Stanley and Dr Ruslan Moukhametzianov who participated in this project.

I want to say thank you for all of the lab members for their support, advice and cheerfully stimulating atmosphere. Especially to Claire Jennings for help with different technics. Her Yorkshire steady attitude to anything she does always sets me into right direction and implements in brilliant results. I am very thankful to Richard Heath and Martina Pastok for useful discussions and advises, to Lan Wang for work she has done for this project and for everyone in a lab for their friendship and ideas.

On a more personal note I would like to thank my family for all their support and patience.
III. Table of Contents

I. ABSTRACT ............................................................................................................. I
II. ACKNOWLEDGEMENTS ......................................................................................... II
III. TABLE OF CONTENTS ............................................................................................ III
IV. LIST OF FIGURES ..................................................................................................... VIII
V. LIST OF TABLES ......................................................................................................... XII
VI. ABBREVIATIONS .................................................................................................... XIII

CHAPTER 1: INTRODUCTION ......................................................................................... 1

1.1 REGULATION OF THE EUKARYOTIC CELL CYCLE .............................................. 1
  1.1.1 Importance of checkpoints in cell cycle regulation ...................................... 3
  1.1.2 CDKs and the cell cycle .............................................................................. 4
  1.1.3 Cyclins as partners that activate CDKs ...................................................... 7
  1.1.4 Role of phosphorylation in CDK activation .............................................. 13
1.2 STRUCTURAL CHARACTERISATION OF THE CDK FAMILY ................................. 14
  1.2.1 A structural paradigm for CDK activation .............................................. 14
  1.2.2 Structural characterisation of the cyclin family ...................................... 17
  1.2.3 CDK requirements for substrate selection .............................................. 20
1.3 OTHER REGULATORS OF CDK ACTIVITY ............................................................ 23
1.4 CDKS AND CANCER .............................................................................................. 26
  1.4.1 CDK4/6 and cancer .................................................................................. 27
  1.4.2 CDK2 and cancer ..................................................................................... 29
  1.4.3 Role of CDK9 in cancer ........................................................................... 31
  1.4.4 CDK1 and cancer ..................................................................................... 32
  1.4.5 Role of pharmacological CDK inhibition for cancer therapy .................. 32
1.5 AIMS OF THE THESIS ........................................................................................... 34

CHAPTER 2: STRUCTURAL INSIGHTS INTO CDK1-CONTAINING COMPLEXES ............... 35

2.1 INTRODUCTION .................................................................................................... 35
  2.1.1 Discovery of CDK1 ................................................................................... 35
  2.1.2 Regulation of CDK1 activity .................................................................... 38
  2.1.3 Structures of CDK-cyclin complexes ...................................................... 40
2.1.4 Structural features of cyclins A and B ................................................................. 42
2.1.5 Analysis of the CDK-cyclin interface ................................................................. 45
2.1.6 Aims and objectives ............................................................................................. 47
2.2 MATERIALS AND METHODS ................................................................................... 48
  2.2.1 Chemicals and Reagents ..................................................................................... 48
  2.2.2 Purification buffers ............................................................................................. 48
  2.2.3 Constructs used for in vitro experiments ............................................................... 49
  2.2.4 Cell culture ......................................................................................................... 50
  2.2.5 Optimisation of protein expression ...................................................................... 50
  2.2.6 Expression tests .................................................................................................. 51
  2.2.7 Bacterial transformation and protein expression in bacterial cells ...................... 51
  2.2.8 Protein purification ............................................................................................. 52
  2.2.9 Protein purification by size-exclusion chromatography ...................................... 53
  2.2.10 SDS-PAGE ...................................................................................................... 53
  2.2.11 Preparation of complexes ................................................................................ 54
  2.2.12 Protein crystallography .................................................................................... 54
  2.2.13 Data collection and analysis ............................................................................ 54
2.3 RESULTS AND DISCUSSION ..................................................................................... 55
  2.3.1 Purification of CDK1, cyclin B and CKS2 and complex assembly ....................... 55
  2.3.2 Structure determination of CDK1-containing complexes ...................................... 59
  2.3.3 Structural details of the CDK1-CKS1 complex .................................................... 62
  2.3.4 Comparison of the CDK1-CKS1 and CDK2-CKS1 structures ............................... 63
  2.3.5 The overall fold of the CDK1-cyclin B-CKS2 complex ....................................... 66
  2.3.6 Comparison of CDK1-cyclin B-CKS2 with CDK2 and CDK4 containing structures .................................................................................................................. 68
  2.3.7 Attempts to determine structures for CDK1-cyclin A and CDK1-cyclin A-CKS2 .................................................................................................................. 70
2.4 CONCLUSIONS ........................................................................................................... 76

CHAPTER 3: FUNCTIONAL CHARACTERISATION OF CDK1 ........................................ 78
3.1 INTRODUCTION ......................................................................................................... 78
  3.1.1 Observations from the cellular studies ................................................................. 78
  3.1.2 Cyclins as CDK1 and CDK2 partners ................................................................. 79
  3.1.3 Sequential comparison of cyclins binding to CDK1 and CDK2 ............................. 83
  3.1.4 CDK substrates ................................................................................................. 84
4.1.11 Aims and objectives .................................................................................................................. 119
4.2 DESCRIPTION OF BIOPHYSICAL METHODS ............................................................................. 119
  4.2.1 Theoretical background of surface plasmon resonance .......................................................... 119
4.3 MATERIALS AND METHODS ....................................................................................................... 121
  4.3.1 Generation of constructs .......................................................................................................... 121
  4.3.2 E. coli transformation ............................................................................................................... 123
  4.3.3 Baculovirus preparation using the MultiBac expression system .............................................. 124
  4.3.4 Insect cell culture and virus amplification .............................................................................. 124
  4.3.5 Expression tests ...................................................................................................................... 124
  4.3.6 Western blotting .................................................................................................................... 125
  4.3.7 Surface plasmon resonance (SPR) ......................................................................................... 126
4.4 RESULTS ...................................................................................................................................... 126
  4.4.1 Generation of Speedy/Ringo A constructs .............................................................................. 126
  4.4.2 Analysis of Ringo A/Speedy affinity to CDK1 and CDK2 ....................................................... 140
  4.4.3 RGC32 binding to CDK1 ......................................................................................................... 141
4.5 CONCLUSIONS ............................................................................................................................. 146

CHAPTER 5: CDK INHIBITION BY ATP-COMPETITIVE INHIBITORS ............................................. 147

5.1 INTRODUCTION ............................................................................................................................ 147
  5.1.1 Role of CDK inhibition for cancer therapy .............................................................................. 147
  5.1.2 ATP binding site .................................................................................................................... 151
  5.1.3 CDK inhibition, exploring the ATP-binding site .................................................................... 153
  5.1.4 Aims of the project ............................................................................................................... 156
5.2 MATERIALS AND METHODS ..................................................................................................... 157
  5.2.1 Protein purification ............................................................................................................... 157
  5.2.2 Preparation of inhibitors ....................................................................................................... 157
  5.2.3 Differential scanning fluorimetry (DSF) ............................................................................... 158
  5.2.4 Isothermal Titration Calorimetry ........................................................................................... 158
  5.2.5 Surface plasmon resonance ................................................................................................ 159
  5.2.6 ADP-Glo™ kinase assay ....................................................................................................... 159
  5.2.7 Protein crystallography ......................................................................................................... 160
5.3 RESULTS AND DISCUSSION ....................................................................................................... 161
  5.3.1 Description of the inhibitors chosen for the study ................................................................. 161
    5.3.1.1 Flavopiridol (Alvocidib) .................................................................................................. 161
    5.3.1.2 AZD5438 ....................................................................................................................... 161
IV. List of Figures

Figure 1-1 Regulation of cell cycle progression. .............................................................. 2
Figure 1-2 Sequence alignment of members of the CDK family. ............................... 5
Figure 1-3 Evolutionary relationships between different mammalian CDK subfamilies. ......................................................................................................................... 7
Figure 1-4 Oscillation of cyclin concentration during the cell cycle. ......................... 8
Figure 1-5 Cell cycle regulation by cyclin dependant kinases (CDKs) and their role in transcription. .................................................................................................................... 11
Figure 1-6 Model for CDK activation in Metazoans.................................................... 14
Figure 1-7 Structural changes during CDK2 activation. ............................................. 16
Figure 1-8 Typical cyclin structure. ............................................................................ 18
Figure 1-9 Structural and topological analysis of human cyclin K, S. pombe cyclin C and human cyclin H. .................................................................................................................... 19
Figure 1-10 CDC6 peptide bound to the recruitment site of pCDK2-cyclin A complex. ................................................................................................................................. 21
Figure 1-11 Possible position of the peptide linker between the catalytic site and the RXL recruitment site. .................................................................................................................... 22
Figure 1-12 Sequence overlay of Human CKS1 and CKS2, S. pombe Suc1 and S. cerevisiae Cks1. ..................................................................................................................... 24
Figure 1-13 Crystal structure of monomeric Suc1 and the β-interchanged dimer conformations. .................................................................................................................... 25
Figure 1-14 Crystal structure of a CDK2-CKSHs1 complex........................................ 25
Figure 1-15 Hallmarks of cancer and therapeutic approaches to its targeting (Hanahan and Weinberg, 2011). .................................................................................................................... 27
Figure 1-16 Deregulation of members of CDK4/6 pathway leading to cancer (Otto and Sicinski, 2017). .................................................................................................................... 28
Figure 1-17 Deregulation of cyclins E1 and E2 leading to cancers (Otto and Sicinski, 2017). .............................................................................................................................. 30
Figure 2-1 Comparison of CDK1 protein sequences from the different species ...... 37
Figure 2-2 Regulation of CDK1 activity by positive and double negative activation loops. .............................................................................................................................. 39
Figure 2-3 Regulation of CDK1 activity at the G2/M transition ..................................... 40
Figure 2-4 The structures of pCDK2-cyclinA, pCDK9-cyclinT, pCDK4-cyclinD ...... 41
Figure 2-5 Comparison of structures of cyclins A, B and E.................................... 43
Figure 3-11 Illustration of the ADP-GloTM assay................................. 99
Figure 3-12 Kinase activity of pCDK1-cyclin B and pCDK2-cyclin A towards the panel of peptides. ........................................................................................................... 100
Figure 3-13 Comparison of CDK1 and CDK2 kinase activity after λ phosphatase treatment................................................................. 101
Figure 4-1 Sequence comparison between members of the Speedy A family from different species................................................................. 104
Figure 4-2 RINGO/Speedy protein family, .......................................................................................................................... 106
Figure 4-3 Comparison of CDK activation by RINGO/Speedy and cyclins. .......... 108
Figure 4-4 Identification of primary and secondary structure of similarities of Speedy A with another proteins. ........................................ 110
Figure 4-5 Overlay of cyclin A with cyclin 10.......................................................... 110
Figure 4-6 Secondary structure comparison of Spy1 and cyclin A..................... 112
Figure 4-7 Comparison of CDK2 secondary structures ................................ 113
Figure 4-8 Activation segments comparison. ...................................................... 114
Figure 4-9 Interactions involved in activation segment positioning. .............. 115
Figure 4-10 Sequence comparison between RGC32 from the different species. ... 118
Figure 4-11 Schematic representation of surface plasmon resonance. .......... 120
Figure 4-12 Protein sequence of Speedy A. ......................................................... 127
Figure 4-13 Sequences of Speedy1/Ringo A cloned into pGEX6P-1 vector........ 128
Figure 4-14 SDS-PAGE analysis of test expression of Speedy A constructs in different E. coli strains in different media.................................................... 129
Figure 4-15 SDS-PAGE gel and Western blotting analysis of Speedy A/Ringo expression test........................................................................ 131
Figure 4-16 Purification of Speedy A (1-313) by size exclusion chromatography... 132
Figure 4-17 Alignment of MBP sequence from pMAL-5x vector with sequence pET21dMBP3C......................................................................................... 133
Figure 4-18 Graphical representation of MBP and Speedy A2 constructs cloned into pET-21d (+)................................................................. 134
Figure 4-19 SDS-PAGE analysis of test expression of Speedy constructs........ 136
Figure 4-20 DH10MultiBacTM and DH10EmBacY baculoviral DNA..................... 138
Figure 4-21 Pull downs of Ringo A/Speedy constructs....................................... 139
Figure 4-22 Investigation of binding between Speedy 64-189 and Speedy 64-206 to CDK2. ................................................................. 141
Figure 4-23 Purified proteins used for SPR experiments................................. 142
Figure 4-24 Analysis of the interactions between CDK1/2 and binding partners by surface plasmon resonance (SPR) ................................................................. 143
Figure 4-25 Sensograms representing kinetic interactions by SPR .................. 145
Figure 4-26 Sensograms representing kinetic interactions by SPR .................. 146
Figure 5-1 Structure (A) and activity (B) of pan- or multitarget- CDK inhibitors .. 149
Figure 5-2 Structure (A) and activity (B) of CDK4 and CDK6 inhibitors ............ 150
Figure 5-3 Conserved residues of ATP binding site in CDK1 and CDK2 ............ 152
Figure 5-4 ATP binding site ........................................................................ 152
Figure 5-5 Type I and Type II kinase inhibitors .......................................... 154
Figure 5-6 Residues that can be exploited by ATP competitive inhibitors ....... 155
Figure 5-7 Comparison of the inhibitor binding profiles of different CDKs in their monomeric and cyclin bound forms ............................................... 156
Figure 5-8 Structures and names of the inhibitors chosen for the study ............ 163
Figure 5-9 Comparison of melting temperatures of pCDK1, pCDK2, pCDK1-cyclin B and pCDK2-cyclin A ................................................................. 164
Figure 5-10 Difference of ΔTm between pCDK1 and pCDK2 and pCDK1-cyclin B and pCDK2-cyclin A ................................................................. 166
Figure 5-11 ITC to measure the affinity of CDK1 and CDK2 to a set of ATP- competitive inhibitors ................................................................. 167
Figure 5-12 Comparison of dissociation constant (Kd) .................................. 168
Figure 5-13 Difference between pCDK1 and pCDK2 and their complexes measured by DSF and ITC ................................................................. 169
Figure 5-14 Sensogram (A) and Response-Concentration plot (B) of AZD5438 binding to GST–CDK1 ................................................................. 170
Figure 5-15 Sensogram (A) and Response-Concentration plot (B) of AZD5438 binding to GST–CDK2 ................................................................. 171
Figure 5-16 Dissociation constant (Kd) of pCDK1 and pCDK2 bound to inhibitors measured by SPR ................................................................. 171
Figure 5-17 Comparison of dissociation constant (Kd) measured by ITC and SPR. 172
Figure 5-18 SDS PAGE of purified complexes CDK1 – CKS2, CDK1 – cyclin B, CDK2 – cyclin A ................................................................. 174
Figure 5-19 Images of crystals used for structure determination .................... 175
Figure 5-20 Structures of AZD5438, Flavopiridol (Alvocidib) and CGP74514A .... 176
Figure 5-21 Overlay of AZD5438, Flavopiridol and CGP 74514A in the ATP binding pocket of CDK1 bound to cyclin B and CKS2 ....................... 177
Figure 5-22 ATP binding pocket of CDK1 in complex with cyclin B and CKS2 occupied by AZD5438.................................................................................................................. 178
Figure 5-23 ATP binding pocket of CDK1 in complex with cyclin B and CKS2 occupied by Flavopiridol .............................................................. 179
Figure 5-24 ATP binding pocket of CDK1 in complex with cyclin B and CKS2 occupied by CGP74514A .............................................................. 180
Figure 5-25 Changes in P-loop position ...................................................... 181
Figure 5-26 ATP binding pocket of CDK1 in complex with CKS2 occupied by Dinaciclib .................................................. 182
Figure 5-27 ATP binding pocket of CDK1 in complex with CKS2 occupied by AZD5438 .............................................................. 184
Figure 5-28 Changes in P-loop position ...................................................... 185
Figure 5-29 Comparison of ATP binding pocket of CDK1 and CDK1-cyclin B occupied by AZD5438 .................................................. 186
Figure 5-30 Comparison of ATP binding pocket of CDK2 and CDK2-cyclin A occupied by AZD5438 .............................................................. 187
Figure 5-31 Overlay of CDK1 and CDK2 structures bound to AZD5438 .... 188

V. List of Tables

Table 1-1 Established and emerging roles of cyclins .................................................. 10
Table 1-2 CDKs of most studied biological models and cyclins activating them at different stages of the cell cycle .......................................................... 12
Table 2-1 Summary of constructs used to express CDK1 and associated regulators .......................................................... 50
Table 2-2 Summary of CDK1-CKS2 and CDK1-cyclinB-CKS2 X-ray data collection and refinement statistics .................................................. 61
Table 3-1 Sequences of p107 mutant peptides .................................................. 89
Table 3-2 Molecular weights and pIs of p107 mutants .................................. 90
Table 3-3 kcat /Km, peptide (s-1) for CDK-cyclin complexes on p107 mutants ..... 100
Table 4-1 Table of primers .............................................................................. 123
Table 4-2 Ringo/Speedy A constructs .......................................................... 128
Table 4-3 Table of constructs with molecular weights .................................. 137
Table 4-4 SpeedyA2 constructs for expression in insect cells ...................... 139
VI. Abbreviations

3C Human Rhinovirus (HRV) 3C Protease
ADP Adenosine di-phosphate
APC/C Anaphase-promoting complex
ATP Adenosine tri-phosphate
BSA Bovine serum albumin
CAK CDK activation kinase
CCP4 Collaborative Computational project 4
CDK Cyclin Dependent Kinase
CDKI Cyclin Dependent kinase inhibitor
Cks Cdc kinase subunit
C-terminal Carboxyl-terminal
Cv column volume
Da Dalton
DNA Deoxyribonucleic acid
dNTP Deoxyribonucleotide triphosphate
DFS  Differential Scanning Fluorimetry
DTT  Dithiothreitol
E.coli  *Escherichia coli*
EDTA  Ethylene Diamine Tetra-acetic Acid
FL  Full Length
GST  Glutathione-S-transferase
HBS  Hapes Buffered Saline
HEPES  4-(2-Hydroxyethyl)piperazine-1-ethanesulphonic acid
Hisx6  Hexahistidine tag
HPLC  High Performance Liquid Chromatography
IPTG  Isopropyl β-D-1-thiogalactoside
ITC  Isothermal Titration Calorimetry
K_{off}  off rate
K_{on}  on rate
K_d  binding constant
kDa  KiloDalton
LB  Lysogeny Broth
M  Molar
MBP  Maltose binding protein
MES  2-(N-morpholino)ethanesulfonic acid
MTG  Monothioglycerol
Ni^{2+}NTA  Nickel-nitriacetic acid
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>N-Terminal</td>
<td>Amino-terminal</td>
</tr>
<tr>
<td>OD&lt;sub&gt;600&lt;/sub&gt;</td>
<td>Optical density at λ=600</td>
</tr>
<tr>
<td>PAGE</td>
<td>Polyacrylamide Gel Electrophoresis</td>
</tr>
<tr>
<td>PBS</td>
<td>Phosphate buffered saline</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase chain reaction</td>
</tr>
<tr>
<td>PDB</td>
<td>Protein Data Bank</td>
</tr>
<tr>
<td>PEG</td>
<td>Polyethylene glycol</td>
</tr>
<tr>
<td>Rb</td>
<td>Retinoblastoma protein</td>
</tr>
<tr>
<td>RNA</td>
<td>Ribonucleic acid</td>
</tr>
<tr>
<td>Rpm</td>
<td>revolutions per minute</td>
</tr>
<tr>
<td>RT</td>
<td>Room Temperature</td>
</tr>
<tr>
<td>SDS</td>
<td>Sodium dodecyl sulphate</td>
</tr>
<tr>
<td>SEC</td>
<td>Size Exclusion Chromatography</td>
</tr>
<tr>
<td>Sf9</td>
<td>Spodoptera frugiperda 9</td>
</tr>
<tr>
<td>SOC</td>
<td>Super optimal broth with catabolic repressor</td>
</tr>
<tr>
<td>SPR</td>
<td>Surface plasmon resonance</td>
</tr>
<tr>
<td>TAE</td>
<td>Tris-acetate-EDTA</td>
</tr>
<tr>
<td>TBS-T</td>
<td>Tris-buffered saline with 0.1% Tween 20</td>
</tr>
<tr>
<td>TCEP</td>
<td>Tris(2-carboxyethyl)phosphine</td>
</tr>
<tr>
<td>T&lt;sub&gt;m&lt;/sub&gt;</td>
<td>Melting temperature</td>
</tr>
<tr>
<td>UV</td>
<td>Ultraviolet</td>
</tr>
<tr>
<td>WT</td>
<td>Wild-type</td>
</tr>
<tr>
<td>Amino acids</td>
<td></td>
</tr>
<tr>
<td>----------------------</td>
<td>---</td>
</tr>
<tr>
<td>Alanine</td>
<td>A</td>
</tr>
<tr>
<td>Arginine</td>
<td>R</td>
</tr>
<tr>
<td>Asparagine</td>
<td>N</td>
</tr>
<tr>
<td>Aspartic acid</td>
<td>D</td>
</tr>
<tr>
<td>Cysteine</td>
<td>C</td>
</tr>
<tr>
<td>Glutamic acid</td>
<td>E</td>
</tr>
<tr>
<td>Glutamine</td>
<td>Q</td>
</tr>
<tr>
<td>Glycine</td>
<td>G</td>
</tr>
<tr>
<td>Histidine</td>
<td>H</td>
</tr>
<tr>
<td>Isoleucine</td>
<td>I</td>
</tr>
</tbody>
</table>
Chapter 1: Introduction

1.1 Regulation of the eukaryotic cell cycle

The cell cycle is the series of cellular events that leads to duplication of genetic material and cell division. This process is the only way to fulfil the reproduction and development of unicellular and multicellular organisms. The oscillating activity of the Cycle Dependant Kinases (CDKs) acts as a master regulator of cell cycle progression in eukaryotes. Since a number of seminal discoveries were made in the 1970s and 1980s using cell lines and model organisms (Hartwell et al., 1970; Johnson and Rao, 1970; Rao and Johnson, 1970; Nurse et al., 1976; Nurse and Thuriaux, 1980; Standart et al., 1987; Hartwell and Weinert, 1989) enormous progress has been achieved in this area. This work lead to an appreciation of the organisation of the cell cycle and the mechanisms that regulate it. For example, an emphasis on the existence of checkpoints and surveillance mechanisms (Weinert and Hartwell, 1988; Hartwell and Weinert, 1989); the activities of different cyclins that can activate CDKs at different stages of the cell cycle to ensure ordered progression (Enoch and Nurse, 1990; Surana et al., 1991) and finally the importance of quantitative oscillations of CDK activity (Surana et al., 1991; Stern and Nurse, 1996).

The cell cycle has five distinct phases (Morgan, 2007). During G1 and G2 phases cells grow and accumulate cellular components, during S phase replication of DNA takes place and during mitosis the cell divides. After mitosis, the cell can continue to proliferate or enter quiescence (G0 phase) (Figure 1-1). Prior to cell cycle entry, mitogenic pathways are triggered by growth factors. The well characterised pathways triggering the cell cycle are mitogen-stimulated protein kinase (MAPK) and phosphatidylinositol-4,5-bisphosphate 3-kinase (PI3K) pathways (Balmanno and Cook, 1999). ERK1/2 is also important as it promotes transcription of D-type cyclins (Balmanno and Cook, 1999). Cyclin D activates CDK4 and CDK6 for phosphorylation of various substrates, most importantly retinoblastoma tumour suppressor protein (Rb) as well as p107 and p130 that collectively bind and repress E2F transcription factors. Rb family phosphorylation relieves E2F inhibition to activate genes important for G1 to S transition, DNA replication, chromatin structure and the spindle assembly checkpoint (SAC). CDK4/6 also bind to CDK inhibitors (CKIs) that prevent activation of CDK2-cyclin E. Two families of CDK inhibitors regulate the cell cycle, the INK4 and Cip/Kip inhibitors (Pavletich, 1999; Jeffrey et al., 2000).
Figure 1-1 Regulation of cell cycle progression.

Mitogenic signals lead to transcription of cyclin D isoforms that form complexes with CDK4 and CDK6 which activate by phosphorylation a number of substrates including Rb. Phosphorylation of Rb leads to activation of the family of E2F transcription factors. Growth inhibitory signals upregulate CIP/KIP and INK4 inhibitors that control progression from one phase to the next. Progress through S phase and G2 phase is controlled by CDK2 complexes together with kinases of the Aurora family and polo-like kinases (PLK). DNA damage is recognised by checkpoint kinases 1 and 2 (CHK1 and 2) with p53 involved. Kinases and cyclins are identified by different colours, grey ovals identify negative regulators of cell cycle. Red P means active phosphorylation, grey P means dephosphorylation. CDC25, cell division cycle 25. Picture adapted from (Otto and Sicinski, 2017).

Proteins of the INK4 family, p16INK4a, p15INK4b, p18INK4c and p19INK4d inhibit mainly members of the CDK4 subfamily and bind preferentially to monomeric CDKs resulting in stabilisation of molecule which decreases CDK affinity for their cyclin partners and ATP (Jeffrey et al., 2000). Proteins of the CIP/KIP family, p21CIP1, p27KIP1 and p57KIP2, in contrast bind to CDKs in complex with their respective cyclins offering another level of CDK regulation (Pavletich, 1999). Activation of E2F1 and FOXM1 by CDK4/6 leads to the increased expression of cyclins E1 and E2 which leads to CDK2 activation. At this point the cell has passed the restriction point and further progress through the cell cycle is no longer dependent on the continuous presence of growth factors. E2F1 also stimulates expression of cyclin A in S phase which becomes the regulator of CDK2 in S phase. The important CDK2 substrates that promote S phase entry are Rb and NPAT and to terminate it – CDC6 and E2F1. Cyclin A also binds CDK1 in late S/G2 stage activating transcription factors FOXM1 and FOXK2.
which are involved in mitotic progression (Laoukili et al., 2008; Sadasivam et al., 2012). Cyclin A is degraded prior to anaphase (Furuno et al., 1999). Cyclin B accumulation at the G2-M transition mainly occurs in the nucleus as CDK1 has a number of nuclear substrates that lead to nuclear re-organisation and fragmentation and also nuclear envelope breakdown (Peter et al., 1990). CDK1 also has a number of targets in the cytoplasm. CDK1 in contrast to CDK2, CDK4 and CDK6 has a different inhibitory mechanism. It is kept in the inactive state mainly by inhibitory phosphorylation by WEE1 and MYT1 kinases. Inhibitory phosphorylation is removed when the cell is ready to enter mitosis by phosphatase CDC25 (Figure 1-1). Polo-like kinase (PLK-1) and Aurora kinases are also involved in G2 phase progression.

1.1.1 Importance of checkpoints in cell cycle regulation

Surveillance mechanisms existing in cells have been named checkpoint pathways (Hartwell and Weinert, 1989; Elledge, 1996). They monitor progression through the cell cycle though the relative importance of different pathways and their ability to impact cell survival differs. For example, in budding yeast, due to the organisation of their cell cycle, where S phase overlaps with M phase, cells regulate the cell cycle primarily at START (Li and Murray, 1991). Consequently if DNA replication advances slower than normal due to DNA damage, mitosis starts anyway before completion (Lengronne and Schwob, 2002). In contrast, the organisation of the cell cycle in fission yeast has distinctive G1 and G2 phases and is primarily controlled by the G2/M checkpoint (Hartwell and Weinert, 1989).

Mammalian cells have rigorous regulatory networks in place to ensure that progression of the cell cycle occurs with a minimal number of errors. One well understood pathway is the DNA damage checkpoint which arrests cells in G1 transition in response to damaged DNA. Depending on the type of damage, ataxia telangiectasia and Rad3-related (ATR) or ataxia telangiectasia mutated (ATM) protein kinases phosphorylate and activate checkpoint kinases 1 or 2 (CHK1, CHK2) (Zhao and Piwnica-Worms, 2001) can activate p53 (Matsuoka et al., 2000) which in its turn induces expression of the CDK2 inhibitor p21CIP1 and leads to G1 arrest (Figure 1-1). Activation of CHK1 can also lead to S/G2 arrest by inactivation of CDC25A, CDC25B and CDC25C (Sanchez et al., 1997) and enhanced activation of WEE1. As a result of WEE1 and CDC25 phosphorylation, inactive CDK1 or CDK2 would remain phosphorylated on the inhibitory residue Thr15 which would lead to G2 arrest (O’Connell et al., 1997). Growth
inhibitory signals can lead to inhibition of CDK2 and CDK4/6 which can cause G1 or S phase arrest.

Another important mitotic checkpoint is the spindle assembly check point (SAC). This sophisticated pathway includes 15 proteins known to date (Musacchio, 2015) and the main function of it is to monitor correct sister chromatid cohesion to prevent premature chromosome segregation. Without it, cells enter anaphase prematurely which causes problems with chromosome segregation and can lead to errors in copy numbers (Dobles et al., 2000; Musacchio, 2015). The SAC is also very important for mitotic exit as it affects the activity of the mitotic checkpoint complex (MCC) which targets APC/C. By inhibiting APC/C, the MCC stabilising its components preventing transition from metaphase to anaphase and the cell leaving mitosis (Chang and Barford, 2014).

1.1.2 CDKs and the cell cycle

CDKs are a family of protein kinases that phosphorylate their substrates on serine/threonine residues. CDKs are regulated by their partner cyclins which control kinase activity and substrate specificity. The important feature responsible for the activation of the CDK family members is a catalytic core which comprises the ATP binding pocket, the activation segment (also called the T-loop) and the cyclin binding surface. The ATP binding pocket is highly conserved (Figure 1-2). For full activation of most CDK-cyclin complexes, phosphorylation by CDK activating kinase (CAK) on a conserved threonine residue (Thr161 in CDK1) within the activation segment (T-loop) is also required (Fisher, 2005). Once active, CDKs mainly have preferences for substrates containing the consensus sequence S/T-P-X-K/R (single letter amino acid abbreviations). Most CDKs also possess inhibitory phosphorylation sites (Thr14 and Tyr15 in CDK1) from the P-loop which interfere with the proper alignment of ATP in the ATP binding site (Mueller et al., 1995). CDK-cyclin complexes are also regulated by CDK inhibitors (CKIs) that can execute checkpoints to stop cell cycle progression if conditions are unfavourable (Morgan, 2007).
Figure 1-2 Sequence alignment of members of the CDK family.
Represented sequences of human CDK2 (Uniprot entry P24941), human CDK1 (Uniprot entry P06493), human CDK4 (Uniprot entry P11802), human CDK6 (Uniprot entry Q00534), human CDK9 (Uniprot entry P50750). P-loop identified with a dark blue bar, αC-helix identified with the gold bar, hinge region is identified with coral bar, activation segment is identified with red bar, DFG motif is coloured pink, APE (in CDK2) motif is coloured magenta, phosphorylation site is identified red. Sequence identities for the examples shown here range from 64% between CDK1 and CDK2 to 32% between CDK1 and CDK9.
The human CDK family is composed of twenty protein kinases (Malumbres, 2014) that play roles in regulation of the cell cycle, gene transcription, protein cellular localisation and protein degradation. CDKs have undergone significant evolutionary divergence and specialisation (Figure 1-3) with their numbers increasing during evolution. For example, yeasts have up to 6 CDKs, whereas flies and Echinodermata contain 11. Their structure has also evolved reflecting adaptations for wider function. Mammalian CDKs range in size from 250 amino acids to proteins that are 1500 residues in length with carboxy- and amino- terminal extensions of variable length and locations.

CDKs can be classified into two large groups, CDKs that have a role in cell cycle promotion and CDKs that regulate transcription. The first group consists of CDKs 1-4 and CDK6. CDKs that belong to the second group – CDK5, CDK8, CDK9, CDK11, CDK12, and CDK13 are mainly involved in transcription regulation. CDK7 is unique as, together with its partners cyclin H and MAT 1, it has roles both in cell cycle regulation (by phosphorylating and activating other kinases) and in initiation of transcription (by phosphorylation of the large subunit of RNA polymerase II) as well as in Nucleotide Excision Repair (NER) (Fisher, 2005).

Regulation of CDK activity is complex and depends on different factors. For example, allosteric regulators including cyclins can influence substrate specificity (Loog and Morgan, 2005; Bloom and Cross, 2007), CKIs can act as positive or negative regulators, and protein degradation via the ubiquitin proteasome pathway helps to control the expression of regulators and substrates (Cheng et al., 1999). Regulation also depends on the timing of substrate accumulation, relative affinities for the different CDKs (Loog and Morgan, 2005) and also intracellular location.
Figure 1-3 Evolutionary relationships between different mammalian CDK subfamilies.
CDKs that primarily regulate the cell cycle are drawn in the orange part of the figure, the ones that are involved in transcription are highlighted by the green background. The name of the sub-family is depicted in bold. The kinase catalytic domain is identified as a red box, proline rich region - green, arginine/serine rich - blue, glutamic acid rich – lilac, glutamine rich – grey. Figure used with permission of (Malumbres, 2014).

1.1.3 Cyclins as partners that activate CDKs
Cyclins were first discovered in embryos of marine invertebrates and amphibians and given the name because of their cyclical accumulation and destruction through the cell cycle (Evans et al., 1983; Murray and Kirschner, 1989). Although it was subsequently discovered that some cyclins do not oscillate. For example, cyclin D accumulates in G1 phase and remains present throughout the cell cycle (Matsushime et al., 1994) (Figure 1-4).
Clarification as to how CDK controls G1-S phase transition in budding yeast came with the discovery of the Cln1-Cln3 cyclins (Nash et al., 1988; Richardson et al., 1989), followed by the identification of Clb1-Clb4 which are required for mitosis and are more similar to metazoan cyclins (Ghiara et al., 1991). Cyclins Clb5 and Clb6 are required for the correct timing of the onset of DNA replication in S phase (Epstein and Cross, 1992; Schwob and Nasmyth, 1993). The requirement for different cyclins at different stages of the cell cycle in yeasts where only one CDK is available suggests that cyclins play a significant role in substrate recognition. It has been shown that budding yeast cyclins modulate the active site of CDK and change its substrate specificity. A model has been proposed according to which binding of different cyclins to CDK1 determine the strength of its activity and direct it to substrates required at specific stages of the cell cycle (Koivomagi et al., 2011). A number of experiments have been performed in budding and fission yeasts to identify the importance of cyclins at different stages of the cell cycle. The main conclusion that emerges is that G1 and S-phase cyclins are replaceable and that their removal by gene deletion can be rescued by mutation or introduction of other genes encoding cyclins. For example, deletion of the CLN1-3 genes in budding yeast leads to cell death and arrest in G1 but can be rescued by deletion of SIC1 (a CDC28 inhibitor) and by ectopic expression of CLB5, a mitotic cyclin (Richardson et al., 1989; Epstein and Cross, 1992). In fission yeast mutation of the G1/S-phase genes puc1, cig1 and cig2 that collectively encode the S. pombe G1 cyclins can lead to delay in G1 but can be rescued by deletion of rum1 (a negative
regulator of CDK1) (Martin-Castellanos et al., 2000). Although deletion of the mitotic cyclin gene *cdc13* prevents cells from entering mitosis (Hagan et al., 1988). In budding yeasts none of the mitotic cyclins are essential but disruption of *CLB1-3, CLB2-4, or CLB1-4* genes together leads to cell death (Fitch et al., 1992).

Cyclins are very diverse in their sequences and are mainly characterised by the presence of a region comprising around 100 amino acids called the cyclin box. The cyclin box encodes a compact fold (the cyclin box fold, CBF) formed of five α-helices that plays a significant role in CDK interactions (Minshull et al., 1989; Pines and Hunter, 1989). Most cyclins have two CBFs but only the N-terminal one interacts with the CDK partner, the C terminal CBF is required for the proper folding of the molecule. Sequences outside the cyclin box allow functional diversity and differences in regulation. The cyclin family comprises 29 proteins involved in the control of a variety of functions (Table 1-1). They are divided into 16 sub-families and three major groups. The first cyclin B group contains cyclins A, B, D, E, F, G, J, I and O. The second, the cyclin Y group includes partners of the CDK5 subfamily; and the third cyclin C group, cyclins C, H, K, L and T (Ma et al., 2013; Cao et al., 2014).
<table>
<thead>
<tr>
<th>Cyclin</th>
<th>Established function</th>
<th>Emerging function</th>
<th>Kinase activity for emerging function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cyclin A</td>
<td>Control of M phase of cell cycle in complex with CDK2 or CDK1</td>
<td>FoxM1 and FoxK2 transcription in complex with CDK2</td>
<td>yes</td>
</tr>
<tr>
<td>Cyclin B</td>
<td>Control of S phase of cell cycle in complex with CDK1</td>
<td>FoxM1 and FoxK2 transcription in complex with CDK1</td>
<td>yes</td>
</tr>
<tr>
<td>Cyclin C</td>
<td>RNAPII transcription in complex with CDK8</td>
<td>Wnt/β-catenin pathway and inhibition of lipogenesis in complex with CDK8</td>
<td>yes</td>
</tr>
<tr>
<td>Cyclin D</td>
<td>Control of G1 phase of cell cycle in complex with CDK4/6; Rb/E2F transcription</td>
<td>HR-mediated DNA damage repair Epigenetic regulation through Mep50</td>
<td>no yes</td>
</tr>
<tr>
<td>Cyclin E</td>
<td>Control of G1-S phase of cell cycle in complex with CDK2; Rb/E2F transcription</td>
<td>Inhibition of neuronal function of CDK5</td>
<td>no</td>
</tr>
<tr>
<td>Cyclin F</td>
<td></td>
<td>DNA damage response SCF mediated proteolysis</td>
<td>no no</td>
</tr>
<tr>
<td>Cyclin H</td>
<td>CDK-activations kinase (CAK) and RNAPII transcription</td>
<td></td>
<td>yes</td>
</tr>
<tr>
<td>Cyclin K</td>
<td></td>
<td>RNAPII transcription in complex with CDK12 and CDK13 DNA damage response in complex with CDK9 and CDK12</td>
<td>yes</td>
</tr>
<tr>
<td>Cyclin L</td>
<td>RNAsplicing in complex with CDK11</td>
<td></td>
<td>yes</td>
</tr>
<tr>
<td>Cyclin T</td>
<td>RNAPII transcription in complex with CDK9</td>
<td></td>
<td>yes</td>
</tr>
<tr>
<td>Cyclin Y</td>
<td></td>
<td>Wnt/β-catenin pathway in complex with CDK14 Spermatogenesis in complex with CDK16</td>
<td>yes</td>
</tr>
</tbody>
</table>

**Table 1-1 Established and emerging roles of cyclins**

Table is adapted from (Lim and Kaldis, 2013). Kinase activity is associated with the established roles of each of the cyclins. This is not the case for all their emerging roles as indicated in column 4.
In metazoans, control of cell cycle progression is orchestrated by different members of the CDK family that bind to cyclins with clear preferences both in vivo and in vitro (Figure 1-5) (Malumbres and Barbacid, 2005). Most cyclins in vertebrates are dispensable for cell cycle progression but some of them have essential roles at particular stages of development. For example, cyclin D1−/− D2−/− D3−/− mice develop until mid/late gestation but then die from heart abnormalities and anemia (Kozar et al., 2004). S-phase cyclins E1 and E2 are essential for embryonic development as cyclin-deficient mice mutants show endoreplication of trophoblast giant cells and megakaryocytes (Geng et al., 2003). CDK1 pairs with cyclin A1 and A2 for activation in G2 phase and with cyclins B1 and B2 to promote mitosis, but only CDK1-cyclin B1 has been shown to be vital for early embryonic cell division (Brandeis et al., 1998).

**Figure 1-5** Cell cycle regulation by cyclin dependant kinases (CDKs) and their role in transcription.

CDKs are represented by the ovals and cyclins as parallelograms of different colours. P represents atom of phosphor.
<table>
<thead>
<tr>
<th>Species</th>
<th>Name</th>
<th>Original name</th>
<th>Size</th>
<th>Cyclins activating CDKs at different stages of the cell cycle</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>S. cerevisiae</em></td>
<td>CDK1</td>
<td>Cdc28, p34 protein kinase</td>
<td>298</td>
<td>G1 - Cln3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1/S – Cln1,2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>S – Clb5.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>M – Clb1,2,3,4</td>
</tr>
<tr>
<td><em>S. pombe</em></td>
<td>CDK1</td>
<td>Cdc2</td>
<td>297</td>
<td>G1 - Puc1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G1/S – Puc1, Sig1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>S – Sig2, Sig1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>M – CDC13</td>
</tr>
<tr>
<td><em>D. melanogaster</em></td>
<td>CDK1</td>
<td>Cdc2</td>
<td>297</td>
<td>S – cyclin A</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>M – cyclin A, B, B3</td>
</tr>
<tr>
<td></td>
<td>CDK2</td>
<td>Cdc2c</td>
<td>314</td>
<td>G1/S – cyclin E</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>S – cyclin E, A</td>
</tr>
<tr>
<td></td>
<td>CDK4</td>
<td>CDK4/6</td>
<td>317</td>
<td>G1 – cyclin D</td>
</tr>
<tr>
<td><em>X. laevis</em></td>
<td>CDK1</td>
<td>Cdc2</td>
<td>301</td>
<td>S – cyclin A</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>M – cyclin A, B</td>
</tr>
<tr>
<td></td>
<td>CDK2</td>
<td></td>
<td>297</td>
<td>G1/S – cyclin E</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>S – cyclin E, A</td>
</tr>
<tr>
<td><em>H. sapiens</em></td>
<td>CDK1</td>
<td>Cdc2</td>
<td>297</td>
<td>S – cyclin A</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>M – cyclin B</td>
</tr>
<tr>
<td></td>
<td>CDK2</td>
<td></td>
<td>298</td>
<td>G1/S – cyclin E</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>S – cyclin A</td>
</tr>
<tr>
<td></td>
<td>CDK4</td>
<td></td>
<td>303</td>
<td>G1 – cyclin D1,2,3</td>
</tr>
<tr>
<td></td>
<td>CDK6</td>
<td></td>
<td>326</td>
<td>G1 – cyclin D1,2,3</td>
</tr>
</tbody>
</table>

Table 1-2 CDKs of most studied biological models and cyclins activating them at different stages of the cell cycle.
1.1.4 Role of phosphorylation in CDK activation

CDK activation minimally requires cyclin binding and also phosphorylation within the activation segment (T-loop) of the kinase domain by CDK activating kinase (CAK). The distribution of a CDK between different phosphorylated and cyclin-bound states differs between organisms and within an organism between different CDKs. These differences reflect for example the accessibility of the CDK activation segment to CAK and phosphatase activities, CDK-cyclin complex stability and levels of CDK and cyclin expression. For example, in budding yeast *S. cerevisiae* that have a single CDK to drive the cell cycle and do not have a sharp G2/M phase transition, monomeric CDK1 is detectable in a phosphorylated state, from which observation it has been proposed that phosphorylation of the CDK1 activation segment occurs before cyclin binding (Ross *et al.*, 2000). In fission yeast *S. pombe*, where a switch-like G2/M transition occurs, CDK1 can be activated by two pathways: as in budding yeasts it can be phosphorylated before binding to cyclin, and also phosphorylation can occur after cyclin binding (Morgan, 2007). In metazoans that contain a number of CDKs and cyclins, different pathways for CDK activation are present. CDKs exist in cells as unphosphorylated monomers and unphosphorylated complexes accessible for phosphorylation by CAK. CDK2 for instance is reported to undergo phosphorylation prior to cyclin binding, but in contrast CDK1 requires phosphorylation by CAK before further activation by cyclin (Figure 1-6) (Desai *et al.*, 1995; Merrick *et al.*, 2008). Monomeric human CDK2 has been shown to be phosphorylated in solution *in vitro* leading to the same conclusion (Merrick *et al.*, 2008). Phosphorylated CDK2-cyclin A (T160pCDK2-cyclin A) is *circa* 300-fold more active towards a model substrate than CDK2-cyclin A (Russo *et al.*, 1996b) as unphosphorylated CDK2 complexes have no measurable activity (Desai *et al.*, 1992; Desai *et al.*, 1995). Nevertheless in some cases the combined effects of cyclin binding and phosphorylation are not sufficient for CDK activation and complexes require additional adjustments. It has been proposed that the activity of CDK4-cyclin D complexes is substrate-assisted and the Michaelis complex is only formed when both ATP and peptide substrates are bound (Day *et al.*, 2009; Takaki *et al.*, 2009). Some CDKs have two phosphorylation sites on the activation segment. For example, CDK7 is phosphorylated on Ser164 and Thr170 (Lolli *et al.*, 2004).
CDKs are also controlled by phosphorylation on a glycine-rich loop also called the P-loop. These phosphorylations normally inhibit CDK activity (Kornbluth et al., 1994; Mueller et al., 1995; Poon et al., 1996).

Figure 1-6 Model for CDK activation in Metazoans.
CDK2 differs in its kinetic of activation from the other CDKs. It is phosphorylated by CAK on the activation segment in a monomeric state with cyclin binding to follow, contrary to CDK1 for example. CDK1 can be phosphorylated only in a complex with cyclin. Picture adapted from (Merrick et al., 2008).

1.2 Structural characterisation of the CDK family
1.2.1 A structural paradigm for CDK activation
The structures of several CDKs have been determined and their comparison has revealed common structural features. Within the kinase catalytic domain the active site is sandwiched between an N terminal lobe mainly consisting of β sheets and a single α helix (often named the C-helix or PSTAIRE helix after the conserved cyclin binding motif within it) connected via a linker to a larger C terminal lobe (Jeffrey et al., 1995; Endicott and Noble, 2013). Another important feature of the N-lobe is a glycine-rich sequence (GXGXXG) also called the P-loop which recognises ATP and includes Thr14 and Tyr15 residues that keep CDK in an inactive state if phosphorylated. These modifications prevent peptide substrate but not ATP binding (Welburn et al., 2007), and subsequent hydrolysis (Solomon et al., 1990). The C-terminal lobe is mainly α helical and includes the activation segment which is located between the conserved
DFG (145-147 in CDK2) and less well conserved APE (170-172 in CDK2) motifs (Figure 1-7) (Endicott et al., 2012). The activation segment includes the phosphorylation sensitive residue (Thr160 in CDK2) consequently it is also termed the T-loop.

Monomeric CDKs normally adopt an inactive conformation, which is defined by the organisation of the C-helix, the DFG motif and the activation segment. Monomeric CDK2 is defined by an inactive “C-helix out, DFG-in” conformation. Residues within the C-helix and DFG motif help create the ATP binding site and contribute to the regulatory and catalytic spine (Figure 1-7 A). These spines are sets of interacting residues that readout the conformation of the active site to the rest of the kinase fold (Kornev and Taylor, 2015). The “DFG-out” conformation has been reported in tyrosine kinases in the inactive conformation where the phenylalanine side chain of the DFG motif is positioned within the ATP-binding pocket. In the “DFG-in” conformation, a conformational switch of this motif displaces the phenylalanine side chain out of the active site to contribute to the catalytic spine, and the aspartate residue is appropriately positioned to interact with a conserved lysine-glutamate pair to coordinate the phosphate groups of ATP (Endicott et al., 2012). All reported structures of inactive CDKs with the exception of CDK8 have the DFG motif in the “in” conformation.

Monomeric CDK2 is in an inactive conformation as a result of C-helix displacement (“C-helix out”). The short α-helical structure at the start of the CDK2 activation loop (αL12) packs against the C helix causing it to adopt a swung out position (De Bondt et al., 1993). This short helix is characteristic of other inactive CDK structures as well (Lolli et al., 2004). The movement results in an increased distance between Glu51 from the C helix and Lys 33 from β3 that leads to positioning of the triphosphate moiety from ATP that isn’t compatible with catalysis. Importantly Leu55 from the regulatory spine and R50 that is a ligand for phosphorylated Thr160 are also displaced. Accompanying these interactions the catalytic spine residues that include Ala31, Val18, Ile87, Leu133, Leu134, Ile135, Ile 192 and Met196 (CDK2 numbering) cannot align appropriately and as a result CDK2 even in the DFG-in conformation is not active (Figure 1-7).
Figure 1-7 Structural changes during CDK2 activation.

Crystal structures of monomeric (A) CDK2 (PDB: 1HCK), (B) CDK2-cyclin A (PDB: 1JST) and (C) pT160 CDK2-cyclin A bound to ATP (PDB: 4EOQ). CDK2 molecule coloured ice blue, cyclin A coloured green, CDK2 activation segment (residues 147-172) coloured red, C-helix (residues 46-55) coloured gold, P-loop (residues 12-16) is coloured dark blue, DFG motif (residues 145-147) coloured pink, APE motif (residues 170-172) is coloured magenta. ATP molecule represented in ball and stick mode with carbon atoms coloured green. Residues forming the regulatory spine are coloured yellow and the location of the regulatory spine is identified with a yellow oval. Residues forming the catalytic spine are coloured green and the position of the catalytic spine is identified with a green oval.

For CDK activation, cyclins bind directly to the C helix and surrounding elements of the kinase N lobe. Generally, the active kinase fold is characterised by having a ‘DFG–in’ conformation where the C helix, which contains a conserved glutamate residue, turns
away from the N terminal lobe, allowing the aspartate of the DFG motif to chelate a Mg\(^{2+}\) ion to position the ATP substrate. The process of kinase activation involves a network of residues from different parts of the molecule (Jura et al., 2011). A conserved glutamate (Glu51 in CDK2) from the C-helix forms a salt bridge with a conserved lysine (Lys33 in CDK2) in strand β3 that coordinates the phosphate groups of ATP (Figure 1-7 B, C). Two more sets of interactions are significant for kinase activation. One set is the interactions of the glycine-rich P-loop with the β and γ phosphates of ATP. The other set positions the catalytic aspartate (Arg127 in CDK2).

Another crucial feature of an active kinase is the activation segment conformation (Figure 1-7 A, B, C). The activation segment is mostly disordered in inactive conformations and requires phosphorylation in most CDKs to compose a substrate binding platform. The phosphorylation promotes closure of the two lobes to the favourable position to accept a peptide for phosphorylation. As more structures have been determined it has become obvious that not all elements of CDK2 activation are conserved across the family. For example, in case of CDK5 (PDB code: 1H4L), binding to its partner p35 is enough for full kinase activation. It adopts the correct conformation through an alternative set of CDK5-p35 interactions (Tarricone et al., 2001).

### 1.2.2 Structural characterisation of the cyclin family

The structures of a number of cyclins have been solved revealing shared common features including the presence of a cyclin box fold (CBF) (De Bondt et al., 1993; Kim et al., 1996; Hoeppner et al., 2005; Baek et al., 2007; Petri et al., 2007). Most cyclin structures comprise two five α-helical domains named CBF1 and CBF2 connected via a linker (Figure 1-8), although some contain only one CBF. CBF1 is relatively well conserved among cyclin A and cyclin E and is responsible for CDK and other protein binding (Petri et al., 2007). CBF2 is very similar to CBF1 but has different lengths of loops and helices. The arrangements of the hydrophobic residues suggest that either domain would be unstable on its own. Four helices within each fold are organised around a H3 or H3’ helix and held in place by Van der Waals and hydrogen bonds. Both CBF domains are stabilized by additional sidechain–main chain hydrogen bonds making the overall structure of the molecule comparatively rigid. The cyclin has been proposed to act as a rigid scaffold for the more malleable CDK (Brown et al., 1995). This stable structure is important for cyclin function as exemplified by cyclin B where
any mutations that destabilise the fold disable binding to its CDK partner (Kobayashi et al., 1992).

**Figure 1-8 Typical cyclin structure.**

Human cyclin B1 (PDB code: 2B9R) is taken as an example. The molecule comprises two parts. CBF1 and CBF2 are each composed of five helices (H1-H5 and H1’-H5’) and coloured in shades of blue or red respectively. The MRAIL sequence is coloured gold and the RRASK motif is coloured yellow. The N terminal sequence is coloured green and the C terminal sequence is coloured red (Petri et al., 2007).

Apart from the CBFs, cyclins can also have additional N-terminal helices, for example two in cyclin K (Baek et al., 2007) and one in cyclins A (Brown et al., 1995) and C (Hoeppner et al., 2005) (Figure 1-9). A distinctive feature of cyclin H is its long C-terminal helix which is present in cyclin A but of much shorter size (Andersen et al., 1996). These additional helices disposed differently in different cyclins. Furthermore, another modification in the cyclin K structure is an additional helix to H4 and H4’, called H4a and H4a’. H4’ and H4a’ from CBF2 differ from their counterparts in CBF1 as they form smaller helices due to insertion of extra residues (225-233). Cyclin H has an extra helix between H1’ and H2’.
Figure 1-9 Structural and topological analysis of human cyclin K, S. pombe cyclin C and human cyclin H.

Corresponding helices are identified by the same colours in the ribbon diagrams. In the topology diagrams CBF1 is coloured green and CBF2 is coloured red (Baek et al., 2007).

On the opposite side to the CDK binding site of cyclins lies the ‘RXL binding groove’ or recruitment site (Schulman et al., 1998; Brown et al., 2007; Petri et al., 2007). The groove is formed by amino acids of CBF1 and is a docking site for RXL motif (x stands for any amino acid)-containing substrates. The RXL motif binding groove has been shown to be essential for phosphorylation of substrates such as pRb and p53.
(Schulman et al., 1998; Adams et al., 1999). It is mainly composed of the MRAIL sequence that is conserved amongst most mammalian cyclins and conserved residues from the H1 helix with contributing residues from the other helices. Interestingly the crystal structure reveals that the analogous region on the CBF2 domain is highly conserved in all types of B-type cyclins and has the sequence RRASK which is important for MYT1 and CDC25c recognition (Goda et al., 2003).

1.2.3 CDK requirements for substrate selection

Protein kinase sequence preferences around the site of phospho-transfer have been well studied (Ewen et al., 1992; Songyang et al., 1994; Cheng et al., 2006). The protein kinase active site cleft can accommodate an extended peptide of circa 7-8 residues permitting sequence-specific interactions to be made with circa 3-4 residues either side of the phosphosite (Brown et al., 2007; Takaki et al., 2009). A study using a CDC6 peptide revealed possibilities of different occupation of the catalytic site by a substrate (Cheng et al., 2006) (Figure 1-10 A). CDK substrates also bind to the cyclin RXL recruitment site through well studied interactions (Brown et al., 1999) (Figures 1-10 B). The fate of the linker between the two sites has not been distinguished in the structures determined by X-ray crystallography due to a lack of supporting electron density (Figure 1-11) (Cheng et al., 2006).

The canonical CDK phosphorylation site consists of a S/T*-P-X-K/R sequence (Ubersax et al., 2003) where S/T is the phosphorylatable residue and X is any amino acid. Although CDKs can target the minimal consensus site comprising only serine or threonine residue followed by proline S/T*-P (Rudner and Murray, 2000). The preference of CDKs for a proline in position P+1 was elaborated by a structure of CDK2-cyclin A bound to a short substrate peptide (Brown et al., 1999). There is a hydrophobic pocket near the catalytic site that can accommodate proline as a result of Val164 in the activation loop adopting an unusual left-handed conformation. CDK2 as well as CDK5 displays strong preferences towards the canonical sequence. However, CDK1 can tolerate residues other than K/R in the P+3 position (Brown et al., 1999). CDK4 has shown similar levels of phosphorylation towards canonical S/T*-P-X-K/R and non-canonical S/T*-P-X-X sequences (Takaki et al., 2009). An explanation of these alternative CDK4 substrate preferences cannot be offered because the CDK4 structure in complex with cyclin D is in an inactive conformation.
Figure 1-10 CDC6 peptide bound to the recruitment site of pCDK2-cyclin A complex.

(A) CDC6 peptide bound to the catalytic site of a pCDK2-cyclin A complex. (B) CDC6 peptide bound to the recruitment site of cyclin A. Figures taken from (Cheng et al., 2006) (PDB code: CCH). Difference maps based on mFo - DFo coefficients were calculated and contoured at 0.1 e/Å³.
Figure 1-11 Possible position of the peptide linker between the catalytic site and the RXL recruitment site.

Linker has been modelled into the pCDK2-cyclin A structure and identified as a green line. Parts of the peptide bound at the catalytic site and at the RXL recruitment site are drawn in ball and stick mode (Cheng et al., 2006).

Transcriptional CDKs, CDK7 and CDK9 have also been reported to have no preference for a positive residue in the P+3 position. This flexibility for CDK7 might be required as it has a dual role in the cell cycle, acting as a CDK activating kinase, and in transcription, phosphorylating the C terminal domain of Pol II (Lolli, 2005). Substrate selectivity for these two processes are known to be different. It has been proposed that for phosphorylation of the activation loop of other CDKs, CDK7 binds to sequences remote from the phosphorylation site and so is not as dependent on residues surrounding the site of phosphotransfer (Garrett et al., 2001; Larochelle et al., 2001). Contrary to the sequence requirements for its transcriptional function CDK7 (as well as CDK9) can autophosphorylate at a YINRVVT sequence which makes specificity to proline redundant (Baumli et al., 2008).

Another important sequence that plays a role in substrate recognition is the substrate recruitment motif RXL. In the presence of this motif CDK2 and CDK4 have been shown to relax their requirement for proline in the P+1 position (Takaki et al., 2009).
1.3 Other regulators of CDK activity

A number of other proteins have been demonstrated to bind directly to CDKs to regulate their activity, although to date, with the exception of the cyclin dependant kinase regulatory subunit (CKS) proteins they are not as well studied as the cyclins and CKIs. They can be generally grouped into cyclin-like regulators that have a cyclin box, for example viral cyclins and Ringo/Speedy (Rapid INducer of G2/M progression in Oocytes), and non-cyclin like regulators.

One well-studied CDK regulator is the CKS1. It was identified as a multi-copy suppressor of a Cdc2 temperature-sensitive mutant in S. pombe and called Suc1 (Hayles et al., 1986a). Loss of the suc1 gene leads to mitotic defects (Hayles et al., 1986b). Homologues of suc1 were subsequently identified in the budding yeast S. cerevisiae (CKS1) (Hadwiger et al., 1989) and in Xenopus eggs (p9) (Patra, D., et al. 1999). Later work identified two homologues in higher eukaryotes, CKS1 and CKS2. Early experiments in Xenopus demonstrated that p9 promotes the phosphorylation of G2/M regulators CDC25, WEE1 and MYT1 (Patra, D., et al. 1999) and Cdc27 (an anaphase promoting complex/cyclosome (APC/C) subunit) by CDK1-cyclin B (Patra, D., and Dunphy, W. G.1998). Mice nullizygous for either CKS1 or CKS2 are smaller than normal but viable (Spruck et al., 2001). However, disruption of both genes simultaneously resulted in embryonic lethality prior to the morula stage, after two or four division cycles (Martinsson-Ahlzen et al., 2008). The role of human CKS1 (CKSHs1) in promoting CDK1 substrate phosphorylation was supported by determination of its structure and the identification of a potential phosphate binding site (Arvai et al., 1995). Further studies performed to investigate the role of CKS1 and its homologues as a phosphoadaptor protein for CDK1 regulation identified hundreds of targets (Ubersax et al., 2003). A model has been proposed in which CKS1 and cyclins coordinate their activities to provide docking sites for multiply phosphorylated CDK substrates to ensure timely G1/S transition (Koivomagi, Valk et al. 2011).

CKS family members share high sequence identity from yeasts to humans (Figure 1-12). The crystal structures of human CKSHs2 (Parge et al., 1993), S. pombe Suc1 (Bourne et al., 1995; Endicott et al., 1995), S. cerevisiae Cks1 (Bourne et al., 2000) and human CKSHs1 (Arvai et al., 1995) have been solved on their own or for HsCKS1 bound to CDK2 (Bourne et al., 1996). Interestingly structures for Suc1 identified two different conformations, a single compact domain fold (Endicott et al., 1995) and β-
strand exchanged dimer (Bourne et al., 1995). The two human homologs CKSHs1 and CKSHs2 have also been reported to adopt respectively the compact fold (Arvai et al., 1995) or the β-interchanged dimer (Parge et al., 1993) (Figure 1-13). Analysis of CKS in both conformations lead to the conclusion that only the compact fold can bind to a CDK (Bourne et al., 1996).

The structure of CDK2 in complex with CKSHs1 has revealed the binding site which is located on the C-lobe of CDK2 (Bourne et al., 1996). Four β strands (β1-β4), the β1-β2 loop also called L1 and the β hinge region of CKSHs2 interact with the α5 helix and L14 loop of CDK2 (Figure 1-13, 1-14). The major difference between the compact fold and the dimer is the conformation of the β hinge region/β-hairpin that plays a role in CDK binding.

Figure 1-12 Sequence overlay of Human CKS1 and CKS2, S. pombe Suc1 and S. cerevisiae Cks1.

Uniprot entries are P61024, P33552, P08463, C7GY17 respectively. β strands and α helices that mediate CDK binding are colored red and those participating in ion binding are blue. β strands and α helices are identified with green and red boxes respectively. Stars represent identical residues, conserved residues marked with (:) or (.)
Figure 1-13 Crystal structure of monomeric Suc1 and the β-interchanged dimer conformations.
(A) Structure represents β interchanged dimer fold (coral and gold subunits) with β-hinge region (His88 to His93) coloured blue. (B) Single domain fold represented by gold and lilac subunits. Zn ions are represented as small spheres and labelled.

Figure 1-14 Crystal structure of a CDK2-CKSHs1 complex.
CDK2 structure color coded ice blue, α5 helix is labeled and color coded green, L14 loop is colored red, CKSHs1 is colored gold, β1 is colored cyan, β2 is colored yellow, β3 is colored light green, β4 is colored light blue, region between β1 and β2 called L1, region between β3 and β4 called L3. C and N terminals are labeled (PDB code: 1BUH).
1.4 CDKs and cancer

The term “cancer” unites a group of diseases that are characterised by common features such as abrupt and uncontrolled proliferation that can spread to other parts of the body. Its causes vary depending on the type of cancer and includes tobacco smoking, exposure to ionising radiation, infections, poor diet and lack of physical activity. Solid tumour cancers usually consist of neoplasms formed from multiple distinct cell types that atypically interact with each other. The process of changing normal cells into ones that can form cancer is called malignant progression and occurs in multiple steps. These cell transformations are the outcome of changes in mitogen signaling caused by either excessive exogenous signaling as growth factors and nutrients (Massague, 2004) or oncogenic mutations in genes (Knudson, 1971).

Cancer has been known to people through history. Nowadays, despite great progress in cancer treatment achieved since the 1970s, it remains a significant research area due to the growing incidence of disease around the world (Cancer Research UK statistics). After analysis of knowledge accumulated through years of research, distinctive and complementary characteristics of cancer cells that promote tumour growth and metastatic dissemination and known as “hallmarks of cancer” have been described (Figure 1-15) (Hanahan and Weinberg, 2011).

One of the most fundamental hallmarks of cancer cells is their ability to initiate and sustain unlimited proliferation as well as evade growth suppression. As proliferation depends on cell cycle progression, mutations that lead to malignancies converge to deregulate CDK activity. Oncogene induced errors in the cell cycle can instigate errors in DNA replication and can lead to genomic instability as well as defects in spindle assembly checkpoints which result in abnormal chromosome segregation (Malumbres and Barbacid, 2009). The functions of key CDKs and their involvement in cell cycle regulation in normal tissue and in malignancies is described below.
1.4.1 CDK4/6 and cancer

In most normal adult tissues, cells exist in a G0 state which can be transient or permanent. From the transient state, cells can re-enter the cell cycle under the influence of mitogenic factors. Most of these factors initiate signalling pathways that result in activation of CDK4 and CDK6. CDK4 and CDK6 have largely overlapping sets of target proteins (Anders et al., 2011) although they are very important for specialised cell types. CDK4-/- mice develop defects in endocrine cells (Rane et al., 1999) and CDK6 is more important for hematopoietic cells (Malumbres et al., 2004; Kollmann et al., 2013). Consequently, amplification and over-expression of these kinases is implicated in tumours often from endocrine and hematopoietic origin (Figure 1-16). Upregulation of CDK4 is also reported in gliomas, sarcomas, breast tumours and carcinomas of the uterine cervix in a co-amplification with the MDM2 locus as part of an amplicon located on chromosome 12(12q13-14) (Lopes et al., 2001; Ortega et al., 2002; Simon et al., 2002). CDK6 has been found overexpressed in lymphoid tumours.
(Raffini et al., 2002) and upregulated by gene amplification in squamous cell carcinomas and gliomas (Lam et al., 2000). In addition to gene amplification and deletion, kinases can also acquire missense mutations according to COSMIC data base,(Bamford et al., 2004). CDK6, for example, in breast carcinoma and hematopoietic and lymphoid tissue, and CDK4 in cases of human familial melanoma (Zuo et al., 1996) and gliomas and several subgroups of sarcomas. Mutations in genes encoding cyclin D can cause mantle cell lymphoma, plasma cell leukaemia, chronic lymphocytic leukaemia and multiple myeloma, according to COSMIC data base, (Bamford et al., 2004).

Figure 1-16 Deregulation of members of CDK4/6 pathway leading to cancer (Otto and Sicinski, 2017).

To study the implications of genomic instability-causing mutations that lead to misregulation of CDK4 and CDK6 in cancer, various genetically engineered mouse models have been developed (Otto and Sicinski, 2017). For example, CDK4 knock-in mice were created which were carrying a miscoding mutation (CDK4R24C) found in
human melanoma (Wolfel et al., 1995). This mutation did not cause any detectable defects during embryonic development, but as adults, mice appeared 10-20% bigger than their normal littermates and by 8-16 months developed multiple tumours of endocrine origin, sarcomas and epithelial hyperplasias of the liver, gut and breasts (Sotillo et al., 2001). Furthermore, mice engineered to overexpress cyclin D1 in mammary glands developed mammary hyperplasia and carcinomas (Wang et al., 1994). These results highlight the oncogenic properties of CDK4/6 and D-type cyclins. Further investigation revealed that the levels of CDK4/6 and cyclin D do not need to be simultaneously elevated for tumorigenesis. For instance, mice lacking cyclin D1 were resistant to mammary cancer induced by some oncogenes, such as vHras (an oncogenic HRAS mutant) but not others (such as Wnt 1 and Myc) (Yu et al., 2001; Bowe et al., 2002). Ccnd3-null mice were resistant to NOTCH1ICD driven T cell acute lymphoblastic leukaemia (T-ALL), whereas CDK6-knockout mice were resistant to lymphoma formation induced by AKT (Sicinska et al., 2003; Hu et al., 2009). Interestingly, total ablation of Ccnd1 or pharmacological inhibition of CDK4/6 kinase activity in mice bearing Erbb2V664E-driven mammary tumours blocked cancer progression without having an effect on normal, non-transformed tissues (Choi et al., 2012). Collectively these results reveal differences in kinase wiring of particular types of cancer in contrast to non-transformed tissues. CDK4 and CDK6 as well as D-type cyclins are essential for some tumour initiation and maintenance as opposed to non-transformed tissue, where the absence of CDK4 and CDK6 catalytic activity has no major effect on progression of the cell cycle (Choi et al., 2012). Taken together, these results suggest that tumours frequently depend on individual cyclins and CDKs and are therefore prone to their targeted inhibition.

1.4.2 CDK2 and cancer

The discoveries that (i) CDK2 has not been found mutated in human cancers, (ii) CDK2 knock down cell lines do not go into cell cycle arrest (Tetsu and McCormick, 2003) and (iii) knock out mice are viable (Berthet et al., 2003; Ortega et al., 2003) led to a doubt about CDK2 being a valid target for cancer therapy. However, recent studies employing chemical genetic approaches have suggested otherwise. In these studies, wild type CDK2 in a human colon cancer cell line and mouth embryonic fibroblasts (MEFs) were replaced with an analog-sensitive version which can be inhibited with ATP analogs. The results demonstrated that in both these cellular settings treatment with a variety of selective CDK2 inhibitors resulted in a significant reduction in cell
growth (Horiuchi et al., 2012a). Additional observations in different cancer types, for example BRCA-deficient cancers (Deans et al., 2006), neuroblastoma (Molenaar et al., 2009) and ovarian cancer (Yang et al., 2015) (Au-Yeung et al., 2017) have identified CDK2 as a potential drug target. These oncogenic features of CDK2 are mainly related to its activators cyclin E1 and cyclin E2 (Figure 1-17). Cyclins E1 and E2 are often overexpressed in multiple tumours, particularly in uterine, ovarian and breast cancers (Karst et al., 2014) (Kuhn et al., 2014; Au-Yeung et al., 2017). Ectopic expression of cyclin E bypasses the physiological requirement for CDK4 and CDK6 to prepare cells for S phase (Caldon et al., 2012) and is assumed to be oncogenic. Overexpression of cyclin E1 has a prognostic effect in some tumour types, as in breast cancer increased expression correlates with the progression of the disease. Likewise, cyclin A, another regulator of CDK2, has been shown to be overexpressed due to genomic amplification in colorectal (Handa et al., 1999) and breast cancers (Michalides et al., 2002) as well as in hepatocellular carcinoma (Chao et al., 1998).

Figure 1-17 Deregulation of cyclins E1 and E2 leading to cancers (Otto and Sicinski, 2017).
1.4.3 Role of CDK9 in cancer

Another CDK which has emerged recently as an important regulator in some cancers is CDK9. CDK9 regulates RNA transcription and it is hypothesized that changes to CDK9 activity have a particular effect on the expression of short lived anti-apoptotic proteins. It has been shown that in some cancers, particularly in lymphocytic leukaemia, CDK9 inhibition reduces the levels of anti-apoptotic proteins such as Mcl-1 and XIAP and restores the ability of these cancer cells to undergo apoptosis (Wang and Fischer, 2008; Bose et al., 2013). A subsequent study concluded that transformed cells in contrast to non-transformed cells are addicted to the expression of anti-apoptotic proteins (Wang et al., 2010). The use of new techniques such as RNAi methods in conjunction with small molecule inhibitors yielded new insights (Lemke et al., 2014). For example, transient knock down of CDK9 sensitized HeLa and A549 cells to tumour necrosis factor-related apoptosis-inducing ligand (TRAIL) yielding a range of phenotypes that might be targeted by available mono and combined therapies (Garriga et al., 2010). In head and neck squamous carcinoma cell lines, knock down of CDK9 led to radio-sensitization (Wang et al., 2014). In ovarian cancer cells (A2780), shRNA knock down of CDK9 or inhibition with flavopiridol induced apoptosis (Lam et al., 2014). The summary of this study suggests that the anti-cancer effect of CDK9 inhibition depends on the methods employed, cell type and kinetics of the inhibition. It also has been shown that the therapeutic potential of CDK9 is strongest if it is co-inhibited with BRD4 (Sonawane et al., 2016). BRD4 targets the P-TEFb complex, that contains CDK9, to mitotic chromosomes promoting expression of growth promoting genes (Yang et al., 2008). It has been found as a component of chromosomal translocation (15;19) of human squamous carcinoma. As a result of this translocation a tandem pair of BRD4 N-terminal bromodomains are expressed as a chimaera together with testis nuclear protein NUT and has been identified as an oncoprotein (French et al., 2001). RNA silencing of BRD4-NUT led to arrest of proliferation (Filippakopoulos et al., 2010). Further experiments in human squamous carcinoma xenograft models demonstrated that the BRD4 fusion oncoprotein bound to the small molecule inhibitor JQ1 can be displaced from chromatin resulting in an antiproliferative effect (Filippakopoulos et al., 2010). These outcomes show the utility and therapeutic potential of CDK9 and BRD4 in cancer treatment.
1.4.4 CDK1 and cancer

CDK1 is the only essential CDK and the only CDK that can initiate mitosis (Santamaria et al., 2007). If mitosis is initiated before S phase is completed it can lead to G2 arrest and cell death (Aarts et al., 2012), whereas inhibition during mitosis results in a quick mitotic exit without cytokinesis (Vassilev et al., 2006). CDK1 is constantly expressed in a cell but its activity is inhibited by a number of mechanisms until DNA replication is completed. In tumorigenesis, CDK1 behaves differently from CDKs that control S phase progression and shows very little evidence of direct genetic alteration. However, other factors such as modifications to DNA damage checkpoints (Bayart et al., 2004) and to p53 signalling pathways (Taylor and Stark, 2001) as well as changes in the level of cyclin B (Nam and van Deursen, 2014) can influence CDK1 activity. These alterations can lead to a rise of CDK1 activity in various cancers, for example colorectal cancer (Sung et al., 2014) and prostate cancer (Willder et al., 2013). Elevated expression of cyclin B is also associated with more aggressive cancer phenotypes and poor survival rates (Aaltonen et al., 2009).

Although inhibition of CDK1 triggers apoptosis of MYC-driven lymphomas and liver tumours (Goga et al., 2007), as well as in human triple-negative breast cancer cells (Horiuchi et al., 2012b), the requirement for CDK1 activity for cell survival of non-transformed cells, and for its elimination to exit mitosis limits its potential as a target for cancer therapy. The only realistic approach for CDK1 inhibition which has been considered so far is its use in combination therapy. It is hypothesized that CDK1 inhibition can induce a BRACA-deficient-like phenotype which would sensitise cells to DNA damaging agents (Johnson et al., 2009). The most useful approach would be to design inhibitors that are selective towards other CDKs but do not inhibit CDK1.

1.4.5 Role of pharmacological CDK inhibition for cancer therapy

In the last twenty years several synthetic CDK inhibitors have been designed as potential anti-cancer drugs and tested in different clinical trials. First generation inhibitors were relatively non-specific although some of them, for example roscovitine, were quite effective for some cancers and it is still in clinical trials for Cushing disease. Another successful early inhibitor was flavopiridol that underwent 60 clinical trials between 1998 and 2014. It is a semi-synthetic compound derived from rohitukine, a chromone alkaloid which has been shown to inhibit CDK1, 2, 4, 6, 7 and 9. It was still used for cancer treatment until 2012 (Byrd et al., 2007).
The second generation of inhibitors emerged from the development of flavopiridol and roscovitine. Their design was focused on increasing potency and selectivity towards CDK1 and CDK2. Many inhibitors of this generation looked promising in preclinical studies but did not progress to Phase 1 clinical trials (Payton et al., 2006; Parry et al., 2010). One of the most studied inhibitors is Dinaciclib which was designed to inhibit CDK1, 2, 5 and 9 and has IC$_{50}$ values in a range of 1-4 nM. It strongly suppresses Rb phosphorylation in cell based assays (Parry et al., 2010). Although results of the Phase 1 clinical trials were promising (Nemunaitis et al., 2013), Phase 2 results in solid tumours were disappointing. Dinaciclib monotherapy in non-small cell lung cancer showed no activity in previously treated patients (Stephenson et al., 2014). In patients with acute myeloid leukaemia or acute lymphoblastic leukaemia no objective changes were observed (Gojo et al., 2010). However, in haematological malignancies it has shown clinical activity and consequently it remains in Phase 3 clinical trials (Blachly and Byrd, 2013). Another inhibitor, AZD5438 (developed by AstraZeneca), was reported to be intolerable when administered to patients with advanced solid tumours (Boss et al., 2010). The Phase 1 study of AG-024322 (developed by Pfizer), a potent inhibitor of CDK1, 2 and 4 was also terminated because of the absence of expected therapeutic effect.

Selectivity is a significant issue for the design of CDK-specific inhibitors. For example some inhibitors target CDK1 which is vital for proliferation and CDK9 which is important for survival of normal cells, thus achieving a therapeutic window is difficult. To overcome this barrier, CDK inhibitors have been used in combination. For example, inhibition of CDK1 with RO3306 (Roche) led to sensitising BRCA proficient cancers to poly (ADP-ribose) polymerase PARP inhibition (Johnson et al., 2011). CDK1 has been shown to participate upstream of DNA damage response pathways (Ira et al., 2004) and in CDK1 depleted cells the function of BRCA1 in the S phase checkpoint was compromised (Johnson et al., 2009). BRCA1 and BRCA2 are tumour suppressor proteins that play crucial roles in DNA repair (Moynahan et al., 1999). PARP in its turn plays a role in single strand DNA repair by homologous recombination. Both BRCA1 and BRCA2 are necessary for homologous recombination and it is known that BRCA-deficient cells are hypersensitive to PARP inhibition (Bryant et al., 2005).
1.5 Aims of the thesis

The central aim of this thesis is to provide a better understanding of the function of CDK1 based on its structure. Structural information can be very beneficial especially in distinguishing highly homologous proteins that are involved into different functions. The structure of CDK1 in its inactive conformation has been solved in complex with its regulator CKS1 and also bound to its activating partner cyclin B and CKS2. This work is described in Chapter 2. Subsequent to the determination of the structure of CDK1-cyclin B, the experiments described in Chapter 3 use biochemical and biophysical techniques to explore insights into CDK1-cyclin B stability, and substrate preferences provided by the structure determination.

To determine the response of CDK1 to association with different cyclin partners and to activation loop phosphorylation, studies to purify various alternative CDK1-containing complexes and CDK1 regulatory proteins, and their subsequent crystallisation trials are reported in Chapter 4. Ringo/Speedy A is an alternative cyclin-like CDK activator that has been characterised in recent years. A series of attempts were made to express it in various expression systems and these are also described in Chapter 4. Chapter 4 also describes preliminary work to characterise another potential CDK1 partner, RGC32 in collaboration with the laboratory of Dr Michelle West (University of Sussex).

CDK upregulation in cancer and in other proliferative diseases has led to interest in them as drug targets. Structural insights can make a significant contribution to the design of potent and specific ATP competitive inhibitors. These compounds can be used as a chemical probes to distinguish CDKs and as leads for inhibitor design. Chapter 5 describes various biophysical and structural studies to understand the differing responses of CDK1 and CDK2 to ATP-competitive inhibitor binding.
Chapter 2: Structural insights into CDK1-containing complexes

2.1 Introduction

2.1.1 Discovery of CDK1

An understanding of cell cycle control is essential for the fields of biology and medicine. Studies to understand the regulation of the cell cycle were started in the 1970s by Leland Hartwell who identified genes that control the cell cycle in *Saccharomyces cerevisiae* (budding yeast), the so-called cdc-genes (cell division cycle genes) (Hartwell et al., 1970; Hartwell et al., 1974). He identified *CDC28*, to play a role in initiation of G1 phase. Hartwell also introduced the idea of checkpoints that was expanded and developed by his successors. The genetic approach used in the laboratory of Hartwell was taken further by Paul Nurse who discovered the *cdc2* gene in *Schizosaccharomyces pombe* (fission yeast) and revealed its essential function in cell division (Nurse et al., 1976; Nurse and Thuriaux, 1980). Paul Nurse also demonstrated that activation of the protein encoded by the *cdc2* gene depends on phosphorylation, and identified the corresponding gene in humans (Simanis and Nurse, 1986; Lee and Nurse, 1987; Lee et al., 1988). The first cyclin molecule was reported in the 1980s by Tim Hunt, who went on to demonstrate its role primarily as a component of maturation promotion factor (MPF) and its importance in controlling the cell cycle (Evans et al., 1983; Standart et al., 1987; Minshull et al., 1989). For these seminal discoveries Leland Hartwell, Paul Nurse and Tim Hunt were awarded a Nobel Prize in Physiology and Medicine in 2001.

To explore the regulation of the cell cycle two approaches have primarily been employed. The first approach was cell biological/biochemical, where cell cycle progression was studied using cell manipulation techniques and proteins were subsequently purified and characterised. Using this approach it was discovered that entry into mitosis is controlled by a cytoplasmic factor in the slime mold *Physarum polycephalum* (Rusch et al., 1966). Similar work was also being performed in frog oocytes where microinjection of the cytoplasm from an oocyte undergoing the first meiotic metaphase into a recipient oocyte not undergoing meiosis was shown to induce maturation (Masui and Markert, 1971; Smith and Ecker, 1971). These studies suggested the presence of an ‘intracellular inducer’ which stimulated oocyte
maturation that has since been named Maturation Promotion Factor (MPF). Further research led to the discovery of MPF activity in mitosis as well (Hara et al., 1980). The MPF complex was purified in 1988 and shown to be composed of two proteins of 45 and 32 kDa molecular weight respectively that was able to induce mitosis and was activated by phosphorylation (Lohka et al., 1988). These proteins were confirmed as CDK1, in both Xenopus and starfish and a cyclin (Labbe et al., 1988). MPF is now known as an active complex of CDK1 and cyclin B, key components of cell cycle regulation.

The second mainly genetic approach was used to identify the regulatory genes that control the cell cycle primarily in yeasts. Hartwell in 1974 had proposed in the budding yeast Saccharomyces cerevisiae the existence of some point in G1 that he named ‘Start’, when a cell commits to division. He subsequently described the gene regulating this event and named it CDC28 (Hartwell, 1974). Work was continued and developed by Paul Nurse and colleagues who identified the network of genes acting through regulation of the cdc2 gene product, CDC2 (also known as CDK1) which controls entry into mitosis (Nurse and Thuriaux, 1980) and S phase (Nurse and Bissett, 1981) in S. pombe. The budding yeast cdc2 orthologue CDC28 was subsequently also shown to have roles required for successful completion of S-phase and progression through mitosis (Piggott et al., 1982).

The human orthologue of the S. pombe cdc2 gene was identified and described in the laboratory of Paul Nurse (Lee and Nurse, 1987). In this work, a gene was identified from a human cDNA library, which could rescue a cdc2 deletion in fission yeast. The encoded protein proved to be of similar molecular weight to S. cerevisiae CDC28 and S. pombe CDC2. Also 62% of the amino acid sequence was identical to S. cerevisiae CDC28 and 57% to S. pombe CDC2. Comparison of CDK1 amongst the different species demonstrates their similarity, overall showing 41% identity (Figure 2-1), suggesting CDK1 conservation through all eukaryotes.
Figure 2-1 Comparison of CDK1 protein sequences from the different species

CDK1 from the different species were aligned using Clustal Omega (Sievers et al., 2011). Origin of the CDK1 is identified on a left hand side with the Uniprot entries in a respective order: P06493 for CDK1 Homo sapiens, P00546 for S. cerevisiae, P04551 for S. pombe, P39951 for Rattus norvegicus, P23572 for Drosophila melanogaster, Q9DG98 for Oryzias luzonensis (Luzon rice fish). Identical sequences marked dark purple, lower similarities identified with the lighter shades of purple.

CDK1 is the only essential cell cycle CDK as it can substitute for other interphase CDKs by pairing with non-cognate cyclins to promote progression from G1 into S phase (Fisher and Nurse, 1996) (Mondesert et al., 1996). Cdk1 knockout mouse embryos fail to develop beyond the morula stage. Also in the absence of interphase CDKs, CDK1 can replace them by binding to their respective cyclins and phosphorylating their substrates. For example, CDK1 can phosphorylate retinoblastoma protein pRb.
p27KIP1 and p57KIP2 to promote cell cycle progression in the developing embryo through organogenesis up to mid-gestation (Santamaria et al., 2007).

### 2.1.2 Regulation of CDK1 activity

A complex regulatory network navigates the activity of the essential cell cycle kinase CDK1. Its activity through G2/M transition is regulated by the concentration of cyclin A and B (Deibler and Kirschner, 2010), by activating phosphorylation on the activation segment Thr161 by CAK and also by inhibitory phosphorylation. The inhibitory phosphorylation is controlled by two kinases in metazoans, WEE1 and MYT1, which phosphorylate residues Thr14 and Tyr15 on the glycine-rich P-loop and phosphatase CDC25 which antagonizes WEE1 and MYT1 activity and de-phosphorylates these residues (Hoffmann et al., 1993). Consequently, WEE1 and MYT1 activity delays CDK1 entry into mitosis and CDC25 promotes it (Dunphy and Kumagai, 1991; Featherstone and Russell, 1991; Gautier et al., 1991). Nuclear WEE1 is present in all eukaryotic organisms but MYT1 is only found in metazoans and is known to be localised to the endoplasmic reticulum (ER) and Golgi membranes (Kornbluth et al., 1994; Liu et al., 1997). A homologue of WEE1 named Swe1 in budding yeasts has been studied to better understand the mechanism of CDK1 regulation by inhibitory phosphorylation. As a result, a model has been proposed where Swe1 binds to Cdc28-Clb2 inhibiting it by phosphorylating Thr 14 in late G2. Cdc28-Clb2 in its turn activates Swe1 by hyper-phosphorylation (Harvey et al., 2005) which allows Swe1 activity towards Cdc28. Removal of the inhibitory phosphate from Cdc28 by Cdc25 allows full hyper-phosphorylation of Swe1 by Cdc28-Clb2 and release of the Cdc28-Clb2 complex. Therefore, Cdc28 is both a positive and negative regulator of Swe1. The large pool of Cdc28-Clb2 in complex with Swe1 maintains it in the inhibited state in G2 phase but it is ready to be rapidly activated to trigger entry into mitosis.

Lower activity of CDK1-cyclin B results in higher activity of WEE1 and MYT1. Activation of CDK1-cyclin B triggers activation of CDC25 phosphatase which is kept inhibited by the kinases CHK1 and CHK2 that operate downstream of the DNA damage checkpoint. CDK1 activation leads to further CDC25 phosphorylation initiating a positive feedback loop. Polo like kinase 1 (Plk1) also activates CDC25 and might downregulate WEE1 and MYT1 (Nigg 2001). The increase in active CDK1-cyclin B leads to deactivation of WEE1 and MYT1 as CDK1-cyclin B phosphorylation inhibits both these enzymes (Figure 2-2), (Okamoto and Sagata, 2007).
It has been proposed that the concerted activity of cyclins A and B can serve as the triggering factor for the activation of positive and double negative feedback loops that leads to mitosis. In somatic cells the threshold of ~400-500 nM was defined for cyclin B above which it causes a dramatic increase in CDK1 activity, mediated by CDC25. At the same time, it has been suggested that active CDK1-cyclin A weakens the WEE1/MYT1 negative feedback promoting cyclin B activation at lower concentrations. CDK1-cyclin A activity is generally much lower than CDK1-cyclin B but the contribution of cyclin A is very important. For example, cyclin A is required for timely nuclear envelope breakdown (NEB) in HeLa cells (D Gong, 2007). It participates in the G1/S transition but also accumulates through S and G2 phases (Deibler and Kirschner 2010). Both cyclins are insufficient inducers of mitosis but complexes of both of them in collaboration generate sufficient signal for mitotic entry (Winn et al., 2011)

**Figure 2-2 Regulation of CDK1 activity by positive and double negative activation loops.**

CDC25 activates CDK1 by dephosphorylation of Thr14 and Tyr15. Active CDK1-cyclin B activates CDC25 in a positive feedback loop. Active CDC25 is pictured in red, inactive in black. WEE1 inactivates CDK1 by phosphorylation on Thr14 and Tyr15.

The phosphorylation of CDC25, WEE1 and MYT1 is mainly catalyzed by CDK1 and is opposed by protein phosphatase 2A(PP2A)-B55δ (Pal et al., 2008). (PP2A)-B55δ activity is reported to be high in interphase and low during mitosis (Yu et al., 2006). It is regulated by Greatwall (Gwl), another kinase which is essential for mitotic entry and
successful maintenance of mitotic progress (Figure 2-3) (Yu et al., 2006). For its activation GwI requires phosphorylation by CDK1, and once phosphorylated it activates a downstream phosphatase Ensa (includes small proteins Arpp19 and endosulfine-α) which switches off 2A(PP2A)-B55δ (Mochida et al., 2010). Therefore, GwI and Ensa are essential for progression of mitosis due to their ability to inhibit phosphatases. In a recently proposed mathematical model (Tyson and Novak, 2015) 2A(PP2A)-B55δ was suggested as a CDK1- counteracting phosphatase for GwI as well. CDK1-cyclin B phosphorylates APC and Cdc20 which form a complex following cyclin B destruction.

![Figure 2-3 Regulation of CDK1 activity at the G2/M transition.](image)

Solid arrows represent chemical reactions, dashed arrow represents catalytic activity, and solid line with two dots on both sides with an arrow at the middle represents binding to form heterodimer. B55 is shorter abbreviation for 2A(PP2A)-B55δ phosphatase (Tyson and Novak, 2015).

### 2.1.3 Structures of CDK-cyclin complexes

CDKs are allosterically regulated by cyclin binding. Monomeric CDKs adopt an inactive conformation, and cyclin association accompanied in most examples by activation loop phosphorylation leads to CDK activation. The inactive CDK conformation is characterised by the αC helix (also called the PSTAIRE helix for the conserved cyclin...
recognition sequence within) being rotated out of the N-lobe (De Bondt et al., 1993). As a result of this shift in the position of the αC helix, a conserved glutamate residue (Glu51 in CDK2) is moved out of the active site which leads to the disruption of its interaction with the lysine side chain that coordinates ATP (Lys33 in CDK2).

Figure 2-4 The structures of pCDK2-cyclinA, pCDK9-cyclinT, pCDK4-cyclinD. CDKs are coloured blue and cyclins are green. CDKs N - and C - terminus and Cyclin Box Folds are identified. The CDK–cyclin binding surface is encircled red. (A) T160pCDK2-cyclin A (PDB code: 1QMZ), (B) T186pCDK9-cyclin T (PDB code: 3BLH), (C) T172pCDK4-cyclinD (PDB code: 2W9F).

The model of CDK2 activation by cyclin A binding and phosphorylation on the activation segment became a paradigm for CDK activation (Jeffrey et al., 1995) as described earlier. Subsequently, the determination of the CDK5-p25 structure revealed that CDK5
can adopt an active conformation without activating phosphorylation (Tarricone et al., 2001); and in contrast, CDK4–cyclin D complexes even when phosphorylated do not fold into an active state in the absence of substrate binding (Day et al., 2009). Apart from these reported differences in activation mechanism, the extent to which the CDK and cyclin subunits interact can also vary. For example, CDK2 forms a very tight complex with cyclin A where residues from both N- and C-lobes participate in cyclin binding. On the other hand, the interaction surface between CDK4 and cyclin D is much smaller with residues from the N-lobe of CDK4 and the cyclin D N terminal cyclin box mediating the association. As structures of transcriptional CDK-containing complexes have been determined, new features that distinguish them from the cell cycle CDKs have been reported revealing that there are significant structural differences between members of the CDK family (Figure 2-4).

### 2.1.4 Structural features of cyclins A and B

CDK1 binds to type A and type B cyclins while CDK2 is found in complexes containing type A and E cyclins. The first monomeric cyclin structure to be solved was that of bovine cyclin A2 (residues 171-433) (Brown et al., 1995). The structure of monomeric human cyclin B1 (residues 165-433) was subsequently published in 2007 (Petri et al., 2007). The structure of human cyclin E1 (residues 81-363) has to date only been reported in a complex with CDK2 (Honda et al., 2005). The N terminal cyclin sequences were not present in the solved structures as sequence analysis suggested they would be unstructured and previous work had demonstrated that they were not required for CDK activity (Kobayashi et al., 1992). Cyclins A, B and E share the typical cyclin structure comprised of two tandem cyclin box folds - CBF1 and CBF2 (Figure 2-6 A) and together with cyclin D share an ‘RXL binding groove’ or “recruitment site” (Figure 2-6 B) (Adams et al., 1996). Both cyclin boxes are organised around a central hydrophobic helix named α3 or α3’. The surrounding helices are held by Van der Waals interactions but also hydrogen bonds. The recruitment site is formed by amino acids from the α1 helix from CBF1 and contains the conserved MRAIL sequence (Figure 2-5). The site is essential for phosphorylation of substrates such as pRb and p53 and for binding of CIP/KIP cyclin-dependent kinase inhibitors (Schulman et al., 1998). In the structurally analogous region in CBF2 the RRASK sequence conserved in all types of type B cyclins, but not in cyclin A or cyclin E replaces the MRAIL motif.
Figure 2-5 Comparison of structures of cyclins A, B and E.

Structures of cyclins A (PDB code: 1VIN), B (PDB code: 2B9R) and E (PDB code: 1W98) in ribbon representation. Helices of CBF1 are colored $\alpha_1$ – cyan, $\alpha_2$ – dark cyan, $\alpha_3$ – light blue, $\alpha_4$ – blue, $\alpha_5$ – dark blue, helices of CBF2 are colored $\alpha_1'$ – purple, $\alpha_2'$ – lilac, $\alpha_3'$ – bright pink, $\alpha_4'$ – pink, $\alpha_5'$ – red. N-terminus colored green, C-terminus – gold (cyclin A), ice blue (cyclin B), orange (cyclin E). The MRAIL motif is highlighted in yellow.

The structural comparison of cyclin A, cyclin B and cyclin E highlights their very similar structures despite there being little sequence similarity between them. For example, the CBF1 sequences of cyclins A and B show 37% amino acid identity and those of cyclins B and E, 29% (Figure 2-7 A) and between cyclin A and cyclin E – 42% (Honda et al., 2005). As cyclin B and cyclin A both interact with CDK1, and CDK2 interacts with both cyclin A and cyclin E, these similarities in the CBF1 folds suggests its role is to facilitate CDK binding. As can be seen from an overlay of the three cyclin structures CBF1 is very similar for all of them, but CBF2 differs more significantly (Figure 2-6 A). The main changes occur in the loop sequences, and in the orientation and lengths of the helices. These observations lead to the hypothesis that CBF2 is mainly responsible for cyclin-specific functions.
Figure 2-6 Comparison of structures of cyclins A, B and E

(A) Structures of cyclin A (gold), cyclin B (ice blue) and cyclin E (coral) were superimposed (using CCP4mg program) (McNicholas et al., 2011). The structures share very good structural homology over the N-terminal cyclin box (CBF1), but the lengths and relative positions of the helices differ in the C-terminal cyclin box (CBF2).

(B) Position of p27 on a surface of cyclin B was modelled by superposition of cyclin B with CDK2-cyclin A-p27 (PDB code: 1JSU). p27 is bound to the RXL binding hydrophobic groove. p27 is colored green and the RXL motif is colored orange. Cyclin A and cyclin B conserved residues involved in contact with p27 are colored ice blue, different residues are colored red. Unfavorable interactions (defined as those where the juxtaposition of atoms in the bound state is less energetically favorable than the solvation of the atoms in the free state) are colored yellow.
Figure 2-7 Sequence alignment.
(A) Schematic representation of the domain organization of cyclins A, B and E. The cyclin box folds (CBFs) are colored grey, the rest of molecule is blue. (B) Multiple sequence alignment of bovine cyclin A2 (UniProt entry P20248), human cyclin B1 (UniProt entry P14635) and human cyclin E1 (UniProt entry P24864) using Clustal Omega (Sievers et al., 2011) to illustrate sequence conservation and diversity between these cyclins. Identical residues are colored in blue. Helices are represented above alignments as red bars and named next to the bar.

2.1.5 Analysis of the CDK-cyclin interface

As mentioned above, an analysis of the structures of various CDK-cyclin complexes reveals substantial differences in the CDK-cyclin interface. However, despite these global differences, a number of structural features are conserved. For example, the part of the CDK structure that is composed of the sequence that precedes the C-helix (β3 to αC), the αC helix and the β4 and β5 sheets, and which makes significant interactions with the cyclin, is conserved in all interfaces known to date (Echali et al., 2010). Eight positions between β3 and β5 have been found in five complexes (CDK2-cyclin A, CDK9-cyclin T, CDK4-cyclin D, CDK5-p25, Pho85-Pho80) to mediate the
interactions between a CDK and its respective cyclin partner. These positions correspond to CDK2 residues Glu41, Ile46, Val49, Val52, Ala53, Arg56, Arg57 and Val84. The interactions are different in different complexes due to the identities of the amino acids involved (Figure 2-8). A conserved feature of the CDK-cyclin interface is the predominance of hydrophobic interactions enhanced by aromatic residues present in the cyclins (Figure 2-8). The residues are not identical but collectively they create a hydrophobic surface that is a characteristic of this part of the CDK-cyclin interface.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Binary complex</strong></td>
<td><img src="image1.png" alt="Image" /></td>
<td><img src="image2.png" alt="Image" /></td>
<td><img src="image3.png" alt="Image" /></td>
<td><img src="image4.png" alt="Image" /></td>
</tr>
<tr>
<td><strong>Interfacial residues on CDK</strong></td>
<td><img src="image5.png" alt="Image" /></td>
<td><img src="image6.png" alt="Image" /></td>
<td><img src="image7.png" alt="Image" /></td>
<td><img src="image8.png" alt="Image" /></td>
</tr>
<tr>
<td><strong>Interfacial residues on cyclin</strong></td>
<td><img src="image9.png" alt="Image" /></td>
<td><img src="image10.png" alt="Image" /></td>
<td><img src="image11.png" alt="Image" /></td>
<td><img src="image12.png" alt="Image" /></td>
</tr>
<tr>
<td><strong>Interface area (Å²)</strong></td>
<td>1,640</td>
<td>970</td>
<td>1,012</td>
<td>1,350</td>
</tr>
<tr>
<td><strong>Active conformation</strong></td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
</tbody>
</table>

**Figure 2-8 Comparison of CDK-cyclin interfaces.**
Top row labelled ‘binary complex’ illustrates CDK-cyclin complexes for which structures have been determined. CDKs are coloured grey and cyclins green – cyclin A2, light green – cyclin T, magenta – cyclin D1, purple – Pho80. Activation segments are shown in red and C-helices in blue. Interface residues of CDKs are represented in ball and stick mode and coloured red. Interface residues of cyclins are coloured black. Interface areas are calculated by Pisa (Krissinel and Henrick, 2007). The Figure is redrawn with permission from (Echalier et al., 2010).
There are also a number of notable differences between the various CDK-cyclin interfaces. For example, the C-terminal part of αE and the activation segments of CDK2, CDK5 and Pho85 contact their respective cyclin partners (Figure 2-9). The first N-terminal residues of CDK9 contact cyclin T, and CDK2 residues between helices αJ and αK interact with cyclin A. CDK2 is notable for interacting with both the N- and C-CBFs of cyclins A, B and E, whereas cyclin D and cyclin T only use the N-terminal CBFs to contact respectively CDK4 and CDK9.

**Figure 2-9 Hydrophobic interactions at the CDK-cyclin interface.**
C-helix of CDKs shown in grey, helix H 5 of cyclins represented in different colours, cyclin A – green, cyclin E – orange, cyclin T – light green, cyclin D – magenta. Figure redrawn with permission from (Echalier et al., 2010).

### 2.1.6 Aims and objectives
Determination of the structures of a number of CDK-cyclin complexes has shown that not all the features of CDK2 activation by cyclin A binding are conserved across the CDK family. The primary aim of the work described in this Chapter was to assess to what extent the structural model for CDK activation proposed from studies of CDK2 is applicable to activation of its close relative CDK1. CDK1 and its binding partners cyclin B and CKS2 were expressed and purified to generate complexes for structural studies.
Structures of monomeric CDK1 and CDK1 bound to a cyclin partner can provide insights into features which make it the essential CDK and explanations for its flexibility towards its substrates. It also would allow investigation of residues involved in cyclin binding. The activation segment and ATP binding site are the important features of all kinases. An examination of the CDK1 active site can provide valuable insights into the mechanisms of CDK1 activation and of substrate phosphorylation.

2.2 Materials and Methods

2.2.1 Chemicals and Reagents

All chemicals listed in the material and methods section were of the analytical grade and obtained from Sigma Aldrich, St Louis, MO, USA, unless otherwise stated.

2.2.2 Purification buffers

**Extraction buffer 1:** 50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 0.01% MTG (monothioglycerol), 0.01 mg/ml DNase, 0.05 mg/ml RNase, 0.25 mg/ml Lysozyme, 5 mM MgCl₂, one EDTA-free complete protease inhibitor tablet (Roche, Basel, Switzerland)/50 ml.

**Extraction buffer 2:** 50 mM Tris-HCl, pH 8.0, 150 mM NaCl, 0.01% monothioglycerol (MTG), 0.01 mg/ml DNase, 0.05 mg/ml RNase, 0.25 mg/ml Lysozyme, 5 mM MgCl₂, one complete protease inhibitor tablet (Roche, Basel, Switzerland)/50 ml.

**Extraction buffer B:** 50 mM Tris-HCl, pH 8.0, 500 mM NaCl, 20 mM imidazole, 0.01% MTG, 0.01 mg/ml DNase, 0.05 mg/ml RNase, 0.25 mg/ml Lysozyme, 5 mM MgCl₂, one EDTA-free complete protease inhibitor tablet (Roche, Basel, Switzerland)/50 ml.

**Binding / SEC buffer 1:** 50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 0.01% MTG.

**Binding / SEC buffer 2:** 50 mM Tris-HCl, pH 8.0, 150 mM NaCl, 0.01% MTG.

**Binding / SEC buffer B:** 50 mM Tris-HCl, pH 8.0, 500 mM NaCl, 0.01% MTG.

**Elution buffer B:** 50 mM Tris-HCl, pH 8.0, 500 mM NaCl, 200 mM imidazole, 0.01% MTG.

**SEC buffer 1/B/2:** 50 mM Tris-HCl, pH 8.0, 1 M NaCl, 0.01% MTG.

**Peptide HBS:** 40 mM Hepes, pH 7.5, 200 mM NaCl, 0.01% MTG.
Binding Peptide MES: 25 mM MES, pH 6.5

Elution Peptide MES: 25 mM MES, pH 6.5, 500 mM NaCl.

All buffers were filtered and degassed before use.

2.2.3 Constructs used for in vitro experiments

CDK1 and CAK1 constructs were kindly generated by N. Brown. CDK1 (Uniprot entry P06493) was cloned by the In-Fusion method (Clontech) into the pVL1393 vector modified by N terminal insertion of the GST-3C sequence. Prior to CDK1 insertion the expression vector was linearized by Sma I and EcoRI digestion (NEB) and the CDK1 insert was amplified by PCR. GSTCAK was cloned in a similar manner although it contains a non-cleavable GST tag. Cyclin B (residues 165-433, Uniprot entry P14635) with mutations C167S/C238S/C350S in the pET28 -(a+) vector with thrombin cleavable 6xHis tag was a kind gift of E. Petri (hereafter called pET28cyclinBcs, (Petri et al., 2007)). CKS1 (Uniprot entry P61024) and CKS2 (Uniprot entry P33552) constructs were kind gifts from Prof. J. Endicott. CKS1 was amplified by PCR and sub-cloned into the pET21-(d+) vector with an N terminal 6xHis tag (pETCKS1). CKS2 was cloned into pGEX6P-1 to generate a construct with a 3C Precision cleavable GST tag (pGEXCKs2). The following vectors were available in the laboratory: CDK2 (Uniprot entry P24941) and Saccharomyces cerevisiae CAK cloned as GST fusions into pGEX6P-1 to generate CDK2 phosphorylated on Thr160 (pGEXCDK2-CAK1, (Brown et al., 1999)), and pET21d modified to express 6xHis-tagged bovine cyclin A2 (pETBvcyclinA, Uniprot entry P30274) (Table 2-1).

<table>
<thead>
<tr>
<th>Protein name</th>
<th>Names of constructs</th>
<th>Uniprot entry</th>
<th>Residue range</th>
<th>Vector backbone</th>
<th>Tag</th>
<th>Molecular weight (Da)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CDK1</td>
<td>pVL1393CDK1</td>
<td>P06493</td>
<td>1 - 297</td>
<td>pVL1393</td>
<td>GST</td>
<td>34095</td>
</tr>
<tr>
<td>Cyclin B1</td>
<td>pET28CyclinBcs</td>
<td>P14635</td>
<td>165- 433</td>
<td>pET28-(a+)</td>
<td>6xHis</td>
<td>30918</td>
</tr>
<tr>
<td>GST CAK1</td>
<td>pVL1393CAK1</td>
<td>P43568</td>
<td>1-368</td>
<td>pVL1393</td>
<td>GST</td>
<td>68186</td>
</tr>
<tr>
<td>CKS1</td>
<td>pET21CKS1</td>
<td>P61024</td>
<td>1-79</td>
<td>pET21-(d+)</td>
<td>6xHis</td>
<td>9660</td>
</tr>
<tr>
<td>CKS2</td>
<td>pGEX6P1CKS2</td>
<td>P33552</td>
<td>1-79</td>
<td>pGEX6P-1</td>
<td>GST</td>
<td>9860</td>
</tr>
<tr>
<td>CDK2</td>
<td>pGEX6P1CDK2</td>
<td>P24941</td>
<td>1 - 298</td>
<td>pGEX6P-1</td>
<td>GST</td>
<td>33930</td>
</tr>
</tbody>
</table>
Table 2-1 Summary of constructs used to express CDK1 and associated regulators

<table>
<thead>
<tr>
<th>Cyclin A2</th>
<th>pET21dCyclinAHis</th>
<th>P30274</th>
<th>171-433</th>
<th>pET21d</th>
<th>6xHis</th>
<th>29833</th>
</tr>
</thead>
<tbody>
<tr>
<td>160T pCDK2</td>
<td>pGEX6P1CDK2CAK1</td>
<td>P24941</td>
<td>1–298</td>
<td>P30274</td>
<td>1-368</td>
<td>GST</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>GST</td>
</tr>
</tbody>
</table>

2.2.4 Cell culture

Human CDK1 and S. cerevisiae CAK1 were expressed using the baculoviral expression system. The pVL1393 vector containing GSTCDK1 or GSTCAK1 was used to prepare recombinant baculovirus with the flashBAC system (OET, Oxford, UK). 1.5 x 10^6 SF9 cells in 2 ml of media per 35 mm dish were left for 30 minutes to form a sub-confluent monolayer. Meanwhile transfection mix was prepared in a polystyrene tube as follows - 1000 μl Lonza InsectXPRESS with L-Glutamine media (Lonza), 100 ng virus DNA from the flashBAC kit (5 µl), 500 ng of transfer vector containing gene of interest, and 5 µl of GeneJuice (Novagen, Merck Group, Darmstadt, Germany) were mixed and left at room temperature for 15 minutes. Once the SF9 cells had formed a monolayer, medium was drained and the transfection mix was added drop by drop to equally distribute it over the plate. 35 mm dishes containing SF9 cells and transfection mix were then sealed by parafilm and placed into sandwich boxes and incubated at 28°C for between 5 and 24 hours. After the initial incubation, an additional 1 ml of SF900 medium was added to each well and the plates were replaced into the sandwich boxes and incubated for a further 4 days at 28°C. As a result of homologous recombination within the insect cells, the GSTCDK1 and GSTCIV1 sequences were inserted into the viral DNA under the control of the polyhedrin promoter. Recombinant viruses were selected by restoration of the essential ORF1629 gene. Culture medium was harvested after five days in total and stored at 4°C. This sample constituted the P0 stock and was used for further amplification. P3 stock was used for protein production.

2.2.5 Optimisation of protein expression

Protein production was optimised by screening different virus stock:cell ratios. Generally, 30 ml of SF9 cells at 1.0 x 10^6 cells per ml were infected with different volumes of viral stock, typically 30 µl, 150 µl or 300 µl and incubated at 27°C with shaking at 120 rpm for a total of 98 hours. Cells were counted after 48 and 72 hours.
to check if they were still doubling every 24 hours. After they have stopped dividing it is assumed that the optimal concentration of virus has been achieved. 2 ml samples of each cell culture infected with viruses were harvested by centrifugation at 646 x g for 15 min after 48, 72 and 98 hours incubation and frozen at -20°C for further expression tests. Optimal conditions for infection were used for subsequent protein expression.

### 2.2.6 Expression tests

The frozen cell pellets were re-suspended in 1ml of BugBuster (Merck Millipore) with addition of 0.25 mg/ml Lysozyme, 50 μg/ml RNase, 10 μg/ml DNase and 5 mM MgCl₂. The cells were incubated with rotation for 20 min at room temperature for cell lysis. 3 μl of each lysate was analysed by SDS-PAGE and the rest was transferred into Eppendorf tubes and centrifuged for 20 min at 6500 x g. 50 μl of Glutathione Sepharose 4B beads (GE Healthcare) were equilibrated in TBS buffer (50 mM Tris-HCl, pH 7.5, 0.15 mM NaCl, 0.01% MTG), and supernatants of each sample were added and incubated for 1h at +4°C. Beads were then washed with ice-cold TBS 3 times in 0.5 ml volume and pelleted by centrifugation at 4°C at 1610 x g. After the final wash, 30 μl of RunBlue LDS Sample buffer (Expedeon) was added to the pelleted beads and proteins were eluted by heating at 95°C for 5 minutes. 15 μl of each sample was analysed by SDS-PAGE as described in Section 2.2.10.

### 2.2.7 Bacterial transformation and protein expression in bacterial cells

For protein expression *E. coli* Rosetta™2(DE3)pLysS cells were transformed using the manufacturer’s protocol. Briefly, circa 10 ng of pET28cyclinBcs or pGEX6PCks2 was incubated with cells (50 μl) on ice for 5 min. The mix was heat shocked for 30 sec and left on ice for a further 2 min. After one hour incubation in SOC media, cells were plated on LB agar containing kanamycin (30 μg/ml) and chloramphenicol (34 μg/ml) and incubated overnight at 37°C. For each construct, one colony was picked and grown overnight in 10 ml of Lysogeny Broth (LB) supplemented with appropriate antibiotics and then added to 1l LB containing kanamycin and chloramphenicol. The cultures were grown at 37°C at 220 rpm in a shaking incubator until the optical density at 600 nm (OD₆₀₀) reached 0.6-0.8 (cyclin B) or 0.5-0.6 (CKS2). The cultures were cooled to 20°C, induced with 0.3 mM IPTG and then left to grow overnight. Cells were harvested by
spinning at 4000 x g and re-suspended in either binding/SEC buffer B (cyclinB) or binding/SEC buffer 2 (CKS2) and then fast-frozen in dry ice at -80°C for future use.

For bovine Cyclin A expression, pET21CyclinA (Brown et al., 1995) was transformed into Rosetta™2(DE3)pLysS cells according to the manufacturer’s instructions and colonies resistant to ampicillin and chloramphenicol were harvested for the initial culture growth in 10 ml of LB o/n. Starter culture was added to 1 L LB containing 50 µg/ml of ampicillin and 34 µg/ml chloramphenicol and inoculated at 37°C at 220 rpm until the OD₆₀₀ nm reached 0.7. The flasks were cooled down to 20°C and cells were induced with 0.1 mM IPTG overnight at 220 rpm. Cells were pelleted by centrifugation at 5000 x g for 20 minutes and re-suspended in 20 ml of 300 mM NaCl, 0.01% MTG, 50 mM Tris-HCl pH7.5, 5 mM MgCl₂, protease Complete tablets minus EDTA (1 tablet for 50 ml). Re-suspended pellets were flash frozen in dry ice and stored at -80°C C until further use.

For the expression of phosphorylated CDK2, Rosetta™2(DE3)pLysS cells were transformed with the pGEX6PCDK2CAK1 (Brown et al., 1999) according to the manufacturer’s instructions. Cells were inoculated in 1 L LB at 37°C and incubated until the optical density at 600nm OD₆₀₀ nm has reached 0.8 and then induced by addition of 80 µM IPTG and left to grow for 20 hours at 20°C. This incubation generates CDK2 phosphorylated on Thr160 by CAK1. Cells were harvested and frozen as described above.

2.2.8 Protein purification

Cell pellets were thawed under running cold water and then kept on ice and sonicated in the presence of 0.25 mg/ml Lysozyme, 50 µg/ml RNase, 10 µg/ml DNase and 5 mM MgCl₂ for 10 cycles totalling 1.5 min with 15 sec intervals for insect cells, and 3 min with the same intervals for bacterial cells. Cells were centrifuged at 60,000 x g, the supernatants collected and kept on ice. Glutathione Sepharose 4B was used to isolate GST-tagged CDK1, CAK1 and CKS2 proteins. The Glutathione Sepharose 4B resin was packed into a gravity flow column and washed with 10 column volumes of binding/SEC buffer 1 for CDK1 and GSTCIV1 or binding/SEC buffer 2 for CKS2. The supernatant was applied to the column under gravity flow and then the column was washed to baseline with 10 column volumes of buffer. Bound protein was eluted using the same buffer supplemented with 20 mM reduced glutathione adjusted to pH 7.4.
CDK1 and CKS2 were cleaved from their GST tags by overnight incubation with 3C protease (1:50 w/w) at +4°C.

Cyclin B and Cyclin A are hexahistidine tagged and were purified by affinity chromatography using Ni-NTA resin. Beads were packed into a column and washed with 10 column volumes of binding/SEC buffer B supplemented with 20 mM imidazole to reduce non-specific binding for cyclin B and with buffer N (300 mM NaCl, 50 mM Tris-HCl, pH 7.5, 0.01% MTG, 25 mM imidazole) for cyclin A. Bacterial lysates supplemented with 20 mM and 25 mM imidazole respectively after thawing were applied under gravity flow to the prepared columns. Beads were washed with 10 column volumes of respective buffers and then cyclin B eluted with binding/SEC buffer B containing 200 mM imidazole. Cyclin A was eluted by step increasing the concentration of imidazole at 100 mM, 150 mM, 200 mM and 500 mM intervals to increase the purity of the protein. MgCl$_2$ was added to cyclin A fractions to a final concentration of 0.1 M to prevent protein aggregation. Thrombin clean cleave kit (Sigma) was used to remove the hexahistidine tag from cyclin B. Beads containing thrombin were equilibrated into binding/SEC buffer B. Cyclin B eluted from the Ni-NTA resin was buffer exchanged into the same buffer as imidazole affects thrombin cleavage. Cyclin B was incubated for a minimum of 3 hours at 4°C with thrombin clean cleave beads (200 µl of 50 % slurry cleaves 1 mg of fusion protein). The efficiency of the cleavage was checked by SDS-PAGE.

### 2.2.9 Protein purification by size-exclusion chromatography

Size-exclusion chromatography (SEC) was used as a final purification step. A Superdex 75 26/60 column was equilibrated into the appropriate buffer: binding/SEC buffer 1 for CDK1 purification, binding/SEC buffer 2 for CKS2 purification and binding/SEC buffer B for cyclin B purification. SEC was carried out at room temperature at a flow rate of 2.5 mls/min. Samples (between 5 and 10 mls) were filtered through a 0.45 μm filter before loading. Four ml fractions corresponding to every peak were collected and the protein content estimated by OD$_{280}$ measurement (Nanodrop) prior to analysis by SDS-PAGE.

### 2.2.10 SDS-PAGE

Protein samples were prepared by mixing with an equal volume of 4xsample buffer (40% glycerol, 4% LDS, 0.8 M Triethanolamine-Cl pH 7.6, 4% Ficoll-400, 0.025% Phenol Red, 0.025% Coomassie Brilliant Blue G250, 2 mM EDTA-2Na) and 0.25 M
DTT as a reducing agent (15 mM final) and then heated for 3 min at 95°C. Between 10-20 µl of sample was loaded onto 12% polyacrylamide gels (Expedeon, Cambridgeshire, UK) alongside a pre-stained protein ladder (Fermentas, Waltham, Massachusetts). Following electrophoresis (1 h at 180 V), proteins were visualised with Instant Blue (Expedeon, Cambridgeshire, UK).

### 2.2.11 Preparation of complexes

Fractions containing CDK1, cyclin B and CKS2 were further purified by affinity chromatography to remove GST and uncleaved proteins respectively. Proteins were concentrated (Sartorius Vivaspin concentrators (10 kDa and 3 kDa cut-off)) and further analysed by SDS-PAGE for the presence of impurities. CDK1-cyclin B-CKS2 and CDK1-cyclin A–CKS2 were mixed in a ratio of 1:1.5:2. The NaCl concentration was again adjusted to 1 M prior to a 1 hr incubation on ice and then the ternary complexes were purified using a Superdex 75 26/60 column. In each case, the peaks corresponding to complexes of CDK1-cyclin B-CKS2, CDK1-cyclin A and CDK1-cyclin A-CKS2 were collected and analysed by SDS-PAGE. Fractions of CDK1-cyclin A were concentrated to 1 mg/ml using 30 kDa cut-off concentrators (Vivaspin) and frozen in 100 µl aliquots. Pooled fractions of CDK1-cyclin B-CKS2 and CDK1-cyclin A-CKS2 were concentrated in the presence of either 0.2 mM AMPPNP or a 3-fold molar excess of ATP-competitive inhibitors to 10 mg/ml using 30kDa cut-off concentrators (Vivaspin) prior to crystallisation trials.

### 2.2.12 Protein crystallography

Sitting drop vapour diffusion crystallization trials were performed using homemade screens designed from an initial hit from the Morpheus screen (Molecular Dimensions, Newmarket, Suffolk, UK) condition B4 (0.1M MES/imidazole buffer (pH 6.7), 6.5% MPD, 5% PEG 4K, 10% PEG1K)). Drops were dispensed by a TPP Labtech Mosquito LCP into MRC 96 wells plates by mixing 0.2 µl of complex and 0.2 µl of mother liquor. Plates were incubated at +4°C and monitored for 4 weeks. Final crystallisation conditions were identified as 0.1 M MES/imidazole buffer (pH 6.7), 6.5% MPD, 5% PEG4K, 10% PEG1K. Crystals were mounted into a nylon crystal loop and flash cooled in liquid nitrogen prior to data collection.

### 2.2.13 Data collection and analysis

Datasets were collected at the Diamond Light Source (Didcot, UK) on I04 and I24 beamlines supplied with detector PILATUS 6M (Broennmann et al., 2006) at 100 K at
the range of wavelength 0.69-2.07Å by Dr A. Basle. Typically 2000 images were collected from a single crystal with an oscillation angle of 0.1° per image and an exposure time of 0.1 s. Data were imported from Xia (Winter, 2010) and processed using programs of the CCP4 suite (Pointless/ Aimless (Evans, 2006; Evans, 2011)). The structures of CKS2 (PDB code 1BUH), CDK2 (PDB code 1HCK), human cyclin B (PDB code 2B9R) and cyclin bound CDK2 (PDB code 1QMZ) were used as search models for molecular replacement by Phaser (McCoy et al., 2007). The structures were refined using cycles of manual rebuilding in COOT (Emsley et al., 2010) followed by refinement by Refmac5 (Murshudov et al., 1997). The final models were evaluated using Molprobity (Chen et al., 2010).

2.3 Results and Discussion

2.3.1 Purification of CDK1, cyclin B and CKS2 and complex assembly

To assemble CDK1-containing complexes, proteins were expressed and purified separately. CDK1 was produced in insect cells following recombinant baculovirus infection (as described in Material and Methods, Section 2.2.4). This virus produced circa 12 mg CDK1/l of SF9 cell culture. The yield was improved by changing the media from SF900 III SFM (Life Technologies) to Lonza Insect XPRESS with L-Glutamine (Lonza). Cyclin B and CKS2 were successfully expressed and purified from recombinant E. coli cells. Typical yields for cyclin B and CKS2 were around 3 mg/l LB and 5 mg/l LB respectively. Representative size-exclusion chromatograms for the final purification step in the preparation of CDK1, cyclin B and CKS2 are shown in Figures 2-10, 2-11 and 2-12 respectively. Dimerization of GST generates a complex of a similar size to CDK1 and as a result the two proteins are not completely resolved by SEC (Figure 2.10 A). After cleavage of the GST tag, two bands can be seen in the expected positions because insect cells express their own endogenous GST of a slightly higher molecular weight of 27614 Da over the tag mass of 26000 Da. Consequently, the fractions containing CDK1 had to be pooled and re-applied to Glutathione Sepharose beads in order to get rid of residual GST.
Figure 2-10 CDK1 purification.
(A) SEC chromatogram of the final CDK1 purification step. Absorbance at 280 nm is plotted against elution volume in ml. Fractions collected for SDS-PAGE analysis are identified with a red bar. (B) Peak fractions (labelled with a red bar) were analysed by SDS-PAGE. Identified proteins are labelled on the side of the gel. The gel was subsequently stained with InstantBlue.

Cyclin B has a thrombin cleavable C-terminal 6 x His tag that was exploited using metal chelate chromatography as a first purification step. After the elution step, cyclin B was buffer-exchanged into SEC/Binding buffer B to remove imidazole that would prevent thrombin cleavage. After the His tag was successfully cleaved, cyclin B was subjected to SEC for further purification (Figure 2-11 A). As can be seen (Figure 2-11 B, C) the His tag was successfully cleaved so all fractions were pooled and concentrated for further use.
Figure 2-11 Cyclin B purification.
(A) SEC chromatogram of cyclin B following cleavage of the 6xHis tag. Absorbance at 280 nm is plotted against elution volume in ml. Fractions identified with a red bar were analysed by SDS-PAGE (B). (C) Comparison of cyclin B tagged with 6xHis and cyclin B cleaved from its tag. Gel has been spliced although positions of the bands relative to the ladder remains unchanged.

CKS2 was also successfully purified using an affinity chromatography step with Glutathione Sepharose 4B beads to pull CKS2 out of solution by its GST tag, followed by size-exclusion chromatography to buffer exchange the sample and remove the excess GST and aggregates (Figure 2-12). All CKS2 containing fractions were pooled together for further experiments.
Figure 2-12 CKS2 purification.
(A) SEC chromatography of CKS2. Absorbance at 280 nm is plotted against elution volume in ml. The first peak comprises GST dimer and is identified with a blue bar. The second peak is pure CKS2 identified with a red bar. (B) Fractions identified by the blue and red bars analysed by SDS-PAGE. Lane 1 contains the cleaved GSTCKS2 complex that has been injected into the Superdex 75 26_60 column.

To assemble the CDK1-cyclin B-CKS2 complex, proteins were individually concentrated to circa 0.5 – 2 mg/ml and then mixed in a ratio of 1:1.5:2 CDK1:cyclin B:CKS1. Subsequent size-exclusion chromatography resolved the complex from the excess of individual components (Figure 2-13). Fractions containing the complex were combined and concentrated to 10 mg/ml for crystallisation trials.
Figure 2-13 Separation of CDK1-cyclin B-CKS2 complex.
(A) Proteins were mixed together in a 1:1.5:2 ratio and subjected to Size Exclusion Chromatography (SEC). Absorbance at 280 nm is plotted against elution volume in ml. The complex elutes first and is identified with a red bar on the chromatogram. Cyclin B and CKS2 were confirmed by SDS-PAGE and are identified with yellow and green bars respectively. (B) Fractions analysed by SDS-PAGE. Proteins are identified on the side of the gel.

2.3.2 Structure determination of CDK1-containing complexes
Sitting drop crystallization trials were initially performed for CDK1-CKS1 and CDK1–cyclin B–CKS2 complexes as described in Materials and Methods using a variety of commercially available screens by Dr N. Brown. Attempts to crystallise CDK1 on its own and CDK1–cyclin B were not successful (Dr N. Brown, personal communication). Addition of proteins of the CKS family improved crystallisation significantly. Drops were set in ratios of 1:1 and 2:1 (protein : well solution) in volumes of 200 nl or 400 nl and incubated at 4°C. Single crystals of CDK1–CKS1 were identified in 16-18% PEG 10K, 0.1 M imidazole pH 8.2- 8.4. Small diffracting crystals were observed in Morpheus (Molecular Dimensions) condition B4 (12.5% w/v PEG 1000, 12.5% w/v PEG 3350, 12.5% v/v MPD, additive mix included 0.03 M of halide, 0.1 M MES/imidazole pH 6.5) which was further optimized by changes to the concentrations of PEG 1000, PEG 3350, MPD and the pH of MES and to variations in the molecular masses of the PEGs. Few crystals grew from the optimized screens. The best diffracting ones came from a
well containing 5% PEG 4K, 10% PEG1K, 6.5% MPD and 0.1 M MES pH 6.7. Crystals of CDK1-CKS2 were cryo-protected by 30% ethylene glycol in mother liquor. The CDK1-cyclin B-CKS2 mother liquor does not require cryoprotection. Crystals were flash cooled in liquid nitrogen before data collection.

**Figure 2-14 Crystallisation trials of CDK1-cyclin B-CKS2 complex**
Crystals from the first hit from Morpheus condition B4. Crystal size around 0.1 mm. (B) Crystals from the optimized screen. Crystal size around 0.25 mm.

Datasets complete to 99.7% for CDK1-CKS1 and 99.6% for CDK1-cyclin B-CKS2 respectively at 2.8 Å were processed using programs from the CCP4 suite (Bailey, 1994) by Dr N. Brown and Professor M. Noble. Structures were solved by molecular replacement using program Phaser (McCoy et al., 2007). CDK2, cyclin B and CKS2 were used as search models. Crystal structures were determined to a resolution of 2.8 Å. A summary of the statistics for data collection and subsequent structure determination and refinement are presented in Table 2-2 and Table 2-3.
<table>
<thead>
<tr>
<th></th>
<th>CDK1-CKS1</th>
<th>CDK-cyclin B-CKS2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data collection</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Space group</td>
<td>P2₁</td>
<td>P2₁ 2₁ 2₁</td>
</tr>
<tr>
<td><strong>Cell dimensions</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a, b, c (Å)</td>
<td>66.8, 147.5, 87.3</td>
<td>69.2, 70.2, 156.2</td>
</tr>
<tr>
<td>α, β, γ (°)</td>
<td>90.0, 92.1, 90.0</td>
<td>90.0, 90.0, 90.0</td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>75.10-2.80 (2.91-2.80)</td>
<td>52.19-2.80 (2.95-2.80)</td>
</tr>
<tr>
<td>Rmerge</td>
<td>0.08 (0.89)</td>
<td>0.15 (0.57)</td>
</tr>
<tr>
<td>I/δ(I)</td>
<td>15.3 (1.8)</td>
<td>7.2 (2.0)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>99.7 (98.8)</td>
<td>99.6 (97.9)</td>
</tr>
<tr>
<td>Redundancy</td>
<td>5.9 (5.2)</td>
<td>3.6 (3.2)</td>
</tr>
<tr>
<td><strong>Refinement</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>75.10 – 2.80</td>
<td>78.10 - 2.80</td>
</tr>
<tr>
<td>No. reflections</td>
<td>41.491 (1.787)</td>
<td>19.278 (1.014)</td>
</tr>
<tr>
<td>R]&lt;sub&gt;work&lt;/sub&gt;/R]&lt;sub&gt;free&lt;/sub&gt;</td>
<td>0.234 (0.269)</td>
<td>0.209 (0.282)</td>
</tr>
<tr>
<td><strong>No. atoms</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein</td>
<td>23.200</td>
<td>10.377</td>
</tr>
<tr>
<td>Ligand/ion</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Water</td>
<td>0</td>
<td>33</td>
</tr>
<tr>
<td><strong>B-factors</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein</td>
<td>36.2</td>
<td>39.2</td>
</tr>
<tr>
<td>Ligand/ion</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Water</td>
<td>NA</td>
<td>14.8</td>
</tr>
<tr>
<td><strong>Root mean-squared deviations</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bond lengths (Å)</td>
<td>0.012</td>
<td>0.013</td>
</tr>
<tr>
<td>Bond angles (°)</td>
<td>1.679</td>
<td>1.696</td>
</tr>
</tbody>
</table>

Table 2-2 Summary of CDK1-CKS2 and CDK1-cyclinB-CKS2 X-ray data collection and refinement statistics
2.3.3 **Structural details of the CDK1-CKS1 complex**

The CDK1-CKS1 crystals contain four heterodimers with a large interface between the pairs in the asymmetric unit (Figure 2-15).

![Figure 2-15 Crystal lattice of the CDK1-CKS2 complex.](image)

(A) CDK1–CKS1 ribbon (PDB code 4YC6) diagram representing the crystal lattice that contains 4 copies of a CDK1-CKS1 dimer in the crystallographic asymmetric unit. CDK1 molecules are coloured blue, with the C-helix (residues 47-57) in gold, the activation segment (146-173) in purple and the pre C-helix hairpin (40-46) coloured pink. CKS1 is coloured orange. (B) Schematic representation of the crystallographic asymmetric unit. Proteins are represented in the same colour scheme (Figure prepared by M. Noble).

The CDK1 and CDK2 sequences show 65% similarity and a high degree of sequence conservation (Figure 2-16). The sequences are conserved throughout their lengths, but there are three regions of around 10 amino acids long, Ser93—Ser103, Lys245-Asn255 and the C terminal tail starting at Leu290, that diverge. These sequences are on the CDK1 surface. The first sequence is found on a loop linking αD and αE and includes the peptide sequence before the start of αH and is involved in CKS1 binding (Figure 2-16). The other two sequences lie away from that binding site but their high degree of conservation across CDK1 orthologues suggests that they may identify a binding site for other proteins.
Figure 2-16 Sequence alignment of human CDK1 and CDK2.
Sequence alignment was performed by Jalview (Waterhouse et al., 2009) and identical sequences are highlighted in dark blue. Sequences that differ between CDK1 (Uniprot entry P06493) and CDK2 (Uniprot entry P24941) Ser93—Ser103, Lys245—Asn255, Leu290—Met297 of CDK1) are labelled with a green bar.

2.3.4 Comparison of the CDK1-CKS1 and CDK2-CKS1 structures
Monomeric CDK1 adopts a classical bi-lobal protein kinase fold in an inactive conformation. The C helix is displaced out of the active site cleft and the non-phosphorylated activation segment (defined as the sequence between the DFG and SPE motifs starting at Asp146 and finishing at Glu173) is positioned in a way which is not capable of substrate recognition (Figure 2-17).

The major differences between CDK1-CKS1 and CDK2-CKS1 were observed in the activation segment, the loop that precedes the C helix and at the C-termini. In most CDK2 structures the activation segment is relatively mobile, but in some it forms a β-hairpin that reaches the glycine-rich loop and occludes the active site from solvent (De Bondt et al., 1993). The glycine-rich loop of CDK1 is similar to CDK2 in its mobility but electron density for residues 156-162 was missing suggesting that the CDK1 activation segment is unstructured in this region. Both molecules form a short α-helical structure (aL12) before the DFG motif that displaces the C–helix from the active site (Figure 2-17). The only difference between the two CDKs in this region is the location of the phenylalanine sidechain of CDK1 it is rotated towards the carbonyl moiety of Val64, whereas in CDK2 it points into the active site towards Asp185. There
are no visible differences between the structures at the end of the activation segments, but most of the CDK1 activation segment cannot be traced as it is not supported by the electron density map. After the activation segment the structures are well ordered and composed of a number of helices. Towards the C-terminus, after Leu290, the CDK1 sequence is disordered in all four CDK1 molecules in the asymmetric unit, but in CDK2 the comparable residues bend along the back of the fold towards the hinge.

**Figure 2-17 Comparison of CDK1-CKS1 and CDK2-CKS1.**
The complexes are shown in identical views and coloured in the same colours, CDK1 (PDB code: 4YC6) (A) and CDK2 (PDB code: 1BUH) (B) are presented blue and CKS1 is coral. Key regulatory elements are highlighted as follows, the C-helix (residues 46-57) is coloured gold, the activation segment (CDK1, 146-173, CDK2, 145-172) is coloured red, and the αL-12 helix at the beginning of the activation segment is identified with an arrow. The pre C-helix hairpin (40-46) and the P-loop (13-16) are coloured pink and dark blue respectively. DFG and SPF motifs at the beginning and end respectively of the activation segment are represented in sticks and balls. Figure prepared using CCP4mg (McNicholas et al., 2011).

The CKS1 binding site is highly conserved between CDK1 and CDK2 with very few different residues participating in the interaction (Figure 2-18). The positioning of CKS1
in both complexes is very similar. CKS1 binds to the CDK C-lobe using all four β strands, the β1-β2 loop and the β-hinge region to form a saddle-shaped interface between helix α5 and loop L14 (Bourne et al., 1996). The localisation of the binding site to the CDK C-terminal lobe suggests that binding to CKS2 is unlikely to trigger any conformational change in the CDK. This has been confirmed for CDK2 by comparison with structures that are not bound to CKS (Cox et al., 1994; Goldsmith and Cobb, 1994). From the CKS1 side, most of one face across the whole molecule contributes residues to the CDK2 interaction.

**Figure 2-18 CDK-CKS2 binding interface.**

(A) The binding interface is highly conserved between CDK1 or CDK2 (PDB entry: 1BUH) and CKS1. CDK residues within 7Å of CKS1 are identified on the surface of CDK1. Identical residues are shown in green, conserved residues are yellow and residues that are different between CDK1 and CDK2 are red. Figure prepared by Prof. M. Noble and used by agreement with (Brown et al., 2015). (B) The CDK molecule is represented in ice blue (illustrated with CDK1). The CKS1 β sheets that interact with the CDK are represented in different colours and labelled. The CDK L14 loop is coloured red and the α5 helix is coloured green. Figure B created using CCP4mg (McNicholas et al., 2011).
2.3.5 The overall fold of the CDK1-cyclin B-CKS2 complex

The structure of CDK1-cyclin B-CKS2 was initially determined by Dr N. Brown at 2.8 Å by molecular replacement as described in the Materials and Methods. Further optimisation of the protein preparation and crystallisation conditions (by addition of an ATP-competitive inhibitors, (Chapter 5) and additional crystal testing improved the resolution of the dataset to 2 Å (Table A1). The overall fold of CDK1-cyclin B is different to CDK2-cyclin A (Jeffrey et al., 1995) and to CDK4-cyclin D (Takaki et al., 2009). Due to changes in the orientation of the C helix in comparison to CDK2-cyclin A, CDK1 makes less extensive interactions with cyclin B resulting in a 30% smaller binding interface.

The most significant difference in the structure of the CDK1-cyclin B-CKS1 complex to that of CDK1-CKS1 is that in the ternary complex the CDK1 C terminal tail forms an amphipathic helix that, due to crystal packing, binds to a symmetry-related cyclin B molecule (Figure 2-18). As discussed earlier, the residues towards the CDK C termini might contain a potential protein binding site as it is a region that differs in sequence between the CDKs. It can be observed that the hydrophobic residues Leu290, Ile294 and Met297 of CDK 1 bind to the cyclin B groove of a neighbouring molecule in the lattice. A similar groove exists on the surface of cyclin A which has previously been shown to bind to the CDK inhibitor p27KIP1 (Russo et al., 1996a) (Figure 2-19). Respective residues Leu41, Leu45 and Cys49 of p27KIP1 bind to cyclin A in a similar manner to CDK1 binding to cyclin B. The lattice contact suggests that the C terminal tail of CDK1 has a tendency to unfold away from the main fold and that this surface of cyclin B has the potential to mediate multiple protein interactions.
**Figure 2-19 Interactions between CDK1 C-terminal tail and cyclin B in the crystal lattice.**

(A) Complexes labelled as Comp1 and Comp2. The C terminal tail of CDK1 (284-297) binds to the groove of cyclin B from the neighbouring molecule in the lattice. CDK1 is coloured ice blue, the cyclin B is coloured green and the CKS1 subunit is orange. The binding site of C-terminal tail and cyclin B from another symmetry unit is magnified. CDK1 hydrophobic residues contributing to the cyclin B interaction, (Leu290, Ile294, and Met297) are drawn in ball and stick mode. (B) Structure of CDK2-cyclin A with p27KIP1 (PDB code IJSU). CDK2 is coloured in ice blue, cyclin A is coloured green and p27KIP is pink. p27KIP residues Leu41, Leu45 and Cys49 bind to cyclin A in a manner similar to that observed between the CDK1 C terminal helix and cyclin B. Figures prepared using CCP4mg (McNicholas *et al.*, 2011).
2.3.6 Comparison of CDK1-cyclin B-CKS2 with CDK2 and CDK4 containing structures

CDK1-cyclin B has a different structure in comparison to CDK2-cyclin A and to CDK4-cyclin D. Cyclin A binding to CDK2 leads to reorganisation of residues forming the active site by melting the αL12 helix and moving the C helix inwards. These re-arrangements bring residues Arg50 on the C helix, Arg126 of the catalytic loop and Arg150 of the activation segment into positions ready to accept peptide substrate within the activation segment. Cyclin binding also re-organises the activation segment to form a platform for peptide substrate recognition. Once bound to cyclin A, CDK2 adopts a conformation that requires very little change to be able to accept the two substrates for phosphorylation (Figure 2-20).

Figure 2-20 Comparison of CDK1-cyclin B-CKS2 structure with CDK2-cyclin A and unphosphorylated CDK4-cyclin D3.

(A, D) CDK1-cyclin B-CKS2, (B, E) CDK2-cyclin A (PDB code 1FIN), (C, F) CDK4-cyclin D3 (PDB code 3G33). All complexes are shown in the same view. Cyclins are depicted green, CDKs are ice blue, C helix of CDKs is coloured gold and activation segment is red. CKS2 is coral. Figures prepared using CCP4mg (McNicholas et al., 2011).
CDK4, in contrast to CDK2, does not adopt an active conformation even when bound to cyclin D3. In the CDK4-cyclin D3 complex, the C helix remains displaced from the active site and the activation segment, although ordered, does not form contacts with the cyclin subunit. Respective residues from the C helix, catalytic loop and activation segment, (Arg55, Arg139, Arg163), as well as Thr172 (the phosphorylated residue equivalent to Thr160 in CDK2), remain dispersed. The activation segment is also not in a conformation that would recognise a peptide for phosphorylation. From these observations it has been proposed that CDK4-cyclin D complexes proceed through a substrate-assisted catalytic mechanism.

CDK1-cyclin B is yet different but more similar to CDK2-cyclin A. One major difference between the two structures is in the activation segment, which forms extensive contacts with cyclin A in the CDK2 complex but shows fewer connections to cyclin B in CDK1-containing structures. The distinctive feature of the CDK1 activation segment is the β hairpin which has residues Phe153 and Gly154 at its tip. The hairpin sits in a hydrophobic groove formed by the aliphatic portion of Arg299, Leu178 and His304 of cyclin B. An interesting feature of CDK1 binding to cyclin B is that the cyclin sequence forms a loop after residues Met261 and Tyr262 at the end of the α3 helix which then occludes Ile155 of CDK1 from solvent (Figure 2-21). This movement of the cyclin behind the activation segment creates a solvent filled cleft also seen in the CDK4-cyclin D and CDK9-cyclin T structures (Baumli et al., 2008) but not in CDK2-cyclin B (Brown et al., 2007).

**Figure 2-21 CDK1 interactions with cyclin B.**

CDK1 forms a β hairpin with residues Phe153 and Gly154 at the tip. These residues make tight interactions with the cyclin B groove formed by Arg299, Leu178 and His304 which are characteristic only for the CDK1-cyclin B structure. Cyclin B is coloured green, structural features of CDK1 identified as follows: N terminal lobe grey, C
terminal lobe, ice blue, activation segment, red, and the C helix is coloured gold. Figure prepared using CCP4mg (McNicholas et al., 2011).

As a result of the movements at the tip of the CDK1 activation loop described above, the loop then bends upwards at Ile157 and causes the isoleucine sidechain to point towards the C helix. To accommodate the sidechain, the C-helix and DFG motif are displaced further out of the active site cleft creating a further difference with the CDK2 structure in this region. The distance between the side chain carboxylate of Glu51 and the Lys33 ε amino group extends to 5.0 Å in CDK1 compared to 3.1 Å in CDK2-cyclin A. The arginines that coordinate the activation segment phosphothreonine (Arg50, Arg127 and Arg151) are also displaced and form interactions centred around CDK1 Tyr181 and also with the backbone carbonyl of cyclin B Lys257.

The activation segment forms the platform that recognises the peptide substrate to either side of the phosphoacceptor residue (Brown et al., 1999). Cyclin A binding to CDK2 remodels the position of the activation segment so that it requires very few changes to adopt the active conformation and be ready to accept and phosphorylate a substrate. The smaller CDK1-cyclin B interface forms fewer interactions between the activation segment and the cyclin unit and as a result the activation segment requires considerable re-arrangement to be able to recognise a peptide substrate.

To summarise, CDK1, CDK2 and CDK4 in complexes with their respective cyclins adopt different structures that may contribute to their different patterns of substrate phosphorylation. The fact that CDK1 forms fewer bonds with cyclin B then CDK2 suggests this feature to be a property of CDK1 rather than the cyclin. A comparative study of CDK2 binding to cyclin A (Lee and Nurse, 1987) and cyclin B (Brown et al., 2007) shows that CDK2 adopts an identical structure bound to either cyclin.

2.3.7 Attempts to determine structures for CDK1-cyclin A and CDK1-cyclin A-CKS2

Bovine cyclin A2 was successfully expressed in E. coli cells. After loading on a prepacked Ni-NTA column cyclin A was washed with increasing concentrations of imidazole to significantly improve its purity (Figure 2-22 A, B). The first peak, eluted after washing the column with 50 mM imidazole, contains a substantial mix of other proteins as well as cyclin A (Figure 2-22 B). SDS-PAGE analysis of fractions from the following two peaks (Figure 2-22 C) shows that cyclin A eluted at 75 mM and 100 mM
imidazole contains significantly smaller amounts of other proteins. Fractions of peaks 2 and 3 were pooled and applied to the Superdex 75 26_60 column for further purification. Although MgCl₂ was added to prevent aggregation after the elution from the Ni-NTA column, slight aggregation did occur as evidenced by the peak around 100 ml. Fractions of the peak eluting around 170 ml were collected and analysed by SDS-PAGE gel (Figure 2-22 C, D). Cyclin A appeared very clean on a gel so fractions were pooled together for complex assembly.

CDK1 was expressed and purified and mixed with cyclin A at a molar ratio of 1:1.3 CDK1:cyclin A and at micromolar concentration so that cyclin A would saturate all of the CDK1 in solution. To separate the CDK1-cyclin A complex from the excess of unbound cyclin A, the mixture was subjected to another run on the Superdex 75 26_60 column (Figure 2-23 A) followed by SDS-PAGE analysis.
Figure 2-22 Purification of bovine cyclin A.

(A) Affinity chromatogram of cyclin A. Absorbance at 280 nm is plotted against elution volume in ml. Cyclin A bound to Ni-NTA beads was washed with increasing concentrations of imidazole (in 25 mM increments). Two eluted peaks, marked by the red and blue bars were collected and analysed by SDS-PAGE (B). (C) SEC of cyclin A after affinity chromatography. The second peak is cyclin A identified with a red bar and fractions run on SDS-PAGE (D).
Figure 2-23 Purification of CDK1-cyclin A.
(A) Size-exclusion chromatogram of the CDK1-cyclin A complex. The CDK1-cyclin A peak is identified with a red bar, excess of cyclin A with a blue bar. Absorbance at 280 nm is plotted against elution volume in ml. The chromatogram below shows the SEC standards (albumin, 66kDa; carbonic anhydrase, 29 kDa; cytochrome C, 12.4 kDa; aprotinin, 6.5 kDa) to calibrate the column. (B) SDS-PAGE analysis of fractions from the peaks indicated by red and blue bars. The lanes labelled 1, 2 and 3 contain respectively pure CDK1, pure cyclin A, and CDK1-cyclin A taken for crystallisation.

The CDK1-cyclin A complex was concentrated to 1 mg/ml and divided into two samples. One was concentrated to 10 mg/ml, the second was mixed with 3-fold excess of compound 23 (Wyatt et al., 2008) and further concentrated to 10 mg/ml, final DMSO concentration 2%. Crystal trays were set using a variety of crystallisation screens. Crystals appeared in Tacsimate 70% (Figure 2-24) but did not diffract. The crystallisation screens were then widened by exploring different concentrations of Tacsimate and addition of NaCl in different concentrations (Figure 2-24 A). In new crystallisation conditions, more crystals grew at around 50% Tacsimate, and the results indicated the importance of pH. pH 7.0 and 7.5 gave the most promising crystallisation conditions. Crystals were harvested but diffraction was very poor. In a subsequent round of optimisation ethylene glycol was introduced into the screen at low concentrations of NaCl and at a pH of 7.0 with the Tacsimate mixture as the buffer (Figure 2-24 B). To stabilize the protein three compounds were added, AZD 5438, RO-2206 and CGP74514A. These compounds were chosen based on crystallisation trials of CDK1-cyclin B-CKS2 where well diffracting crystals were grown with these inhibitors. Crystals were harvested but unfortunately did not diffract.
Figure 2-24 Home designed crystallisation screens for the CDK1-cyclin A complex.

(A) Screen was designed by changing the concentration of Tacsimate across the plate from row A to H from 30 to 70% exponentially, NaCl from 0 to 0.5 M from column 1 to column 4, 5 to 8, and 9 to 12 with the plate split into 3 fields with a different pH of Tacsimate. (B) Sample of crystals grown in well C8. (C) Homemade screen. Screen was set into four fields of repeats, where Tacsimate pH 7.0 was varying from 30 to 60% in columns 1 to 6 and 7 to 12 and ethylene glycol was screened from 0 to 25% in rows A to D and E to H exponentially. The concentration of NaCl was 0.16 M across the plate. Different inhibitors were used in each quarter of a plate. (D) Crystals produced in well D4.

Because CDK1 in complex with cyclin B or on its own did not form crystals, but addition of CKS2 lead to complex crystallisation, the same strategy was applied in an attempt to crystallise a CDK1-cyclin A-CKS2 complex. To improve the quality of the crystals and increase the possibility of diffraction, CKS2 was introduced to the complex. Purified proteins were mixed in a molar ratio of 1:1.5:1.5 CDK1:cyclin B:CKS2 and the complex was separated on a Superdex 75 26_60 size exclusion column. The ternary complex was successfully separated from the individual components (Figure 2-25).
Figure 2-25 Purification of a CDK1-cyclin A-CKS2 complex.

(A) Size-exclusion chromatogram of CDK1-cyclin A-CKS2. The elution position of the complex is identified by the red bar. Absorbance at 280 nm is plotted against elution volume in ml. The chromatogram below shows the SEC standards (albumin, 66 kDa; carbonic anhydrase, 29 kDa; cytochrome C, 12.4 kDa; aprotinin, 6.5 kDa) to calibrate the column. (B) SDS-PAGE analysis of fractions identified by the purple bar. The first three samples labelled on the gel are the individual proteins (1, CDK1; 2, cyclin A; 3, CKS2) and the column load (mix of CDK1, cyclin A and CKS2 in ratio 1:1.5:1.5) labelled 4 on the gel. (C) Crystals grown in the PACT premier TM HT96. Crystal size around 0.1 mm.

Fractions of the CDK1-cyclin A-CKS2 peak were pooled together and concentrated with or without the inhibitors AZD5438, RO-2206 and CGP74514A. Protein complex without added inhibitor formed crystals in three wells of the PACT premier TM HT96 screen (Molecular Dimensions) incubated at RT. They were E8 (0.2M Sodium sulphate, 20% w/v PEG 3350), G8 (0.2M Sodium sulphate, 0.1M Bis-Tris propane pH 7.5, 20% w/v PEG 3350) and F8 (0.2M Sodium sulphate, 0.1M Bis-Tris propane pH 6.5, 20% w/v PEG 3350). Crystals were cryo-protected in 30% ethylene glycol in mother liquor of a corresponding well followed by harvesting by a nylon loop and flash
cooling in liquid nitrogen. Despite their promising appearance, the crystals did not diffract well and a dataset was not collected.

2.4 Conclusions

This chapter focuses on novel aspects of the structures of CDK1 and CDK1-containing complexes. One significant difference between CDK1 and CDK2 bound to their cognate cyclin partners is the conformation of the activation segment. CDK1 and CDK2 have a highly conserved activation segment sequence but there are reported differences in activatory phosphorylation of both kinases on Thr160 in CDK2 and Thr161 in CDK1 by CAK, the CDK7-cyclin H complex. CAK phosphorylates CDK1 and CDK2 in a different way. It has been reported that CDK2 can be phosphorylated as a monomer before cyclin binding whereas CDK1 requires cyclin binding for activation segment phosphorylation (Merrick et al., 2008). The structure of monomeric CDK2 is well ordered at the termini of the activation loop but the corresponding region of CDK1 is disordered. When it comes to the complexes, CDK1 when bound to cyclin B undergoes significant changes in this region and its activation segment is better ordered but it still cannot be traced completely suggesting that it remains disordered around Thr161. CDK1-cyclin B is a better substrate for CDK7-cyclin H than monomeric CDK1 (Larochelle et al., 2007). CDK2 in complex with cyclin A is more susceptible to phosphorylation as well (Merrick et al., 2008). This observation leads to the model that the conformational flexibility of the activation segment is a factor in the efficiency of its phosphorylation by CAK.

The structures of CDK1 that have been solved are not phosphorylated on the activation segment. The superposition of the activation segment of CDK2 phosphorylated on T160 and unphosphorylated CDK1 also revealed some differences (Figure 2-26). The Cα atoms of the threonine phosphoacceptor residues are located 5 Å apart which suggests that for engaging CDK7-cyclin H for phosphorylation effectively they have to adjust their positions.
Figure 2-26 Overlay of the CDK1 and pT160CDK2 activation segments. CDK1 is represented in ice blue and pCDK2 in lilac. The CDK1 activation segment is dark blue and the CDK2 activation segment is red (PDB entry:1YC3 and 1QMZ respectively). Figure prepared using CCP4mg (McNicholas et al., 2011).

Another interesting feature of CDK1-cyclin B that is not present in CDK2-cyclin A but is reported in CDK4-cyclin D and CDK9-cyclin T structures (Baumli et al., 2008) is a cleft between the C terminal lobes of CDK1 and cyclin B. A similar cleft in CDK9-cyclin T has been shown to play role in binding the HIV regulatory protein TAT (Tahirov et al., 2010). The role (if any) in CDK1 is still to be discovered, but it does suggest a possible novel protein binding site that can be investigated in the future.

To summarise all the data presented, the structures of CDK1-CKS2 and CDK1-cyclin –CKS2 have revealed distinct features that can be connected to its function in a cell and explain the mechanism of its activation. The flexibility of the CDK1 activation segment and its general plasticity may be characteristics that are required for it to carry out its unique functions in the cell.
Chapter 3: Functional characterisation of CDK1

3.1 Introduction

3.1.1 Observations from the cellular studies

The main function of CDKs is phosphorylation of their substrates to orchestrate cell cycle progression. Some CDKs can phosphorylate many substrates on few or many phosphorylation sites (for example CDK1), whereas others are very specific, such as CDK4/6 and their substrate specificity is very limited. Regulation of CDK activity towards their substrates has been investigated by different groups. One hypothesis proposed based on investigation in fission yeast is that of a threshold or quantities model (Stem and Nurse, 1996) (Coudreuse and Nurse, 2010). The regulation of substrate specificity based on this model can depend on the affinity of binding sites and on their availability (Schulman et al., 1998). For example, where higher affinity sites when phosphorylated do not show any effects and only ones phosphorylated later do have a biological effect (Kim and Ferrell, 2007). Experiments performed in budding yeast which contain only one CDK that is activated by different cyclins has led to another model that emphasizes the role of the cyclin partner in substrate choice. A number of studies have shown that cyclins in budding yeasts are mostly interchangeable. For example, overexpression of Clb5 rescues cells lacking Clns but cells in the absence of Clb5 are also viable (Epstein and Cross, 1992). A more recent quantitative study has shown that CDK1-Clb5 phosphorylates 24% of chosen substrates more efficiently then CDK1-Clb2 (Loog and Morgan, 2005). The mechanisms controlling the selection of specific multiple phosphorylation sites by CDK1-cyclin-CKS1 have been described (Koivomagi et al., 2013). The importance of (i) distances between the phosphorylation sites, (ii) distribution of phosphoacceptors along the substrate, and (iii) positioning of cyclin docking motifs were all analysed. Different phosphorylation sites on the same protein can be targeted by different kinases and result in different outcomes. For example, cyclin E contains three major sites for phosphorylation: Ser372, Tyr380 and Ser384 (Welcker et al., 2003). All three can be phosphorylated by CDK2 in vitro but in vivo Tyr380 phosphorylation has been shown not to be CDK2 dependent (Welcker et al., 2003). It has been shown that for cyclin E turnover by Fbw7 it has to be phosphorylated on Ser384 and by GSK3 on Tyr380.
Animal cells have evolved to have a higher number of CDKs than the yeasts *S. cerevisiae* and *S. pombe*, but also to have a number of cyclins with more strictly defined functions. With CDK1 being the only essential CDK (Santamaria *et al.*, 2007) its partners cyclin A and B are important cyclins. cyclin A is required for DNA replication and cyclin B for mitosis (Brandeis *et al.*, 1998; Santamaria *et al.*, 2007; Kalaszczyńska *et al.*, 2009). Cyclin E and cyclin D are not essential for the cell cycle (Geng *et al.*, 2003; Kozar *et al.*, 2004), however cyclin E is required for chromosome endoreduplication and for cell cycle re-entry from quiescence (Geng *et al.*, 2003).

### 3.1.2 Cyclins as CDK1 and CDK2 partners

CDK1 is partnered by cyclins A and B, and CDK2 is partnered by cyclins A and E. In G2 phase both cyclin A2 and cyclin B1 bind CDK1 as both CDK1-cyclin A2 and CDK1-cyclin B1 complexes are required for induction of mitosis. Cyclin B is generally considered as an essential CDK1 activator for mitotic entry although studies focusing on the role of cyclin A have demonstrated its importance for nuclear envelope breakdown (D Gong, 2007). At the start of mitosis, cyclin A is degraded and CDK1 activity is maintained by cyclin B. The cyclins that promote mitotic transition are cyclin A2 and cyclin B1 and they cannot be substituted by other cyclins in mammalian cells (Satyanarayana and Kaldis, 2009). Type B cyclins are degraded in late mitosis by the anaphase-promoting complex also known as the cyclosome (APC/C), a core component of the ubiquitin-dependent proteolytic machinery that controls timely destruction of securins and cyclins (Hershko and Ciechanover, 1998; Kramer *et al.*, 2000; Peters, 2002). Degradation of cyclin B leads to inactivation of CDK1 which allows chromosome separation, cytokinesis and completion of mitosis.

In mammalian cells two cyclin A paralogues exist, cyclin A2 is normally expressed in most cell types, and cyclin A1 is mainly associated with testis (Sweeney *et al.*, 1996). Within the testis where both cyclins are present, they exhibit very different patterns of expression (Ravnik and Wolgemuth, 1999). Cyclin A1 in comparison to cyclin A2 has a distinctive influence on the biochemical properties of CDK1 and CDK2 and on regulation of their activity (Joshi *et al.*, 2009). It also has been demonstrated that the activities of CDK1 and CDK2 bound to both cyclins can vary depending on the nature of the substrate (Almon *et al.*, 1993). p53 is expressed in pachytene spermatocytes and consequently was used as a substrate to compare the kinase activities of CDK1 and CDK2 in complexes with cyclin A1 and A2. This study showed that *in vitro* CDK1-cyclin A1 phosphorylates p53 more efficiently that CDK1-cyclin A2 suggesting that the
cyclin can influence the kinase activity towards the substrate, and providing an explanation for cyclin A1 expression mainly in testis (Joshi et al., 2009).

Figure 3-1 Alignment of human cyclin A1 and human cyclin A2 sequences.

Human cyclin A1 (Uniprot entry P78396) and human cyclin A2 (Uniprot entry P20248) sequence alignment. Residues in α helices from the structure of cyclin A2 (PDB: 1VIN) are marked red with the helices labelled. MRAIL residues are underlined and the crystalized fragment (171-433) is identified by arrows («). Identical residues are marked with a star underneath the sequence and conserved residues are highlighted with a colon (:) or dot (.). Highly homologous CBF1 is marked with a red box and less homologous CBF2 is marked with a blue box. Sequences aligned using Clustal Omega (Sievers et al., 2011). Overall, human cyclins A1 and A2 share 43% identity.

Despite playing different roles, the sequences of human cyclin A1 and cyclin A2 are relatively conserved, sharing 208 identical amino acids (43%) as well as 114 similar ones (Sievers et al., 2011). The sequences have 88% identity in a region of CBF1 (Figure 3-1). However, there are several non-conserved amino acids in the area of CDK binding that have been proposed to modulate the CDK-cyclin interaction (Figure 3-1).
Cyclin B, in contrast to cyclin A, only forms a complex with CDK1. Two types of cyclin B were identified in mammals, B1 (Chapman and Wolgemuth, 1992) (Pines and Hunter, 1989) and B2 (Chapman and Wolgemuth, 1993) and a third more distant cyclin B3 in chicken, nematodes, frogs and flies (Kreutzer et al., 1995). Human cyclin B1 and cyclin B2 compared by the program Clustal Omega (Sievers et al., 2011) show identity at 213 residues (49%) and share similar residues at a further 106 aligned positions (Figure 3-2). Although there is not much similarity in the first ~100 residues, the following ~300 that encode the tandem CBF1 and CBF2 folds are well conserved.

Cyclin B3 does not align as well to cyclin B1 and cyclin B2 (8.6% and 8.5% of identical positions and 37% and 33% similarity between cyclin B3 and cyclins B1 and B2 respectively). The presence of cyclin B1 and cyclin B2 in frogs suggests that these genes have diverged at an early stage of evolution. In the majority of dividing cells the cyclins are co-expressed but at different locations: cyclin B1 is normally associated with microtubules and cyclin B2 with intracellular membranes (Tombes et al., 1991; Ookata et al., 1993; Jackman et al., 1995). The fact that cyclin B1 enters the nucleus in late G2 phase but cyclin B2 does not, as well as their different expression patterns during murine spermatogenesis (Chapman and Wolgemuth, 1993) and in Xenopus oocytes (Kobayashi et al., 1991) suggests that these two variants play different roles in preparing cells to enter mitosis.

<table>
<thead>
<tr>
<th></th>
<th>B1</th>
<th>B2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MALRVTNRKNAENKAKIMAGAKRVPTAPAAT-SKGPLRPTALGDIGNKVESEQLOAK</td>
<td>MAL-——LRPVSSDL-ENI-------DTGVNSKSVHTTIRTVLEEIGNEVRTRAAQV</td>
</tr>
<tr>
<td></td>
<td>MPMKKEAKPSTAKVIDKLPKPLEKVMPMLVPVPSEPVEPEPEPEPEPEPVMEKLPK-</td>
<td>AK-------------KAQNTKVPVQ-------------PTKTNVKLPTASVKPVMEKLPK-</td>
</tr>
<tr>
<td></td>
<td>ILVDTASPSPMETSAGPAEDLCQAFSDVIL- AVNDVDAEDGAPNLCSEYVKDIYAYL</td>
<td>----GPSPTP-----EDSVKKEENLCQAFSDALLCKIEDIDNEDWNPQLCDYVKDIYQL</td>
</tr>
<tr>
<td></td>
<td>αN-term</td>
<td>α1</td>
</tr>
<tr>
<td></td>
<td>RQLEEEQAVRPKYLGGRTVMNRAIILDVWLVQVMKFLRLLQETMYTVSDIFROMQC</td>
<td>RQLEVQVSHPHLDGDRDINGMRAIILDVWLVQVHSKFRLLQETLMCVDQIMDFRLOQVP</td>
</tr>
<tr>
<td></td>
<td>VPKKMLQLVGYTAMFIASKYEMYPPEIGDFAFVTDNTYTKHQRMEMKILRALNFGLG</td>
<td>VSRKLQLVGATALLASKYKEMFPNIEDFYITDNYTSSQIREMTILKELKFLG</td>
</tr>
<tr>
<td></td>
<td>RPLHLFRRASKIGEVDFVEQHTLAKYMELTMDMVHPPSQIAAAGFCLALKILDN</td>
<td>RPLHLFRRASKAGEVDFVEQHTLAKYLMELTIDYDMVHPKAAASCLSLQKVLQG</td>
</tr>
<tr>
<td>358</td>
<td>******************************************</td>
<td>322  ******************************************</td>
</tr>
</tbody>
</table>

```
Figure 3-2 Alignment of human cyclin B1 and human cyclin B2 sequences. 

Human cyclin B1 (Uniprot entry P14635) and human cyclin B2 (Uniprot entry O95067) sequences aligned. Residues in α helices from the structure of cyclin B1 (PDB code: 2B9R) are marked red with the helix name above. The MRAIL motif is underlined and the crystalized fragment (residues 165-433) is identified by the » at the beginning and « at the end. Identical residues are marked with a star underneath the sequence and conserved are highlighted with a colon (:) or dot (.). Highly homologous CBF1 is marked with a red box and less homologous CBF2 is marked with a blue box. Sequences aligned using Clustal Omega (Sievers et al., 2011). Overall, human cyclins B1 and B2 share 49% identity.

Cyclin E also plays a crucial role in CDK2 regulation. Cyclin E is known as a regulator of CDK2 but is also reported to be associated with other CDKs (Moroy and Geisen, 2004). Its functions in cells are both CDK dependent and independent. In complex with CDK2 it is necessary for the formation of pre-replication complexes on DNA as cells re-enter the cell cycle from quiescence (Geng et al., 2003). It also interacts with centromeres and stimulates their duplication (Moroy and Geisen, 2004), and plays roles in histone synthesis (Ma et al., 2000) and cell cycle progression (Hoffmann et al., 1994). Cyclin E has a 20 amino acid sequence which plays a role in centosomal localisation and promotes DNA synthesis in a CDK-independent manner (Matsumoto and Maller, 2004; Geng et al., 2008). Cyclin E has two paralogues in mammalian cells, cyclin E1 and cyclin E2. Cyclin E1 was first described in 1991 (Koff et al., 1991) and cyclin E2 in 1998 (Lauper et al., 1998). Alignment by Clustal Omega identifies 46.5% identity between the two paralogues and 25% similarity (Figure 3-3).
Figure 3-3 Alignment of human cyclin E1 and human cyclin E2 sequences.

Human cyclin E1 (Uniprot entry 24864) and human cyclin E2 (Uniprot entry O96020) sequences are aligned. Residues in α helices from the structure of cyclin E1 (PDB: 1W98) are marked red with helix name above the sequence. The MRAIL motif is underlined and the crystalized fragment (residues 88-357) are marked » at the beginning and « at the end. Identical residues are marked with a star underneath the sequence and conserved residues are highlighted with a colon (:) or dot (.). Highly homologous CBF1 is marked with a red box and less homologous CBF2 is marked with a blue box. Sequences were aligned using Clustal Omega (Sievers et al., 2011). Overall, human cyclins E1 and E2 share 46.5% identity.

Cyclin E is critically required for development of Drosophila melanogaster (Knoblich et al., 1994) and Caenorhabditis elegans (Fay and Han, 2000) and essential for mammalian embryogenesis (Geng et al., 2003). Double knockout mice E1−/− appeared normal through their life and could not be distinguished from wild type. Mice lacking E2 were initially indistinguishable from control littermates although males displayed lower fertility with around 50% sterility. Female littermates did not show any differences. The lethality of double cyclin E knockout (E1−/−;E2−/−) is not due to embryo abnormalities, but rather to placental dysfunction developed as a result of a defect in trophoblast development.

3.1.3 Sequential comparison of cyclins binding to CDK1 and CDK2

The determination of the structure of CDK1-cyclin B-CKS2 provided an opportunity to compare the residues at the CDK-cyclin interfaces in various CDK1 and CDK2-cyclin complexes to attempt to understand the observed pairing specificities. The residues within the cyclin A structure that are within 3.6 Å of a cyclin B residue contacting CDK1 were identified and compared (Figure 3-4). Of the 18 residues identified, 10 are identical in cyclin A and only 5 are identical with the equivalent amino acids of cyclin E. It was noted that cyclin B residues Tyr170, Tyr177 and Tyr258 are conserved in cyclin A and have large side chains that pack against the CDK1 surface, in contrast to the small amino acids Asn112, Ile119 and Leu202 present in the cyclin E sequence. The
ability of CDK2 to form a larger interface with its cognate cyclins may lead to the smaller side chains present in cyclin E being better tolerated. These observations offer an explanation as to why CDK1 prefers to pair with cyclin B and cyclin A rather than with cyclin E.

![Figure 3-4 Analysis of cyclin residues that participate in CDK binding.](image)

Amino acids that mediate the CDK1-cyclin B interface were identified using the CCP4 program CONTACT (Winn et al., 2011). Identical amino acids present in cyclin A and cyclin E were identified by superimposing the CDK2-cyclin A (PDB code: 1QMZ) and CDK2-cyclin E (PDB code: 1W98) structures in CCP4MG.

### 3.1.4 CDK substrates

A proteomic study focusing on analysis of substrates of CDK in complex with cyclin A2, B1 and E1 has been performed to understand the dynamics of the cell cycle. In this study CDK-cyclin complexes were purified from different stages of the cell cycle and analysed by quantitative mass spectrometry using SILAC (stable isotope labelling by amino acid in cell culture) (Ong et al., 2002) and bound substrates were recovered by immunoprecipitation (Archambault et al., 2004). A number of substrates were identified for complexes of CDK1 with cyclins A2, B1 and E1 that are unique for each of these cyclins, but also overlap between cyclins E1 and B1 to some extent, between A2 and E1 more and a large number of substrates are phosphorylated by cyclin A2 and cyclin B1 (Figure 3-5). The overlapping substrates were demonstrated to be bound by different cyclins at different times in the cell cycle even if another potential cyclin that can bind it is present in the same cell cycle phase (Pagliuca et al., 2011) (Figure 3-5). These results support a model in which cyclins remodel the cognate CDK for phosphorylation of specific substrates, and where there is coordination between cyclins in targeting the same set of substrates but at a different phases of the cell cycle.

Most CDK substrates contain S/T*-P-X-K/R sequence (where X is any amino acid) (Ubersax et al., 2003) with a different requirements around the site of phospho transfer.
Figure 3-5 Partners of cyclin A1, B2 and E1 in interphase.

(A) Representation of substrates that overlap between cyclins A1, B2 and E1. (B) Substrates that bind cyclin A in S phase represented in red and blue in G2, substrates that bind cyclin B in mitosis represented green (Pagliuca et al., 2011).
3.1.5 Substrate recruitment

The canonical sequence which CDKs phosphorylate is S/TPXK/R, where S is the phosphorylatable serine and X is any amino acid. The RXL motif is present in a number of CDK substrates including Rb, E2F-1, Cdh1, Cdc6, and the inhibitor proteins p21CIP1,p27KIP1 and p57KIP2 (Adams et al., 1999) (Saha et al., 1997; Takeda et al., 2001; Wohlschlegel et al., 2001; Willder et al., 2013). The RXL motif binds to a hydrophobic patch on the surface of cyclins A, B, D and E called the recruitment site (Brown et al., 1999; Petri et al., 2007). The presence of an RXL motif is important but arginine can be substituted for lysine with a small loss of binding affinity (Lowe et al., 2002). One of the proposed roles of the RXL motif is in localising the substrate to the CDK in a non-specific manner increasing the substrate concentration (Schulman et al., 1998). The addition of it has been shown to increase phosphorylation of otherwise poorly phosphorylated substrates (Adams et al., 1999). The linker between the phosphorylatable residue and the RXL motif has been shown to be flexible and doesn’t have a preferred path on the cyclin surface (Takeda et al., 2001).

The dependence on the presence of certain residues in the canonical phosphorylation site and an RXL motif in different CDK-cyclin complexes for efficient substrate phosphorylation has been investigated (Brown et al., 1999). CDK4-cyclin D which does not have a strict preference for position P+3 shows a strong dependence on the R/KXL motif, in the absence of which CDK4-cyclin D loses activity completely (Takaki et al., 2009). The special feature of CDK4–cyclin D that it may require substrate binding to adopt an active conformation might explain the importance of both recruitment sites. An investigation of other CDK-containing complexes would be beneficial.

3.1.6 Aims and objectives

The structure of CDK1-cyclin B revealed different features at the CDK-cyclin interface in comparison to the other CDK–cyclin pairs. As the interface between CDK1 and cyclin B is smaller than that between CDK2 and cyclin A, it is of interest to determine the affinity and nature of this interaction. The structure of CDK1-cyclin B revealed that the non-phosphorylated activation segment is more flexible and less well-ordered than that of CDK2-cyclin A. Unlike CDK2-cyclin A, CDK4 bound to cyclin D also features a disordered activation segment (Day et al., 2009) and it was of interest to determine whether a disordered activation segment is also a feature of the phosphorylated CDK1-
cyclin B complex. Such a structure could influence both the stability of the complex (Brown et al., 1999) and its substrate specificity.

3.2 Description of biophysical methods

3.2.1 Isothermal titration calorimetry to measure affinity between proteins

Isothermal titration calorimetry is a biophysical technique that can be used to measure protein-ligand, protein–nucleic acid and protein – protein interactions. One component is injected at higher concentration in small aliquots into a solution of its interacting partner at lower concentration and the energy associated with the interaction is measured. Measurement of heat transfer allows accurate determination of change of enthalpy (ΔH) which derives from the ITC isotherm. It can be used to determine other thermodynamic parameters such as the association constant (K_a), changes in entropy (ΔS) and stoichiometry of the interactions (n) in solution using Origin 7.0 software (OriginLab Corp.)

Change of enthalpy can be used for the further determination of thermodynamic parameters by using the Wiseman isotherm (stoichiometry 1:1 is specified for simplicity):

\[
\frac{dQ}{d[X]t} = \Delta H^0 V_0 \left[ \frac{1}{2} \frac{1 - X_R - r}{2\sqrt{(1 + X_R + r)^2 - 4X_R}} \right]
\]

where

\[
\frac{1}{r} = c = K_a [M]_t = \frac{[M]_t}{K_d}
\]

Equation 3-1 Wiseman isotherm for ligand – receptor binding at 1:1 stoichiometry.

Abbreviations: X – ligand, M – receptor, V_0 – volume of the calorimeter cell, [M]_t – total receptor concentrator, \( \frac{dQ}{d[X]t} \) – change of heat/moles ligand added, ΔH – change in enthalpy, X_R=[X]/[M]_t ratio of ligand to receptor concentration, K_a – association constant.
The Gibbs-Helmholtz equation is regularly used to determine the entropic contribution to binding.

\[ \Delta G = -RT \ln K_a = \Delta H - T \Delta S \]

**Equation 3-2 Gibbs-Helmholtz equation.**

Abbreviations: G – change in Gibbs energy, R – ideal gas constant, T – absolute temperature, \( \Delta H \) – changes in enthalpy, \( K_a \) – association constant, \( \Delta S \) – changes in entropy.

### 3.3 Materials and Methods

#### 3.3.1 Protein production

CDK1 and CAK1 each with an N terminal GST tag, cleavable by 3C protease for GSTCDK1 and non-cleavable for GSTCAK1, were expressed in SF9 insect cells as described in Section 2.2.4. Phosphorylated CDK2 is generated by CDK2 co-expression with CAK which phosphorylates CDK2 in bacterial culture as described in Section 2.2.7 using the construct described in Table 2-1. Constructs pGEX6P1CDK2CAK1, pET21dCyclinAHis, pET28CyclinBcs described in section 2.2.3 were used to express phosphorylated CDK2, His-tagged cyclin A and cyclin B in Rosetta2™(DE3)pLysS E. coli cells as outlined in Section 2.2.3. All proteins were purified by standard protocols outlined in Sections 2.2.8 and 2.2.9.

#### 3.3.2 Peptide synthesis, expression and purification by ion exchange chromatography

To measure kinase activity a panel of eight peptides derived from the previously described sequence of p107 (Leng et al., 2002) were prepared. The peptides differ systematically at the P+1 and P+3 positions around the site of phosphotransfer (where the phosphorylated residue is denoted P) and at the cyclin binding RXL motif (Table 3-1). All the peptide sequences also contain the following mutations: (i) Pro651 to alanine so that the peptides contain only one SP motif and (ii) Tyr648 and Ile674 to tryptophans for more accurate determination of protein concentration at 280 nm. DNA sequences were gene synthesised by GenScript and subsequently inserted into pGEX6-P1 by In-Fusion cloning (Clontech, USA) using BamH1 and EcoR1 restriction sites. Plasmid DNA was transformed into Stella E. coli cells, following the manufacturer’s protocol.
then plated on to LB plates supplemented with ampicillin and incubated overnight at 37°C. Plasmid DNA was purified using a QIAprep Spin Miniprep kit (Qiagen, Hilden, Germany) from a single colony and verified by sequencing. The peptides were expressed in the Rosetta™(DE3)pLysS E. coli strain at 20°C overnight after induction by 0.1 mM IPTG. Pelleted by centrifugation at 5000 x g cells were re-suspended in 40 mM Hepes, pH 7.5, 200 mM NaCl, 0.01% MTG, protease inhibitors Complete tablet and flash frozen in dry ice and stored at -80°C. Cell pellets were thawed under running cold water and then kept on ice and sonicated in the presence of 0.01 mg/ml DNase, 0.05 mg/ml RNase, 0.25 mg/ml Lysozyme, 5 mM MgCl₂, for 15 cycles totalling 3 min with 20 sec intervals. Cells were centrifuged at 60,000xg, the supernatants collected and kept on ice. Proteins were purified by sequential affinity chromatography exploiting the GST tag followed by ion exchange chromatography. The choice of ion-exchange resin (HiTrap Q FF or HiTrap SP FF) was made with reference to the calculated pI of each peptide (Table 3-2). Peptides were buffer exchanged into 25 mM MES, pH 6.5 prior to IEC and eluted by a one step increase in the NaCl concentration to 1M. Peak fractions were collected and confirmed by SDS-PAGE, pooled and flash cooled in liquid nitrogen in 100 μl aliquots and stored at -80°C.

<table>
<thead>
<tr>
<th>Peptide</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPIS+RXL</td>
<td>GPLGSMHPRVKEVRTDSGLRDMQPLSPISVHERWSSATAGSAKRRLFGEDPPKEMLMDKW</td>
</tr>
<tr>
<td>SPIK+RXL</td>
<td>GPLGSMHPRVKEVRTDSGLRDMQPLSPIKVHERWSSATAGSAKRRLFGEDPPKEMLMDKW</td>
</tr>
<tr>
<td>SPIS-RXL</td>
<td>GPLGSMHPRVKEVRTDSGLRDMQPLSPISVHERWSSATAGSAAAAAFGEDPPKEMLMDKW</td>
</tr>
<tr>
<td>SPIK-RXL</td>
<td>GPLGSMHPRVKEVRTDSGLRDMQPLSPIKVHERWSSATAGSAAAAAFGEDPPKEMLMDKW</td>
</tr>
<tr>
<td>SAIS+RXL</td>
<td>GPLGSMHPRVKEVRTDSGLRDMQPLSAISVHERWSSATAGSAKRRLFGEDPPKEMLMDKW</td>
</tr>
<tr>
<td>SAIK+RXL</td>
<td>GPLGSMHPRVKEVRTDSGLRDMQPLSAIKVHERWSSATAGSAAAAAFGEDPPKEMLMDKW</td>
</tr>
<tr>
<td>SPIS-RXL</td>
<td>GPLGSMHPRVKEVRTDSGLRDMQPLSPISVHERWSSATAGSAAAAAFGEDPPKEMLMDKW</td>
</tr>
<tr>
<td>SAIK-RXL</td>
<td>GPLGSMHPRVKEVRTDSGLRDMQPLSAIKVHERWSSATAGSAAAAAFGEDPPKEMLMDKW</td>
</tr>
</tbody>
</table>

**Table 3-1 Sequences of p107 mutant peptides.**

The sequence of p107 (Uniprot entry P28749), residues 618-674 was used as the starting point for the peptide set. The sequences contain systematic mutations at the SPXK consensus sequence around the site of phospho-transfer and at the RXL motif that binds to the cyclin recruitment site (highlighted in bold and underlined). The GPLGS sequence at the start of each peptide is a cloning artefact that remains following cleavage of the N-terminal GST tag by 3C protease. The mutations Tyr648Trp and Ile674Trp (highlighted in blue) were introduced to aid the determination of peptide concentrations and Pro651Ala removes the second authentic SP motif in the p107 sequence. Construct SPIS+RXL is the authentic p107 sequence at these motifs.
<table>
<thead>
<tr>
<th>No.</th>
<th>Peptides</th>
<th>Molecular Weight</th>
<th>pI (GST fusion)</th>
<th>pI (peptide)</th>
<th>pI of GST</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>SPIS+RXL</td>
<td>7847</td>
<td>6.47</td>
<td>9.9</td>
<td>5.73</td>
</tr>
<tr>
<td>21</td>
<td>SPIK+RXL</td>
<td>7536.5</td>
<td>6.72</td>
<td>10.1</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>SPIS-RXL</td>
<td>7536.5</td>
<td>5.95</td>
<td>6.9</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>SPIK-RXL</td>
<td>7577.6</td>
<td>6.1</td>
<td>8.44</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>SAIS+RXL</td>
<td>7779.9</td>
<td>6.47</td>
<td>9.97</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>SAIK+RXL</td>
<td>7821</td>
<td>6.72</td>
<td>10.1</td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>SPIS-RXL</td>
<td>7510.5</td>
<td>5.95</td>
<td>6.94</td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>SAIK-RXL</td>
<td>7551.6</td>
<td>6.1</td>
<td>8.44</td>
<td></td>
</tr>
</tbody>
</table>

Table 3-2 Molecular weights and pIs of p107 mutants.
Molecular weights and isoelectric points of peptides on their own and GST bound calculated by ExPASy Bioinformatics Resource portal (Artimo et al., 2012).

### 3.3.3 SDS-PAGE

SDS-PAGE was performed as described in Section 2.2.10.

### 3.3.4 Phosphorylation of CDK1

GSTCAK1 (S. cerevisiae) expressed in insect cells was used to phosphorylate CDK1 in vitro. Purified CDK1 and GSTCAK1 were mixed in a ratio of 1:4 in binding/SEC Buffer 1 (recipe is described in chapter 2.2.2) in the presence of 1 mM ATP and 10 mM MgCl$_2$ and incubated overnight at 4°C. Phosphorylated CDK1 (referred as pCDK1 from this point) was separated from GSTCAK1 by SEC using a Superdex 200 16/60 column. The extent of phosphorylation was determined by HPLC.

### 3.3.5 Reverse phase HPLC

0.1ml of CDK1 (at 0.07mg/ml in binding/SEC buffer 1) was supplemented to 0.01% TFA and injected onto a 5 um C4 300A column (Phenomenex, Torrance, California, US) equilibrated into 0.01% TFA in water (Buffer A). Proteins were separated by applying a gradient of 50% to 60% acetonitrile /0.01% TFA in water over 20 column volumes.
3.3.6 Preparation of the complexes

For preparation of phosphorylated CDK1-cyclin B complex (referred as pCDK1-cyclin B from this point), pCDK1 prepared as described in 3.2.4 was first separated from GSTCAK1 by SEC, and then mixed with cyclin B in a ratio of 1:1.5, CDK1:cyclin B. The NaCl concentration was adjusted to 1 M to stabilise the pCDK1-cyclin B complex, and then the complex was concentrated using a Sartorius Vivaspin concentrator 10kDa cut-off before further separation of pCDK1-cyclin B from cyclin B on a Superdex 75 26/60 column. The peak corresponding to the pCDK1-cyclinB complex was collected and analysed by SDS-PAGE. Fractions of pCDK1-cyclin B were concentrated to 1 mg/ml using 10 kDa cut-off concentrators (Vivaspin) and frozen in 100 µl aliquots.

Phosphorylated CDK2-cyclin A (referred as pCDK2-cyclin A from this point) was prepared essentially as described (Brown et al., 1999) with minor modifications. Briefly, a cell lysate of GSTpCDK2 phosphorylated in bacterial cells by co-expression with CAK1 was mixed with a cell lysate of cyclin A in a ratio of 1:2 GSTpCDK2:cyclin A to ensure binding of all pCDK2 to cyclin A. The complex was loaded onto a Glutathione Sepharose 4B column equilibrated in modified HBS (40 mM HEPES, 200 mM NaCl, 1 mM DTT, pH 7.0, 0.01% MTG) and washed to baseline. The GSTpCDK2-cyclin A was eluted by 20 mM glutathione in mHBS and cleaved overnight with 3C protease in a ratio of 1:50 (w/w). The cleaved sample was then loaded onto a Superdex 75 26_60 size-exclusion column equilibrated in mHBS to remove aggregates and glutathione. Fractions containing pCDK2-cyclin A (identified by SDS-PAGE) were pooled and re-applied to a Glutathione Sepharose 4B column to remove the GST dimer. The complex was concentrated using a Sartorius Vivaspin concentrator 10kDa cut-off to 12-15 mg/ml and fast cooled in 25 µl aliquots in liquid nitrogen and stored at -80°C.

3.3.7 Isothermal Titration Calorimetry (ITC)

ITC experiments were carried out at 30°C using a MicroCal ITC 200 (Malvern Instruments). Cyclin B1 (165-433) and cyclin A2 (169-430) were used as samples in the cell and unphosphorylated CDK1 and CDK2 were used as ligands. Titrations were performed in two buffer conditions; 50 mM HEPES pH 8.0, 500 mM NaCl, 0.01% MTG (high salt), and 40 mM HEPES, 200 mM NaCl, 0.01% MTG, pH7.4 (low salt). For ITC experiments in the higher salt conditions cyclin B and cyclin A were both in the cell at 140 µM, and CDK1 and CDK2 were added from the syringe at 16 and 18 µM,
respectively. In the lower salt and lower pH buffer, cyclin B and cyclin A were in the cell at 60 and 200 μM, and CDK1 and CDK2 were added from the syringe at 18 and 30 μM, respectively. The thermograms were analysed by Origin 7.0 (OriginLab Corp.) software.

### 3.3.8 ADP-Glo™ Kinase assay

Kinase reactions (ADP-Glo assay format (Promega) were carried out in 5 µl total volume at room temperature using 4 nM pCDK1-cyclin B or 1.5 nM CDK2-cyclin A in 40 mM Tris-HCl pH 7.5, 20 mM MgCl₂, 0.1 mg ml⁻¹ BSA buffer. Ten points of two times dilutions (94 to 0 μM) of the panel of peptides (Table 3-1) were added to the CDK-cyclin solution and the reaction was initiated by the addition of 75 μM ATP. Assays were performed in white 384-well plates in duplicate. Luminescence was detected using a PheraStar plate reader (BMG). Kₘ and Vₘₙₐₓ values were derived using PRISM (GraphPad) and were calculated using Michaelis-Menten equation:

\[ \gamma = \frac{V_{max}}{1+(K_m[S])} \]

**Equation 3-3 Michaelis-Menten equation.**

Abbreviations: \( \gamma \) is a reaction rate or speed of reaction, \([S]\) is concentration of a substrate, \(K_m\) is a Michaelis constant which is the substrate concentration at which a reaction rate is half of \(V_{max}\), \(V_{max}\) represents the rate achieved by the system at saturating substrate concentration.

### 3.3.9 CDK de-phosphorylation

CDK-cyclin complexes were dephosphorylated using λ phosphatase. Phosphorylated CDK1-cyclin B (4 nM) and CDK2-cyclin A (2.5 nM) complexes were incubated for 30 min at 30 °C in total volume of 60 µl with 32 units of lambda phosphatase in 50 mM Tris, pH 8.0, 500 mM NaCl, 0.01% MTG (CDK1-cyclin B) and 50 mM Tris, pH 7.4, 200 mM NaCl, 0.01% MTG (CDK2-cyclin A) buffers supplemented with 400 μM MnCl₂. Mock-treated samples were incubated with buffer only. The extent of dephosphorylation was assessed by subsequently measuring the residual kinase activity by ADP-Glo™ assay.

### 3.3.10 Differential Scanning Fluorimetry (DSF)

Experiments were performed using a ViiA7 Real-Time PCR system (Applied Biosystems). 4 μM pCDK1-cyclin B and 6 μM pCDK2-cyclin A were mixed with SyPRO Orange (x10) (Invitrogen) fluorescent probe in DMSO (supplied by Sigma Aldrich). The
plates were first equilibrated at 25°C for 5 min in the PCR machine before starting the thermal melting experiment. The plates were heated at 1°C per 90 sec from 25 to 85°C. The fluorescence intensity was recorded using the ROX filters (excitation λ=470 nm, emission λ = 586nm) every 90 sec. Results were plotted in PRISM (GraphPad) and inflection point ™ was calculated using Boltzmann equation.

3.4 Results and Discussion

3.4.1 Determination of CDK-cyclin binding

The principals of ITC have been described in Section 3.1.4. It is a method that measures the heat that is released or absorbed as a result of chemical or physical changes in a sample. This method has been employed to estimate the binding affinity between unphosphorylated CDK1 and cyclin B in comparison to CDK2 and cyclin A. The experiments have been done in collaboration with Dr. Will Stanley. The observation that the binding interface between CDK1-cyclin B and CDK2-cyclin A differs instigated the idea that this might result in differences in stability of the complexes in vitro. All four proteins were prepared separately and subsequently mixed to generate the complexes. The direct comparison was difficult as both complexes were stable under different buffer conditions. CDK2-cyclin A was stable at near physiological conditions (pH 7.4, 200 mM NaCl, T=30°C), yielding a $K_d$ value of 21 nM (Figure 3-6 A) but cyclin B was unstable at this salt concentration undergoing precipitation which made it hard to interpret the thermograms quantitatively (Figure 3-6 B). The experiment was repeated in the reverse direction where the $K_d$ of CDK1 binding to cyclin B was determined at the higher salt concentration and higher pH (pH 8.0, 500 mM NaCl, T=30°C). This experiment yielded a $K_d$ value of 28 nM, but these conditions were not favourable to study interactions between CDK2 and cyclin A. These results show that both kinases bind to their respective cyclins with nanomolar affinity but that binding can only be reproducibly measured for CDK1–cyclin B in the presence of high salt.
Figure 3-6 Isothermal titration calorimetry to assess the formation of CDK1-cyclin B and CDK2-cyclin A complexes in vitro.
Thermograms generated by ITC during formation of CDK1-cyclin B (B, C) and CDK2-cyclin A (A, D) complexes. Experiments were performed at the different buffer conditions, (A, B) at a lower salt and lower pH (40 mM Hepes, pH 7.4, 0.01 % MTG and (C,D) at a higher salt concentration and higher pH (40 mM Hepes, pH 8.0, 500 mM NaCl, 0.01 % MTG).
The technique employed to explore the stability of the complexes was a DSF assay. For this experiment pCDK1-cyclin B and pCDK2-cyclin A were prepared in buffer containing 200 mM NaCl. pCDK1 and cyclin B were purified separately as described earlier and the complex was assembled by mixing pure proteins in a ratio of 1 (CDK1):2 (cyclin B). The protein mix was applied to a Superdex 75 26_60 column and fractions from peak 1 were collected. Fractions were analysed by SDS-PAGE and then pooled together and concentrated to 1 mg/ml then frozen for use in the assay (Figures 3-7, 3-8).

Figure 3-7 Preparation of a pCDK1-cyclin B complex.
(A) Size-exclusion chromatography chromatogram of pCDK1-cyclin B. Peak of pCDK1-cyclin B labelled 1 and fractions identified with a red bar. Peak 2 contains excess of cyclin B. Absorbance at 280 nm is plotted against elution volume in ml. The lower panel shows the elution positions of a standard mix of protein markers to calibrate the column. (B) Sample of pooled fractions analysed by SDS-PAGE. (C) Size-exclusion chromatography chromatogram of standards.
Figure 3-8 Separation of pCDK2-Cyc A complex by SEC.

(A) Chromatogram of pCDK2-cyclin A complex. Peak of pCDK2-cyclin A identified with a red bar. Absorbance at 280 nm is plotted against elution volume in ml. (B) Fractions labelled with a red bar were pulled together and identified by the SDS-PAGE. (C) Size-exclusion chromatography chromatogram of standards.

Analysis by DSF assay showed that pCDK2-cyclin A complex is 11°C more stable than pCDK1-cyclin B (Figure 3-9 A). The melting temperature depends on the thermal stability of individual proteins as well as the energy of their binding. However, the results do support the hypothesis based on the structural comparison (Section 2.3.5) that the extended interface between CDK2 and cyclin A generates a complex that is more thermally stable than CDK1-cyclin B.
To find out if phosphorylation influences protein stability, pCDK1-cyclin B was treated with λ phosphatase or a buffer control and then re-tested in the DSF assay. This experiment revealed a small but reproducible difference of 1°C between the phosphorylated and non-phosphorylated complexes suggesting that the phosphorylated CDK1 complex is more stable (Figure 3-9 B).

![Comparison of thermal stability by Differential Scanning Fluorimetry.](image)

(A) Thermal stability comparison of phosphorylated CDK1-cyclin B and CDK2-cyclin A complexes. (B) Thermal stability comparison of phosphorylated CDK1-cyclin B complex and λ-phosphatase treated CDK1-cyclin B. Experiment performed in duplicate.

### 3.4.2 Investigation of CDK1 activity towards the panel of p107 derivatives

As the main function of CDKs is to phosphorylate their substrates, the activities of phosphorylated CDK1 and CDK2 in complexes with their cognitive cyclins towards a panel of peptides was compared. The panel is composed of seven derivatives from retinoblastoma protein p107 to explore the importance of amino acids around the site of phospho-transfer and in the RXL motif which plays a role in cyclin recognition. Seven mutants were created by replacing proline in position P+1 and the positively charged residue in position P+3 of a peptide (Table 3-1). The RXL motif was also eliminated in some of the variants. Peptides were expressed and purified as described in Materials and Methods. Mutants 20, 21, 23, 24, 25 (Table 3-1) were purified by HiTrap SP Sepharose FF column and mutants 22 and 26 by HiTrap Q Sepharose FF (Table 3-2). Peptides were eluted from the column by a step increase in the NaCl concentration.
from 0 to 500 mM to help concentrate the samples. Figure 3-10 shows the purification of peptide mutant 23 as an illustrative example.

**Figure 3-10 Separation of p107 mutant 23 from GST by HiTrap SP Sepharose FF column.**

(A) Chromatogram of p107 mutant 23. For ion-exchange chromatography 25 mM MES pH 6.5 buffer was applied to the column without NaCl and after 35 ml 1M NaCl in 25 mM MES pH 6.5 was introduced to the column to ensure quick protein release. Fractions from both peaks labelled with the blue bar and the red bar were collected and analysed by SDS-PAGE. The increase of NaCl concentration is shown as a green curve. (B) Peaks from ion exchange chromatography identify GST and p107 mutant 23 by the SDS-PAGE.

The ADP Glo™ assay (Promega) was employed to measure kinase activity. This assay is a luminescent ADP detection assay that measures the amount of ADP produced as a result of a kinase reaction (Figure 3-11). This assay is performed in two steps. After the kinase reaction is completed, the ADP-Glo reagent is added in equal amounts to all sample wells to deplete the remaining ATP. In the next step kinase detection reagent is added to transform all ADP into ATP which can produce luminescence as a result of a luciferase/luciferin reaction. Luminescence is measured by the luminometer. Luminescence is correlated to ADP concentration by using an ATP-to-ADP conversion curve which can be prepared for the different proportion of ADP in an ADP+ATP mixture. As the luminescence signal is generated according to the ADP concentration produced, it is correlated with kinase activity.
Figure 3-11 Illustration of the ADP-GloTM assay.

Generally, the results of the assay show that both CDK1 and CDK2 are active against the peptides with canonical phosphorylation site SPIK and intact RXL motif and demonstrate no activity against the peptides highly compromised in the phosphorylation site SAIS with or without an RXL motif. Also in the absence of the RXL motif both kinases phosphorylate peptides that have canonical phosphorylation sites SPIK (Figure 3-12).

Despite their similarities, the complexes do have differences in their activity towards the mutants (Table 3-3). pCDK2-cyclin A requires proline in position P+1 and a positively charged residue in position P+3 as well as the RXL motif for activity (SPIS+RXL). However, CDK1-cyclin B remains active even if only proline (SPIS+RXL) is present in position P+1 or a positively charged residue (SAIK+RXL) is present at position P+3. In the absence of the RXL motif, pCDK1 retains activity towards substrates with an impaired phosphorylation site (SPIS-RXL) but CDK2 –cyclin A is almost inactive against these substrates. Consequently pCDK1-Cyclin B is more accommodating towards variations in the substrate sequence around the site of phospho transfer in comparison to pCDK2-Cyclin A.
Figure 3-12 Kinase activity of pCDK1-cyclin B and pCDK2-cyclin A towards the panel of peptides.

Seven peptides were generated with introduced mutations around the site of phospho-transfer in positions P+1 and P+3 and with or without KRXL motif that binds to cyclin recruitment site to test the activity of pCDK1-cyclin B (A) and pCDK2-cyclin A (B). The side panel identifies amino acids in site of phosphor-transfer and the presence or absence of RXL motif. Graphs representing the activity of different peptides shown in different colours and calculated using GraphPad Prizm 6 software. Experiments were repeated twice and error bars correspond to the range of different values.

<table>
<thead>
<tr>
<th>Phosphorylated CDK1-cyclin B</th>
<th>Phosphorylated CDK2-cyclin A</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPIK</td>
<td>SAIK</td>
</tr>
<tr>
<td>+RXL</td>
<td>0.52</td>
</tr>
<tr>
<td>-RXL</td>
<td>0.45</td>
</tr>
</tbody>
</table>

Table 3-3 kcat/Km, peptide (s-1) for CDK-cyclin complexes on p107 mutants.
ND – not determined.
3.4.3 Role of activation segment phosphorylation in determining the observed substrate preferences

According to previous work (Takaki et al., 2009) the CDK2-cyclin A activation segment is buried in a cluster of positively charged residues and after treatment with λ phosphatase CDK2 remains as active as before treatment. However, active CDK4 after the same treatment loses its activity suggesting that the phosphorylated Thr172 is accessible to λ phosphatase and the activation segment is flexible. A structural comparison of CDK1-cyclin B and CDK2-cyclin A shows that the activation segment is more open and flexible in CDK1. To investigate this observation pCDK1-cyclin B was treated with λ phosphatase under the conditions where the CDK2-cyclin A activity was minimally affected and then the treated complex was tested in a kinase assay towards the peptide SPIK+RXL (Figure 3-13). The results reveal a significant decrease in pCDK1-cyclin B activity after de-phosphorylation which supports the hypothesis that the activation segment is locally flexible around Thr161 and therefore accessible to λ phosphatase treatment. The results of this experiment suggest that the relaxed substrate specificity of the activation segment of pCDK1-cyclin B is a consequence of its flexibility.

![Comparison of CDK1 and CDK2 kinase activity after λ phosphatase treatment.](image)

**Figure 3-13 Comparison of CDK1 and CDK2 kinase activity after λ phosphatase treatment.**
Phosphorylated CDK1-cyclin B and CDK2-cyclin A complexes were incubated with λ phosphatase or incubation buffer at 37°C for 1 h and kinase activity was determined towards the optimal peptide (SPIK+RXL) by the ADP-Glo assay. Results were plotted in GraphPad Prizm 6. Blue bar indicates kinase activity of pCDK2-cyclin A complex against the optimal peptide, red bar – pCDK1-cyclin B complex, green bar – pCDK2-cyclin A treated with λ phosphatase, purple bar – pCDK1-cyclin B treated with λ phosphatase. Error bars indicate the standard deviation of two independent measurements.
3.5 Conclusions

Experiments described in this Chapter were mainly stimulated by the structural findings and included in vitro assays for biochemical characterisation of CDK1 and CDK1-cyclin B complex in comparison with CDK2 and CDK2-cyclin A. CDK2 is a well studied kinase which despite its similarities in sequence with CDK1 differs substantially in function. Structural analysis of CDK1–cyclin B revealed a 30% smaller interface between CDK1 and cyclin B in comparison to CDK2-cyclin A and a smaller number of interactions involved in binding. The results of DSF confirmed the higher stability of pCDK2-cyclin A in comparison to pCDK1-cyclin B although it should be noted that monomeric CDK2 has a higher melting temperature then CDK1 by circa 5°C (DSF data from Chapter 5). The comparison of the residues involved in CDK-cyclin binding between cyclin A, cyclin B and cyclin E has offered a possible explanation as to why CDK1 binds preferably to cyclin B and cyclin A. It has been noticed that cyclin E has smaller residues Asn112, Ile119 and Lue202 in the positions where cyclin B and cyclin A have larger more hydrophobic tyrosine residues (Tyr170, Tyr177 and Tyr258 in cyclin B). In the context of the smaller CDK1 interface, these residues do not contribute as much binding energy with the result that a CDK1-cyclin E complex would not be stable. However, in the context of the larger CDK2-mediated interface, additional interactions help stabilise the CDK2-cyclin A complex.

The investigation of pCDK1-cyclin B substrate preferences and comparison with pCDK2-cyclin A yielded trends that again agreed with the structural findings. pCDK1-cyclin B is more relaxed in its substrate preferences in comparison to pCDK2-cyclin A which offers an explanation as to why CDK1 can phosphorylate a much wide range of substrates than CDK2-cyclin A.
Chapter 4: Diverse CDK regulators

4.1 Introduction

As has been discussed in Chapter 1 CDK activity is regulated directly by both protein binding and reversible phosphorylation, and indirectly by, for example, mechanisms that control its localisation within the cell and the level of expression and degradation of its regulators (e.g. via the ubiquitin proteasome pathway). As an example, a number of sequence diverse but structurally conserved protein families activate CDKs by a mechanism shared with members of the cyclin family, but their levels do not oscillate through the cell cycle. Some cyclin-like proteins are expressed in terminally differentiated cells that have exited the cell cycle such as p35 that binds to CDK5 in neurons, while viruses encode cyclin-like proteins that can hijack the host cell cycle machinery to promote an environment that supports viral replication. A number of non-cyclin CDK binding proteins have been discovered but have not been the subject of structural studies. In this chapter, preliminary work to characterise two CDK binding proteins, Speedy/RINGO and RGC32 and to establish conditions to study the role of sumoylation in controlling CDK activity is described.

4.1.1 Ringo/Speedy – a family of atypical CDK regulators

Ringo/Speedy proteins are a family of atypical regulators of CDK1 and CDK2 which have no sequence homology with the cyclins, but are predicted to contain a cyclin box fold. Speedy was first described as a gene in *Xenopus* that can rescue the UV and γ -radiation sensitivity of the *Rad1* mutant of *S. pombe* (Lenormand et al., 1999). RINGO (Rapid INducer of G2/M progression in Oocytes) was independently discovered in a screen of regulators of G2/M progression in *Xenopus* oocytes (Ferby et al., 1999). Both of these genes were reported to induce mitotic maturation in *Xenopus* oocytes. Subsequently a human homologue was identified by RT-PCR in 293T cells that was only expressed during the G1/S phase of a cell cycle and it was named Spy1 (Porter et al., 2002). The human sequence has 40% similarity to RINGO/Speedy and it is also able to induce oocyte maturation. Additional mammalian homologs have been discovered (Porter et al., 2002; Cheng et al., 2005b; Dinarina et al., 2005), although most of them, apart from Speedy A orthologues differ substantially in their amino acid sequences between different species (Figure 4-1) (Chauhan et al., 2012). RINGO/Speedy mRNAs demonstrate different expression patterns in different mouse and human tissues, but in both species show high expression in testis suggesting a
role in a reproduction process. Interestingly, sequence analyses have failed to identify RINGO/Speedy homologues in flies, worms, plants or yeasts.

Figure 4-1 Sequence comparison between members of the Speedy A family from different species.

Speedy A proteins from zebrafish, chicken, rat, mouse, human, african frog, ghost shark (Uniprot entries F1QWB8, E1C1Y3, Q8R496, Q5IBH7, Q5MJ7O, Q9PU13, I3PG92) were aligned using Clustal Omega (Sievers et al., 2011) and visualised using the Jalview suite (Waterhouse et al., 2009). Identical amino acids are coloured blue, conserved residues within the family are represented by the range of colours from yellow to brown and identified by the number below. Speedy box, K-, L-, H- and PRGP motifs are boxed in different colours. Figure adopted from (Chauhan et al., 2012).

The most significant characteristic of the family is the existence of a conserved region termed the ‘speedy domain’ or ‘speedy box’ (Cheng et al., 2005b; Dinarina et al., 2005; Chauhan et al., 2012). The sequence that includes the speedy domain, between residues 66-206 in human Speedy A, was found to be responsible for *Xenopus* oocyte maturation and also for CDK2 activation (Dinarina et al., 2005). It has 51-67%
sequence similarity amongst the family of mammalian proteins (Figure 4-2 A). A phylogenetic tree of full length mammalian RINGO-Speedy presented 85% identity between RINGO/Speedy A in human and mice (Figure 4-2 B). The other family members are not present in all the species, with RINGO/Speedy C and E present in humans but not in mice although RINGO/Speedy B and D are not present in humans.

The Speedy domain has a different location in different members of the family. For example in Speedy A it is located almost in the centre of the sequence, whereas in Speedy B it is closer to the C-terminus (Figure 4-2 A). Two motifs that are 100% conserved from human to fish were identified within the speedy box and named the K-motif and the L-motif (Figure 4-1). Their high degree of sequence conservation signifies their potential importance for function. Speedy A proteins from different species also share two other highly conserved motifs: The H- and PRGP motifs are located at the C-terminal end of the speedy domain.

The flanking regions to either side of the speedy box are reported to be dispensable for RINGO/Speedy activity. The N-terminal sequence was reported to control expression levels in transfected cells (Cheng et al., 2005b) and the C-terminal sequence has in some cases been reported to contribute to CDK regulation (Dinarina et al., 2005). As a consequence, the presence and location of the conserved motifs might influence the expression and function of Speedy A and B.
Figure 4-2 RINGO/Speedy protein family.
(A) Cyclin box region of mammalian RINGO/Speedy family members have been compared and similarity within the speedy domain with XRINGO/Speedy has been identified. Cyclin box region coloured yellow and flanking region blue. (B) Phylogenetic tree is adopted from (Nebreda, 2006). It identifies 85% similarity between human and mouse Speedy A. Not all family members present in both mouse and human.

4.1.2 Importance of RINGO/Speedy in meiosis
In Xenopus oocytes Speedy and RINGO can induce meiotic maturation much faster than for example progesterone (Ferby et al., 1999; Lenormand et al., 1999). Ablation of endogenous XRINGO/Speedy mRNAs also delayed progesterone-induced maturation suggesting that it is important for this process (Ferby et al., 1999). Work performed in mice confirmed this finding. Ringo A knock out mice are born without obvious abnormalities but are sterile and exhibit hypoplastic ovaries and testes (Mikolcevic et al., 2016). Interestingly, mouse oocytes and spermatocytes deficient in Ringo A show the same phenotype as those deficient in CDK2 (Viera et al., 2009). CDK2 is essential for male and female meiosis, and cells deficient in CDK2 arrest in a pachytene-like stage of prophase I. They also contain a high amount of γH2AX which indicates unrepaired DNA and undergo apoptosis due to the pachytene checkpoint (Mikolcevic et al., 2016). Recently, it has been suggested that CDK2-Ringo A is involved in the regulation of telomere tethering to the inner nuclear membrane by phosphorylation of the nuclear membrane protein Sun1 (Mikolcevic et al., 2016). The possibility that Ringo controls chromatin methylation has also been investigated as Ringo A deficient spermatocytes show decreased levels of histone H3 trimethylated at Lys9 (H3K9-3me) and increased methylation on Lys4 (H3K4-3me).

4.1.3 RINGO/Speedy and cell proliferation
Human RINGO/Speedy A/Spy1 (hereafter referred to as Speedy A, uniprot entry Q5MJ70) has been proposed to play a role in proliferation: it is expressed during the
G1/S phase transition and activates CDK2 (Porter et al., 2002). Judging by parameters such as increased cell number, mitochondrial activity and level of histone H3, overexpression of human Speedy A leads to cell proliferation in a number of different cell lines. siRNA knockdown of Speedy A resulted in a higher cell population in G1/S compared to G2/M supporting a requirement for CDK2 activation for Speedy A-induced cell proliferation.

### 4.1.4 CDK1 and CDK2 activation by RINGO/Speedy proteins

Both CDK1 and CDK2 can be activated by RINGO/Speedy family members *in vitro* and *in vivo*, though binding of Speedy A to CDK2 is more efficient than its association with CDK1 *in vitro* (Porter et al., 2002; Cheng et al., 2005a). A screen of other CDKs has shown that CDK5 binds to RINGO/Speedy but CDK4 and CDK6 do not (Cheng et al., 2005b; Dinarina et al., 2005).

Despite the fact that cyclins and RINGO/Speedy proteins bind to overlapping sites on CDKs, the mechanism of CDK activation appears to be different (Nebreda, 2006) (Figure 4-3). The canonical CDK activation pathway is cyclin binding accompanied by activation segment phosphorylation. Interestingly, activation of CDK1 and CDK2 by RINGO/Speedy proteins does not require this additional phosphorylation step. CDK1 and CDK2 mutants Thr161Ala and Thr160Ala respectively can be activated by RINGO/Speedy as effectively as the wild type proteins (Karaiskou et al., 2001). However, phosphorylated CDK1/2 can also be isolated bound to Ringo A suggesting that Ringo A can work as a conventional cyclin with activation segment phosphorylation to regulate CDK activity.

Another feature of CDK1 or CDK2 bound to RINGO/Speedy is a lower susceptibility to inhibition by the CKI p21CIP1 (Karaiskou et al., 2001). In complex with Speedy A CDK1 and CDK2 can more efficiently phosphorylate MYT1. Phosphorylated MYT1 in turn loses its activity towards Thr14 and Tyr15 that are inhibitory CDK1 phosphorylation sites. This result has led to speculation about the role of RINGO/Speedy proteins in regulating the maturation of *Xenopus* oocytes (Gutierrez et al., 2006; Ruiz et al., 2008; Ruiz et al., 2010).
Figure 4-3 Comparison of CDK activation by RINGO/Speedy and cyclins.
The cartoon shows the different ways in which CDKs can be regulated by cyclins and RINGO/Speedy. CDKs in complex with RINGO/Speedy do not require additional phosphorylation on the activation segment for full activation in contrast to CDK-cyclin complexes. They can phosphorylate the same substrates but also have specific substrates. CDKs are depicted blue, RINGO/Speedy, dark orange, and the cyclin is drawn in green. Substrates are coloured differently (Nebreda, 2006).

Investigations of substrate specificity of CDK1 and CDK2 activated by RINGO/Speedy has shown different preferences in comparison to cyclin activated complexes (Cheng et al., 2005b). It has been suggested that CDK2-RINGO/Speedy would phosphorylate different substrates due to the difference in preference in the site of phosphotransfer in comparison to CDK2-cyclin A (Figure 4-3). CDK2-cyclin A strongly favours a lysine residue in position +3 of the phospho-acceptor site ((S/T)PX(K/R)) of canonical substrates, as the phosphate moiety of phosphorylatedT160 contacts it directly (Brown et al., 1999). Consequently peptide KSPRK and histone H1 are excellent substrates. In contrast, CDK2 in complex with RINGO/Speedy can tolerate most amino acids in the +3 position and shows a preference for tyrosine, arginine and tryptophan (Cheng et al., 2005a) and therefore can phosphorylate KSPRY and KSPRR peptides. Activation loop phosphorylation has no significant effect on substrate recognition. CDK2-cyclin A and CDK2-RINGO/Speedy can phosphorylate the same substrates as
well. For example, phosphatase CDC25 (A, B and C) have many CDK phosphorylation sites which can be phosphorylated with different efficiency by CDK2-cyclin A and CDK2-RINGO/Speedy A2 (Cheng et al., 2005a).

4.1.5 Sequence and structural analysis of Ringo A/ Speedy

When this work started the structure of CDK2 in complex with Spy1 has not been reported. Previous analysis had suggested that the speedy box in human Speedy A starts at residue 74 and finishes at residue 173 (Chauhan et al., 2012). In earlier papers (Dinarina et al., 2005) the region responsible for oocyte maturation had been mapped to residues 64 to 206. The Speedy A conserved region (Figure 4-1) finishes at residue 234. HHpred analysis (Soding et al., 2005; Alva et al., 2016) was performed on the Speedy A sequence to identify similarities in secondary structure to other proteins. HHpred is able to identify very remotely related proteins, by searching alignment databases such as Pfam or SMART which reduces the number of hits to families rather than single sequences. Cyclin-10 that activates S. cerevisiae CDK PHO85 was predicted to share 85% similarity in its secondary structure to Speedy A in a region between residues 83 and 189, although there is overall only 7.5 % sequence identity between the two proteins (Figure 4-4 A, B).
Figure 4-4 Identification of primary and secondary structure of similarities of Speedy A with another proteins.

(A) Region of Speedy A1 sequence with high secondary structure homology to cyclin 10 (Uniprot entry 4krd_B) identified by HHpred (Soding et al., 2005). (B) Sequence overlay of Speedy A and cyclin 10. Identical residues are identified by a star under the sequence alignment, similar residues identified by two dots. Uniprot entries for Speedy A is Q5IBH7 and for cyclin 10 is P53124.

The structure of cyclin-10 has been solved (PDB code: 4KRD) and was compared to that of cyclin A (PDB code: 1VIN) using CCP4MG (Figure 4-5). The CBF of cyclin -10 overlays well with the cyclin A N-terminal CBF.

Figure 4-5 Overlay of cyclin A with cyclin 10.
Cyclin A (PDB code: 1VIN) is coloured light green, cyclin-10 (PDB code: 4KRD) is coloured lilac. Cyclin A helices are labelled from α1 to α5. Structures were superposed and image created using CCP4mg (Winn et al., 2011).
4.1.6 Structural analysis of RINGO/Speedy in complex with CDK2

While this project was in progress, the structure of human Speedy A referred to as Spy 1 (McGrath et al., 2017) and comprising the speedy domain (residues 61-213) was solved bound to CDK2. The structure revealed some new features of CDK activation and provided structural evidence to show that CDK2 could adopt an active structure in the absence of activation segment phosphorylation.

Spy 1 has one cyclin box fold (CBF) comprised of four smaller helices (α1, α2, α4, and α5) surrounding a longer central one (α3) (Noble et al., 1997). The CBF of Spy 1 is book-ended by one additional helix (α1’) and two shorter additional helices (α6 and α7) at its N- and C-termini respectively (Figure 4-6 A, C). Spy1 and cyclin A adopt a very similar binding orientation towards CDK2. The Spy1 α helices α3 and α5, the α3-α4 and α5-α6 loops form contacts with the C-helix of CDK2, α2, and the α3-α4 loop, and α6 directs the position of the activation segment (Figure 4-6 B).
Figure 4-6 Secondary structure comparison of Spy1 and cyclin A.

(A) Alignment of Spy1 (light blue) with cyclin A2 (light green) (PDB entries: 5UQ2 and 1JST respectively). Cyclin A2 box folds are marked, helices are identified by numbers. The MRAIL motif of cyclin A is coloured purple. (B) Comparison of overall Spy1 and cyclin A2 binding to CDK2 by alignment. CDK2 is coloured coral, CDK2 C helix is coloured gold, activation segment is coloured red, Spy1 helices involved in the interaction are identified by numbers. (C) Sequence alignment of Spy1 homologs within the speedy domain, Spy1, Spy1C, Spy1E4 (Uniprot entries Q5MJ70, Q5MJ68, A6NLX3). Identical residues are indicated with an apteryx underneath the sequences. Sequences folding into helices are identified as red boxes and labelled. Residues of Spy1 that bind to the CDK C helix are coloured blue and residues that interact with the CDK activation segment are red. Structures were superposed and image created using CCP4mg (Winn et al., 2011).

The overall structure of CDK2-Spy1 is very similar to that of CDK2 bound to cyclin A (Jeffrey et al., 1995; Russo et al., 1996b; Welburn et al., 2007). It revealed the features of an active CDK conformation where the C-helix is rotated inwards and the activation segment is moved away from the active site by 5-6 Å when compared to a monomeric CDK2 structure (Figure 4-6B, 4-7A). The C-helix of CDK2 forms similar hydrogen bonds with Spy1 as with cyclin A (Jeffrey et al., 1995) although there are notable differences in the ways in which Spy1 and cyclin A bind to CDK2. Despite the same location on secondary structures of Spy1 A and cyclin A residues that interact with CDK2 their chemical properties are different. Spy1 primarily uses Trp168 but cyclin A uses Phe304 and Leu299 (Figure 4-7 B). The overall structure supports the findings that Spy1 can substitute for the cyclins as a CDK activator (Karaiskou et al., 2001; Cheng et al., 2005a).
Figure 4-7 Comparison of CDK2 secondary structures

(A) Structures of monomeric CDK2 (PDB entry: 1HCK), CDK2-cyclin A (PDB entry: 1JST) and CDK2-Spy1 (PDB entry: 5UQ2) were aligned and the CDK2 structures extracted. Monomeric CDK2 is coloured ice blue, CDK2 in complex with Spy1 is coral and CDK2 in complex with cyclin A is sea green. Movements of the C-helix and the activation segment (T-loop) are shown by colour-coded arrows. (B) C-helix (also termed PSTAIRE helix) interactions with Spy1 and cyclin A. CDK2 is coloured coral and Spy1 is blue, Cyclin A is green. Figures prepared using CCP4mg (McNicholas et al., 2011).

However, the most significant functional difference between the CDK2-cyclin A and CDK2-Spy1 structures is the conformation of the activation loop. Cyclin A binding to CDK2 changes the position of the activation segment moving it out of the active site (De Bondt et al., 1993; Jeffrey et al., 1995) (Figure 4-8) and phosphorylation of Thr160 opens it even further. The position of the non-phosphorylated activation segment of CDK2 bound to Spy1 positioned the way that it resembles activated CDK2 although in slightly different position then either activated or fully activated CDK2 (Figure 4-8).
Figure 4-8 Activation segments comparison.
Comparison of activation segments of CDK2 in complexes with cyclin A (coloured sea green, PDB code: 1JST), phosphorylated on Thr160 CDK2 in complex with cyclin A (coloured pink, PDB code: 4EOQ) and bound to Spy1 (coloured coral, PDB code: 5UQ2). Thereonine 160 is represented in cylindrical form. Figures prepared using CCP4mg (McNicholas et al., 2011).

A complex set of interactions have been proposed to stabilise the CDK2 activation segment conformation. Three acidic residues Glu134, Glu135 and Asp136 of the α3-α4 loop interact within a CDK2 cleft formed between the C-helix and activation segment. Important interactions are made between Asp136 and Arg50 of the CDK2 C-helix and Arg150 of the CDK2 activation segment (Figure 4-9). D136 mimics the phosphate moiety of the phosphorylated Thr160 side chain in pCDK2-cyclin A that is coordinated by three arginine side chains. Spy1 also interacts with residues N terminal to Thr160 to pull the activation segment even further out of the active site cleft. Glu135 of Spy1 forms hydrogen bonds with the amide and side chain oxygen of Thr158 of CDK2 and Asp97 forms a salt bridge with Arg157, for which there are no analogous interactions observed in the CDK2-cyclin A structure. Interestingly, the Spy1 residues involved in T-loop positioning are conserved across the Spy1 paralogs suggesting its importance for CDK activation (Figure 4-6C) (McGrath et al., 2017).
CDK2 is coloured coral and Spy1 coloured blue, CDK2 activation segment is coloured red (PDB code: 5UQ1). Figures prepared using CCP4mg (McNicholas et al., 2011).

The structure of a CDK2-Spy1-p27KIP1 complex has also provided a structural explanation for why CDK2 bound to Spy1 is not as effectively inhibited by CKI CIP/KIP inhibitors as CDK2 bound to cyclin A or cyclin E (McGrath et al., 2017). Spy1 does not contain an MRAIL motif that is embedded in the cyclin recruitment site located within the N-terminal CBF that binds to CDK substrates and inhibitors that contain a (K/R)XLF or (K/R)XLXF motif (Russo et al., 1996b), (Figure 4-6 C). Notably, the α1 and α3 helices that help form the recruitment site are shorter in Spy1 (Figure 4-6 A). However, despite the structural differences, CDK2-Spy1 can form a weak complex with p27KIP1 and is inhibited by the association (McGrath et al., 2017).

4.1.7 Aims and objectives

Structural analysis of a CDK2-Spy1 structure has revealed how Spy1 can activate CDK2 in the absence of activation segment phosphorylation. However, a number of questions remain. Previous work has shown how the structure of the activation segment can dictate CDK sequence preferences around the site of phosphotransfer (Brown et al., 2015). The expression of Spy1 in specialised tissues suggests that it may alter CDK substrate specificity for tissue-specific functions. Speedy A also activates CDK1 for which there is no CDK1-Spy1 structure. The first aim of this project was to express and purify a fragment of Speedy A in order to characterise its binding towards CDK1 and CDK2 and to compare the activity of CDK1 and CDK2 bound to
cyclin A, cyclin B and Speedy A. The second aim was to determine the structures of a CDK1/2-Spy1 complex and compare it to structures of CDK1/2-cyclin B and CDK1/2-cyclin A.

### 4.1.8 RGC-32 as a CDK1 binding partner

Another CDK activator for which there is little structural information is the protein (Response gene to complement 32) (RGC32). RGC32 was discovered in a screen for genes that respond to complement activation in primary rat oligodendrocytes (Badea et al., 1998). The complement system is a part of the body’s immune response that involves around 20 proteins that act against foreign agents, trigger inflammation and remove debris from cells and tissues (Chaplin, 2005). This system must be very carefully regulated so that it does not attack the body’s healthy cells.

The human RGC32 gene is located on chromosome 13 and encodes a 117 amino acid protein with 92% sequence similarity to the rat and mouse proteins (Badea et al., 2002). RGC32 mRNA is over expressed in placenta, skeletal muscle, liver, pancreas, kidney and endothelial cells (Badea et al., 2002). The protein has been detected by Western blotting in human liver, brain and heart (Badea et al., 2002). It is expressed in a wide range of cell lines and in response to different stimuli (Vlaicu et al., 2008). Its overexpression can be induced by complement and subsequent formation of the C5b-9 terminal complement complex, by growth factors (Fosbrink et al., 2003), and by steroid hormones (Chen et al., 2005). As a consequence it has been linked to inflammation, tissue regeneration and cancer. It has also been implicated in cell cycle activation (Badea et al., 1998; Badea et al., 2002) and apoptosis (Tegla et al., 2013).

To demonstrate its relevance to the cell cycle, experiments have been performed in cultured smooth muscle cells, where it was shown that exposure to activated complement triggered expression of RGC32 in S-phase and G2/M entry (Badea et al., 2002). Accordingly, RGC32 silencing in the same cells blocked DNA synthesis induced by C5b-9 and growth factors, confirming the importance of RGC32 for S-phase entry. Furthermore, the knock down of RGC32 by siRNA resulted in a reduction of C5b-9 induced activation and Akt phosphorylation (Fosbrink et al., 2009). From these findings in cultured cells it can be hypothesized that cell cycle progression induced by C5b-9, depends on RGC32 and involves regulation of Akt activity and growth factor release (Fosbrink et al., 2009).
The inhibitory effect on cell cycle activation induced by receptor/CD28 binding has been shown in T-cells of mice (Tegla et al., 2015) (Vlaicu et al., 2010). Because RGC32 induces cell cycle in primary cells (Badea et al., 1998; Fosbrink et al., 2009) but inhibits cell cycle progression in T-cells (Tegla et al., 2015) this protein might play a dual role.

It also has been shown that RGC-32 expression is dependent on age. Fibroblasts from older and younger people were compared and revealed downregulation with age suggesting that decreasing levels of RGC32 are associated with cellular senescence (Kyng et al., 2003). RGC32 has also been reported to form a complex with CDK1-cyclin B and increase its kinase activity. It can be phosphorylated by CDK1 and its phosphorylation on Thr91 is critical for its ability to stimulate CDK1 activity (Badea et al., 2002).

4.1.9 RGC32 sequence analysis

RGC32 is a small protein of 137 amino acids in humans. It is well conserved through the mammalian species with 86% identity (Figure 4-10 A). It has been analysed using secondary structure prediction software HHpred (Soding et al., 2005; Alva et al., 2016) and Ronn v3.2. The analysis has shown mainly disordered structure by Ronn v3.2 (Figure 4-10 B) and HHpred suggested probability of 65% of RGC32 forming a helical structure between aminoacids 14 and 41 (Figure 4-10 C) due to its similarity to proteasome-associated ATPase (Uniprot entry 3M91_B).
Figure 4-10 Sequence comparison between RGC32 from the different species. (A) RGC32 from human (Uniprot entry: Q9H4X1), mouse (Uniprot entry: Q9DBX1) and rat (Uniprot entry: Q9Z2P4) have been compared using Clastal Omega (Sievers et al., 2011). Identical positions coloured dark blue, similar positions are coloured light blue. Conservation level is represented by yellow bars. (B ) Results of protein sequence analysis by the structure prediction software Ronn v3.2 (Yang et al., 2005). (C) RGC32 can form helical structure between aminoacids 14 and 41 with probability of 65% calculated by HHpred (Alva et al., 2016).

4.1.10 RGC32 in cancer

Since RGC-32 promotes cell proliferation it is perhaps not surprising that its expression is upregulated in many cancers, for example in colon (Fosbrink et al., 2005), ovarian (Dassen et al., 2007), prostate (Fosbrink et al., 2005) and breast cancers (Kang et al., 2003). However, in astrocytoma, osteoblastic and colon cancer cell lines overexpression of RGC32 was found in contrast to delay G2/M progression. In these cell lines, RGC32 was preferentially localised in the cytoplasm during interphase, but on centrosomes and spindle poles during prometaphase and metaphase (Saigusa et al., 2007). In this regard, RGC32 can also form a complex with polo-like kinase 1 (Plk1) suggesting it may in certain settings be a negative regulator of the cell cycle. Supporting this potential dual role, studies using gene microarrays have indicated RGC32 mRNA expression to be upregulated in some tumours, for example drug
resistant glioblastoma (Bredel et al., 2006) or multiple myeloma (Zhan et al., 2006) and downregulated in others including ovarian (Donninger et al., 2004) and breast cancer (Chen et al., 2006). RGC32 has also been proposed to be a tumour suppressor protein (Saigusa et al., 2007). However, data from a colon cancer study that detected RGC32 in epithelial tissue in tumours but not in normal adjacent tissue argues against this hypothesis (Fosbrink et al., 2005).

4.1.11 Aims and objectives
RGC-32 is not a very well-studied protein in terms of its structure and interactions with members of the CDK family. It is known as an activator of CDK1 but neither its binding site on CDK1, nor its mechanism of action have been characterised. In collaboration with the laboratory of Professor Michelle West the aim of this project was to characterise the interaction of RGC32 with CDK1 and CDK1-cyclin B and assess its effect on CDK1 activity.

4.2 Description of biophysical methods
4.2.1 Theoretical background of surface plasmon resonance
Surface plasmon resonance is a very sensitive biophysical method that can measure label-free real time interactions between molecules. It measures the association and dissociation rates of protein interactions from which dissociation constants can be calculated. The basis of this technique lies in the measurement of changes in the refractive index of the liquid close to a metal surface to which a protein bait (conjugated ligand) is bound. The presence of weight in one position resonates to electrons across the whole surface. This change in electron resonance affects the refractive index which can be measured by the detector as the change in the angle of the refracted light.
Figure 4-11 Schematic representation of surface plasmon resonance.
An antibody is attached to the gold surface to act as a bait for the protein of interest. The protein of interest (which is also called the ligand) in its turn can bind the binding partner (analyte) flowed over the chip surface. Due to the change of mass, refractory properties of the gold surface change accordingly resulting in the alteration of the refractive angle. Changes of refractive angle are measured by the detector.

The first SPR instrument was developed in 1991 using BiaCore technology. Modern instruments now use sensor chips with a gold surface that can be modified with different matrices depending on the application (Figure 4-11 A). For example, dextran can be activated to bind directly to molecules of interest. Molecules of interest that attached to the surface of the chip are called analytes and the ones that flow through the cell are called ligand. Polarised light from the laser source is directed at a certain angle through the half circular prism to the gold surface where the plasmon resonance is generated. As ligand binds to the surface it changes the refractive index that affects
the angle of reflection recorded by the detector. The mass of bound ligand is proportional to the angle of the plasmon resonance. The results of SPR experiments are normally represented by a sensogram (Figure 4-11 B). The shape of the curve can be used to measure the rate of association (\(k_a\)). Once the binding reaches an equilibrium and the sensor gram plateaus, the rate of dissociation (\(k_d\)) can be measured as the ligand is released during a wash step. The \(K_d\) can be determined from the \(k_a\) and \(k_d\) values or alternatively from the equilibrium binding determined from the injection of a ligand concentration series, the concentration of the bound ligand relating to the response units (RU).

\[
K_d = \frac{k_d}{k_a}
\]

Equation 4-1 Binding constant determination.

Abbreviations: \(K_d\) – dissociations constant, \(k_d\) – dissociation rate, \(k_a\) – association rate.

### 4.3 Materials and Methods

#### 4.3.1 Generation of constructs

The full length human Speedy A sequence (uniprot entry Q5MJ70) was ordered from Open Biosystems, (Clone ID 4839330) in the pBluescriptR vector. Custom designed primers were purchased from Eurofins Genomics, purified by a standard desalting protocol and re-suspended in nuclease free water (Ambion). The full list of primers is provided in Table 4-1.

PCR reactions were performed to generate DNA inserts for subsequent cloning. Each PCR reaction contained 1 x Phusion HF buffer, 200 \(\mu\)M dNTPs, 0.2 \(\mu\)M of forward primer and 0.2 \(\mu\)M of reverse primer, 50 ng template DNA and 1U Phusion Hot Start DNA Polymerase (NEB). Reactions were started by heating to 98°C for 30 seconds and then cycled through the following steps 30 times: (i) 5 sec denaturation step at 98°C, (ii) 5 sec annealing step at 65°C and (iii) an extension step depending on the length of DNA to be amplified, calculated at 30 sec per kb pair. The final extension at 72°C lasted for 7 minutes before reactions were held at 4°. To remove methylated template DNA, the samples were digested by addition of 0.5 \(\mu\)l Dpn1 (20,000 U/ml) at 37°C for 1 hour and subsequently purified using a Qiagen QIAquick PCR purification
The DNA concentration was estimated by Nanodrop Spectrophotometer 2000 (Thermoscientific).

Different plasmids were used depending on the expression system required (Table 4-2, Table 4-3), pGEX6P-1 and pET21dMBP were used for expression in E. coli and pACEBac1 was used for virus generation and future expression in insect cells. Plasmids were prepared by digesting with the appropriate restriction enzymes in the presence of the appropriate restriction enzyme buffer for 3 hours all together but with new addition of the restriction enzyme every hour at 37°C followed by enzyme heat inactivation at 65°C for 20 minutes. Plasmids were purified using Qiagen QIAquick PCR purification kits.

Purified inserts and plasmids were validated by agarose gel electrophoresis and subsequently cloned by in-fusion cloning (Takara Bio USA). 100 nM of DNA fragment of interest and 100 nM of linearized vector were incubated for 15 minutes at 50°C in the presence of 1x in-fusion enzyme premix. After the incubation reactions were stored on ice.

<table>
<thead>
<tr>
<th>Construct</th>
<th>Cloning site</th>
<th>Forward primer</th>
<th>Reverse primer</th>
</tr>
</thead>
<tbody>
<tr>
<td>HsSpeedyA21-313/pGEX6P-1</td>
<td>BamH1 EcoR1</td>
<td>GGGGCCCTGGGATCA TGAGCACAATCAGATG</td>
<td>GTGCACCGGGAATT TGGGGCCATCTCATT</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TGTGTTGAGACAC</td>
<td>CTTCACTTCTGG</td>
</tr>
<tr>
<td>HsSpeedyA264-313/pGEX6P-1</td>
<td>BamH1 EcoR1</td>
<td>GGGGCCCTGGGATCC TGGTTTACAGCCTCAG</td>
<td>GTGCACCGGGAATT TGGGGCCATCTCATT</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GATAGACTGC</td>
<td>CTTCACTTCTGG</td>
</tr>
<tr>
<td>HsSpeedyA264-234/pGEX6P-1</td>
<td>BamH1 EcoR1</td>
<td>GGGGCCCTGGGATCC TGGTTTACAGCCTCAG</td>
<td>GTGCACCGGGAATT TGGGGCCATCTCATT</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GATAGACTGC</td>
<td>CTTCACTTCTGG</td>
</tr>
<tr>
<td>HsSpeedyA21-189/pET21dMBP3C</td>
<td>BamH1 Xho1</td>
<td>AGGGGCCCTGGGATCA TGAGGACAATCAGA</td>
<td>GGTGGGTTGCTCGA TCAAGCTCCACTGTGA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TGTGTTGAGACAC</td>
<td>TGAACAGACGTGTTTCT</td>
</tr>
<tr>
<td>HsSpeedyA21-206/pET21dMBP3C</td>
<td>BamH1 Xho1</td>
<td>AGGGGCCCTGGGATCA TGAGGACAATCAGA</td>
<td>GGTGGGTTGCTCGA TCAAGCTCCACTGTGA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TGTGTTGAGACAC</td>
<td>TGAACAGACGTGTTTCT</td>
</tr>
<tr>
<td>HsSpeedyA21-313/pET21dMBP3C</td>
<td>BamH1 Xho1</td>
<td>AGGGGCCCTGGGATCA TGAGGACAATCAGA</td>
<td>GGTGGGTTGCTCGA TCAAGCTCCACTGTGA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TGTGTTGAGACAC</td>
<td>TGAACAGACGTGTTTCT</td>
</tr>
<tr>
<td>HsSpeedyA227-189/pET21dMBP3C</td>
<td>BamH1 Xho1</td>
<td>AGGGGCCCTGGGATCA GCCUAAAAAGGCCAUU</td>
<td>GGTGGGTTGCTCGA TCAAGCTCCACTGTGA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ACUCUGAAGC</td>
<td>TGAACAGACGTGTTTCT</td>
</tr>
<tr>
<td>HsSpeedyA227-206/pET21dMBP3C</td>
<td>BamH1 Xho1</td>
<td>AGGGGCCCTGGGATCA GCCUAAAAAGGCCAUU</td>
<td>GGTGGGTTGCTCGA TCAAGCTCCACTGTGA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ACUCUGAAGC</td>
<td>TGAACAGACGTGTTTCT</td>
</tr>
</tbody>
</table>
4.3.2 E. coli transformation

DNA from the in-fusion reaction were transformed into Stellar™ competent cells (Clontech) following the manufacturer’s protocol. Briefly, 2µl of the in-fusion reaction were added to 100 µl of thawed competent cells and incubated on ice for 30 minutes. Cells were heat shocked for 45 seconds at 42°C followed by cooling down on ice for 2 minutes. 900 µl of SOC media was added before further incubation at 37°C for 1 hour. Cells were plated on agar containing the appropriate antibiotic and incubated at 37°C overnight. Single colonies were used for inoculation in 10 ml of LB media containing appropriate antibiotic at 37°C while shaking overnight. 5 ml of culture was pelleted by centrifugation at 6797 x g for 3 minutes and DNA extracted using a Miniprep kit (Qiagen). The DNA concentration was measured by Nanodrop 2000 (Thermo Scientific) at 280 nm wavelength. DNA samples were sent for sequencing analysis by Eurofins Genomics. The results were analysed using the ExPASy translation tool (Gasteiger et al., 2003) and sequence alignment by Clustal Omega (Sievers et al., 2011).

Bacmids were produced by transformation of DH10EMBacY cells with 1 µg pACEBac1 vector containing MBP-tagged truncated variants of the human Speedy A sequence. Transformation reactions were incubated for 16 hours at 37°C and then plated on LB agar plates containing 50 µg/ml ampicillin, 30 µg/ml kanamycin, 7 µg/ml gentamycin, 10 µg/ml tetracycline, 100 µg/ml X-Gal and 40 µg/ml IPTG. White colonies were selected after 24 hours incubation and used to inoculate 10 ml LB without added
antibiotic. Bacteria were pelleted by centrifugation and bacmid DNA prepared using a MiniPrep kit (Qiagen) with the modification that the column filtration step was substituted by isopropanol precipitation followed by a 70% ethanol wash. Recombinant bacmid was re-suspended in 20 µl of sterile water and used straight away or stored at -20°C for later use to generate recombinant baculovirus.

### 4.3.3 Baculovirus preparation using the MultiBac expression system

20 µl of recombinant bacmid was mixed with 200 µl of Sf900 media warmed to 27°C. 100 µl of GeneJuice™:Sf900 mixture in a ratio of 1:10 was added to each tube and left to incubate for 30 minutes. 3 ml of Sf9 cells were seeded into 6-well plates to give a total cell number of 0.5 x 10⁶ cells. 150 µl of the DNA/GeneJuice™/Sf900 cocktail was added to each well and left to incubate at 27°C in the dark for 60 hours. Because the bacmid contains the YFP gene, the success of transfection can be monitored by fluorescence microscopy (Nikon Eclipse TE2000-U). Further expression tests were also required to check the efficiency of transfection. When fluorescence was observed, the media was harvested separately from each well as the P0 stock and used for the next generation of virus.

### 4.3.4 Insect cell culture and virus amplification

P0 virus was used to generate P1 stock by infecting 1 x 10⁶ Sf9 cells. 0.5 ml of P0 stock was used for every 50 ml of cells. Cells where monitored every day and harvested 24 hours after arrest of cell proliferation which typically occurred after 96 hours incubation. P2 was generated from P1 viral stock and was routinely used for protein expression.

### 4.3.5 Expression tests

To identify the optimal volume of viral stock for infection and length of infection all viruses were subjected to expression tests. 50 ml of cells were infected with 50, 100 and 500 µl of P2 virus and incubated with shaking at 27°C. 10 ml aliquots were collected after 72, 96 and 120 hours and pelleted by centrifugation at 646 x g for 15 minutes. Pellets were frozen to test expression at the end of the trial. Optimal conditions for infection were used for subsequent protein expression.

To optimise protein expression in recombinant *E. coli* cells, DNA sequences of interest were cloned into bacterial expression vectors and transformed into Rosetta
2\textsuperscript{TM}(DE3)pLysS cells. 10 ml of starter culture was prepared from a single colony in LB containing appropriate antibiotic grown overnight at 37°C in a shaking incubator. 1 ml overnight cultures were used to inoculate 100 ml of LB media supplemented with antibiotic at 37°C with shaking (200 rpm) until the optical density at 600 nm (OD\textsubscript{600}) reached 0.6-0.8. The temperature was then reduced to 20°C for 30 minutes before induction with 100 µM IPTG followed by incubation with shaking for 16-20 hours. Bacterial cells were pelleted by centrifugation (JLA 8.1000 rotor at 6238 x g, 20 mins) and pellets were stored at -20°C for further expression tests.

Frozen cell pellets were re-suspended in 1 ml of BugBuster (Merck Millipore) with addition of 0.25 mg/ml Lysozyme, 50 µg/ml RNase, 10 µg/ml DNase and 5 mM MgCl\textsubscript{2}. The cells were incubated with rotation for 20 min at room temperature for cell lysis. 3 µl of each lysate was taken for SDS-PAGE analysis and the rest was transferred into Eppendorf tubes and centrifuged for 20 min at 6500 x g. 3 µl of each supernatant was analysed by SDS-PAGE. Amylose beads were equilibrated in TBS buffer (50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 0.01% MTG) as a 50% slurry and supernatants of each sample were added to 50 µl bead aliquots and incubated for 1h at +4°C on a rotating wheel. Beads were washed with ice cold TBS 3 times in 0.5 ml volume and pelleted by centrifugation at +4° at 1610 x g. After the final wash 30 µl of RunBlue LDS Sample buffer (Expedeon) was added to the pelleted beads and denaturated through heating at 95°C for 5 minutes. 15 µl of each sample was analysed by SDS-PAGE as described in section 2.2.10.

**4.3.6 Western blotting**

After separation by SDS-PAGE, proteins were transferred by wet transfer onto Hybond\textsuperscript{TM} –C Extra nitrocellulose membrane (Amersham Bioscience, GE Healthcare, Little Chalford, UK) at 200 V for 75 min in Tris-Glycine transfer buffer (Expeodon Protein Solutions, Cambridge, UK). The nitrocellulose membrane was blocked in 5% milk in Tris-buffered saline with 0.1% Tween 20 (TBS-T) for 20 min followed by 15 hour incubation at +4°C with Speedy polyclonal rabbit primary antibody (Bio-Technne, Abingdon, UK) in 2.5% milk TBS-T, dilution 1:1000. After the incubation, the membrane was washed twice in TBS-T and left at RT on a rocker with gentle agitation for 1 hour in Goat Anti-Rabbit IgG, HRP conjugate secondary polyclonal antibody (Dako) diluted in TBS-T 1:2000. Unbound secondary antibody were removed by washing in TBS-T.
twice. Proteins were detected by enhanced chemiluminescence using Amersham ECL Western blotting detection reagent (GE Healthcare, Little Chalfont, UK) and visualised and analysed in a FujiFil LAS-3000 (Fujifilm, Tokyo, Japan).

4.3.7 Surface plasmon resonance (SPR)

All SPR experiments were carried out on a Biacore S200 (GE Healthcare) system. All runs were performed at 25°C using buffer containing 10 mM HEPES, 150 mM NaCl, 3 mM EDTA, 0.05% TWEEN. CM5 BIACore sensor chips were prepared by amine coupling goat anti-GST antibody to a chip surface using a GST capture kit and following the standard protocol provided by GE Healthcare. Protein samples were clarified by centrifugation at 10000 x g at 4°C (Beckman Coulter Allegra 22R) and diluted to the required concentrations before use. 20 µg/ml GSTCDK1, GSTCDK2 or GSTRGC32 were captured at 5 µl/min for 400 s (GSTCDK2) or 500 s (GSTCDK1 and GSTRGC32). CDK1, RGC32 or CKS1 were used as analytes at different concentrations. CDK1 was used at 40 nM, 200 nM, 1000 nM, 2000 nM and 5000 nM; RGC32 at 62.5 nM, 250 nM, 1000 nM, 2500 nM, 4000 nM; and CKS1 at 0.8 nM, 4 nM, 20 nM, 100 nM, and 500 nM. Analytes were flowed over the GST-captured protein at 30 µl/min for 380 s followed by dissociation over 720 s. Between each run GST-captured protein was removed from the antibody surface by 20 mM glycine pH 2.2 at 30 µl/min for 240 s and then recaptured. GST was used as a negative control on a separate lane and was captured at 10 µg/ml for 60 s. GST sensogram readings were subtracted from the sample sensograms to eliminate bulk effects and non-specific binding. $K_d$ values together with association $k_a$ and dissociation rates $k_d$ were calculated by Biacore evaluation software.

4.4 Results

4.4.1 Generation of Speedy/Ringo A constructs

Based on secondary structure prediction and published data (Dinarina et al., 2005) a number of sequences were selected to be cloned into vectors for protein expression (Table 4-2 and Figure 4-13). During the first round of cloning, bacterial protein expression was assessed. Three different constructs to express the full length protein, the speedy box (residues 64-313) and the C-terminal domain (residues 64-234 that contains the speedy box and H- and PRGP-motifs) were cloned by in-fusion into
pGEX6P-1 which contains an N terminal GST tag. The cyclin box of Speedy A according to HHpred starts at residue 83 but earlier published data suggested that residues between 64 and 83 were important for function (Dinarina et al., 2005). Consequently we chose to start at residue 64 (Figure 4-12).

<table>
<thead>
<tr>
<th>10</th>
<th>20</th>
<th>30</th>
<th>40</th>
<th>50</th>
</tr>
</thead>
<tbody>
<tr>
<td>MRHNQ MCCET</td>
<td>PPTTVVYVKS</td>
<td>GSNRSHQPKK</td>
<td>PITLKRPICK</td>
<td>DNWQAFEKNT</td>
</tr>
<tr>
<td>60</td>
<td>70</td>
<td>80</td>
<td>90</td>
<td>100</td>
</tr>
<tr>
<td>HNNKSKRPK</td>
<td>GPCLVIQRQD</td>
<td>MTAFFKLFD</td>
<td>DL IQDFLWMD CCCKIADKYL</td>
<td></td>
</tr>
<tr>
<td>110</td>
<td>120</td>
<td>130</td>
<td>140</td>
<td>150</td>
</tr>
<tr>
<td>LAMTFVYFKR</td>
<td>AKFTISEHTR</td>
<td>INFFIALYLA</td>
<td>NTVEEDEET</td>
<td>KYEIFPWALG</td>
</tr>
<tr>
<td>160</td>
<td>170</td>
<td>180</td>
<td>190</td>
<td>200</td>
</tr>
<tr>
<td>KNWRKLFPNF</td>
<td>LKLRDQLWDR</td>
<td>IDYRAIVSRR</td>
<td>CCEEVMAIA</td>
<td>P THYIWQRERS</td>
</tr>
<tr>
<td>210</td>
<td>220</td>
<td>230</td>
<td>240</td>
<td>250</td>
</tr>
<tr>
<td>V HHSGAVRN</td>
<td>NRDEVQL PRG PS ATPVDCSL</td>
<td>CGKKRRYVRL</td>
<td>GLSSSSSLS</td>
<td></td>
</tr>
<tr>
<td>260</td>
<td>270</td>
<td>280</td>
<td>290</td>
<td>300</td>
</tr>
<tr>
<td>HTAGVTEKHS</td>
<td>QDSYNLSMD</td>
<td>IIGDPSQAYT</td>
<td>GSEVVNDHQS</td>
<td>NKGKKTNFLK</td>
</tr>
<tr>
<td>310</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KDKSMEWFTG</td>
<td>SEE</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Figure 4-12 Protein sequence of Speedy A.**
Speedy box predicted by HHpred software. Residues 83 – 189 coloured orange, H motif – purple and PRGP motif – red (Soding et al., 2005).
Figure 4-13 Sequences of Speedy1/Ringo A cloned into pGEX6P-1 vector.

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Tag</th>
<th>Vector</th>
<th>Host E. coli strain</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-313</td>
<td>GST</td>
<td>pGEX6P-1</td>
<td>Rosetta\textsuperscript{TM} (DE3 )pLysS</td>
</tr>
<tr>
<td>64-313</td>
<td>GST</td>
<td>pGEX6P-1</td>
<td>Rosetta\textsuperscript{TM} (DE3 )pLysS</td>
</tr>
<tr>
<td>64-234</td>
<td>GST</td>
<td>pGEX6P-1</td>
<td>Rosetta\textsuperscript{TM} (DE3 )pLysS</td>
</tr>
<tr>
<td>1-313</td>
<td>MBP</td>
<td>pET21dMBP3C</td>
<td>Rosetta\textsuperscript{TM} (DE3 )pLysS</td>
</tr>
<tr>
<td>1-189</td>
<td>MBP</td>
<td>pET21dMBP3C</td>
<td>Rosetta\textsuperscript{TM} (DE3 )pLysS</td>
</tr>
<tr>
<td>1-206</td>
<td>MBP</td>
<td>pET21dMBP3C</td>
<td>Rosetta\textsuperscript{TM} (DE3 )pLysS</td>
</tr>
<tr>
<td>27-189</td>
<td>MBP</td>
<td>pET21dMBP3C</td>
<td>Rosetta\textsuperscript{TM} (DE3 )pLysS</td>
</tr>
<tr>
<td>27-206</td>
<td>MBP</td>
<td>pET21dMBP3C</td>
<td>Rosetta\textsuperscript{TM} (DE3 )pLysS</td>
</tr>
<tr>
<td>27-313</td>
<td>MBP</td>
<td>pET21dMBP3C</td>
<td>Rosetta\textsuperscript{TM} (DE3 )pLysS</td>
</tr>
<tr>
<td>64-189</td>
<td>MBP</td>
<td>pET21dMBP3C</td>
<td>Rosetta\textsuperscript{TM} (DE3 )pLysS</td>
</tr>
<tr>
<td>64-206</td>
<td>MBP</td>
<td>pET21dMBP3C</td>
<td>Rosetta\textsuperscript{TM} (DE3 )pLysS</td>
</tr>
<tr>
<td>64-313</td>
<td>MBP</td>
<td>pET21dMBP3C</td>
<td>Rosetta\textsuperscript{TM} (DE3 )pLysS</td>
</tr>
</tbody>
</table>

Table 4-2 Ringo/Speedy A constructs
Constructs have been generated for expression in E.coli and insect cells. GST, glutathione-S-transferase; MBP, maltose binding protein.
Prepared to identify optimal protein expression conditions, vectors were transformed into a set of expression strains, ArcticExpress(DE3) (Agilent Technologies), Lemo21(DE3) (NEB), and Rosetta™ (DE3)pLysS (Novagen), and grown in Lysogeny (LB), Terrific broth (TB) or Auto-induction media (AIM) (Figure 4-14 A,B).

Figure 4-14 SDS-PAGE analysis of test expression of Speedy A constructs in different E. coli strains in different media.

(A) Comparison of soluble protein expression in Arctic Express (DE3), marked with a red square, Rosetta™ (DE3)pLysS, marked with a green square, Lemo21(DE3) (NEB), marked with a yellow square. Buffers identified below the gels. AIM is auto induction media, TB is terrific broth and LB is Lysogeny broth. Speedy A (residues 1-313) constructs are in lanes 1-3, 10-12, 19-21. Molecular weight of GST tagged protein is 62.5 kDa. Speedy A (residues 64-313) constructs are in lanes 4-6, 13-15, 22-24. Molecular weight of GST tagged protein is 55.2 kDa. Speedy A (residues 64-234) constructs are in lanes 7-9, 16-18, 25-27. Molecular weight of GST tagged protein is 46.4 kDa. Expected molecular weights of truncated Speedy A constructs are identified at the side of the gels. (B) Comparison of protein bound to glutathione beads assessed by pull-downs. Lanes 1-3, 10-12, 19-21 contain Speedy A (residues 1-313), lanes 4-6, 13-15, 22-24 contain Speedy A (residues 64-313), lanes 7-9, 16-18, 25-26 contain Speedy A (residues 64-234) proteins.

Unfortunately, the results of this experiment (Figure 4-14) were difficult to interpret: there are no obvious over-expressed proteins running in the expected places. Analysis
of the proteins bound to the glutathione beads revealed multiple species bound for each construct, although constructs Speedy A (1-313) and Speedy A (64-313) appeared to express better in Arctic Express and Rosetts™ (DE3)pLysS cells (Compare lanes 1—6 and 10-15 in Figure 4-14 A, B).

Western blotting was subsequently used to check for Speedy A expression (Figure 4-15). The full-length protein and construct 64-234 containing the speedy box were detected in the whole cell extract (Lanes 1 and 3), as well as in the supernatant (Lanes 4 and 6) and the pull downs (Lane 7 and 9). Construct 64-313 had a weaker signal in the supernatant and could be detected only after a longer exposure time. As expected, this sample then generated a very weak signal by pull down which could be detected only after 2 hours development with the FUJI film (Figure 4-15 B, Lane 5).
Figure 4-15 SDS-PAGE gel and Western blotting analysis of Speedy A/Ringo expression test.

GST tagged Speedy A/Ringo constructs 1-313, 64-313, 64-234 were inoculated as described and to determine its level of expression the whole cell fraction, supernatant and the bead bound protein was run on two gels. One gel was stained with Instant Blue (Expedeon), (A) and the another was immunoblotted with polyclonal antibody against Speedy A (B). Films were developed for a different length of time, as indicated. Expected molecular weights of constructs identified on the side of the gels and films.

To further assess protein expression, the full-length construct was selected for larger scale expression. The protein was expressed in 1L of LB media and purified by sequential affinity, ion exchange (results not shown) and size exclusion chromatography. Ion exchange has been employed to remove GST and chaperone that was identified by SDS-PAGE. Although a peak corresponding to the expected size of monomeric Speedy A (1-313) can be seen on the chromatogram (Figure 4-16) the yield was very low. From 1L, circa 0.5 mg of GST-Speedy A fusion was prepared.
Figure 4-16 Purification of Speedy A (1-313) by size exclusion chromatography.

(A) Size-exclusion chromatogram. Fractions from peaks eluting between 100 and 160 ml were collected. 1 ml of fractions 22-26 (lines 1-5, identified orange) and 34-37 (lines 6-9, identified red) were incubated for 1 min with 4 µl of protein extracting beads. Beads were heated for 2 min at 95°C with LDS to elute the bound proteins which were then analysed by SDS-PAGE. (B) A band of the expected molecular weight is eluted as a shoulder on the main peak.
The first round of test expressions demonstrated that further optimisation was required. Different protein tags, an increased number of truncations and the alternative insect cell based expression system were tested. The maltose binding protein (MBP) tag was selected for its ability to improve protein solubility. The MBP sequence was amplified from the pMAL-p5X vector together with the polylinker and Factor Xa cleavage site and cloned into pET-21d(+). To improve purification options, the 3C cleavage site was introduced between the factor Xa cleavage site and the hexahistidine tag (Figure 4-17).

![Alignment of MBP sequence from pMAL-5x vector with sequence pET21dMBP3C.](image)

End of MBP sequence marked with the grey line, factor Xa cleavage site IEGR and 3C cleavage site underlined red.

The following MBP-Speedy A proteins were prepared: (i) The speedy box (64-189), (ii) the speedy box and the sequence important for oocyte maturation (64-206), (iii) the speedy box and C-terminal sequence important for CDK activation (64-313); (iv) the speedy box and N-terminal sequences reported to influence protein expression in transfected cells (27-189, 27-206, 27-313 and 1-189, 1-206, 1-313) (Cheng et al., 2005b) (Figure 4-18).
Figure 4-18 Graphical representation of MBP and Speedy A2 constructs cloned into pET-21d (+).
The Speedy box is represented by the red bar and the N- and C-terminal flanking regions are green. MBP coloured yellow.

Constructs were again transformed into different bacterial strains, this time Tuner\textsuperscript{TM}(DE3) and Rosetta\textsuperscript{TM} (DE3)pLysS, and inoculated in LB and TB media. As described earlier, whole cell lysate, supernatant and pull down samples were analysed by SDS-PAGE to detect protein expression. As can be seen from the gel (Figure 4-19) there was no obvious overexpression in the whole cell lysates or the clarified lysates. Although there is an overexpressed protein in the Speedy A (64-313) sample, it is at 46 kDa and so corresponds to MBP rather than the MBP-Speedy A (64-313) fusion (71.02 kDa). More promising are the bands in lane 7 (Speedy A (64-189)) and lane 8 (Speedy A (64-206)) of the pull downs from Rosetta\textsuperscript{TM} (DE3)pLysS competent cells. These are running at circa 55 kDa, close to the expected size of the fusion proteins. These constructs were selected for larger scale expression and testing for CDK2 binding \textit{in vitro}.
A

Whole cell lysate

Rosetta™ (DE3)pLysS in LB

Tuner™ (DE3)pLysS in LB

Rosetta™ (DE3)pLysS in TB

Tuner™ (DE3)pLysS in TB

B

Clarified lysate

Rosetta™ (DE3)pLysS in LB

Tuner™ (DE3)pLysS in LB

Rosetta™ (DE3)pLysS in TB

Tuner™ (DE3)pLysS in TB
Figure 4-19 SDS-PAGE analysis of test expression of Speedy constructs.
Constructs described and numbered in Table 4.3 were transformed into Rosetta™ (DE3)pLysS and Tuner™(DE3)pLysS E. coli strains for test expression. A – Whole cell lysate, each gel has a name of the competent cell strain where protein has been expressed. B – Clarified cell lysate, C – Pull-downs of Speedy A2 constructs.

<table>
<thead>
<tr>
<th>Lane</th>
<th>Construct</th>
<th>MW of Speedy A construct</th>
<th>MW of MBP tagged Speedy A</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Speedy A2 1-189</td>
<td>22.6 kDa</td>
<td>64.6 kDa</td>
</tr>
<tr>
<td>2</td>
<td>Speedy A2 1-206</td>
<td>24.6 kDa</td>
<td>66.6 kDa</td>
</tr>
<tr>
<td>3</td>
<td>Speedy A2 1-313</td>
<td>36.7 kDa</td>
<td>78.7 kDa</td>
</tr>
<tr>
<td>4</td>
<td>Speedy A2 27-189</td>
<td>19.5 kDa</td>
<td>61.5 kDa</td>
</tr>
<tr>
<td>5</td>
<td>Speedy A2 27-206</td>
<td>21.5 kDa</td>
<td>63.5 kDa</td>
</tr>
<tr>
<td>6</td>
<td>Speedy A2 27-313</td>
<td>33.4 kDa</td>
<td>75.4 kDa</td>
</tr>
<tr>
<td>7</td>
<td>Speedy A2 64-189</td>
<td>15.2 kDa</td>
<td>57.2 kDa</td>
</tr>
</tbody>
</table>
Because only two constructs showed low, detectable expression, the decision was taken to assess expression using insect cells. The cassettes to express the MBP-Speedy A constructs were amplified using newly designed universal primers (Table 4-1) and cloned into the pACEBac1 vector digested with BamH1 and EcoR1. pACEBac1 contains elements required for integration into a baculovirus via Tn7 transposition. It also carries the gentamycin resistance gene and contains a very late polyhedrin promoter and a terminator sequence from SV40 which can be substituted if required. The baculoviral genome called MultiBac is specially engineered to improve protein production properties. The two important genes, v-cath and chiA have been disrupted. Disruption of v-cath leads to elimination of proteolytic activity triggered by V-CATH protease. Disruption of chiA, which encodes chitinase, provides an opportunity to use chitin-affinity chromatography to isolate intein-chitin binding domain fused proteins without interference from chiA gene product. The disrupted DNA sequence has been replaced with a LoxP site to enable cre-lox recombination. However, in these experiments LoxP was not used for insertion of the gene of interest. Instead the sequences were transferred into the bacmid via transposition into the mini Tn7 attachment site located in the LacZα subunit coding sequence (Figure 4-19). Disruption of the LacZ gene leads to clones carrying inserted DNA to appear white. White clones containing baculoviral DNA were inoculated in LB and then the DNA was purified and used for virus production as described in Materials and Methods.
Figure 4-20 DH10MultiBacTM and DH10EmBacY baculoviral DNA.
Baculoviral genome is shown in schematic representation. Tn7 attachment site on LacZ gene shown as a black bar. Disrupted genes chiA and V-cath replaced with antibiotic resistance marker and gene encoding YFP (Fitzgerald et al., 2006).

In contrast to the DNA generated in DH10MultiBac™, baculoviral DNA produced in DH10EmBacY cells expresses yellow fluorescent protein and so infected cells could be detected using a fluorescent microscope. Speedy A sequences of interest produced using MultiBac™ baculoviral DNA technology (Table 4-4) were amplified to P2 generation and were then used to test expression of recombinant proteins in insect cells. Unfortunately, as summarised in Figure 4-21, analysis of the whole cell, supernatant and bead-bound fractions for each of these constructs failed to identify conditions from which a significant amount of Speedy A could be purified.

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Tag</th>
<th>Vector</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-313</td>
<td>GST</td>
<td>pACEBac1</td>
</tr>
<tr>
<td>64-313</td>
<td>GST</td>
<td>pACEBac1</td>
</tr>
<tr>
<td>64-234</td>
<td>GST</td>
<td>pACEBac1</td>
</tr>
<tr>
<td>1-313</td>
<td>MBP</td>
<td>pACEBac1</td>
</tr>
<tr>
<td>1-189</td>
<td>MBP</td>
<td>pACEBac1</td>
</tr>
<tr>
<td>1-206</td>
<td>MBP</td>
<td>pACEBac1</td>
</tr>
</tbody>
</table>
Table 4-4 SpeedyA2 constructs for expression in insect cells.

<table>
<thead>
<tr>
<th>Construct</th>
<th>Sample</th>
<th>Type of baculoviral DNA</th>
<th>MW of MBP tagged Speedy A</th>
</tr>
</thead>
<tbody>
<tr>
<td>27-189 MBP pACEBac1</td>
<td>Whole cell lysate</td>
<td>EmBacY</td>
<td>78.7 kDa</td>
</tr>
<tr>
<td>27-206 MBP pACEBac1</td>
<td>Clarified lysate</td>
<td>EmBacY</td>
<td>78.7 kDa</td>
</tr>
<tr>
<td>27-313 MBP pACEBac1</td>
<td>Pull down</td>
<td>EmBacY</td>
<td>78.7 kDa</td>
</tr>
<tr>
<td>64-189 MBP pACEBac1</td>
<td>Whole cell lysate</td>
<td>MultiBac™</td>
<td>78.7 kDa</td>
</tr>
<tr>
<td>64-206 MBP pACEBac1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>64-313 MBP pACEBac1</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 4-21 Pull downs of Ringo A/Speedy constructs.
(A) SDS-PAGE gel of Ringo A/Speedy pull downs expressed in insect cells. Constructs are labelled and described in Table 4-5. (B) Ringo A/Speedy pull downs transferred to a nitrocellulose membrane and blotted with Ringo A/Speedy primary rabbit antibody.
Table 4-5 Table of constructs with molecular weights

<table>
<thead>
<tr>
<th></th>
<th>Speedy A2 1-313</th>
<th>Clarified lysate</th>
<th>MultiBac™</th>
<th>78.7 kDa</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>Speedy A2 1-313</td>
<td>Pull down</td>
<td>MultiBac™</td>
<td>78.7 kDa</td>
</tr>
<tr>
<td>7</td>
<td>Speedy A2 1-206</td>
<td>Whole cell lysate</td>
<td>MultiBac™</td>
<td>66.6 kDa</td>
</tr>
<tr>
<td>8</td>
<td>Speedy A2 1-206</td>
<td>Clarified lysate</td>
<td>MultiBac™</td>
<td>66.6 kDa</td>
</tr>
<tr>
<td>9</td>
<td>Speedy A2 1-206</td>
<td>Pull down</td>
<td>MultiBac™</td>
<td>66.6 kDa</td>
</tr>
<tr>
<td>10</td>
<td>Speedy A2 1-206</td>
<td>Whole cell lysate</td>
<td>EmBacY</td>
<td>66.6 kDa</td>
</tr>
<tr>
<td>11</td>
<td>Speedy A2 1-206</td>
<td>Clarified lysate</td>
<td>EmBacY</td>
<td>66.6 kDa</td>
</tr>
<tr>
<td>12</td>
<td>Speedy A2 1-206</td>
<td>Pull down</td>
<td>EmBacY</td>
<td>66.6 kDa</td>
</tr>
<tr>
<td>13</td>
<td>Speedy A2 1-189</td>
<td>Whole cell lysate</td>
<td>EmBacY</td>
<td>66.4 kDa</td>
</tr>
<tr>
<td>14</td>
<td>Speedy A2 1-189</td>
<td>Clarified lysate</td>
<td>EmBacY</td>
<td>66.4 kDa</td>
</tr>
<tr>
<td>15</td>
<td>Speedy A2 1-189</td>
<td>Pull down</td>
<td>EmBacY</td>
<td>66.4 kDa</td>
</tr>
<tr>
<td>16</td>
<td>Speedy A2 27-189</td>
<td>Whole cell lysate</td>
<td>EmBacY</td>
<td>61.5 kDa</td>
</tr>
<tr>
<td>17</td>
<td>Speedy A2 27-189</td>
<td>Clarified lysate</td>
<td>EmBacY</td>
<td>61.5 kDa</td>
</tr>
<tr>
<td>18</td>
<td>Speedy A2 27-189</td>
<td>Pull down</td>
<td>EmBacY</td>
<td>61.5 kDa</td>
</tr>
<tr>
<td>19</td>
<td>Speedy A2 27-189</td>
<td>Whole cell lysate</td>
<td>MultiBac™</td>
<td>61.5 kDa</td>
</tr>
<tr>
<td>20</td>
<td>Speedy A2 27-189</td>
<td>Clarified lysate</td>
<td>MultiBac™</td>
<td>61.5 kDa</td>
</tr>
<tr>
<td>21</td>
<td>Speedy A2 27-189</td>
<td>Pull down</td>
<td>MultiBac™</td>
<td>61.5 kDa</td>
</tr>
<tr>
<td>22</td>
<td>Speedy A2 27-206</td>
<td>Pull down</td>
<td>EmBacY</td>
<td>63.5 kDa</td>
</tr>
<tr>
<td>23</td>
<td>Speedy A2 27-206</td>
<td>Pull down</td>
<td>MultiBac™</td>
<td>63.5 kDa</td>
</tr>
</tbody>
</table>

4.4.2 Analysis of Ringo A/Speedy affinity to CDK1 and CDK2

The most successful conditions to express Speedy A were constructs containing residues 64-189 and 64-206 expressed from the pET21dMBP vector in Rosetta™ (DE3)LysS cells. Protein was bound to MBP resin and washed prior to incubating with an excess of CDK2 (0.06 mg total). After incubation, bead bound protein was analyzed by SDS-PAGE (Figure 4-22). Although the gels exhibited a few proteins that were non-specifically bound to the amylose beads, in addition to MBP and the MBP Speedy A constructs, a band consistent with the molecular weight of CDK2 appears after incubation of MBP-Speedy A (64-189) and MBP-Speedy A (64-206) with CDK2 (Figure 140).
4-22A, lanes 4 and 8, etc). To check the non-specific binding of CDK2 to the amylose resin or to MBP, a negative control was also run. Results show that CDK2 exhibits very weak non-specific binding to the amylose beads (Figure 4-22 B, lane 1). This binding is significantly weaker than in the binding experiment and no binding was observed to MBP (Figure 4-22). These results agree with binding between RINGO/Speedy and CDK2 reported in vivo (Dinarina et al., 2005), but require further confirmation by more sensitive methods.

**Figure 4-22 Investigation of binding between Speedy 64-189 and Speedy 64-206 to CDK2.**

Lanes are numbered according to the experiments described below the gel.

**4.4.3 RGC32 binding to CDK1**

Previously it has been reported that RGC32 can bind and activate CDK1 in vivo (but not other CDKs) in a manner dependant on Thr91 phosphorylation of RGC32 (Badea et al., 2002). To investigate binding in vitro, recombinant proteins were expressed in different expression systems. RGC32 and GSTRGC32 were expressed and purified in the lab of Professor M. West (Sussex University). CDK1, GSTCDK1, GSTCDK2 and CKS1 were purified in our lab as described in Materials and Methods. The purity of the proteins was verified by SDS-PAGE (Figure 4-23).
Pull down experiments were performed to test the binding of GSTCDK1 to RGC32 and GSTRGC32 to CDK1 using cyclin H as a negative control. Results of that experiment were not conclusive (results not shown) probably due to the weak affinity between the two proteins. Consequently, SPR was chosen as an alternative more sensitive method that can detect kinetics of binding in real time and can determine dissociation constants (Kd).

The interactions between GSTCDK1 and GSTCDK2 with their well-characterised partner CKS1 were studied as positive controls. Sensograms of binding were analysed by the BiaCore S200 software and results revealed tight binding of 28.9 ± 11.2 nM between GSTCDK1 and CKS1 and weaker binding of 200 ± 70.7 nM between GSTCDK2 and CKS1. Experiment has been performed twice with two different biological replicates. As can be seen from the sensograms, CKS1 binds very rapidly to both proteins although to a greater extent to GSTCDK2 in comparison to GSTCDK1. Consequently the time of association represented in the association constant is slightly shorter for GSTCDK2 in comparison to GSTCDK1 and association constant for GSTCDK1 binding to CKS1 is 0.74 M⁻¹s⁻¹x10⁵± 0.05x10⁵ in comparison to 0.56 M⁻¹s⁻¹x10⁵.
$1 \times 10^{-5} \pm 0.06 \times 10^5$ for GSTCDK2-CKS1 binding. The level of dissociation of CKS1 differs quite significantly between two proteins: $0.01 \text{s}^{-1} \pm 0.004$ for GSTCDK1-CKS1 and $0.0007 \text{s}^{-1} \pm 0.0008$ for GSTCDK2-CKS1. It can be seen from the shape of the sensorgrams that CKS1 dissociates more slowly and gradually from GSTCDK1 than from GSTCDK2 (Figure 4-24).

**Figure 4-24 Analysis of the interactions between CDK1/2 and binding partners by surface plasmon resonance (SPR).**

(A) SPR plot of response units against time to measure CKS1 binding to GSTCDK1. Different colours of sensograms represent different concentrations of analyte. Red represents 40 nM of CDK1, green – 200 nM, blue – 1000 nM, pink – 2000 nM, and turquoise – 5000 nM. (B) SPR plot of response units against time of CKS1 binding to GSTCDK2. Red represents 62.5 nM of CDK1, green – 250 nM, blue – 1000 nM, pink – 2500 nM, turquoise – 5000 nM. Measurements were performed once.

Kinetic interactions between CDK1 and RGC32 were studied in two ways. In one set of experiments, GSTCDK1 was captured on the anti GST- antibody and RGC32 was used as an analyte. In the second set, GSTRGC32 was captured and CDK1 was flowing over. Both sets of sensograms (Figure 4-25) have similar profiles characterised by slow binding (over 400 s) and slow dissociation, after a dissociation phase of 700 s most of the protein remains bound. Association constants are comparable, $0.06 \text{M}^{-1}\text{s}^{-1}$.
10^4 for CDK1 binding to GSTRGC32, and 0.03 M^{-1}s^{-1}10^4 for RGC32 binding to GSTCDK1. The calculated Kd values are almost identical respectively, 0.18 s^{-1}10^{-3} in the first case and 0.13 s^{-1}10^{-3} in the second. The kd values calculated from these experiments are similar as well demonstrating the weaker binding of 290 8 nM (CDK1 – GSTRGC32) and 448 nM (RGC32 – GSTCDK1) in comparison to interactions between GSTCDK1 and GSTCDK2 with CKS1.

<table>
<thead>
<tr>
<th>GST captured protein</th>
<th>Analyte</th>
<th>ka (M^{-1}s^{-1}10^5)</th>
<th>kd (s^{-1}10^{-4})</th>
<th>KD(nM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GSTCDK1</td>
<td>CKS1</td>
<td>0.74 ± 0.05</td>
<td>147 ± 45</td>
<td>28.9 ± 11.1</td>
</tr>
<tr>
<td>GSTCDK2</td>
<td>CKS1</td>
<td>0.57 ± 0.06</td>
<td>7 ± 8.5</td>
<td>200 ± 70.7</td>
</tr>
<tr>
<td>GSTCDK1</td>
<td>RGS32</td>
<td>0.3</td>
<td>1.3</td>
<td>448</td>
</tr>
<tr>
<td>GSTRGC32</td>
<td>CDK1</td>
<td>0.6</td>
<td>1.8</td>
<td>290</td>
</tr>
</tbody>
</table>

Table 4-6 Summary of SPR kinetic analysis.
Ka and Kd values are estimated using BiaCore S200 software. SPR to estimate binding of GSTCDK1 and GSTCDK2 to CKS1 were performed in biological duplicate. Standard deviation is calculated by Excel.
Figure 4-25 Sensograms representing kinetic interactions by SPR.

(A) SPR plot of response units against time of CDK1 binding to GSTRGC32. Different colours of sensograms represent different concentrations of analyte. Red represents 40 nM of CDK1, green – 200 nM, blue – 1000 nM, pink – 2000 nM, turquoise – 5000 nM. (B) SPR plot of response units against time of GSTCDK1 binding to RGC32. Red represents 62.5 nM of CDK1, green – 250 nM, blue – 1000 nM, pink – 2500 nM, turquoise – 5000 nM. Measurements performed once.

Cyclin H and CDK2 have not been reported to bind to RGC32 and were used as negative controls (Dinarina et al., 2005). Neither cyclin H nor CDK2 bind to GSTRGC32 in vitro (Figure 4-26).
Figure 4-26 Sensograms representing kinetic interactions by SPR.

(A) SPR plot of response units against time of cyclin H and CDK2 binding to GSTRGC32. Different colours of sensograms represent different concentrations of analyte. Red represents 8 nM cyclin H, green – 40 nM, blue – 200 nM, pink – 1000 nM, turquoise – 5000 nM. (B) SPR plot of response units against time of CDK2 binding to GSTRGC32. Representation of colours of sensograms are the same as for cyclin H and GSTRGC32 experiment. Measurements performed only once.

4.5 Conclusions

This chapter describes experiment to express the two CDK partner proteins RINGO A/Speedy and RGC32 and to investigate their binding properties. The aim of the Speedy project was to identify parts of the protein containing the speedy box which would be stable and could be expressed in sufficient amounts for biochemical, biophysical and structural characterisation. Unfortunately, despite a number of attempts to clone different protein truncations and using alternative expression systems, the expression of stable protein was not achieved. According to a recently published article (McGrath et al., 2017) a truncated version of human Speedy A containing the speedy box (residues Gly61-Asp213) can be expressed using a codon optimised sequence inserted into the pMBP vector in the SoluBL21 E. coli strain. The next step for this project would be to repeat this experiment. Further work is planned that includes CDK1-Speedy A crystallisation trials and a more systematic analysis of CDK1 and CDK2 sequence preferences around the site of phosphotransfer when bound to cyclin A, cyclin B and Speedy A.
Using purified proteins, the interaction between RGC32 and CDK1 has been confirmed. No binding was detected between RGC32 and CDK2. The results of the SPR experiments have shown that RGC32 binds to CDK1 with lower affinity than for example CKS1 to CDK1 or CDK2. Further work is planned to investigate binding using alternative assays such as homogenous time-resolved fluorescence (HTRF).

Chapter 5: CDK inhibition by ATP-competitive inhibitors

5.1 Introduction

5.1.1 Role of CDK inhibition for cancer therapy

In recent years a number of synthetic CDK inhibitors have been designed for cancer therapy. Most such inhibitors are ATP-competitive type I inhibitors that target the active, ‘DFG in’ CDK conformation, although type II inhibitors that target a non active CDK conformation, known as ‘DFG out’ and believed to be offer more opportunities for selectivity (Treiber and Shah, 2013) have started to emerge (Rzymski et al., 2015).

First generation inhibitors, pan-CDK inhibitors, were relatively non-specific and acted against most CDKs (Figure 5-1). They were designed using different scaffolds such as flavonoid, purine, aminothiazole, hymenialdisine, indenopyrasole, indirubin and paullone derivatives (Asghar et al., 2015). Some of them, such as roscovitine and flavopiridol, are quite effective for some cancers and have been used in over 60 clinical trials carried out between 1998 and 2014, although responses were limited due to high toxicity prompted by general CDK inhibition in normal tissue (Aklilu et al., 2003; Benson et al., 2007; Byrd et al., 2007; Le Tourneau et al., 2010).

A second generation of the inhibitors emerged from the development of flavopiridol and roscovitine. The design of the new generation was focused on the increase of potency and selectivity towards CDK1 and CDK2. Many inhibitors of this generation gave high expectations in a preclinical study but not many passed Phase 1 clinical trials (Payton et al., 2006; Parry et al., 2010). One of the most studied inhibitors is Dinaciclib which was designed especially for CDK1, 2, 5 and 9 inhibition, with IC\textsubscript{50} values in a range of 1-4 nM. It has demonstrated strong suppression of Rb phosphorylation in cell based assays (Parry et al., 2010). Although results of the Phase 1 clinical trials were promising (Nemunaitis et al., 2013), Phase 2 results in solid tumours came up as a
disappointment. Monotherapy of non-small cell lung cancer with Dinaciclib has shown no activity in previously treated patients (Stephenson et al., 2014). In patients with acute myeloid leukaemia or acute lymphoblastic leukaemia no objective changes were observed (Gojo et al., 2010). However, in other haematological malignancies Dinaciclib has shown evidence of clinical activity, and consequently it remains in a Phase 3 clinical trials (Blachly and Byrd, 2013). Another inhibitor, AZD5438 (developed by AstraZeneca), was reported to be poorly tolerated when administered to patients with advanced solid tumours (Boss et al., 2010). A Phase 1 study of AG-024322 (developed by Pfizer), a potent inhibitor of CDK1, 2 and 4, was also terminated because of absence of expected therapeutic effect.

<table>
<thead>
<tr>
<th>Inhibitor</th>
<th>CDK1-cyclinB</th>
<th>CDK2-cyclinA</th>
<th>CDK4-cyclinD</th>
<th>CDK6-cyclinD</th>
<th>CDK9-cyclinT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alvociclib (flavopiridol)</td>
<td>27</td>
<td>405</td>
<td>132</td>
<td>395</td>
<td>11</td>
</tr>
<tr>
<td>Seliciclib (roscovitine)</td>
<td>2100</td>
<td>100</td>
<td>13500</td>
<td>23500</td>
<td>950</td>
</tr>
<tr>
<td>Dinaciclib</td>
<td>3</td>
<td>1</td>
<td></td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>AT7519</td>
<td>190</td>
<td>44</td>
<td>67</td>
<td>660</td>
<td>&lt;100</td>
</tr>
<tr>
<td>RGB-286638</td>
<td>480</td>
<td>38</td>
<td>925</td>
<td>&gt;1000</td>
<td>4</td>
</tr>
<tr>
<td>ZK-304709</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>55</td>
<td>1</td>
</tr>
</tbody>
</table>
The third generation of CDK inhibitors came as a result of development targeting more selective inhibitors. Palbociclib was designed in 2004 and has shown a high degree of specificity towards CDK4 and CDK6, demonstrating 26-fold tighter binding to CDK4 and CDK6 (Figure 5-2) in comparison to the other CDKs (Cho et al., 2010) and additional kinases (Rae et al., 2007). Palbociclib went through many rounds of experiments and has been evaluated in several cancers with elevated CDK4 and CDK6 levels, finally reaching Phase III trial in patients with ER+/HER2-advanced breast cancer (Finn et al., 2016). It has now been approved to be used in combination with letrozole for the treatment of postmenstrual women with ER+/HER2-advanced breast cancer. Another two promising inhibitors of CDK4 and CDK6 are abemaciclib and ribociclib, which have been put into clinical trials in patients with advanced solid tumors and lymphomas (Patnaik et al., 2014; Infante et al., 2016). Most recently, ribociclib has been licensed for a similar patient population to that specified for treatment by palbociclib. All three inhibitors were crystalized with monomeric CDK6 (Chen et al., 2016), while Palbociclib has also been crystalized with a CDK6–cyclin V (Lu and Schulze-Gahmen, 2006) complex, revealing binding to both the inactive and active forms of CDK6 mainly through amino acids in the hinge region. Drug interactions with residues His100 and Thr107 have been proposed as key elements providing selectivity to the inhibitors (Chen et al., 2016).
Inhibitor | CDK1-cyclinB | CDK2-cyclinA | CDK4-cyclinD | CDK6-cyclinD | CDK9-cyclinT |
--- | --- | --- | --- | --- | --- |
Palbociclib | >10000 | >10000 | 11 | 15 | |
Abemacliclib | 1627 | 2 | 10 | 57 | |
Ribociclib | >100000 | >100000 | 10 | 39 | |

**Figure 5-2** Structure (A) and activity (B) of CDK4 and CDK6 inhibitors.
Table indicates IC50 (nM) values. Figure adopted from (Whittaker et al., 2017).

Recently CDK7 and CDK9 emerged as potential cancer drug targets. Inhibition of CDK7 has led to ablation of phosphorylation of CTD of RNA polymerase at Ser2, Ser5 and Ser7 that lead to antitumor effect in neuroblastoma, triple-negative breast cancer, small cell lung cancer shown on cell lines and in the animal models (Chipumuro et al., 2014; Christensen et al., 2014; Kwiatkowski et al., 2014). Inhibiting CDK9 in CLL cells is known to induce apoptosis and increase sensitivity to fludarabine which is used as a standard treatment for CLL (Walsby et al., 2014). At the moment there a number of CDK7 and CDK9 inhibitors have been described with the scope to develop as therapeutic targets.

The problem of many kinase inhibitors is that they display similar toxicity to cancerous and healthy tissue, leading to a difficulty to achieve a therapeutic window. For achieving therapeutic effect, the approach of combining drugs aiming at the inhibition of different targets is becoming popular for some cancers treatment. For example inhibition of CDK1 with RO3306 compound (Roche) led to sensitising BRCA proficient cancers to poly(ADP-ribose) polymerase PARP inhibition (Johnson et al., 2011). CDK1 has been shown to participate upstream of DNA damage response pathways (Ira et al., 2004) and, in CDK1 depleted cells, function of BRCA1 in S phase checkpoint was found ot be compromised (Johnson et al., 2009). BRCA1 and BRCA2 are tumour suppressor proteins that play crucial roles in DNA repair (Moynahan et al., 1999). PARP, in its turn, plays a role in single strand DNA break repair by homologous recombination. Both BRCA1 and BRCA2 are necessary for the process of homologous
recombination to go through and it is known that cells deficient in BRCA are hypersensitive to PARP inhibition (Bryant et al., 2005), which opens a new opportunity for the treatment of BRCA proficient cancers.

Because CDKs play a crucial role in the cell cycle and transcription regulation, there are sound arguments for targeting them for cancer therapy in suitably chosen patient segments. Due to them being easily accessible to the drugs, a number of successful inhibitors have been developed into clinical tools; where early compounds have been unsuccessful, they have often provided starting points for the development of better ones. The positive results where the ATP–competitive inhibitors palbociclib and ribociclib have been licensed for treatment provide encouragement for the further development of other potential drugs that target the CDK family.

5.1.2 ATP binding site

ATP is used by kinases for phosphorylation of their protein substrates and is bound in the ATP binding pocket between the N- and C-lobes (Figure 5-3). This site is also used by ATP competitive inhibitors for binding. A number of residues from different parts of molecule are involved into this process. The adenine base is positioned by the hinge sequence and forms two hydrogen bonds with backbone carbonyl and amide moieties (respectively Glu81 and Leu83 in CDK2). The roof and floor of the adenine binding cleft is lined with small hydrophobic residues that interact with the planar adenine ring. The ribose ring makes connections with a conserved aspartate (D86 in CDK2) towards the C-terminal end of the hinge. Triad of residues, Lys33, Glu51 and Asp134 are important in coordination of Mg²⁺ ions. Lys33 organises the α and β phosphates of ATP as well as positioning the αC helix by forming a salt bridge with conserved Glu51. Lys33 is crucially important as its mutation leads to kinase inactivation (Iyer et al., 2005). The glycine rich loop lies between β1 and β2 strands, and is formed by the GXGXXG motif comprised from conserved Gly12-Gly17 (in CDK2, Figure 5-3). This loop plays a role in positioning the adenine ring and accommodating the phosphates by holding the β- and γ-phosphates in place. In the absence of ATP the P-loop becomes very flexible. One more important residue which plays a role in formation of the catalytic spine is Asp127 (in CDK2) (Jura et al., 2011). One Mg²⁺ ion is crucial for efficient transfer of ATP phosphates to the substrates. The side chain of Asp145 (in CDK2), part of DFG motif, points towards the phosphate groups of ATP and is responsible for Mg²⁺ coordination. The DFG motif is located at the beginning of the activation segment and is known to exist in ‘in’ and ‘out’ states. In the ‘in’ state, which is thought to be
entropically less favourable (Shan et al., 2009), the aspartate is positioned in the active site, ensuring correct magnesium positioning required for catalysis. Consequently this position is more characteristic of an active kinase conformation.

**Figure 5-3 Conserved residues of ATP binding site in CDK1 and CDK2.**
The CDK1 and CDK2 sequences share 93% sequence identity amongst aminoacids participating in ATP binding.

**Figure 5-4 ATP binding site.**
(A) CDK1 ATP binding site (PDB: 4YC3), (B) CDK2 ATP binding site (PDB: 5if1) (C) CDK2 ATP binding site with ATP and Mg\(^{2+}\). The conserved residues in both CDKs are
identified. Activation segment is coloured red, C – helix is gold, G-loop is blue, hinge region is coral, DFG residues are marked in lilac, Glu51 in light green, Leu133 (CDK2) Leu134 (CDK1) in pink.

5.1.3 CDK inhibition, exploring the ATP-binding site

The ATP binding site is a part of a kinase molecule which participates in transfer of the terminal phosphate of ATP to a kinase substrate. To date, most described kinase inhibitors are ATP competitive, of which type I inhibitors target the active DFG-in kinase conformation and present one to three hydrogen bonds to the amino acids located in a hinge region similarly to the hydrogen bonds formed by adenine ring of ATP (Liu and Gray, 2006). ATP-competitive inhibitors generally have additional functional groups that form contacts with less conserved regions of ATP-binding pocket. Type I inhibitors can generally bind both active and inactive enzyme forms as they are normally quite small and do not require special arrangements of the αC helix, the DFG motif, or the Asp motifs. Type II inhibitors target the inactive DFG-out conformation of a kinase. Similar to inhibitors of type I, type II inhibitors form connections to the hinge region but also get engaged into DFG and C helix binding (Dar and Shokat, 2011). They also tend to bind preferentially to kinases with a small or medium ‘gatekeeper’ residue (Hasegawa et al., 2007; Kufareva and Abagyan, 2008). Type II inhibitors, being normally bigger, target the ATP binding site of kinases in a DFG-out conformation as they need to occupy not only the hinge region where adenine of ATP normally sits but also adjacent to an hydrophobic pocket (Janne et al., 2009).

Because type II inhibitors form more interactions within the ATP binding pocket than do type I inhibitors they are expected to be more specific but harder to design and develop. A further type of inhibitor, termed type III, bind allosterically to sites that do not overlap the ATP-binding site. Type I and II inhibitors represent most of the FDA-approved kinase inhibitors.

The CDK inhibitors described in the introduction to this chapter are mainly type I inhibitors, and target both active and inactive conformations of the kinase (Whittaker et al., 2017). The first successful drug that targets an inactive kinase conformation was approved by the FDA in 2001. This drug, which targets BCR-Abl-driven chronic myelogenous leukaemia and gastrointestinal stromal tumours, is Imatinib (Gleevec) and was developed by Novartis (Druker et al., 1996). The crystal structure of Imatinib with the structurally similar kinases Abl and Src revealed binding characteristic of Type
II inhibitors (Schindler et al., 2000). Remarkably, despite near identical binding pockets Imatinib binds Abl almost 1000-fold more potently then Src (Druker et al., 1996). These differences confirm the potential role of conformational flexibility in determining inhibitor selectivity. There are not many type II inhibitors designed to the date, although some interesting results have emerged in development of inhibitors of CDK8, an oncogene in colorectal cancer (Rzymski et al., 2015), from Sorafenib (Schneider et al., 2013).

**Figure 5-5 Type I and Type II kinase inhibitors**
Type I inhibitors form bonds with hinge region (PDB code: 1QCF). Type II bind also to amino acids of DFG motif and C Helix (PDB code: 1FPU) (Dar and Shokat, 2011).

For the design of ATP competitive inhibitors that target active and non-active kinase conformations, the residues that line the ATP binding pocket and structural plasticity of kinase fold are important. The kinase active site is highly flexible and can remodel to accommodate binding of a wide variety of different inhibitors that instigate changes in the position of the activation loop (Knight et al., 2006; Liu and Gray, 2006). In addition, the active site of kinases in general, and CDKs in particular, is highly conserved in primary, secondary and tertiary structure, which creates a difficulty for generating inhibitor selectivity between kinases. For example, there are 33 amino acids that participate in forming the active site in CDK1, CDK2, CDK4, CDK7, CDK9 (Figure 5-6) of which 55% are conserved across the family. If the comparison is narrowed to the closely related CDK1 and CDK2 family members, there are only two amino acids that differ, suggesting complications in designing selective inhibitors which would effectively distinguish between the two.
One proposed method to evaluate the effective differences between kinase active sites is the “fingerprinting” approach (Echalier et al., 2014). According to this approach a wide panel of inhibitors was chosen to investigate chemical diversity of CDK active sites. In this study, the melting temperature of selected CDKs (CDK2, CDK4, CDK7 and CDK9) and their complexes with cognate cyclins, was measured in the presence of different inhibitors. ΔTm (the difference between melting temperature of inhibitor bound form and non bound) was plotted to compare monomeric CDKs, phosphorylated forms and fully activated. From this analysis, it was found that CDKs in their monomeric forms significantly differ in their inhibitory profile with a lowest correlation coefficient of 0.39 (CDK4 vs CDK7) and highest correlation coefficient of 0.74 (CDK2 vs CDK7) (Figure 5-6). In the fully activated, cyclin-bound state the correlation coefficient between the different profiles are substantially higher. The lowest correlation (0.73) is measured for pCDK2-cyclin A vs pCDK9-cyclin T and the highest (0.78) for pCDK2-cyclin A vs pCDK4-cyclin D. These findings suggest that inhibitor binding properties do not only depend on the sequence of the CDK ATP-binding site but also on the kinase conformational state. The results also support the suggestion that, in an active state, two different CDKs become very similar in their inhibitor binding properties. Taken together, these studies lead to the conclusion that protein kinases in general, and CDKs in particular, are more available for selective inhibition in their monomeric, inactive state.
Figure 5.7 Comparison of the inhibitor binding profiles of different CDKs in their monomeric and cyclin bound forms.
Numbers within the block indicate internal consistency of inhibitory fingerprints for each CDK or complex measured over two or three repeats. Numbers between the blocks shows the correlation of fingerprints between CDKs or CDK-cyclin complexes (Echalier et al., 2014).

5.1.4 Aims of the project
As described in the CDK1 structural and functional study chapters, monomeric CDK1 demonstrates conformational differences from CDK2. Apparent flexibility in the CDK1-cyclin B complex has been proposed to explain functional flexibility of CDK1-cyclin B in comparison to CDK2-cyclin A. We propose that these plastic differences can result in different response to the binding of ATP-competitive inhibitors. CDK2 has recently been identified as a potential cancer target in cyclin E amplified ovarian cancer. However, CDK1, its close homologue, is essential for cell cycle progression and can generate toxic off-target effects when inhibited. Consequently, the development of selective inhibitors that can discriminate CDK2 from CDK1 would be both highly desirable and extremely challenging, since the active sites of CDK1 and CDK2 differ only by two amino acids. To explore the feasibility of selectively inhibiting CDK2, a number of inhibitors with different scaffolds were chosen to investigate the relative importance of sequence similarity versus conformational differences in binding to CDK1 versus CDK2.

To structurally characterise the active sites of CDK1 and CDK2 in their monomeric and cyclin-bound states, monomeric and binary proteins were subjected to crystallisation
trials in the presence of the panel of inhibitors. To assess and compare chemical similarity of CDK1 and CDK2 and their complexes biophysical methods including DSF, ITC, SPR and DSC were employed. To explore the influence of the inhibitors on the kinase activity ADP-Glo™ assay has been performed.

5.2 Materials and Methods

5.2.1 Protein purification

CDK1 that was phosphorylated on Thr161 (pCDK1) and non-phosphorylated CDK1 were expressed and purified as described in Sections 2.2.5 and 2.2.6 and phosphorylated in a kinase incubation as described in Section 2.2.9. Cyclin B and CKS2 were expressed and purified as described in Sections 2.2.5 and 2.2.6. pCDK1-cyclin B, CDK1-cyclin B-CKS2 and CDK1-CKS2 complexes were assembled as described in Section 2.2.11. Monomeric CDK2 was expressed in Rosetta2™(DE3)pLysS cells using pGEX6p-1 as expression vector (Table in Annex) following a well established protocol (Welburn and Endicott, 2005). Phosphorylated CDK2 (on Thr160, pCDK2) and bovine cyclin A were expressed as described in Section 2.2.4. To prepare the pCDK2-cyclin A complex, cell lysates were mixed post sonication in a culture ratio 1 : 2 pCDK2 : cyclin A to saturate the pCDK2 with excess cyclin A and then spun at 60,000 x g for 1h. The supernatant was collected and loaded onto a glutathione-sepharose column equilibrated in HBS (40mM HEPES, pH7.0, 200 mM NaCl, 0.01% MTG). The pCDK2-cyclin A complex was eluted with 20 mM glutathione in HBS and the GST tag subsequently cleaved with 3C protease in 1:50 w:w ratio, at 4°C overnight. The complex was further purified by size exclusion chromatography using a Superdex 75 26_60 column equilibrated in HBS at room temperature. As the complex elutes from the column with the GST dimer, a subsequent subtractive step on a glutathione-sepharose column to remove the GST was required, the pCDK2-cyclin A being collected in the flow-through.

5.2.2 Preparation of inhibitors

All inhibitors were dissolved in 100% DMSO to a concentration 50 mM and stored at -20°C. The inhibitors used in this study are listed in Table 5-1.
<table>
<thead>
<tr>
<th>Name</th>
<th>Scaffold</th>
<th>Chemical formulae</th>
<th>Smiles string</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flavopiridol</td>
<td>Chromone</td>
<td>C_{21}H_{20}ClNO_{5}</td>
<td>O=C1C=C(C2=CC=CC=C2Cl)OC3=C1C(O)=CC(O)=C3[C@H]4<a href="O">C@H</a>CN(C)CC4</td>
</tr>
<tr>
<td>AZD5438</td>
<td>Pyrimidamine</td>
<td>C_{18}H_{21}N_{5}O_{2}S</td>
<td>O=S(C1=CC=C(NC2=NC=CC(C3=CN=C(C)N3C(C)C)=N2)C=C1)(C)=O</td>
</tr>
<tr>
<td>CGP74514A</td>
<td>Purine</td>
<td>C_{18}H_{24}ClN_{7}</td>
<td>NC1=C2N=CN(CC)C2=NC(N([C@H]3<a href="N">C@@H</a>CCCC3)C4=CC=CC(Cl)=C4)=N1</td>
</tr>
<tr>
<td>SU9516</td>
<td>Indolinone</td>
<td>C_{13}H_{11}N_{5}O_{2}</td>
<td>O=C1NC2=C(C=NC(C)C=C2)/C1=C/C3=CN=CN3</td>
</tr>
<tr>
<td>Dinaciclib</td>
<td>Pyrazole</td>
<td>C_{21}H_{28}N_{6}O_{2}</td>
<td>OCC[C@H]1N(C2=NC3=C(CC)C=NN3C(NCC4=C<a href="%5BO-%5D">N+</a>=CC=C4)=C2)CCCC1</td>
</tr>
</tbody>
</table>

Table 5-1 ATP competitive inhibitors of CDK1 and CDK2.

5.2.3 Differential scanning fluorimetry (DSF)

Experiments were performed using a Real-Time PCR system (QS7Flex, Applied Biosystems). 3 μM of pCDK1, pCDK2, pCDK1–Cyclin B or pCDK2–Cyclin A were mixed with SyPRO Orange (5 times dilution of the stock) (Invitrogen) fluorescent probe in DMSO (Sigma Aldrich) and the eight inhibitors of interest were added to final concentrations of 250, 100, 50, 10 or 2 μM in a total volume of 25 ul. The plates were equilibrated at 25°C for 5 min followed by heating at 1°C per 90 sec from 25 to 85°C. The fluorescence intensity was recorded using the ROX filters (excitation λ=470 nm, emission λ=586nm) every 90 sec. Results were analysed by Protein Thermal Shift Software 1.3 and the inflection point was calculated using the Boltzmann method.

5.2.4 Isothermal Titration Calorimetry

ITC experiments were carried out at 30°C using a MicroCal ITC 200 (Malvern Instruments). CDK1, CDK1–cyclin B, CDK2 and CDK2–cyclin A at 10 μM were used as samples in the cell and five inhibitors were titrated as ligands in the syringe (100 μM). Titrations were performed in 40 mM HEPES pH 8.0, 500 mM NaCl, 0.25% TCEP. For each experiment, 20 injections of inhibitor were performed, 1x 0.5 μL and 19x 2 μL.
with 120 s interval between injections. The thermograms were analysed by Origin 7.0 (OriginLab Corp.) software using single binding-site model.

5.2.5 Surface plasmon resonance

Experiments were performed on a Biacore S200 (GE Healthcare) at 20°C in buffer containing 20 mM HEPES pH 7.4, 150 mM NaCl, 0.01 % Tween. A CM5 BIACore sensor chip was used for amine coupling of goat anti-GST antibody according to the GE Healthcare standard protocol. 50 μg/ml of GSTCDK1 or GSTCDK2 were immobilised on a surface of antibody at 5 μl/min for 480 s. Inhibitors were diluted 1:4 in 20mM HEPES pH 7.4, 150 mM NaCl, 0.01 % Tween buffer to give a dilution series of 50000 nM, 12500 nM, 3125 nM, 781.25nM, 195.31 nM, 12.2 nM, 3.05 nM0.76 nM, 0.19 nM and 0.05 nM using an Echo550 liquid handler (Labcyte). The inhibitors were flowed over the bound GSTCDK1 and GSTCDK2 at 30 μl/min for 60 seconds. Dissociation was measured for 360 seconds. GST was loaded at a concentration of 20 μg/ml for 60 seconds at 30 μl/min and was used as a negative control. Binding to the GST control was subtracted from inhibitor binding sensorgrams to account for non-specific interactions and bulk effects. K_d values and dissociation rates were calculated using the rate equation implemented within the Biacore analysis software.

\[ K_d = \frac{k_d}{k_a} \]

Equation 5-1 Binding constant determination.

Abbreviations: K_d – dissociations constant, k_d – dissociation rate, k_a – association rate.

5.2.6 ADP-Glo\textsuperscript{TM} kinase assay

Kinase reactions were carried out in 5 μl total volume at room temperature containing 40 mM Tris-HCl pH 7.5, 20 mM MgCl₂, 0.1 mg/ml BSA, using 4 nM pCDK1-Cyclin B or 1.5 nM CDK2-Cyclin A. Kinase activity was measured towards 25 μM peptide 21 (GPLGSMHPRVKEVRTDSGSLRRDMQPSPIKVHERWSATAGSAKKRLFGEFPPKEMLDMKW) added to the CDK-Cyclin reaction in the presence of inhibitor (Table 5-1) at concentrations 300, 240, 150, 90, 48, 24, 12, 6, 1.5, 0.75 nM. Reactions were initiated by the addition of 25 μM ATP. Assays were performed in white 384-well plates in duplicate. Luminescence was detected using a PheraStar plate reader (BMG). K_m
and $V_{\text{max}}$ values were derived using GraphPad PRISM 6 and were calculated using the Michaelis-Menten equation 3-1 as described in 3.2.15

### 5.2.7 Protein crystallography

Sitting drop vapour diffusion crystallization trials were performed initially using a homemade screen designed from an initial hit from the Morpheus screen, (Molecular Dimensions, Newmarket, Suffolk, UK) condition B4 (0.1M MES/imidazole buffer (pH 6.7), 6.5% MPD, 5% PEG 4K, 10% PEG1K). Further trials using a variety of commercially available screens, Index (Hampton Research, Aliso Viejo, USA), PACT premier™ HT-96, Proplex, Morpheus and Structure (all from Molecular Dimensions, Newmarket, UK) were also carried out. Drops were dispensed by a TPP Labtech Mosquito LCP into MRC 96 wells plates by mixing 0.2 µl of complex and 0.2 µl of mother liquor. Plates were incubated at 4°C and monitored for 2 weeks. Optimised crystallisation conditions for each complex are listed in Table 5.2. Crystals were mounted into a nylon crystal loop and cryo-cooled in liquid nitrogen prior to data collection.

<table>
<thead>
<tr>
<th>Protein</th>
<th>Inhibitors</th>
<th>Crystallisation condition</th>
</tr>
</thead>
<tbody>
<tr>
<td>CDK1-cyclin B-CKS2</td>
<td>Flavopiridol</td>
<td>MESpH6.7, PEG1K 5%, PEG4K 10%, MPD 6.5%</td>
</tr>
<tr>
<td></td>
<td>Dinaciclib</td>
<td>Poor diffraction</td>
</tr>
<tr>
<td></td>
<td>CGP74514A</td>
<td>MESpH6.7, PEG1K 10%, PEG 4K 19.5%, MPD 6.5</td>
</tr>
<tr>
<td></td>
<td>SU9516</td>
<td>Poor diffraction</td>
</tr>
<tr>
<td></td>
<td>AZD5438</td>
<td>MES pH6.3, PEG1K 5%, PEG4K 19%, MPD 18.5%</td>
</tr>
<tr>
<td>CDK1-CKS2</td>
<td>Flavopiridol</td>
<td>Poor diffraction</td>
</tr>
<tr>
<td></td>
<td>Dinaciclib</td>
<td>PACT premier TM (Molecular Dimensions), F5</td>
</tr>
<tr>
<td></td>
<td>CGP74514A</td>
<td>Morpheus (Hampton Research), D10</td>
</tr>
<tr>
<td></td>
<td>SU9516</td>
<td>Dataset is not collected</td>
</tr>
<tr>
<td></td>
<td>AZD5438</td>
<td>Morpheus (Hampton Research), F10</td>
</tr>
<tr>
<td>CDK2-cyclin A</td>
<td>Flavopiridol</td>
<td>0.1 M HEPES pH7.0, 1 M AmSO4, 0.75 M KCl</td>
</tr>
<tr>
<td></td>
<td>Dinaciclib</td>
<td>Dataset is not collected</td>
</tr>
<tr>
<td></td>
<td>CGP74514A</td>
<td>0.1 M HEPES pH7.0, 1.1 M AmSO4, 0.8 M KCl</td>
</tr>
<tr>
<td></td>
<td>SU9516</td>
<td>0.1 M HEPES pH7.0, 0.9 M AmSO4, 0.7 M KCl</td>
</tr>
<tr>
<td></td>
<td>AZD5438</td>
<td>0.1 M HEPES pH7.0, 0.6 M AmSO4, 0.9 M KCl</td>
</tr>
<tr>
<td>CDK2</td>
<td>Flavopiridol</td>
<td>Crystals were not obtained</td>
</tr>
</tbody>
</table>

160
<table>
<thead>
<tr>
<th>Inhibitor</th>
<th>PDB code</th>
<th>Growth Conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dinaciclib</td>
<td>4KD1</td>
<td>0.1 M HEPES pH 7.0, 1 M AmSO4, 0.9 M KCl</td>
</tr>
<tr>
<td>CGP74514A</td>
<td></td>
<td>0.1 M HEPES pH 7.0, 1 M AmSO4, 0.9 M KCl</td>
</tr>
<tr>
<td>SU9516</td>
<td>3PY0</td>
<td></td>
</tr>
<tr>
<td>AZD5438</td>
<td>0.1 M HEPES pH 7.0, 0.8 M AmSO4, 0.9 M KCl</td>
<td></td>
</tr>
</tbody>
</table>

Table 5-2 Crystal growth conditions.

5.3 Results and Discussion

5.3.1 Description of the inhibitors chosen for the study

The type I inhibitors that were chosen for this study have diverse scaffolds and collectively probe different parts of the ATP binding site. Though they are all potent CDK ATP-competitive inhibitors they show differences in potency towards different CDK family members (Figure 5.7).

5.3.1.1 Flavopiridol (Alvocidib)

2-(2-chlorophenyl)-5,7-dihydroxy-8-[(3S,4R)-3-hydroxy-1-methyl-4-piperidinyl]-4-chromenone is an ATP-competitive inhibitor of CDK1, CDK2, CDK4 and CDK6 with some activity also against CDK7 (Table 5.3) (Whittaker et al., 2017). Flavopiridol is an effective cytotoxic compound in human tumor models (Chipumuro et al., 2014). It was the first CDK inhibitor to be used in a clinical trial (Senderowicz, 1999), and was under development for treatment of acute myeloid leukemia and arthritis (Sekine et al., 2008) but its development was discontinued in 2012.

5.3.1.2 AZD5438

4-(1-isopropyl-2-methyl-1H-imidazol-5-yl)-N-(4-(methylsulfonyl)phenyl)pyrimidin-2-amine was developed by AstraZeneca Ltd (UK) as a potent ATP competitive inhibitor of CDK1, CDK2 and CDK9 (Byth et al., 2009). It exhibits anti-proliferative activity in a range of human cell lines including lung, colorectal, breast, prostate, and hematologic tumors (Natrajan et al., 2012).

5.3.1.3 CGP74514A

N-(cis-2-Aminocyclohexyl)-N-(3-chlorophenyl)-9-ethyl-9H-purine-2,6-diamine was developed by Novartis Pharma (Switzerland) as an ATP competitive CDK inhibitor that is cell penetrant, potent and most selective for CDK1 (Imbach et al., 1999). It has been
reported to induce apoptosis in several leukaemia cell lines (Yu et al., 2003) and to sensitize breast cancer cells to tumour necrosis factor-related apoptosis-inducing ligand (TRAIL) (Park et al., 2014).

5.3.1.4 Su9516

3-[1-(3H-imidazol-4-yl)-meth-(Z)-ylidene]-5-methoxy-1,3-dihydro-indol-2-one is described as a highly selective ATP competitive CDK inhibitor, in particular targeting CDK2 and CDK1 and, to a lesser extent, CDK4, it has been explored as a potential cancer drug (Gao et al., 2006) (Lane et al., 2001).

5.3.1.5 Dinaciclib (MK-7965, formerly SCH727965)

2-[(2S)-1-[3-ethyl-7-[(1-oxidopyridin-1-ium-3-yl)methylamino]pyrazolo[1,5-a]pyrimidin-5-yl][piperidin-2-yl]ethanol is a potent CDK2, CDK5, CDK1, CDK9 inhibitor developed by Merck & Co. Inc (Kenilworth, USA). It is active against a broad spectrum of human tumour cell lines (Parry et al., 2010) and acts in a dose dependant manner (Guzi et al., 2011). It has been in a number of clinical trials since 2012 for different types of cancer and shows promising results for haematological malignancies (Blachly and Byrd, 2013).
Figure 5-8 Structures and names of the inhibitors chosen for the study

<table>
<thead>
<tr>
<th></th>
<th>CDK1-cyclin B</th>
<th>CDK2-cyclin A</th>
<th>CDK4-cyclin D</th>
<th>CDK7-cyclin H</th>
<th>CDK9-cyclin T</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flavopiridol</td>
<td>27</td>
<td>405</td>
<td>132</td>
<td>514</td>
<td>11</td>
</tr>
<tr>
<td>AZD5438</td>
<td>16</td>
<td>3</td>
<td>449</td>
<td>821</td>
<td>20</td>
</tr>
<tr>
<td>CGP74514</td>
<td>25</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Su9516</td>
<td>40</td>
<td>22</td>
<td>200</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dinaciclib</td>
<td>3</td>
<td>1</td>
<td></td>
<td></td>
<td>4</td>
</tr>
</tbody>
</table>

Table 5-3 IC$_{50}$s of the inhibitor panel towards different CDK complexes (Byth et al., 2009; Whittaker et al., 2017).

5.3.2 Effect of inhibitor binding on CDK1 and CDK2 thermal stability.

The thermal stability of a protein in the presence of a diverse range of inhibitors can be used to characterise the inhibitor binding properties and consequently probe the physicochemical features of the active site (Najmanovich et al., 2007). Differential scanning fluorimetry (DSF) was employed to characterise changes in thermal stability of pCDK1 and pCDK2 alone and in a complex with their respective cyclins, in the presence of a diverse, small set of inhibitors (Figure 5-9). The melting temperatures of the CDKs and their complexes increase with increasing concentration of inhibitors until the CDK1 or CDK2 ATP-binding sites are saturated with inhibitor. For all inhibitor binding, saturation is reached at an inhibitor concentration of 100 μM. Consequently the T$_{m}$ value at this inhibitor concentration was taken for comparative analysis of inhibitor-bound proteins (Figure 5-9).
Figure 5-9 Comparison of melting temperatures of pCDK1, pCDK2, pCDK1-cyclin B and pCDK2-cyclin A

Melting temperature measured by thermal denaturation assay using 3 μM of phosphorylated CDK1, CDK1 – cyclin B, CDK2 and CDK2-cyclin A represented by bars of a different colour in absence and after addition of 100 μM of the inhibitors. Bars representing melting temperature measured in the absence of the inhibitor, as a control for each inhibitor separately, contoured yellow. Bars representing melting temperature of the proteins in the presence of AZD5438 compound are coloured red, Dinaciclib are coloured purple, SU9515 are coloured black, Flavopiridol are coloured green, CGP745147 are coloured dark blue.

The ΔT_m values presented in Table 5-4 were calculated as the difference between the apparent melting temperature of the protein in the presence of an inhibitor at 100 μM and the protein alone. All the inhibitors stabilised pCDK2 significantly. The highest values (13.6°C and 12.8°C) were measured in the presence of Dinaciclib and AZD5438 respectively. The lowest ΔT_m (8.4°C) was recorded in the presence of Flavopiridol. The melting temperature of pCDK2 rises by at least 6°C when it is bound to cyclin A confirming the model that CDKs are stabilised by cyclin binding (Brown et al., 2015). Moreover, pCDK2 in complex with cyclin A is stabilised further by inhibitor binding, although to a slightly lower extent then is monomeric pCDK2. As is the case for
monomeric pCDK2, Dinaciclib and AZD5438 provide the highest stabilization of pCDK2-cyclin A (Table 5-4).

pCDK1 demonstrates a very different response in comparison to pCDK2. The monomeric form of pCDK1 is already less stable than monomeric pCDK2 by 5°C. In the presence of all inhibitors apart from Flavopiridol, its melting temperature decreases, indicating destabilisation of the protein (Table 5-4, Figure 5-9). When it is bound to cyclin B, CDK1 stability increases dramatically in the presence of all five inhibitors. For example by around 3°C bound to either CGP74514A or SU9516 and by 9.6°C bound to Dinaciclib.

<table>
<thead>
<tr>
<th>Inhibitors</th>
<th>CDK2 ΔTmB</th>
<th>CDK2-cyclinA ΔTmB</th>
<th>CDK1 ΔTmB</th>
<th>CDK1-cyclinB ΔTmB</th>
</tr>
</thead>
<tbody>
<tr>
<td>AZD5438</td>
<td>12.8±0.11</td>
<td>7.9±4.83</td>
<td>-1.8±0.11</td>
<td>4.8±7.92</td>
</tr>
<tr>
<td>Dinaciclib</td>
<td>13.6±0.12</td>
<td>6.6±9.59</td>
<td>-0.4±0.12</td>
<td>9.6±6.6</td>
</tr>
<tr>
<td>SU9516</td>
<td>10.6±0.21</td>
<td>4.8±3.48</td>
<td>-0.8±0.21</td>
<td>3.5±4.76</td>
</tr>
<tr>
<td>Flavopiridol</td>
<td>8.4±0.25</td>
<td>6.1±7.44</td>
<td>1.4±0.36</td>
<td>7.4±6.15</td>
</tr>
<tr>
<td>CGP74514A</td>
<td>9.7±0.11</td>
<td>3±3.27</td>
<td>-3.5±0.11</td>
<td>3.2±3.0</td>
</tr>
</tbody>
</table>

Table 5-4 ΔTm values of pCDK2, pCDK2-cyclin A, CDK1, CDK2-cyclin B.

ΔTm is the difference between the melting temperature of the protein in the presence of the inhibitor at the concentration 100μM and that in the presence of DMSO. Values are calculated from results obtained from processing the DSF results in triplicate. Standard error of the mean calculated by Protein Thermal Shift Software 1.3.

These results demonstrate that ATP competitive inhibitors stabilise pCDK2 and pCDK2-cyclin A. Remarkably, this stabilising effect of inhibitor binding is not observed for monomeric CDK1, but does occur when CDK1 is bound to cyclin B. As expected, binding of cyclin B to CDK1 generates a complex that is more stable than is monomeric CDK1. Taken together, monomeric pCDK1 differs from pCDK2 in its inhibitor binding profile while complexes respond to inhibitor binding in a very similar manner (Figure 5-10).
Figure 5-10 Difference of ΔTm between pCDK1 and pCDK2 and pCDK1-cyclinB and pCDK2-cyclinA

5.3.3 Determining the affinity of CDK1-Cyclin B for inhibitors

In order to provide more quantitative measures of the interactions between pCDK1 and pCDK2 and the panel of inhibitors, the more sensitive techniques of ITC and SPR were attempted. ITC is a reliable method to determine the binding affinity between a protein and a small molecule. ITC experiments were completed by Daniel Woods and Mathew Martin as described in Section 5.2.6. Binding isotherms for pCDK1, pCDK1-cyclin B, pCDK2 and pCDK2-cyclin A binding to compound AZD5438 are displayed in Figure 5.11. Binding isotherms for the rest of the chosen panel of inhibitors are presented in Appendix A.1 (Figures A-1, A-2, A-3, A-4). Control titrations were also carried out to confirm that there was no heat exchange accompanying the dilution of the inhibitor into the buffer (Appendix A.1, Figure A-5). The results (Figure 5.12) show that the inhibitor Kd values vary substantially, although all inhibitors bind very weakly to monomeric pCDK1. pCDK1 binding to cyclin B significantly improves inhibitor binding. Monomeric pCDK2 has greater affinity for each member of the panel of chosen inhibitors than pCDK1, and cyclin A binding further enhances those affinities. The Kd values for inhibitor binding to CDK1-cyclin B vary considerably but are similar for pCDK2–cyclin A. Generally it can be concluded that inhibitors bind more weakly to monomeric CDKs in comparison to the cyclin-bound forms confirming the results of the DSF experiments (Figure 5-13).
Figure 5-11 ITC to measure the affinity of CDK1 and CDK2 to a set of ATP-competitive inhibitors.

Binding isotherms are shown for titrations of monomeric pCDK1 (A), monomeric pCDK2 (B) pCDK1–cyclin B (C) and pCDK2–cyclin A (D) to AZD5438. Experiments were carried out at 30°C, with 10 μM CDK or CDK–cyclin in the cell and inhibitor at 100 μM in the syringe. The measured enthalpy (ΔH) and stoichiometry of the reaction (N – number of binding sites) can be measured from which the equilibrium binding constants and changes of entropy (ΔS) can be derived.
Figure 5-12 Comparison of dissociation constant (Kd).

K_d was measured by ITC for pCDK1, pCDK1-cyclin B, pCDK2 and pCDK2 – cyclin A bound to inhibitors. Values are provided in nM.

To compare the results of DSF and ITC, values from each experiment were collated and plotted in Excel. The resulting analyses show similar trends when the relative affinity of monomeric and binary complex binding is evaluated by either ITC (Figure 5-13 A) or DSF (Figure 5-13 B): the binary complexes, pCDK2-cyclin A and CDK1-cyclin B, bind to inhibitors with similar affinities, while monomeric pCDK2 and CDK1 differ in their affinities markedly. While this general trend is preserved between DSF and ITC studies, the more subtle differences in affinity of each protein species for the panel of inhibitors are not perfectly reproduced between the two techniques.
Figure 5-13 Difference between pCDK1 and pCDK2 and their complexes measured by DSF and ITC.

Comparison of difference in melting temperature from DSF experiment (A) and $K_d$ from ITC (B) between pCDK2 and pCDK1 and pCDK2 – cyclin A and pCDK1 – cyclin B. Fold difference between pCDK2 and pCDK1 identified blue and between pCDK2 and pCDK1 – orange.

To confirm the ITC results, inhibitor binding to CDK1 and CDK2 was also assessed by SPR. GST-tagged CDK1 and CDK2 were immobilised to goat anti-GST antibody on the surface of a CM5 chip and inhibitors were subsequently flowed over to calculate the binding and dissociation kinetics. The character of GSTCDK1 and GSTCDK2 binding to chosen compounds differ substantially, displaying different dissociation constants and different kinetics (Figure 5-14, also see Appendix A.1 Figures from A-6.
to A-13). Nevertheless, subsets of inhibitor can be grouped in terms of the kinetics of their interactions with CDK2 and CDK1. For example, SU9516 and CGP 74514A (Appendix A.1, Figures A-10, A-12) bind with a fast-on fast-off kinetic to GSTCDK1, but a slow-on slow-off kinetic to GSTCDK2 (Appendix A.1, Figures A-11, A-13). This is reflected in different K\textsubscript{d}s for the different proteins: the affinities of GSTCDK2 for SU9516 and CGP 74514A are 56 nM and 312 nM respectively, while the affinities of GSTCDK1 for these inhibitors are 2 \mu M and over 2 \mu M.

The affinity of GSTCDK1 is higher to AZD5438, Dinaciclib and Flavopiridol than to SU9516 and CGP 74514A but still substantially lower than are the affinities of the same inhibitors for CDK2. According to the SPR results the GST-CDK2 binds with the highest affinity to AZD5438, Dinaciclib and SU9516. Interestingly, flavopiridol shows K\textsubscript{d} value for GSTCDK1 and GSTCDK2 that differ least. Comparison of K\textsubscript{d}s for both proteins to the panel of inhibitors is shown in figure 5-16.

**Figure 5-14 Sensogram (A) and Response-Concentration plot (B) of AZD5438 binding to GST–CDK1.**
Figure 5-15 Sensogram (A) and Response-Concentration plot (B) of AZD5438 binding to GST–CDK2.

Figure 5-16 Dissociation constant (K\text{d}) of pCDK1 and pCDK2 bound to inhibitors measured by SPR. Values are provided in nM.
Figure 5-17 Comparison of dissociation constant ($K_d$) measured by ITC and SPR.

$K_d$ is represented by the bars of the different colours, bars representing $K_d$ for CDK1 are coloured blue with red outline if measured by ITC and yellow if measured by SPR. Bars representing $K_d$ for CDK2 are coloured green, with red outline if measured by ITC and yellow if measured by SPR.

According to ITC and SPR experiments measuring affinity of selected inhibitors to CDK1 and CDK2 and their complexes with cognate cyclins, compounds show tighter binding in general to CDK2 in comparison to CDK1. Results of the ITC show that cyclin binding improves the inhibitor $K_d$ values for both CDK family members, although more significantly for CDK1 in comparison to CDK2. Consequently, the affinity of inhibitor binding to monomeric forms differs substantially between CDK1 and CDK2 but when they are cyclin bound the differences are significantly smaller. These results confirm the idea of conformational changes that happen in CDKs due to cyclin binding that stabilise the protein in a form that is better predisposed toward inhibitor binding (Echalier et al., 2014).

5.3.4 Role of inhibition of pCDK1 and pCDK2 for its activity

To investigate how structural differences and changes in thermal stability can be connected to the catalytic activity of both proteins when treated with selected inhibitors we have performed a kinase activity assay against a preferred peptide substrate derived from retinoblastoma protein p107 which has a canonical phosphorylation site, SPIK, and an intact RXL motif (Brown et al., 2015).
Table 5-5 IC<sub>50</sub> (nM) of pCDK1 and pCDK2 treated with ATP-competitive inhibitors

<table>
<thead>
<tr>
<th>Inhibitors</th>
<th>pCDK1</th>
<th>Standard deviation</th>
<th>pCDK2</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flavopiridol</td>
<td>8.84</td>
<td>3.02</td>
<td>25.67</td>
<td>4.74</td>
</tr>
<tr>
<td>Dinaciclib</td>
<td>23.83</td>
<td>9.61</td>
<td>3.53</td>
<td>0.21</td>
</tr>
<tr>
<td>CGP74514A</td>
<td>833.4333</td>
<td>561.73</td>
<td>33.22</td>
<td>8.48</td>
</tr>
<tr>
<td>SU9516</td>
<td>186.33</td>
<td>70.67</td>
<td>37.73</td>
<td>4.54</td>
</tr>
<tr>
<td>AZD5438</td>
<td>25.16</td>
<td>13.03</td>
<td>4.09</td>
<td>1.14</td>
</tr>
</tbody>
</table>

Both complexes show differences in their inhibition by the panel of compounds, although all of the inhibitors, apart from flavopiridol, appear more selective against pCDK2 in comparison to pCDK1. The unusual CDK1-selectivity of flavopiridol was not anticipated from the previous DSF and SPR experiments, although flavopiridol, in contrast to the other compounds shows very similar K<sub>d</sub> and melting temperature for both pCDK1–cyclin B and pCDK2–cyclin A complexes, suggesting similar selectivity. Also, according to the kinase assay, Dinaciclib and AZD5438 affect kinase activity of pCDK2–cyclin A the most and they are proved to be the tightest binders on ITC and also show highest melting temperature in DSF data. Interestingly, the activity of pCDK1–cyclin B and pCDK2–cyclin A in the presence of SU9516 and CGP74514 is the lowest amongst them all, consistent with the tight binding indicated by ITC and DSF as well.

5.3.5 Process of structure determination.

The biophysical and biochemical experiments suggested that there may be differences between CDK1 and CDK2 in their inhibitor binding modes. To assess these potential changes a systematic crystallisation screen was carried out to co-crystallise each inhibitor with monomeric CDK1 and CDK2 and with CDK1-Cyclin B-CKS2 and CDK2-Cyclin A. Proteins were purified and complexes of CDK1-CKS2, CDK1–cyclin B–CKS2, CDK2–cyclin A were prepared as described in Section 5.3.1 (Figure 5.18). All complexes produced crystals after a few rounds of optimisation (Figure 5.19). Normally
crystals appeared over one to two weeks in the different crystallisation conditions (see Table 5.2). However, despite several repeats, not all of them diffracted to a resolution that allowed the structure to be solved.

![SDS PAGE of purified complexes CDK1 – CKS2, CDK1 – cyclin B, CDK2 – cyclin A.](image)

**Figure 5-18 SDS PAGE of purified complexes CDK1 – CKS2, CDK1 – cyclin B, CDK2 – cyclin A.**

<table>
<thead>
<tr>
<th>Inhibitors</th>
<th>CDK1 - CKS2</th>
<th>CDK1 – B – CKS2</th>
<th>CDK2 – cyclin A</th>
</tr>
</thead>
<tbody>
<tr>
<td>AZD5438</td>
<td><img src="image" alt="AZD5438" /></td>
<td><img src="image" alt="AZD5438" /></td>
<td><img src="image" alt="AZD5438" /></td>
</tr>
<tr>
<td></td>
<td>Data collected at 2.75Å</td>
<td>Data 2.65Å</td>
<td>Data collected at 1.99Å</td>
</tr>
<tr>
<td>Dinaciclib</td>
<td><img src="image" alt="Dinaciclib" /></td>
<td><img src="image" alt="Dinaciclib" /></td>
<td><img src="image" alt="Dinaciclib" /></td>
</tr>
<tr>
<td></td>
<td>Data collected at 2.33Å</td>
<td>Data not collected</td>
<td>Data not collected</td>
</tr>
<tr>
<td>Compound</td>
<td>Data Collection Details</td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>-------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Flavopiridol</td>
<td>Data collected at 2.73Å</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Data collected at 2Å</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Data collected at 2.52Å</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SU9516</td>
<td>Data collected although ligand is not bound</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Data not collected</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Data collected at 2Å</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CGP74514A</td>
<td>Data not collected</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Data collected at 2.73Å</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Data collected at 2.65Å</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Figure 5-19 Images of crystals used for structure determination.**

Data collection was performed at the Diamond Light Source (Didcot, UK) on beamlines I04-1 and I24 at 100 K using oscillation range 0.1° for a total oscillation of 200°. For the data processing Xia2 (Winter, 2010) and programs from CCP4i2 suite (Winn et al., 2011) were used. Structures were determined by molecular replacement using Molrep (Vagin and Teplyakov, 1997) with CDK1–cyclin B–CKS2 (PDB code: 4YC3) as a search model for inhibitor bound complexes to CDK1–cyclin–CKS2, CDK1–CKS1 (PDB code: 4YC6) as a search model for inhibitor bound CDK1–CKS2 and CDK2–cyclin A (PDB code: 5IF1) for inhibitor bound CDK2 complexes. Solved structures underwent a few rounds of rebuilding in COOT (Emsley et al., 2010) and refinement by REFMAC5 (Murshudov et al., 1997), followed by final validation by MolProbity (Chen et al., 2010). Data collection, processing and statistics are provided in Appendix A.2 (Tables A-1, A-2, A-3 and A4).
5.3.6 Structural analysis of CDK1-cyclin B-CKS2 complex bound to AZD5438, Flavopiridol and CGP74514A

Structures have been solved at different resolution for different complexes of CDK1-cyclin B-CKS2: at 2.7 Å for the AZD5438 bound structure, at 2.7 Å for the CGP74514A bound structure and at 2Å for Flavopiridol bound structure. The electron density for AZD5438 (Figure 5-20 A), Flavopiridol (Figure 5-20 B) and CGP7451A (Figure 5-20 C) was visible in the ATP binding sites of CDK1 and ligands were modelled using CCP4 suites programs. Interactions between inhibitors and amino acids were analysed using CCP4mg within 5.5Å of a compound. Hydrogen bonds were assigned by CCP4mg based on the optimal distance of 3Å. Furthermore compound-aminoacid bonds were analysed by Poseview to represent key interactions in 2D (Stierand et al., 2006).

![AZD5438](image)

AZD5438

![Flavopiridol (Alvocidib)](image)

Flavopiridol (Alvocidib)

![CGP74514A](image)

CGP74514A

Figure 5-20 Structures of AZD5438, Flavopiridol (Alvocidib) and CGP74514A. Structures drawn in Chemdraw.
All three inhibitors occupy ATP binding pocket in a similar manner which is typical for Type I inhibitors. Inhibitors sit in a common plane and share a pose such that their pirimidineamine, chromone and purine groups pack towards the top of the pocket where gatekeeper residue Phe80 is located (Figure 5-21). All three inhibitors make a conserved hydrogen bond to the Leu83 of the hinge.

Figure 5-21 Overlay of AZD5438, Flavopiridol and CGP 74514A in the ATP binding pocket of CDK1 bound to cyclin B and CKS2.
AZD5438 is coloured blue, Flavopiridol is coloured green, CGP74514A coloured yellow. Residues of the CDK1 ATP binding pocket that form interactions with the inhibitors are identified.

AZD5438 and CGP 74514A form hydrogen bonds to carbonyl moiety of Leu83 of CDK1-cyclin B-CKS1 and also to the hinge backbone, although the chromone core of Flavopiridol has only one hydrogen bond to the main chain of Leu83 (Figures 5-23) and is held in a hydrophobic pocket between Leu135, Ala31 and Ile10 by van der Waals forces. The cyclohexane moiety of CGP74514A occupies the ribose binding pocket of CDK1, forming interactions with Gln132 while Flavopiridol and AZD5438 each form a network of interactions with Lys33 and Asp146 within the phosphate binding pocket through their piperidol and imidazole moieties. Towards the solvent exposed region of the active site, the methylsulfone group of AZD5438 forms a network of interactions between the two kinases lobes through Asp86 and the main chain of Ile10. Both Flavopiridol and CGP74514a harbour a chlorophenyl group in this region, however they adopt a different position. In Flavopiridol the chlorine atom is flipped in towards the inhibitor core to take advantage of interaction with Val18 and Ile10. With
CGP74514a the chlorophenyl group is twisted out, sitting up against the main chain of Met85.

Figure 5-22 ATP binding pocket of CDK1 in complex with cyclin B and CKS2 occupied by AZD5438.
(A-B) Amino acids that form interactions with the inhibitor are identified. The inhibitor is displayed in ball and stick form and coloured green, oxygen is coloured red, nitrogen is coloured dark blue, and sulphur is coloured yellow. Electron density map 2F_o-F_c around the inhibitor shown as a wire in light blue with contour at 1 sigma.(C) The surface of CDK1 active site occupied by AZD5438 is coloured by the electrostatic potential with positive in blue and negative in red (D) Poseview created two dimensional diagram. Direct bonds are represented as dashed lines and interacting residues visualised as structure diagrams. Hydrophobic and van der Waals interactions are represented as green spines with amino acids involved identified (Stierand et al., 2006)
Figure 5-23 ATP binding pocket of CDK1 in complex with cyclin B and CKS2 occupied by Flavopiridol.

(A-B) Amino acids that form interactions with the inhibitor identified. The inhibitor is displayed in ball and stick form and coloured green, oxygen is coloured red, nitrogen is coloured dark blue, and sulphur is coloured yellow. Electron density map 2Fo-Fc around the inhibitor shown as a wire in light blue with contour at 1 sigma. (C) The surface of CDK1 active site occupied by Flavopiridol coloured by the electrostatic potential with positive in red and negative in blue. (D) Poseview created two dimensional diagram. Direct bonds are represented as dashed lines and interacting residues visualised as structure diagrams. Hydrophobic and van der Waals interactions represented as green spines with amino acids involved identified (Stierand et al., 2006).
Figure 5-24 ATP binding pocket of CDK1 in complex with cyclin B and CKS2 occupied by CGP74514A.

(A-B) Amino acids that form interactions with the inhibitor identified. The inhibitor is displayed in ball and stick form and coloured green, oxygen is coloured red, nitrogen is coloured dark blue, and sulphur is coloured yellow. Electron density map 2Fo-Fc around the inhibitor shown as a wire in light blue with contour at 1 sigma. (C) The surface of CDK1 active site occupied by AZD5438 coloured by the electrostatic potential with positive in red and negative in blue (D) Poseview created two dimensional diagram. Direct bonds are represented as dashed lines and interacting residues visualised as structure diagrams. Hydrophobic and van der Waals interactions represented as green spines with amino acids involved identified (Stierand et al., 2006).

In the complexes bound to AZD5438 and Flavopiridol, the position of the P-loop follows the trend as for apo-CDK1-cyclin B-CKS2 with a phenyl ring accommodated within active site (Figure 5-25) and threonine 14 carbonyl group pointing out. However, the
position of the P-loop in CDK1-cyclin B-CKS2 bound to CGP74514A is different (Figure 5-25), with the phenyl ring of Tyr15 pointing out of the active site and helping coordination of Glu163 on the activation segment. This conformation of the P-loop is also apparent in apo monomeric CDK1.

![Figure 5-25 Changes in P-loop position.](image)

CDK1 in complex with cyclin B and CKS2 bound to AZD5438 compound is coloured ice blue, CDK1 bound to Flavopiridol is coloured cyan, CDK1 bound to CGP74511A is coloured pink.

### 5.3.7 Structural analysis of CDK1-CKS2 complex bound to AZD5438 and Dinaciclib

Two structures of monomeric CDK1 have been solved bound to AZD5438 and Dinaciclib. Overall, the structure of CDK1 bound to Dinaciclib is very similar to the structure of the apo CDK1-CKS2 complex, apart from a small movement of the P-loop (about 2Å) towards the ATP binding site and a movement of the main chain of the hinge region towards the active site (about 1.5Å). The activation segment does not have supporting electron density in either structure. Dinaciclib is reportedly the most potent monomeric CDK1 inhibitor (Parry et al., 2010). It binds to CDK1 in a similar manner to the previously reported structures in complex with monomeric CDK2 and fully active complex of phosphorylated CDK2-cyclin E (Martin et al., 2013; Chen et al., 2016). The aminopyrazolopyrimidine moiety of Dinaciclib forms two hydrogen bonds with the main chain Leu83 within the hinge region. The ethyl group adjoined to the pyrazolo core is positioned up against Phe80, while the core is sandwiched between the two lobes through hydrophobic and van der Waals interactions with Ala31, Ile10, and Leu135. The piperidine ring adopts the chair conformation and the pendent 2-hydroxyethyl group of the piperidine ring interacts with the conserved Lys33 residue and potentially forms an intramolecular hydrogen bond with the pyrazolo core. The pyridine oxide
group is solvent exposed due to its position in front of binding pocket, stacking up against the main chain of Met85 such that it seems to form interactions with Lys89 (Figure 5-26).

Figure 5-26 ATP binding pocket of CDK1 in complex with CKS2 occupied by Dinaciclib.

(A-B) Amino acids that form interactions with the inhibitor identified. The inhibitor is displayed in ball and stick form and coloured green, oxygen is coloured red, nitrogen is coloured dark blue, and sulphur is coloured yellow. Electron density map 2Fo-Fc around the inhibitor shown as a wire in light blue with contour at 1 sigma. (C) The surface of CDK1 active site occupied by Dinaciclib is coloured by the electrostatic potential with positive in red and negative in blue (D) Poseview created two dimensional diagram. Direct bonds are represented as dashed lines and interacting residues visualised as structure diagrams. Hydrophobic and van der Waals interactions represented as green spines with amino acids involved identified (Stierand et al., 2006)
Overall, the structure of CDK1 bound to AZD5438 is almost identical to the apo-monomeric CDK1 structure, with a similar position of a P-loop and the main chain of the hinge region (Figure 5-27). Due to the relatively small nature of AZD5438, this inhibitor does not extend enough towards the P loop to form the hydrogen bonds in that region. Instead, it forms connections along the hinge region. However, the AZD5438 pirimidineamine scaffold forms two hydrogen bonds with the main chain of Leu83 and the methylsulfonyl group makes additional hydrogen bonds towards the solvent exposed area. The AZD5438 imidazole moiety is involved in making hydrophobic and van der Waals connections to amino acids in the phosphate binding pocket.

The position of the P-loop of CDK1 bound to Dinaciclib is pulled in towards the activation pocket with the phenyl ring of Tyr15 pointing out, as is the case when CDK1 is bound to CGP74514A and apo monomeric CDK1. However, the P-loop of monomeric CDK1 bound to Dinaciclib is moved even further, by about 2Å, into the active site in comparison with apo monomeric CDK1. The conformation of the P-loop of CDK1 bound to AZD5438 resembles closely that of CDK1 in complex with cyclin B and CKS2, with the phenylalanine ring being tucked into the ATP binding site and the P loop pushed out. Although it has been hypothesised that the P-loop position towards the active site leads to inhibitors being more potent (Patel and Doerksen, 2010), that hypothesis is not supported by the structures reported here: CDK1 binds to two inhibitors of substantially different potency (Dinaciclib and CGP74514A) with P-loops in very similar conformations, although in the case of Dinaciclib the structure has been solved bound to the monomeric form and in case of CGP74514A to the CDK1-cyclin B complex. The structures of CDK1-cyclin B bound to either inhibitor were not obtained because crystals were too small.
Figure 5-27 ATP binding pocket of CDK1 in complex with CKS2 occupied by AZD5438.

(A-B) Amino acids that form interactions with the inhibitor identified. The inhibitor is displayed in ball and stick form and coloured green, oxygen is coloured red, nitrogen is coloured dark blue, and sulphur is coloured yellow. Electron density map $2F_o-F_c$ around the inhibitor shown as a wire in light blue with contour at 1 sigma. (C) The surface of CDK1 active site occupied by AZD5438 coloured by the electrostatic potential with positive in red and negative in blue. (D) Poseview created two dimensional diagram. Direct bonds are represented as dashed lines and interacting residues visualised as structure diagrams. Hydrophobic and van der Waals interactions represented as green spines with amino acids involved identified (Stierand et al., 2006).
Monomeric CDK1 bound to AZD5438 is coloured sea green, monomeric CDK1 bound to Dinaciclib is coloured gold. The distance between the respective αC atoms of Tyr15 of each of the P-loops is measured by CCP4mg to be 5.3 Å.

**5.3.8 Comparison of ATP binding pocket of monomeric CDK1 and CDK1 in complex with cyclin B**

This study was aiming to investigate how inhibitors with different scaffolds mould the ATP-binding sites of CDK1 and CDK2. Because of the differing plasticity of monomeric and cyclin bound CDKs, one objective of the study was to compare protein conformations and ligand poses when a specific inhibitor is bound to two different CDKs in their monomeric and cyclin-bound states. After a lot of effort, this objective was achieved for the inhibitor AZD5438, for which structures were determined in complex with CDK1-CKS2, CDK1-cyclin B-CKS2 and CDK2, CDK2-cyclin A was solved.

Overall structures of monomeric CDK1 and CDK1 in complex with cyclin B bound to AZD5438 are similar, apart from some differences which are mainly related to cyclin B binding (Figure 5-29). As expected, cyclin B causes the CDK1 activation segment to become more ordered in the complex and “unwinds” the small α helix at the beginning thereof. Similarly, in the cyclin B-containing complex the C helix is moved towards the ATP binding pocket and the P-loop is better ordered. The main chain of the P-loop is tacked in towards the ATP binding site by 3.7Å. The main chain of the hinge region follows the same path in the monomeric and binary structures. The compound overlays very well in both structures and forms similar interactions.
Figure 5-29 Comparison of ATP binding pocket of CDK1 and CDK1-cyclin B occupied by AZD5438.

(A) Overlay of CDK1 (ice blue) and CDK1-cyclin B-CKS2 (green) bound to AZD5438. (B) Amino acids from ATP binding pocket of CDK1-CKS2 that form interactions with the inhibitor. Inhibitor is displayed in ball and stick form and coloured green, oxygen is coloured red, nitrogen is coloured dark blue, and sulphur is coloured yellow. Electron density map 2Fo-Fc around the inhibitor shown as a wire in light blue with contour at 1 sigma. (C) ATP binding pocket of CDK1-cyclin B-CKS2.

Comparison of the AZD5438 complexes of monomeric CDK2 and CDK2 in complex with cyclin A follows the same trend as for CDK1 and CDK1-cyclin B (Figure 5-30). The ATP binding site forms similar interactions with AZD5438. The P-loop, however, displays larger differences between the monomeric and binary structures, moving by 4.8 Å depending on whether cyclin A is present or not.
Figure 5-30 Comparison of ATP binding pocket of CDK2 and CDK2-cyclin A occupied by AZD5438.

(A) Overlay of CDK2 and CDK2-cyclin A bound to AZD5438. (B) Amino acids from ATP binding pocket of CDK2 that form interactions with the inhibitor. Inhibitor is displayed in ball and stick form and coloured green, oxygen is coloured red, nitrogen is coloured dark blue, and sulphur is coloured yellow. Electron density map $2F_o-F_c$ around the inhibitor shown as a wire in light blue with contour 1 sigma. (C) ATP binding pocket of CDK2-cyclin A.

From this comparative analysis, it can be concluded that the inhibitor bound structures of CDK1 and CDK2 are very similar in their ATP binding site in the monomeric form as well as in cyclin bound in terms of interactions made (Figure 5-31). Furthermore comparison of CDK1 and CDK2 structures reveal very close similarity. Consequently
the differences in biochemical properties of two must be deriving from factors other than the specific interactions made by the inhibitors in the ATP binding site.

Figure 5-31 Overlay of CDK1 and CDK2 structures bound to AZD5438. CDK1 is displayed in ice blue and CDK2 is displayed in gold. AZD5438 compound represented in form of sticks and balls. Compound bound to CDK1 is coloured ice blue, compound bound to CDK2 is coloured gold.

5.4 Conclusions

The work produced for this chapter revealed very exciting results. The comparison of melting temperatures of proteins bound to different inhibitors has shown not only the different level of stabilisation provided by different inhibitors, but also i) that cyclin-binding enhances inhibitor binding, presumably by stabilising the CDK in conformation(s) compatible with optimising the inhibitor’s interactions, and ii) that monomeric CDK1 and CDK2 differ markedly in the affinity with which they are able to bind to inhibitors. ITC experiments measuring affinity between proteins and compounds have confirmed the preference of the inhibitors evaluated here to bind to CDK2 rather than CDK1. They have also confirmed and quantified the enhancement of inhibitor binding affinity in a cyclin bound form. These results also agree with published data that monomeric CDKs differ more in their inhibitor binding profile than
do CDKs when in complex with their cognate cyclins. We have also observed not only different affinity to inhibitors between CDK1 and CDK2 but also different kinetics of binding and dissociation by using Surface Plasmon Resonance. X-ray structure determination revealed only minor differences in compound binding between CDK1 and CDK2 as well as between CDK1 and CDK1 – cyclin B and CDK2 and CDK2 – cyclin A. This absence of significant differences in the bound structures supports the hypothesis that differences in affinity derive, instead, from differences in the conformational landscape of the whole CDK molecule in the bound and unbound states, rather than from the detailed geometries of the contacts made in the bound state.

Therefore, the conclusion can be drawn that ATP binding sites can be explored by ATP competitive inhibitors to reveal both quantitatively and qualitatively different properties. Furthermore, we conclude that inhibitors bind much more weakly to monomeric CDK1 in comparison to CDK2, while that differential interaction is largely removed by cyclin binding. Consequently, cyclin bound complexes do not differ so much in their ability for inhibitor binding. Therefore, it may be possible to achieve selective inhibition of CDK2 by targeting the monomeric form of the enzyme, exploiting conformational flexibility of the whole kinase molecules and not only that of the ATP binding site.
Chapter 6: Conclusions and further directions

6.1 Conclusions

CDK1 is the founding member of the cyclin dependent kinase family and the only essential CDK, which can substitute for the other interphase CDKs to ensure progression of the eukaryotic cell cycle. It has been proposed that this ability of CDK1 to step in for another CDKs is driven, in part, by its flexibility in substrate selection. A few structures of CDKs have been solved in recent years to assist with understanding of their function as well as help with the design of drugs for use in disease therapy. The crystal structure of CDK1 has been reported in this study in a monomeric state as well as bound to its activator cyclin B. Monomeric CDK1 adopts a typical bi-lobal inactive CDK fold where the C-helix is moved out of the active site and the activation segment is disordered. The main dissimilarities between monomeric forms of CDK1 and CDK2 can be noticed in the area of activation segment, αC helix and C-termini. The activation segments of both CDK2 and CDK1 begin with a small α helical structure that occludes the site occupied by the C-helix in the activated, cyclin-bound states of the enzymes. Residues 156-162 that encompass the site of phosphorylation within the activation segment of CDK1 do not have supporting electron density, suggesting that these residues are likely to be flexible and/or disordered in the monomeric state. A major difference between monomeric CDK1 and monomeric CDK2 can be observed in the conformation of the DFG motif, with phenylalanine side chain in CDK1 pointing out towards carbonyl moiety of Val64, while in CDK2 this side chain is rotated into the active site towards Asp185. A further structural feature specific to monomeric CDK1 is a small β hairpin formed by residues 40-46 that form the loop that preceded the C-helix. The C terminus of CDK1, beginning at Leu290, is disordered but in CDK2 it forms the tail which spreads at the back of the structure towards the hinge.

Overall, the structure of CDK1-cyclin B-CKS2 appeared to be distinctive from both the CDK2-cyclin A and CDK4-cyclin D structures, with a 30% smaller binding interface between CDK and cyclin in comparison to CDK2-cyclin A which leaves the activation segment more exposed to solvent. An interesting feature of the CDK1-cyclin B structure, that is shared with CDK4-cyclin D and CDK9-cyclin T but not present in CDK2-cyclin A, is a cleft between the C terminal lobe of CDK1 and cyclin B. The equivalent cleft in CDK9-cyclin T structure has been shown to play a role in binding of HIV regulatory protein TAT. By analogy, the cleft in CDK1 may play a role in mediating
interaction with as yet unidentified partners, a hypothesis which might be investigated in the future. The structure of CDK1-cyclin B-CKS2 was found not to be in a fully active conformation, although with activation segment being relatively well ordered. Comparison of positions of Thr 161 of CDK1-cyclin B-CKS2 and Thr of CDK2-cyclin A shows 5 Å distance between them in the superimposed structures. The change of position in two complexes suggests that one or both have to move for engaging CDK7-cyclin H for phosphorylation productively.

The structural findings prompted a series of experiments directed towards determination of the affinity between CDK1 and cyclin B in comparison to CDK2 and cyclin A by Isothermal Titration Calorimetry as well as the stability of the two complexes by Differential Scanning Fluorimetry. The results came into agreement with the structural outcomes, showing lower stability by about 5°C of the CDK1-cyclin B complex in comparison to CDK2-cyclin A.

It has been demonstrated in the past that the Thr160 phosphorylation site of CDK2 is buried in the pCDK2-cyclin A complex and not accessible to dephosphorylation by λ phosphatase. As expected from the more open character of the CDK1-cyclin B complex, we found that the equivalent phosphorylated residue in the activation segment of CDK1 remains accessible to λ phosphatase-mediated dephosphorylation when in complex with cyclin B.

The comparison of residues involved in CDK-cyclin binding between CDK1 and cyclin A, cyclin B and cyclin E has offered a possible explanation for the preferential binding of CDK1 to cyclin B and cyclin A. It has been noticed that the smaller residues Asn112, Ile 119 and Leu202 of cyclin E cannot contribute as much binding energy as the larger hydrophobic residues Tyr170, Tyr177 and Tyr258 in cyclin B.

Experiments focused on investigating substrate preferences have confirmed a prediction made from the structure, namely that CDK1-cyclin B displays a more relaxed specificity towards amino acids around the site of phosphotransfer in comparison to CDK2-cyclin A.

Another direction was taken to describe the binding properties of two CDK partner proteins RINGO/Speedy A and RGC32. The aim of the Speedy project was to express and purify a protein fragment containing a speedy box in amounts sufficient for biochemical, biophysical and structural characterisations. Despite extensive efforts,
protein was not expressed in the required amounts. Subsequently, the structure of a Speedy fragment in complex with CDK2 has been published recently.

In agreement with cell-based experiments performed in the lab of Prof. Michelle West we have been able to demonstrate the binding between RGC32 and CDK1 by SPR. No binding has been detected between RGC32 and CDK2.

The last chapter presents exciting work to dissect structural and functional differences between CDK1 and CDK2 in respect of their binding to a panel of inhibitors. Inhibitors were chosen to span diverse different scaffolds so as to probe the ATP binding site in different ways. In this collaborative study a number of biophysical and biochemical techniques were performed to investigate i) the inhibitors' influence on stability of monomeric proteins and complexes by DSF, ii) the affinity of proteins towards the inhibitor panel by SPR and ITC, iii) the influence of the inhibitors on the kinase activity of CDK-cyclin complexes by ADP-Glo™ assay, and iv) similarities and differences in inhibitor binding modes by X-ray crystallography. Results revealed that, despite relatively minor differences in binding interactions demonstrated by crystallography, inhibitors show markedly different affinities and kinetics of binding to the monomeric form of the CDKs, but tight and consistent binding to the cyclin-bound forms. All these findings lead to the conclusion that ATP-competitive inhibitors might exploit the conformational energetic landscape of the monomeric forms of CDKs to achieve potentially exquisitely selective inhibition.

6.2 Future directions

Some of the work attempted during the last few years has yielded promising leads that open up new lines of experimentation. The conclusions from it could lead to future possibilities to crystallise CDK1 in a fully active conformation, phosphorylated on Tyr 161 which would give more understanding of its function in comparison to the other CDKs. Also, determination of the structure of CDK1 bound to cyclin A could be beneficial for understanding the impact of the cyclin partner on the choice of substrates by CDK1. Furthermore work is planned to investigate the binding of CDK1 to non-cyclin regulators as Ringo /Speedy A and perform CDK1-Speedy A crystallisation trials. Regarding RGC32 protein, alternative assays such as homogenous time resolved fluorescence (HTRF) can be performed to take the project forward. Finally, the work
presented here clearly demonstrates that whereas inhibitors struggle to discriminate between CDK2-cyclin A and CDK1-cyclin B, they can readily distinguish monomeric CDK2 from monomeric CDK1. This observation opens up exciting possibilities for future development of inhibitors that target one or the other family member, which might in turn be used as pharmacological probes for their cellular roles and as starting points for drug discovery.
Appendix

A.1 Additional results

A.1.1 Results of ITC

Binding isotherms of pCDK1, pCDK1-cyclin B, pCDK2 and pCDK2-cyclin A to Dinaciclib, SU9516, Flavopiridol, CGP7451.

Figure A-1 ITC to measure the affinity of CDK1 and CDK2 to a set of ATP-competitive inhibitors. Binding isotherms are shown for titrations of monomeric pCDK1 (A), monomeric pCDK2 (B) pCDK1–cyclin B (C) and pCDK2–cyclin A (D) to Dinaciclib.
Experiments were carried out at 30°C, with 10 μM CDK or CDK-cyclin in the cell and inhibitor at 100 μM in the syringe. The measured enthalpy (ΔH) and stoichiometry of the reaction (N – number of binding sites) can be measured from which the equilibrium binding constants and changes of entropy (ΔS) can be derived.

Figure A-2 ITC to measure the affinity of CDK1 and CDK2 to a set of ATP-competitive inhibitors. Binding isotherms are shown for titrations of monomeric pCDK1 (A), monomeric pCDK2 (B) pCDK1–cyclin B (C) and pCDK2–cyclin A (D) to SU9516. Experiments were carried out at 30°C, with 10 μM CDK or CDK-cyclin in the cell and inhibitor at 100 μM in the syringe. The measured enthalpy (ΔH) and stoichiometry of the reaction (N – number of binding sites) can be measured from which the equilibrium binding constants and changes of entropy (ΔS) can be derived.
Figure A-3 ITC to measure the affinity of CDK1 and CDK2 to a set of ATP-competitive inhibitors. Binding isotherms are shown for titrations of monomeric pCDK1 (A), monomeric pCDK2 (B) pCDK1–cyclin B (C) and pCDK2–cyclin A (D) to Flavopiridol. Experiments were carried out at 30°C, with 10 μM CDK or CDK-cyclin in the cell and inhibitor at 100 μM in the syringe. The measured enthalpy (ΔH) and stoichiometry of the reaction (N – number of binding sites) can be measured from which the equilibrium binding constants and changes of entropy (ΔS) can be derived.
Figure A-4 ITC to measure the affinity of CDK1 and CDK2 to a set of ATP-competitive inhibitors. Binding isotherms are shown for titrations of monomeric pCDK1 (A), monomeric pCDK2 (B) pCDK1–cyclin B (C) and pCDK2–cyclin A (D) to CGP74514A. Experiments were carried out at 30°C, with 10 μM CDK or CDK–cyclin in the cell and inhibitor at 100 μM in the syringe. The measured enthalpy (ΔH) and stoichiometry of the reaction (N – number of binding sites) can be measured from which the equilibrium binding constants and changes of entropy (ΔS) can be derived.
Figure A-5 Binding isotherms of compound in 1% DMSO injected into buffer.
A.1.2 SPR sensograms

Figure A-6 Sensogram (A) and Response-Concentration plot (B) of Dinaciclib binding to GSTCDK1.
Figure A-7 Sensogram (A) and Response-Concentration plot (B) of Dinaciclib binding to GSTCDK2.
Figure A-8 Sensogram (A) and Response-Concentration plot (B) of Flavopiridol binding to GSTCDK1.
Figure A-9 Sensogram (A) and Response-Concentration plot (B) of Flavopiridol binding to GSTCDK2.
Figure A-10 Sensogram (A) and Response-Concentration plot (B) of Su9516 binding to GSTCDK1.
Figure A-11 Sensogram (A) and Response-Concentration plot (B) of Su9516 binding to GSTCDK2.
Figure A-12 Sensogram (A) and Response-Concentration plot (B) of CGP74514A binding to GSTCDK1.
A.2 Data collection and phasing statistics

Datasets dls150516_CDK1_15_AZD5438 for CDK1-cyclin B-CKS2 AZD5438, dls250616_CDK1B_1_FLP for CDK1-cyclin B-CKS2 Flavopiridol and dls150516_CDK1B_11_CGP74514A CDK1-cyclin B-CKS2 CGP74514A were collected from single crystals on beamlines I03, I04 and I24 respectively, operating at 0.969 Å and equipped with Pilatus 6M detectors (Broennimann et al., 2006). Typically 2000 images were collected with an oscillation angle of 0.1° per image and an exposure time of 0.1 s. Data were imported from Xia (Winter, 2010) and processed
using programs of the CCP4 suite (Pointless/ Aimless (Evans, 2006; Evans, 2011)). The structures of CKS2 (PDB code 1BUH), CDK2 (PDB code 1HCK), human cyclin B (PDB code 2B9R) and cyclin-bound CDK2 (PDB code 1QMZ) were used as search models for molecular replacement by Phaser (McCoy et al., 2007). The structures were refined using cycles of manual rebuilding in COOT (Emsley et al., 2010) followed by refinement by Refmac5 (Murshudov et al., 1997). The final models were evaluated using Molprobity (Chen et al., 2010). Statistics for each data set are presented in Table A1.

<table>
<thead>
<tr>
<th></th>
<th>CDK1-cyclinB-CKS2 AZD5438</th>
<th>CDK1-CyclinB-CKS2 Flavopiridol</th>
<th>CDK1-CyclinB-CKS2 CGP74514A</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data collection</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Beamline</td>
<td>I03</td>
<td>I04</td>
<td>I03</td>
</tr>
<tr>
<td>Space group</td>
<td>P 21 21 21</td>
<td>1 1 2 1</td>
<td>P 21 21 21</td>
</tr>
<tr>
<td><strong>Cell dimensions</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a, b, c (Å)</td>
<td>64.6, 68.2, 167.3</td>
<td>78.1, 53.8, 180.2</td>
<td>64.5, 68.2, 166.1</td>
</tr>
<tr>
<td>α, β, γ (°)</td>
<td>90, 90, 90</td>
<td>90.0, 98.3, 90.0</td>
<td>90, 90, 90</td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>83.80-2.65 (2.65-2.65)</td>
<td>89.17-2.00 (2.0-2.0)</td>
<td>83.18 (2.73-2.73)</td>
</tr>
<tr>
<td>R-factor/R-free</td>
<td>0.194 /0.258</td>
<td>0.189 /0.240</td>
<td>0.197 /0.252</td>
</tr>
<tr>
<td>Rmerge</td>
<td>0.122 (0.824)</td>
<td>0.087 (1.012)</td>
<td>0.267 (1.458)</td>
</tr>
<tr>
<td>I/δ(i)</td>
<td>9 (1.9)</td>
<td>6.9 (1.1)</td>
<td>5.2 (1.3)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>98.2 (95.3)</td>
<td>100 (100)</td>
<td>100 (100)</td>
</tr>
<tr>
<td>Redundancy (%)</td>
<td>3.4 (3.4)</td>
<td>3.7 (3.7)</td>
<td>6.7 (5.5)</td>
</tr>
<tr>
<td><strong>Refinement</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. reflections</td>
<td>21743 /1128</td>
<td>50412 /2485</td>
<td>20165 /989</td>
</tr>
<tr>
<td>R_work/R_free</td>
<td>0.194 /0.258</td>
<td>0.189 /0.240</td>
<td>0.197 /0.252</td>
</tr>
<tr>
<td><strong>B-factors</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein (Chains A, B,C)</td>
<td>48.2, 42.7, 55.6</td>
<td>45.1, 51.4, 57.1</td>
<td>54.4, 47.7, 65.4</td>
</tr>
<tr>
<td>Ligand/ion</td>
<td>39.2</td>
<td>40.3</td>
<td>47.7</td>
</tr>
<tr>
<td>Water</td>
<td>33.8</td>
<td>47.4</td>
<td>37.1</td>
</tr>
<tr>
<td><strong>No. atoms</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein (Chains A, B,C)</td>
<td>4651, 4216, 1260</td>
<td>4699, 4226, 1239</td>
<td>4651, 4216, 1260</td>
</tr>
<tr>
<td>Ligand/ion</td>
<td>48</td>
<td>45</td>
<td>52</td>
</tr>
<tr>
<td>Water</td>
<td>62</td>
<td>194</td>
<td>76</td>
</tr>
<tr>
<td><strong>Root mean-squared deviations</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bond lengths (Å)</td>
<td>0.0115</td>
<td>0.0163</td>
<td>0.0109</td>
</tr>
<tr>
<td>Bond angles (°)</td>
<td>1.552</td>
<td>1.839</td>
<td>1.534</td>
</tr>
</tbody>
</table>

Table A1 Summary of CDK1-cyclin B-CKS2 bound to AZD5438, Flavopiridol, CGP74514A X-ray data collection refinement statistics.

CDK1-CKS2 AZD5438 data collected as described earlier on a single crystal at 100 K on a beamline I04-1, DLS, dataset called dls080517_CDK1_18_AZD. CDK1-cyclin B-CKS2 Dinaciclib, data collected on a single crystal at 100 K on a beamline I24, DLS, dataset called dls240916_CDK1_6_dinaciclib.
<table>
<thead>
<tr>
<th></th>
<th>CDK1-CKS2 AZD5438</th>
<th>CDK1-CKS2 Dinaciclib</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data collection</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Beamline</td>
<td>I04-1</td>
<td>I24</td>
</tr>
<tr>
<td>Space group</td>
<td>P 1 2 1 1</td>
<td>I 2 2 2</td>
</tr>
<tr>
<td><strong>Cell dimensions</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a, b, c (Å)</td>
<td>68.0, 149.2, 87.3</td>
<td>93.7, 97.2, 108.6</td>
</tr>
<tr>
<td>α, β, γ (°)</td>
<td>90, 90, 90</td>
<td>90, 90, 90</td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>87.35-2.75 (2.75-2.75)</td>
<td>72.43-2.33 (2.33-2.33)</td>
</tr>
<tr>
<td>Rmerge</td>
<td>0.091 (0.842)</td>
<td>0.052 (0.42)</td>
</tr>
<tr>
<td>I/σ(I)</td>
<td>8.8 (1.4)</td>
<td>10.8 (2.4)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>99.8</td>
<td>99.8</td>
</tr>
<tr>
<td>Redundancy</td>
<td>3.8 (3.8)</td>
<td>1.9 (1.9)</td>
</tr>
<tr>
<td><strong>Refinement</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. reflections</td>
<td>45022 /2187</td>
<td>21580 /1077</td>
</tr>
<tr>
<td>R_work/R_free</td>
<td>0.225 /0.271</td>
<td>0.199 /0.250</td>
</tr>
<tr>
<td><strong>B-factors</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein (Chains A, B)</td>
<td>74.0, 72.1</td>
<td>47, 56</td>
</tr>
<tr>
<td>Ligand/ion</td>
<td>83.8</td>
<td>47.7</td>
</tr>
<tr>
<td>Water</td>
<td>52.2</td>
<td>40.7</td>
</tr>
<tr>
<td><strong>No. atoms</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein (Chains A, B)</td>
<td>4536, 1186</td>
<td>4586, 1164</td>
</tr>
<tr>
<td>Ligand/ion</td>
<td>47</td>
<td>56</td>
</tr>
<tr>
<td>Water</td>
<td>14</td>
<td>86</td>
</tr>
<tr>
<td><strong>Root mean-squared deviations</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bond lengths (Å)</td>
<td>0.0117</td>
<td>0.0158</td>
</tr>
<tr>
<td>Bond angles (°)</td>
<td>1.601</td>
<td>1.887</td>
</tr>
</tbody>
</table>

**Table A2 Summary of CDK1-CKS2 bound to AZD5438 and Dinaciclib X-ray data collection refinement statistics.**

CDK2-cyclin A AZD5438 data collected on a single crystal at 100 K on a beamline I03, DLS, dataset called dls250616_CDK2A_2_AZD. CDK2-cyclin A- Flavopiridol, data collected on a single crystal at 100 K on a beamline I04-1, DLS, dataset called dls100716_CDK2A_1_flavopiridol. CDK2-cyclin A Su9516 data collected on a single crystal at 100 K on a beamline I04-1, DLS and dataset called dls100716_CDK2A_5_SU9516.
<table>
<thead>
<tr>
<th></th>
<th>CDK2-cyclinA AZD5438</th>
<th>CDK2-CyclinA Flavopiridol</th>
<th>CDK2-CyclinA Su9516</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data collection</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Beamline</td>
<td>I03</td>
<td>I04-1</td>
<td>I04-1</td>
</tr>
<tr>
<td>Space group</td>
<td>P 21 21 21</td>
<td>P 21 21 21</td>
<td>P 21 21 21</td>
</tr>
<tr>
<td>Cell dimensions</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a, b, c (Å)</td>
<td>74.3,133.2,148.1</td>
<td>75.2,136.2,150.0</td>
<td>75.4,135.8,149.6</td>
</tr>
<tr>
<td>α, β, γ (°)</td>
<td>90.90,90.90</td>
<td>90.90,90.90</td>
<td>90.90,90.90</td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>99.26-1.99</td>
<td>101.04-2.52</td>
<td>100.77-2.00</td>
</tr>
<tr>
<td>Rmerge</td>
<td>0.099 (1.252)</td>
<td>0.275 (2.581)</td>
<td>0.085 (0.984)</td>
</tr>
<tr>
<td>I/σ(I)</td>
<td>12.1 (1.2)</td>
<td>8.5 (1.5)</td>
<td>12.1 (2.1)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>98.6 (93.2)</td>
<td>100 (100)</td>
<td>100 (100)</td>
</tr>
<tr>
<td>Redundancy</td>
<td>6.5 (4.6)</td>
<td>7.5 (7.7)</td>
<td>7.3 (7.3)</td>
</tr>
<tr>
<td><strong>Refinement</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. reflections</td>
<td>99925 / 4909</td>
<td>52814 /2638</td>
<td>104286 / 5195</td>
</tr>
<tr>
<td>R&lt;sub&gt;work&lt;/sub&gt;/R&lt;sub&gt;free&lt;/sub&gt;</td>
<td>0.191 / 0.229</td>
<td>0.194 /0.236</td>
<td>0.187 / 0.220</td>
</tr>
<tr>
<td><strong>B-factors</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein (Chains A, B)</td>
<td>31.6, 34.2</td>
<td>38.0, 43.5</td>
<td>34.4, 36.7</td>
</tr>
<tr>
<td>Ligand/ion</td>
<td>32.0</td>
<td>31.5</td>
<td>30.5</td>
</tr>
<tr>
<td>Water</td>
<td>37.2</td>
<td>30.1</td>
<td>43.2</td>
</tr>
<tr>
<td><strong>No. atoms</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein (Chains A, B)</td>
<td>4687, 4146</td>
<td>4690, 4136</td>
<td>4628, 4136</td>
</tr>
<tr>
<td>Ligand/ion</td>
<td>48</td>
<td>45</td>
<td>30</td>
</tr>
<tr>
<td>Water</td>
<td>545</td>
<td>72</td>
<td>538</td>
</tr>
<tr>
<td><strong>Root mean- squared deviations</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bond lengths (Å)</td>
<td>0.0178</td>
<td>0.0141</td>
<td>0.0188</td>
</tr>
<tr>
<td>Bond angles (°)</td>
<td>1.899</td>
<td>1.698</td>
<td>1.804</td>
</tr>
</tbody>
</table>

Table A3 Summary of CDK2-cyclin A bound to AZD5438, Flavopiridol and Su9516 X-ray data collection refinement statistics.

CDK2-cyclin A AZD5438 data collected on a single crystal at 100 K on a beamline I04, DLS, dataset called dls040317_CDK2_25_AZD5438. CDK2-cyclin A CGP74514A data collected on a single crystal at 100 K on a beamline I04, DLS, dataset called dls040317_CDK2_29_CGP74514A.
<table>
<thead>
<tr>
<th>CDK1-CKS2 AZD5438</th>
<th>CDK1-CKS2 CGP74514A</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data collection</strong></td>
<td></td>
</tr>
<tr>
<td>Beamline</td>
<td>I04</td>
</tr>
<tr>
<td>Space group</td>
<td>P 21 21 21</td>
</tr>
<tr>
<td>Cell dimensions</td>
<td></td>
</tr>
<tr>
<td>$a, b, c$ (Å)</td>
<td>53.3, 71.3, 72.1</td>
</tr>
<tr>
<td>$\alpha, \beta, \gamma$ (°)</td>
<td>90, 90, 90</td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>50.75-1.50</td>
</tr>
<tr>
<td>Rmerge</td>
<td>0.05 (1.14)</td>
</tr>
<tr>
<td>I/\sigma(I)</td>
<td>13.5 (1.5)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>100 (100)</td>
</tr>
<tr>
<td><strong>Redundancy</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Refinement</strong></td>
<td></td>
</tr>
<tr>
<td>No. reflections</td>
<td>44735 /2227</td>
</tr>
<tr>
<td>$R_{work}/R_{free}$</td>
<td>0.189 /0.234</td>
</tr>
<tr>
<td><strong>B-factors</strong></td>
<td></td>
</tr>
<tr>
<td>Protein (Chain A)</td>
<td>34.0</td>
</tr>
<tr>
<td>Ligand/ion</td>
<td>26.9</td>
</tr>
<tr>
<td>Water</td>
<td>38.6</td>
</tr>
<tr>
<td><strong>No. atoms</strong></td>
<td></td>
</tr>
<tr>
<td>Protein (Chain A)</td>
<td>4681</td>
</tr>
<tr>
<td>Ligand/ion</td>
<td>47</td>
</tr>
<tr>
<td>Water</td>
<td>9</td>
</tr>
<tr>
<td><strong>Root mean-squared deviations</strong></td>
<td></td>
</tr>
<tr>
<td>Bond lengths (Å)</td>
<td>0.0215</td>
</tr>
<tr>
<td>Bond angles (°)</td>
<td>2.237</td>
</tr>
</tbody>
</table>

Table A4 Summary of CDK2 bound to AZD5438 and CGP74514A X-ray data collection refinement statistics.
Bibliography


p220(NPAT) by cyclin E/Cdk2 in Cajal bodies promotes histone gene transcription', *Genes Dev*, 14(18), pp. 2298-313.


