Targeting the Aurora A/N-Myc Protein-Protein Interaction in Neuroblastoma

Sarah Louise Ward

Submitted in accordance with the requirements for the degree of
Doctor of Philosophy

The University of Leeds

School of Chemistry and School of Molecular and Cellular Biology

November 2020

The candidate confirms that the work submitted is her own and that appropriate credit has been given where reference has been made to the work of others.

This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

The right of Sarah Ward to be identified as Author of this work has been asserted by Sarah Ward in accordance with the Copyright, Designs and Patents Act 1988.

Acknowledgements

Targeting the Aurora A/N-Myc Protein-Protein Interaction in Neuroblastoma

Abstract

Neuroblastoma is a rare cancer affecting children, with around one hundred new cases each year in the UK. Mortality rates are around 40% but current intensive treatment regimens result in long-term health effects in the majority of survivors.[1] For these reasons, novel approaches to neuroblastoma treatment are required to improve survival rates and prevent chronic health problems following treatment.

The protein-protein interaction between Aurora kinase A and the transcription factor N-myc has been shown to be important in neuroblastoma. There is evidence that neuroblastoma cells become "addicted" to N-myc stabilisation by interaction with Aurora A and the disruption of this interaction leads to N-myc degradation and neuroblastoma cell death.[2]

The identification of compounds which disrupt the Aurora A/N-myc complex, either directly or by acting at an allosteric site, could have therapeutic potential and may avoid the significant toxicity of current treatments. In this work, various methods, including virtual high-throughput screening, high-throughput crystallography, *in vitro* screening and traditional medicinal chemistry approaches, were combined to assess the feasibility of several approaches to the modulation of Aurora A activity and the potential to disrupt the Aurora A/N-myc protein-protein interaction.

Computational analysis of possible binding sites of Aurora A provided some guidance in directing the project to the most promising strategies to develop a small molecule inhibitor of the key protein-protein interaction. *In silico* screening proved a useful tool to guide the selection of compounds taken forwards for crystal soaking and *in vitro* assays, which allowed the identification of a small number of fragments as potential starting points for the development of the first type III Aurora A kinase inhibitor.

Table of Contents

List of Figures

List of Tables

Abbreviations

AIR	Aurora A interaction region
AlphaScreen	Amplified luminescent proximity homogeneous assay screen
AML	Acute myeloid leukaemia
APC	Anaphase-promoting complex
ATP	Adenosine triphosphate
BLAST	Basic local alignment search tool
CDK1	Cyclin-dependant kinase 1
CENP-A	Centromere protein A
Coot	Crystallographic object-oriented toolkit
CNS	Central nervous system
CYP	Cytochrome P450
DCM	Dichloromethane
DIPEA	*N*,*N*-diisopropylethylamine
DMAP	4-dimethylaminopyridine
DMSO	Dimethyl sulfoxide
ERK	Extracellular signal-regulated kinase
EtOAc	Ethyl acetate
FLT3	FMS-related tyrosine kinase
FOLFOX	Folinic acid, fluorouracil and oxaliplatin
FP	Fluorescence polarisation
GF	Gel filtration
GSK3β	Glycogen synthase kinase-3β
HCC	Hepatocellular carcinoma
HMM	Hidden Markov Model
HP	High performance
HPLC	High performance liquid chromatography
IMAC	Immobilised metal ion affinity chromatography
IMAP	Immobilised metal ion affinity particle
IPTG	Isopropyl β-D-1-thiogalactopyranoside

ITC	Isothermal titration calorimetry
LCMS	Liquid chromatography mass spectrometry
MAPK	Mitogen-activated protein kinase
MAX	Myc-associated factor X
MEK	MAPK/ERK kinase
MES	Morpholino ethanesulfonic acid
MIBG	Meta-iodobenzylguanidine
MLCK	Myosin light chain kinase
MLK2	Mixed lineage kinase 2
MRD	Minimal residual disease
MUSCLE	Multiple sequence comparison by log-expectation
MWCO	Molecular weight cut-off
NaHMDS	Sodium bis(trimethylsilyl)amide
NCT	National clinical trial
NHL	Non-Hodgkin's lymphoma
NMR	Nuclear magnetic resonance
OD	Optical density
PAIN	Pan-assay interference
PE	Petroleum ether
PDK1	Pyruvate dehydrogenase kinase
PEG	Polyethylene glycol
PLK	Polo-like kinase
PP1	Protein phosphatase 1
ROCS	Rapid overlay of chemical structures
SD	Stable disease
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SP	Standard precision
TACC3	Transforming acidic coiled-coil-containing protein 3
TES	N-tris(hydroxymethyl)methyl-2-aminoethanesulfonic acid
TEV	Tobacco etch virus
THF	Tetrahydrofuran

TPX2	Targeting protein for *Xenopus* centrosomal kinesin-like protein 2
ULK	Unc-51-like kinase
XP	Extra precision

Chapter 1 Introduction

1.1 Neuroblastoma

Neuroblastoma is a rare cancer first defined in 1910[3] which affects the sympathetic nervous system, is diagnosed in childhood (adults are affected exceedingly rarely) and is identified in approximately 100 children each year in the UK.[1] It is the fourth most common cancer type in children under 15 years of age, behind leukaemias, lymphomas and intracranial solid tumours,[4,5] a ranking which has remained consistent since the 1970s,[6] but is the most common cancer for new-borns in the UK.[4] However, neuroblastoma is responsible for a disproportionate number of childhood deaths (under 15 years) from cancer,[1] with a mortality rate second only to malignant bone tumours[4] despite some improvement in survival in recent decades.[6]

The average age at diagnosis for these patients is 17 months and symptoms can vary greatly depending on the site of the primary tumour and subsequent organ involvement.[1] Family history of neuroblastoma is seen in a very small number of cases (1-2%) and these patients are often diagnosed earlier, at an average age of 9 months.[1,7]

Neuroblastoma has a highly variable prognosis depending on disease stage at diagnosis, affected sites, age and any underlying conditions.[7] Overall, neuroblastoma has a mortality rate close to 40%,[4] but a subset of neuroblastoma patients, known as "high-risk", have a higher mortality rate (close to 60%) and are identified by widespread metastasis, amplification of the MYCN gene and age at diagnosis over 18 months.[8]

1.1.1 Diagnosis

Diagnosis of neuroblastoma is usually made form stage 2 onwards, when symptoms often become more specific. At the early stages, generally only non-

specific pain or malaise are reported. Symptoms can include constipation, diarrhoea, swelling or twitching,[1] depending on the affected organs (Table 1.1).

Table 1.1 Possible symptoms of neuroblastoma depending on site or type of tumour[1]

Affected Area or Tumour Type	Possible Symptoms
Abdomen	Swelling
	Constipation or urinary retention (due to compression)
Orbital region	Proptosis and "raccoon eyes" (periorbital ecchymosis)
Liver	Jaundice
Vasoactive intestinal peptide secreting tumours	Diarrhoea
Catecholamine-secreting tumours	Flushing and sweating
Cerebellum	"Opsoclonus-myoclonus" syndrome (eye twitches, gait problems, mutism, vomiting)
Extensive metastasis	Generally unwell, probably combinations of the above symptoms

Neuroblastoma may also be detected during routine testing of apparently healthy children if increased levels of catecholamine derivatives in urine are observed, although these increased levels can be caused by other factors (including consumption of olives) and can produce false positives.[9] Some biomarkers can also be measured in serum, which are not specific to neuroblastoma and are therefore not useful for diagnosis, but they do play a role in prognosis prediction.[1]

Treatment decisions in neuroblastoma depend on the stage of disease. Thus, accurate staging techniques play a vital role in determining the outcome and long-term effects for patients. An initial ultrasound scan and subsequent CT scans give an indication of the extent of metastasis and an initial assessment of tumour resectability. Where bone metastasis is suspected, Meta-iodobenzylguanidine (MIBG) bone scintigraphy is performed for an accurate indication of disease stage.[1,10]

1.1.2 Treatment

Stage 4S disease, an unusual neuroblastoma sub-type, often only requires observation. Exceptions occur when patients experience symptoms associated with tumour positioning in the affected organs (for example jaundice) or in rare cases where stage 4S disease progresses to a more malignant phenotype.[1]

In general, "higher-risk" patients undergo more intensive treatment regimens. In low-risk cases surgery is often all that is required and radiotherapy is actively discouraged as the risk of long-term effects outweigh any benefit gained.[1]

By definition intermediate-risk patients have tumours which are not completely resectable. In these cases chemotherapy including etoposide and platinum-based compounds are used, after which surgery may be possible. If there is progression despite the use of these treatments, radiotherapy may be considered provided there is no spinal involvement.[1]

For high-risk patients intensive, multimodal therapy is employed. In these cases chemotherapy involves etoposide and carboplatin, as well as doxorubicin, cyclophosphamide and vincristine accompanied by radiotherapy, myeloblative therapy and haematopoietic stem cell transplants (Figure 1.1).[7]

The most recent new treatment for neuroblastoma was approved in the EU in 2017 for high-risk neuroblastoma patients over 12 months old.[11] Dinutuximab is an anti-GD2 antibody and targets the GD2 ganglioside which is overexpressed in neuroblastoma cells.[12,13]

Following treatment, a common problem in high-risk patients is Minimal Residual Disease (MRD), which triggers relapse in around 50% of cases, usually from remaining neuroblastoma cells in the bone marrow.[1] This form of relapse has a very poor prognosis as the tumours are often resistant to chemotherapy and patients have not recovered from the initial treatment, leaving very few options. New treatments targeting MRD involve immunotherapy to target diganglioside on any remaining neuroblastoma cells.[1,13]

Chronic health problems are a significant concern for survivors of childhood cancers, with more intensive treatments more likely to cause long-term effects such as infertility,[14] pulmonary effects[15] and mental health concerns including post-traumatic stress and suicide ideation.[16] The majority of late deaths of neuroblastoma survivors are due to neuroblastoma relapse and 15% are due to the occurrence of a second malignant cancer, most commonly breast, thyroid and renal carcinoma.[1]
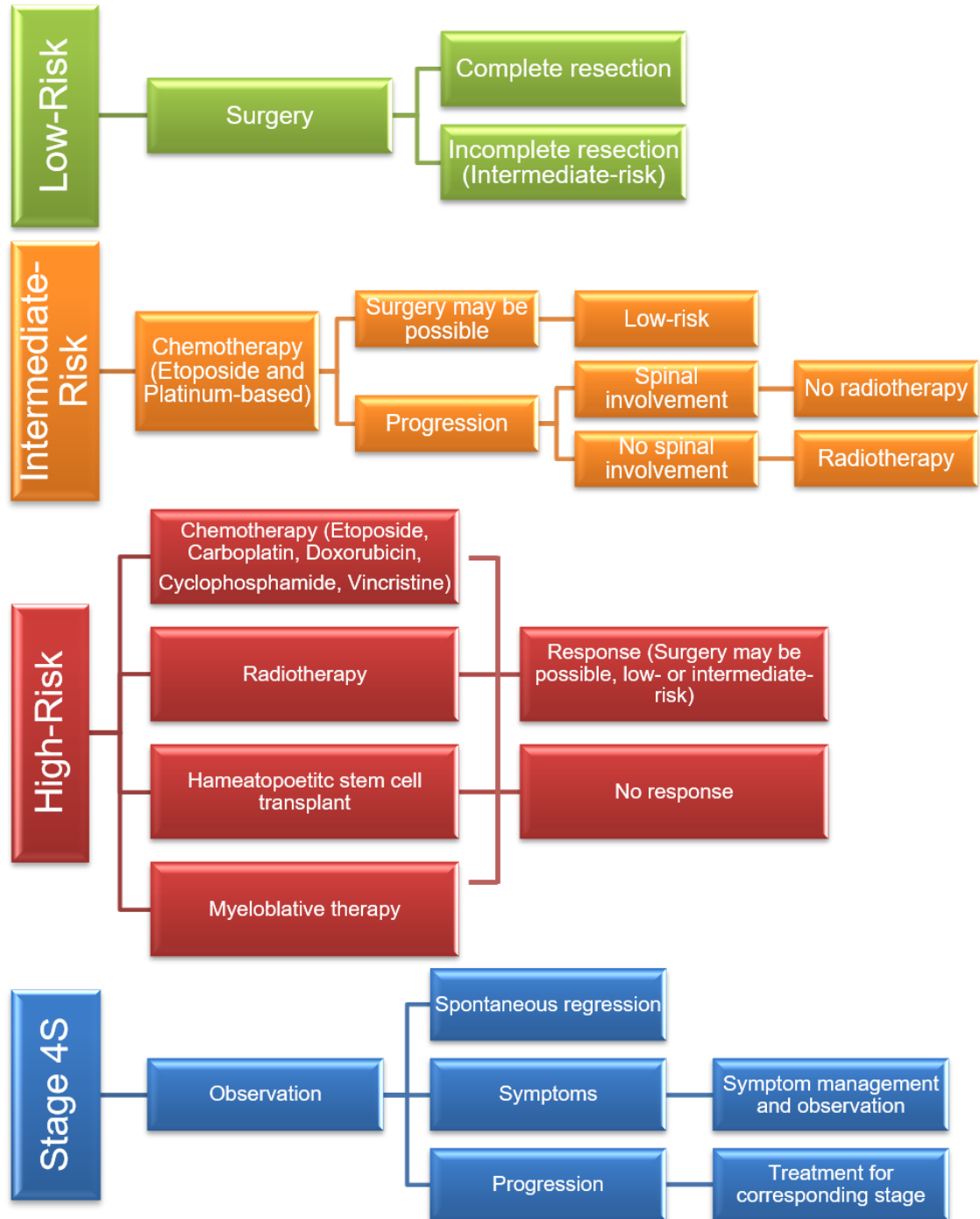
Figure 1.1 Treatment options and outcomes based on neuroblastoma risk characterisation and the Stage 4S special case[1,7]

1.2 Aurora A and Interacting Proteins

A number of proteins have been identified with the potential for involvement with neuroblastoma. Aurora A and N-myc have both been suggested as possible targets in the development of novel treatments for neuroblastoma.[2] The Aurora A/N-myc protein-protein interaction is made up of two proteins which both perform vital roles, both in healthy cells and in cancer.[2,17]

1.2.1 Aurora A

Aurora A is a serine/threonine protein kinase made up of a β-sheet domain at the N-terminus and 6 α-helices, with the two domains connected by a "hinge" region. Aurora A interacts with an array of proteins and substrates, in both health and disease.[17–19] Until recently, the Aurora kinases were not grouped with any of the kinase families and were instead grouped with "other" kinases. Using Hidden Markov Models (HMM), the Aurora kinases were able to be grouped with CAMK kinases.[20]

All human kinases have multiple common features. All kinases contain an ATP-binding site, which is highly conserved across all members of the kinase family. The tertiary structures of kinases are also similar, with N-terminal and C-terminal lobes formed of a β-sheet region and multiple α-helices, respectively.[21]

For an active kinase, further features must be present. The flexible activation loop contains a DFG motif, which must be in a DFG-in conformation in the active kinase, with the key Asp side chain directed towards the active site.[21] This feature alone is not sufficient for kinase activity, so a kinase could be inactive whilst still in the DFG-in conformation. A further key feature of an active kinase is the formation of a salt bridge between a Lys residue from the β-sheet region and a Glu side chain on the αC-helix.[21]

Finally, the correct alignment of two hydrophobic spines is required for activity. The regulatory R-spine is responsible for ensuring the correct positioning of substrates and is formed of the Phe side chain of the DFG motif and residues from both the α-

helix and β-sheet regions. In an active kinase the R-spine should be mostly linear. The catalytic C-spine is involved in alignment of ATP in preparation for catalysis and again involves residues from both the C- and N-lobes of the kinases and the adenine portion of ATP.[22]

1.2.1.1 Aurora A in Healthy Cells

Aurora A was first discovered in 1990 as one of a family of three proteins (Aurora A, B and C).[23] All three Aurora proteins function as kinases, interacting with a variety of other proteins and phosphorylating a range of substrates. When active, Aurora A is autophosphorylated at Thr288.[24]

Aurora kinase A is inactive for most of the cell cycle and is found in centrosomes.[18] During the late G_2 phase Aurora A interacts with and is activated by the protein Ajuba. Ajuba is both a substrate and activator for Aurora A: only phosphorylated Ajuba can form a complex with Aurora A.[25] Active Aurora A then forms a complex with TPX2, which is also both a substrate and activator of Aurora A and this interaction can trigger mitosis (Figure 1.3).[17,18, 24,26]

During the early stages of mitosis Aurora A moves to the mitotic spindles and the Aurora A/TPX2 complex helps to ensure correct bipolar spindle assembly and loss of either of these proteins results in loose or unstable spindles.[17,24,25]

During metaphase Aurora A phosphorylates the protein CENP-A at Ser7 (Figure 1.3). It is believed that phosphorylated CENP-A plays a crucial role in attaching microtubules to the kinetochore, which provides a framework to align chromosomes before separation. Therefore, the phosphorylation of CENP-A by Aurora A indirectly affects chromosome alignment and separation.[17]

Levels of Aurora A are reduced from the metaphase-anaphase transition onwards. The anaphase-promoting complex (APC) targets Aurora A for proteasomal degradation using adapter proteins which target the "D-box" (destruction box) of Aurora A, a short sequence (RxxL) highly conserved across all three members of the Aurora family (

Table **1.2**).[17,27] The D-box sequence is crucial for reducing levels of active Aurora A to complete mitosis.
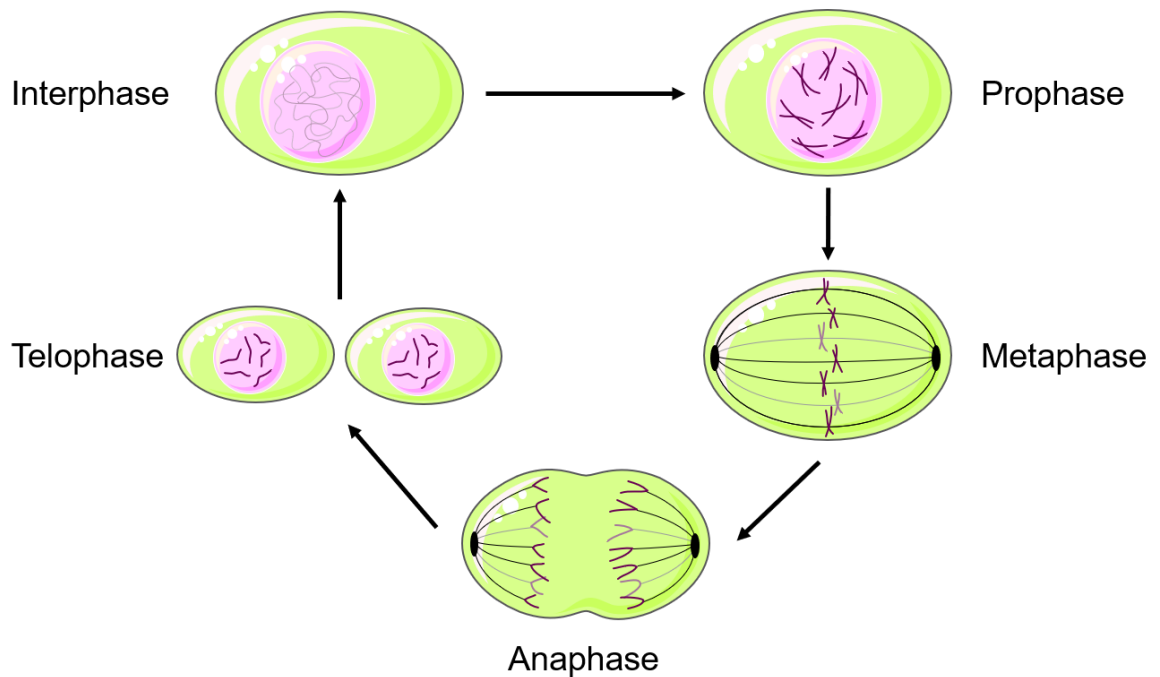
Figure 1.2 Stages of mitosis

Table 1.2 D-box sequences for Aurora A, B and C in different species

Protein	Species	Residue numbers	Sequence
Aurora A	*Homo sapiens*	371-374	...HNPSQ**RPML**REVLEHPW...
	Xenopus laevis	378-381	...HNPNH**RLPL**KGVLEHPW...
	Mus musculus	384-387	...HNASQ**RLTL**AEVLEHPW...
Aurora B	*Homo sapiens*	315-318	...HNPSE**RLPL**AQVSAHPW...
	Xenopus laevis	331-334	...YHPPQ**RLPL**KGVMEHPW...
	Mus musculus	320-323	...HNPWQ**RLPL**AEVAAHPW...
Aurora C	*Homo sapiens*	281-284	...YQPLE**RLPL**AQILKHPW...
	Mus musculus	244-247	...YHPSE**RLSL**AQVLKHPW...
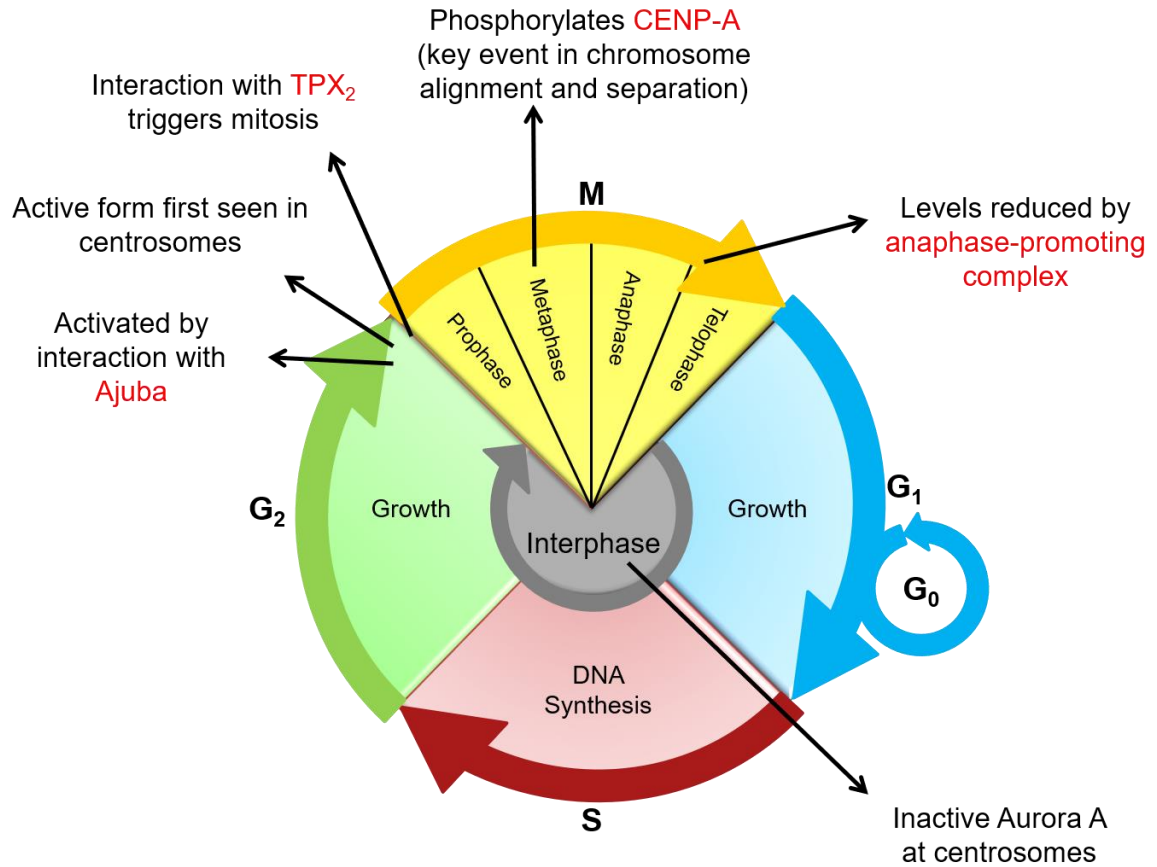
Figure 1.3 Key events and interactions of Aurora A during the cell cycle; proteins which either form interactions with, or are substrates of, Aurora A are red.[17,18]

1.2.1.2 Aurora A Kinase Activity in Cancer

Aurora A has been implicated in cancer on its own, regardless of interaction with N-myc.[28,29] As mentioned above, levels of activated Aurora A increase from the late G_2 phase into mitosis.[17] However, some cancers show increased Aurora A at all stages of the cell cycle and Aurora A is not usually restricted to the nucleus in cancer.[17] For this reason, phosphorylation of various proteins is seen in cancer as well as increased levels of phosphorylation of the normal substrates of Aurora A.[17,30]

Furthermore, there is a cell cycle checkpoint between the G_2 and M phases, which can prevent cells with damaged DNA from entering mitosis.[17] When Aurora A protein levels are unusually high, DNA-damage signals, which would normally inhibit Aurora A, still leave high levels of active Aurora A which can override the checkpoint and allows the cell to enter mitosis.[17]

Finally, Aurora A appears to have a complicated relationship with the tumour suppressor p53.[31] It has been reported that p53 can abrogate the oncogenic activity of Aurora A. However, there is also evidence that p53 is a substrate of Aurora A and phosphorylation at Ser215 inhibits the transcriptional role of p53. In addition, phosphorylation of p53 at Ser315 prepares p53 for export from the nucleus, where it is a substrate for MDM2-mediated degradation (Figure 1.4).[17]
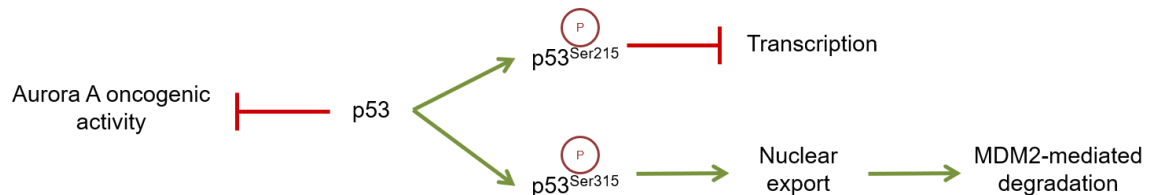
Figure 1.4 A summary of the relationship between p53 and Aurora A: p53 can block the oncogenic activity of Aurora A but phosphorylation of p53 at Ser215 or Ser315 by Aurora A results in loss of transcriptional activity or nuclear export and degradation, respectively[17]

Inhibition of the kinase activity of Aurora A could be a useful target for new cancer treatments, preventing the involvement of Aurora A in the cancerous pathways described. Aurora A kinase activity is essential for mitosis so cell death arising from kinase inhibition might be expected.

1.2.1.3 Aurora A Interactions in Cancer

As mentioned above, the interaction between Aurora A and TPX2 is important in initiating mitosis (Figure 1.3). When TPX2 binds, the activation loop of Aurora A stretches slightly so substrates are better able to bind in the active site. The presence of TPX2 also protects Aurora A from dephosphorylation and inactivation

at Thr288 by protein phosphatase 1 (PP1).[32] TPX2 binding also stimulates Aurora A autophosphorylation at Thr288 and subsequent Aurora A activation.[33,34] In the event of DNA damage (especially double strand breakage) mitosis is inhibited by decreasing TPX2 levels and activating the anaphase promoting complex which promotes Aurora A degradation and G_2 arrest.[32]

In some hepatocellular carcinoma (HCC) cell lines an interaction between Aurora A and the transcription factor c-myc has been seen, especially in cases with p53 mutations. Aurora A helps to stabilise c-myc, allowing c-myc to support tumour cell proliferation and survival. Liver cirrhosis is the biggest risk factor for HCC. Although the underlying mechanisms are not fully understood, a possible explanation proposed by Dauch *et al.* suggests the requirement for liver regeneration results in increased c-myc, eventually leading to evasion of the G_2/M cell cycle arrest.[35]

1.2.2 N-myc

N-myc is one member of a family of three proteins (N-myc, c-myc and L-myc), which all function as transcription factors and together the myc family are thought to regulate the transcription of one third of the genome.[31] They are particularly involved in promoting the transcription of proteins involved in cell growth and proliferation, as well as angiogenesis and inhibition of differentiation, all key events in cancer.[36]

N-myc is absolutely essential for proper development: knockout of the MYCN gene is embryonic lethal in mice due to widespread organ failure resulting from insufficient proliferation. However, myc proteins are only detected at low levels in non-dividing cells, if at all, so although myc proteins are important, they are not required in all cell types at all times.[36]

As a transcription factor, N-myc binds to another protein, MAX (myc-associated factor X, Figure 1.5), via helix-loop-helix-leucine zipper regions on each of the proteins. Together, the N-myc/MAX complex binds to a variety of DNA sites, recognising the "E-box" DNA sequence (CACGTG) (Figure 1.5). MAX also binds to

other proteins with similar helix-loop-helix-leucine zipper regions, which are believed to act as inhibitors of N-myc/MAX-mediated transcription.[36,37]
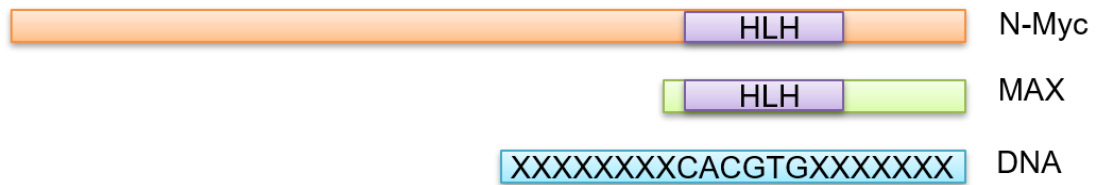
Figure 1.5 Schematic of the interaction between N-myc and MAX: interaction between the two proteins is via helix-loop-helix regions (HLH, purple box); DNA is recognised by the complex at the "E-box" sequence of DNA (CACGTG) [36]
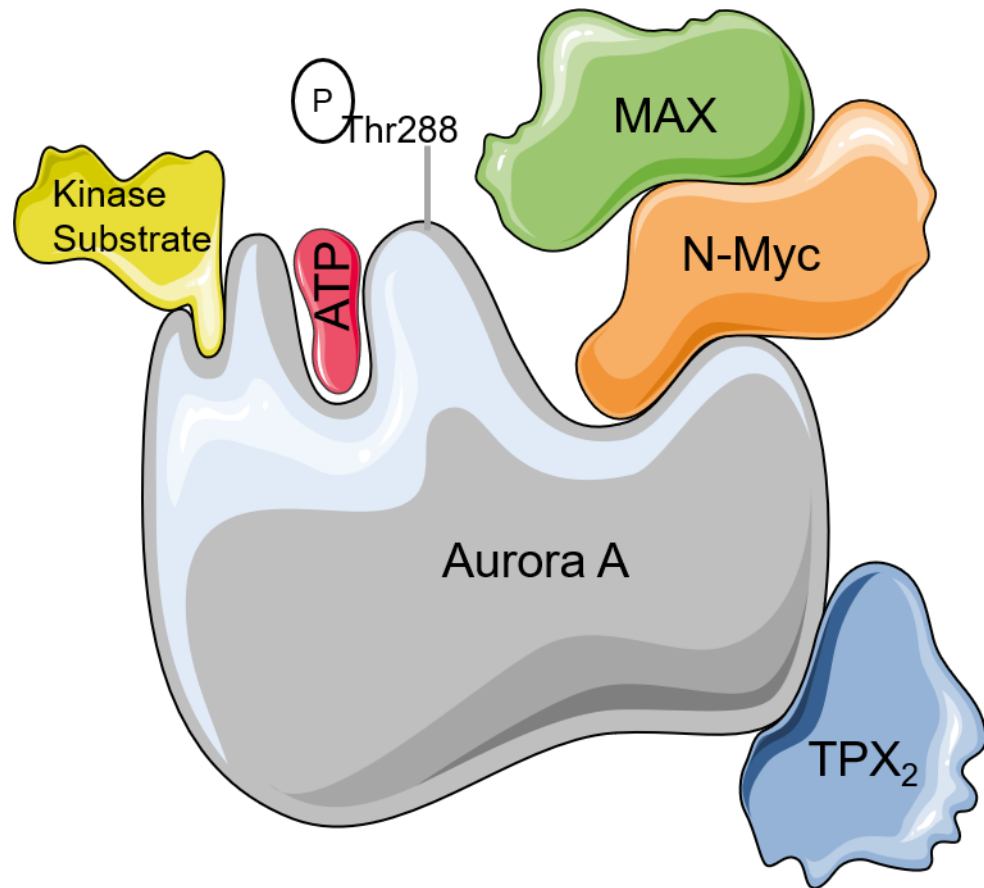
Figure 1.6 Diagram of the interactions of Aurora A with other key proteins, in reality the different binding sites are less distinct and may overlap

1.2.3 TACC3

TACC3 is a substrate and activator of Aurora A and this interaction is important for successful mitosis by ensuring correct central spindle formation.[38] Overexpression of TACC3 has been implicated in various cancers and there has been some work towards identifying modulators of TACC3.[39] A crystal structure of the Aurora A/TACC3 complex has been published, with the inactive form of Aurora A (Figure 1.7).[40] TACC3 is phosphorylated at Ser558 by Aurora A and the phosphorylated form of TACC3 is able to form an interaction with ch-TOG (colonic and hepatic tumor overexpressed gene), which in turn can form a TACC3-chTOG-clathrin complex which is believed to stabilise microtubules by crosslinking different chains.[41,42]
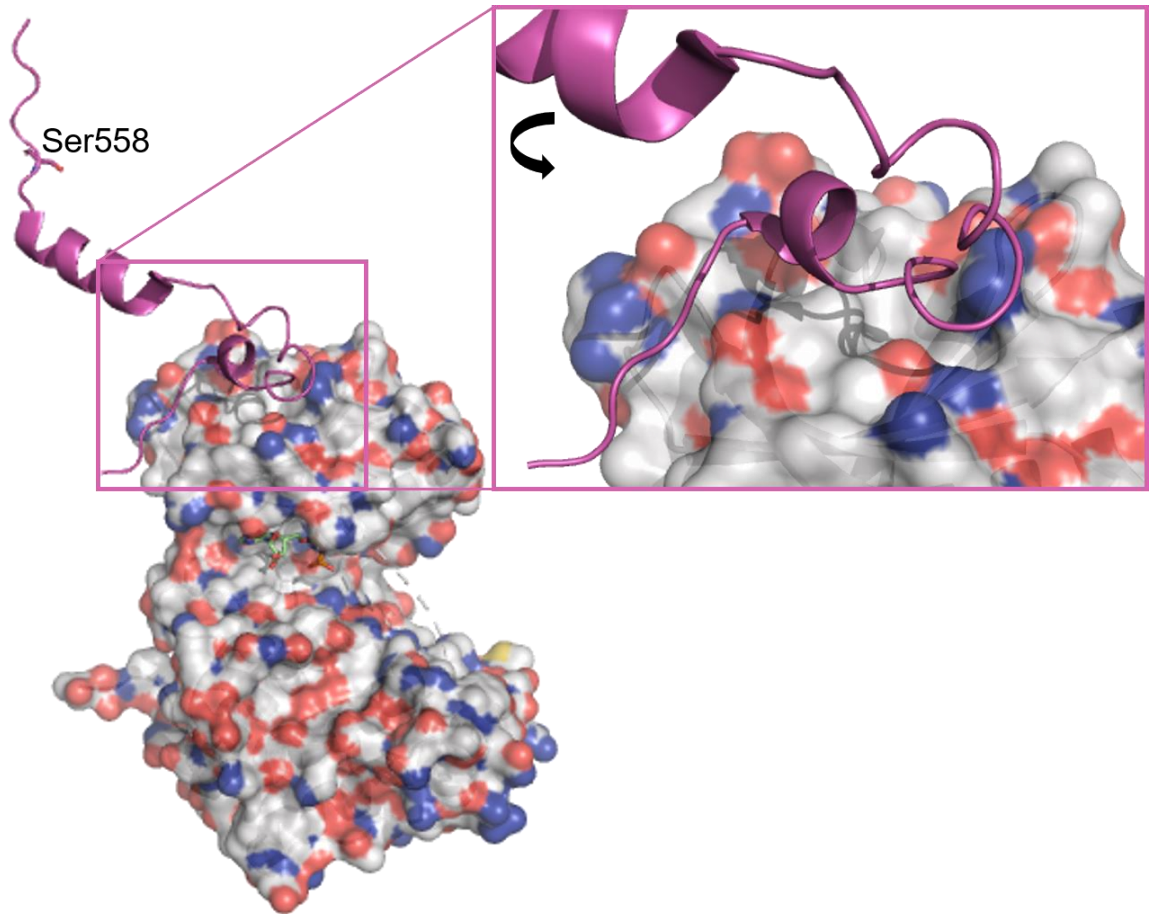
Figure 1.7 The Aurora A/TACC3 interaction (PDB 5ODT)[40] with Aurora A carbons shown in grey and TACC3 carbon in pink; ADP is shown (green sticks) in its binding site

1.3 Aurora A/N-myc in Neuroblastoma

It has been demonstrated that increased levels of the Aurora A/N-myc complex correlates with a poorer prognosis in neuroblastoma. The complex protects N-myc from degradation, allowing it to continue to function as a transcription factor.[43] There is also evidence that neuroblastoma cells can become "addicted" to high levels of N-myc and reduction in the amount of N-myc in the cell can lead to apoptosis. Recently, the Aurora A/N-myc interaction has also been implicated in neuroendocrine prostate cancer.[44]

When in a complex with N-myc, both the ATP-binding site and the kinase active site of Aurora A are unaffected. The protein-protein interaction is shown in Figure 1.8.
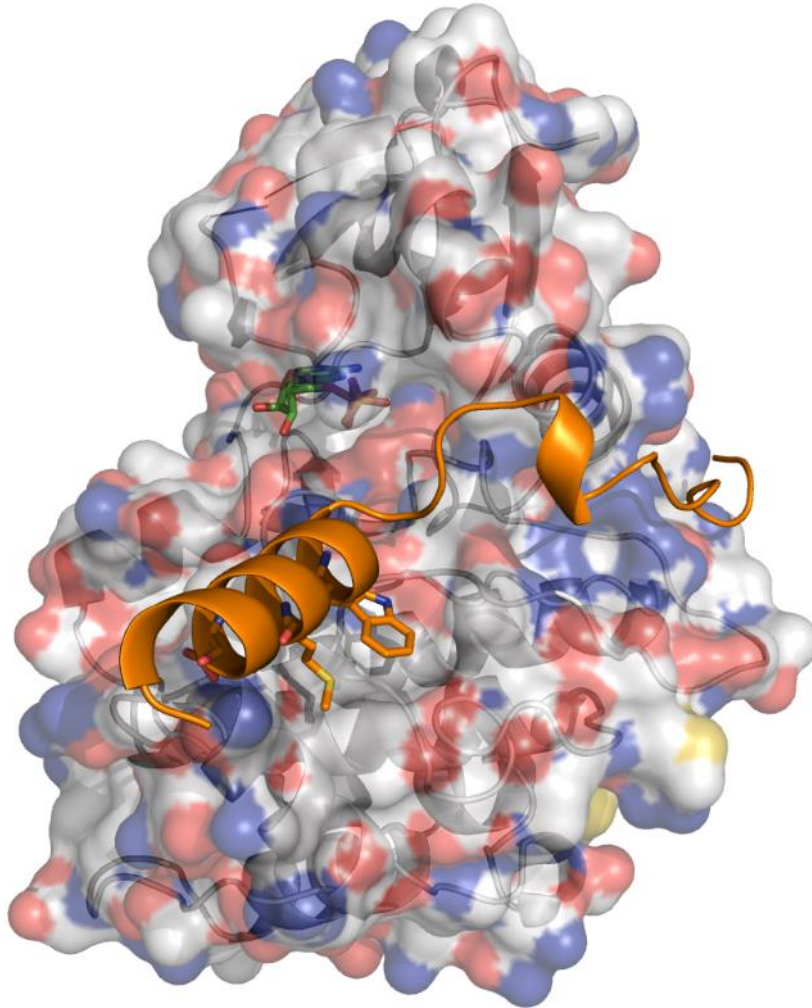
Figure 1.8 The Aurora A/N-myc interaction with Aurora A carbons shown in grey and N-myc carbon and backbone in orange (PDB 5G1X).[45] ADP is also shown (green sticks) in the ATP binding site.

N-myc is protected from degradation by its interaction with Aurora A, which conceals and prevents access to two key sites of phosphorylation on N-myc.[43] When the Aurora A/N-myc complex separates, N-myc is first phosphorylated at Ser62 by cyclin B/CDK1, then at Thr58 by GSK3β.[2] Once both sites are

phosphorylated, N-myc is ubiquitinated and targeted for proteasomal degradation by Fbxw7, a ubiquitin ligase (Figure 1.9).[19, 37,45]

In vitro, shRNA (short hairpin RNA) has been used to knockdown MYCN, the gene encoding N-myc, in neuroblastoma cell lines.[37] IMR32 and SH-EP cells were used which express amplified and single copy MYCN respectively. IMR32 cells appear to depend on high levels of N-myc, as MYCN knockdown gives a reduction in the number of cells. Conversely, no effect is seen when MYCN is disrupted in SH-EP cells.[2]
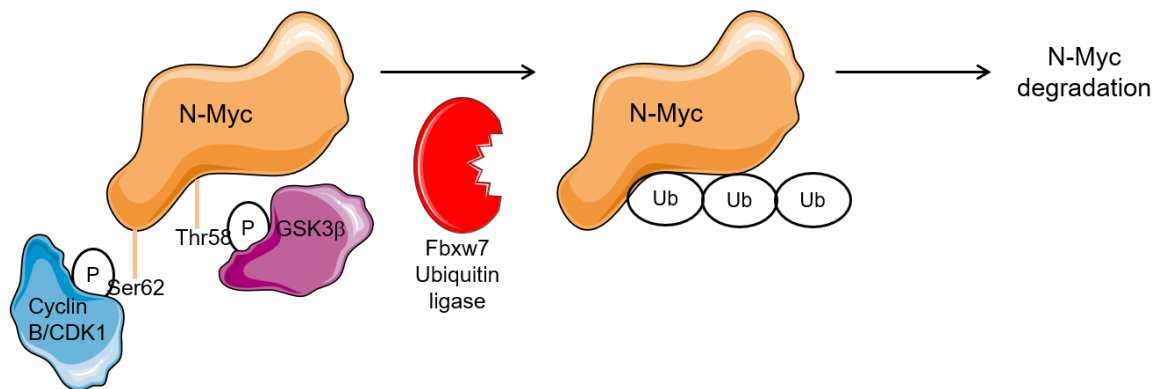
Figure 1.9 Dissociation of the N-myc/Aurora A complex allows phosphorylation of N-myc at Thr58 and Ser62 by GSK3β and cyclin B/CDK1 respectively; phosphorylated N-myc is ubiquitinated by Fbxw7 and targeted for proteasomal degradation[2,37]

A similar effect is seen when siRNAs (small interfering RNA) targeting AURKA (encoding Aurora A) are transfected into the same cell lines. IMR32 cells show lower proliferation rates whereas SH-EP cells are unaffected. In these studies AURKA knockdown did not induce apoptosis in IMR32 cells but did make them more sensitive to other chemotherapy agents including doxorubicin.[46]

1.4 Aurora A/N-myc

The Bayliss lab has produced a crystal structure of the Aurora A/N-myc interaction at 1.72 Å resolution (Figure 1.8).[45] Only one of the two α-helices of N-myc are present, but the crystal structure is able to provide an insight into the relative positions of important binding sites of Aurora A and highlights some hydrogen bonding interactions between N-myc and Aurora A. When Aurora A forms an interaction with N-myc, the active conformation of the kinase is observed.[47]

A crystal structure has been published of Aurora A in complex with vNAR-D01,[48] a single variable domain of the antigen receptor from sharks, which stabilised a conformation of Aurora A in which the salt bridge between Lys162 and Glu181 was broken, resulting in an inactive kinase state. This break in the salt bridge can also be seen in the crystal structures of Aurora A with some active site kinase inhibitors,[49–54] so the identification of a drug-like compound which stabilises this inactive conformation of Aurora A may be useful. As mentioned, the Aurora A/N-myc interaction sees Aurora A in the active conformation, so confining Aurora A to the inactive conformation, with a misaligned R spine and αC-helix, could generate an Aurora A conformation which is incompatible with N-myc binding, thus leading to N-myc degradation and neuroblastoma cell death.[47]

1.5 Kinase Inhibitors

With over 500 kinases in the human kinome,[55] many of which have been linked to disease, it is unsurprising that this class of proteins has become a major target class for new drugs. The high level of interest in small molecule modulators of kinases has necessitated the classification of kinase inhibitors into six distinct "types," characterised by their mode of binding to the target protein (Table 1.3).[21,56]

There are many examples of successful type I, I½ and II inhibitors, including crizotinib (**1**), a type I inhibitor of ALK and type I½B inhibitor of c-Met, gefitinib (**2**), a type I inhibitor of EGFR, and pazopanib (**3**, Figure 1.10), a type I inhibitor of VEGFR. As mentioned above, the ATP-binding regions of kinases are highly

conserved so achieving selectivity for a single kinase can be a challenge with ATP-competitive inhibitors. Type A inhibitors, which extend beyond the main cleft of the ATP-binding site and are a subtype of type I, I½ and II inhibitors, may provide an opportunity to introduce selectivity.[57] Conversely, the type B subgroup inhibitors occupy only the main cleft of the ATP-binding site.[21]

There is currently only one type III kinase inhibitor, approved for use in advanced melanoma and non-small cell lung cancer with mutations in the BRAF gene.[58] Trametinib (**4**) is an inhibitor of MEK and computational models have suggested trametinib binds close to the ATP-binding site without competing with ATP, although there is no crystal structure of MEK with trametinib bound to confirm this.[59]

Type I, I½, II

1

2

3

Type III

4
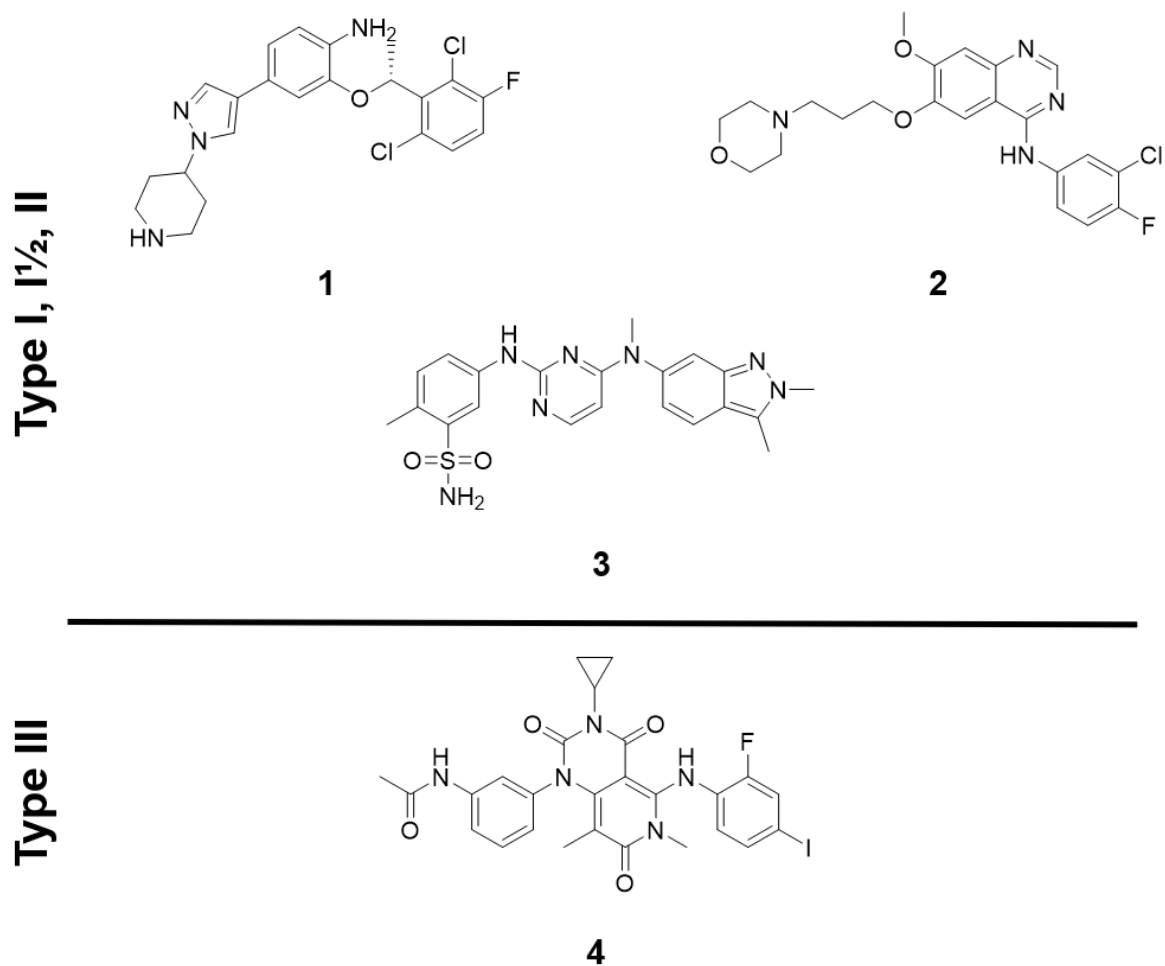
Figure 1.10 Type I-IV kinase inhibitor structures mentioned in this section, grouped by inhibition type

Type IV inhibitors include sirolimus (**5**), temsirolimus (**6**) and everolimus (**7**), all of which act on the same target in the mTOR pathway.[21] There are currently no clinical examples of type V inhibitors. Ibrutinib (**8**) and afatinib (**9**, Figure 1.11) are both examples of covalent type VI inhibitors. Both compounds react via Michael addition with their target proteins.[21]

Type IV

5

6

7

Type VI

8

9
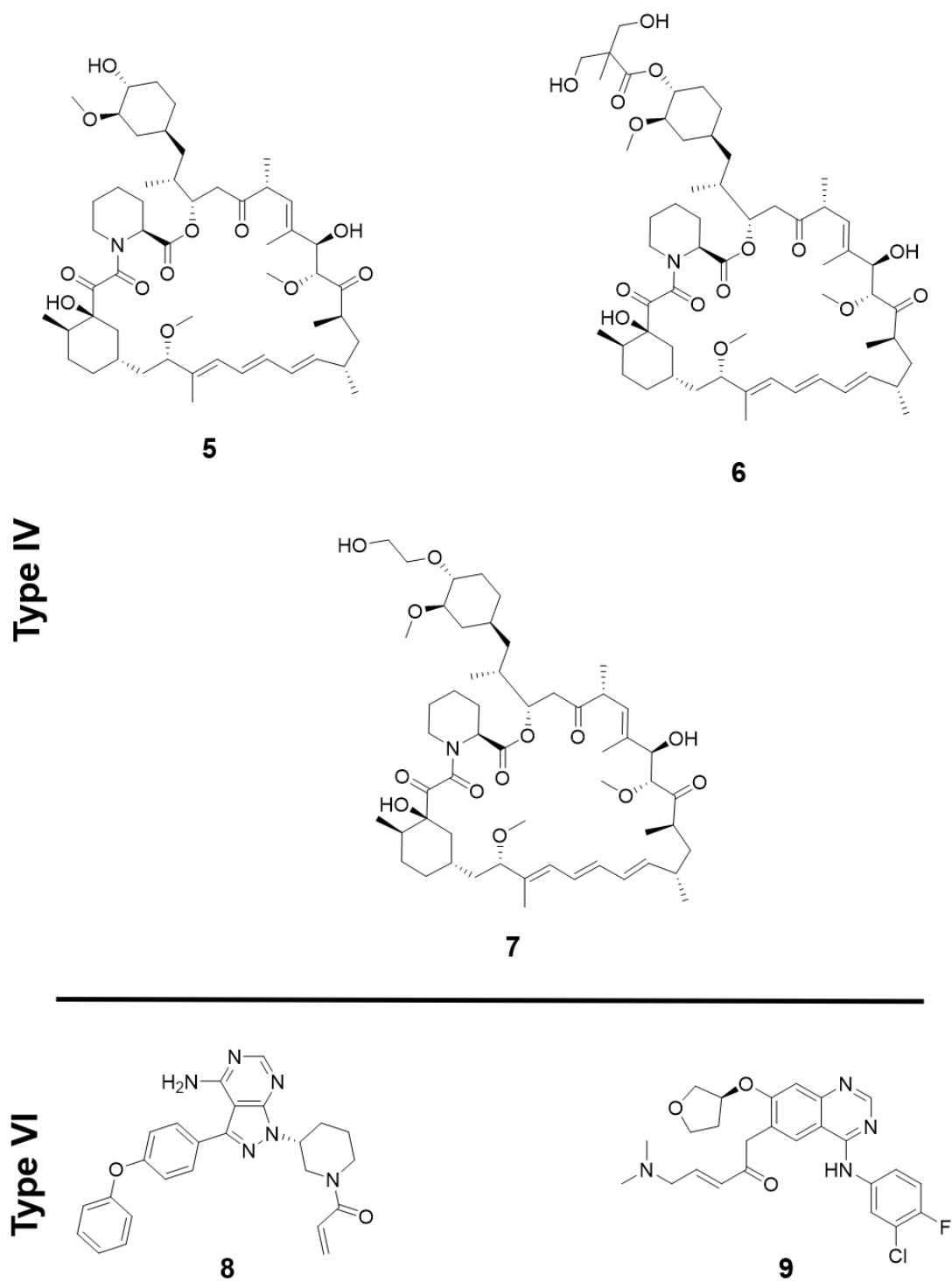
Figure 1.11 Type IV and VI kinase inhibitor structures mentioned in this section, grouped by inhibition type

Table 1.3 A summary of kinase inhibitor types

Type	Subtype	Binding Site	Kinase Conformation	Mode of Inhibition
Type I	-	ATP pocket	Active	ATP-competitive
Type I½	A	ATP pocket and extend to back cleft	Inactive conformation (DFG-in)	ATP-competitive
	B	ATP pocket (only front cleft)	Inactive conformation (DFG-in)	ATP-competitive
Type II	A	ATP pocket and extend to back cleft	Inactive (usually DFG-out)	ATP-competitive
	B	ATP pocket (only front cleft)	Inactive (usually DFG-out)	ATP-competitive
Type III	-	Next to the ATP pocket	Inactive	Allosteric (Non-competitive with ATP)
Type IV	-	Distant from substrate and ATP binding pockets	Any conformation	Allosteric
Type V	-	Two different regions	Any conformation	Allosteric
Type VI	-	Any binding site	Any conformation	Covalent

1.6 Existing Aurora Inhibitors

A number of Type I ATP-competitive inhibitors of Aurora A have been reported[60–67]: the MLN series of compounds are selective Aurora A kinase inhibitors and CD532 is an Aurora A kinase inhibitor and an allosteric Aurora A/N-myc interaction inhibitor.[31] The third compound which will be discussed, VX-680, is a non-selective Aurora kinase inhibitor.

1.6.1 MLN Compounds

MLN8054 **(10)** and its successor, MLN8237 **(11)** (Figure 1.12) inhibit the kinase activity of Aurora A and the crystal structure of Aurora A with MLN8054 shows this compound bound in the ATP binding site (PDB 2X81).[49]

MLN8237 **(11)** is a specific Aurora kinase A inhibitor which has advanced to clinical trials for treatment of neuroblastoma under the name Alisertib.[68] MLN8237 is an ATP-competitive inhibitor of Aurora A and is selective for Aurora A over Aurora B with IC_{50} values of 1.2 nM and 396.5 nM respectively.[31] Evidence for selectivity within the aurora family is obtained by measurement of the levels of phosphorylated histone H3, which can be phosphorylated at Ser10 by both Aurora A and Aurora B.[31] With a nonselective inhibitor, phosphorylated H3 levels decrease in a dose-dependent manner, as expected. However, with an Aurora A selective inhibitor, there is an initial increase in phosphorylated H3 at low concentrations of the inhibitor as Aurora B activity increases in response to Aurora A inhibition. At higher concentrations, there is a sudden drop in phosphorylated H3 as Aurora B is no longer able to compensate. Aurora A inhibition by MLN8237 causes a G_2/M arrest in the cell cycle as Aurora A kinase activity is vital for mitosis.[31]

The MLN compounds **(10, 11)** also show a reduction in the levels of MYCN in SK-N-BE(2) cells, a neuroblastoma cell line with amplified MYCN.[37] However, the effect on MYCN levels are limited, even at much higher doses than the IC_{50}, suggesting that inhibition of the kinase function of Aurora A does not affect the protein-protein interaction between Aurora A and N-myc. The small decrease in N-

myc observed with these compounds is believed to arise either from prolonged kinase inhibition or a small change in Aurora A conformation.[31]
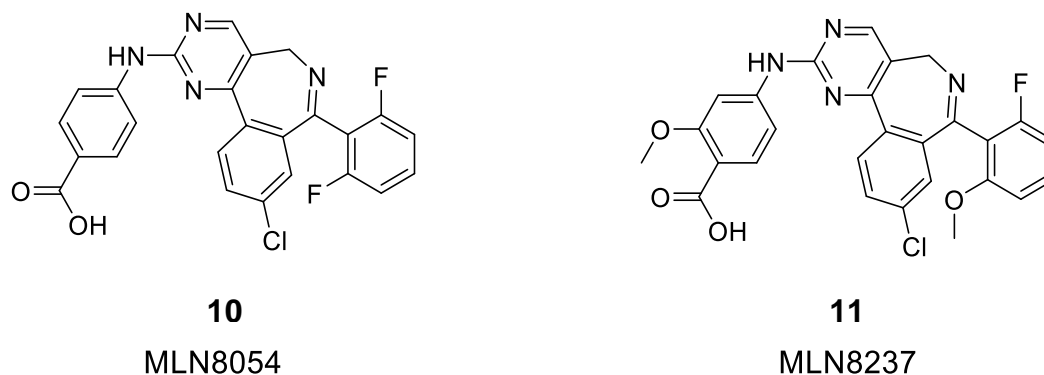
Figure 1.12 Structures of MLN series of compounds

MLN8237 (alisertib) **(11)** was considered promising enough to warrant testing in clinical trials and entered a phase I trial for relapsed neuroblastoma patients testing the combination of MLN8237 with irinotecan and temozolamide.[69] Irinotecan and temozolamide were chosen as the regime is generally well-tolerated though response rates are very low. MLN8237 has been shown to increase the sensitivity of neuroblastoma cells to standard chemotherapies including doxorubicin *in vitro*. However, the combination of MLN8237, irinotecan and temozolamide showed no improvement in survival.[68] Significant toxicity was also seen, resulting in the need to modify the protocol to include myeloid growth factor support.[68] Currently, the compound **(11)** is in clinical trials for advanced breast cancer and CNS tumours in adults.[70]

Alisertib has also been assessed for its effect on other Aurora A-driven cancers. In a phase I study of advanced gastrointestinal cancers, alisertib was given along with a modified FOLFOX regimen (folinic acid, fluorouracil and oxaliplatin). The treatment regime was only tolerated at the lowest dose of alisertib, which only achieved a partial response (PR) in one out of fourteen patients involved in the trial, and this treatment did not progress beyond phase I.[71]

In a 2019 phase I trial in combination with pazopanib in advanced solid tumours in adults, 15 of 27 patients experienced stable disease (SD) and two achieved PR.[72] Alisertib has been shown to be metabolised by CYP 3A4,[73] indicating a potential for drug interaction with pazopanib, also metabolised by CYP 3A4.[74] Indeed, pharmacokinetic analysis showed an increase in alisertib plasma concentration when administered with pazopanib.[72]

Alisertib has entered clinical trials for many other cancers in combination with many other existing regimens (Table 1.4) with limited success. There are still ongoing trials to try to find a useful setting for alisertib.

Table 1.4 Clinical trials using alisertib[70, 72,75–82]

NCT Identifier	Target Disease	Combination Therapy	Outcome
NCT02114229 **(Phase II)**	CNS rhabdoid tumours	With or without radiotherapy and/or standard chemotherapy	Ongoing
NCT02860000 **(Phase II)**	Advanced endocrine-resistant breast cancer	With or without fulvestrant	Ongoing
NCT02719691 **(Phase Ib)**	Refractory solid tumours or metastatic triple-negative breast cancer	Sapanisertib	Ongoing
NCT00807495 **(Phase II)**	Aggressive B-and T-cell non-Hodgkin lymphoma (NHL)	Monotherapy	Completed 2011 (no further investigation into this use)

NCT01482962 (Phase III)	Relapsed peripheral T-cell lymphoma	Monotherapy	Terminated 2015 (no improvement in survival)
NCT01639911 (Phase I)	Advanced solid tumours	Pazopanib	Completed 2016 (partial response in 2 of 27 patients)
NCT02319018 (Phase I)	Advanced gastrointestinal cancers	Modified FOLFOX	Completed 2018 (no results published)
NCT01601535 (Phase II)	Relapsed neuroblastoma	Irinotecan and temozolamide	Completed 2018 (partial response in 21% of participants, but currently no progression to phase III)
NCT01154816 (Phase II)	Young patients (<21 years) with refractory solid tumours or leukaemia	Monotherapy	Completed 2015 (response rate <5%)
NCT01397825 (Phase I)	Relapsed or refractory B-cell lymphoma	Rituximab or rituximab and vincristine	Completed 2016 (seven complete

			responses and seven partial responses in 45 patients, 89% experienced various adverse events)
NCT01848067 **(Phase I/II)**	Hormone-resistant prostate cancer	Abiraterone acetate and prednisone	Completed 2016 (Dose-limiting toxicity in all nine patients and no response at doses used)

It has been suggested that clinical trials for investigational neuroblastoma treatments should be modified in response to the lack of suitable patients (generally only relapsed or refractory patients are eligible) and to enable the assessment of a treatment's effect at an earlier stage of disease, as it is especially difficult to achieve any response in relapsed neuroblastoma.[83] One proposal is to identify more groups of high-risk patients to use in clinical trials, including those presenting with MYCN amplification, before relapse, with small groups of patients each trialling a variation of a new treatment in the hopes of more quickly identifying the most useful treatment regimens.[84]

1.6.2 CD532

CD532 **(12)** (Figure 1.13) is a selective Aurora A inhibitor developed by Gustafson *et al.*[31] Even at low concentrations of CD532, a loss of N-myc could be detected. CD532 inhibits the kinase activity of Aurora A (IC_{50} = 45 nM), though not as potently as MLN8237 (IC_{50} = 1.2 nM) and does not cause a G_2/M cell cycle arrest.[31] Instead, a G_0/G_1 arrest is seen, which accompanies the loss of N-myc. Furthermore, a nonphosphorylateable N-myc mutant was not degraded to the same extent as wild-type N-myc, supporting the idea that phosphorylation of N-myc after disruption of the interaction is a key step in N-myc degradation.[31,43]

The crystal structure of Aurora A bound to CD532 helps to explain the apparent disruption of the Aurora A/N-myc interaction. CD532 binds at the ATP binding site and induces a conformational change which appears to affect the binding of N-myc to Aurora A. The crystal structure shows a 6.5-7.2 Å shift of β1 and β2, causing a twist in the β-sheet region (Figure 1.14). There is also a 6.2 Å shift of the α-helix closest to the β-sheet region and a 180º flip of the activation loop. For this reason the authors describe CD532 as an "amphosteric" inhibitor, both competing with ATP at the ATP-binding site and acting at an allosteric site to disrupt protein-protein interactions.[31]
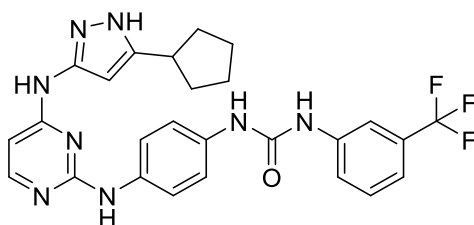
12

Figure 1.13 Structure of CD532
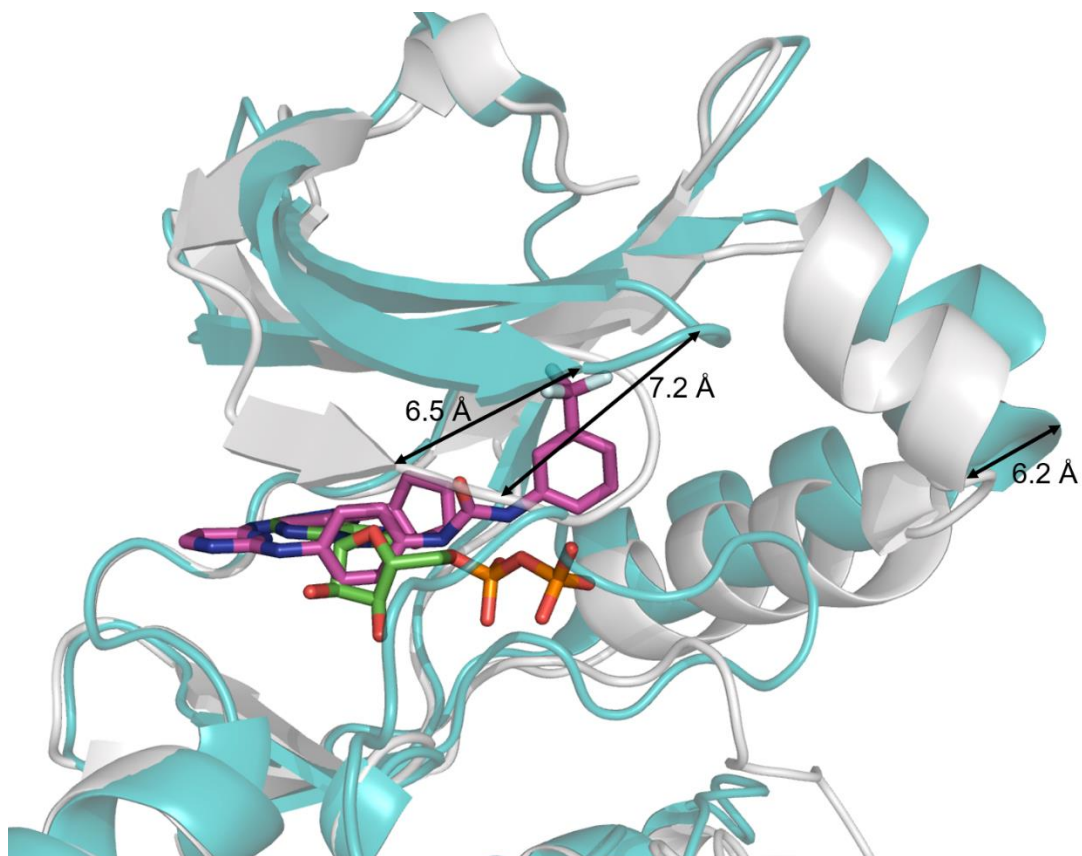
Figure 1.14 Aurora A with N-myc (grey, PDB 5G1X), with ADP bound (green sticks), overlaid with Aurora A crystallised with CD532 (turquoise and pink sticks, PDB 4J8M). Shifts in conformation are shown as measurements.[31]

There is limited published SAR data for the development of CD532 but published data highlight the importance of the trifluoromethyl group (Figure 1.15) for the disruption of the Aurora A/N-myc complex. Other groups appear to be important in strengthening the interaction with Aurora A.[31]
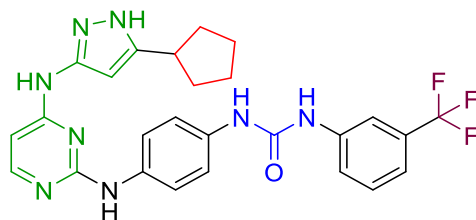
Figure 1.15 The structure of CD532 with groups involved in interaction with Aurora A highlighted; the cyclopentyl group (red) sits in a hydrophobic pocket of Aurora A; the aminopyrazole-pyrimidine section (green) facilitates binding to the ATP binding hinge and the urea feature (blue) interacts with an important catalytic residue, Asp274. Finally, the trifluoromethyl (purple) is believed to be large enough to displace surrounding chains and trigger a conformational change in Aurora A.[31]
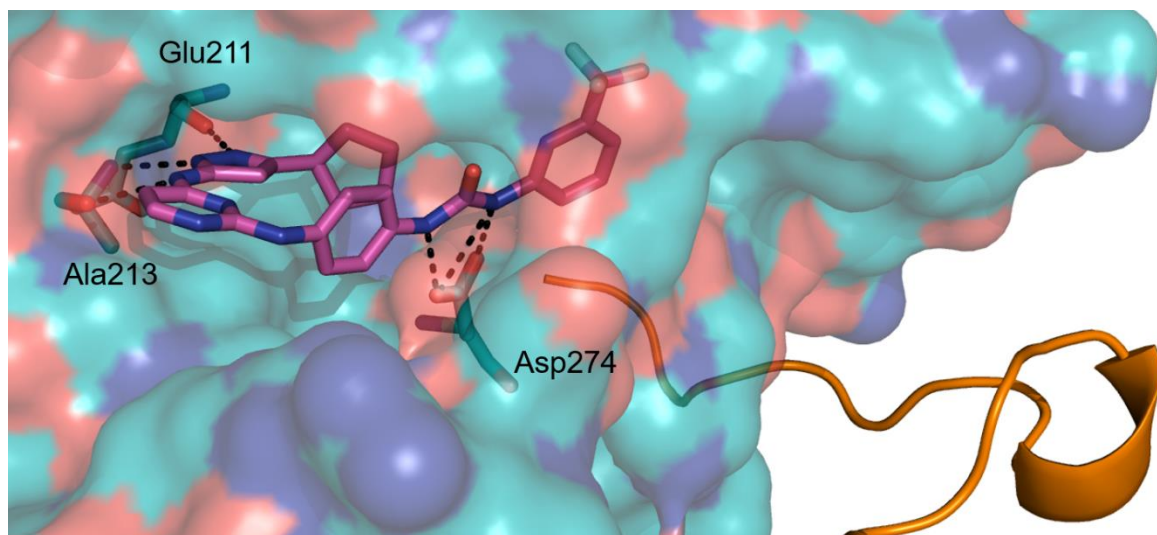
Figure 1.16 CD532 (pink sticks) in the ATP binding site of Aurora A; hydrogen bonds are shown in black (PDB 4J8M); interacting side chains are labelled and shown as sticks; N-myc (orange) is overlaid (from PDB 5G1X) for reference

Preclinical trials of CD532 using neuroblastoma xenografts in mice showed a decrease in N-myc levels with addition of CD532. The mouse models showed significantly improved survival and a significant reduction in total tumour volumes.[31]

Using slowly increasing CD532 concentrations, the extent of the difference between CD532 and MLN8237 was measured: CD532 promoted dose-dependent loss of the protein-protein interaction, with MLN8237 producing a greatly reduced response. CD532 also appears to selectively disrupt the Aurora A/N-myc interaction, as the N-myc/MAX binding was not affected.[31]

1.6.3 VX-680

VX-680 **(13)** (Figure 1.17), developed by Vertex Pharmaceuticals, inhibits all three members of the Aurora kinase family with K_i values of 0.6, 18 and 4.6 nM for Aurora A, B and C respectively.[85] In cell-based assays, VX-680 has been shown to inhibit proliferation of tumour cell lines with IC_{50} values in the range of 15-130 nM.[86] Although VX-680 is not selective for a specific member of the Aurora family, it does favour the Aurora family over other, similar kinases, [28] with only FLT3 kinase activity inhibited with an IC_{50} of 30 nM.[87] Dual inhibition of the Aurora family and FLT3 may even be beneficial in certain cancers, such as Acute Myeloid Leukaemia (AML), which can also be driven by FLT3 activity.[87] While VX-680 is a potent inhibitor of the kinase activity of the Aurora family, it does not affect the interaction between Aurora A and N-myc.[31]

13
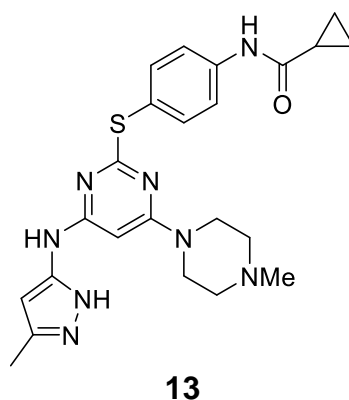
Figure 1.17 Structure of VX-680

In vivo studies showed promising results in mice with human AML xenograft and appeared to be well tolerated.[85] However the compound did not progress beyond phase II clinical trials due to a prolonged QT interval observed in one patient.[88]

1.7 Aims of the Project

The project aims to identify novel inhibitors of Aurora A which may also be capable of disrupting the Aurora A/N-myc complex using several potential target sites (Figure 1.18). The sites include the ATP binding pocket, the kinase salt bridge, a site directly involved in the Aurora A/N-myc protein-protein interaction and two sites identified by the Institute of Cancer Research (ICR). These options will be narrowed down by various methods to focus on the most promising target sites.
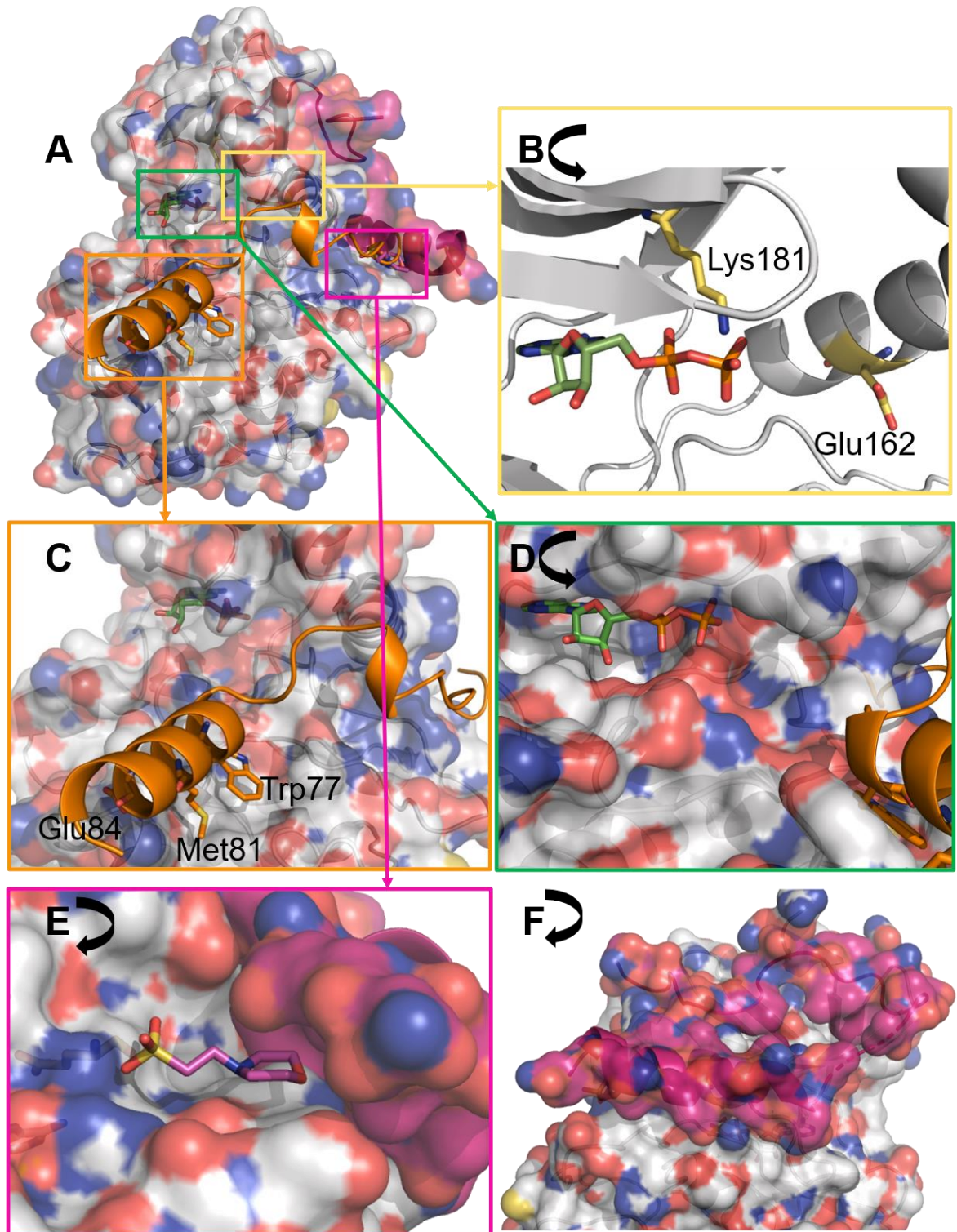
Figure 1.18 The potential target sites for small molecule interaction with Aurora A; (A) The crystal structure of Aurora A (grey) and TPX2 (purple) (PDB 5LXM)[24] with N-myc (orange) overlaid from the crystal structure of Aurora A bound to N-myc (PDB 5G1X);[45] (B) The broken salt bridge between Glu162 and Lys181 (pale yellow sticks) from the crystal structure of Aurora A with vNAR-D01 (PDB 5L8L)[48] to stabilise the inactive conformation; (C) Crystal structure of Aurora A (grey) interacting with N-myc (orange). Only one of the α-helical regions of N-myc was present in the crystal structure; (D) The ATP binding site of Aurora A with ADP bound (green sticks); (E) Crystal structure of Aurora A (grey) and TPX2 (purple) with the fragment MES bound (pink sticks); (F) Closer view of TPX2 bound to Aurora A

1.7.1 Salt Bridge Site (Type III)

The salt bridge formed between residues Lys162 and Glu181 is important for the active conformation of Aurora A and in the crystal structures of Aurora A with various inhibitors this salt bridge is broken.[26, 31, 49–54, 62,67] A compound which prevents the formation of the salt bridge might stabilise the inactive conformation of Aurora A which may be useful for neuroblastoma treatment and would act as a type III kinase inhibitor if it is non-competitive with regard to ATP. Computational methods could help to narrow down the search for a starting point here.

Type III kinase inhibitors are generally believed to be able to achieve greater selectivity between kinases as they do not rely on binding at the highly conserved ATP binding site.[21] This would be a major advantage of pursuing this approach, but to answer the question of whether the ATP binding site is more highly conserved than the area surrounding the salt bridge, the most similar kinases to Aurora A will be analysed for conservation of the surrounding amino acids in an attempt to justify this claim.

1.7.2 ATP Site (Type I)

The Aurora A ATP binding site is already known to bind small molecules, including existing ATP-competitive type I inhibitors such as CD532 **(12)** (Figure 1.13),[31] and VX-680 **(13)** (Figure 1.17). Some of these inhibitors are selective for Aurora A and have been shown to induce a conformational change to cause some disruption to the Aurora A/N-myc complex. These existing inhibitors could be used as a starting point to develop compounds with a greater effect on the Aurora A/N-myc interaction, hopefully whilst maintaining the selectivity over other kinases.

Whilst these existing inhibitors might provide a useful starting point for further development, there is some toxicity associated with current ATP-competitive Aurora A inhibitors.[81] For this reason, and to identify potentially more novel compounds, computational methods can be employed to investigate other possible starting points for inhibitor design.

1.7.3 MES Site (Type IV)

Crystallographic experiments conducted at ICR revealed a fragment (MES, morpholino ethanesulfonic acid, **(14,** Figure 1.20) which appeared to show some affinity for the Aurora A/TPX2 complex (Figure 1.18E). This interaction is yet to be validated with biophysical data and the importance of this site in disrupting the Aurora A/N-myc complex for neuroblastoma treatment remains to be proven. However, overlaying N-myc on this crystal structure shows a clash with the MES binding pocket (Figure 1.19), indicating N-myc binding may be incompatible with a small molecule inhibitor in the MES pocket. Additional crystallographic experiments could be supplemented by fragment growth programs (such as SPROUT[89] or OpenGrowth[90]) or virtual screening of compound libraries to develop MES into a selection of drug-like compounds.
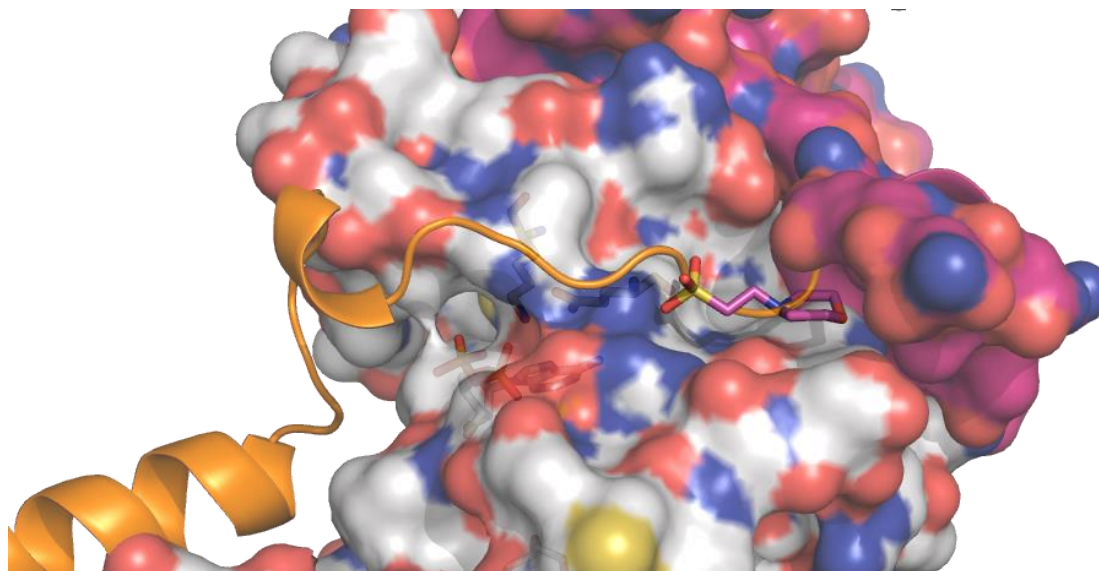
Figure 1.19 Aurora A/TPX2 (grey and purple) crystal structure with MES bound (pink sticks). N-myc (orange) is overlaid to demonstrate the clash with MES.

Alternatively, fragment and compound libraries can be screened virtually at the same site as MES to look for alternative starting points at this site. The same libraries can be analysed for similarity to MES to identify whether any MES-like fragments and compounds are predicted to bind strongly at this site.
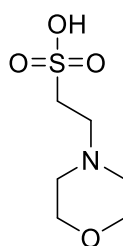
14

Figure 1.20 The structure of the fragment MES, identified at ICR

1.7.4 TPX2 Site (Type IV)

Building from work at the MES site, a key cavity involved in TPX2 binding will be explored (Figure 1.21). Fragment and compound libraries will be docked to identify those predicted to bind. Fragments from the TPX2 and MES sites could be joined in an effort to improve binding strength and selectivity, hoping to induce a conformational change to prevent N-myc binding. TPX2 is also involved in activation of Aurora A so targeting this interaction could affect the activity of Aurora A.[47]
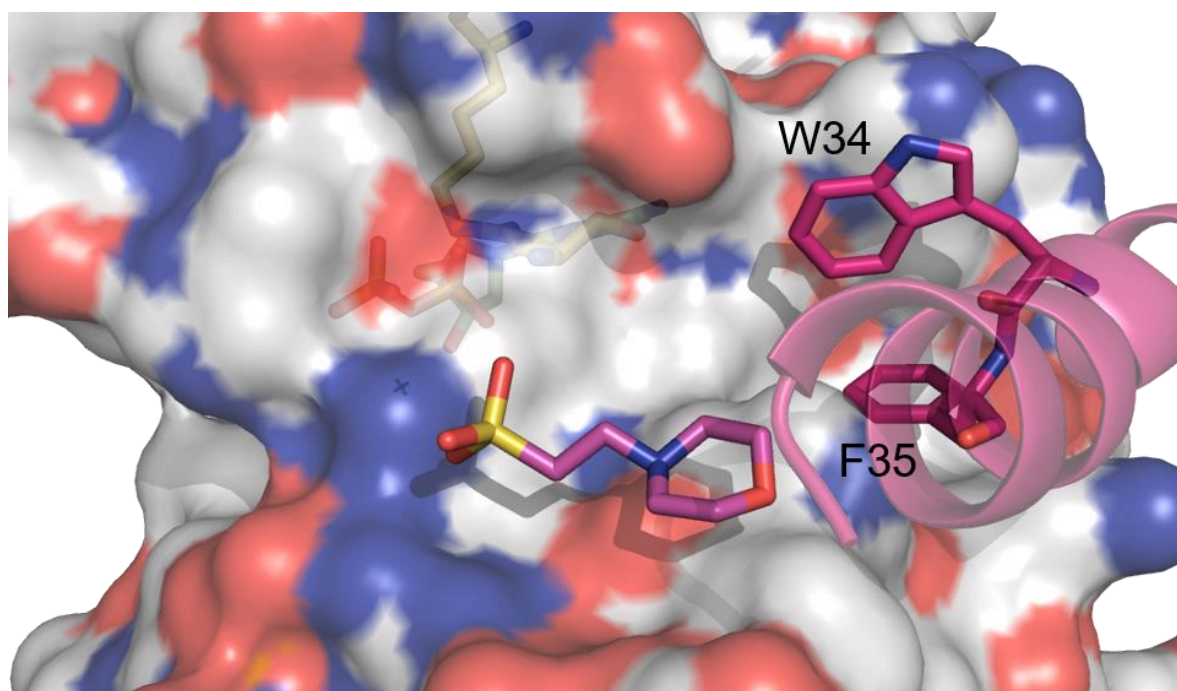
Figure 1.21 The TPX2 site explored in this project, with MES overlaid for reference (pink sticks). The key TPX2 side chains (purple) defining the site are highlighted (Trp34 and Phe35).

1.7.5 N-myc Site

Directly blocking the protein-protein interaction at the N-myc site might be possible in a number of ways. First, Trp77 of the α-helix appears to sit in a shallow cavity in the crystal structure and binds to Aurora A via both hydrophobic and H-bond interactions (Figure 1.18C). Although this does not appear to be a typical binding

pocket for small molecules, docking of fragment and small molecule libraries may provide an idea of the viability of the development of direct Aurora A/N-myc interaction inhibitors. The shallow Trp77 binding pocket can be analysed by SiteMap[91] to provide an idea of how likely this site might be to accommodate a small molecule. In addition, successful soaking of fragments and small molecules into Aurora A crystals could give further support of this approach.

A second strategy could be to design a mimetic of the α-helix based on the key residues of N-myc which interact with Aurora A. A rigid scaffold supporting structures similar to the important side chains could provide good selectivity and promote N-myc degradation by preventing binding to Aurora A.[92–94]

The crystallised N-myc peptide could also be used as a basis to search within compound libraries for anything with some similarity to the key N-myc side chains. Alternatively, curated peptidomimetic libraries, such as the Enamine α-helix mimetic library,[95] could be virtually screened against the Aurora A crystal structure to help in the selection of compounds which may interact with Aurora A in this area.

1.7.6 Identifying Compounds

Where possible, virtual screening conditions for each site were validated by redocking co-crystallised ligands into their crystal structures to provide confidence in docking results. Docking was performed using Glide[96] (Schrödinger), alongside which a prediction of "druggability" at each site by small molecules was obtained using SiteMap[91] assessment (Schrödinger). Various virtual libraries were screened at different sites, including diversity-oriented,[97] fragment[98,99] and peptidomimetic libraries.[95,100]

Where appropriate, other computational methods were used, for example ROCS (Rapid Overlay of Chemical Structures) screening of crystallised molecules to identify compounds within screening libraries with similar shapes and electronics. Alternatively, fragment growth programs such as SPROUT[89] were used to explore some options for developing fragments into larger, more drug-like molecules.

Computational work guided the selection of sets of compounds to test *in vitro*, with these results providing further validation of the docking models, which helped inform iterative cycles of virtual and *in vitro* testing. One possible workflow is shown (Figure 1.22).

Figure 1.22 The potential workflow for any hits found through virtual screening

High throughput crystal soaking used Aurora A crystals, possibly prepared in complexes with other binding partners (for example vNAR-D01), into which many possible inhibitors can be soaked and analysed using X-Ray crystallography to identify compounds which bind to Aurora A in the expected binding sites. The availability of each site for soaking in fragments is affected by crystal packing and the conformation of the protein.

For example, the crystal packing arrangement with Aurora A alone exposes the face required to probe the TPX2 binding site, but obscures the N-myc binding site. When Aurora A is co-crystallised with TPX2 the crystal packing is altered to expose the N-myc binding site. This co-crystal structure also creates the MES binding site. Co-crystallising Aurora A with vNAR-D01 stabilises the inactive conformation of the kinase with the salt bridge broken and available for crystal soaking.

Crystallography experiments would give information about the binding pose of fragments, but would not reveal the effects of fragments and small molecules on the kinase activity of Aurora A. To probe this second question, the Caliper mobility shift assay[101] was used to quantify Aurora A kinase inhibition. This could be used either to screen large numbers of compounds at a single concentration or to generate IC_{50} values of individual molecules. The assay can also be used to determine whether compounds compete for binding with ATP, which will be especially important in the development of type III inhibitors.

To explore the question of whether compounds disrupt the Aurora A/N-myc protein-protein interaction, previous work has made use of fluorescence polarisation assays[45] to assess the effect of CD532 and other type I inhibitors of Aurora A on the Aurora A/N-myc interaction. AlphaScreen has also been explored, which uses singlet oxygen to detect the presence or absence of an interaction between two entities.[102] In this case, the effect of the addition of small molecules on the extent of Aurora A/N-myc interaction could be observed. Finally, K_d values can be calculated using NMR experiments, including STD-NMR. Some work was in this project was impacted by Covid-19.

Chapter 2 N-Myc Site

Perhaps the most direct method of blocking the interaction between Aurora A and N-myc would be to use a competitive inhibitor of N-myc which binds in the same region as the helical portion of N-myc. A competitive inhibitor used in this way could destabilise the interaction, which would result in the Fbxw7-mediated degradation of N-myc.[103] This approach might be achieved through traditional small molecule inhibition or from a designed α-helix mimetic.

Studies suggest the Aurora A Interaction Region (AIR), the key section of N-myc which are responsible for Aurora A binding is between residues 28-89. The Aurora A/N-myc crystal structure used for docking (PDB 5G1X) contains only part of this region (residues 61-89), but this has been shown to be sufficient for binding to Aurora A, albeit with a lower binding affinity than the full AIR.[45]

2.1 Binding Site Analysis

Visual inspection of the Aurora A/N-myc interaction interface indicated the three key residues of the N-myc α-helix involved in binding to Aurora A were Trp77, Met81 and Glu84, which correspond to the *i*, *i+4*, and *i+7* side chains which project in the same direction from an α-helix (Figure 2.1). Trp77 appeared to sit in the most promising looking potential binding site for small molecule of the three side chains, although by eye this was not immediately convincing as a binding site for small molecules.
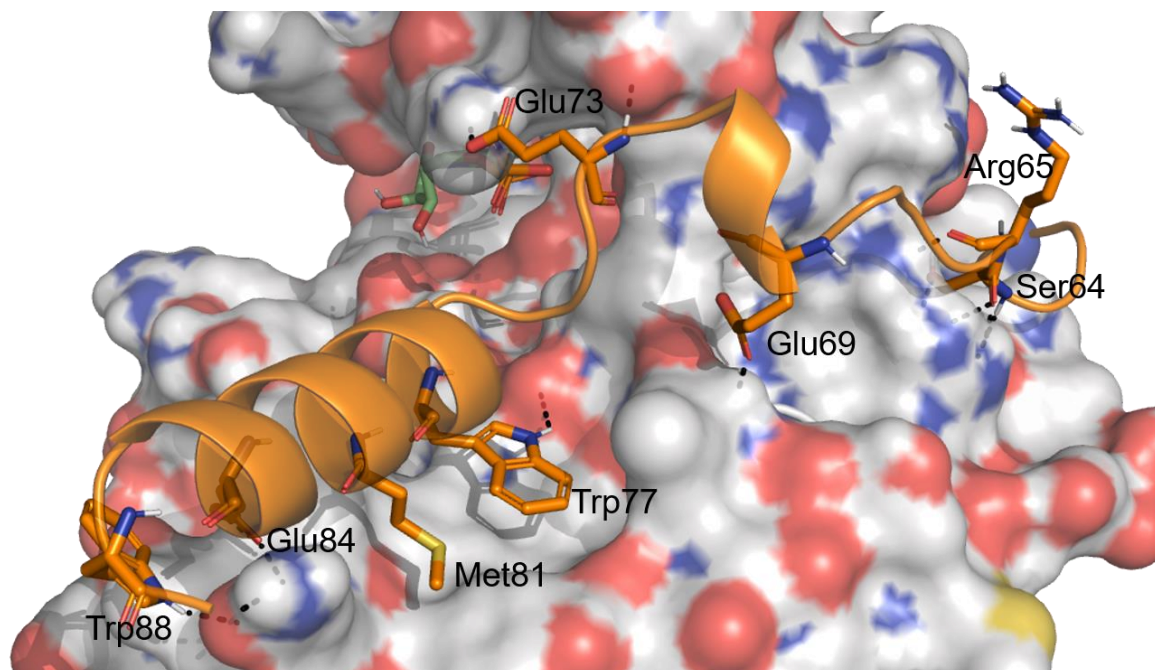
Figure 2.1 Predicted interactions between Aurora A (grey) and N-myc (orange) (PDB 5G1X)[45]

To add to the visual inspection of this area, a prediction of the "druggability" of the Trp77 binding pocket was generated using SiteMap (Schrödinger).[91] SiteMap analyses potential binding sites by using "site points" spaced 1 nm apart within the binding area. Each site point is analysed for its potential to be involved in interaction with a ligand, by assessing properties such as the exposure to solvent and proximity to the protein. Site points which could contribute to binding can be linked together to highlight a potential binding site.[104]

SiteMap also generates a SiteScore based on this analysis. SiteScores are calculated by taking into account the size of the site, depth of the pocket, hydrogen bond formation capability and the balance between hydrophobic and hydrophilic regions,[104] incorporated into the equation:

$$SiteScore = 0.0733\sqrt{n} + 0.6688e - 0.20p$$

Where n is the size (number of site points), e is the overall score for solvent exposure and protein enclosure and p is the extent of hydrophilicity, with a cap to

avoid very high scores for the most polar sites. A SiteScore above 0.8 indicates the area is a binding site for small molecules, whilst a SiteScore below 0.8 means the site is unlikely to bind ligands.[91]

A second score, the Dscore or "druggability" score which is designed to focus on the ability of the site to bind a drug-like ligand, as opposed to any kind of small molecule. The Dscore is calculated in a similar way to the SiteScore, with the main difference of the removal of the limit on the hydrophilicity value. This means the requirement for highly charged ligands to bind highly charged sites, which may be unfavourable drugs, is penalised. The Dscore equation is:

$$Dscore = 0.094\sqrt{n} + 0.60e - 0.324p$$

Where n, e and p each take the same properties as SiteScore.[91] In the case of Dscore, a value above 0.98 is "druggable", 0.83-0.98 means a site is "difficult" and below 0.83 is "undruggable".[91]

Analysis of the Trp77 binding site by SiteMap from the Aurora A/N-myc crystal structure (PDB 5G1X) calculated SiteScores of 0.525 and 0.536 with and without water, respectively. These scores are both much lower than the minimum score of 0.8 for a "druggable" binding site suggested by Schrödinger. Dscores were similarly disappointing, at 0.457 and 0.502 with and without water, again much lower than even the "difficult" druggability score of 0.83.

2.2 Docking

Despite the low scores from computational binding site analysis, libraries of compounds were screened virtually to further explore the feasibility of using small molecule inhibitors to directly block the Aurora A/N-myc interaction (Table 2.1). The docking libraries used in the project included fragment libraries (Asinex[99], Bionet[98] and Maybridge), larger compound libraries with a focus on diversity (MCCB and Chembridge[97]) and α-helix mimetic libraries (Enamine[95] and ChemDiv[100], used only at the N-myc site). Overall, around 180 000 compounds were available for virtual screening.

Especially for the α-helix mimetic libraries, the goal would be to mimic the three amino acid side chains of N-myc which project towards and interact with Aurora A. In the case of N-myc those side chains are Trp77, Met81 and Glu84, making up the *i*, *i+4*, and *i+7* side chains which project in the same direction from an α-helix (Figure 2.1). Currently, no helical peptide based inhibitor of the Aurora A/N-myc protein-protein interaction has been reported in the literature.

Table 2.1 Compound libraries used for virtual screening

Library	Library Type	Number of Compounds (approx.)
Asinex[99]	Fragments	21000
Bionet[98]	Fragments	1100
Maybridge	Fragments	1000
MCCB	Drug-like compounds	30 000
Chembridge DIVERSet[97]	Diversity	100 000
Enamine[95]	α-helix mimetics	14 000
ChemDiv[100]	Peptidomimetics	14 000

At the N-myc site libraries of compounds were first docked using Glide (Schrödinger) with the coordinates of Trp77 as the centre point for docking grid generation, both with and without crystallised water molecules. All seven compound libraries were screened using this docking grid. Docking results were ranked by Glide score, which takes into account properties including hydrogen bonding, distances between ligand and receptor and bond angles. The ranked compounds were then filtered to remove any PAINs scaffolds. These included Michael acceptors, quinones, rhodanines, and others mentioned in the literature (Figure 2.2).[105–108]

| Enone | Ene-rhodanine | Quinone |
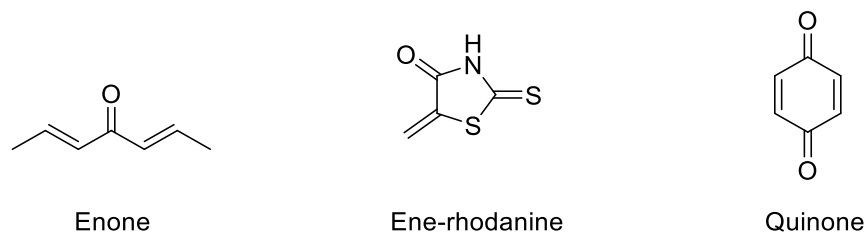
Figure 2.2 Some of the PAINs structures from docking results

2.2.1 Trp77 Grid Centre

With the water molecules included, a good proportion of fragments appeared to show good overlap with the Trp77 side chain and mimicked the hydrogen bond interaction identified between the tryptophan side chain and the carbonyl of Gly291 of Aurora A. A number of fragments were also predicted to form hydrogen bonding interactions with the backbone of Leu293 (Figure 2.3).

Glide score: -6.181
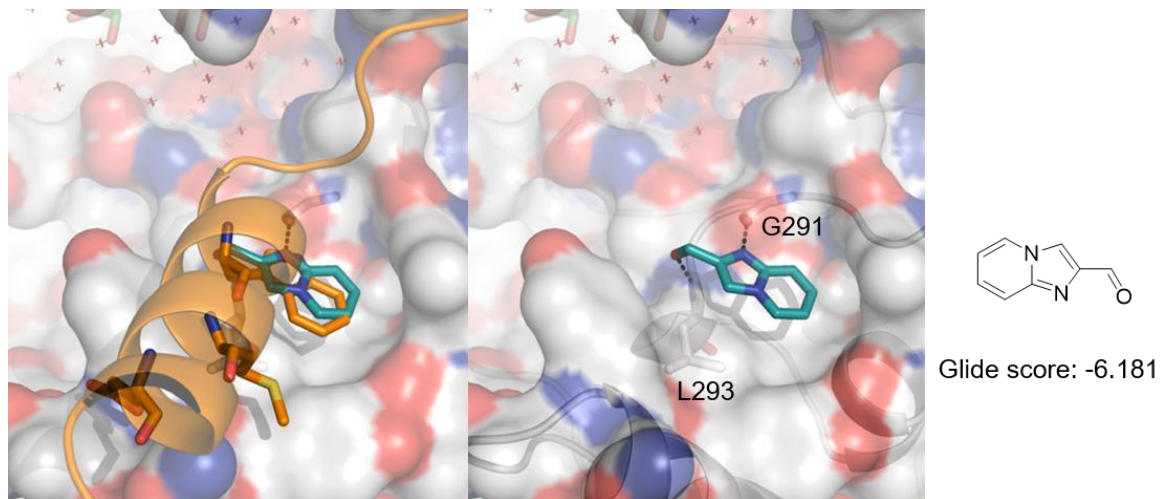
Figure 2.3 A fragment from the Maybridge library (compound ID CC31304) showing a docking pose with some overlap with the Trp77 side chain of N-myc (orange) and predicted hydrogen bonding interaction with Leu293 of Aurora A (grey, PDB 5G1X)

Without crystallised water molecules in the docking grid, fragments were able to dock in positions closer to the ATP binding site, often with predicted interactions with magnesium ions and surrounding Aurora A residues (Figure 2.4). Under these

docking conditions, Glide scores were much higher, a possibly unsurprising result considering the predictions from the docking poses of more hydrogen bonding interactions between high-scoring fragments and Aurora A amino acids. Unfortunately, with the crystallised water excluded, fragments were not predicted to overlap with the key side chains of N-myc.
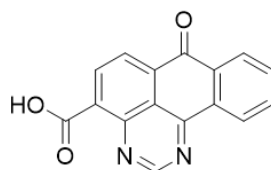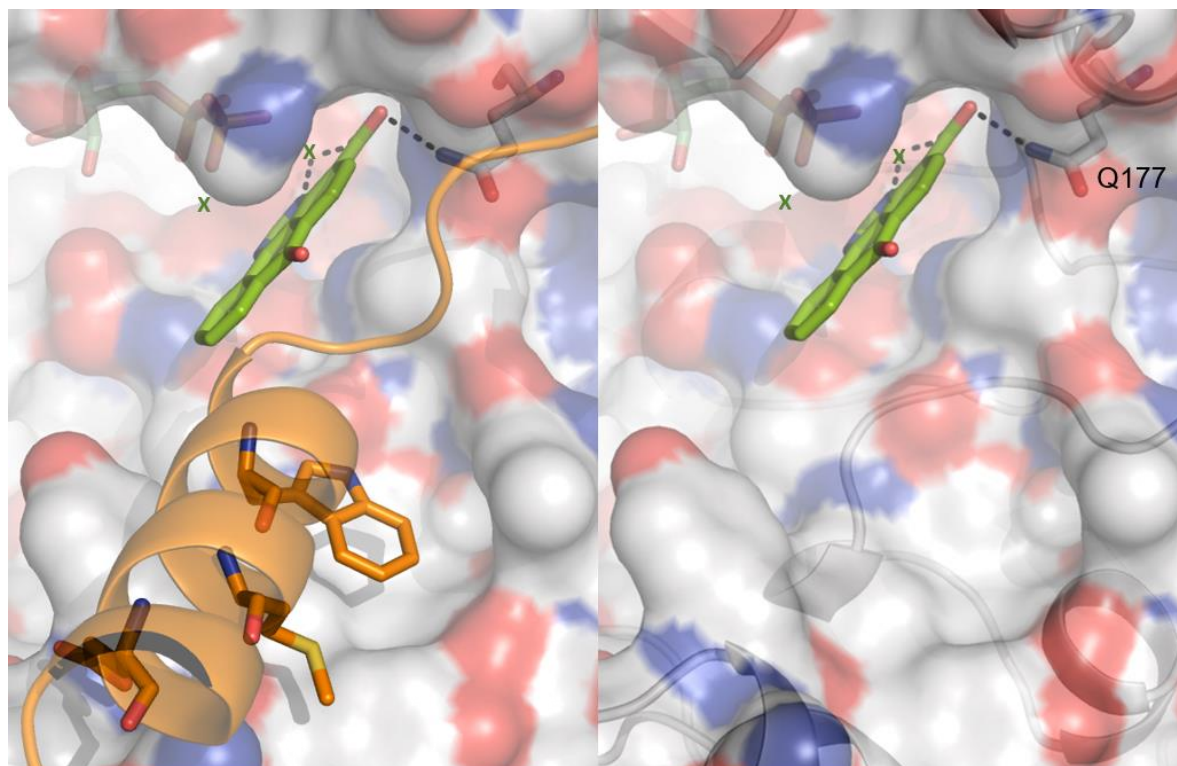
Glide score: -10.400

Figure 2.4 One of the highest ranked fragments by Glide score (compound ID BAS 00368055, Asinex) using the Trp77-centred Glide grid, showing predicted interactions with Mg^{2+} ions (green cross) with N-myc shown (orange, PDB 5G1X)

Larger compounds docked using N-myc Trp77 as the docking grid centre tended to extend away from the N-myc helix binding region and towards the ATP binding

site, as opposed to the desired mimicking of the N-myc helix side chains. This result was observed regardless of the presence or absence of crystallised water. As with the fragment libraries, docking the larger compound libraries with a Trp77 grid centre also generated predicted poses in which there was often a good overlap between the aromatic side chain of Trp77 and an aromatic ring in the docked compounds. Unfortunately, no compounds were predicted to mirror all three of the desired N-myc helix side chains (Figure 2.5). A few compounds were also predicted to reach into the kinase substrate binding site, providing another alternative to the desired mimicking of the three key N-myc side chains (Figure 2.6).
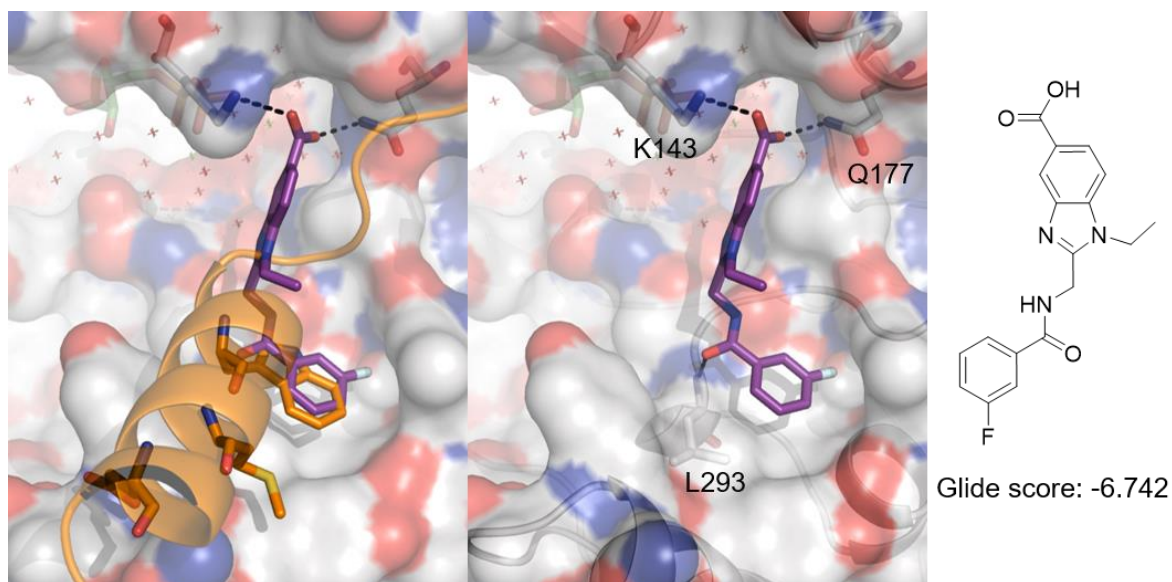
Figure 2.5 A docking pose of a compound (compound ID ALB-H01519346, MCCB) from docking drug-like compounds showing good overlap with the Trp77 side chain of N-myc (orange, PDB 5G1X). Aurora A side chains predicted to form interactions with the compound are labelled.
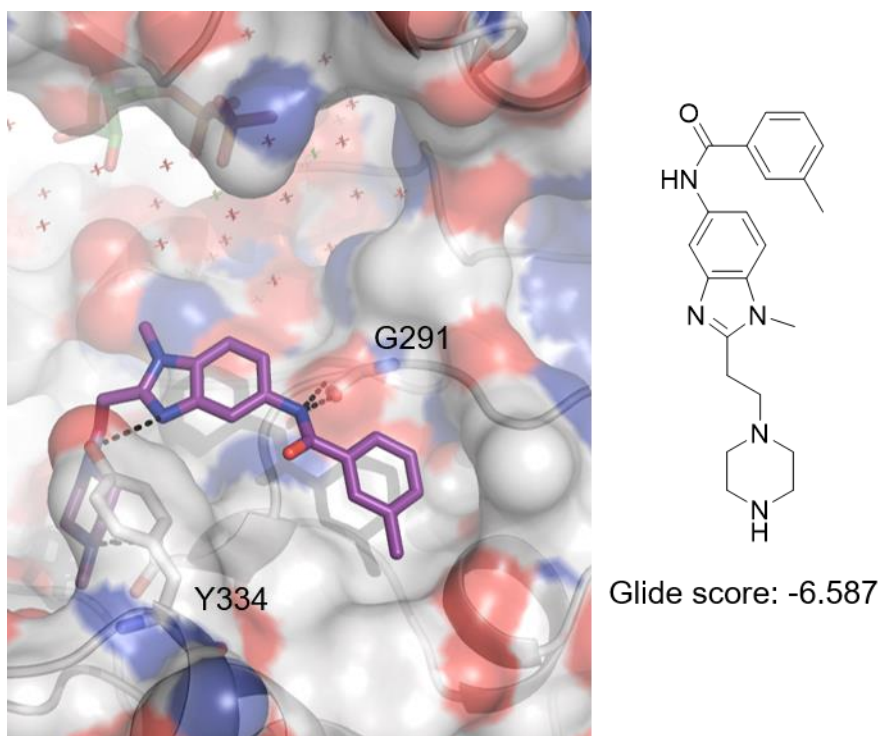
Figure 2.6 One of the larger compounds (compound ID 9107029, MCCB) predicted to extend towards the substrate binding region of Aurora A (grey, PDB 5G1X). Aurora A side chain predicted to form interactions with the compound are labelled.

The final commercial libraries docked at this site were made up of larger compounds with the potential to act as peptidomimetics. An α-helix mimetic of N-myc would be designed to mimic the positions and interactions of the side chains Trp77, Met81 and Glu84 as *i*, *i*+4 and *i*+7 positions of an α-helix. Screening libraries containing compounds selected for helix mimetic properties were docked in the same way as the previous compound libraries, using Trp77 as the centre point for generation of the docking grid. The size of the grid was checked to ensure the whole of the N-myc helix was included.

A similar pattern of results was seen with these compounds as was seen for other libraries using Trp77 as the grid centre. These compounds were not predicted to mimic the three key N-myc side chains. At best, a small number showed some overlap with Trp77 and Met81, but none extended to Glu84. Furthermore, as seen

with other virtual screens at this site, many compounds extended towards the ATP and substrate binding sites (Figure 2.7).
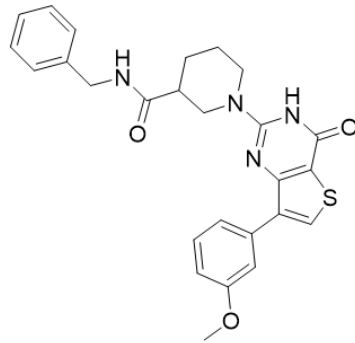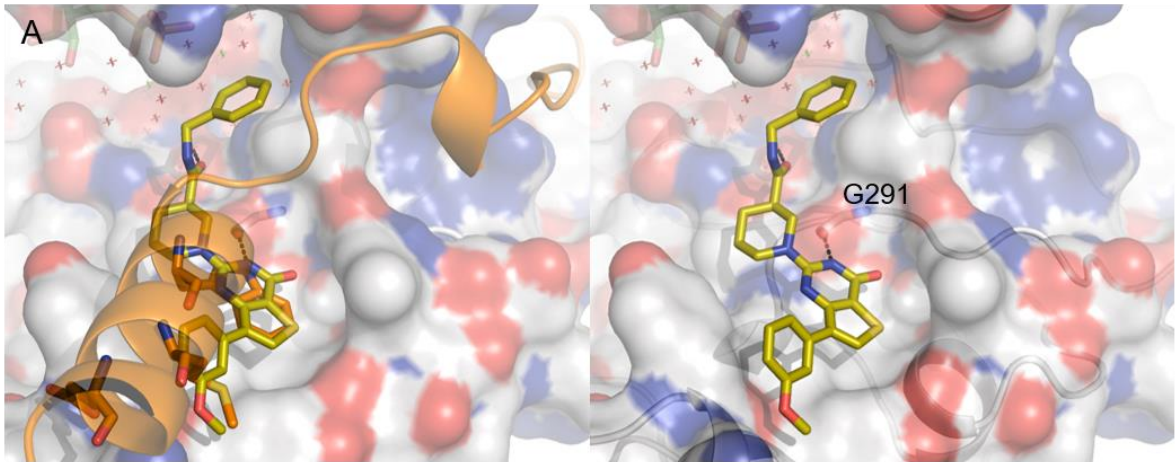
Glide score: -6.533
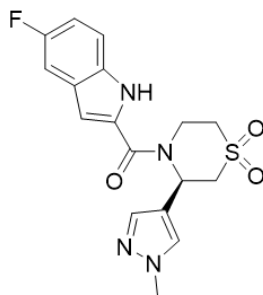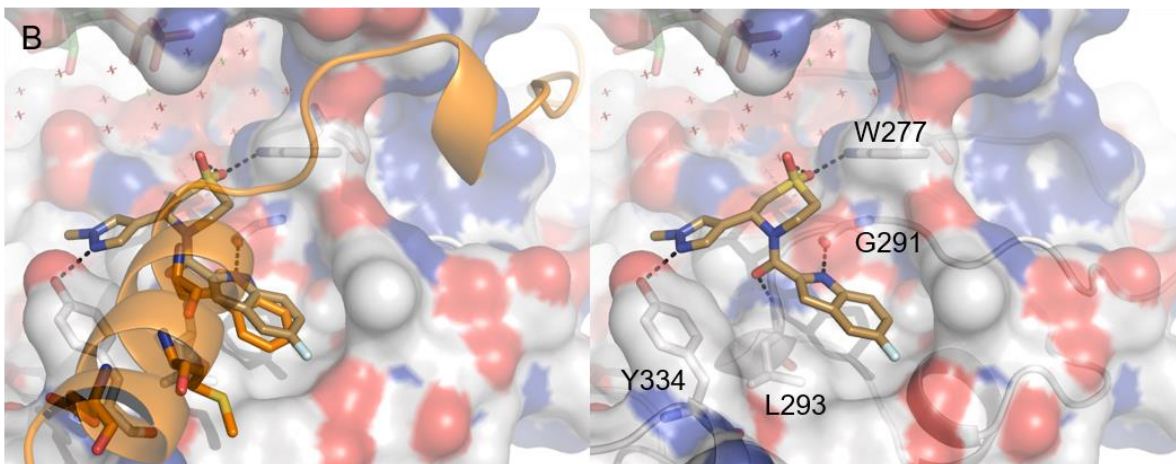
Glide score: -6.973

Figure 2.7 Docking results using peptidomimetic compound libraries using the
Trp77-centred Glide grid; (A) One of the docking poses (compound ID Z250-
1612, ChemDiv) with some overlap with N-myc side chains Trp77 and Met81
(orange); (B) The more common type of pose generated from these libraries,
showing one compound (ID Z1757451390, Enamine)extending towards the
ATP binding site of Aurora A (grey, PDB 5G1X). Amino acids predicted to
form interactions with the compounds are labelled.

With the exclusion of crystallised water, fewer compounds showed any overlap
with the side chains of N-myc. Instead, as seen with fragment libraries, compounds
seemed to favour binding close to the ATP binding site, often predicting
interactions with magnesium ions and surrounding residues and generating higher
Glide scores.

Using Trp77 as the centre point for grid generation in the hope of finding
compounds which mimic N-myc interactions meant the dimensions of the grid had
to be large enough to encompass the whole of the helix. With Trp77 as the centre
point, this meant the grid also extended to cover a portion of the ATP binding site.
If the centre of the docking grid was changed to Met81, the second of the key N-
myc residues, the docking grid dimensions could be altered in the hope of avoiding
the generation of poses extending towards the ATP binding site and improving the
chances of finding compounds which overlap with the helix side chains as desired.

2.2.2 Met81 Grid Centre

The helix mimetic libraries were docked using the Met81-centred docking grid, a
total of 28000 compounds. With Met81 as the docking grid centre compounds were
again predicted to approach the ATP binding site. However, in contrast to using the
Trp77 centre, more compounds were predicted to overlap with one or two of the
three target residues, usually Trp77 and Met81 (Figure 2.8). Unfortunately, no
compounds were predicted to match all three side chains.
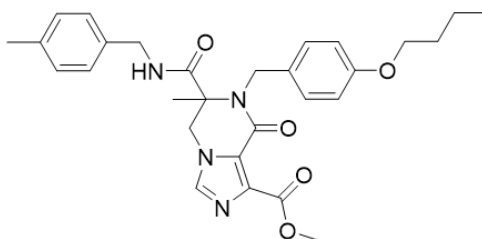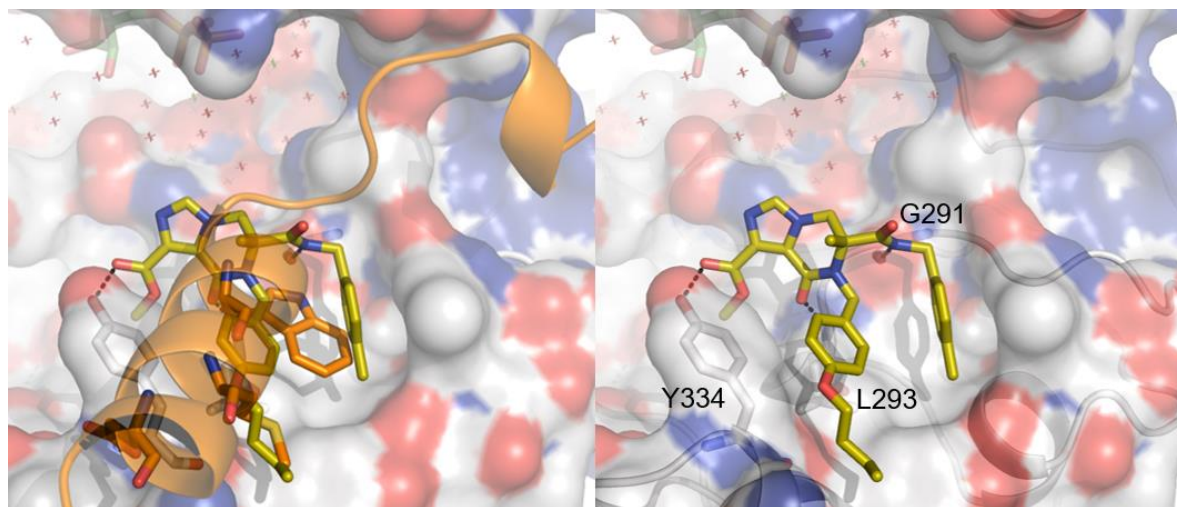
Glide score: -6.693

Figure 2.8 A docking pose from the peptidomimetic libraries (compound ID G389-1190, ChemDiv) using a Met81-centred Glide grid with N-myc overlaid (orange, PDB 5G1X), showing some overlap with the N-myc side chains Trp77 and Met81. Aurora A residues predicted to form interactions with the compound are labelled (grey sticks).

Moving the grid centre to Met81 did improve the proportion of top results following the N-myc helix but did not completely overcome the problem of compounds extending in the direction of the ATP binding site, especially with the removal of crystallised water (Figure 2.9). Reassuringly, Glide scores of the top results were not greatly affected by the change of grid (remaining roughly between -6.5 and -7.0 with water). To further explore the effect of changing the grid centre and avoiding the ATP binding site interactions still observed to an extent, the grid dimensions

were altered once more, using Glu84 as the centre point, even further from the ATP binding site.
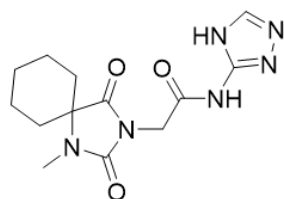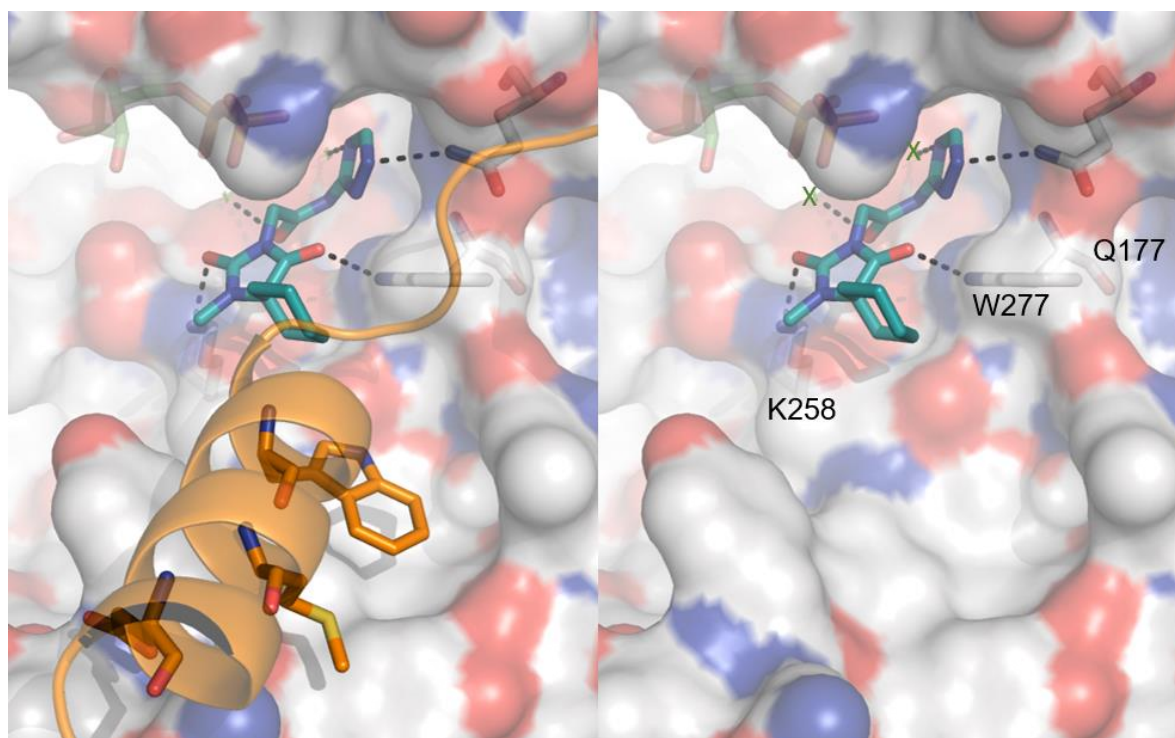
Glide score: -8.800

Figure 2.9 Docking peptidomimetic compounds with the Met81-centred Glide grid without water often generated poses in which compounds extended away from the N-myc helix (orange, PDB 5G1X) and towards the ATP binding site (compound ID Z352538160, Enamine). Interacting side chains of Aurora A (grey) are labelled.

2.2.3 Glu84 Grid Centre

In an attempt to minimise the number of compounds attempting to bind in the ATP binding site and increase the chance of identifying compounds which overlapped

with all three target side chains, Glu84 was used as the centre point for the docking grid. Glu84 is at the end of the N-myc α-helix, furthest from the ATP binding site, so the docking grid could cover the N-myc helix without reaching towards the ATP binding site. The two peptidomimetic libraries were screened at this site, a total of 28000 compounds.

This approach was successful in reducing the proportion of compounds extending towards the ATP binding site but again did not identify any compounds predicted to mimic all three key N-myc side chains. Instead, with a Glu84 grid centre more compounds were predicted to reach into the kinase substrate binding site than either of the other two grid centres (Figure 2.10).
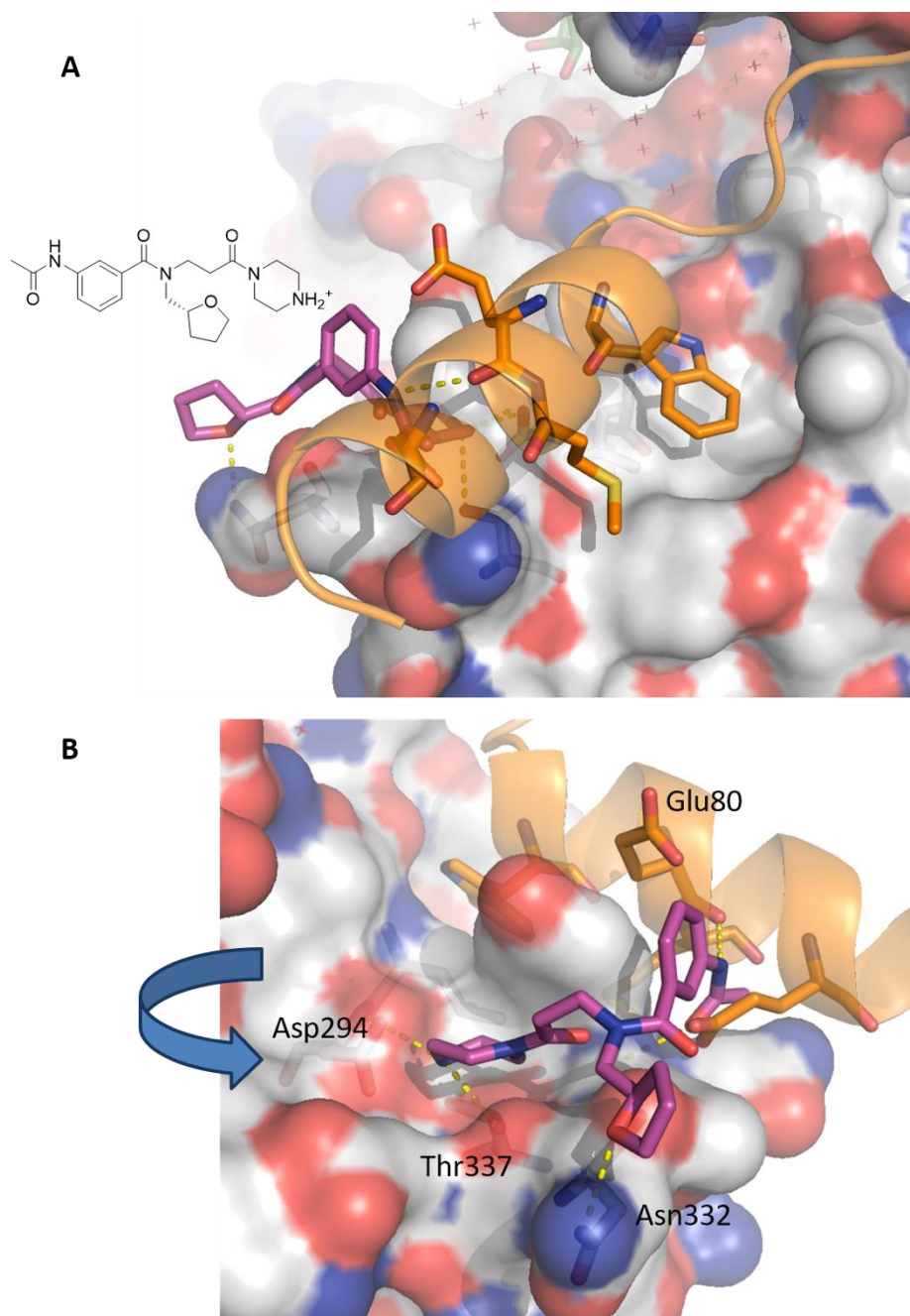
Figure 2.10 Docking results at the N-myc site with Glu84 grid centre; (A) The second result from the ChemDiv helix mimetic library (pink sticks, Glide score -5.403, ID F594-0384) docked with water using with N-myc (orange) overlaid from the Aurora A/N-myc crystal structure (PDB 5G1X); (B) Rotated view of the same docking result; the compound is predicted to extend over the ridge of Aurora A shown and towards the kinase substrate binding site

In addition, Glide scores were also much lower under these conditions, with the highest score at only -5.625. Using Glu84 as the centre point did not give the kind of docking results hoped for with commercial peptidomimetic libraries. An alternative option could be to use designed helix mimetics based more specifically on the N-myc helix in particular.

2.2.4 Helix Mimetic Scaffolds

Published helix mimetic scaffolds in which the required side chains of an α-helix mimetic can replace specific R groups of a designed core (Figure 2.11) were also docked using each of the target side chains as the centre point for docking in the hope of finding a scaffold predicted to bind in the same orientation as the side chains of N-myc. A bespoke virtual library was designed for this purpose, using the published scaffolds and the specific side chains required based on the N-myc α-helix.
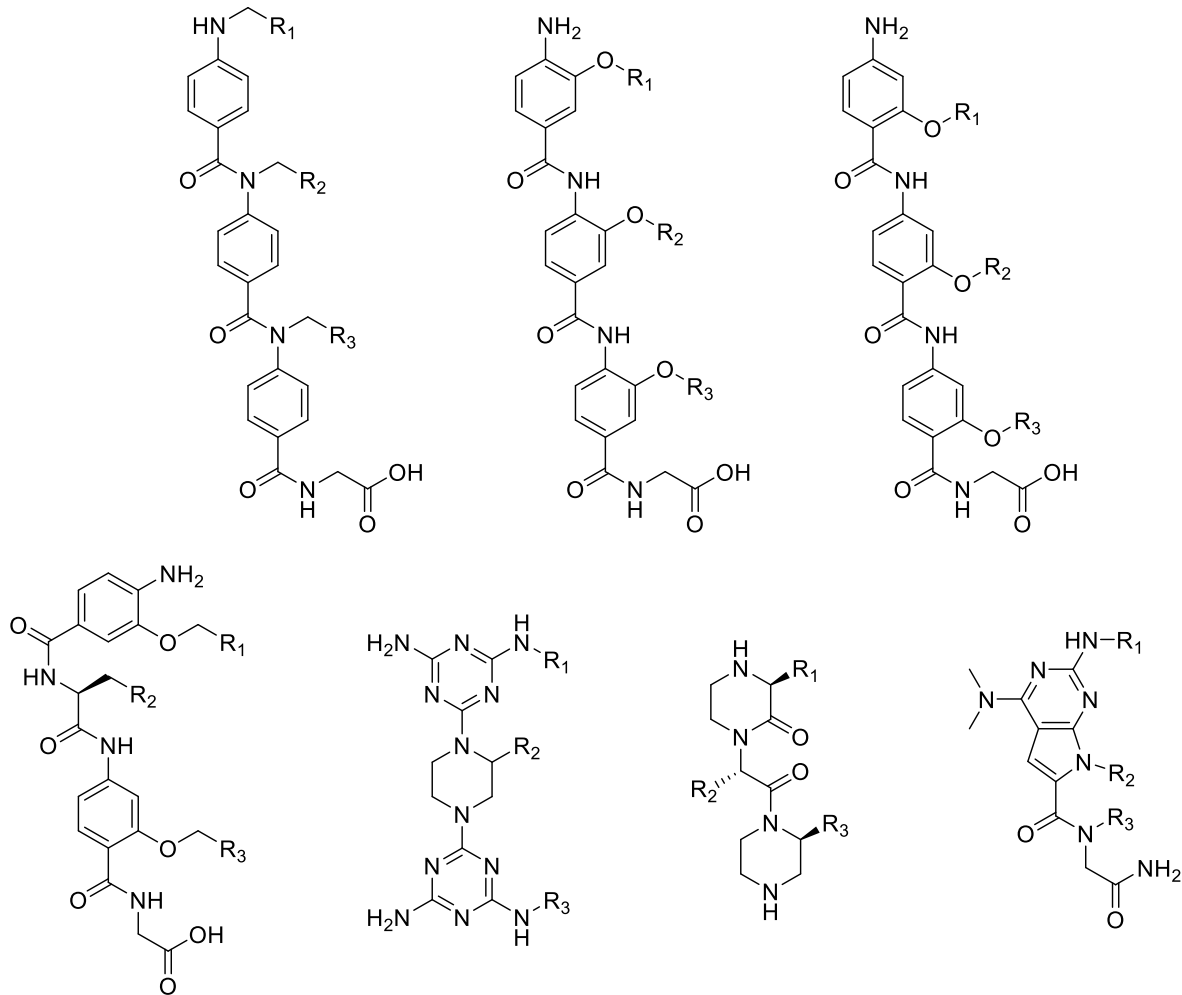
Figure 2.11 Helix mimetic scaffolds used; the groups R1, R2 and R3 were replaced with the side chains of Trp77, Met81 and Glu84, respectively

As with the previously discussed docking experiments, each of the three side chains were used as the centre point for grid generation, resulting in three sets of docking results. Each set of compounds was also screened with and without crystallised water molecules.

With the Trp77-centred grid, some good overlap with the Trp77 side chain was observed, especially in the absence of crystallised water, along with some good Glide scores (between -6.3 and -9.5) for the top results. Unfortunately, these high scores were based on poses in which the compounds extended towards the ATP binding site, the same problem observed with commercial library virtual screening

(Figure 2.12). As seen with other screens at this site, the problem was worse with the removal of crystallised water, and these poses also gave the highest Glide scores.
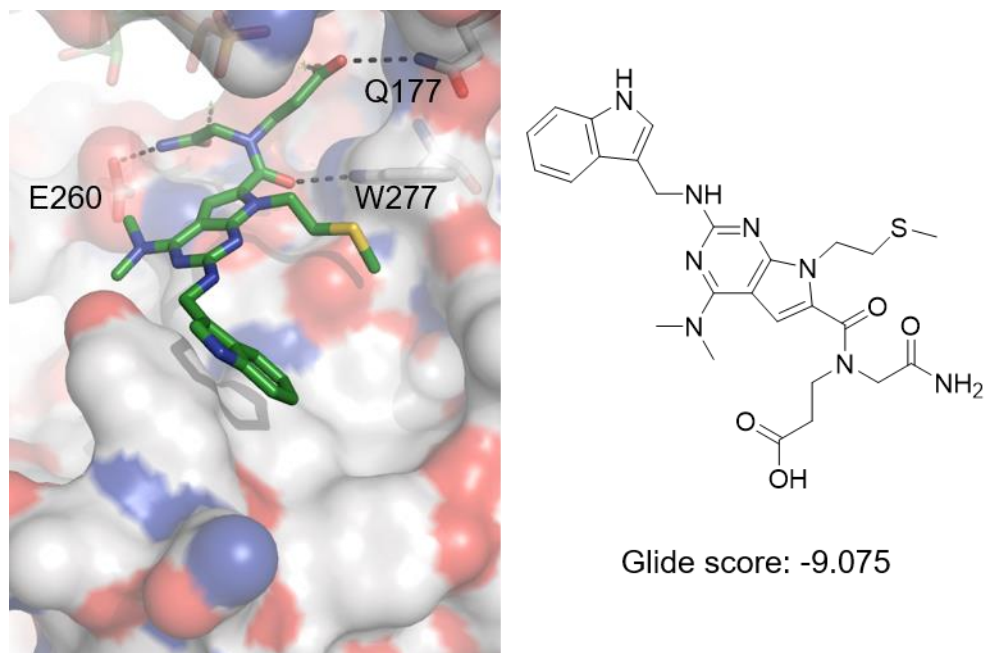
Figure 2.12 One of the highest ranked results (by Glide score) of the set of designed helix mimetic compounds docked with a Trp77 grid centre, showing the compounds extending towards the ATP binding site. Interacting side chains of Aurora A (grey) are labelled.

Moving along the helix and using Met81 as the grid centre produced the same problems observed with the Trp77 grid; the top results all extended towards the ATP binding site. Some aromatic rings were predicted to overlap with the aromatic Trp77 side chain, but the tryptophan side chain added to designed scaffolds did not mirror the N-myc side chain (Figure 2.13).
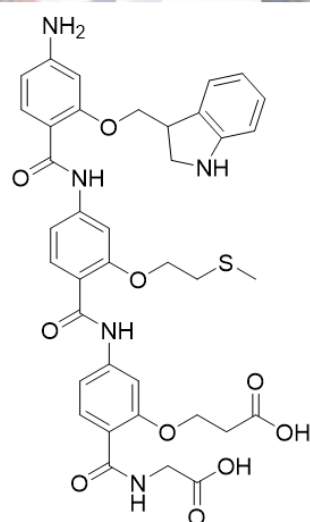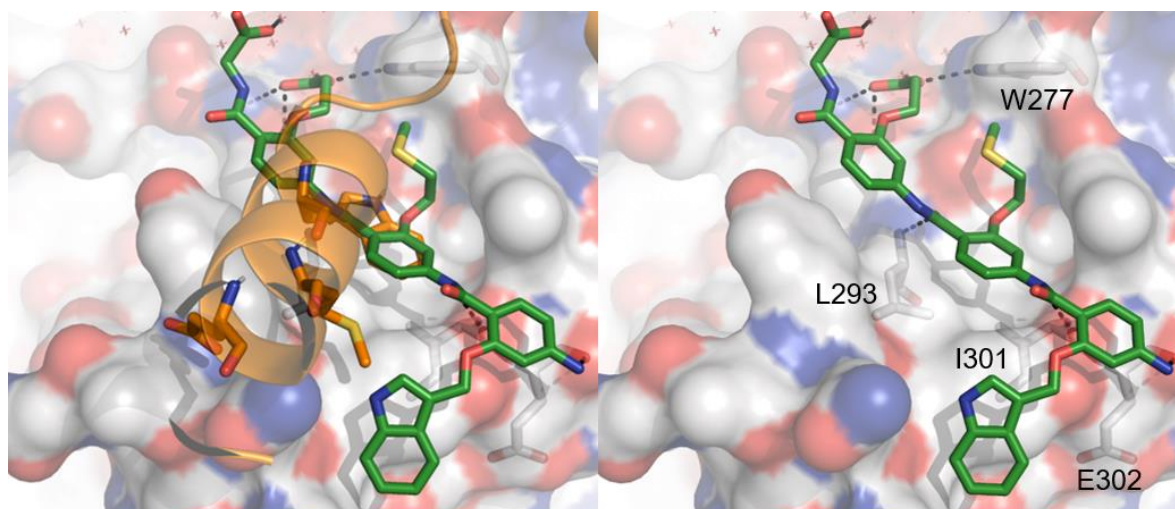
Glide score: -7.512

Figure 2.13 Helix mimetic scaffold docking using Met81 as the grid centre. The N-myc Trp77 side chain (orange, PDB 5G1X) often showed some overlap with aromatic rings from the designed compounds, but this was not with the intended tryptophan-like portion of the designed compound. Predicted interacting amino acids of Aurora A (grey) are labelled.

To complete the docking set, the library of designed scaffolds was also screened virtually against the Glu84-centred grid. Again, no compounds were predicted to mimic all three key side chain of N-myc, although there was a pronounced reduction in the number of compounds extending towards the ATP binding site.

Instead, docking poses showed hydrogen bonding interactions with a variety of other Aurora A residues not seen with previous virtual screens (e.g. K339), but this was accompanied by a significant worsening of Glide scores associated with these poses (Figure 2.14).
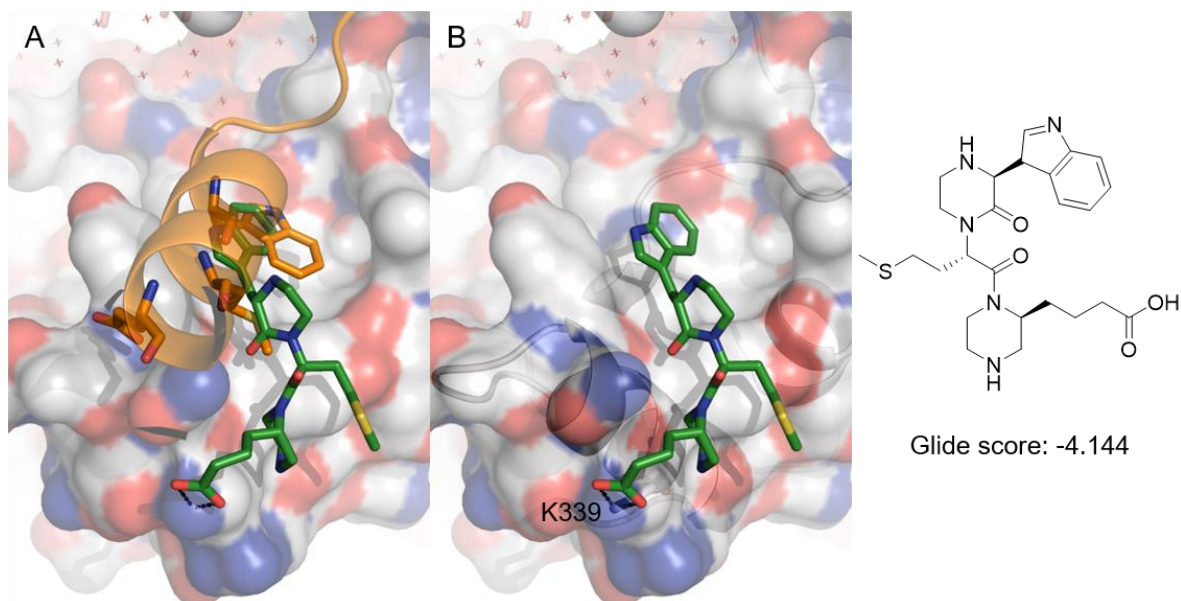
Glide score: -4.144

Figure 2.14 N-myc Glu84-centred docking results from designed helix mimetics, showing a lack of cover of the three key N-myc side chains (orange) and lower Glide scores than other virtual screening conditions

Most compounds were not predicted to mimic the side chains and Glide scores were generally much lower for these compounds than for the helix mimetic libraries. The compound closest to overlapping with all three target residues is shown (**Error! Reference source not found.**). Given the poor predictions from virtual screening, this group of compounds was not pursued further in the project.
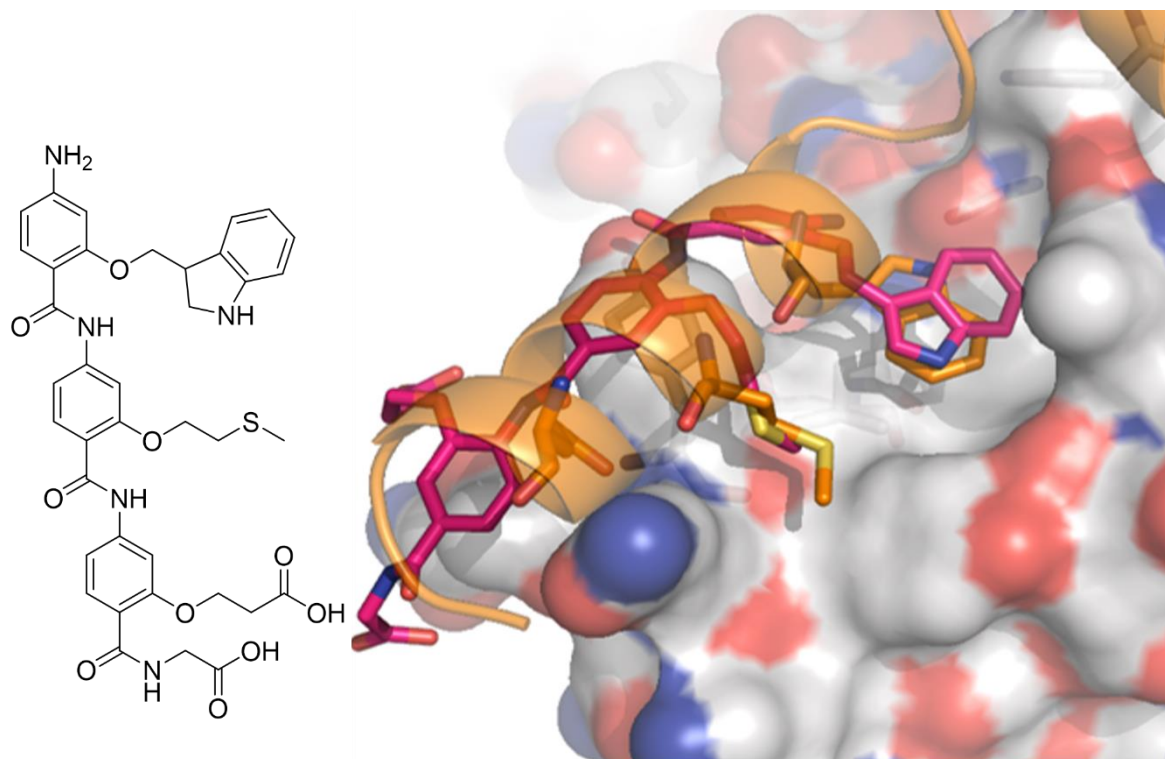
Figure 2.15 Docking of helix mimetic scaffolds at the N-myc site; the 20th result by Glide score (pink sticks, Glide score -3.110) from the library of designed helix mimetics docked at the N-myc site without water with Met81 as the centre point for grid generation

2.3 ROCS

The ligand-based design tool ROCS (Rapid Overlay of Chemical Structures) was used to complement the identification of α-helix mimetic inhibitors at the N-myc site found by Glide docking. ROCS can be used to identify compounds in a given library which have a similar 3D shape to a known ligand.

In a ROCS screen to search for compounds within the Enamine α-helix mimetic library with a similar shape to the N-myc helix (residues 77-84), a number of compounds were identified which closely matched two of the three target residues, usually Trp77 and Glu84, but none were found which matched all three side chains (Figure 2.16).
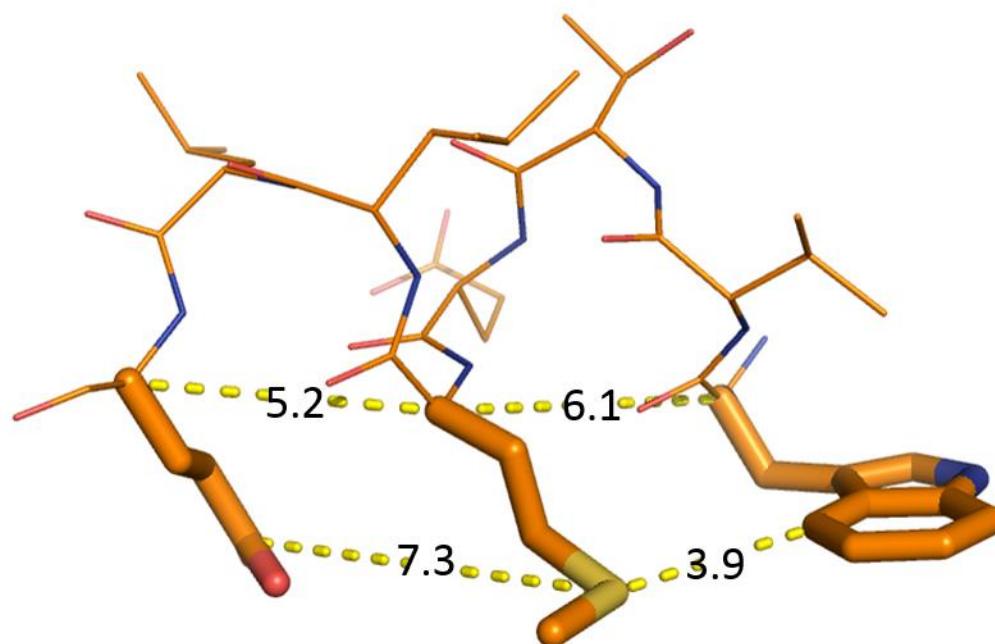
Figure 2.16 The core N-myc helix used for ROCS screening with distances shown
in Å and the three key side chains highlighted

Many compounds overlapped with only one of the target residues and overlaid with
sections of the backbone instead of the key side chains. To avoid these kinds of
compounds dominating the top ROCS results, a second ROCS screen was
performed using only the desired side chains. The distances between each side
chain from the crystallised N-myc helix were measured and the side chains alone
were used, connected by carbon chains the right length to give similar distances
between groups as the crystal structure.

The ROCS screen based on this structure was more successful in finding
compounds which favoured following the side chains over the backbone.
Unfortunately, as with the previous screen, compounds overlaid well with only two
of the side chains at most, with the majority of compounds overlapping with only
one side chain.

2.4 Summary

Targeting the N-myc binding site would be the most direct way to prevent formation of the Aurora A/N-myc protein-protein interaction. However, analysis of the binding site did not support this region as a binding site for small molecules. Docking results from the screening of commercial compound libraries supported the site analysis, with low Glide scores and unconvincing docking poses. Designed mimetics of the N-myc helix could be a route to disrupting the protein-protein interaction, but the predicted binding poses generated did not predict these compounds would occupy the same space as the N-myc helix. Overall, whilst the directness of this approach was appealing, the discouraging computational predictions resulted in this approach being set aside to allow greater focus on other options.

Chapter 3 Type IV Inhibition

Interaction with TPX2 helps to activate Aurora A and trigger mitosis so targeting this site could prevent Aurora A activation. The TPX2 site is defined by the residues Trp34 and Phe35 of TPX2 (Figure 1.21). TPX2 is a substrate and activator of Aurora A. By preventing the activation of Aurora A by TPX2 using small molecule inhibitors binding in the cavity defined by Trp34 and Phe35, the effect of Aurora A kinase activity in cancer could be avoided.[30] This would be especially useful in cancers presenting with Aurora A overexpression.[29,43] Furthermore, preventing the activation of Aurora A by TPX2 could contribute to an inactive conformation of Aurora A which is incompatible with N-myc binding,[47] subsequently leading to N-myc degradation and cell death in neuroblastoma.[45]

Previous crystallography experiments yielded MES, a fragment seen to bind at the interface of Aurora A and TPX2 in a co-crystal structure (Figure 1.21). Overlaying N-myc from the Aurora A/N-myc co-crystal structure[45] indicates a clash between MES and N-myc (Figure 1.19). It might be possible to develop MES into a compound capable of disrupting the Aurora A/N-myc protein-protein interaction.

3.1 Binding Site Analysis

SiteMap analysis of the TPX2 site, using the co-ordinates of the side chains Trp34 and Phe35 from the Aurora A/TPX2 crystal structure (PDB 5LXM) gave a SiteScore of 0.708 with water and 0.573 without water. Like the Trp77 site, these scores were lower than the "druggable" score of 0.8.[91]

At the MES binding site, crystallised MES was used to identify the site in the Aurora A/TPX2 crystal structure, generating SiteScores of 0.828 and 0.798 with and without crystallised water molecules. These scores provided further hope that the MES binding site was "druggable" and it may be possible to develop small molecules capable of binding in this region.

3.2 Docking

Compound and fragment libraries were docked at the TPX2 binding site using the centre of two Aurora A residues (His187 and His280) as the centre point for virtual screening. The key TPX2 residues (Trp34 and Phe35) lie between these two histidines in the Aurora A/TPX2 crystal structure (Figure 3.1). For all TPX2 site screens the Aurora A/N-myc crystal structure (PDB 5G1X) was used with N-myc removed. Three fragment libraries and two drug-like compound libraries were docked at this site, a total of around 152 000 compounds.
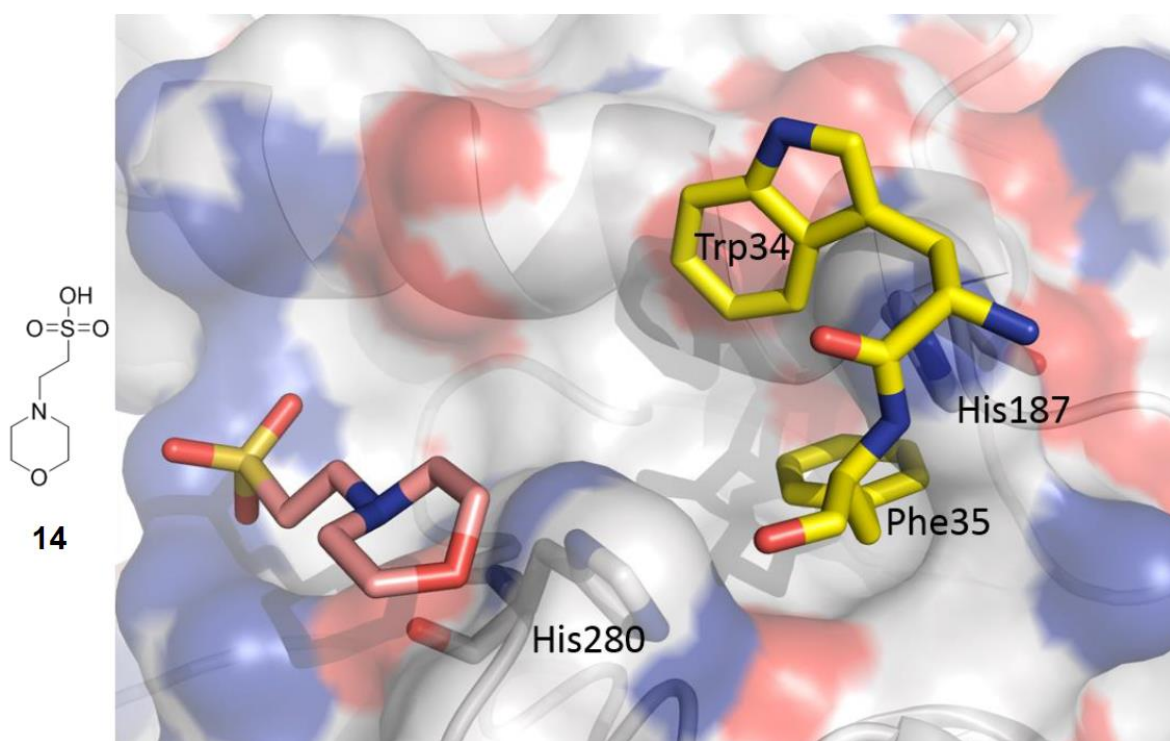
Figure 3.1 The TPX2 binding site: TPX2 residues used to define the site (yellow sticks) with MES (pink sticks) shown for reference and the two Aurora A histidines used as a centre point for docking highlighted (grey sticks, labelled)

A large number of compounds contained cyclic regions (often aromatic) which were predicted to sit in between His187 and His280 and may show opportunities for π-stacking. The vast majority of compounds were then directed towards the

MES binding site, with a large number of diverse compounds from the MCCB library overlapping with MES both with and without water molecules (Figure 3.2).
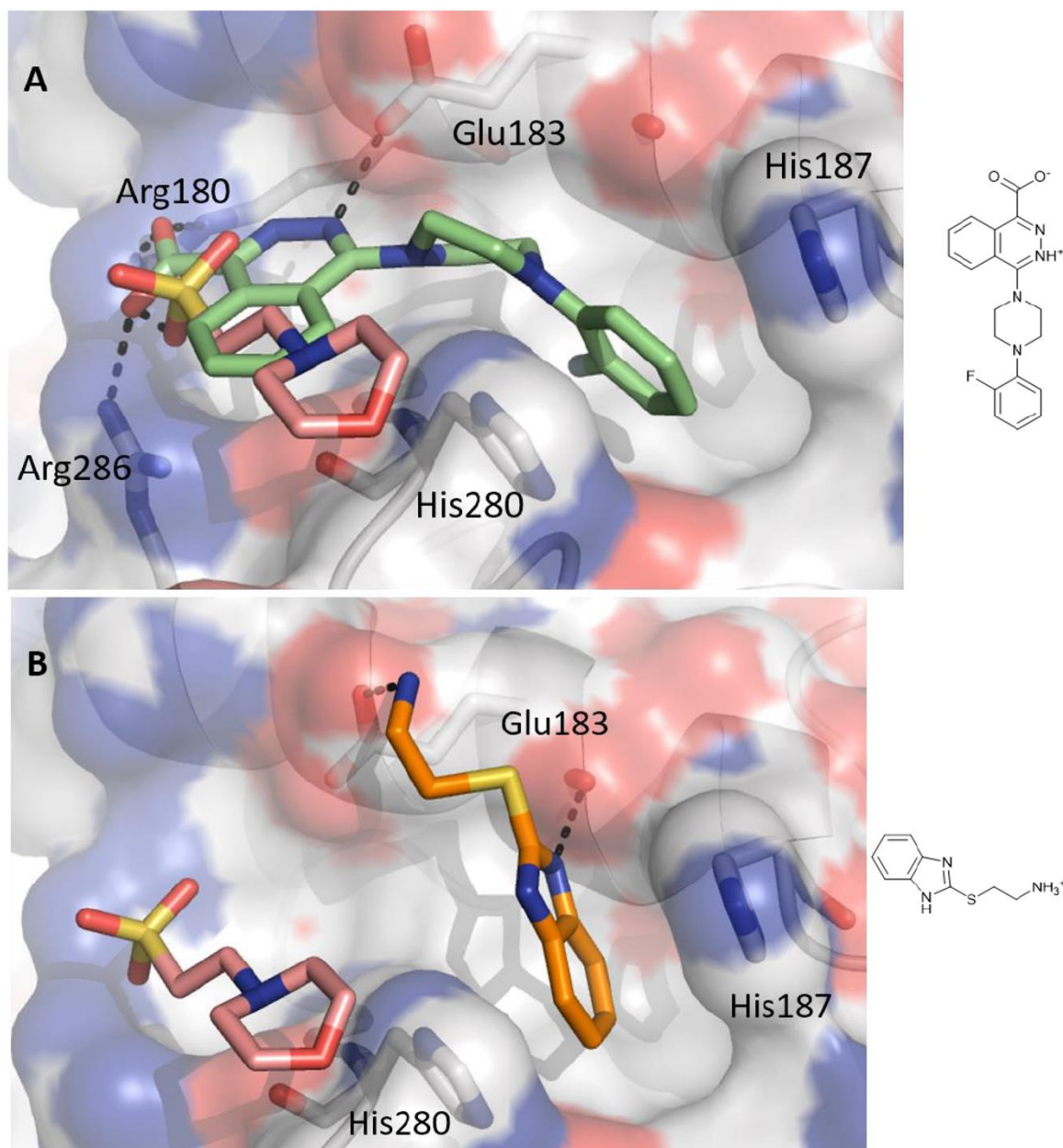
Figure 3.2 Library docking at the TPX2 site without water with MES overlaid as pink sticks; (A) The top result from the MCCB library (green sticks, docking score -6.533, ID ALB-H10304495) is shown reaching from the TPX2 site into the MES site; (B) The top result from the Asinex fragment library (orange sticks, docking score -6.757, ID BAS02937058) is shown extending from the TPX2 site without approaching the MES site

Docking was also performed using MES as the centre point for grid generation. With crystallised water included many hydrogen bonding interactions to water were predicted, in both fragment and drug-like compound libraries (Figure 3.3).

Glide score: -7.887   Glide score: -6.973
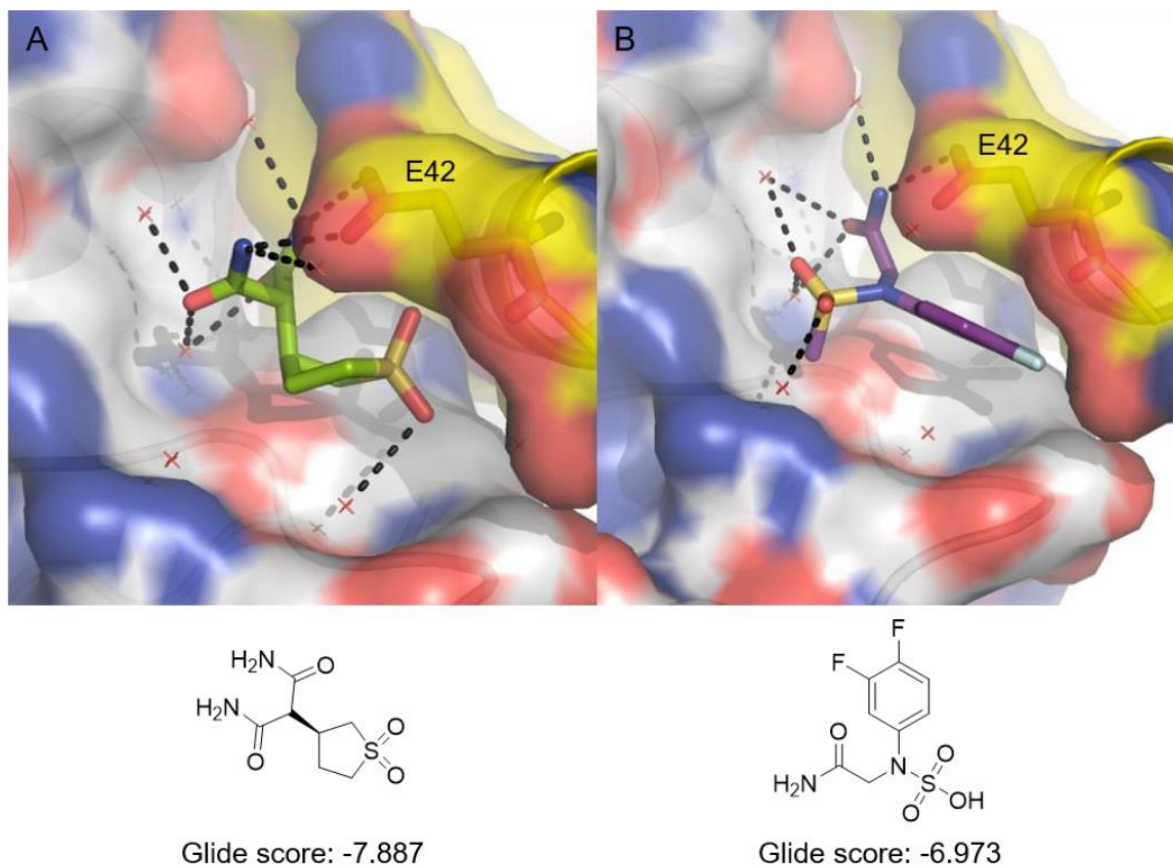
Figure 3.3 Docking at the MES binding site with water, showing predicted hydrogen bonding interactions with water molecules (red crosses) and TPX2 (yellow); (A) Fragment library result (ID B05437278, Asinex); (B) Drug-like compound library result (ID 9041792, MCCB)

With water excluded, more direct interactions with the side chains of both Aurora A and TPX2 were predicted (Figure 3.4). A set of 125 of the most interesting docking results were collected from in-house and commercial libraries and prepared for crystal soaking (See section 3.4).

Glide score: -6.359   Glide score: -6.247
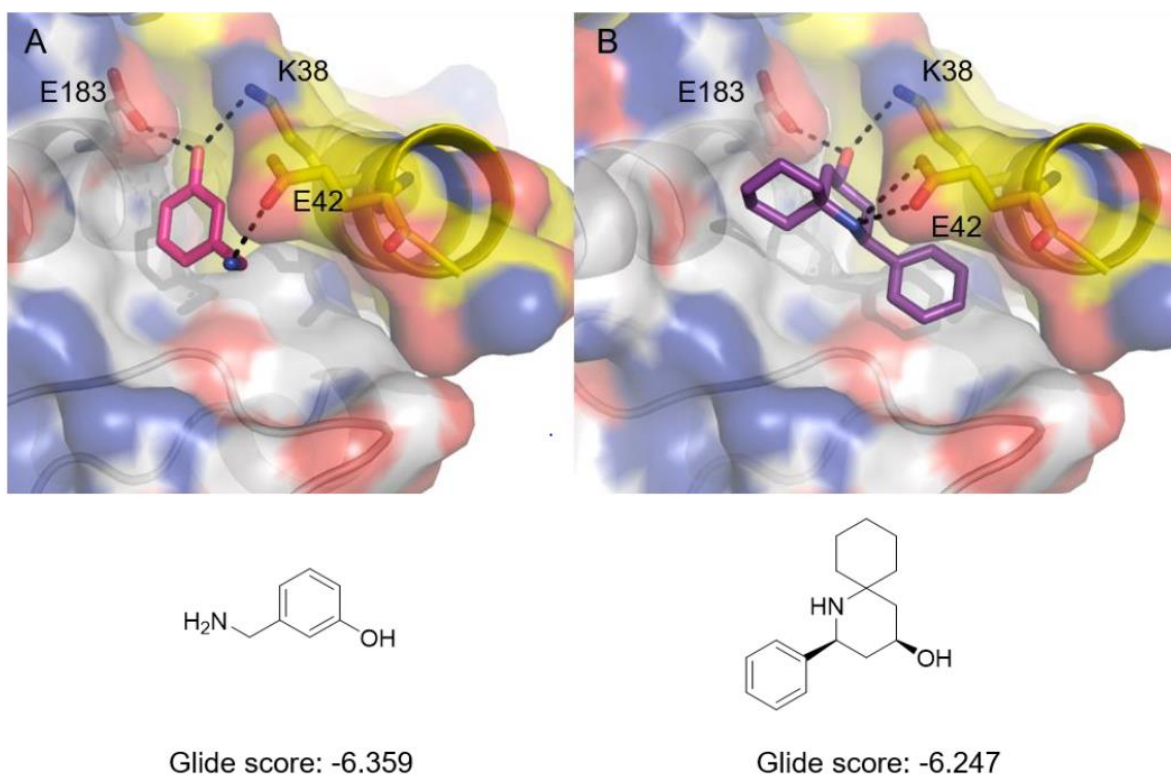
Figure 3.4 Docking at the MES binding site without water, showing predicted hydrogen bonding interactions to labelled side chains of Aurora A (grey) and TPX2 (yellow); (A) Fragment library result (ID PS-3756, Bionet); (B) Drug-like compound library result (ID 4333-1690, MCCB)

Finally, a library of compounds based on the structure of MES was created from commercial and designed compounds and screened *in silico* (Appendix 1). This library was derived from a combination of searches within commercial databases for structural similarity to MES and designed compounds with small structural changes to MES, such as changes in the ring size and changes to the sulfonic acid group. Many of these compounds generated docking poses which appeared similar to the docking pose of MES in its crystal structure and the predicted interactions of these designed compounds mirrored the kinds of interactions predicted for the commercial compound libraries (Figure 3.5).
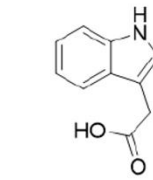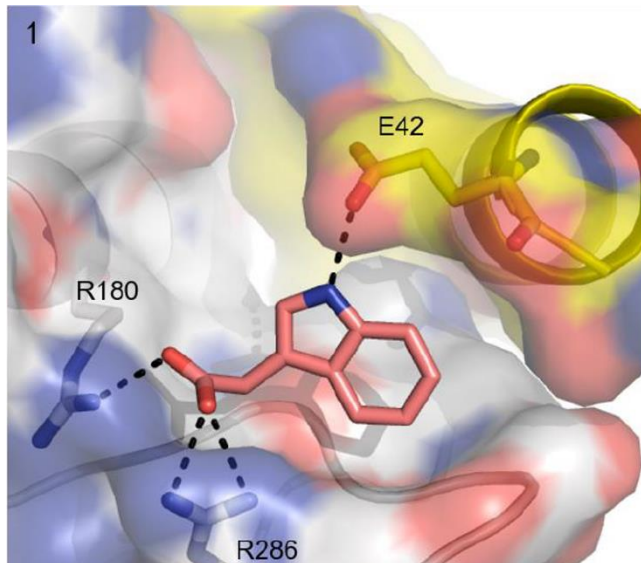
Glide score: -5.459
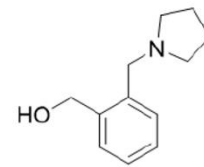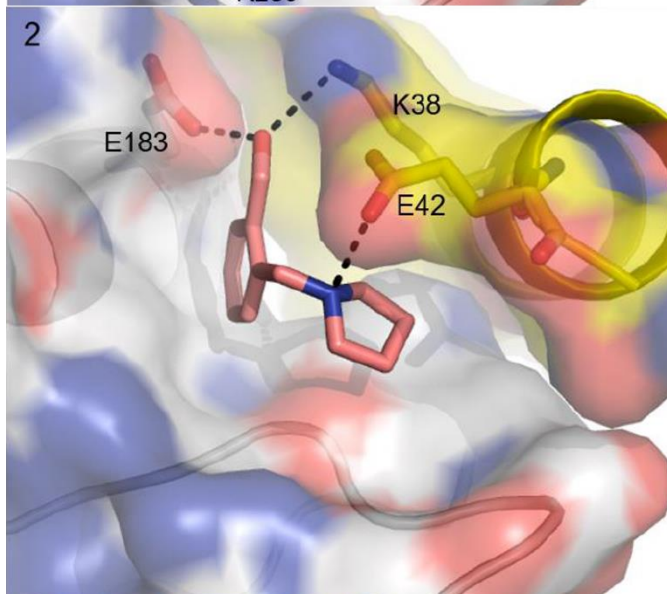
Glide score: -5.347
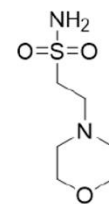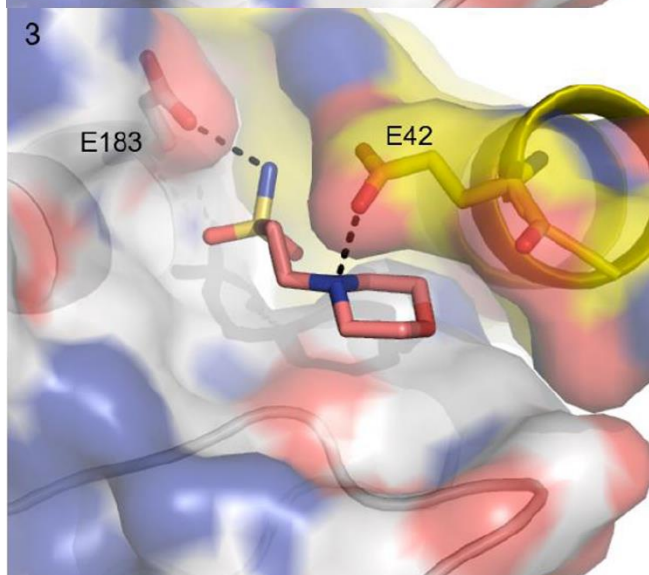
Glide score: -5.313

Figure 3.5 The top three results (by Glide score) from the designed MES-based library without water. Predicted interacting residues are labelled on Aurora A (grey) and TPX2 (yellow)

3.3 SPROUT

The fragment growth program SPROUT was used to try to improve the crystallised MES fragment. SPROUT uses a crystal structure of the target protein and identifies hydrogen bond donors, acceptors, and hydrophobic regions. Specific regions can then be chosen to extend towards by selecting potential interacting groups (e.g. hydrogen bond donors to interact with a hydrogen bond acceptor on the protein) and a selection of linking groups to connect the fragment to the new interacting group to generate a "tree" of possible extended fragments. After repeating this process multiple times, a "forest" of possibilities can be established from multiple "trees".

In the case of MES, the crystallised fragment failed to grow towards any chosen regions without rotating the predicted binding position of the MES portion. When MES was restricted to remain in its crystal pose, it failed to produce options for growing the fragment. The sulfonic acid group was also removed in an attempt to find possible alternative group in this position (Figure 3.6). Similar problems were also encountered in this setup.
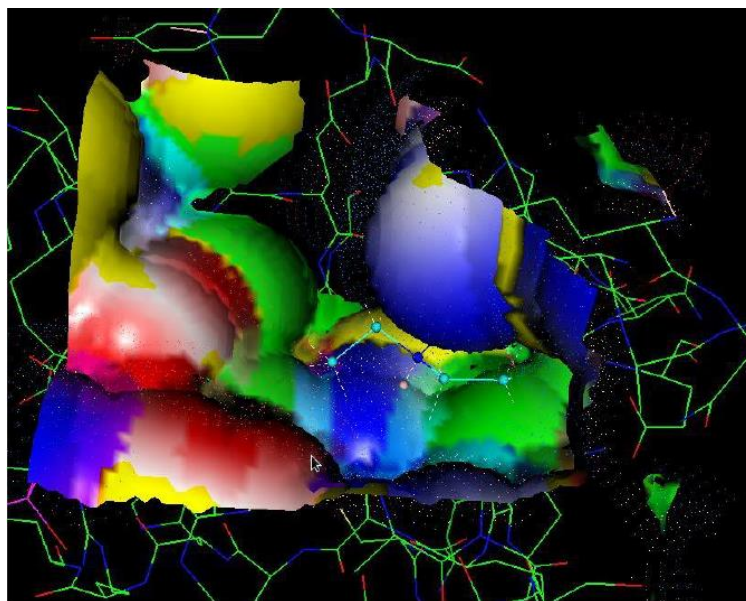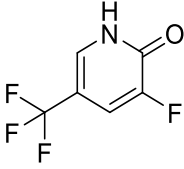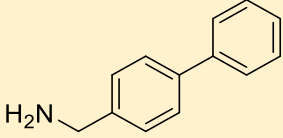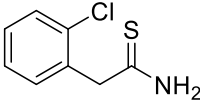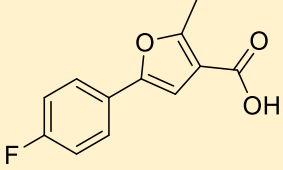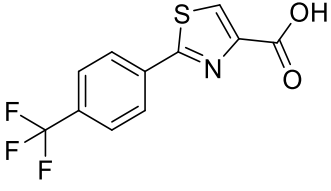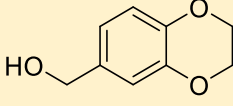
Figure 3.6 SPROUT setup, with a starting fragment shown in the binding site (blue lines)

3.4 Crystallography

A set of 125 of the most promising docking results were selected and sent to be soaked into Aurora A crystals for x-ray crystallography, performed by Patrick McIntyre of the Bayliss group. Compounds were selected based on a combination of Glide score, predicted interactions with the protein and predicted binding poses. Of these 125, only seven resulted in a crystal structure with a fragment clearly bound anywhere on Aurora A. Unfortunately, none of these compounds matched the binding prediction from Glide and did not bind at any of the sites of interest in the project.

Table 3.1 Compound details for the seven fragments present in the Aurora A crystal structure

Compound ID	Library	Structure
KM01547	Maybridge	

MAY00212	Maybridge	
CC21113	Maybridge	
DP01383	Maybridge	
MO01305	Maybridge	
MO07064	Maybridge	
HE-0207	Bionet	

3.5 Summary

Similarly to targeting the N-myc binding site directly, this type IV approach to Aurora A kinase inhibition did not provide a convincing starting point for further development. Crystallography experiments failed to show any of the chosen compounds binding in the predicted positions. Although further investigation into these results could be interesting, the focus of the project began to narrow at this time to type I and III inhibition (see Chapters 4 and 5).

Chapter 4 Type I Inhibition

A small molecule inhibitor at the ATP binding site would inhibit the kinase activity of Aurora A. As seen with CD532 **(12,** Figure 1.13), a compound binding at the ATP site can also produce a shift in the protein to disrupt the Aurora A/N-myc interaction.[31] These would be type I kinase inhibitors and could be developed from existing inhibitors of Aurora A of this type, or *de novo* using high throughput or virtual screening of fragments or small drug-like molecules.

4.1 Binding Site Analysis

Analysis by SiteMap of the ATP binding pocket from the Aurora A/N-myc crystal structure (PDB 5G1X) gave SiteScores of 1.14 and 1.08 with and without water, respectively. These were the best Sitemap results of all the target sites, indicating this is the "most druggable" of the five identified sites discussed in this report, an unsurprising result since this pocket is already known to interact well with the small molecule ADP and existing type I kinase inhibitors. The difference in the score between the inclusion and exclusion of water could be due to the potential for making additional water-mediated interactions between the protein and an inhibitor.

4.2 Development of Existing Inhibitors

VX-680 **(13)** and CD532 **(12)** are existing type I ATP-competitive inhibitors of Aurora A. Both inhibitors contain the same hinge binding region so could be combined to form a hybrid (Figure 4.1) with the hope that the designed hybrid might combine the beneficial features of each. CD532 is a potent and selective Aurora A inhibitor, but has shown toxicity and solubility issues. However, CD532 does have a disrupting effect on the Aurora A/N-myc protein-protein interaction, with the Aurora A/CD532 crystal structure suggesting the large trifluoromethyl group may be instrumental in this disruptive effect.[31] This group is not present in VX-680, which is a more soluble inhibitor selective for the Aurora family.[85]

By keeping the trifluoromethyl group of CD532, which appeared to be important for disrupting the Aurora A/N-myc interaction by SAR,[31] the hybrid compound would hopefully retain the ability to disrupt the protein-protein interaction.

13

VX-680

12

CD532

15a
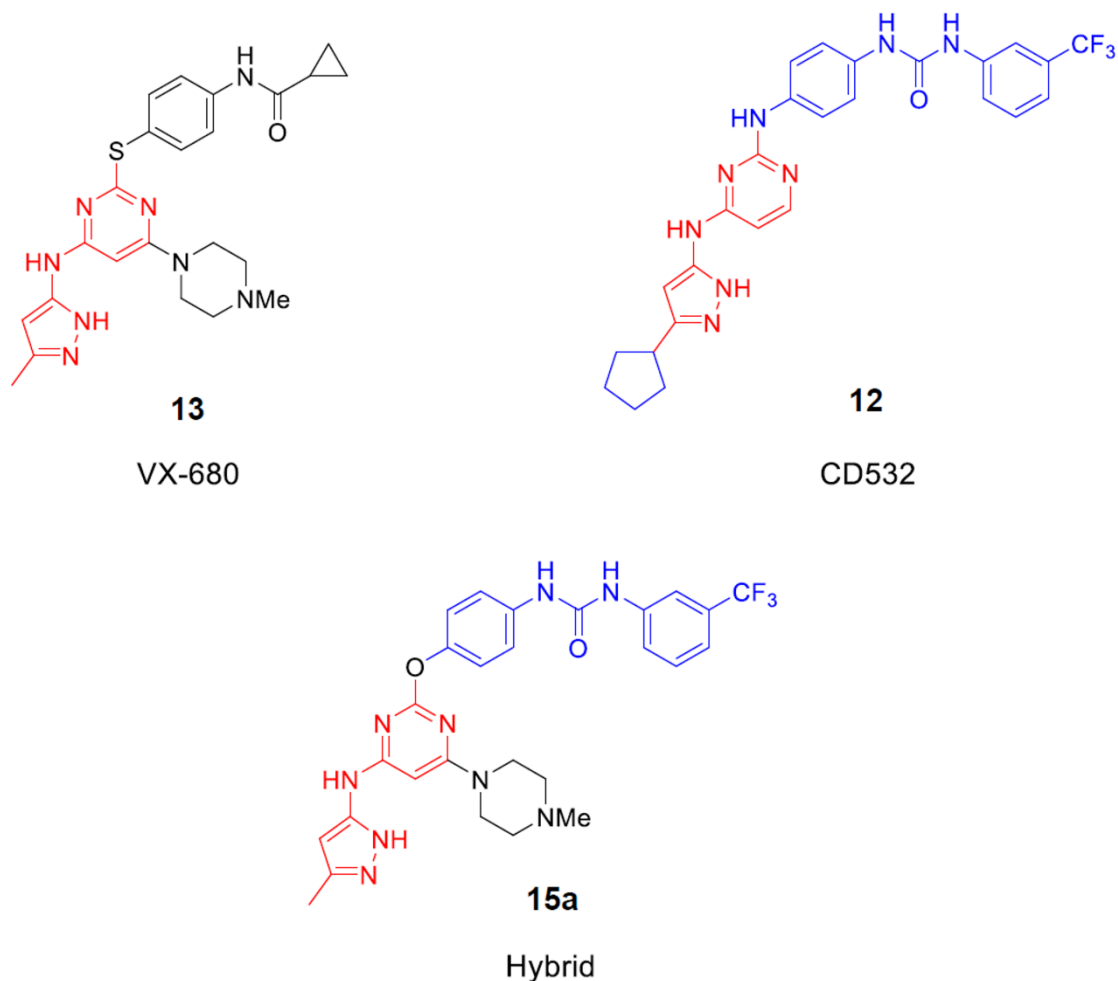
Hybrid

Figure 4.1 CD532, VX-680 and the designed hybrid: the hinge binding region (red) is common to both CD532 and VX-680; urea moiety is from CD532 (blue) and methylpiperazine group is from VX-680 (black)

Incorporating the protonatable amine from the methylpiperazine group of VX-680, hopefully the solubility of the new hybrid would improve without compromising potency. The designed hybrid would offer a structurally novel class of compounds, based on the scaffold shown (Figure 4.2).
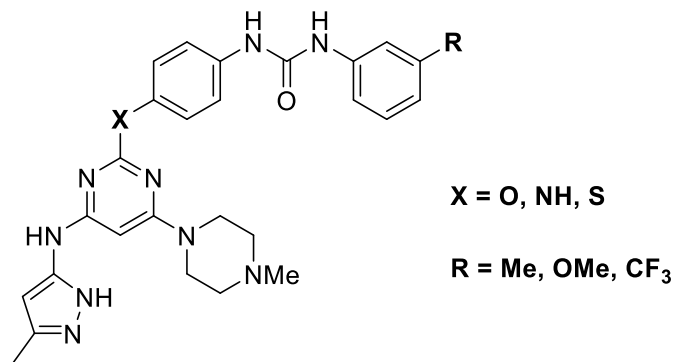
Figure 4.2 Planned CD532/VX-680 hybrid compound template

Synthesis of the hybrid compound was adapted from a literature procedure[109] (Scheme 1). The thiopyrimidine **(16)** was oxidised with mCPBA to give the sulfone **(17)** in good yield which was then able to undergo S_NAr substitution at the 2-position to generate the substituted pyrimidine **(18a-c)**. Further nucleophilic aromatic substitution with 3-amino-5-methylpyrazole and N-methylpiperazine at the 4- and 6- positions produced the substituted pyrimidine **(19a-c)** in good yield. The pyrazole nitrogen was then *Boc*-protected **(20a-c)**, followed by reduction of the nitro group to the amine **(21a-c)**. The amine was then reacted with the corresponding isocyanate to form the urea moiety and deprotection of the pyrazole nitrogen yielded the desired compounds **(15a-c)**, which were tested for kinase inhibition in the assay described below (Section 4.4).
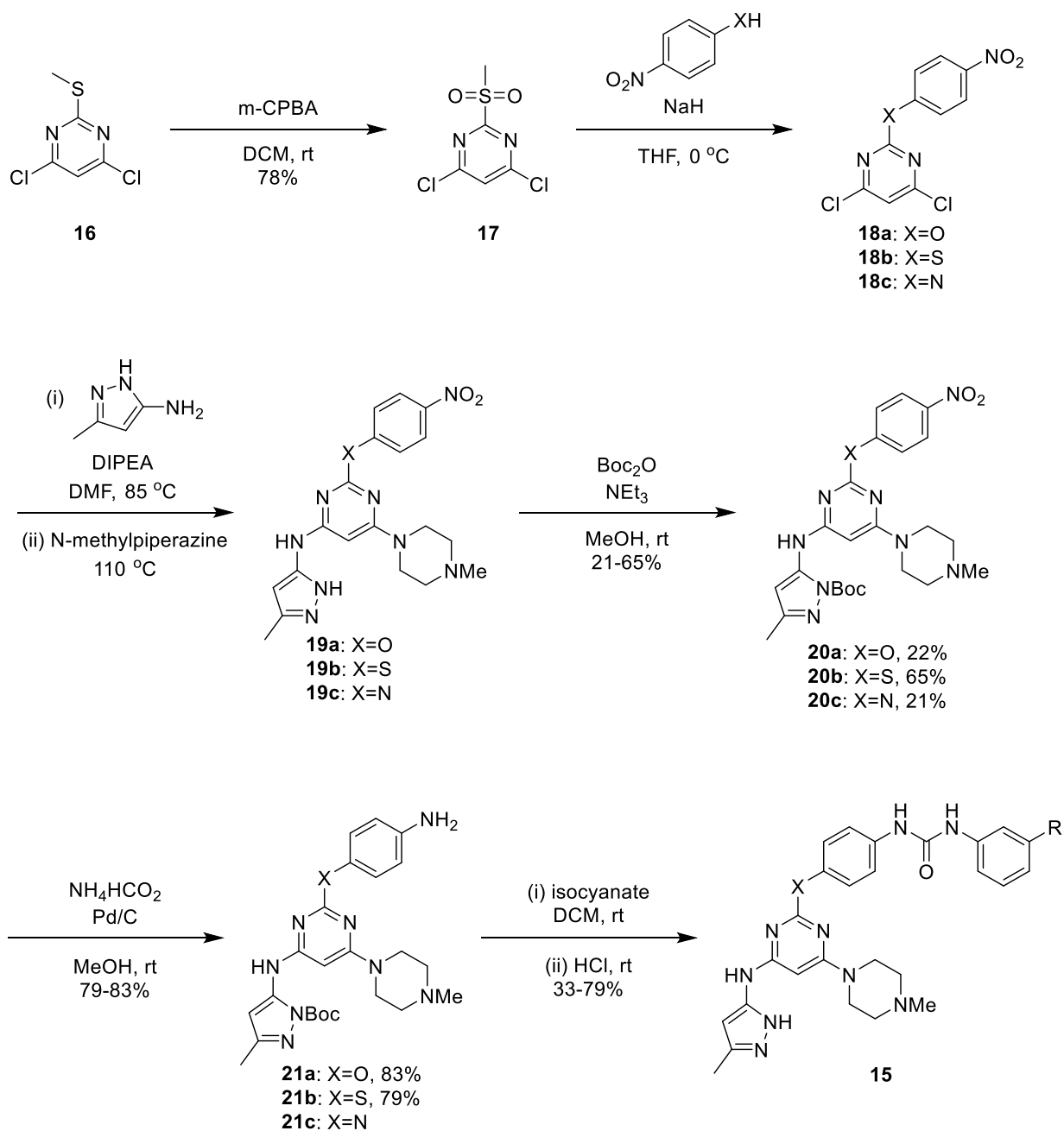
Scheme 1 Synthesis of CD532/VX-680 hybrids **(15)**

Table 4.1 Summary of final compounds and individual yields for final step

Compound	X	R	Yield for final step
15a	O	CF_3	79%
15b	S	Me	33%
15c	S	OMe	36%

4.3 Docking

4.3.1 Screening Libraries

As a complementary approach to the development of ATP-competitive compounds based on existing inhibitors, both fragment and larger compound libraries were docked at the ATP binding site using the co-ordinates of crystallised ADP as the docking grid centre point. Although this was not the main focus for work at the ATP site, any promising results from this docking could generate new scaffolds for the development of ATP-competitive inhibitors which disrupt the Aurora A/N-myc interaction, similar to CD532 **(12)**. Docking results may also help to identify interactions predicted to be important at this site.

ADP was redocked in both the Aurora A/N-myc (PDB 5G1X) and Aurora A/TPX2 (PDB 5LXM) crystal structures. The Aurora A/N-myc structure gave a much better RMSD value (0.41 Å versus 2.40 Å) for redocked ADP under SP conditions and was, therefore, used for screening libraries. Although XP (extra precision) docking did give a better RMSD for both crystal structures, XP docking did take significantly more time. The SP RMSD value of 0.41 Å was considered sufficiently low to be able to give useful predictions from screening libraries. The 152 000 compounds from fragment and drug-like compounds were docked at this site, both with and without crystallised water.

Docking with the crystallised water molecules in place did result in a number of predicted water-mediated interactions, providing support for the higher druggability score from SiteMap with water included. However, docking scores for both the fragment and drug-like compound libraries were similar regardless of whether water was included or not, suggesting these additional water-mediated interactions were not predicted to produce significantly more favourable binding.

The fragment libraries tended to produce poses which overlapped with where the adenine portion of ADP would sit. Lys162 and Lys143 were popular Aurora A side chains with predicted interactions to fragments, often a carboxylic acid. Without the presence of water molecules, the highest scored fragments showed only these lysine-carboxylic acid interactions and some charge-charge interactions between the same carboxylic acid and a magnesium ion (Figure 4.3). Lower ranked fragments did show greater variability in the interactions predicted without water, with hydrogen bonding predicted to Ala213, Arg220, Glu260, Asn261 or the key enzymatic residue Asp274, but none approached the number of interactions shown for ADP (Figure 4.4).

Glide score: -12.375   Glide score: -11.255
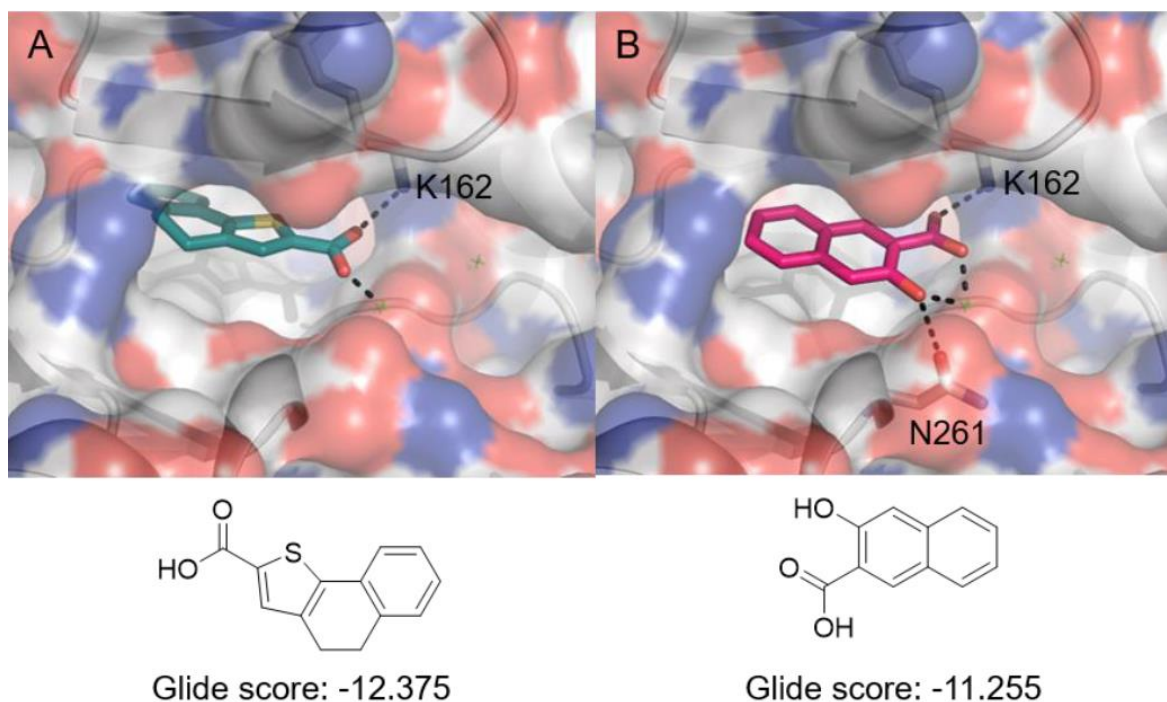
Figure 4.3 Fragment docking results at the ATP binding site without water; (A) One of the highest ranked results (by Glide score), showing the fragment positioned in the ATP binding site and interactions to a magnesium ion and a lysine side chain (compound ID 12N-508S, Bionet), as was seen for many of the top results; (B) One of the slightly lower ranked results predicting interactions to other side chains, including Asn261 in this case (compound ID DK-0222, Bionet)
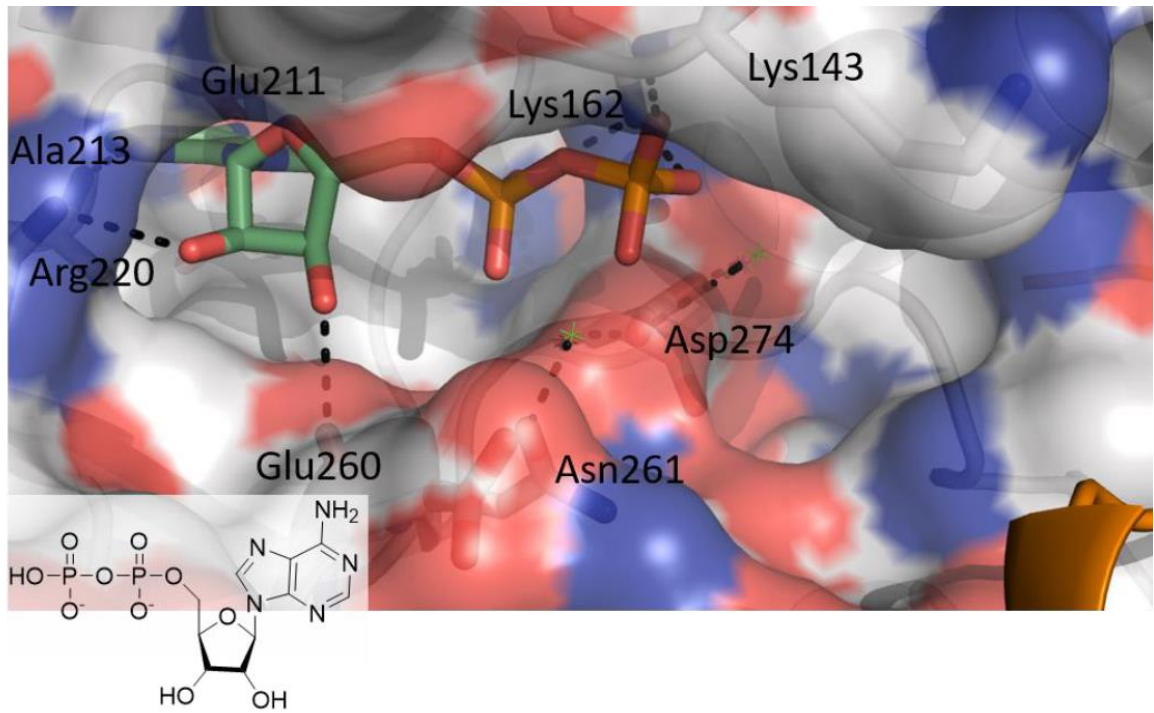
Figure 4.4 Crystallised ADP in its binding site with N-myc (orange) shown (PDB 5G1X) with water removed

With the crystallised water included, direct hydrogen bonds were predicted almost exclusively to Lys143 and Lys162 among the highest scoring fragments (Figure 4.5). Again, almost all fragments were predicted to bind close to where ADP would be found.

Glide score: -11.411   Glide score: -11.411
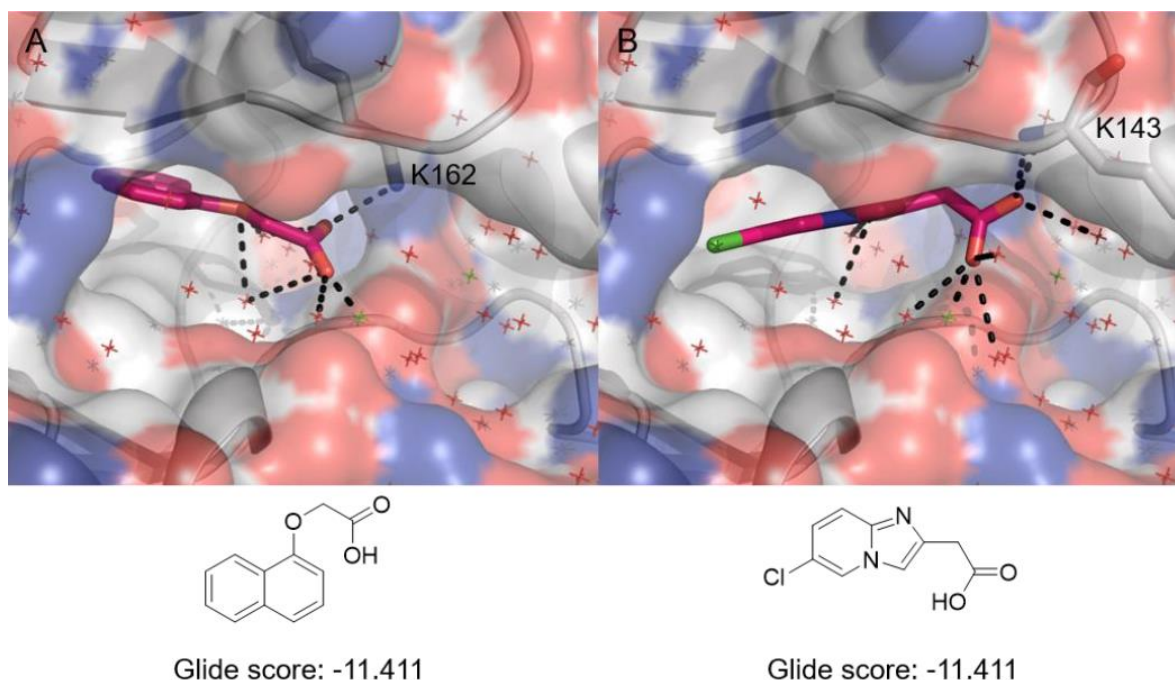
Figure 4.5 Docking results at the ATP binding site with water included, showing predicted interactions with (A) Lys162 (compound ID 2M-936, Bionet) and (B) Lys143 (compound ID 7X-0811, Bionet), along with predicted hydrogen bonding to water (red cross) and interactions with a magnesium ion (green cross)

For the libraries made up of larger, more drug-like compounds, again the vast majority of the highest ranked compounds were predicted to overlap with where ADP would be found. There were a few similarities between these libraries and fragment libraries, notably the recurrence of interactions with Lys143 and Lys162 and a carboxylic acid regardless of whether crystallised water was included or not (**Error! Reference source not found.**). With the inclusion of water, the highest scoring results all showed some hydrogen bonding interactions with the surrounding water molecules, although these additional interactions did not affect the scores produced when compared to the exclusion of water.

Glide score: -13.170   Glide score: -12.747
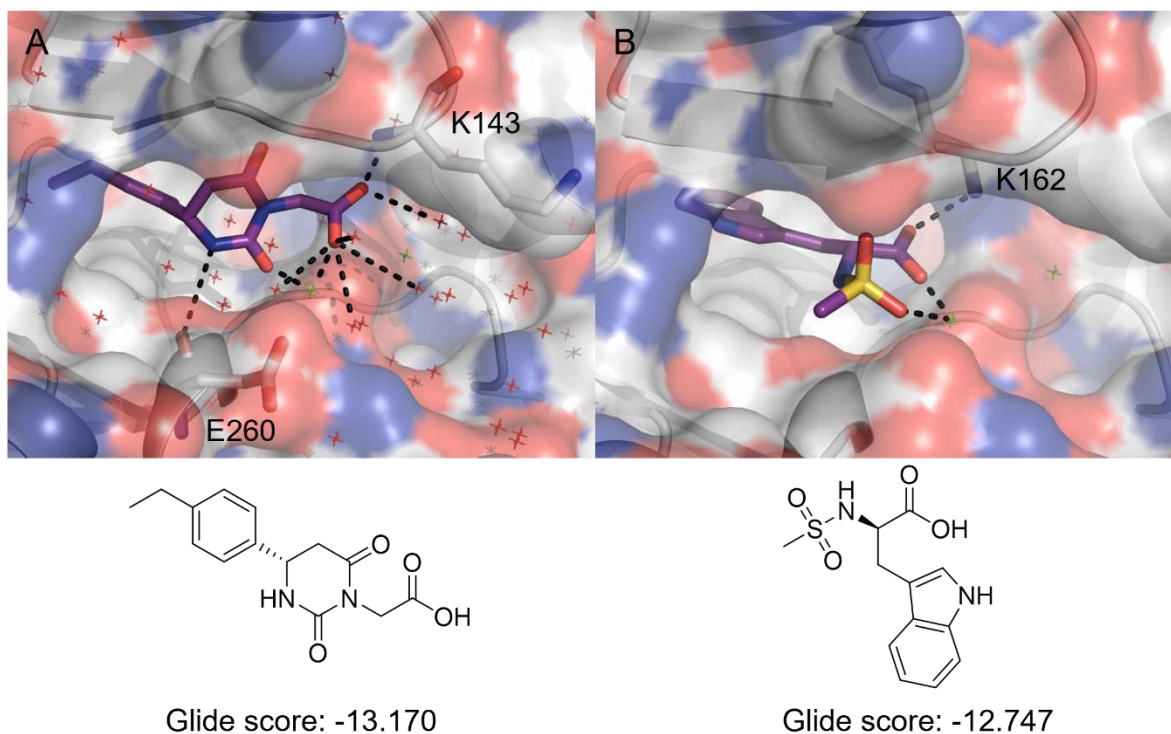
Figure 4.6 Docking results from drug-like libraries screened at the ATP binding
site; (A) with water (compound ID ALB-H04810563, MCCB); (B) without water
(compound ID 9199753, MCCB)

One difference between fragment libraries and these drug-like libraries was the
predicted pose of many results. Whereas for the fragment libraries a number of
highly ranked compounds were predicted to interact with the common lysine side
chains and moved away from the hinge, the larger compounds in the MCCB and
Chembridge libraries allowed compounds to both form predicted interactions with
the favoured lysine residues whilst also extending towards the hinge. A few of the
highest ranking compounds showed predicted hydrogen bonding to key hinge
residues including Ala213 (Figure 4.7), as opposed to fragment libraries where only
lower ranked results showed these interactions.

Glide score: -12.958
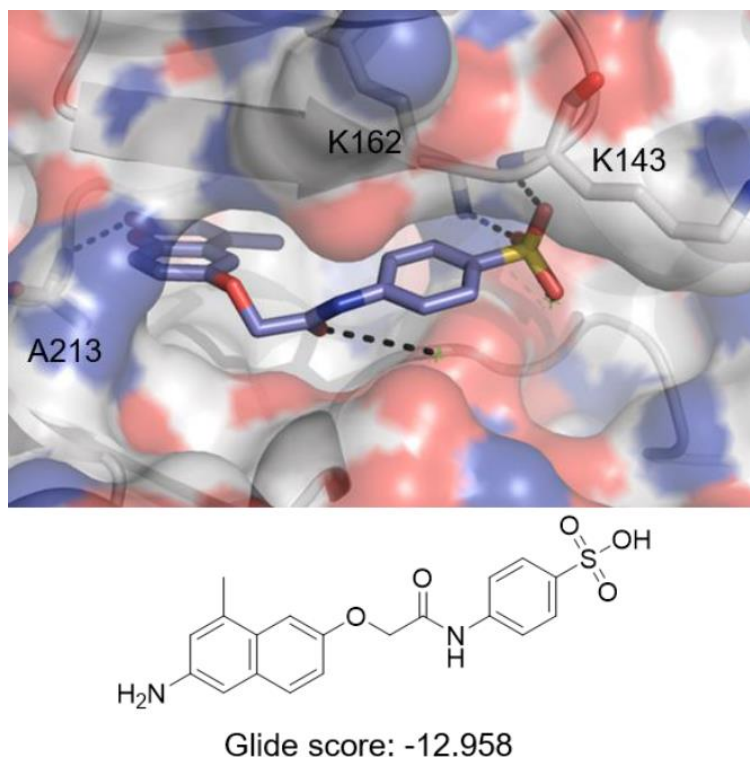
Figure 4.7 One of the top results from docking the drug-like libraries at the ATP binding site, showing predicted hydrogen bonding interactions to Ala213 (compound ID omega_86847, Chembridge)

These virtual screening results have given some potential ideas for investigating new type I kinase inhibitors of Aurora A which could be further explored using enzyme activity assays or protein crystallography.

4.3.2 Designed CD532/VX680 Hybrid Compounds

The existence of a good quality crystal structure of Aurora A with CD532 bound[31] presented the opportunity to explore the docking poses of CD532/VX680 hybrid compounds. CD532 was first re-docked into its binding site until conditions were found which produced a re-docked pose similar to that observed in the crystal (RMSD 0.53 Å). The optimum conditions for re-docking CD532 were found using crystallised CD532 as the grid centre point under SP (standard precision) with a Glide grid consisting of a 10 Å inner box and 24 Å outer box. A selection of hybrid compounds was then docked under these conditions in the same site. The hybrid

compounds were docked with and without the cyclopentyl group of CD532 with the hope of exploring the role of this group in hybrid compounds.

Some compounds produced poses which overlaid well with CD532 (Figure 4.8) with the additional piperazine ring from VX-680 predicted to be directed towards solvent. Unfortunately, this was not the case for the majority of compounds docked. In most cases, the docked compounds were "flipped," with the cyclopentyl region directed towards solvent and the piperazine group towards the back of the pocket, overlapping with the position of the CD532 cyclopentyl group.
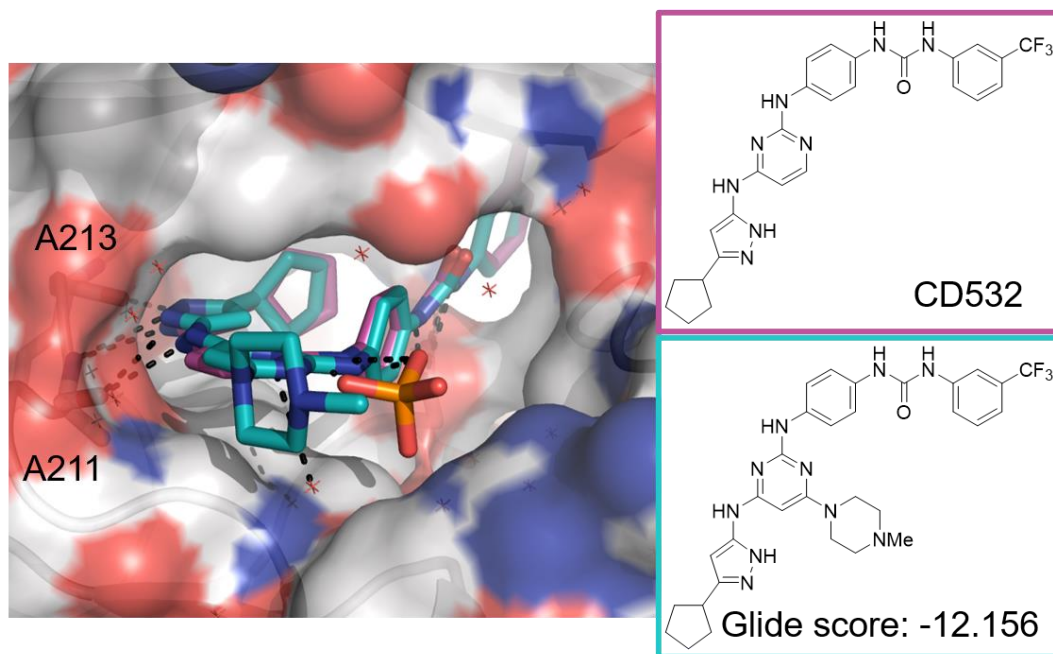
Figure 4.8 The docking pose produced for one of the designed compounds (blue) with crystallised CD532 overlaid (pink) (PDB 4J8M). Aurora A side chains predicted to form hydrogen bonding interactions are labelled and highlighted (grey sticks).

The predicted poses produced by this docking did not show a clear prediction of whether the designed compounds would be likely to bind to Aurora A in a similar manner to CD532. The reported docking poses showed big differences even between similar compounds. Without a clear prediction from docking studies, the

designed compounds were still synthesised and tested using the methods described (Sections 4.2 and 4.4).

4.4 Kinase Inhibition Assay

The ability of the designed compounds based on CD532 and VX-680 to inhibit the kinase activity of Aurora A was assessed using the Caliper Mobility-Shift Assay. This assay used the mutant Aurora A C290A:C293A, which was prepared in *E. coli* using a construct available in the Bayliss lab. The expressed His-tagged protein was purified by IMAC (Immobilised metal ion affinity chromatography) and size exclusion chromatography. Successful purification was confirmed by SDS-PAGE (Figure 4.9).
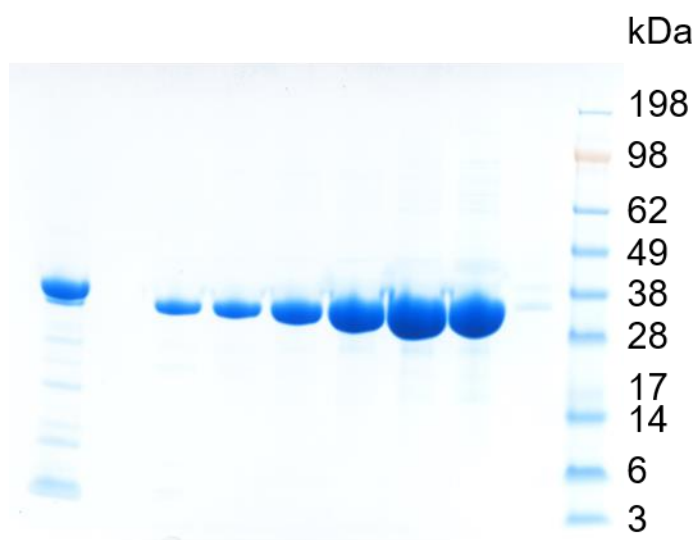
Figure 4.9 SDS-PAGE results following final purification of Aurora A C290A:C393A mutant. The left lane shows a sample before gel filtration and subsequent lanes show eluted fractions.

The Caliper Mobility-Shift Assay uses a fluorescent peptide substrate of the enzyme, with the fluorescence intensity eventually used to determine enzyme activity. At set time intervals, samples of the enzyme/substrate mixture are taken, which pass into a capillary where the fluorescent substrate and any phosphorylated

product are separated by charge using the positive and negative electrodes. The detector measures fluorescence over time, with the negatively charged phosphorylated product detected first, followed by any unreacted substrate. The ratio of the area under the curve (AUC) for these two observed peaks allows the determination of the extent of substrate conversion (Figure 4.10). By taking these same measurements over time, enzyme kinetics can be studied. In this case, synthesised analogues of CD532 and VX-680 were added to the enzyme/substrate mixture at a range of concentrations to analyse their ability to inhibit the activity of Aurora A.
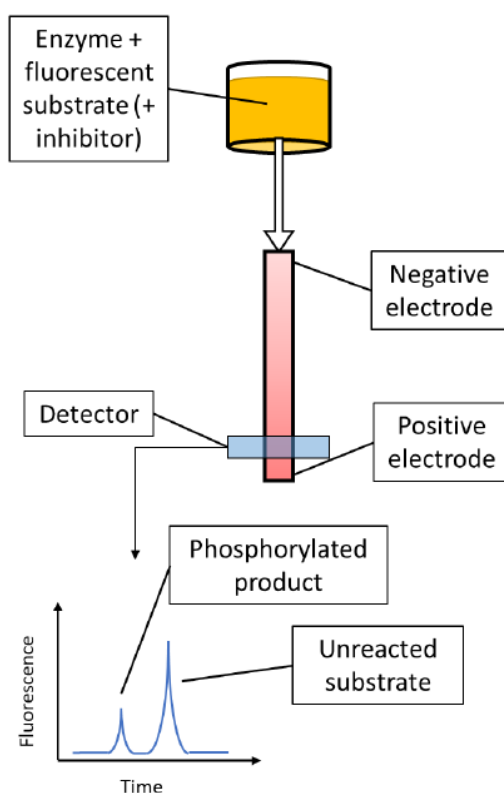
Figure 4.10 Schematic of the Caliper Mobility-Shift Assay

An enzyme titration (Figure 4.11) and ATP K_m determination (Figure 4.12) were first performed to identify the optimum concentrations of Aurora A and ATP to use to achieve the greatest sensitivity. Concentration used for IC_{50} experiments were 50 nM Aurora A and 80 μM ATP. Substrate conversion was measured at various

concentrations of each of the designed compounds, with the hope that higher concentrations of inhibitor would show a decreased substrate turnover. Assays were performed in duplicate, and each fluorescence measurement was taken twice. Unfortunately, for each of the designed compounds (**15a-d**), no effect was seen on the activity of Aurora A, regardless of the inhibitor concentration (Figure 4.13).
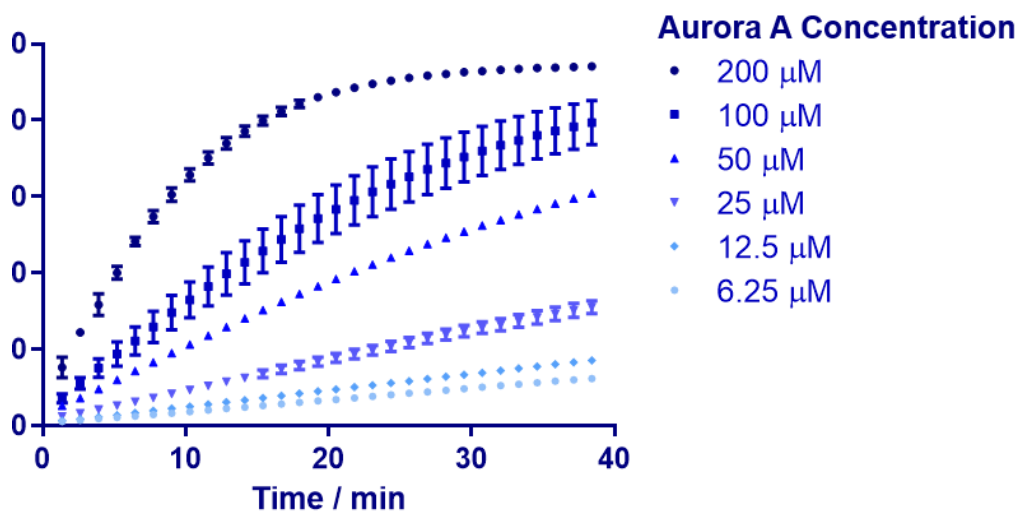
Figure 4.11 Aurora A kinase titration; a graph of substrate conversion over time for various concentrations of Aurora A. Error bars indicate two standard deviations from the mean.

V_{max}: 1.02 ± 0.03 µmol min^{-1}
K_m: 0.07 ± 0.01 mM
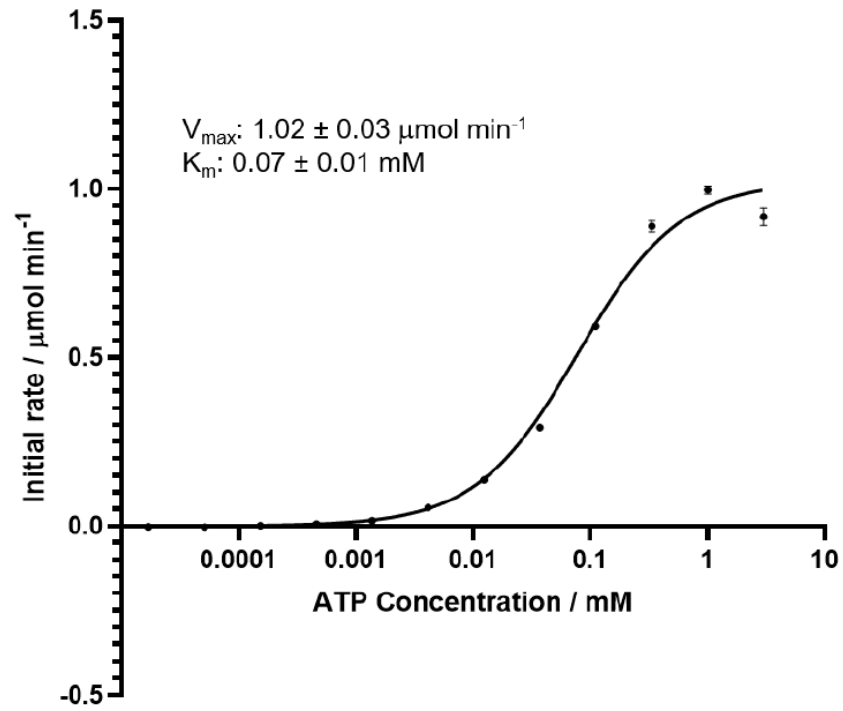
Figure 4.12 A graph showing the initial rate of reaction with increasing ATP concentration, used to determine V_{max} and ATP K_m of Aurora A protein used
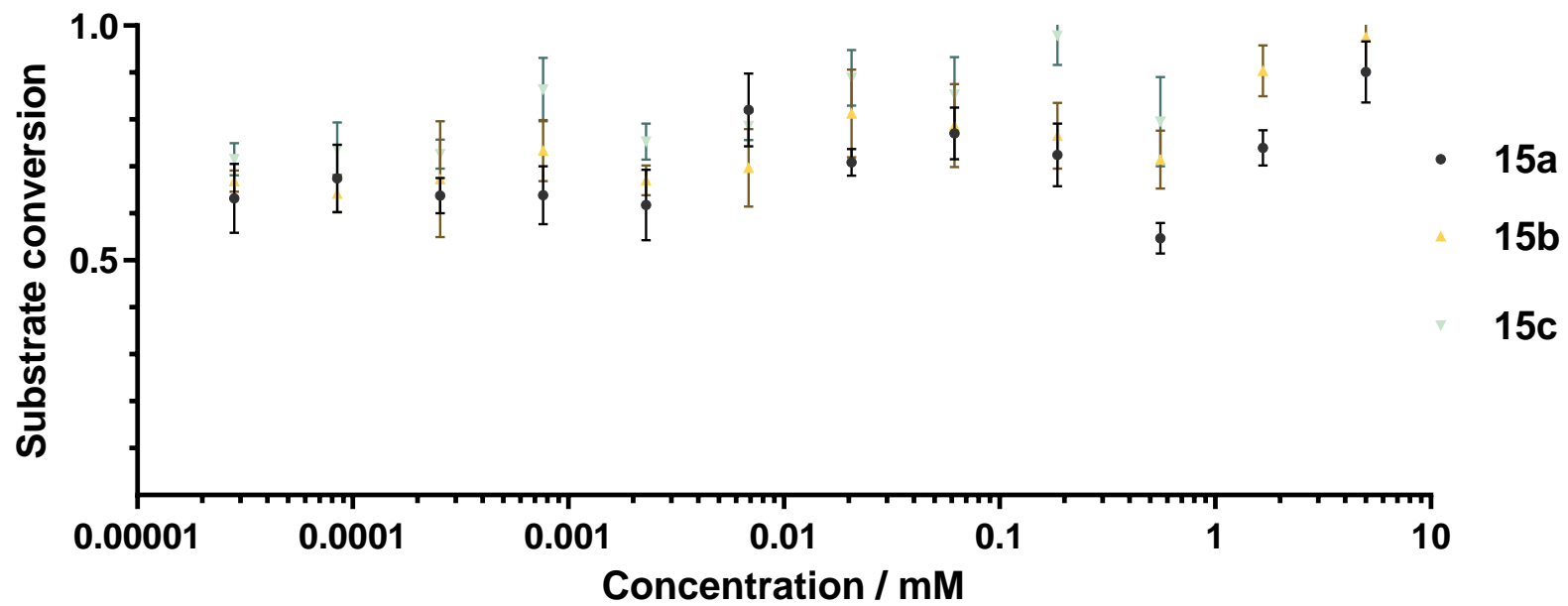
Figure 4.13 Results of proportion of substrate conversion over increasing concentrations of **15a-c**. Error bars indicate two standard deviations from the mean.

Although these results were disappointing, type I inhibition of Aurora, even if successful, may not have been the ideal mode of Aurora A inhibition. Existing type I inhibitors of Aurora A often suffer from toxicity,[81,110] so a different mode of Aurora A inhibition which is not competitive with ATP may be useful in overcoming selectivity challenges of type I inhibition and toxicity issues arising from the inhibition of Aurora A activity in healthy cells.

4.5 Summary

The type I mode of inhibition was the branch of the project with the largest amount of previous research. Existing compounds in the literature, some of which showed a disruptive effect on the Aurora A/N-myc protein-protein interaction, provided a starting point for the development of novel inhibitors with improved properties. Unfortunately, results from the Caliper mobility shift assay showed the synthesised hybrid compounds lost the inhibitory activity on Aurora kinase A. Compounds library docking generated a number of small molecules with good Glide scores and docking poses which predicted binding at the ATP binding site, although selectivity over other kinases and the potential for toxicity would likely be challenges in the development of any type I inhibitor.

Chapter 5 Type III Inhibition

The active conformation of Aurora A contains a salt bridge between Lys162 and Glu181 (Figure 5.1). A compound which disrupts this salt bridge interaction could lead to development of inhibitors of activation of Aurora A kinase and may affect N-myc binding, especially if this distortion induces a wider conformational change incompatible with N-myc interaction. A compound which disrupts the salt bridge interaction while ADP is still able to interact with its binding site would be classified as a type III kinase inhibitor. Type III kinase inhibitors may be able to offer greater selectivity over other, similar, kinases than existing type I, ATP-competitive inhibitors which rely on binding to the highly conserved ATP binding site. To date, no type III kinase inhibitors of Aurora A have been described in the literature.
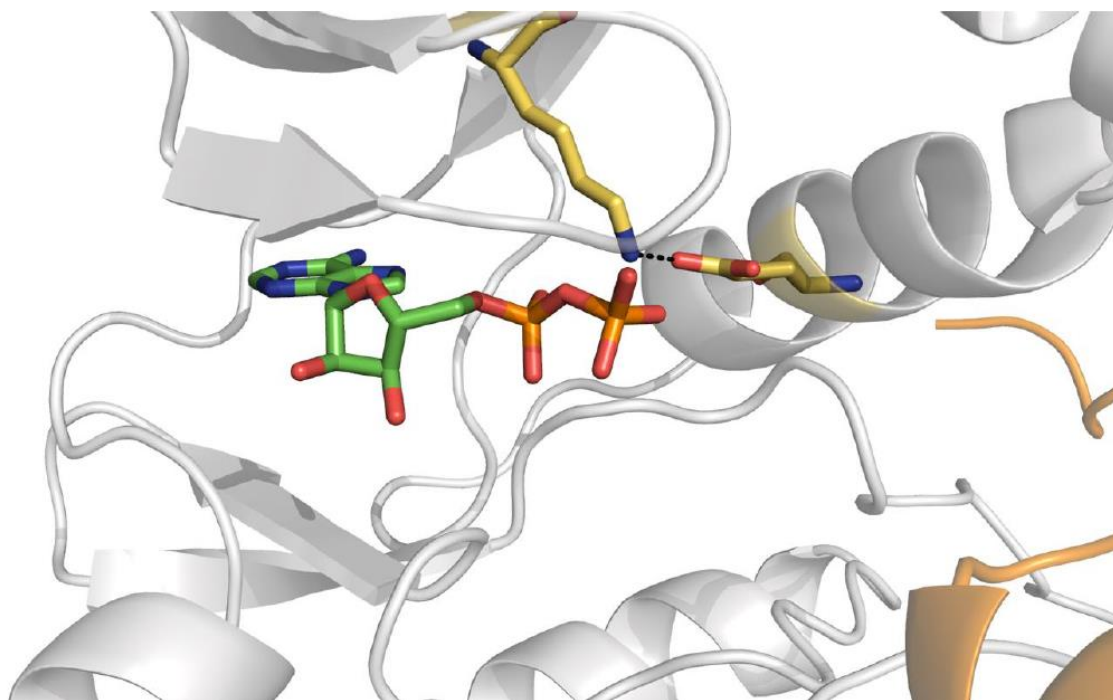
Figure 5.1 The salt bridge formed in the active conformation of Aurora A (yellow sticks) from the crystal structure of Aurora A (grey) with N-myc (orange) (PDB 5G1X)

The inactive conformation of Aurora A, with the salt bridge between Lys162 and Glu181 broken, has been crystallised, stabilised with a nanobody, vNAR-D01[48]

(Figure 5.2). In order to identify starting points for inhibitor design, this crystal structure could be used for virtual screening of compound libraries to identify compounds binding in the region between the salt bridge residues. Compounds could also be soaked into crystals of this stabilised inactive structure.
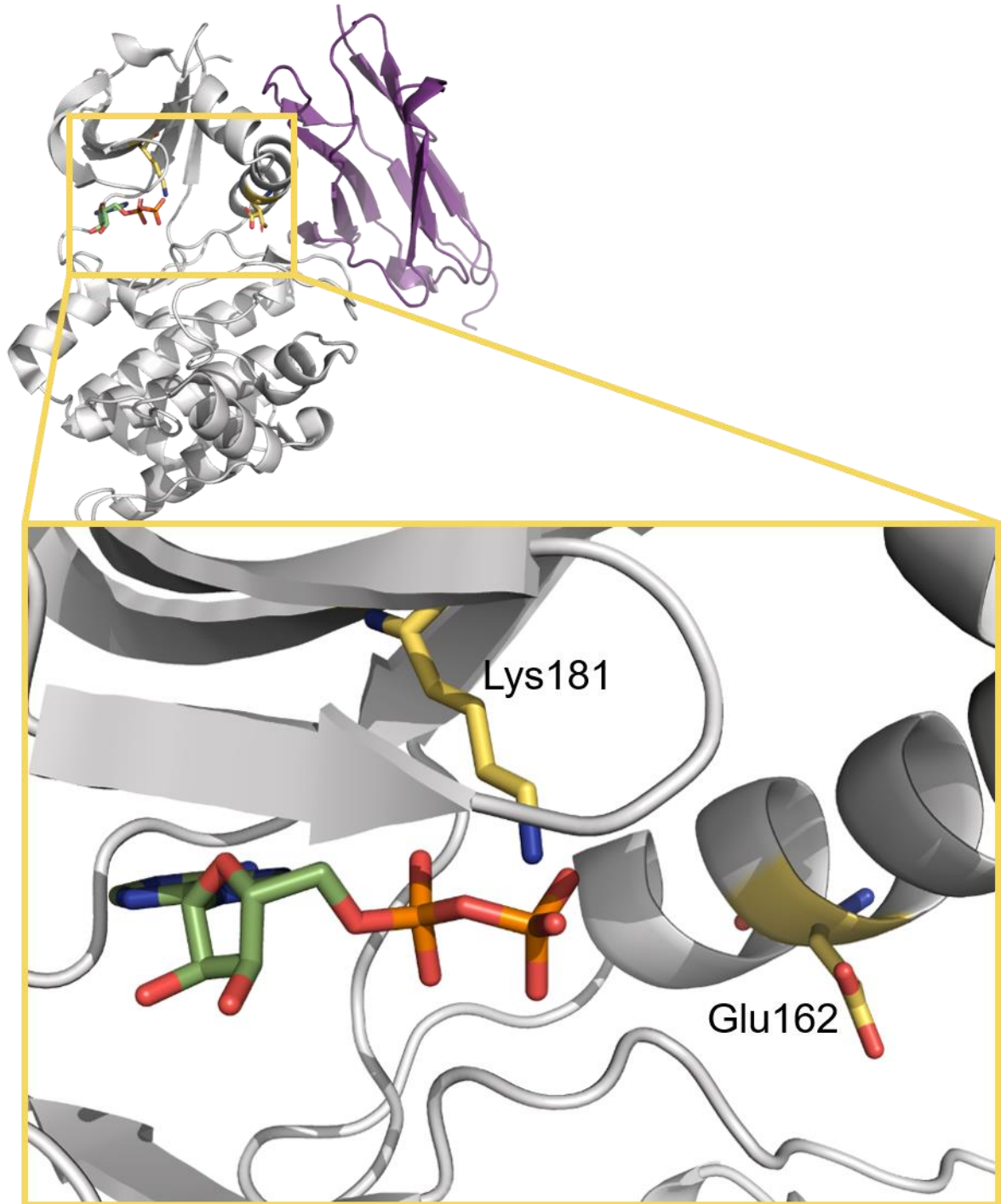
Figure 5.2 Aurora A crystal structure (grey) with the salt bridge residues (yellow sticks) and ADP (green sticks) shown. The broken salt bridge is stabilised by the nanobody vNAR-D01 (purple, PDB 5L8L).

Analysis of the binding site by SiteMap from the Aurora A/vNAR-D01 crystal structure (PDB 5L8L) (retaining the crystallised water molecules) produced SiteScores of 1.10 and 1.01 with and without ADP, respectively. Without the water molecules, SiteScores were 1.06 and 1.05 with and without ADP. In general, scores above 0.8 indicate a "druggable" site.[91] Although the score is only a prediction of the "druggability" of a target site, the high scores returned is reassuring for the potential development of small molecule type III kinase inhibitors at this site.

5.1 Kinase Alignment

To investigate the question of whether type III kinase inhibitors could offer greater selectivity over similar kinases than type I inhibitors, the two proposed binding areas were analysed for conservation of amino acids found within each area in the most similar kinases to Aurora A. A protein BLAST search was performed using the full-length Aurora A sequence to identify the most similar human protein sequences for subsequent alignment studies (Table 5.1). The similarity of proteins to Aurora A considered the percentage identity, or the proportion of amino acids of the same type in the same position as the query protein, as well as the Expectation (E) value, which assesses the probability of a similarly scoring sequence being found by chance in a database. The total score is calculated using the number of matches between sequences and deducting any penalties for mismatches and the introduction of gaps. A well-matched protein would show a high total score and percent identity, and a low E value.

Table 5.1 Protein BLAST results with highest similarity to Aurora A and a good
quality crystal structure

Protein	Accession Number	% Identity	E value	Total score
Aurora A	NP_001310234.1	Query (100%)	Query (0)	-
Aurora B	NP_001300881.1	71	7e-139	401
Aurora C	NP_001015878.1	68	1e-141	408
PLK4	BAB69958.1	38	2e-58	207
PDK1	pdb|3PWY|A	37	9e-49	170
ULK3	pdb|6FDY|U	35	1e-48	169
Ribosomal S6	AAH06106.3	39	1e-48	174
PLK2	pdb|4I5M|A	31	9e-47	165

Before embarking on a more comprehensive analysis, initially the eight kinases
with highest sequence identity with Aurora A and available good quality crystal
structures were aligned based on their ATP binding pockets using Maestro
(Schrödinger) for a visual inspection of binding site similarity (Figure 5.3).
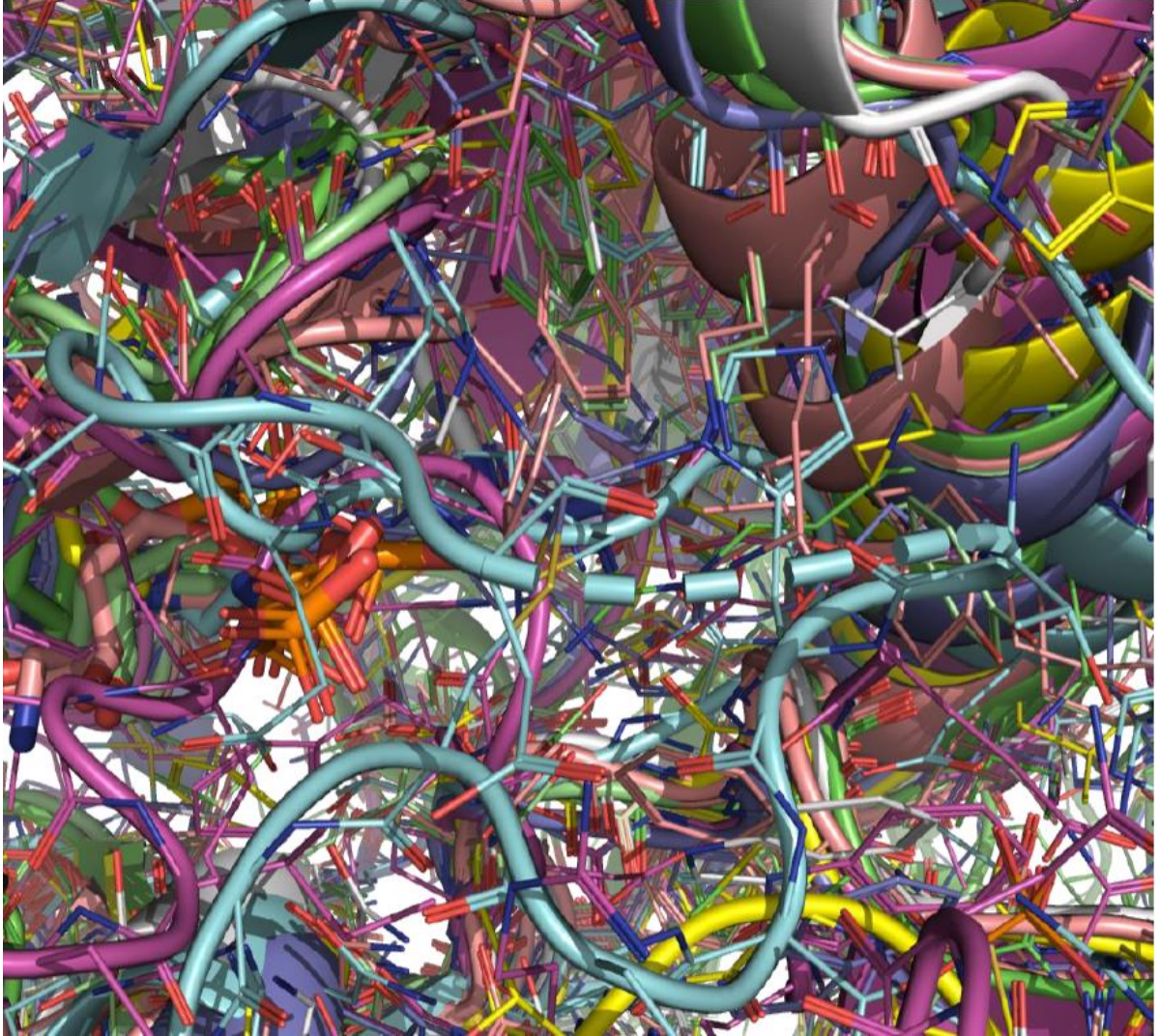
Figure 5.3 Crystal structure alignment of kinases similar to Aurora A, with Aurora A
(grey), Aurora B (blue/purple), PLK4 (green), PDK1 (dark pink), Ribosomal S6
protease (blue), ULK (light pink), PLK2 (yellow) and FLT3 (teal).

Aligning kinases in this way has limited use when working with large numbers of
kinases. Although it is possible to identify the phosphate end of ADP, little useful
information about the conservation of different areas of the kinases from this visual
inspection, even with a limited number of crystal structures included. Furthermore,
this relied on the availability of good quality crystal structures, which were not
available for all kinases. However, there is some use to this method, for example
when comparing very specific sections of a very limited number of structures.

Comparing only Aurora A and FLT3 kinases, it is possible to see the presence of Phe691 of FLT3 in place of the Aurora A residue Leu210 (Figure 5.4). Differences like these could present an opportunity to introduce selectivity over related kinases by exploiting key differences in structure. Although this direct comparison of two structures could be applied multiple times to study all of the related kinases with crystal structures, a more "high-throughput" approach, which didn't rely on the availability of good quality crystal structures, was desired.
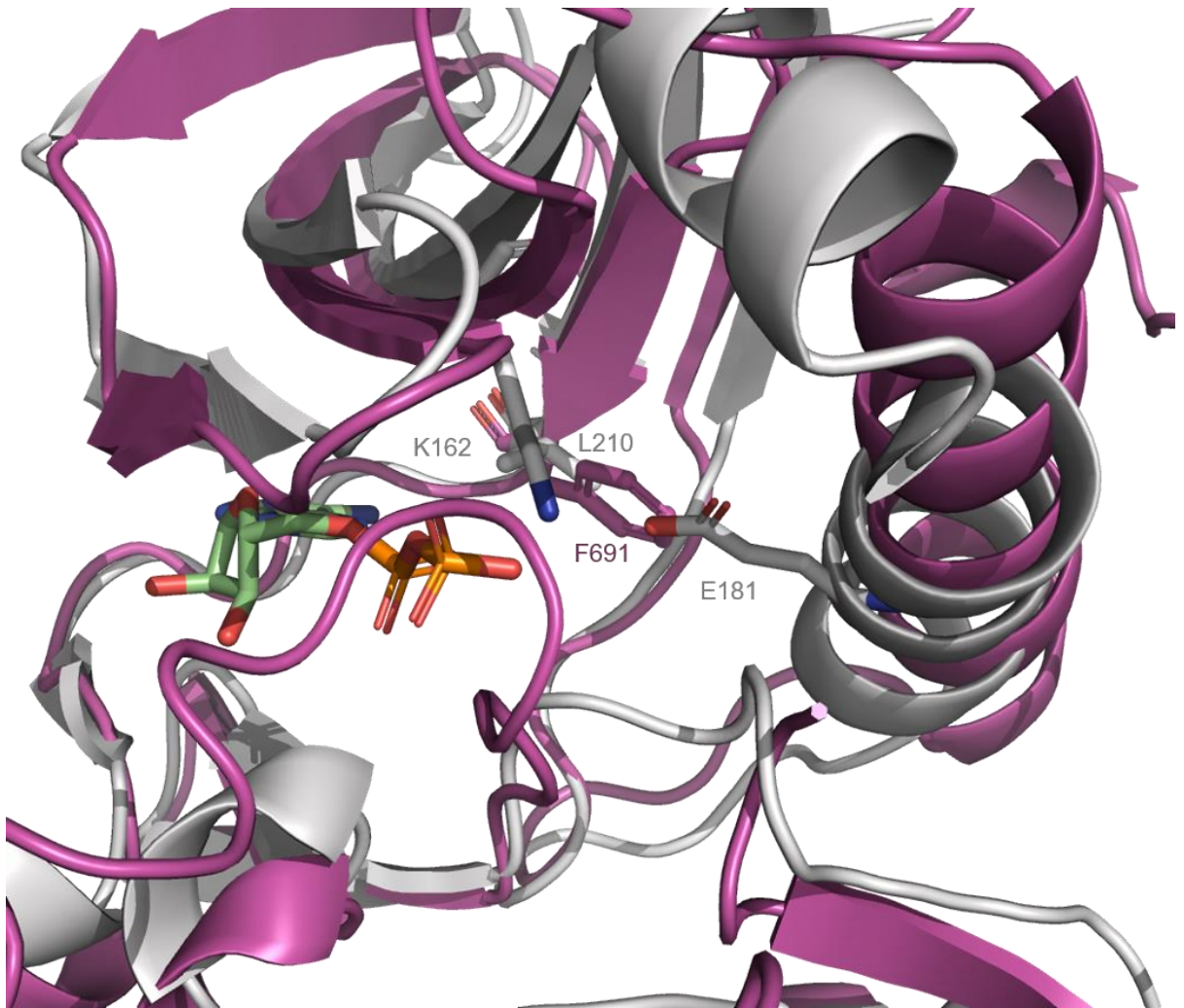
Figure 5.4 Alignment of Aurora A (grey, PDB 5G1X) and FLT3 (pink, PDB 5XO2) kinases with ADP (green sticks) and key side chains labelled

For a more quantitative analysis of the degree of conservation between related kinases, the amino acid sequences of Aurora A and the 109 most closely related kinases (Appendix) were aligned using the multiple sequence alignment programs MUSCLE (Multiple Sequence Comparison by Log-Expectation) and ClustalW.[111] The results from both methods were similar throughout the length of the protein, and agreed in almost all cases in the specific regions of interest (ATP-binding region and salt bridge). The conservation of the DFG loop (a subset is shown in Figure 5.5) was also used as a quick check to ensure the alignment was reasonable. If this highly conserved motif was not aligned, there may have been an error in the selection of sequences or a problem in the alignment. For all sequences chosen, the DFG loop was aligned as expected.

Protein																											
Aurora A	N	L	L	L	G	S	-	-	-	-	-	-	A	G	E	L	K	I	A	-	-	-	D	F	G	W	S
Aurora B	N	L	L	L	G	L	-	-	-	-	-	-	K	G	E	L	K	I	A	-	-	-	D	F	G	W	S
Aurora C	N	L	L	L	G	F	-	-	-	-	-	-	R	G	E	V	K	I	A	-	-	-	D	F	G	W	S
PLK4	N	L	L	L	T	R	-	-	-	-	-	-	N	M	N	I	K	I	A	-	-	-	D	F	G	L	A
SNRK	N	V	V	F	F	E	-	-	-	-	-	K	Q	G	L	V	K	L	T	-	-	-	D	F	G	F	S
MARK3	N	L	L	L	D	A	-	-	-	-	-	-	D	M	N	I	K	I	A	-	-	-	D	F	G	F	S
MARK4	N	L	L	L	D	A	-	-	-	-	-	-	E	A	N	I	K	I	A	-	-	-	D	F	G	F	S
Greatwall	N	M	L	I	S	N	-	-	-	-	-	-	E	G	H	I	K	L	T	-	-	-	D	F	G	L	S
PDK1	N	I	L	L	N	E	-	-	-	-	-	-	D	M	H	I	Q	I	T	-	-	-	D	F	G	T	A
ULK3	N	I	L	L	S	-	-	-	-	S	L	E	K	P	H	L	K	L	A	-	-	-	D	F	G	F	A
Ribo. S6	N	I	M	L	S	S	-	-	-	-	-	-	Q	G	H	I	K	L	T	-	-	-	D	F	G	L	C
MARK2	N	L	L	L	D	A	-	-	-	-	-	-	D	M	N	I	K	I	A	-	-	-	D	F	G	F	S
PLK2	N	F	F	I	N	E	-	-	-	-	-	-	S	M	E	L	K	V	G	-	-	-	D	F	G	L	A
SIK3	N	L	L	L	D	A	-	-	-	-	-	-	N	L	N	I	K	I	A	-	-	-	D	F	G	F	S
DCLK2	N	L	L	V	C	E	-	-	Y	P	D	G	T	K	S	L	K	L	G	-	-	-	D	F	G	L	A
PLK1	N	L	F	L	N	E	-	-	-	-	-	-	D	L	E	V	K	I	G	-	-	-	D	F	G	L	A
DCLK3	N	L	L	V	Q	R	-	-	N	E	D	K	S	T	T	L	K	L	A	-	-	-	D	F	G	L	A
CAMK1	N	L	L	Y	Y	S	-	-	L	D	E	D	S	K	-	I	M	I	S	-	-	-	D	F	G	L	S
AKT2	N	L	M	L	D	K	-	-	-	-	-	-	D	G	H	I	K	I	T	-	-	-	D	F	G	L	C
AKT3	N	L	M	L	D	K	-	-	-	-	-	-	D	G	H	I	K	I	T	-	-	-	D	F	G	L	C
SGK1	N	I	L	L	D	S	-	-	-	-	-	-	Q	G	H	I	V	L	T	-	-	-	D	F	G	L	C

Figure 5.5 A subset of alignment results using ClustalW, with the DFG loop shown to be conserved for all proteins. Amino acids are grouped using the MEGA-X colouring (A, F, I, L, M, V in yellow; N, Q, S, T, W in green; D, E in red; K, P, R in blue; H in pale blue; C in beige; Y in lime; G in purple).

All of the amino acids within 10 Å of either ADP or the salt bridge binding site of Aurora A were recorded. Using the alignments by MUSCLE and ClustalW, the corresponding amino acids for all other kinases were also noted. The percentage of proteins containing each amino acid at each position studied were calculated (a subset of this data is shown in Table 5.2).

Table 5.2 A selection of data from kinase alignment analysis showing the percentage of each amino acid at some of the positions studied in proximity of the salt bridge binding site. Highlighted figures show the percentage of proteins with the same amino acid as Aurora A in that position.

Res. No. (AurA)	162	163	164	165	166	167	170	172	173	175	176	178	179	180	181
Aurora A	K	V	L	F	K	A	E	A	G	E	H	L	R	R	E
A	0	3.6	0	1.8	1	7	9.5	6.6	5.8	0.9	13	12	0.9	4.5	0
C	0	2.7	0	0	0	0	0	0.9	0	0	0.9	3.7	0	4.5	0
D	0	0	0	21	2	9.3	8.1	10	15	0	5.5	0	1.9	0.9	0
E	0	7.3	0	5.5	2	8.1	18	7.5	10	10	28	0	10	10	100
F	0	5.5	1.8	2.7	3	0	0	1.9	0	0.9	0	4.6	10	0	0
G	0	0	0	0	0	3.5	6.8	0.9	16	1.9	2.8	0	1.9	0	0
H	0	1.8	0	0.9	1	1.2	0	0.9	2.3	0	4.6	0	1.9	0	0
I	0	29	42	0	2	1.2	4.1	1.9	0	9.4	0.9	22	14	0	0
K	100	14	0	18	61	16	19	17	14	4.7	5.5	0	6.5	14	0
L	0	4.5	26	0	11	10	4.1	3.8	2.3	13	2.8	24	7.4	3.6	0
M	0	2.7	10	0	1	0	0	1.9	0	0.9	2.8	1.8	7.4	1.8	0
N	0	0	0	13	0	2.3	2.7	2.8	9.3	0	1.8	0	0	13	0
P	0	0	0	10	0	8.1	5.4	1.9	2.3	5.7	0.9	0	2.8	0	0
Q	0	7.3	0	7.3	0	4.7	4.1	4.7	0	0.9	17	0	8.3	7.3	0
R	0	1.8	0	11	4	9.3	5.4	21	2.3	17	3.7	0	15	24	0
S	0	0.9	0	7.3	1	9.3	8.1	10	14	6.6	8.3	1.8	0.9	4.5	0
T	0	1.8	0	0	2	1.2	5.4	3.8	2.3	7.5	0.9	7.3	1.9	6.4	0
V	0	17	20	0.9	6.1	8.1	0	0	3.5	17	0.9	22	3.7	6.4	0
W	0	0	0	0	0	0	0	1.9	0	2.8	0	0	1.9	0	0
Y	0	0	0	0.9	3	0	0	0	0	0	0	0.9	3.7	0	0
Total	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100

(Row label for amino acid column: Amino Acids)

This analysis highlighted a few areas which were very highly conserved among the proteins studied. The salt bridge residues (Lys162 and Glu181 in Aurora A) were

lysine and glutamic acid in all other proteins in this study. This was an unsurprising result since this is a key feature across the kinase family. Other positions showed more variation. In the set shown above, position 165 of Aurora A is phenylalanine, but only 2.7% of the proteins studied also had phenylalanine in this position. Other kinases had a range of amino acids in this position, with the most common groups being aspartic acid (21%) and lysine (18%). Both of these more common side chains at this position are charged at physiological pH and these differences could be exploited in later development of inhibitors to acquire selectivity over related kinases.

His176 in Aurora A was another position which gave an interesting result in this analysis. Only 4.6% of proteins studied showed histidine in this position, with almost all amino acids seen at least once in other kinases. The most frequently seen amino acid was glutamic acid (28%). At physiological pH glutamic acid will be negatively charged, whereas the basic histidine side chain will be partially positively charged. Again, these differences could be exploited to improve selectivity at later stages of inhibitor design.

This information was used to generate a "heatmap" based on conservation of amino acids in Aurora A, with the most highly conserved residues shown in red and in green the Aurora A amino acids which were seen in the fewest other kinases (Figure 5.6). These maps highlighted some highly conserved areas, and some more flexible areas.
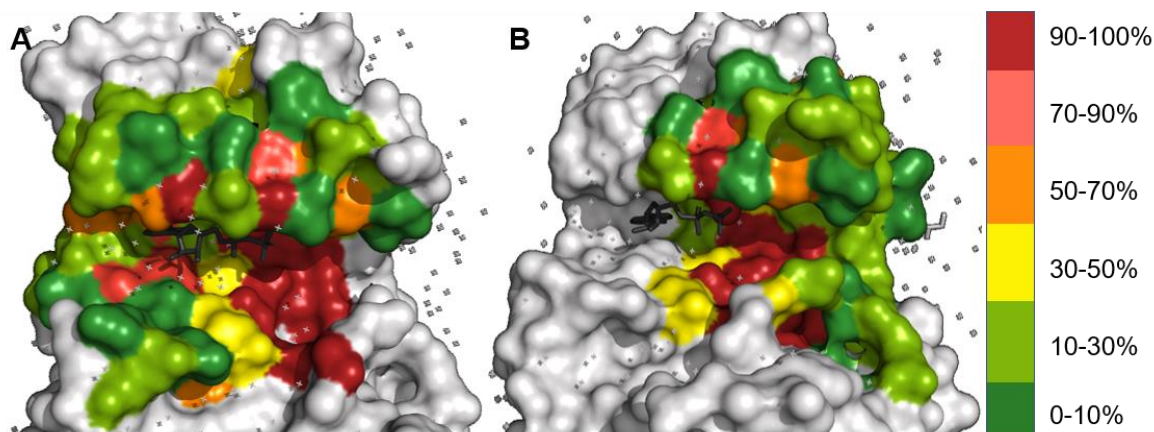
Figure 5.6 Amino acid conservation within 10 Å of (A) ADP and (B) the Glu-Lys salt bridge. The percentage of proteins studied with the same amino acid as Aurora A at each position is shown by different colours.

However, the properties of each amino acid may also play a role. For example, a negatively charged acidic side chain such as glutamic acid or aspartic acid could both interact with a positively charged component of a small molecule inhibitor. In this case, it might be more useful to group amino acids by their properties as opposed to treating each amino acid individually. A number of grouping methods were considered and the percentages of proteins with each position's amino acid group was calculated (a subset is in Table 5.3). A similar heatmap was generated at each target site using this grouping (Figure 5.7) and showed more conservation of amino acid type, as expected. The ATP binding site is highly conserved between kinases, so a small molecule which relies on this binding may also be able to interact with a number of other kinases, making selectivity difficult to achieve. At the salt bridge site, there are still a number of highly conserved amino acids, but there is greater variation between kinases starting a short distance from the salt bridge site. It could be possible to grow a fragment which binds to the salt bridge site to a larger, more drug-like molecule which could be capable of exploiting these variations to gain selectivity over other kinases.

Table 5.3 A selection of kinase alignment analysis showing the percentage of proteins with amino acids grouped as shown in proximity of the salt bridge binding site. Highlighted figures show the group Aurora A is found in at that position.

	Res. No. (AurA)	162	163	164	165	166	167	170	172	173	175	176	178	179	180	181
	Aurora A	K	V	L	F	K	A	E	A	G	E	H	L	R	R	E
Amino Acids	HKR	100	17	0	30	66	27	24	39	19	22	14	0	23	37	0
	DE	0	7.3	0	26	4	17	26	18	26	10	34	0	12	11	100
	AILMFWYV	0	63	100	6.4	27	27	18	18	12	45	20	87	49	16	0
	STNQ	0	10	0	27	3	17	20	22	26	15	28	9.2	11	31	0
	C	0	2.7	0	0	0	0	0	0.9	0	0	0.9	3.7	0	4.5	0
	G	0	0	0	0	0	3.5	6.8	0.9	16	1.9	2.8	0	1.9	0	0
	P	0	0	0	10	0	8.1	5.4	1.9	2.3	5.7	0.9	0	2.8	0	0
	Total	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
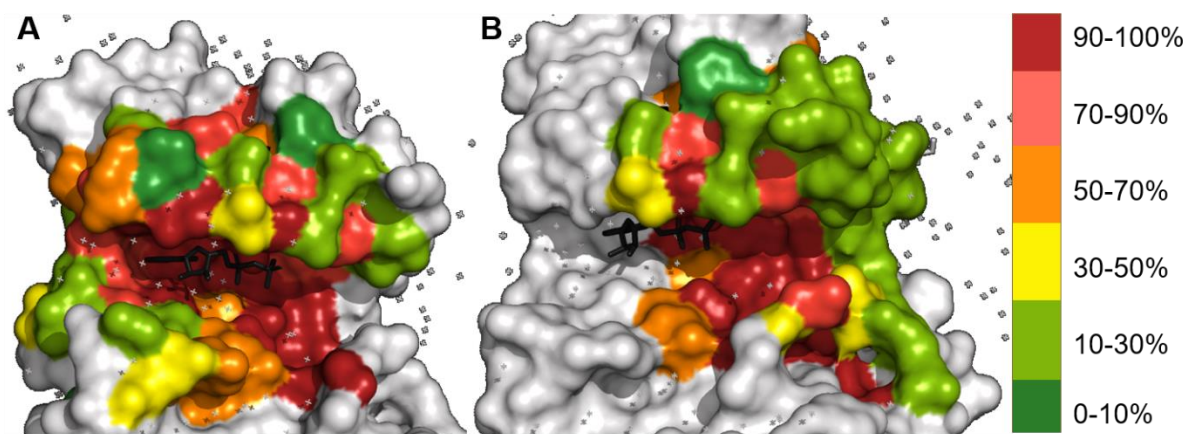
Figure 5.7 Conservation of amino acids using the grouping system described. Percentages of proteins with amino acids of the same group as Aurora A at each position are shown by different colours.

One of the potentially interesting findings from this grouping was Lys143. This residue was lysine for only 4.5% of the proteins studied (Aurora A, B and C, Myosin light-chain kinase and Mixed lineage kinase 2). When this was combined with other basic amino acids providing positively charged side chains at physiological pH, only 8.2% of proteins were in this group. This particular amino

acid was even more interesting as it was also identified through virtual screening as being able to make hydrogen bonding interactions with multiple fragments docked at the salt bridge site (Section 5.2, Figure 5.13).

5.2 Docking

Compounds were docked at the salt bridge site using the crystal structure of Aurora A in a complex with vNAR-D01, a nanobody used to stabilise the inactive conformation (PDB 5L8L). Both fragment and larger compound libraries (around 152 000 compounds) were docked with and without water to explore the predicted impact of any water-mediated interactions and to compare the results from these two conditions which are predicted to bind at the site regardless of the presence or absence of certain water molecules. At this site, docking was also performed both with and without ADP (**Error! Reference source not found.**). Docking compounds with crystallised ADP present would hopefully avoid the identification of ATP-competitive inhibitors which bind to the hinge region of the ATP-binding site and may help to identify compounds which are predicted to bind at the same time as ADP. Comparing results (docking poses and Glide scores) for compounds with and without ADP could enable the filtering of compounds which are predicted to preferentially bind to the hinge region, over the intended target site at the salt bridge.
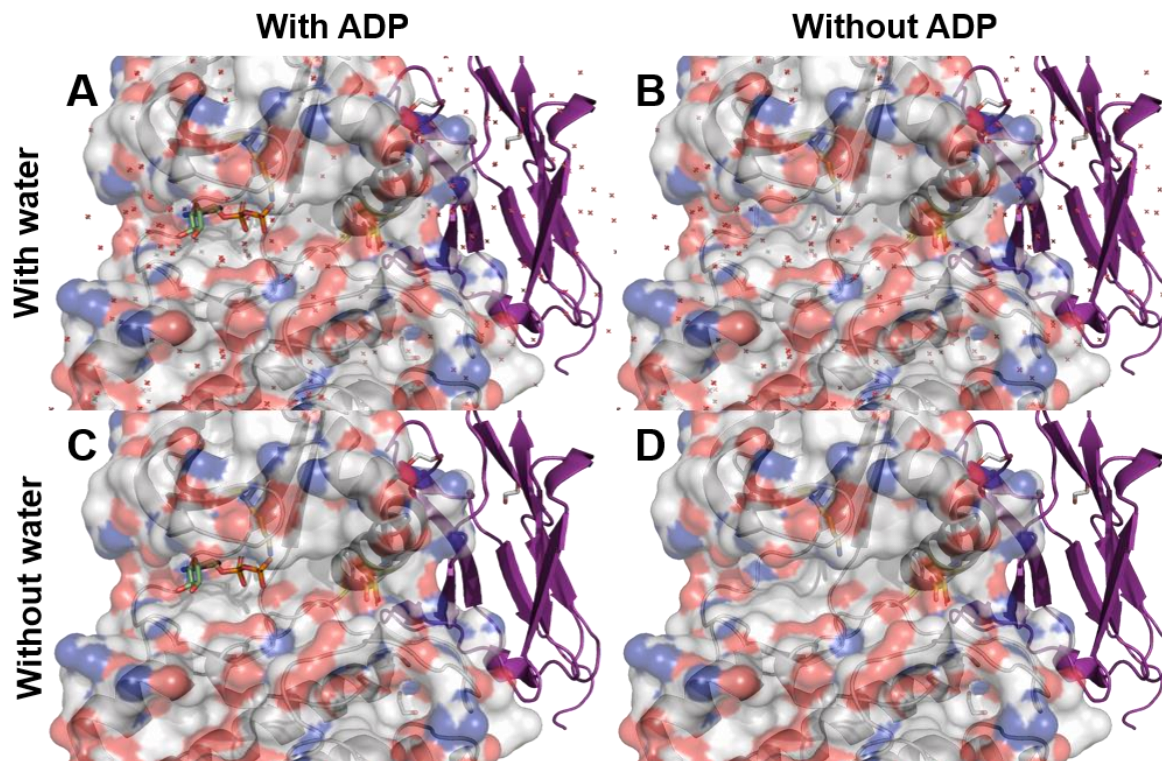
Figure 5.8 The docking conditions used in docking at the salt bridge site, with salt bridge residues highlighted (yellow sticks), and ADP (green sticks), vNAR-D01 (purple) and water (red crosses) shown

For the first round of virtual screening, crystallised ADP was kept in place. When fragment libraries were docked most compounds were predicted to bind close to the desired target site, near the salt bridge residues Lys162 and Glu181. Some poses even predicted hydrogen bonding interactions with the salt bridge side chains (Figure 5.9). Binding poses and Glide scores were similar regardless of whether crystallised water molecules were included or excluded. When water was included, most predicted hydrogen bonding interactions were with Aurora A side chains, with only a small number of water-mediated interactions. This lack of reliance on water-mediated interactions could explain the similar results between docking conditions.
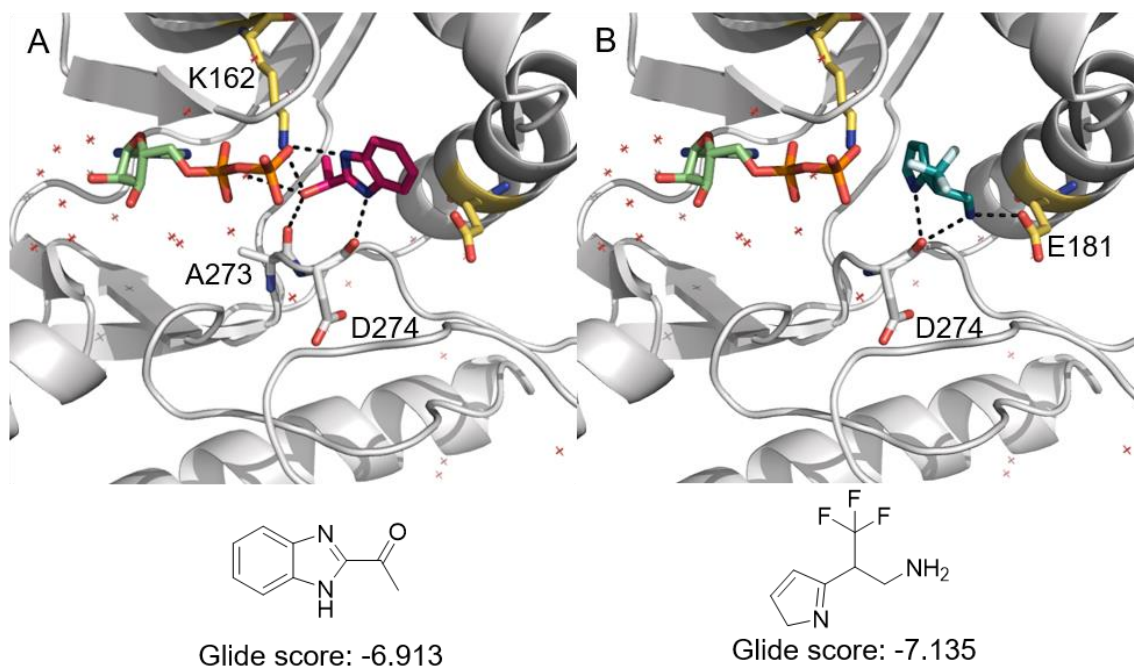
Figure 5.9 Selected docking poses from fragment libraries with water, showing predicted hydrogen bonding interactions with the salt bridge residues Lys162 and Glu181 (yellow sticks). Other predicted interacting residues are labelled. (A) Predicted interaction with Lys161 (compound ID 12P-660, Bionet); (B) Predicted interaction with Glu181 (compound ID SEW04444, Maybridge).

A number of fragments were predicted to form interactions to residues within the flexible activation loop of Aurora A (Figure 5.10). Under virtual screening conditions, the crystallised conformation of the activation loop was not flexible, so compounds capable of interacting with these residues could be favoured. However, these kinds of interactions may not be seen *in vitro*, when this loop is flexible. When filtering compounds using these results, docking predictions might be more reliable for those which do not rely on activation loop interactions.
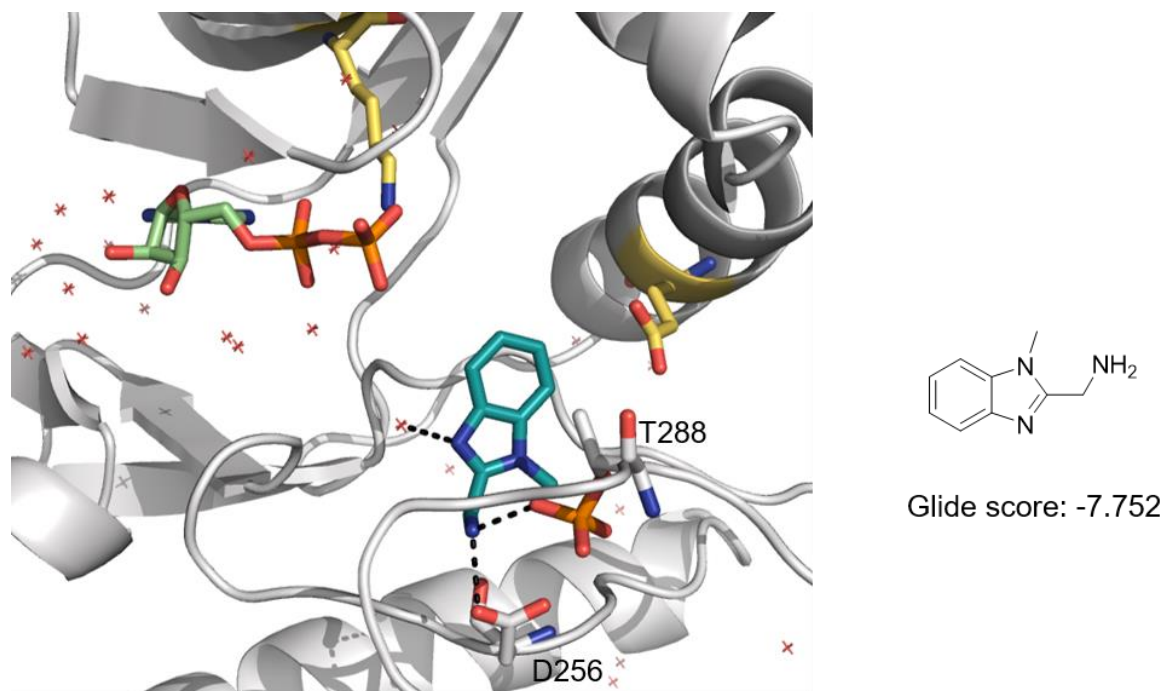
Figure 5.10 One example of a fragment docking pose predicting interactions with the flexible activation loop (compound ID BAS 00226633, Asinex). Salt bridge residues are highlighted (yellow sticks) and interacting residues are labelled.

When larger (non-fragment) compound libraries were docked, similar interactions were seen with the activation loop as were seen with fragment docking. Again, the removal of crystallised water had little effect on the Glide scores or poses generated. As these libraries contained larger compounds, most extended from the salt bridge site either towards the activation loop or towards the back of the ATP binding pocket (Figure 5.11).
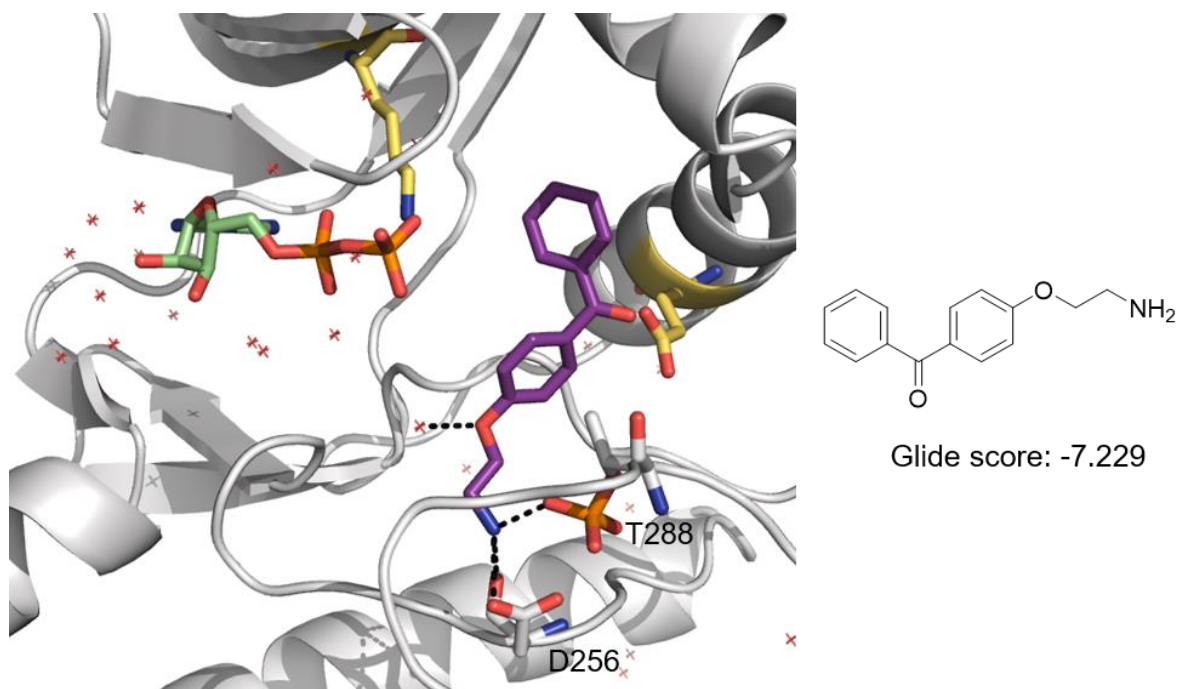
Figure 5.11 One example of drug-like compounds docking with poses extending from the salt bridge region (compound ID 9192554, MCCB). Salt bridge residues are highlighted (yellow sticks) and interacting residues labelled.

Docking was also performed with ADP removed for two reasons. Primarily, this was to investigate whether compounds were predicted to bind more strongly in the ATP binding pocket than the target salt bridge site. With ADP in place, compounds might have generated docking poses at the target site which might preferentially bind to the ATP binding pocket when given the option. To do this, the same docking grid coordinates and dimensions were used which covered both areas so poses at either site were possible.

Secondly, these conditions could generate interesting results in their own right. Particularly for larger, more drug-like compound libraries to find poses predicted to disrupt the salt bridge whilst also able to occupy the ATP binding site. This would be more likely to be ATP-competitive, but may still be worth taking forward for *in vitro* testing.

As expected, a lot of the highest scoring results gave poses at the ATP binding site, for both fragments and larger compounds. The Glide scores were generally in the same range as when ADP was included, with a few of the highest results reaching scores of -9. Reassuringly, some of the top fragment poses were still centred at the salt bridge target site and did not favour the ATP binding hinge (Figure 5.12).

Glide score: -7.973   Glide score: -7.257
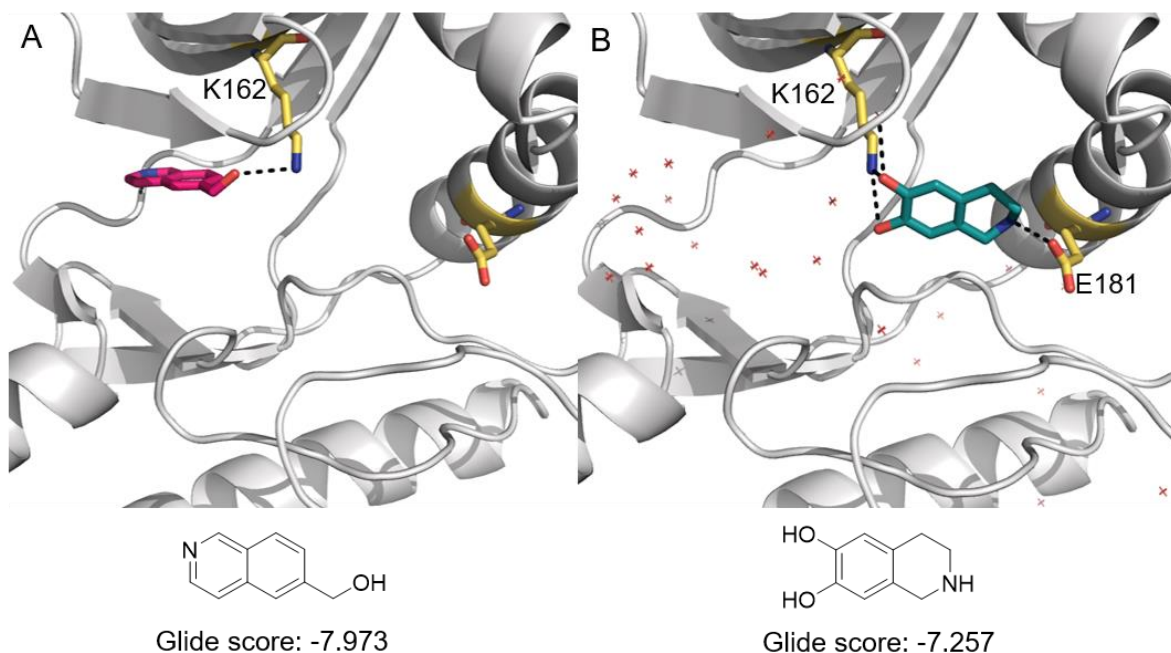
Figure 5.12 Docking results with ADP removed showing; (A) predicted pose from the Bionet fragment library (ID PS-4737) in the ATP-binding pocket; (B) A Maybridge library fragment (ID AC23506) predicted to bind in the salt bridge region. Interacting residues are labelled and salt bridge residues are highlighted (yellow sticks).

Many of the larger compounds were predicted to bind to the ATP binding site and extend to the salt bridge. With water included, a number of water-mediated hydrogen bonding interactions were predicted, whereas without water, more direct interactions to the hinge region Ala213 were predicted (Figure 5.13).
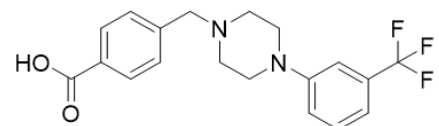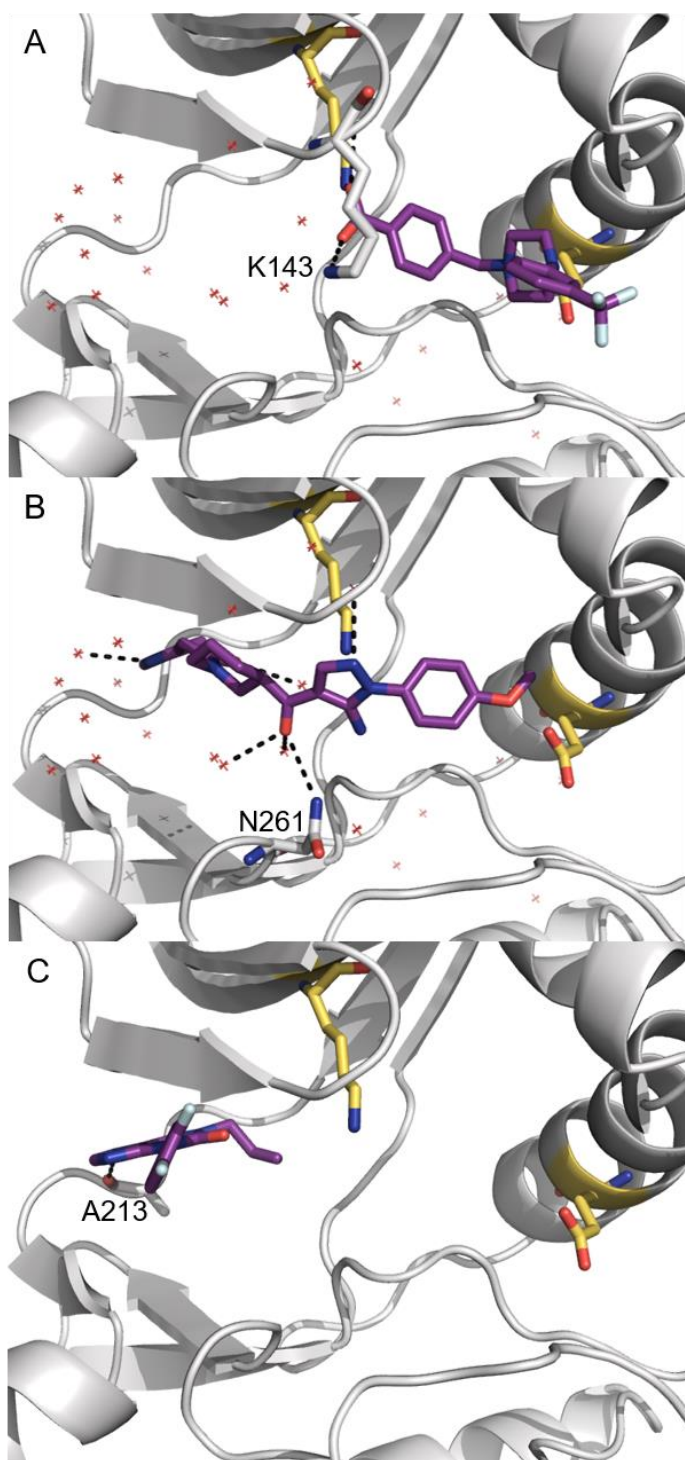
Glide score: -8.513
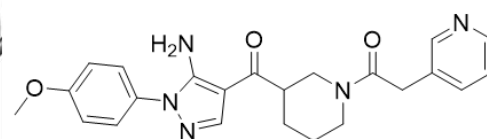
Glide score: -8.311
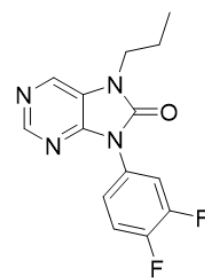
Glide score: -8.420

Figure 5.13 Docking results from drug-like compound libraries with ADP removed;
(A) Some results were predicted to extend towards and even beyond the salt
bridge region (compound ID ALB-H05615918, MCCB); (B) Many compounds
were predicted to form hydrogen bonding interaction with water (compound ID
LMK 22209945, MCCB); (C) With water removed, more interactions to Aurora
A amino acids were predicted (compound ID ALB-H05231315, MCCB). Salt
bridge residues are highlighted (yellow sticks) and interacting residues are
labelled.

If these predictions are accurate there could be a number of fragments able to bind
at the salt bridge region without interfering with ADP binding. The docking poses
and Glide scores were used to identify a subset of small molecules to analyse
further using crystallography and the Caliper mobility shift assay. The first identified
subset included fragments with the highest Glide scores which satisfied a number
of prerequisites. First, any fragments with PAINs scaffolds were removed from
consideration. Fragments also had to show good docking poses when docked with
ADP in place, with hydrogen bonding interactions predicted with surrounding
residues without relying on interactions with the flexible activation loop, which may
not be viable when this loop is free to move. Finally, poses and Glide scores were
compared between docking with and without ADP present. If a fragment looked
promising with ADP present, but showed a higher Glide score and a pose which
overlapped with ADP when removed, this fragment was not added to the final
subset. Overall, around 130 fragments were taken forward in the first set.

5.3 Crystallography

Current published crystal structures of Aurora A in complex with vNAR-D01 have
all used a mutated form of Aurora A in which the two surface-exposed cysteine
residues were mutated to alanine (C290A:C393A, PDB 5L8L).

The formation of Aurora A/vNAR-D01 crystals was attempted using wild-type
Aurora A. Each of the two proteins were expressed separately in *E. coli* using

constructs available in the Bayliss lab. Both of the proteins were His-tagged, facilitating purification by IMAC, with a final purification by GF (Gel filtration) chromatography. SDS-PAGE was used to confirm the purification of each protein (Figure 5.14).
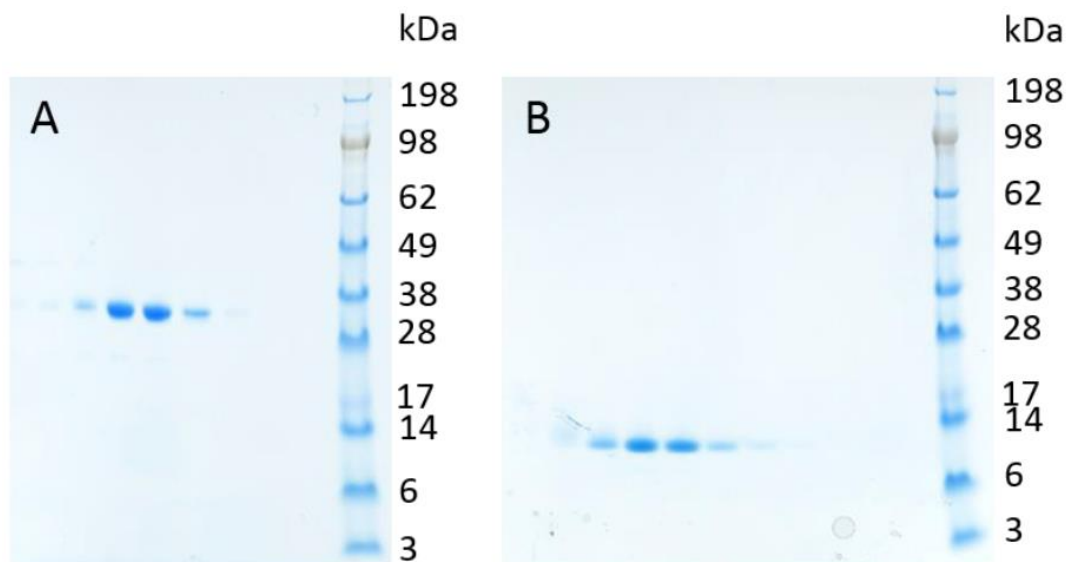
Figure 5.14 SDS-PAGE results after final purification; (A) Aurora A following purification (expected mass 32.8 kDa); (B) vNAR-D01 following purification (expected mass 12.7 kDa)

The proteins were combined with an excess of vNAR-D01 and the resulting complex was isolated by GF (Figure 5.15). The complexes were co-crystallised with either ADP, the Aurora family inhibitor VX-680 or acetyl CoA in the ATP binding site.
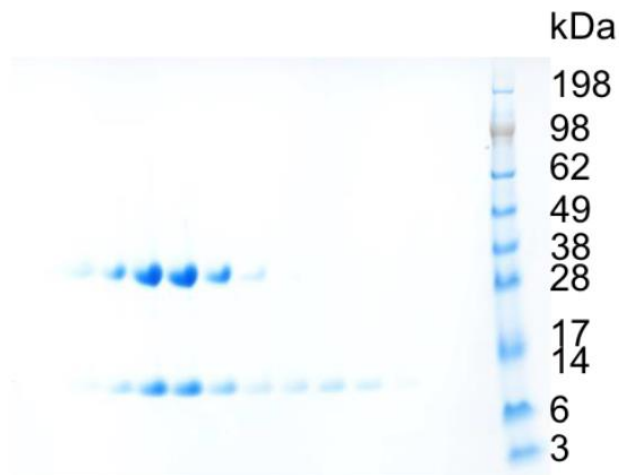
Figure 5.15 SDS-PAGE results of Aurora A/vNAR-D01 complex: the larger Aurora A/vNAR-D01 complex can be seen along with an excess of the smaller protein, vNAR-D01

Crystallisation conditions were tested using JCSG-*plus*, PACT premier and Classics Lite crystallisation screens.[112,113] The chosen screens assess a variety of conditions including pH ranging from 4.0 to 10.5, various salts and buffers at a range of concentrations and many different precipitants, again at a range of concentrations.

Following incubation of the complexes for a few days at 25 °C to allow crystal formation, viable crystals were only observed with ADP co-crystallised under one set of conditions (0.1 M citrate buffer, pH 5.0, 20% w/v PEG 6000).
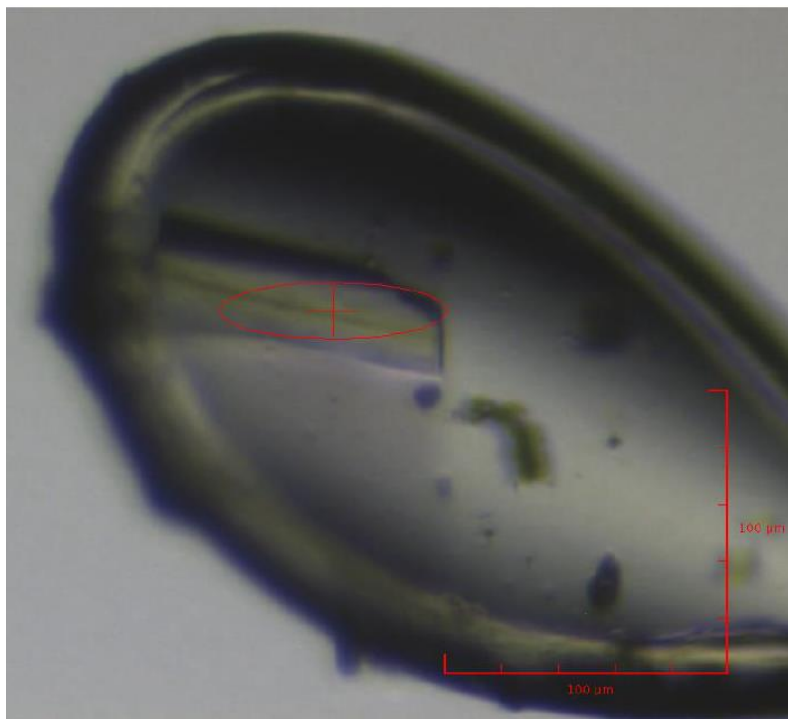
Figure 5.16 Wild-type Aurora A/vNAR-D01 crystal as observed mounted on beam
I04 (beam size 80.0 x 20.0 μm)

X-ray crystallography, performed at Diamond, of crystals grown under these
conditions produced data at 1.70 Å. This data was refined in *Coot* (Crystallographic
Object-Oriented Toolkit).[114] The conformation of the kinase was not altered by the
addition of the C290A:C393A mutations. Wild-type crystals were also more
sensitive to damage at lower concentrations of DMSO than the mutated version,
which is stable even at 20% DMSO. The stability of the crystals at higher DMSO
concentration means a higher concentration of fragments (stored in DMSO) can be
soaked into the C290A:C393A crystals. The fact that these mutations do not cause
changes to the protein conformation meant crystal soaking experiments were
performed using the more stable, mutated version of Aurora A/vNAR-D01 crystals.

Based on virtual screening results, an initial set of 130 fragments was selected to
send for crystal soaking at 5 mM fragment concentration, exposing the crystals to
5% DMSO. The majority of fragments did not bind to the crystals. Four compounds
showed binding to other areas of the protein, but did not appear in the salt bridge
site (Figure 5.17). One fragment (BTB10042, **22**) bound at the salt bridge site. The

binding mode of this fragment showed similarity to the predicted pose from virtual screening (Figure 5.18), giving some confidence in the ability of the virtual screening to predict the binding pose of fragments. R-free and B-factor values are provided in appendix 4.
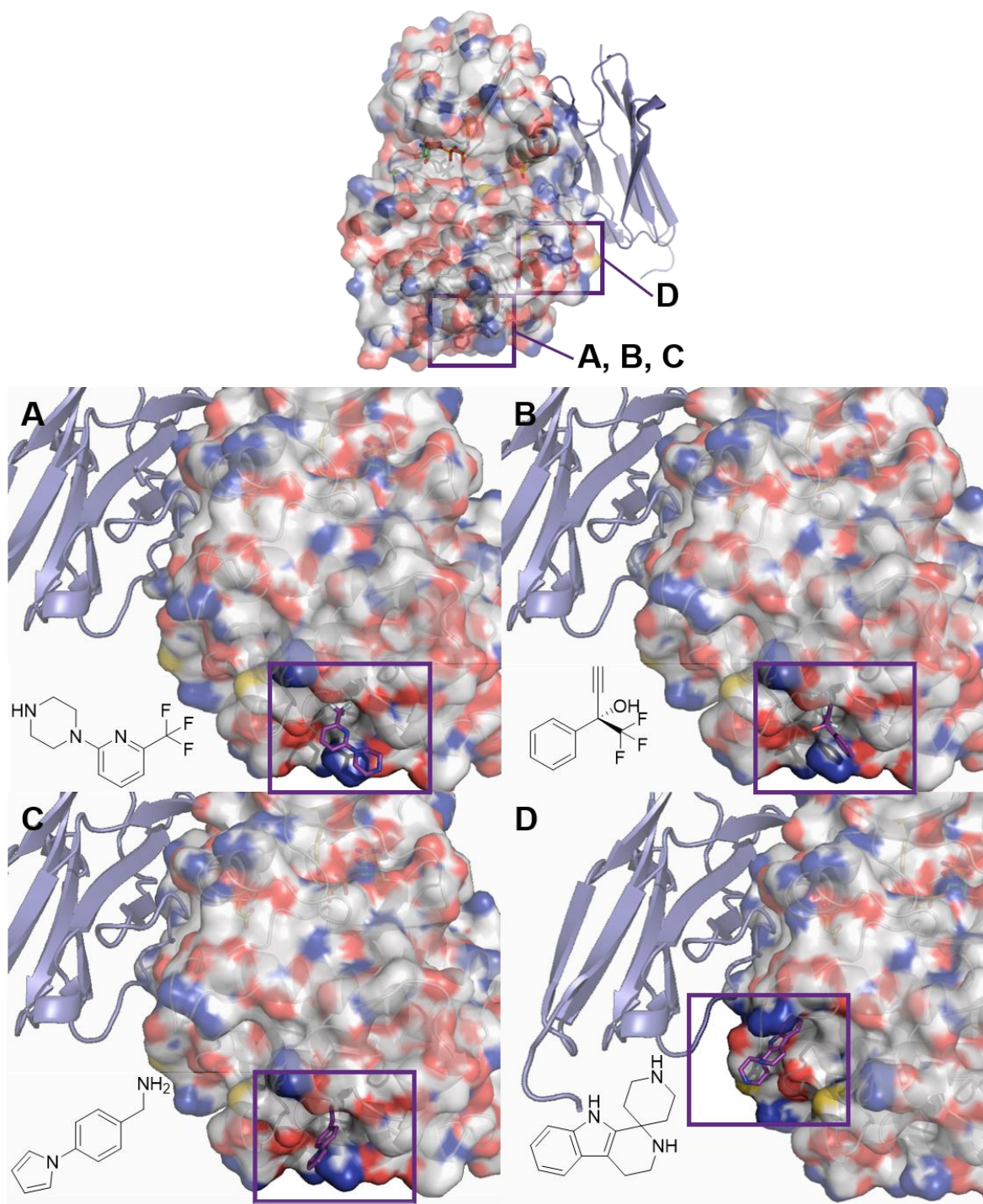
Figure 5.17 Crystal soaking results from the first set of fragments soaked into Aurora A/vNAR-D01 crystals (Compound IDs (A) SEW03804, (B) AC36409, (C) CC25513, (D) HTS09269, all Maybridge library)
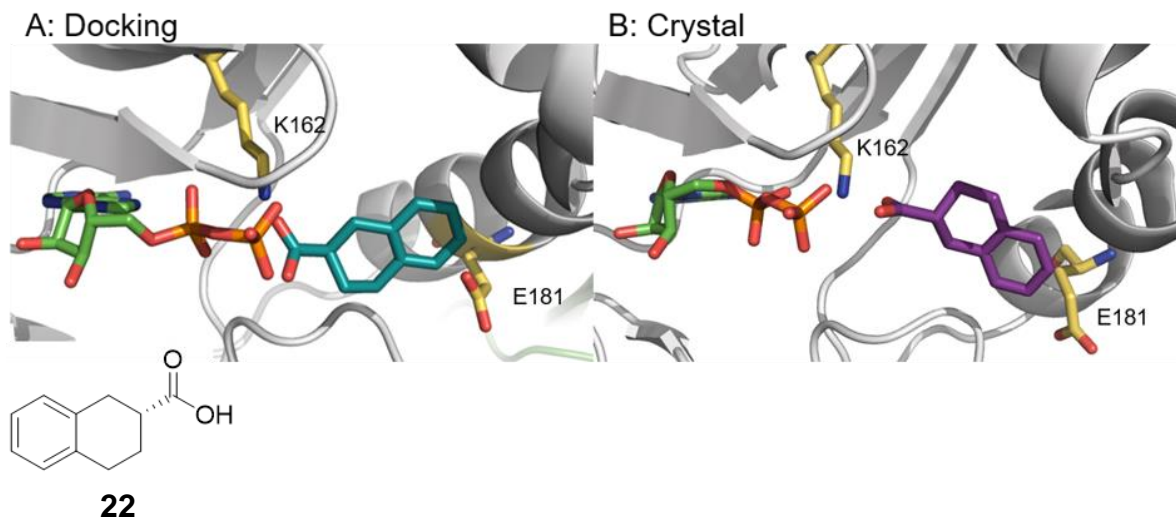
Figure 5.18 A comparison of; (A) the docking pose and (B) crystal soaking binding of **22**

Following the identification of **22** bound at the target site, this crystal structure was used to revisit the binding site analysis. This time, the binding site of interest was identified using **22** as the ligand and gave a SiteScore of 1.12. This score is well above the "druggable" 0.8 score and is close to the SiteScore of the ATP binding site (1.14, see Chapter 4), again giving confidence that this site may be a useful target site for small molecules.

The binding pose of **22** indicated the carboxylic acid group may be important for forming interactions with Lys162. Armed with this information, the virtual screening results were revisited and fragments which matched all of the previous constraints (limited predicted interaction with the activation loop residues, etc.), hadn't been tested in the first set of compounds and contained a carboxylic acid group positioned in a similar orientation were chosen. These measures, combined with the fact that the most promising, highly scoring compounds had been tested in the first round of soaking, meant these compounds often had significantly lower Glide scores. This resulted in a set of 85 compounds for the next round of crystal soaking. A random selection of other in-house fragments which had not scored highly enough in virtual screening to be prioritised was also chosen. If the virtual

screening approach was useful, this set should not show any binding to Aurora A. In total, 159 fragments were sent for crystal soaking at this point, which had not been selected from virtual screening.

Unfortunately, this second set did not show any compounds binding at the salt bridge site. However, five compounds did bind at other areas. Some of these fragments bound at sites known to be involved in Aurora A function (e.g. TACC3, Figure 5.19A), but some were found at previously uninvestigated sites (Figure 5.19B). The impact of small molecule interactions at these sites on the activity of Aurora A was unknown, so although these fragments did not bind as predicted, they were still taken forward for further investigation of their effects on the kinase activity of Aurora A.
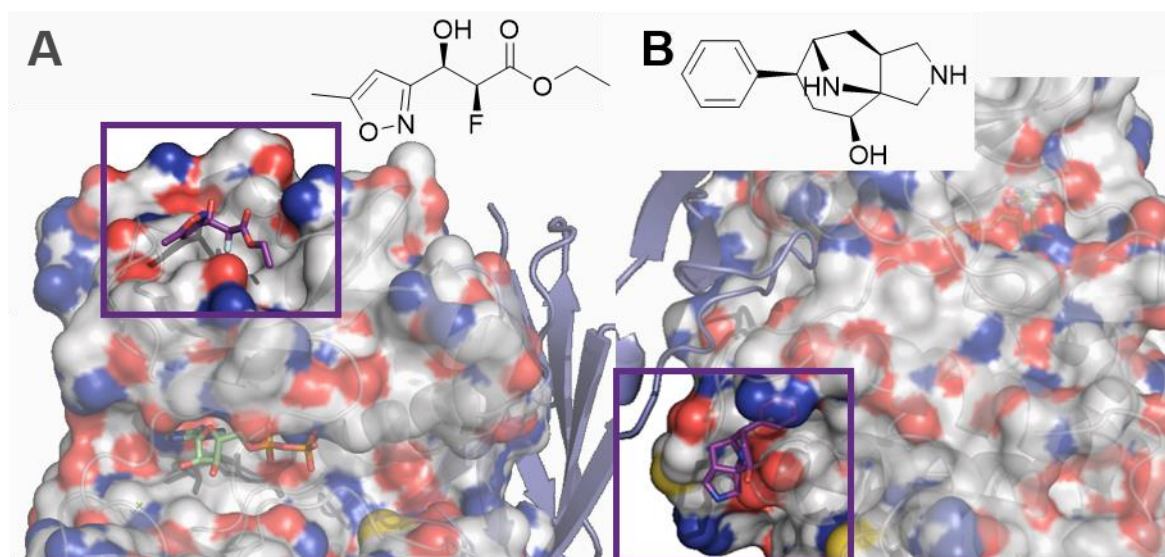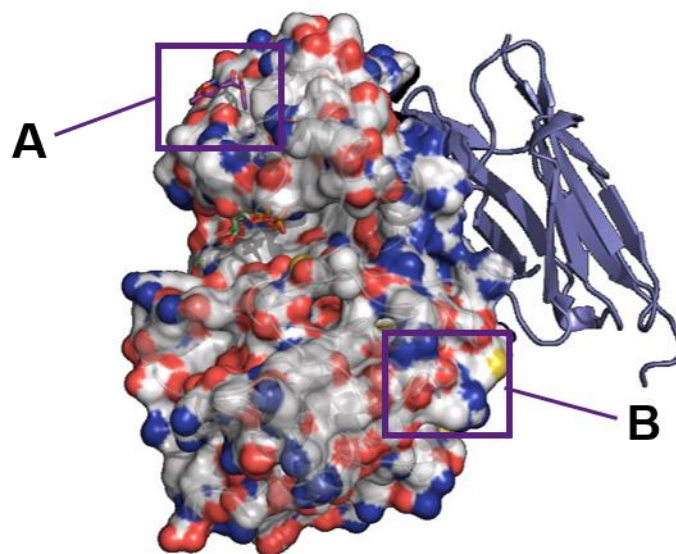
Figure 5.19 Crystal soaking results from the second round of fragments. (A) One of the fragments bound near the TACC3 binding site (compound ID LDS-034188); (B) One of the fragments bound at a previously uninvestigated site (compound ID LDS-034251).

A set of 300 compounds from an in-house fragment library chosen with a focus on diversity were all sent for crystal soaking. This library had not been used for docking studies. These compounds were all stored at a lower concentration of 50 mM and were soaked at a final concentration of 2.5 mM. This set provided no hits

from crystallography bound at the salt bridge site. Only one fragment bound anywhere on Aurora A (Figure 5.20). Although this was disappointing, it did help to illustrate the advantages of the virtual screening in identifying useful sets of compounds to help narrow down larger libraries.
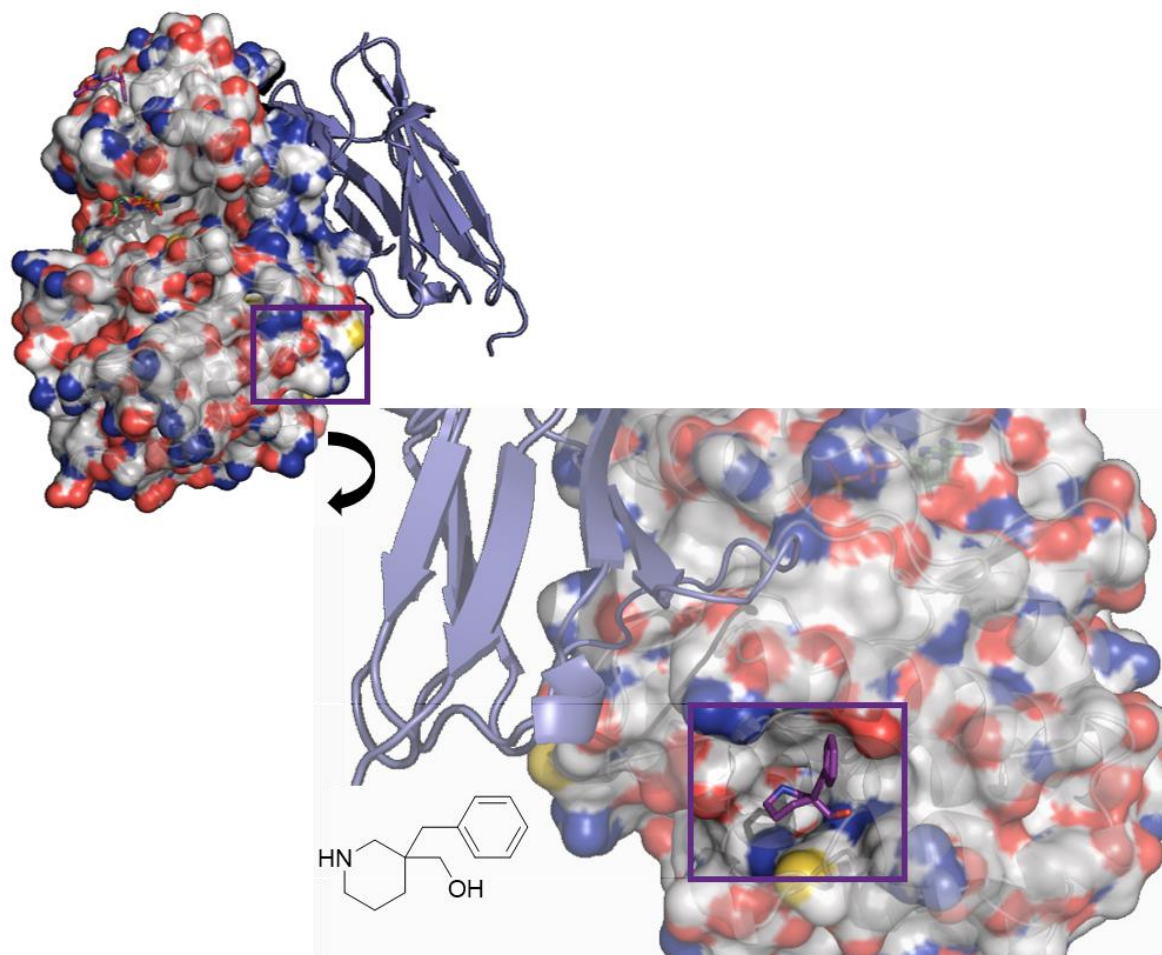
Figure 5.20 The only fragment found through crystal soaking using the diversity-focussed in-house fragment library (compound ID Z1270137020)

Based on the one hit bound at the salt bridge site by crystallography, and a hit identified through enzyme activity assays run in parallel (Section 5.4), a set of analogues were chosen to begin to explore structure activity relationships through crystallography. This work was subsequently disrupted by Covid-19.

5.4 Enzyme Activity Assays

The effect of interesting virtual screening hits and any fragments successfully soaked into Aurora A/vNAR-D01 crystals on the kinase activity of Aurora A was studied using the Caliper Mobility-Shift Assay, which uses electrophoresis to monitor enzyme kinetics. A fluorescent peptide substrate of Aurora A was used. Multiple "sips" are taken over time and any phosphorylated product and unreacted substrate are separated by charge. Negatively charged phosphorylated product is detected first, followed by unreacted substrate, and the ratio of fluorescence intensity in these two peaks provides a percentage conversion. Monitoring the change in product formation over time can be used to generate enzyme kinetics data. Alternatively, reactions can be quenched after a set time and measurements taken at a single time point.

The 127 virtual screening hits identified in the first round of selection (based on Glide score, docking pose and the comparison of docking results for the same fragment when ADP was included and excluded) were assessed first. Type III inhibitors at this site should not compete with ATP binding, so fragments which gave better Glide scores without ADP present, or whose binding poses showed an overlap with where ADP would bind if it were present, were removed.

The most interesting virtual screening hits based on these criteria were tested at a single concentration (5 mM) where possible for their effect on Aurora A activity. The more stable C290A:C393A Aurora A mutant was used in all Caliper tests and an enzyme titration was performed to identify the optimal protein concentration for greatest sensitivity (Figure 4.11). An ATP K_m determination was performed to identify the concentration of ATP to use, again for greatest sensitivity (Figure 4.12).

Following a 30 minute incubation of Aurora A, ATP and fluorescent peptide substrate with each of the 127 fragments, the extent of substrate conversion to phosphorylated product was measured. Aurora A activity without the addition of a fragment was normalised to 100% (Figure 5.21). Plates were set up in duplicate with each measurement taken twice.

In the first set of virtual screening hits tested, most of which were from the in-house fragment library, a range of activities was seen (Figure 5.21), usually with small error. Virtual screening only provided predictions of a compound's ability to bind in a given area, so it is perhaps unsurprising that only a small proportion of those tested showed inhibition of Aurora A activity.
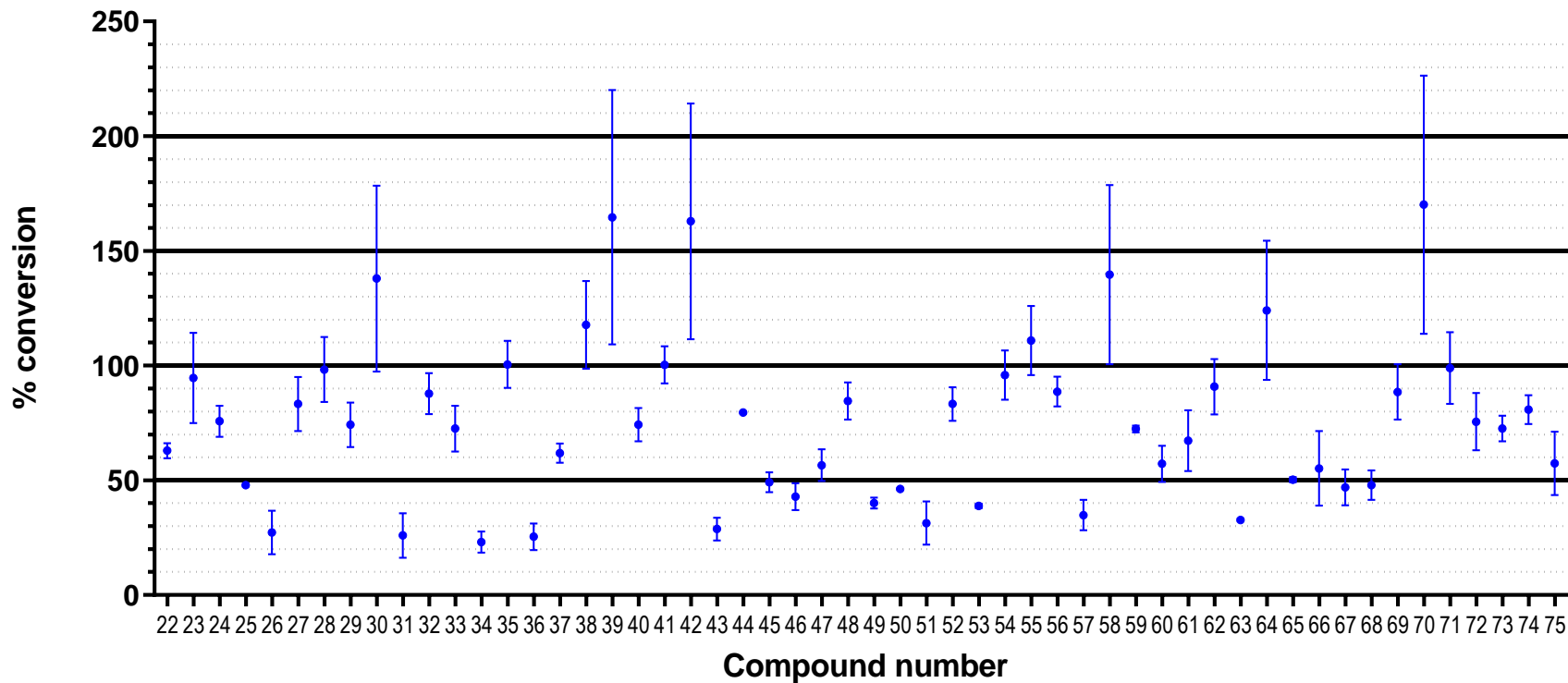
Figure 5.21 Substrate conversion results with the first set of compounds tested in the Caliper mobility shift assay. 100% conversion was normalised to uninhibited Aurora A control.

In the second set of virtual screening hits (Figure 5.22), some of the in-house compounds studied had been stored at lower concentrations, so the intended 5 mM could not be used without increasing the final DMSO concentration to above 5%. These compounds were tested twice; once with a 5% DMSO concentration, which caused the inhibitor concentration to drop below 5 mM, and once at 5 mM, which resulted in a DMSO exposure above 5%. In this set, some of the errors were much larger, especially for those with higher DMSO concentration (Figure 5.23).
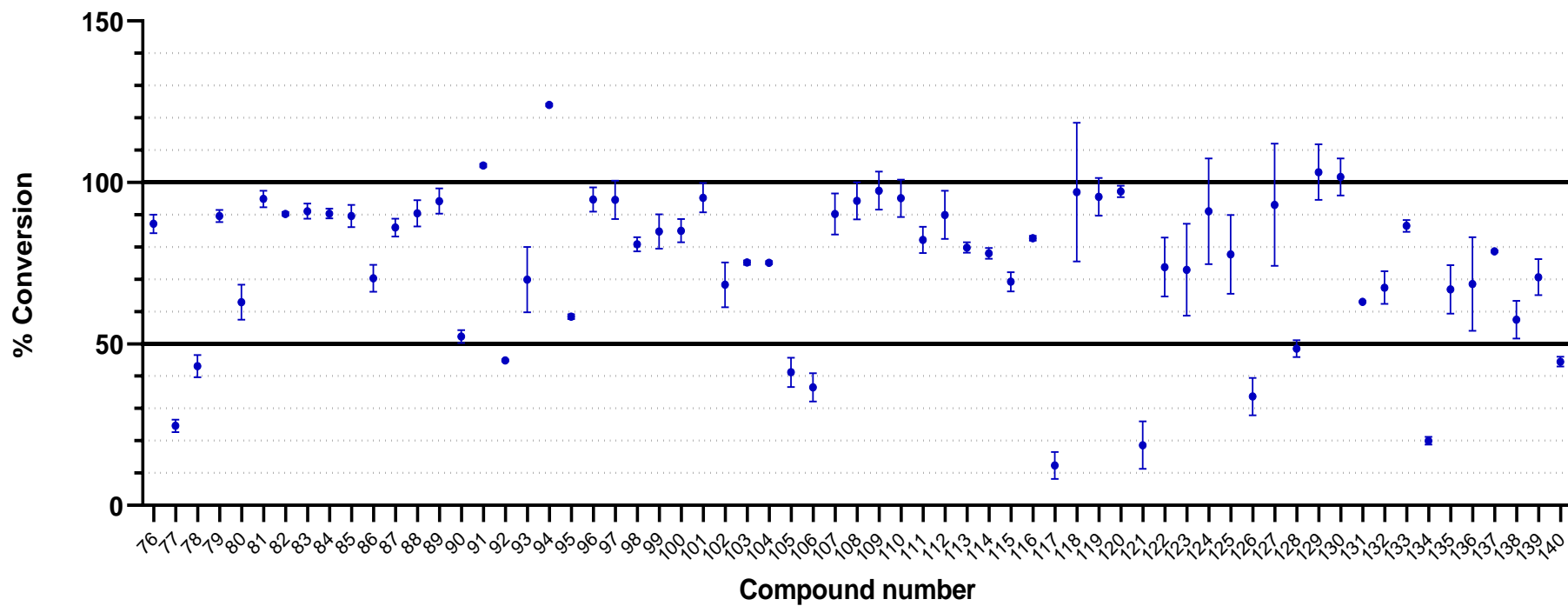
Figure 5.22 Substrate conversion results for the second set of compounds (stored at 100 mM) tested in the Caliper mobility shift assay at a final concentration of 5 mM. 100% conversion was normalised to uninhibited Aurora A control.
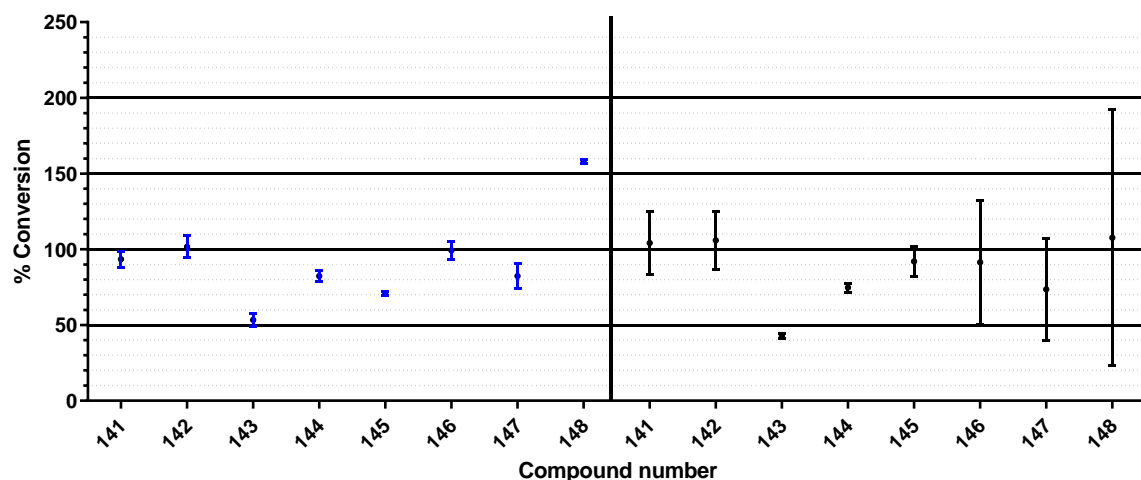
Figure 5.23 Caliper mobility shift assay results for compounds stored below 100 mM; (Left, blue) Compounds tested at 5 mM and higher DMSO concentration; (Right, black) Compounds tested at 5% DMSO and lower than 5 mM. 100% conversion was normalised to uninhibited Aurora A control.

Overall, 28 compounds did reduce Aurora A kinase activity to below 50%, giving a first "hit-rate" of 22%. This is higher than might be expected from a purely randomly selected set of fragments, suggesting the filtering by virtual screening was helpful.

Interestingly, a few compounds produced Aurora A activity above 100%. Some of these had very large error bars, perhaps suggesting a problem in the setup or measurement of some samples. However, some did show high Aurora A activity without large errors. These results were interesting enough to put forward for further investigation, along with all compounds which resulted in Aurora A activity under 25% (a hit rate of 6% for all compounds tested in the single point assay).

IC_{50} experiments were performed, again using the Caliper Mobility Shift Assay, this time with ten 3-fold serial dilutions of each fragment with a top concentration of 5 mM (when fragments were soluble at 5 mM). The three fragments which showed a possible activation of Aurora A activity in the single point assay (**39**, **42** and **148**) failed to show the same relationship in this assay. For most of the fragments which

showed an inhibition of activity in the single point assay, it was not possible to generate an IC_{50} curve. In some cases, some inhibition was seen at the highest concentration. However, without testing at even higher concentrations, which was not possible due to solubility issues, the curve could not be completed or an IC_{50} value determined.

Three of the fragments investigated did generate data which could be used to find an IC_{50} value (Figure 5.24, Table 5.4). These values were in the millimolar range, as might be expected from small fragments. The crystal soaking hit, **22**, was also tested in this assay and gave an IC_{50} of 0.6 mM. This fragment had the additional benefit of a crystal structure with Aurora A.
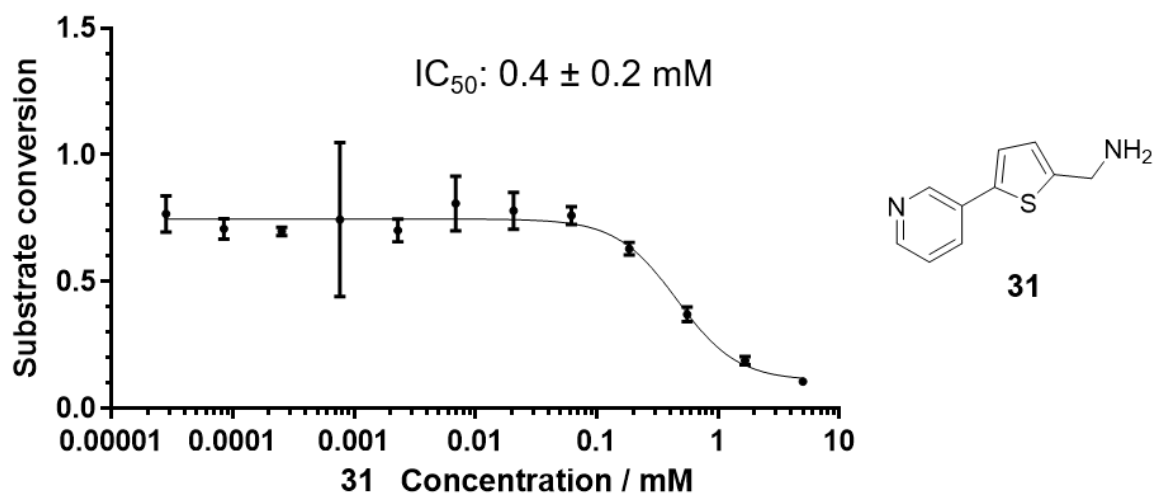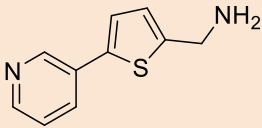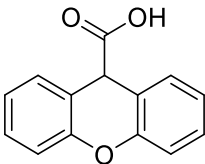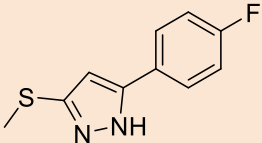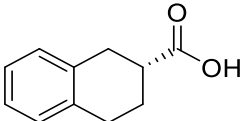
Figure 5.24 Graph of proportion of substrate converted to phosphorylated product over increasing concentration of **31**

Table 5.4 IC$_{50}$ values for fragments which gave a full curve

Compound	IC$_{50}$	Structure
31	0.4 ± 0.2 mM	
57	0.6 ± 0.2 mM	
126	0.3 ± 0.2 mM	
22 (BTB10042)	0.6 ± 0.1 mM	

As these experiments only showed inhibition of Aurora A activity, the next step was to attempt to determine whether this inhibition was from direct competition with ATP, or whether there was a non-competitive or uncompetitive mode of inhibition.

5.5 ATP Competition

ATP competition experiments were conducted for the three tested compounds which gave clear IC$_{50}$ curves. The ATP K$_m$ was determined for Aurora A with each fragment included at the IC$_{50}$ concentration and at double the IC$_{50}$ concentration. A competitive inhibitor of Aurora A would raise the ATP K$_m$ with higher inhibitor concentration whilst maintaining the maximum rate of substrate conversion, V$_{max}$. This was seen for two of the fragments tested (**31** and **126**, Figure 5.25-7), with an increase of ATP K$_m$ with inhibitor present from Aurora A without inhibitor (Table 5.5).
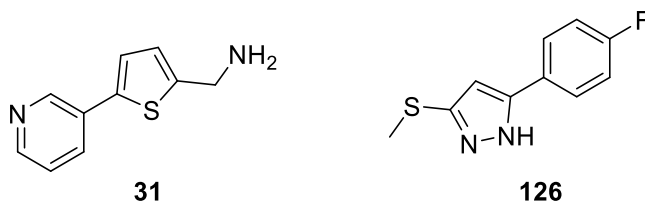
Figure 5.25 Structures of the fragments **31** and **126**
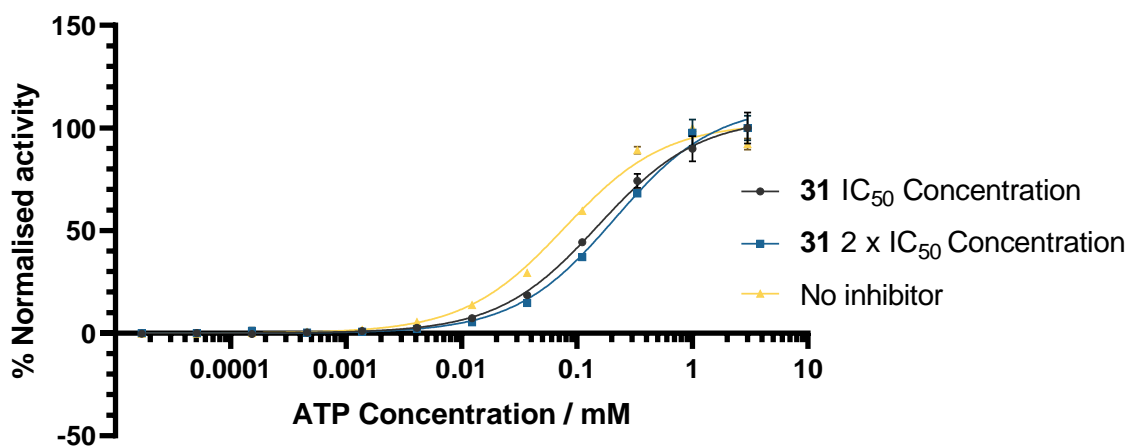
Figure 5.26 ATP K_m determination with **31** at its IC_{50} concentration, and at $2 \times IC_{50}$ concentration
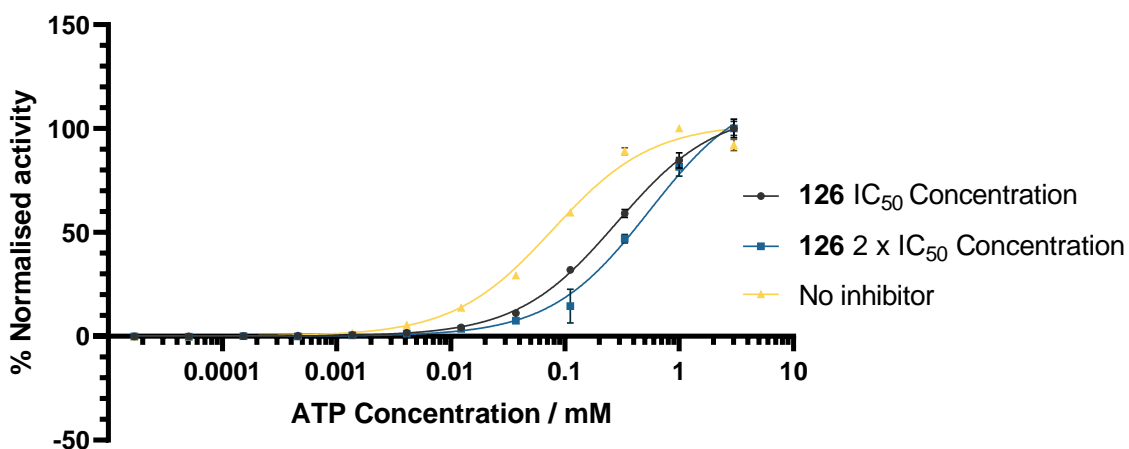
Figure 5.27 ATP K_m determination with **126** at its IC_{50} concentration, and at $2 \times IC_{50}$ concentration

Table 5.5 ATP K_m results for the compounds analysed for ATP competition

Compound	ATP K_m (mM)		
	No inhibitor	Inhibitor at IC_{50} concentration	Inhibitor at $2 \times IC_{50}$ concentration
31	0.07 ± 0.01	0.15 ± 0.01	0.21 ± 0.02
57	0.07 ± 0.01	0.05 ± 0.01	0.07 ± 0.01
126	0.07 ± 0.01	0.28 ± 0.01	0.55 ± 0.05

For the third fragment, the addition of inhibitor did not alter the ATP K_m from that observed with uninhibited Aurora A (Figure 5.29, Table 5.5). This suggested this fragment (**57**) was a non-competitive inhibitor of Aurora A. The docking pose of this fragment showed some similarity to that of **22** (Figure 5.28), and the crystal structure of **22** soaked into Aurora A/vNAR-D01 (Figure 5.18). Unfortunately, when **57** was soaked with these crystals, no crystal structures with the inhibitor bound were produced. The docking poses suggest the carboxylic acid groups in each of the most promising fragments (**22** and **57**) might have a key role in their activity. The development of some SAR data around these two hits became the next focus.

Glide score: -6.227   Glide score: -5.948
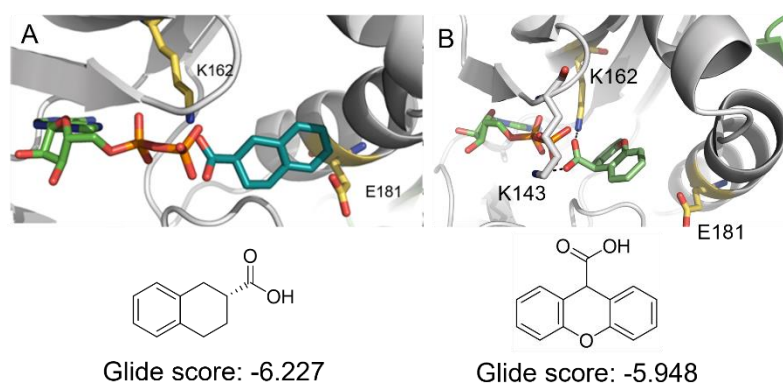
Figure 5.28 A comparison of the docking poses of (A) **22**; and (B) **57**
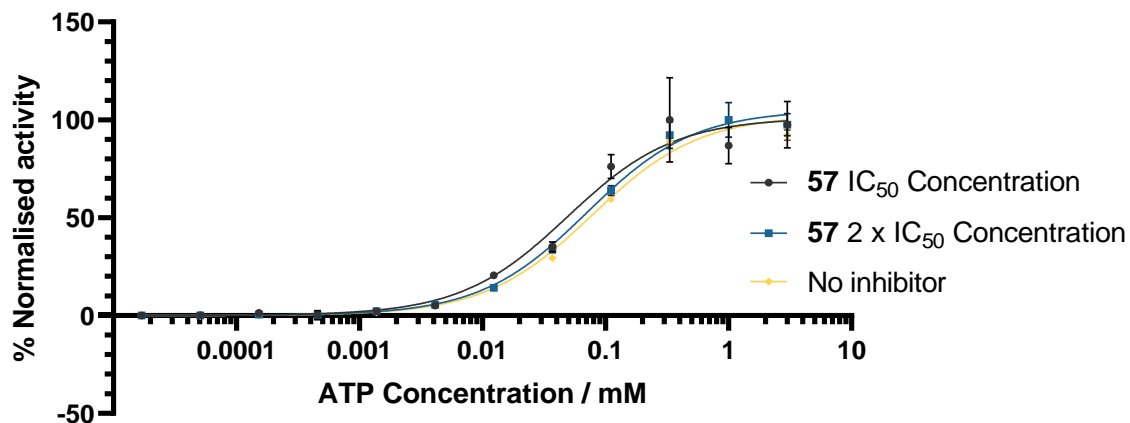
Figure 5.29 ATP K_m determination with **57** at its IC_{50} concentration, and at $2 \times IC_{50}$ concentration

5.6 Further Investigation of Hit Compounds

From the two hits identified above, a selection of commercially available analogues were chosen (Figure 5.30, Figure 5.31) to explore the potential for further development of these fragments. Quick, one- or two-step syntheses were also performed to expand this selection where possible.
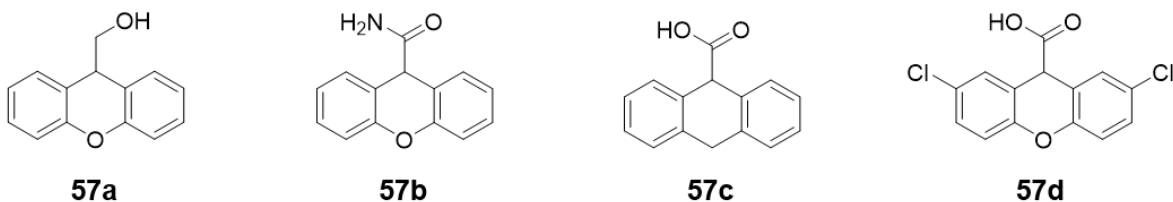
Figure 5.30 Commercial compounds based on **57**
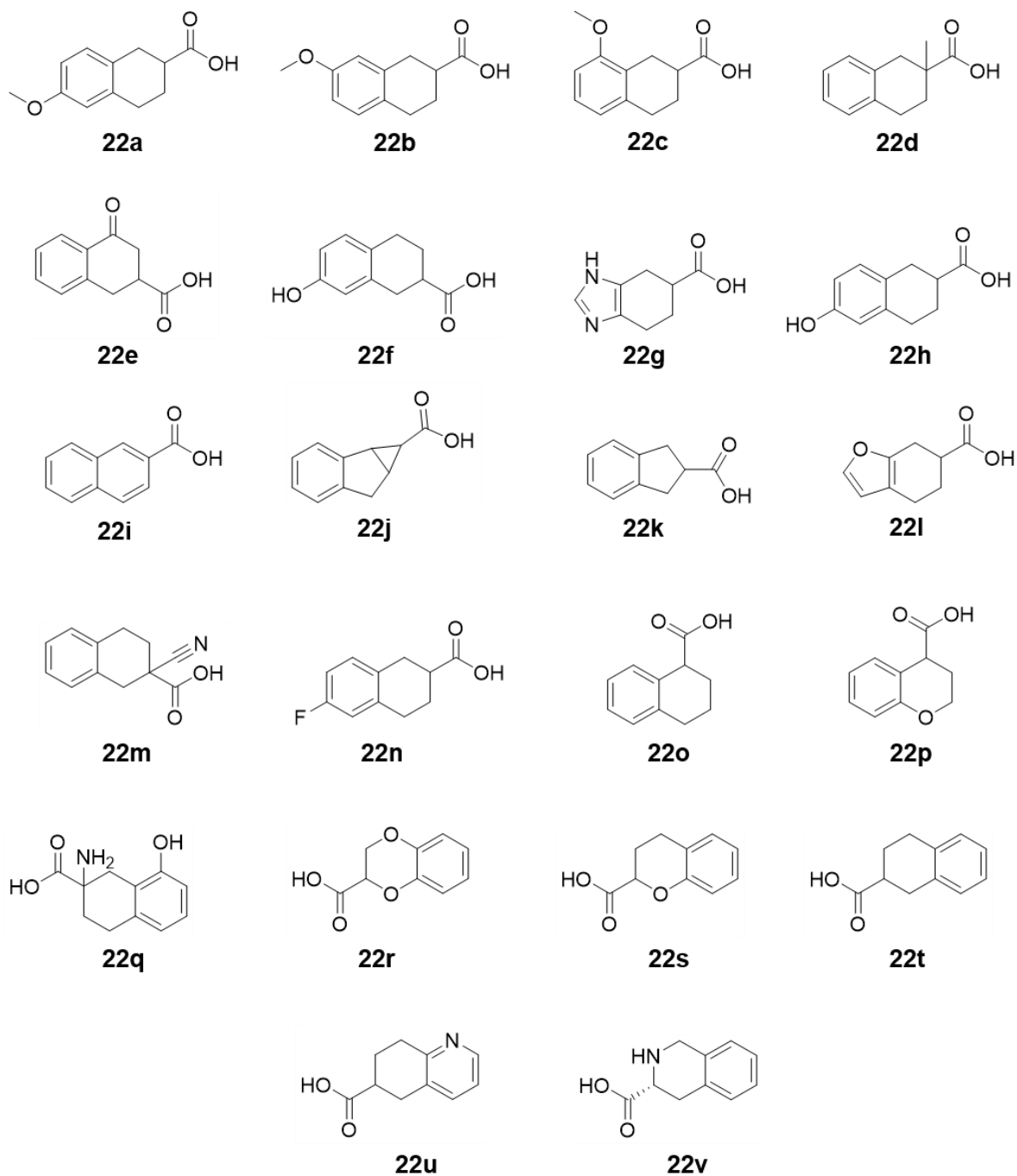
Figure 5.31 Commercial compounds based on **22**

Synthesised compounds included an ester variant of **57**, with the hypothesis that removing the carboxylic acid, which appeared to be important from the docking poses, would produce an inactive compound, thus demonstrating the importance of

the carboxylic acid group. Purchased **57** was reacted with methanol in 1,4-dioxane with catalytic hydrochloric acid at room temperature overnight (Scheme 2).
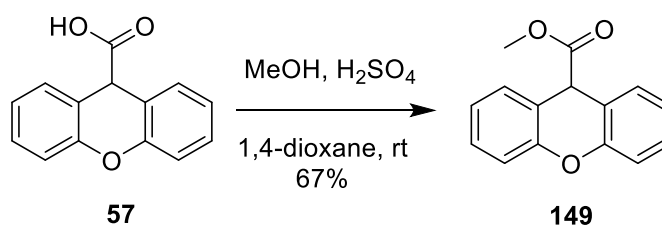
Scheme 2: Esterification of 57

This compound was assessed for its effect on Aurora A kinase activity using the Caliper mobility shift assay, and did not generate a dose-response curve or IC$_{50}$ value. This indicated an importance of the carboxylic acid group, as predicted by the docking poses and suggested by the crystal structure of **22**.

The synthesis of a series of analogues of **22** was also planned, using commercially available starting materials (Scheme 3), to expand the SAR knowledge surrounding the fragment, with the hope of identifying further fragments with the potential to begin to grow **22** into a more drug-like compound with increased activity. Although these reactions were completed, final products were not fully purified before Covid-19-induced lab shutdown.
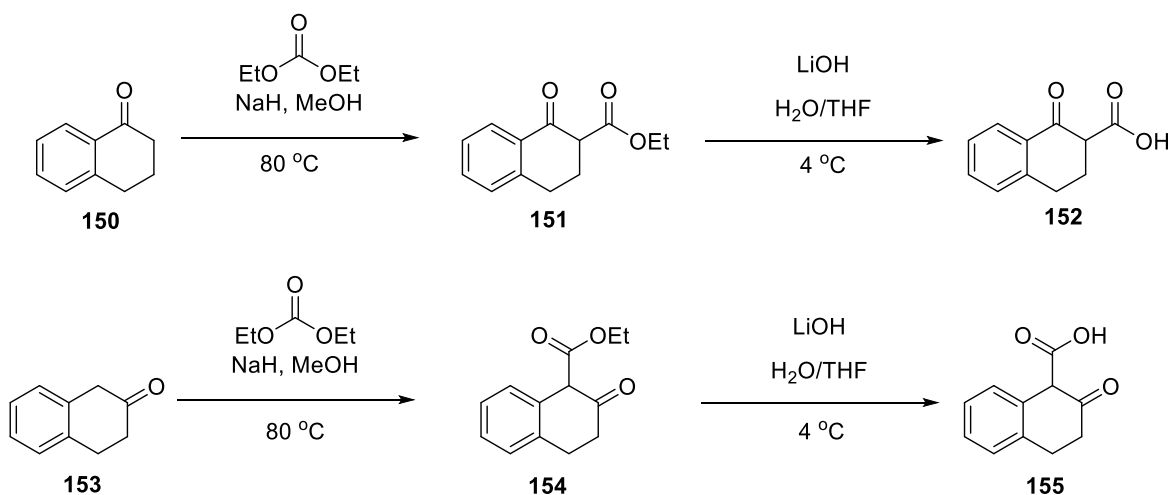
Scheme 3: Planned synthesis of analogues based on 22

Commercially available analogues of both **22** and **57** were purchased and prepared as 100 mM stock solutions in DMSO. Three of these compounds were screened using the Caliper mobility shift assay (Figure 5.32, experiments performed by Christopher Arter upon lab re-opening during Covid-19) and IC_{50} values determined (Table 5.6).
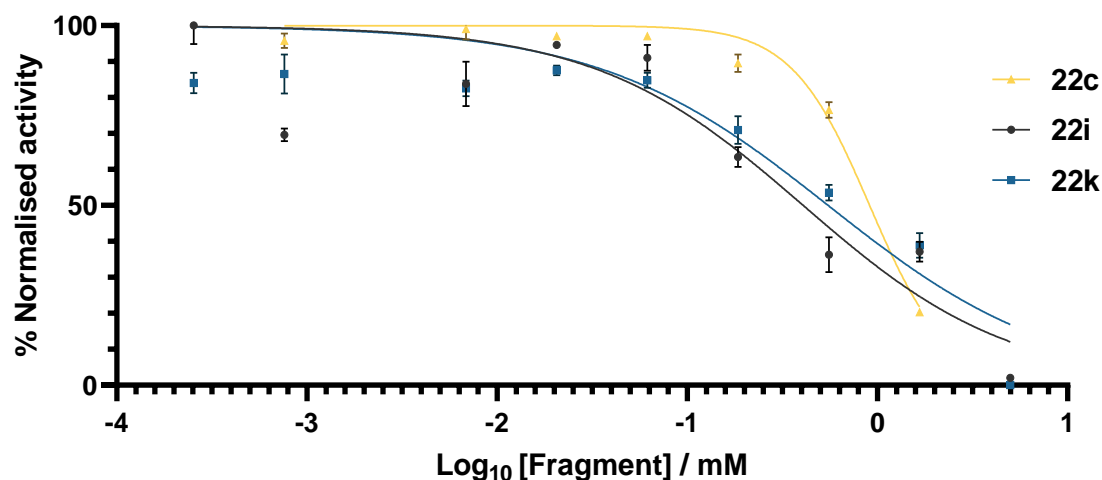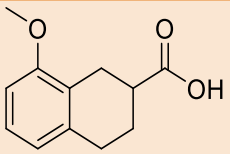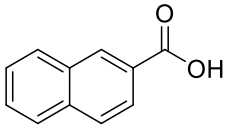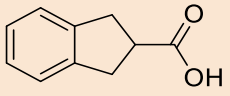
Figure 5.32 Graph of Aurora A activity over increasing concentrations of **22c**, **22i** and **22k**

Table 5.6 IC_{50} values for the tested analogues of **22**

Compound	IC_{50}	Structure
22c	0.9 ± 1 mM	
22i	0.4 ± 1 mM	
22k	0.5 ± 1 mM	

Errors in these values were much higher than had been seen in other experiments, partly due to the incomplete IC_{50} curves generated. It would be useful to repeat these experiments with higher fragment concentrations if possible. If these compounds do show inhibition of Aurora A kinase, especially in a non-competitive manner with respect to ATP, this would be further support of the viability of **22** as a starting point for the design of a type III Aurora A inhibitor.

This section of work was disrupted by Covid-19 and the resulting lab shutdown, so the remaining purchased and synthesised compounds were not screened for IC_{50} determination or ATP competition investigation, and some synthesis plans could not be completed.

5.7 Summary

Based on literature findings and the kinase alignment performed here, type III inhibition showed promise as an approach to kinase inhibition with the potential to introduce selectivity over other kinases and may present opportunities to disrupt the Aurora A/N-myc protein-protein interaction. Docking studies helped to narrow down the number of compounds taken forward for crystal soaking and *in vitro* assays to a manageable number. The fragment hits identified, one of which was shown to be non-competitive with ATP, could be useful starting points for SAR studies, which were slowed down in this work by Covid-19 and the resulting lab shutdown.

Chapter 6 Conclusions and Future Work

6.1 Conclusions

The current, variable treatments for neuroblastoma, ranging from observation through to aggressive multimodal treatment involving classic, non-specific chemotherapy including doxorubicin, cyclophosphamide, and etoposide, among others, reflect the variability of this disease, from symptoms, disease progression and ultimate long-term outcomes. Unsurprisingly, the most intensive treatment pathways in such young patients result in the highest toxicity, with adverse effects sometimes persisting way beyond the end of treatment and leading to chronic health effects for survivors of this cancer.[8, 15,115]

More targeted treatments with fewer toxic effects are required to improve survival, particularly for high-risk neuroblastoma, and to minimise long- and short-term side effects. Disrupting the Aurora A/N-myc interaction could be a useful approach to solving this problem, with neuroblastoma cells shown to become "addicted" to N-myc stabilisation by Aurora A through this protein-protein interaction.[2]

Whilst there is an abundance of possible approaches to identifying compounds capable of modulating the activity of Aurora A or disrupting the Aurora A/N-myc interaction, one challenge in this project has been to identify the most promising methods on which to focus. Computational methods proved helpful in narrowing down libraries of compounds to a manageable number of compounds for crystal soaking and single point *in vitro* Caliper mobility shift assays. Further crystal soaking of diverse sets of fragments from commercial and in-house libraries helped to provide some validation for the *in silico* screening process as well as highlighting other interesting compounds which soaked into sites of interest in this project and other Aurora A binding sites which weren't intended to be the focus of this project.

Another challenge in this project was to identify which of the possible target sites showed the most potential for development of small molecule inhibitors of Aurora A, possibly with the ability to also disrupt the Aurora A/N-myc protein-protein

interaction. A number of factors considered during the course of the project helped inform the decision to ultimately focus on the development of type III inhibitors of Aurora A, discussed below.

The initial foray into the design of novel compounds capable of disrupting the Aurora A/N-myc interaction involved computational analysis of potential target binding sites of Aurora A and *in silico* screening of commercial and in-house libraries. Binding site analysis of the five sites considered highlighted potential challenges associated with some approaches. SiteScores below the "druggable" level of 0.8 were generated for the type IV and N-myc binding sites considered in this project (Table 6.1). The highest SiteScore came from the ATP-binding site, which was already known to allow the binding of small molecules, with a number of type I inhibitors already reported.[45,116,117] The potential for the development of a type III inhibitor of Aurora A also showed some promise at this point, with SiteScores comfortably above the 0.8 threshold. A further encouraging SiteScore was obtained at this site following the soaking of the fragment **22** into Aurora A and the subsequent co-crystal structure produced. The SiteScore generated from using this fragment to define the binding site for analysis gave more confidence in the predicted druggability of this region in developing a type III inhibitor of Aurora A.

Table 6.1 SiteScores at each of the target sites considered

Site	SiteScore (with water)	SiteScore (without water)
N-myc (Trp77)	0.525	0.536
MES	0.828	0.798
TPX2	0.708	0.573
ATP	1.14	1.08
Salt bridge	1.12 (**22** bound)	1.06 (PDB 5L8L)

Compound library docking at each of the target sites highlighted further challenges at the N-myc binding site, with very few compounds predicted to bind in the desired region and low Glide scores predicting very weak binding to Aurora A. With these results, despite the attractiveness of directly blocking the binding of N-myc to Aurora A to disrupt the interaction, work on this approach did not progress further. The type IV approach to Aurora A inhibition also stalled at this time, with the focus shifting to type I and type III inhibition, the two areas with the most encouraging computational predictions.

This decision did not only rely on results from computational modelling. A number of type I Aurora A inhibitors had already been reported in the literature,[45, 50, 87,110] with alisertib also in clinical trials.[118] Efforts at combining two of these inhibitors (Figure 6.1), CD532 and VX-680, in the hope of finding a compound with the ability to disrupt the Aurora A/N-myc interaction (as with CD532) whilst avoiding the solubility challenges associated with CD532, ultimately failed to produce a hybrid compound which retained the inhibitory activity.
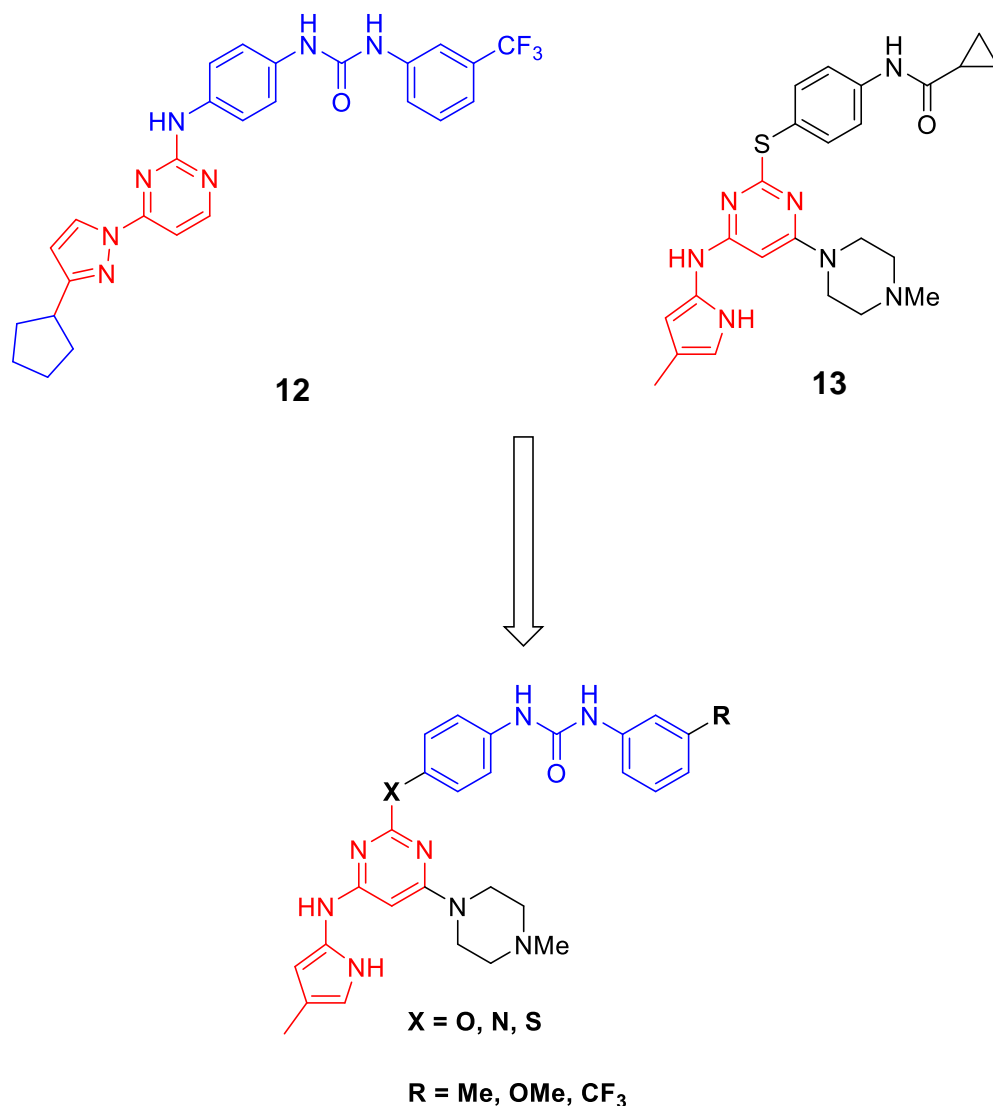
Figure 6.1 Hybrid compound design based on published inhibitors CD532 (**12**) and VX-680 (**13**)

The type III approach to Aurora A inhibition eventually proved to be one of the most promising branches of this project. Alignment of the most closely related human proteins to Aurora A from a BLAST search against the Aurora A amino acid sequence suggested more conservation of amino acid types around the ATP binding site. By contrast, the proposed type III binding site, around the salt bridge of Aurora A, showed more variability within the 10 Å radius investigated. Although there were still some highly conserved amino acids, particularly those closest to and including the salt bridge residues, the surrounding areas showed higher

variability in amino acid properties. This could indicate the potential to build selectivity for Aurora A over other kinases as any fragments are grown into more drug-like compounds in the future.

In silico screening of compound libraries enabled more focussed *in vitro* screening, which resulted in the identification of two fragments, one of which was successfully soaked into Aurora A/vNAR-D01 crystals and showed binding which closely matched the prediction generated through virtual high-throughput screening. This compound also showed activity in the Caliper mobility shift assay with an IC_{50} of 0.6 mM.

Although the second fragment, **57**, did not successfully soak into protein crystals, an IC_{50} of 0.6 mM was identified through Caliper and subsequent ATP competition experiments indicated the kinase inhibition was non-competitive with respect to ATP. Both fragments contained a carboxylic acid group which might interact with the positively charged magnesium ions at the ATP binding site.

The *in silico* screening results were revisited to focus on fragments containing carboxylic acids with docking poses similar to those of **22** and **57**. Unfortunately, crystal soaking experiments based on this selection of carboxylic acid-containing fragments failed to yield any further compounds shown to bind at the desired type III binding site.

Disappointingly, the extent of SAR investigated was restricted by Covid-19. Fragments based on the structures of **22** and **57** were purchased or synthesised with the intention of determining IC_{50} values using the Caliper mobility shift assay and crystal soaking to assess any binding modes. Three of these fragments were tested and IC_{50} values determined by another PhD student upon lab re-opening, the results of which indicated fragments **22c**, **22i** and **22k** might have a similar activity to **22**, although the large errors in this value means a clear conclusion cannot be drawn. ATP competition experiments were also not completed on these compounds.

6.2 Future Work

Previously described ATP-competitive inhibitors have been shown to potently inhibit the kinase activity of Aurora A.[50] Although the changes made have not been successful in improving the potency of this class of inhibitors, there may be useful modifications to be made to these structures to improve solubility whilst maintaining potency, or increasing the disruptive effect on the Aurora A/N-myc interaction. Virtual screening of compound libraries identified some compounds and fragments with structures unrelated to any of the previously published Aurora A inhibitors. These compounds could be candidates for further *in vitro* investigation, although the reliability of these computational predictions remains to be demonstrated for type I inhibitors of Aurora A.

Although the relatively low potency of the fragments identified in this project prevented investigation into their effect on the key protein-protein interaction, future work with more potent type III inhibitors of Aurora A would include an analysis of their effect on the stability of the Aurora A/N-myc interaction. Other work has used fluorescence polarisation to study the effect of an inhibitor on kinase activity or the disruption of protein-protein interactions, including the Aurora A/N-myc interaction[39] and may be used to analyse the effect of any promising compounds.[119] When measuring phosphorylation of a kinase substrate, Immobilised Metal Ion Affinity Particle (IMAP) technology allows a metal complex to form with phosphate groups on the labelled substrate. A potent kinase inhibitor would prevent phosphorylation and formation of this large complex, and a high polarisation value would be observed.[120] For protein-protein interactions, the effective disruption of the complex results in two smaller components, one of which will be fluorescently labelled and will give a lower polarisation than the complex.[119]

Alternatively, the protein-protein interaction could be analysed using AlphaScreen, which uses donor and acceptor beads attached to the two interacting proteins. A laser is used to excite the donor bead to induce the formation of singlet oxygen, which can travel 200 nm before decaying. If the acceptor bead is within 200 nm of the donor bead, energy transfer allows luminescence of the acceptor bead at a

wavelength of 615 nm. The donor and acceptor beads are attached to proteins *via* antibodies secured to the beads by streptavidin and biotin.[102]

Further SAR investigation around the two key fragments (**22** and **57**) identified as potential type III inhibitors of Aurora A would be a priority in continuing this work. With further SAR data from Caliper experiments and crystallography it may be possible to begin to grow these fragments into more drug-like molecules, potentially increasing potency by targeting key interactions with nearby amino acid side chains. This could be achieved using crystal structures with fragments bound near the Aurora A salt bridge. By referring to the kinase alignment performed, it may be possible to build some selectivity over other kinases by targeting Aurora A residues which are not conserved amongst closely related kinases.

Chapter 7 Experimental

7.1 Computational Methods

7.1.1 Screening Libraries

The Asinex library[99] contains 21872 fragments. The library aims for diversity as well as containing groups known to be good synthetic starting points.

The 2nd Generation Bionet library[98] contains 1166 fragments with 445 fragments found in currently used pharmaceuticals. The library excludes reactive or promiscuous compounds and anything known to aggregate.

The Enamine α-helix mimetic library[95] contains around 14000 compounds with most obeying Lipinski's Rule of 5 (molecular weight <500 Da, ≤5 H-bond donors, ≤10 H-bond acceptors, logD ≤5). The library was designed to contain a diverse range of 3D shapes with three groups similar to amino acid side chains projected from the core.

The Enamine fragment library contains 300 fragments under 300 Da with a focus on diversity. This library is available in-house.

The ChemDiv library contains just over 14000 peptidomimetic compounds aimed at targeting protein-protein interactions with α-helix or β-turn mimetics containing amino acid side chain-like groups projected.[100]

The Chembridge DIVERSet library contains around 100000 compounds selected for structure and pharmacophore diversity, filtered for desirable drug-like properties (adhering to the Lipinski's rule of 5).[97]

The Maybridge libraries together contain 1000 fragments with a focus on diversity. The library complies with the Astex Rule of 3 (molecular weight <300 Da, ≤3 H-bond donors, H-bond acceptors, ClogP ≤3). The Maybridge compounds are available in-house at the University of Leeds.

The MCCB library contains around 30000 larger compounds from the University of Leeds with a typical molecular weight range of 300-500 g mol^{-1}.

All of the libraries were prepared using Ligprep (Schrödinger) with default settings (pH 7.0 ± 2.0, specified chirality maintained)[121] apart from the number of stereoisomers generated, which was reduced from 32 to 2 prior to screening.

7.1.2 Docking

All docking was performed using Glide (Schrodinger)[122] and either the OPLS2005 or OPLS3 force fields, with results ranked by docking score. For each screen, default Glide SP docking settings were used with the addition of strain correction terms. Grid sizes were changed to accommodate each binding site but the maximum length of docked compounds was kept at 14 Å. At each site, libraries were docked with and without the crystallised water molecules.

The N-myc site used PDB 5G1X with N-myc removed. The coordinates of N-myc residues Met81 and Glu84 were used as centres for the Met81 and Glu84 Glide grids, respectively. In the crystal structure the N-myc side chain of Trp77 sits between the Aurora A residues Leu189 and Leu193, so these were used to generate the Trp77 Glide grid. Grid inner and outer box sizes and centre point coordinates (where applicable) for all sites are summarised in Table 7.1.

At the MES site, PDB 5LXM was used and MES was redocked to give RMSD values of 1.21 Å for SP and 1.31 Å for XP (heavy atoms, calculated in place). MES was identified as the ligand to generate the Glide grid. The grid size was based on docking ligands with a similar size to MES.

The TPX2 site used PDB 5G1X with N-myc removed and used the Aurora A residues His187 and His280 as the centre point for Glide grid generation.

ADP was redocked at the ATP binding site in each of the chosen crystal structures (PDB 5LXM and 5G1X). RMSD values for PDB 5LXM were 2.4505 Å (SP) and 1.7030 Å (XP). For PDB 5G1X the RMSD values were 0.41 Å (SP) and 0.21 Å (XP). Crystallised ADP coordinates were used to make the Glide grid.

CD532 was redocked into the ATP binding site in the crystal structure of Aurora A with CD532 bound (PDB 4J8M). With SP, the RMSD value was 0.53 Å. Crystallised CD532 was used as the centre point for the Glide grid.

The salt bridge site used the midpoint of Glu162 and Lys181 as the centre for grid generation.

Table 7.1 Dimensions and centre coordinates of Glide grids

Site	Inner box dimensions (Å)	Outer box dimensions (Å)	Grid centre coordinates
N-myc Trp77 (PDB 5G1X)	10, 10, 10	24, 24, 24	-
N-myc Met81 (PDB 5G1X)	10, 10, 10	25, 25, 25	39.808, -33.781, 46.472
N-myc Glu84 (PDB 5G1X)	10, 10, 10	23, 23, 23	42.560, -34.096, 52.154
N-myc Trp77 (PDB 5LXM)	10, 10, 10	24, 24, 24	-
MES	10, 10, 10	20.82, 20.82, 20.82	-
TPX2	10, 10, 10	30, 30, 30	-
ATP (PDB 5G1X)	10, 10, 10	25, 25, 25	44.272, -8.127, 41.224
ATP (PDB 5LXM)	10, 10, 10	30, 30, 30	45.455, -8.135, 14.624
CD532 (PDB 5J8M)	10, 10, 10	24, 24, 24	-
Salt bridge	10, 10, 10	30, 30, 30	-

A summary of libraries docked at each site is shown in Tables 7.2-4.

Table 7.2 The libraries docked at each site with water. Libraries marked (F) are fragment libraries. The Maybridge fragment library was screened using XP at the ATP site with PDB 5LXM due to a better RMSD for redocked ADP than SP. All other libraries were screened using SP.

Site	PDB Code		MCCB	Chembridge	Maybridge (F)	Asinex (F)	Bionet (F)
MES	5LXM		✓	✓	✓	✓	✓
TPX2	5G1X		✓	✓	✓	✓	✓
N-myc	5G1X		✓	✓	✓	✓	✓
	5LXM		–	–	–	–	✓
ATP	5G1X		✓	✓	✓	✓	✓
	5LXM		–	–	XP	–	–
Salt bridge	5L8L	No ADP	✓	✓	✓	✓	✓
	5L8L	With ADP	✓	✓	✓	✓	✓

Table 7.3 The libraries docked at each site without water

Site	PDB Code	MCCB	Chembridge	Maybridge (F)	Asinex (F)	Bionet (F)
MES	5LXM	✓	✓	✓	✓	✓
TPX2	5G1X	✓	✓	✓	✓	✓
N-myc	5G1X	✓	✓	✓	✓	✓
	5LXM	–	–	–	–	–
ATP	5G1X	✓	✓	✓	✓	✓

	5LXM		–	–	–	–	–
Salt bridge	5L8L	No ADP	✓	✓	✓	✓	✓
	5L8L	With ADP	✓	✓	✓	✓	✓

Table 7.4 The peptidomimetic libraries docked at each site

Site	PDB Code	Enamine	ChemDiv
Trp77 with water	5G1X	✓	✓
Trp77 without water	5G1X	✓	✓
Met81 with water	5G1X	✓	✓
Met81 without water	5G1X	✓	✓
Glu84 with water	5G1X	✓	✓
Glu84 without water	5G1X	✓	✓

7.1.3 SiteMap

A prediction of "druggability" for each of the proposed binding sites was acquired by SiteMap (Schrödinger).[104] Ligands were chosen either from co-crystallised compounds (MES and ADP) or a result from virtual screening which sat in the intended binding site. Table 9.4 shows the ligand and crystal structure used for each binding site.

Table 7.5 Crystal structures and ligands used at each site in SiteMap to generate SiteScores

SITE	PDB CODE	LIGAND USED	
		With water	Without water
N-MYC	5G1X	Maybridge library top result (Trp77 grid centre)	Enamine library 13th result (Met81 grid centre)
MES	5LXM	MES	MES
TPX2	5G1X	Bionet library top result	Bionet library top result
ATP	5G1X	ADP	ADP
SALT BRIDGE	Crystal structure with fragment **22** bound	Crystallised fragment	Crystallised fragment
	5L8L, without ADP	MCCB library top result	MCCB library top result

7.1.4 SPROUT

SPROUT fragment growth analysis used PDB 5LXM with MES defined as the ligand and the Aurora A/TPX2 structure as the receptor.

7.1.5 ROCS

ROCS version 3.2.1.4[123] was used with the Enamine α-helix mimetic compound library. Query molecules included both crystallised N-myc (residues 77-84) and the

three N-myc side chains Trp77, Met81 and Glu84 connected by a carbon chain linker to ensure the correct distance between side chain groups.

7.1.6 Kinase Alignment

Proteins with highest sequence similarity to Aurora A were found *via* a BLAST (Basic Local Alignment Search Tool) search using the catalytic domain of Aurora A (accession NP_001310234.1, residues 1-403) as the search query against all human protein entries on the NCBI (National Center for Biotechnology Information) database. Results were ranked by sequence similarity and E value and the highest ranked 107 discrete protein sequences were collected as FASTA (Fast Adaptive Shrinkage-Threshold Algorithm) format. Amino acid sequences were aligned in MEGA-X (Molecular Evolutionary Genetics Analysis) using ClustalW and MUSCLE (MUltiple Sequence Comparison by Log-Expectation) functions (?).

Amino acids within the vicinity of the Aurora A salt bridge and ATP binding sites were identified through a visual inspection of the surrounding amino acid side chains, the identification of side chains within 14 Å of each site and the generation of ligand interaction diagrams using Glide (Schrödinger).

7.2 Protein Expression and Purification

7.2.1 Physical Methods

Centrifugation used either a Beckman Avanti J-20XP or J-26XP instrument for larger volumes (40 mL and above). Concentration of purified protein used a Beckman Allegra X-12R centrifuge.

Immobilised metal ion affinity chromatography (IMAC) was performed using a HisTrap High Performance (HP) column with ÄKTA Prime. His-tagged proteins were eluted with HisB buffer (Table 7.6).

Gel filtration (GF) chromatography was performed using a HiLoad 16/600 Superdex 75 prep grade column with ÄKTA Prime and Aurora A GF buffer (Table 7.6).

SDS-PAGE was performed using Novex Tris-Glycine precast gels with Novex Tris-Glycine running buffer at 200 V and 200 mA. Novex SeeBlue Plus2 Pre-Stained protein standard was used to assess results.

7.2.2 General Procedure

All proteins were expressed in BL21 DE3 RIL *E. coli* cells grown in LB medium at 37ºC to OD (optical density) 0.6. BL21 DE3 RIL cells were chloramphenicol resistant and protein constructs contained kanamycin resistance genes. Colonies were therefore grown in the presence of kanamycin and chloramphenicol. Expression was induced by addition of isopropyl b-D-1-thiogalactopyranoside (IPTG, 1M, final concentration 0.6 mM). Following centrifugation (3500 rpm, 4ºC, 15 min) cell pellets were either purified immediately (section 10.2.3-4) or stored at -80ºC. Buffers specific to this project are described in Table 9.5.

Table 7.6 Buffer components for protein expression and purification

Buffer	Components
HisA	50 mM Tris pH 7.5, 0.3 M sodium chloride, 5 mM magnesium chloride, 10% glycerol
HisB	50 mM Tris pH 7.5, 0.3 M sodium chloride, 5 mM magnesium chloride, 0.5 M imidazole, 10% glycerol
Aurora A GF	20 mM Tris pH 7.0, 0.2 M sodium chloride, 5 mM magnesium chloride, 5 mM β-mercaptoethanol, 10% glycerol
TEV protease	50 mM Tris pH 7.5, 0.2 M sodium chloride, 5 mM magnesium chloride, 5 mM β-mercaptoethanol, 10% glycerol
TES	50 mM Tris pH 8.0, 0.2 mM ethylenediaminetetraacetic acid (EDTA), 0.2 M sucrose

7.2.3 Aurora A

Purification of both wild-type Aurora A and the mutant C290A:C393A followed the same protocol. Both variants contained a cleavable C-terminal His-tag. Pellets

were resuspended in HisA buffer with cOmplete™, EDTA-free protease inhibitor cocktail (Roche). Cells were sonicated (6 cycles of 15 s on, 30 s off at 55% amplitude), centrifuged (19000 rpm, 4ºC, 45 min) and filtered (0.45 μm pore). Filtered lysate was purified by IMAC and presence of the protein was confirmed by SDS-PAGE. Relevant fractions were dialysed using SnakeSkin Dialysis Tubing (MWCO 3500 Da) with TEV protease buffer and the His-tag cleaved by TEV protease (1 mg mL^{-1}, 450 μL aliquots). Protein was again purified by IMAC and confirmed by SDS-PAGE. Relevant fractions were concentrated (Vivaspin Turbo 15 (Sartorius) MWCO 5000 Da, 3500 rpm, 4ºC) and any aggregated protein removed by GF. Protein presence was confirmed by SDS-PAGE; key fractions were concentrated to 103 μM (wild-type) and 800 μM (C290A:C393A) and stored at -80ºC.

7.2.4 vNAR-D01

The vNAR-D01 construct contained a non-cleavable C-terminal His-tag for purification and expression was targeted to the periplasm. Cell pellets were resuspended in TES buffer and centrifuged (10000 g, 4ºC, 30 min). The supernatant was filtered and dialysed using SnakeSkin Dialysis Tubing (MWCO 3500 Da) with HisA buffer. Initial purification was performed by IMAC and presence of protein was confirmed by SDS-PAGE. Relevant fractions were concentrated (Vivaspin Turbo 15 (Sartorius) MWCO 3000 Da, 3500 rpm, 4ºC). Final purification was performed by GF and confirmed by SDS-PAGE. Key fractions were concentrated to 43 μM and stored at -80ºC.

7.3 Crystallography

7.3.1 Wild-Type Aurora A/vNAR-D01 Crystallisation

Purified vNAR-D01 and wild-type Aurora A proteins were combined with an excess of vNAR-D01 and the resulting complex was isolated by GF with Aurora A GF buffer. Crystals were obtained with 0.1 M citrate, pH 5.0 with 20% PEG6000 as precipitant. X-Ray diffraction data was collected at Diamond Light Source.

7.3.2 Fragment Soaking

Fragments stored at a stock concentration of either 50 or 100 mM in DMSO were soaked into Aurora A/vNAR-D01 crystals to obtain a final DMSO concentration of 5% where possible. In some cases, higher DMSO concentrations were used in an effort to detect any weak binding fragments by soaking at a higher fragment concentration. Most fragment soaking was performed by Mohd Syed Ahanger.

7.4 Assays

7.4.1 Caliper

7.4.1.1 Single Point Assay

Caliper mobility shift assays were performed using the LabChip EZ Reader. Fragments were tested at 5 mM concentration and a maximum DMSO concentration of 5% where possible. Reaction mixtures also contained 30 nM cysteine-less Aurora A mutant (C290A:C393A), 80 μM ATP and 1.5 μM fluorescent substrate (FL-peptide 21) in Aurora A reaction buffer (50 mM Tris pH 7.5, 200 mM NaCl, 5 mM MgCl$_2$, 10% glycerol and 10 mM β-mercaptoethanol). Reactions were stopped after 30 minutes with the addition of Caliper Separation Buffer (containing EDTA) before analysis of substrate phosphorylation using the LabChip EZ Reader. Percentage conversion was normalised to uninhibited Aurora A. Each reaction was performed at least in duplicate and measurements of conversion for each reaction were also taken in duplicate.

7.4.1.2 IC$_{50}$ Determination

For IC$_{50}$ determination, compound stocks were diluted with 10 3-fold serial dilutions in DMSO to obtain a final maximum concentration of 5 mM, with 5% DMSO in the final reaction mixture. As above, reaction mixtures also contained 1.5 μM fluorescent substrate (FL-peptide 21), 80 μM ATP and 50 nM Aurora A (C290A:C393A) in Aurora A reaction buffer. In some cases, compounds were insoluble at the highest concentrations, in which case this top concentration was discarded and a further dilution made to retain 10 different inhibitor concentrations.

Controls included one sample containing all reaction mixture components without any inhibitor, and one sample containing all reaction mixture components save Aurora A kinase. Each reaction was again performed at least in duplicate.

Substrate conversion was monitored using the LabChip EZ Reader with 40 cycles of measurements taken over ca. 40 minutes. Blockages in some channels occasionally caused the delay of fluorescence detection and the misidentification of fluorescence peaks by the automated system. This necessitated the manual adjustment of this identification to generate useful data.

7.4.1.3 ATP K_m

ATP K_m experiments were performed using 12 3-fold serial dilutions of ATP with a maximum assay concentration of 3 mM. Reaction mixtures also contained 50 nM Aurora A (C290A:C393A), 1.5 μM fluorescent substrate (FL-peptide 21) and inhibitor at either the IC_{50}, half the IC_{50} or double the IC_{50} concentration, where inhibitor was included, in Aurora A reaction buffer.

Substrate conversion was monitored in the same way as for IC_{50} determination (see 7.4.1.2). The initial, linear portion of the reaction was used to determine ATP K_m using Michaelis Menten kinetics in GraphPad Prism versions 5.0-9.0.[124]

7.5 Synthesis

Some synthesis and analysis was affected by Covid-19 and the resulting lab shutdown.

7.5.1 Materials

All purchased chemicals and solvents were used without further purification unless otherwise stated.

7.5.2 Physical Methods

^{1}H Nuclear Magnetic Resonance spectra were recorded at 298 K on either a Bruker AV 500 or a Bruker AV-NEO spectrometer at 500 MHz (^{1}H) and 125 MHz (^{13}C), or a Bruker AV3HD-400 instrument at 400 MHz (^{1}H) and 100 MHz (^{13}C). NMR data was analysed and coupling constants calculated using TopSpin version

3.0 or 4.0 (Bruker). Coupling constants (J) are reported to the nearest 0.1 Hz for ^{1}H and to the nearest 1 Hz for ^{13}C, with ^{1}H chemical shift reported to the nearest 0.01 ppm and ^{13}C chemical shift to the nearest 0.1 ppm. NMR splitting assigned as: singlet (s), doublet (d), triplet (t), apparent (app), multiplet (m).

LCMS was performed on a Thermo Scientific Dionex UltiMate with a gradient of MeCN (5-95%) in water, each containing 0.1% formic acid, with a flow rate of 1 mL min[-1] on a C_{18} reverse phase column. Compounds detection used a diode array detector and a Bruker amaZon speed mass spectrum analyser.

Infrared spectroscopy was performed using a Bruker Alpha Platinum ATR FTIR instrument. Biotage silica column chromatography used an Isolera Four EXP with Spektra.

7.5.3 Type I Inhibitors

General Procedure 1

Sodium hydride (60% dispersion in mineral oil) was added to a solution of the relevant 4-nitrophenyl compound in anhydrous tetrahydrofuran (0.1-0.2 M) at 0 °C and stirred for 30 min. This mixture was then added to a solution of **17** in tetrahydrofuran and stirred at 0 °C. The reaction mixture was diluted with ethyl acetate, washed with saturated sodium bicarbonate and extracted with ethyl acetate (3 × 100 mL). The combined organic layers were dried over sodium sulfate, filtered and the solvent removed under reduced pressure. The resulting crude material was purified by recrystallisation from methanol.

General Procedure 2

3-Amino-5-methylpyrazole was stirred with diisopropylethylamine in anhydrous dimethylformamide at 85 °C for 10 min. Compound **18** was added and the reaction stirred at room temperature. The mixture was diluted with ethyl acetate, washed with sodium bicarbonate and extracted with ethyl acetate. The combined organic layers were dried over sodium sulfate, filtered and the solvent removed under reduced pressure. The intermediate product was dissolved in 1-methylpiperazine and stirred at 110 °C. The reaction mixture was diluted with ethyl acetate, washed

with saturated sodium bicarbonate, and extracted with ethyl acetate (3 × 50 mL). The combined organic layers were dried over sodium sulfate, filtered and the solvent removed under reduced pressure.

General Procedure 3

Di-*tert*-butyl-dicarbonate was added to a solution of **19** in methanol with triethylamine and stirred at room temperature. Dichloromethane and saturated sodium bicarbonate were added to the reaction mixture before extracting with dichloromethane (3 × 50 mL). The combined organic layers were dried over sodium sulfate, filtered and the solvent removed under reduced pressure.

4,6-dichloro-2-(methylsulfonyl)pyrimidine (17)
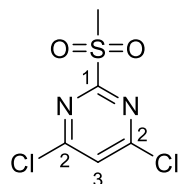
4,6-Dichloro-2-(methylthio)pyrimidine (546 mg, 2.80 mmol) was dissolved in dichloromethane (10 mL) at 0 °C with addition of *m*-chloroperoxybenzoic acid (1.82 g, 10.5 mmol) over 20 min and stirred at room temperature for 12 h. The reaction was diluted with ethyl acetate (20 mL), washed with a mixture of aqueous sodium thiosulfate/sodium bicarbonate (2:1, 30 mL) and extracted with ethyl acetate (3 × 30 mL). The combined organic layers were dried over sodium sulfate, filtered and the solvent removed under reduced pressure to give the crude product, a pale yellow solid, which was purified using flash column chromatography on silica gel (50-70% ethyl acetate/petroleum ether) to produce the *title compound* **(17)** as a colourless powder (496 mg, 2.18 mmol, 78%). ^{1}H NMR δ_H/ppm (400 MHz, DMSO-d_6) 8.44 (1H, s, 3-H), 3.45 (3H, s, methyl C-H$_3$); ^{13}C NMR δ_C/ppm (100 MHz, DMSO-d_6) 165.3 (C2), 163.1 (C1), 125.8 (C3), 39.7 (SO$_2$Me)

This compound was synthesised *via* an unmodified literature procedure.[109] All analytical data agreed with the literature data.

4,6-dichloro-2-(4-nitrophenoxy)pyrimidine (18a)
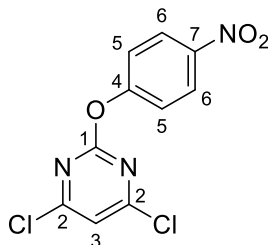
Following General Procedure 1, sodium hydride (120 mg, 5.20 mmol) was added to a solution of 4-nitrophenol (412 mg, 2.96 mmol, 0.20 M) in anhydrous tetrahydrofuran (15 mL) at 0 °C and stirred for 30 min. This mixture was then added to a solution of **17** (480 mg, 2.11 mmol, 0.14 M) in tetrahydrofuran (15 mL) at 0 °C and stirred for 4 h. The reaction mixture was diluted with ethyl acetate (30 mL), washed with saturated sodium bicarbonate (50 mL) and extracted with ethyl acetate (3 × 100 mL). The combined organic layers were dried over sodium sulfate, filtered and the solvent removed under reduced pressure. The resulting crude material was purified by recrystallisation from methanol to yield the *title compound* as colourless crystals (129 mg, 0.45 mmol, 21%). ^{1}H NMR δ_H/ppm (400 MHz, DMSO-d_6) 8.37 (2H, d, J = 7.2 Hz, 6-H), 7.88 (1H, s, 3-H), 7.62 (2H, d, J = 6.8 Hz, 5-H); ^{13}C NMR δ_C/ppm (100 MHz, DMSO) 162.8 (C1), 162.7 (C2), 156.7 (C4), 145.2 (C7), 125.8 (C6), 122.8 (C5), 117.2 (C3)

This compound was synthesised *via* an unmodified literature procedure.[109] All analytical data agreed with the literature data.

4, 6-dichloro-2-(4-nitrothiophenoxy)pyrimidine (18b)
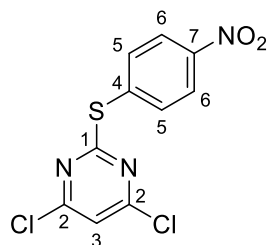
Following General Procedure 1, sodium hydride (150 mg, 6.60 mmol) was added to a solution of 4-nitrothiophenol (430 mg, 2.77 mmol, 0.18 M) in anhydrous tetrahydrofuran (15 mL). This was added to a solution of **17** (459 mg, 2.42 mmol, 0.16 M) in tetrahydrofuran (15 mL) and stirred for 17 h. The reaction mixture was

diluted with ethyl acetate (30 mL), washed with saturated sodium bicarbonate (50 mL) and extracted with ethyl acetate (3 × 100 mL). The combined organic layers were dried over sodium sulfate, filtered and the solvent removed under reduced pressure. The crude material was taken forward to the next step of the reaction. LCMS *m/z* [M]⁻ found 299.99, requires 299.95

4, 6-dichloro-2-(4-nitroaniline)pyrimidine (18c)
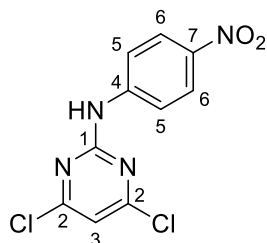
4-Nitroaniline (175 mg, 1.27 mmol) was dissolved under nitrogen in anhydrous tetrahydrofuran (10 mL) and cooled to -70 ºC. Sodium bis(trimethylsilyl)amide (1 M solution in tetrahydrofuran, 1.8 mL, 1.80 mmol) was added, followed by a solution of **17** (205 mg, 0.90 mmol, 0.09 M) in anhydrous tetrahydrofuran (10 mL). The resulting mixture was stirred for 16 h, followed by the addition of glacial acetic acid (10 mL). The crude product was extracted with brine (20 mL) and a mixture of ethyl acetate/hexane (1:1, 30 mL). The organic layers were dried over sodium sulfate, filtered and the solvent removed under reduced pressure. The crude material was taken forward to the next step of the reaction. LCMS *m/z* [M+H]⁺ found 284.99, requires 284.99

N-(5-methyl-1H-pyrazol-3-yl)-6-(4-methylpiperazin-1-yl)-2-(4-nitrophenoxy)pyrimidin-4-amine (19a)
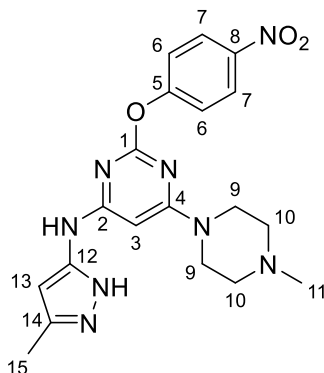
3-Amino-5-methylpyrazole (83 mg, 0.85 mmol) was stirred with diisopropylethylamine (0.1 mL, 0.5 mmol) in anhydrous dimethylformamide (5 mL) at 85 ºC for 10 min. Compound **18a** (104 mg, 0.36 mmol) was added and the reaction stirred at room temperature for 16 h. The mixture was diluted with ethyl acetate (10 mL), washed with saturated sodium bicarbonate (20 mL) and extracted with ethyl acetate (3 × 50 mL). The combined organic layers were dried over sodium sulfate, filtered and the solvent removed under reduced pressure. The intermediate product was dissolved in 1-methylpiperazine (5 mL) and stirred at 110 ºC for 30 min. The reaction mixture was diluted with ethyl acetate (20 mL), washed with saturated sodium bicarbonate (30 mL) and extracted with ethyl acetate (3 × 50 mL). The combined organic layers were dried over sodium sulfate, filtered and the solvent removed under reduced pressure. The crude material was purified by flash column chromatography on silica gel (0-20% methanol/dichloromethane) to produce the *title compound* **(19a)** as a pale yellow solid (125 mg, 0.30 mmol, 44%). ^{1}H NMR δ_H/ppm (400 MHz, DMSO-d_6) 11.84 (1H, s, NH), 9.34 (1H, s, pyrazole NH), 8.30 (2H, d, J = 9.0 Hz, 7-H), 7.43 (2H, d, J = 9.0 Hz, 6-H), 6.36 (1H, s, 3-H), 5.64 (1H, s, 13-H), 3.44 (s, 9-H$_2$), 2.39 (4H, s, 10-H$_2$) 2.23 (3H, s, 15-H$_3$), 2.07 (3H, s, 11-H$_3$); HRMS *m/z* (M+H)$^+$ found 411.1884, requires 411.1888

This compound was synthesised via an unmodified literature procedure.[109] All analytical data agreed with the literature data.

N-(5-methyl-1H-pyrazol-3-yl)-6-(4-methylpiperazin-1-yl)-2-(4-nitrothiophenoxy)pyrimidin-4-amine (19b)
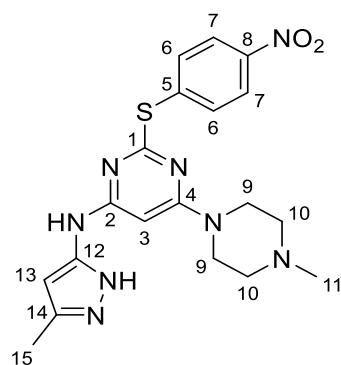
3-Amino-5-methylpyrazole (100 mg, 1.03 mmol) was stirred with diisopropylethylamine (0.25 mL, 1.25 mmol). Compound **18b** (273 mg, 0.90 mmol) was added and the reaction stirred at 85 ⁰C for 48 h. The intermediate product was dissolved in 1-methylpiperazine (5 mL) and stirred at 110 ⁰C for 30 min. The crude material was taken forward to the next step of the reaction. LCMS *m/z* [M]⁻ found 425.53, requires 425.50

N-(5-methyl-1H-pyrazol-3-yl)-6-(4-methylpiperazin-1-yl)-2-(4-nitroaniline)pyrimidin-4-amine (19c)
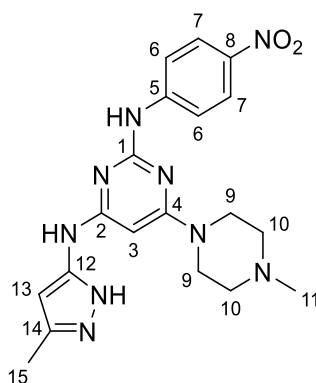
3-Amino-5-methylpyrazole (37 mg, 0.35 mmol) was stirred with diisopropylethylamine (0.6 mL, 0.35 mmol). Compound **18c** (84 mg, 0.29 mmol) was added and the reaction stirred at 85 ⁰C for 48 h. The intermediate product was dissolved in 1-methylpiperazine (5 ml) and stirred at 110 ⁰C for 16 h. The crude material was taken forward to the next step of the reaction. LCMS *m/z* [M+H]⁺ found 410.20, requires 410.20

***Tert*-butyl 5-methyl-3-(6-(4-methylpiperazin-1-yl)-2-(4-nitrophenoxy)pyrimidin-4-yl)amino)-1H-pyrazole-1-carboxylate (20a)**
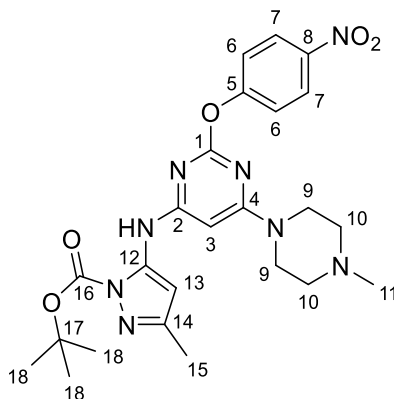
Di-*tert*-butyl-dicarbonate (206 mg, 0.94 mmol) was added to a solution of **3** (81 mg, 0.20 mmol, 0.02 M) in methanol (10 mL) with triethylamine (0.15 mL, 1.1 mmol) and stirred at room temperature overnight. Dichloromethane (20 mL) and saturated sodium bicarbonate (50 mL) were added to the reaction mixture before extracting with dichloromethane (3 × 50 mL). The combined organic layers were dried over sodium sulfate, filtered and the solvent removed under reduced pressure. The crude material was purified by flash column chromatography on silica gel (0-20% methanol/dichloromethane) to produce the *title compound* **(20a)** as a yellow oil (16 mg, 0.03 mmol, 16%). ^{1}H NMR δ_H/ppm (400 MHz, DMSO-d_6) 8.19 (2H, d, J = 9.1 Hz, 7-H), 7.49 (1H, s, NH), 7.26 (2H, d, J = 9.1 Hz, 6-H), 6.53 (1H, s, 3-H) 5.82 (1H, s, 13-H), 3.53 (4H, t, J = 4.8 Hz, 9-H$_2$), 2.36 (4H, t, J = 4.8 Hz, 10-H$_2$), 2.34 (3H, s, 15-H$_3$), 2.24 (3H, s, 11-H$_3$), 1.54 (9H, s, *boc* CH$_3$); ^{13}C NMR δ_C/ppm (100 MHz, DMSO-d_6) 164.9 (C2), 163.5 (C4), 160.5 (C1), 158.6 (C5), 150.3 (C12), 148.6 (C16), 144.5 (C8), 144.3 (C14), 125.1 (C7), 122.5 (C6), 102.2 (C13), 84.4 (C3), 81.7 (C17), 54.6 (C10), 45.2 (C11), 44.1 (C9), 28.0 (C18), 14.5 (C15)

This compound was synthesised via an unmodified literature procedure.[109] All analytical data agreed with the literature data.

Tert-butyl 5-methyl-3-(6-(4-methylpiperazin-1-yl)-2-(4-nitrothiophenoxy)pyrimidin-4-yl)amino)-1H-pyrazole-1-carboxylate (20b)
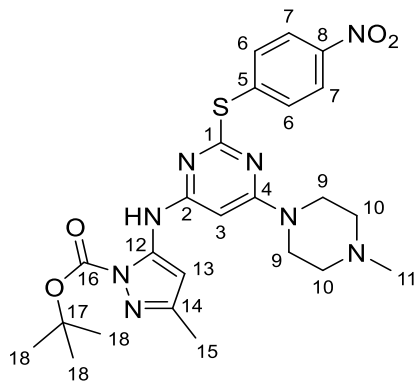
To a solution of **19b** (270 mg, 0.63 mmol, 0.06 M) in 1,4-dioxane (10 mL) were added triethylamine (0.4 mL, 2.8 mmol), di-*tert*-butyl-dicarbonate (610 mg, 2.8 mmol) and 4-dimethylaminopyridine (8 mg, 0.07 mmol). The reaction mixture was heated at reflux temperature for 16 h. Dichloromethane (20 mL) was added to the cooled mixture, which was washed with saturated sodium bicarbonate (50 mL) and extracted with dichloromethane (3 × 50 mL). The organic layers were dried over sodium sulfate, filtered and the solvent removed under reduced pressure. The crude material was purified by flash column chromatography on silica gel (0-5% methanol/dichloromethane) to produce the *title compound* **(20b)** as a yellow oil (215 mg, 0.41 mmol, 65%). This compound was taken forward without full purification. LCMS *m/z* [M]⁻ found 525.20, requires 525.21

This compound was synthesised *via* a literature procedure.[109] Analytical data agreed with the literature data.

***Tert*-butyl 5-methyl-3-(6-(4-methylpiperazin-1-yl)-2-(4-nitroaniline)pyrimidin-4-yl)amino)-1H-pyrazole-1-carboxylate (20c)**
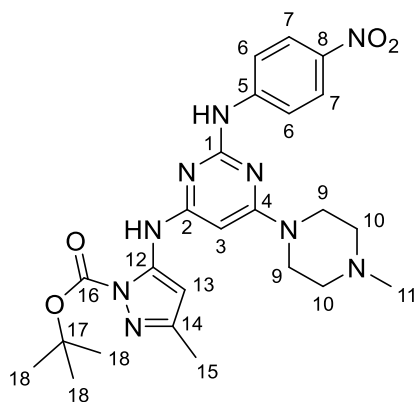
To a solution of **19c** (156 mg, 0.38 mmol, 0.04 M) in 1,4-dioxane (10 mL) were added triethylamine (0.4 mL, 2.8 mmol), di-*tert*-butyl-dicarbonate (610 mg, 2.8 mmol) and 4-dimethylaminopyridine (8 mg, 0.07 mmol). The reaction mixture was heated at reflux temperature for 16 h. Dichloromethane (20 mL) was added to the cooled mixture, which was washed with saturated sodium bicarbonate (50 mL) and extracted with dichloromethane (3 × 50 mL). The organic layers were dried over sodium sulfate, filtered and the solvent removed under reduced pressure. The crude material was purified by flash column chromatography on silica gel (0-5% methanol/dichloromethane) to produce the *title compound* **(20c)** as a yellow oil (42 mg, 0.08 mmol, 21%). This compound was taken forward without full purification. LCMS *m/z* [M+H]$^+$ found 510.12, requires 510.25

***Tert*-butyl 5-methyl-3-(6-(4-methylpiperazin-1-yl)-2-(4-aminophenoxy)pyrimidin-4-yl)amino)-1H-pyrazole-1-carboxylate (21a)**
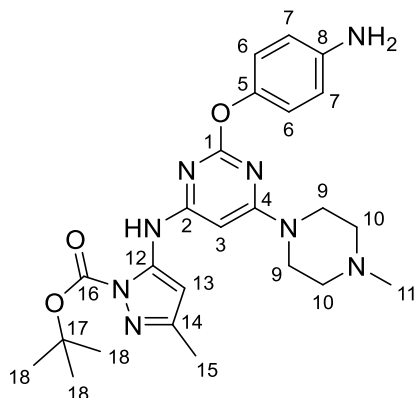
Ammonium formate (16 mg, 0.2 mmol) and **20a** (33 mg, 0.06 mmol) were dissolved in ethanol (5 mL) and treated with 10% palladium on carbon (80 mg, 0.03 mmol). The reaction mixture was stirred at reflux temperature for 2 h. The resulting mixture was filtered over Celite and the solvent removed under reduced pressure. The crude material was dissolved in dichloromethane (10 mL) and washed with saturated sodium bicarbonate (20 mL) and water (20 mL). The organic layers were dried over sodium sulfate, filtered and the solvent removed under reduced pressure to give the *title compound* **(21b)** as a dark yellow solid (26 mg, 0.05 mmol, 83%) which was taken forward without full purification. LCMS *m/z* [M]⁻ found 479.25, requires 479.26

***Tert*-butyl 5-methyl-3-(6-(4-methylpiperazin-1-yl)-2-(4-aminothiophenoxy)pyrimidin-4-yl)amino)-1H-pyrazole-1-carboxylate (21b)**
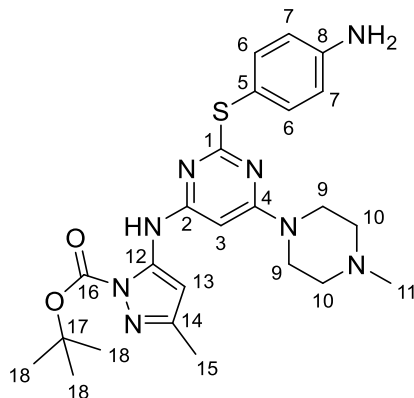
Ammonium formate (150 mg, 2.3 mmol) and **20b** (203 mg, 0.38 mmol) were dissolved in ethanol (10 mL) and treated with 10% palladium on carbon (100 mg, 0.04 mmol). The reaction mixture was stirred at reflux temperature for 2 h. The resulting mixture was filtered over Celite and the solvent removed under reduced pressure. The crude material was dissolved in dichloromethane (10 mL) and washed with saturated sodium bicarbonate (30 mL) and water (30 mL). The organic layers were dried over sodium sulfate, filtered and the solvent removed under reduced pressure to give the *title compound* **(21b)** as a dark yellow solid (148 mg, 0.30 mmol, 79%) which was taken forward without full purification. LCMS *m/z* [M+H]$^+$ found 497.24, requires 497.24

***Tert*-butyl 5-methyl-3-(6-(4-methylpiperazin-1-yl)-2-(4-nitroaniline)pyrimidin-4-yl)amino)-1H-pyrazole-1-carboxylate (21c)**
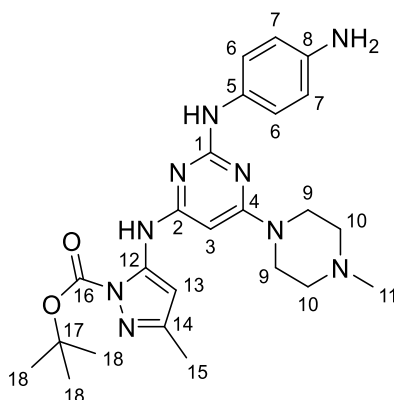
Ammonium formate (38 mg, 0.60 mmol) and **20c** (35 mg, 0.07 mmol) were dissolved in ethanol (10 mL) and treated with 10% palladium on carbon (30 mg, 0.01 mmol). The reaction mixture was stirred at reflux temperature for 2 h. The resulting mixture was filtered over Celite and the solvent removed under reduced pressure. The crude material was dissolved in dichloromethane (10 mL) and washed with saturated sodium bicarbonate (30 mL) and water (30 mL). The organic layers were dried over sodium sulfate, filtered and the solvent removed under reduced pressure to give the crude compound, which was taken forward without full purification. LCMS *m/z* [M+H]$^+$ found 481.14, requires 481.28

1-(4-((4-((5-methyl-1H-pyrazol-3-yl)amino)-4-(4-methylpiperazin-1-yl)pyrimidin-2-yl)phenoxy)-3-(trifluoromethyl)phenyl)urea (15a)
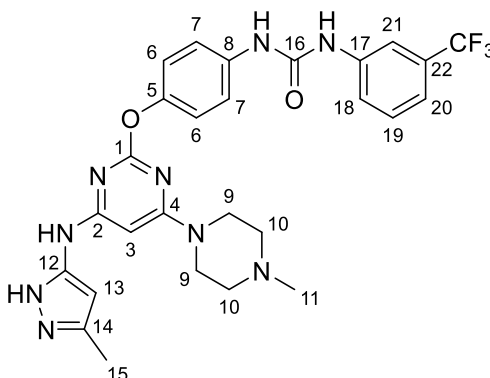
21a (22 mg, 0.04 mmol) and 3-trifluoromethylphenyl isocyanate (0.10 mL, 0.5 mmol) were stirred in dichloromethane (6 mL) at room temperature for 20 h. The resulting mixture was washed with saturated sodium bicarbonate (25 mL) and extracted with dichloromethane (3 × 25 mL). The organic layers were dried over sodium sulfate, filtered and the solvent removed under reduced pressure. The crude material was purified by flash column chromatography (0-5% methanol/dichloromethane) to produce the Boc-protected version of the *title compound* **(15a)** as a yellow oil. This intermediate (11 mg, 0.036 mmol) was dissolved in dichloromethane (2 mL), to which was added hydrochloric acid (2M in diethyl ether, 2 mL) and the mixture stirred at room temperature for 2 h. The solvent was removed under reduced pressure and the crude material was purified by flash column chromatography on silica gel (0-5% methanol/dichloromethane) to produce the *title compound* **(15a)** as the hydrochloride salt, a colourless solid (7 mg, 0.01 mmol, 31%). ^{1}H NMR δ_H/ppm (400 MHz, CD$_3$OD) 8.12 (1H, s, NH), 7.92 (1H, s, 21-H), 7.64-7.71 (3H, m, 18-H, 19-H, 20-H), 7.08-7.19 (4H, m, 6-H, 7-H), 6.81 (1H, s, 13-H), 6.56 (1H, s, 3-H), 3.32-3.50 (8H, m, 9-H, 10-H), 2.31 (3H, s, 15-H), 2.21 (3H, s, 11-H); LCMS *m/z* (M+H)$^+$ found 568.59, requires 568.58

This compound was synthesised *via* a literature procedure.[109] Analytical data agreed with the literature data.[31]

1-(4-((4-((5-methyl-1H-pyrazol-3-yl)amino)-4-(4-methylpiperazin-1-yl)pyrimidin-2-yl)thiophenyl)-3-(methyl)phenyl)urea (15b)
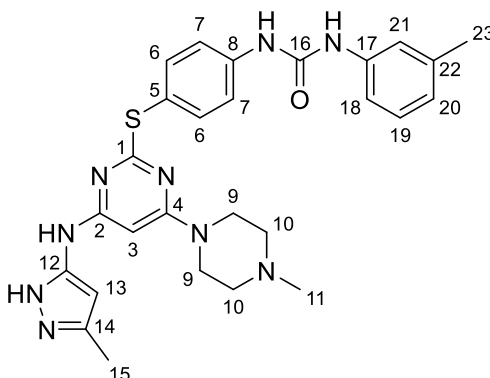
21b (33 mg, 0.066 mmol) and *m*-tolyl isocyanate (0.15 mL, 2.1 mmol) were stirred in dichloromethane (10 mL) at room temperature for 24 h. The resulting mixture was washed with saturated sodium bicarbonate (25 mL) and extracted with dichloromethane (3 × 25 mL). The organic layers were dried over sodium sulfate, filtered and the solvent removed under reduced pressure. The crude material was purified by flash column chromatography on silica gel (0-5% methanol/dichloromethane) to produce the Boc-protected version of the *title compound* **(15b)** as a yellow oil. This intermediate (23 mg, 0.036 mmol) was dissolved in dichloromethane (2 mL), to which was added hydrochloric acid (2M in diethyl ether, 2 mL) and the mixture stirred at room temperature for 2 h. The solvent was removed under reduced pressure and the crude material was purified by flash column chromatography (0-5% methanol/dichloromethane) to produce the *title compound* **(15b)** as the hydrochloride salt, a yellow solid (10 mg, 0.018 mmol, 33%). ^{1}H NMR δ_H/ppm (400 MHz, CDCl$_3$) 8.03-8.18 (1H, m, 21-H), 7.43-7.84 (3H, m, 18-H, 19-H, 20-H), 7.08 (1H, s, NH), 6.90-7.08 (4H, m, 6-H, 7-H), 6.77 (1H, s, 13-H), 6.35 (1H, s, 3-H), 2.60-2.90 (11H, m, 9-H, 10-H, 11-H), 2.28 (3H, s, 15-H); LCMS *m/z* [M+H]$^+$ found 530.24, requires 529.23

This compound was synthesised *via* a literature procedure.[109] Analytical data agreed with the literature data.[31]

1-(4-((4-((5-methyl-1H-pyrazol-3-yl)amino)-4-(4-methylpiperazin-1-yl)pyrimidin-2-yl)thiophenyl)-3-(methoxy)phenyl)urea (15c)
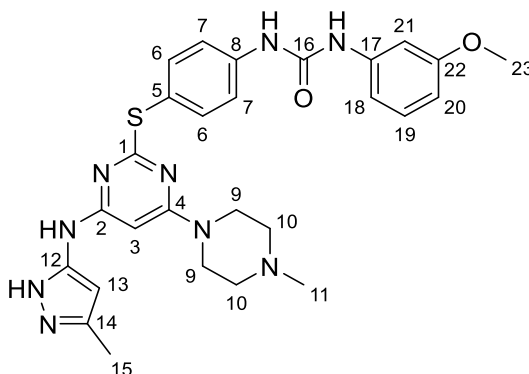
Compound **21b** (29 mg, 0.058 mmol) and 3-methoxy phenyl isocyanate (0.10 mL, 1.1 mmol) were stirred in dichloromethane (10 mL) at room temperature for 18 h. The resulting precipitate was filtered and washed with diethyl ether (10 mL). The crude material was purified by flash column chromatography on silica gel (0-5% methanol/dichloromethane) to produce the Boc-protected version of the *title compound* as a yellow oil. This intermediate (17 mg, 0.026 mmol) was dissolved in dichloromethane (2 mL), to which was added hydrochloric acid (2M in diethyl ether, 2 mL) and the mixture stirred at room temperature for 2 h. The solvent was removed under reduced pressure and the crude material was purified by flash column chromatography on silica gel (0-5% methanol/dichloromethane) to produce the *title compound* **(15c)** as the hydrochloride salt, a yellow solid (8 mg, 0.013 mmol, 36%). ^{1}H NMR δ_H/ppm (400 MHz, CDCl$_3$) 8.03-8.08 (1H, m, 21-H), 7.58-7.72 (3H, m, 18-H, 19-H, 20-H), 7.25 (1H, s, NH), 6.80-6.94 (4H, m, 6-H, 7-H), 6.73 (1H, s, 13-H), 6.56 (1H, s, 3-H), 3.37-3.73 (8H, m, 9-H, 10-H), 2.35 (3H, s, 15-H), 2.25 (3H, s, 11-H); LCMS *m/z* [M+H]$^+$ found 546.24, requires 546.23

This compound was synthesised *via* a literature procedure.[109] Analytical data agreed with the literature data.[31]

7.5.4 Type III Inhibitors

Methyl 9H-xanthene-9-carboxylate (149)
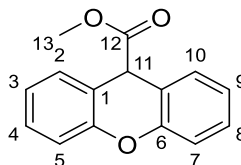
Xanthene-9-carboxylic acid (162 mg, 0.72 mmol) was dissolved in 1,4-dioxane (20 mL), followed by the addition of methanol (0.1 mL, 2.5 mmol) and sulfuric acid (0.1 mL). The reaction mixture was stirred at reflux for 24 h. The reaction mixture was poured into ice water (100 mL) before extracting with diethyl ether (3 × 50 mL) and washing with water (2 × 50 mL), ice-cold saturated sodium bicarbonate (3 × 20 mL) and saturated sodium chloride (1 × 50 mL). The combined organic layers were dried over sodium sulfate, filtered and the solvent removed under reduced pressure to yield the *title compound* as an off-white solid (115 mg, 0.48 mmol, 67%). ^{1}H NMR δ_H/ppm (400 MHz, CD$_3$OD) 7.30-7.36 (4H, m, 2, 5, 7, 10-H), 7.09-7.14 (4H, m, 3, 4, 8, 9-H), 5.10 (1H, s, 11-H), 3.64 (3H, s, 13-H)

All analytical data agreed with the literature data.[125]

Ethyl 1-oxo-1, 2, 3, 4-tetra-hydronaphthalene-2-carboxylate (151)
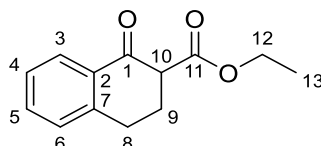
α-Tetralone (0.18 mL, 1.3 mmol), diethyl carbonate (2 mL, 1.6 mmol), sodium hydride (60% dispersion in oil, 110 mg) and anhydrous methanol (0.1 mL) were stirred together under nitrogen at 80 ºC for 2 h. The mixture was cooled before adding hydrochloric acid (5 mL, 1 M) and then extracted with ethyl acetate (3 × 30 mL). The organic layers were dried over sodium sulfate and the solvent removed under reduced pressure to produce a yellow oil (320 mg), which was carried forward without further purification.

2-oxo-1, 2, 3, 4-tetrahydronaphthalene-2-carboxylic acid (152)
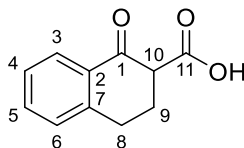
A solution of **151** (58 mg) in tetrahydrofuran (5 mL) was combined with a solution of lithium hydroxide (62 mg, 2.6 mmol, 0.52 M) in water (5 mL) at 4 °C. Hydrochloric acid (3 mL, 1 M) was added and the mixture stirred for 1 h to form a yellow precipitate.

Ethyl 2-oxo-1, 2, 3, 4-tetrahydronaphthalene-1-carboxylate (154)
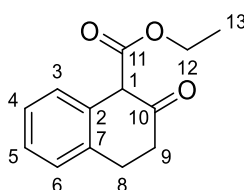
β-Tetralone (0.18 mL, 1.3 mmol) diethyl carbonate (2 mL, 1.6 mmol), sodium hydride (60% dispersion in oil, 113 mg) and anhydrous methanol (0.1 mL) were stirred together under nitrogen at 80 °C for 2.5 h. The mixture was cooled before adding hydrochloric acid (5 mL, 1 M) and then extracted with ethyl acetate (3 × 30 mL). The organic layers were dried over sodium sulfate and the solvent removed under reduced pressure to produce a yellow oil (255 mg) which was carried forward without further purification.

2-oxo-1, 2, 3, 4-tetrahydronaphthalene-1-carboxylic acid (155)
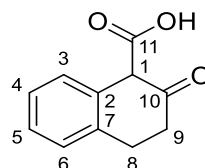
A solution of **154** (51 mg) in tetrahydrofuran (5 mL) was combined with a solution of lithium hydroxide (53 mg, 2.2 mmol, 0.44 M) in water (5 mL) at 4 °C. Hydrochloric acid (3 mL, 1 M) was added and the mixture stirred for 1 h to form a yellow precipitate.

Chapter 8 References

1.   Mullassery, D., and Losty, P.D. 2016. Neuroblastoma. *Paediatrics and Child Health*. **26** (2), 68–72.

2.   Otto, T., Horn, S., Brockmann, M., Eilers, U., Schüttrumpf, L., Popov, N., Kenney, A.M., Schulte, J.H., Beijersbergen, R., Christiansen, H., Berwanger, B., and Eilers, M. 2009. Stabilization of N-myc is a critical function of Aurora A in human neuroblastoma. *Cancer Cell*. **15** (1), 67–78.

3.   Wright, J.H. 1910. Neurocytoma or neuroblastoma, a kind of tumor not generally recognized. *The Journal of Experimental Medicine*. **12** (4), 556–561.

4.   Public Health England 2018. Childhood cancer statistics, England annual report 2018.

5.   Siegel, R.L., Miller, K.D., and Jemal, A. 2016. Cancer statistics, 2016. *CA: A Cancer Journal for Clinicians*. **69** (1), 7–34.

6.   Young, J.L., Ries, L.G., Silverberg, E., Horm, J.W., and Miller, R.W. 1986. Cancer incidence, survival, and mortality for children younger than age 15 years. *Cancer*. **58** (S2), 598–602.

7.   Hara, J. 2012. Development of treatment strategies for advanced neuroblastoma. *International Journal of Clinical Oncology*. **17** (3), 196–203.

8.   Smith, V., and Foster, J. 2018. High-risk neuroblastoma treatment review. *Children*. **5** (9), 114.

9.   Flynn, N., LeFebvre, A., Messahel, B., and Hogg, S.L. 2018. Olive ingestion causing a false suspicion of relapsed neuroblastoma: A case of "oliveblastoma?" *Pediatric Blood and Cancer*. **65** (10), Article no: e27248 [no pagination].

10.   Sharp, S.E., Trout, A.T., Weiss, B.D., and Gelfand, M.J. 2016. MIBG in neuroblastoma diagnostic imaging and therapy. *RadioGraphics*. **36** (1), 258–278.

11. European Medicines Agency 2017. Assessment report - Dinutuximab beta Apeiron. *Science Medicines Health*. **44** (March), 1–129.

12. Mora, J. 2016. Dinutuximab for the treatment of pediatric patients with high-risk neuroblastoma. *Expert Review of Clinical Pharmacology*. **9** (5), 647–653.

13. Ozkaynak, M.F. 2017. Dinutuximab in the treatment of neuroblastoma. *Expert Opinion on Orphan Drugs*. **5** (3), 277–284.

14. Barton, S.E., Najita, J.S., Ginsburg, E.S., Leisenring, W.M., Stovall, M., Weathers, R.E., Sklar, C.A., Robison, L.L., and Diller, L. 2013. Infertility, infertility treatment, and achievement of pregnancy in female survivors of childhood cancer: A report from the Childhood Cancer Survivor Study cohort. *The Lancet Oncology*. **14** (9), 873–881.

15. Bhakta, N., Liu, Q., Ness, K.K., Baassiri, M., Eissa, H., Yeo, F., Chemaitilly, W., Ehrhardt, M.J., Bass, J., Bishop, M.W., Shelton, K., Lu, L., Huang, S., Li, Z., Caron, E., Lanctot, J., Howell, C., Folse, T., et al. 2017. The cumulative burden of surviving childhood cancer: an initial report from the St Jude Lifetime Cohort Study (SJLIFE). *The Lancet*. **390** (10112), 2569–2582.

16. Brinkman, T.M., Zhang, N., Recklitis, C.J., Kimberg, C., Zeltzer, L.K., Muriel, A.C., Stovall, M., Srivastava, D.K., Sklar, C.A., Robison, L.L., and Krull, K.R. 2014. Suicide ideation and associated mortality in adult survivors of childhood cancer. *Cancer*. **120** (2), 271–277.

17. Marumoto, T., Zhang, D., and Saya, H. 2005. Aurora-A - a guardian of poles. *Nature Reviews Cancer*. **5** (1), 42–50.

18. Dodson, C.A., and Bayliss, R. 2012. Activation of Aurora-A kinase by protein partner binding and phosphorylation are independent and synergistic. *Journal of Biological Chemistry*. **287** (2), 1150–1157.

19. Maris, J.M. 2009. Unholy matrimony: Aurora A and N-myc as malignant partners in neuroblastoma. *Cancer Cell*. **15** (1), 5–6.

20. Modi, V., and Dunbrack, R.L. 2019. A Structurally-Validated Multiple Sequence Alignment of 497 Human Protein Kinase Domains. *Scientific*

Reports. **9** (1), 1–16.

21. Roskoski, R. 2016. Classification of small molecule protein kinase inhibitors based upon the structures of their drug-enzyme complexes. *Pharmacological Research*. **103**, 26–48.

22. Kornev, A.P., Taylor, S.S., and Ten Eyck, L.F. 2008. A helix scaffold for the assembly of active protein kinases. *Proceedings of the National Academy of Sciences*. **105** (38), 14377–14382.

23. Paris, J., and Philippe, M. 1990. Poly(A) metabolism and polysomal recruitment of maternal mRNAs during early Xenopus development. *Developmental Biology*. **140** (1), 221–224.

24. Rennie, Y.K., McIntyre, P.J., Akindele, T., Bayliss, R., and Jamieson, A.G. 2016. A TPX2 proteomimetic has enhanced affinity for aurora-A due to hydrocarbon stapling of a helix. *ACS Chemical Biology*. **11** (12), 3383–3390.

25. Garrido, G., and Vernos, I. 2016. Non-centrosomal TPX2-dependent regulation of the Aurora A kinase: functional implications for healthy and pathological cell division. *Frontiers in Oncology*. **6**, article no: 88 [no pagination].

26. Kilchmann, F., Marcaida, M.J., Kotak, S., Schick, T., Boss, S.D., Awale, M., Gonczy, P., and Reymond, J.L. 2016. Discovery of a selective Aurora A kinase inhibitor by virtual screening. *Journal of Medicinal Chemistry*. **59** (15), 7188–7211.

27. Arlot-Bonnemains, Y., Klotzbucher, A., Giet, R., Uzbekov, R., Bihan, R., and Prigent, C. 2001. Identification of a functional destruction box in the Xenopus laevis Aurora-A kinase pEg2. *FEBS Letters*. **508** (1), 149–152.

28. Keen, N., and Taylor, S. 2004. Aurora-kinase inhibitors as anticancer agents. *Nature Reviews Cancer*. **4** (12), 927–936.

29. Alcaraz-Sanabria, A.L., Nieto-Jimenez, C., Corrales-Sanchez, V., Serrano-Oviedo, L., Serrano-Heras, G., Pretel, F.A., Montero, J.C., Burgos, M., Llopis, J., Galan-Moya, E.M., Amir, E., Pandiella, A., and Ocana, A. 2018.

Synthetic lethality interaction between aurora kinases and CHEK1 inhibitors in ovarian cancer. *Journal of Clinical Oncology*. **35** (15_suppl), e17089–e17089.

30. Shah, K.N., Bhatt, R., Rotow, J., Rohrberg, J., Olivas, V., Wang, V.E., Hemmati, G., Martins, M.M., Maynard, A., Kuhn, J., Galeas, J., Donnella, H.J., Kaushik, S., Ku, A., Dumont, S., Krings, G., Haringsma, H.J., Robillard, L., et al. 2019. Aurora kinase A drives the evolution of resistance to third-generation EGFR inhibitors in lung cancer. *Nature Medicine*. **25** (1), 111–118.

31. Gustafson, W.C., Meyerowitz, J.G., Nekritz, E.A., Chen, J., Benes, C., Charron, E., Simonds, E.F., Seeger, R., Matthay, K.K., Hertz, N.T., Eilers, M., Shokat, K.M., and Weiss, W.A. 2014. Drugging MYCN through an allosteric transition in Aurora kinase A. *Cancer Cell*. **26** (3), 414–427.

32. Bhatia, P., Menigatti, M., Brocard, M., Morley, S., and Ferrari, S. 2010. Mitotic DNA damage targets the Aurora A/TPX2 complex. *Cell Cycle*. **9** (22), 4592–4599.

33. Dodson, C.A., Yeoh, S., Haq, T., and Bayliss, R. 2013. A kinetic test characterizes kinase intramolecular and intermolecular autophosphorylation mechanisms. *Science Signaling*. **6** (282), ra54.

34. Eyers, P.A., Erikson, E., Chen, L.G., and Maller, J.L. 2003. A novel mechanism for activation of the protein kinase Aurora A. *Current Biology*. **13** (8), 691–697.

35. Dauch, D., Rudalska, R., Cossa, G., Nault, J.-C., Kang, T.-W., Wuestefeld, T., Hohmeyer, A., Imbeaud, S., Yevsa, T., Hoenicke, L., Pantsar, T., Bozko, P., Malek, N.P., Longerich, T., Laufer, S., Poso, A., Zucman-Rossi, J., Eilers, M., et al. 2016. A MYC–aurora kinase A protein complex represents an actionable drug target in p53-altered liver cancer. *Nature Medicine*. **22** (7), 744–753.

36. Hurlin, P.J. 2005. N-myc functions in transcription and development. *Birth*

Defects Research Part C - Embryo Today: Reviews. **75** (4), 340–352.

37. Brockmann, M., Poon, E., Berry, T., Carstensen, A., and Hedwig, E. 2013. Small molecule inhibitors of Aurora-A induce proteasomal degradation of N-myc in childhood neuroblastoma. *Cancer Cell.* **24** (1), 75–89.

38. Lioutas, A., and Vernos, I. 2013. Aurora A kinase and its substrate TACC3 are required for central spindle assembly. *EMBO Reports.* **14** (9), 829–836.

39. Campo, L., and Breuer, E.K. 2018. Inhibition of TACC3 by a small molecule inhibitor in breast cancer. *Biochemical and Biophysical Research Communications.* **498** (4), 1085–1092.

40. Burgess, S.G., Mukherjee, M., Sabir, S., Joseph, N., Gutiérrez-Caballero, C., Richards, M.W., Huguenin-Dezot, N., Chin, J.W., Kennedy, E.J., Pfuhl, M., Royle, S.J., Gergely, F., and Bayliss, R. 2018. Mitotic spindle association of TACC3 requires Aurora-A-dependent stabilization of a cryptic α-helix. *The EMBO Journal.*, e97902.

41. Burgess, S.G., Peset, I., Joseph, N., Cavazza, T., Vernos, I., Pfuhl, M., Gergely, F., and Bayliss, R. 2015. Aurora-A-dependent control of TACC3 influences the rate of mitotic spindle assembly. *PLoS Genetics.* **11** (7), article no: e1005345 [no pagination].

42. Thakur, H.C., Singh, M., Nagel-Steger, L., Kremer, J., Prumbaum, D., Fansa, E.K., Ezzahoini, H., Nouri, K., Gremer, L., Abts, A., Schmitt, L., Raunser, S., Ahmadian, M.R., and Piekorz, R.P. 2014. The centrosomal adaptor TACC3 and the microtubule polymerase chTOG interact via defined C-terminal subdomains in an Aurora-A kinase-independent manner. *Journal of Biological Chemistry.* **289** (1), 74–88.

43. Romain, C., Paul, P., Kim, K.W., Lee, S., Qiao, J., and Chung, D.H. 2014. Targeting Aurora kinase-A downregulates cell proliferation and angiogenesis in neuroblastoma. *Journal of Pediatric Surgery.* **49** (1), 159–165.

44. Lee, J.K., Phillips, J.W., Smith, B.A., Park, J.W., Stoyanova, T., McCaffrey, E.F., Baertsch, R., Sokolov, A., Meyerowitz, J.G., Mathis, C., Cheng, D.,

Stuart, J.M., Shokat, K.M., Gustafson, W.C., Huang, J., and Witte, O.N. 2016. N-Myc drives neuroendocrine prostate cancer initiated from human prostate epithelial cells. *Cancer Cell*. **29** (4), 536–547.

45. Richards, M.W., Burgess, S.G., Poon, E., Carstensen, A., Eilers, M., Chesler, L., and Bayliss, R. 2016. Structural basis of N-Myc binding by Aurora-A and its destabilization by kinase inhibitors. *Proceedings of the National Academy of Sciences of the United States of America*. **113** (48), 13726–13731.

46. Shang, X., Burlingame, S.M., Okcu, M.F., Ge, N., Russell, H. V, Egler, R.A., David, R.D., Vasudevan, S.A., Yang, J., and Nuchtern, J.G. 2009. Aurora A is a negative prognostic factor and a new therapeutic target in human neuroblastoma. *Molecular Cancer Therapeutics*. **8** (8), 2461–2469.

47. Bayliss, R., Burgess, S.G., and McIntyre, P.J. 2017. Switching Aurora-A kinase on and off at an allosteric site. *FEBS Journal*. **284** (18), 2947–2954.

48. Burgess, S.G., Oleksy, A., Cavazza, T., Richards, M.W., Vernos, I., Matthews, D., and Bayliss, R. 2016. Allosteric inhibition of Aurora-A kinase by a synthetic vNAR domain. *Open Biology*. **6** (7), article no: 160089 [no pagination].

49. Sloane, D.A., Trikic, M.Z., Chu, M.L., Lamers, M.B., Mason, C.S., Mueller, I., Savory, W.J., Williams, D.H., and Eyers, P.A. 2010. Drug-resistant Aurora A mutants for cellular trget validation of the small molecule kinase inhibitors MLN8054 and MLN8237. *ACS Chemical Biology*. **5** (6), 563–576.

50. de Groot, C.O., Hsia, J.E., Anzola, J. V, Motamedi, A., Yoon, M., Wong, Y.L., Jenkins, D., Lee, H.J., Martinez, M.B., Davis, R.L., Gahman, T.C., Desai, A., and Shiau, A.K. 2015. A cell biologist's field guide to Aurora kinase inhibitors. *Frontiers in Oncology*. **5**, article no: 285 [no pagination].

51. Cee, V.J., Schenkel, L.B., Hodous, B.L., Deak, H.L., Nguyen, H.N., Olivieri, P.R., Romero, K., Bak, A., Be, X., Bellon, S., Bush, T.L., Cheng, A.C., Chung, G., Coats, S., Eden, P.M., Hanestad, K., Gallant, P.L., Gu, Y., et al.

2010. Discovery of a potent, selective, and orally bioavailable pyridinyl-pyrimidine phthalazine aurora kinase inhibitor. *Journal of Medicinal Chemistry*. **53** (17), 6368–6377.

52. Karra, S., Xiao, Y., Chen, X., Liu-Bujalski, L., Huck, B., Sutton, A., Goutopoulos, A., Askew, B., Josephson, K., Jiang, X., Shutes, A., Shankar, V., Noonan, T., Garcia-Berrios, G., Dong, R., Dhanabal, M., Tian, H., Wang, Z., et al. 2013. SAR and evaluation of novel 5H-benzo[c][1,8]naphthyridin-6-one analogs as Aurora kinase inhibitors. *Bioorganic and Medicinal Chemistry Letters*. **23** (10), 3081–3087.

53. Prime, M.E., Courtney, S.M., Brookfield, F.A., Marston, R.W., Walker, V., Warne, J., Boyd, A.E., Kairies, N.A., Von Der Saal, W., Limberg, A., Georges, G., Engh, R.A., Goller, B., Rueger, P., and Rueth, M. 2011. Phthalazinone pyrazoles as potent, selective, and orally bioavailable inhibitors of Aurora-A kinase. *Journal of Medicinal Chemistry*. **54** (1), 312–319.

54. Martin, M.P., Zhu, J.Y., Lawrence, H.R., Pireddu, R., Luo, Y., Alam, R., Ozcan, S., Sebti, S.M., Lawrence, N.J., and Schönbrunn, E. 2012. A novel mechanism by which small molecule inhibitors induce the DFG flip in Aurora A. *ACS Chemical Biology*. **7** (4), 698–706.

55. Manning, G., Whyte, D.B., Martinez, R., Hunter, T., and Sudarsanam, S. 2002. The protein kinase complement of the human genome. *Science*. **298** (5600), 1912–1934.

56. Bhullar, K.S., Lagarón, N.O., McGowan, E.M., Parmar, I., Jha, A., Hubbard, B.P., and Rupasinghe, H.P.V. 2018. Kinase-targeted cancer therapies: Progress, challenges and future directions. *Molecular Cancer*. **17** (1), 1–20.

57. Zuccotto, F., Ardini, E., Casale, E., and Angiolini, M. 2010. Through the "gatekeeper door": Exploiting the active kinase conformation. *Journal of Medicinal Chemistry*. **53** (7), 2681–2694.

58. British National Formulary 2019. Trametinib.

59. Roskoski, R. 2017. Allosteric MEK1/2 inhibitors including cobimetanib and trametinib in the treatment of cutaneous melanomas. *Pharmacological Research*. **117**, 20–31.

60. Bavetsias, V., and Linardopoulos, S. 2015. Aurora kinase inhibitors: Current status and outlook. *Frontiers in Oncology*. **5**, article no: 278 [no pagination].

61. Bavetsias, V., Large, J.M., Sun, C., Bouloc, N., Kosmopoulou, M., Matteucci, M., Wilsher, N.E., Martins, V., Reynisson, J., Atrash, B., Faisal, A., Urban, F., Valenti, M., De Haven Brandon, A., Box, G., Raynaud, F.I., Workman, P., Eccles, S.A., et al. 2010. Imidazo[4,5-b]pyridine derivatives as inhibitors of Aurora kinases: Lead optimization studies toward the identification of an orally bioavailable preclinical development candidate. *Journal of Medicinal Chemistry*. **53** (14), 5213–5228.

62. Bavetsias, V., Crumpler, S., Sun, C., Avery, S., Atrash, B., Faisal, A., Moore, A.S., Kosmopoulou, M., Brown, N., Sheldrake, P.W., Bush, K., Henley, A., Box, G., Valenti, M., De Haven Brandon, A., Raynaud, F.I., Workman, P., Eccles, S.A., et al. 2012. Optimization of imidazo[4,5- b]pyridine-based kinase inhibitors: Identification of a dual FLT3/aurora kinase inhibitor as an orally bioavailable preclinical development candidate for the treatment of acute myeloid leukemia. *Journal of Medicinal Chemistry*. **55** (20), 8721–8734.

63. Cancilla, M.T., He, M.M., Viswanathan, N., Simmons, R.L., Taylor, M., Fung, A.D., Cao, K., and Erlanson, D.A. 2008. Discovery of an Aurora kinase inhibitor through site-specific dynamic combinatorial chemistry. *Bioorganic and Medicinal Chemistry Letters*. **18** (14), 3978–3981.

64. Kong, Y., Bender, A., and Yan, A. 2018. Identification of novel Aurora kinase A (AURKA) inhibitors via hierarchical ligand-based vrtual screening. *Journal of Chemical Information and Modeling*. **58** (1), 36–47.

65. Meyerowitz, J.G., Gustafson, W.C., Weiss, W.A., Hertz, N.T., and Shokat, K.M. 2014. Aurora kinase inhibitors.

66. Park, H., Jung, H.Y., Mah, S., and Hong, S. 2018. Systematic computational design and identification of low picomolar inhibitors of Aurora kinase A.

67. Carry, J.C., Clerc, F., Minoux, H., Schio, L., Mauger, J., Nair, A., Parmantier, E., Le Moigne, R., Delorme, C., Nicolas, J.P., Krick, A., Abécassis, P.Y., Crocq-Stuerga, V., Pouzieux, S., Delarbre, L., Maignan, S., Bertrand, T., Bjergarde, K., et al. 2015. SAR156497, an exquisitely selective inhibitor of Aurora kinases. *Journal of Medicinal Chemistry*. **58** (1), 362–375.

68. DuBois, S.G., Marachelian, A., Fox, E., Kudgus, R.A., Reid, J.M., Groshen, S., Malvar, J., Bagatell, R., Wagner, L., Maris, J.M., Hawkins, R., Courtier, J., Lai, H., Goodarzian, F., Shimada, H., Czarnecki, S., Tsao-Wei, D., Matthay, K.K., et al. 2016. Phase I study of the Aurora A kinase inhibitor alisertib in combination with irinotecan and temozolomide for patients with relapsed or refractory neuroblastoma: A NANT (new approaches to neuroblastoma therapy) trial. *Journal of Clinical Oncology*. **34** (12), 1368–1375.

69. DuBois, S.G., Marachelian, A., Fox, E., Kudgus, R.A., Reid, J.M., and Groshen, S. 2016. Alisertib combined with chemotherapy achieves response in neuroblastoma. *Cancer Discovery*. **6** (4), 339.

70. National Cancer Institute 2019. Clinical trials using alisertib.

71. Goff, L.W., Azad, N.S., Stein, S., Whisenant, J., Vaishampayan, U.N., Hochster, H.S., Connolly, R.M., Weise, A.M., LoRusso, P., El-Rifai, W., and Berlin, J. 2018. Phase I study combining the Aurora kinase A (AURKA) inhibitor alisertib (Ali) with mFOLFOX in gastrointestinal (GI) cancer. *Journal of Clinical Oncology*. **35** (15_suppl), 2593–2593.

72. Shah, H.A., Fischer, J.H., Venepalli, N.K., Danciu, O.C., Christian, S., Russell, M.J., Liu, L.C., Zacny, J.P., and Dudek, A.Z. 2019. Phase I study of Aurora A kinase inhibitor alisertib (MLN8237) in combination with selective VEGFR inhibitor pazopanib for therapy of advanced solid tumors. *American Journal of Clinical Oncology: Cancer Clinical Trials*. **42** (5), 413–420.

73. Zhou, X., Pant, S., Nemunaitis, J., Craig Lockhart, A., Falchook, G., Bauer,

T.M., Patel, M., Sarantopoulos, J., Bargfrede, M., Muehler, A., Rangachari, L., Zhang, B., and Venkatakrishnan, K. 2018. Effects of rifampin, itraconazole and esomeprazole on the pharmacokinetics of alisertib, an investigational Aurora A kinase inhibitor in patients with advanced malignancies. *Investigational New Drugs*. **36** (2), 248–258.

74.   Tan, A.R., Gibbon, D.G., Stein, M.N., Lindquist, D., Edenfield, J.W., Martin, J.C., Gregory, C., Suttle, A.B., Tada, H., Botbyl, J., and Stephenson, J.J. 2013. Effects of ketoconazole and esomeprazole on the pharmacokinetics of pazopanib in patients with solid tumors. *Cancer Chemotherapy and Pharmacology*. **71** (6), 1635–1643.

75.   Friedberg, J., Mahadevan, D., Jung, J., Persky, D.O., Lossos, I.S., Danaee, H., Zhou, X., Leonard, E.J., and Bernstein, S.H. 2011. Phase 2 trial of alisertib (MLN8237), an investigational, potent inhibitor of Aurora A kinase (AAK), in patients (pts) with aggressive B- and T-cell non-Hodgkin lymphoma (NHL). *Blood*. **118** (21), 95.

76.   Takeda 2015. Takeda announces termination of alisertib phase III trial in relapsed or refractory peripheral T-cell lymphoma.

77.   Barr, P.M., Li, H., Spier, C., Mahadevan, D., LeBlanc, M., Ul Haq, M., Huber, B.D., Flowers, C.R., Wagner-Johnston, N.D., Horwitz, S.M., Fisher, R.I., Cheson, B.D., Smith, S.M., Kahl, B.S., Bartlett, N.L., and Friedberg, J.W. 2015. Phase II intergroup trial of alisertib in relapsed and refractory peripheral T-cell lymphoma and transformed Mycosis fungoides: SWOG 1108. *Journal of Clinical Oncology*. **33** (21), 2399–2404.

78.   Mosse, Y.P., Fox, E., Teachey, D.T., Reid, J.M., Safgren, S.L., Carol, H., Lock, R.B., Houghton, P.J., Smith, M.A., Hall, D., Barkauskas, D.A., Krailo, M., Voss, S.D., Berg, S.L., Blaney, S.M., and Weigel, B.J. 2019. A phase II study of alisertib in children with recurrent/refractory solid tumors or leukemia: Children's Oncology Group Phase I and pilot Consortium (ADVL0921). *Clinical Cancer Research*. **25** (11), 3229–3238.

79. Kelly, K.R., Friedberg, J.W., Park, S.I., McDonagh, K., Hayslip, J., Persky, D., Ruan, J., Puvvada, S., Rosen, P., Iyer, S.P., Stefanovic, A., Bernstein, S.H., Weitman, S., Karnad, A., Monohan, G., VanderWalde, A., Mena, R., Schmelz, M., et al. 2018. Phase I study of the investigational Aurora A kinase inhibitor alisertib plus rituximab or rituximab/vincristine in relapsed/refractory aggressive B-cell lymphoma. *Clinical Cancer Research*. **24** (24), 6150–6159.

80. Borisa, A.C., and Bhatt, H.G. 2017. A comprehensive review on Aurora kinase: Small molecule inhibitors and clinical trial studies. *European Journal of Medicinal Chemistry*. **140**, 1–19.

81. DuBois, S.G., Mosse, Y.P., Fox, E., Kudgus, R.A., Reid, J.M., McGovern, R., Groshen, S., Bagatell, R., Maris, J.M., Twist, C.J., Goldsmith, K., Meaghan Granger, M., Weiss, B., Park, J.R., Macy, M.E., Cohn, S.L., Yanik, G., Wagner, L.M., et al. 2018. Phase II trial of alisertib in combination with irinotecan and temozolomide for patients with relapsed or refractory neuroblastoma. *Clinical Cancer Research*. **24** (24), 6142–6149.

82. Lin, J., Patel, S.A., Sama, A.R., Hoffman-Censits, J.H., Kennedy, B., Kilpatrick, D., Ye, Z., Yang, H., Mu, Z., Leiby, B., Lewis, N., Cristofanilli, M., and Kelly, W.K. 2016. A phase I/II study of the investigational drug alisertib in combination with abiraterone and prednisone for patients with metastatic castration-resistant prostate cancer progressing on abiraterone. *The Oncologist*. **21** (11), 1296-1297e.

83. London, W.B., Castel, V., Monclair, T., Ambros, P.F., Pearson, A.D.J., Cohn, S.L., Berthold, F., Nakagawara, A., Ladenstein, R.L., Iehara, T., and Matthay, K.K. 2011. Clinical and biologic features predictive of survival after relapse of neuroblastoma: A report from the International Neuroblastoma Risk Group Project. *Journal of Clinical Oncology*. **29** (24), 3286–3292.

84. Fletcher, J.I., Ziegler, D.S., Trahair, T.N., Marshall, G.M., Haber, M., and Norris, M.D. 2018. Too many targets, not enough patients: Rethinking neuroblastoma clinical trials. *Nature Reviews Cancer*. **18** (6), 389–400.

85. Harrington, E.A., Bebbington, D., Moore, J., Rasmussen, R.K., Ajose-Adeogun, A.O., Nakayama, T., Graham, J.A., Demur, C., Hercend, T., Diu-Hercend, A., Su, M., Golec, J.M.C., and Miller, K.M. 2004. VX-680, a potent and selective small-molecule inhibitor of the Aurora kinases, suppresses tumor growth in vivo. *Nature Medicine*. **10** (3), 262–267.

86. Kollareddy, M., Zheleva, D., Dzubak, P., Brahmkshatriya, P.S., Lepsik, M., and Hajduch, M. 2012. Aurora kinase inhibitors: Progress towards the clinic. *Investigational New Drugs*. **30** (6), 2411–2432.

87. Doggrell, S.A. 2004. Dawn of Aurora kinase inhibitors as anticancer drugs. *Expert Opinion on Investigational Drugs*. **13** (9), 1199–1201.

88. Boss, D.S., Beijnen, J.H., and Schellens, J.H.M. 2009. Clinical experience with Aurora kinase inhibitors: A review. *The Oncologist*. **14** (8), 780–793.

89. Gillet, V., Johnson, P., Mata, P., Sike, S., and Williams, P. 1993. SPROUT: A program for structure generation. *Journal of Computer-Aided Molecular Design*. **7** (2), 127–153.

90. Chéron, N., Jasty, N., and Shakhnovich, E.I. 2016. OpenGrowth: An automated and rational algorithm for finding new protein ligands. *Journal of Medicinal Chemistry*. **59** (9), 4171–4188.

91. Halgren, T.A. 2009. Identifying and characterizing binding sites and assessing druggability. *Journal of Chemical Information and Modeling*. **49** (2), 377–389.

92. Wilson, A.J. 2015. Helix mimetics: Recent developments. *Progress in Biophysics and Molecular Biology*. **119** (1), 33–40.

93. Wilson, A.J. 2009. Inhibition of protein-protein interaction using designed molecules. *Chemical Society Reviews*. **38** (12), 3289–3300.

94. Azzarito, V., Rowell, P., Barnard, A., Edwards, T.A., Macdonald, A., Warriner, S.L., and Wilson, A.J. 2016. Probing protein surfaces: QSAR analysis with helix mimetics. *ChemBioChem*. **17** (8), 768–773.

95. Enamine 2016. Alpha-helix mimetics.

96. Friesner, R.A., Banks, J.L., Murphy, R.B., Halgren, T.A., Klicic, J.J., Mainz, D.T., Repasky, M.P., Knoll, E.H., Shelley, M., Perry, J.K., Shaw, D.E., Francis, P., and Shenkin, P.S. 2004. Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *Journal of Medicinal Chemistry.* **47** (7), 1739–49.

97. ChemBridge 2017. DIVERSet diverse screening libraries.

98. KeyOrganics 2016. BIONET fragment libraries.

99. Asinex 2016. Building blocks/Fragments.

100. ChemDiv 2016. Peptidomimetic library.

101. PerkinElmer 2014. Mobility shift assay development guide.

102. Eglen, R.M., Reisine, T., Roby, P., Rouleau, N., Illy, C., Bossé, R., and Bielefeld, M. 2008. The use of AlphaScreen technology in HTS: Current status. *Current Chemical Genomics.* **1**, 2–10.

103. Bayliss, R., Burgess, S.G., Leen, E., and Richards, M.W. 2017. A moving target: Structure and disorder in pursuit of Myc inhibitors. *Biochemical Society Transactions.* **45** (3), 709–717.

104. Halgren, T. 2007. New method for fast and accurate binding-site identification and analysis. *Chemical Biology and Drug Design.* **69** (2), 146–148.

105. Baell, J., and Walters, M.A. 2014. Chemistry: Chemical con artists foil drug discovery. *Nature.* **513** (7519), 481–483.

106. Dahlin, J.L., Nissink, J.W.M., Strasser, J.M., Francis, S., Higgins, L., Zhou, H., Zhang, Z., and Walters, M.A. 2015. PAINS in the assay: Chemical mechanisms of assay interference and promiscuous enzymatic inhibition observed during a sulfhydryl-scavenging HTS. *Journal of Medicinal Chemistry.* **58** (5), 2091–2113.

107. Dahlin, J.L., and Walters, M.A. 2016. How to Triage PAINS-Full Research.

Assay and Drug Development Technologies. **14** (3), 168–174.

108. Baell, J.B., and Nissink, J.W.M. 2018. Seven Year Itch: Pan-Assay Interference Compounds (PAINS) in 2017 - Utility and Limitations. *ACS Chemical Biology*. **13** (1), 36–44.

109. Richters, A., Basu, D., Engel, J., Ercanoglu, M.S., Balke-Want, H., Tesch, R., Thomas, R.K., and Rauh, D. 2015. Identification and further development of potent TBK1 inhibitors. *ACS Chemical Biology*. **10** (1), 289–298.

110. Michaelis, M., Selt, F., Rothweiler, F., Löschmann, N., Nüsse, B., Dirks, W.G., Zehner, R., and Cinatl, J. 2014. Aurora kinases as targets in drug-resistant neuroblastoma cells. *PLoS ONE*. **9** (9), article no: e108758 [no pagination].

111. Edgar, R.C. 2004. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*. **32** (5), 1792–1797.

112. Newman, J., Egan, D., Walter, T.S., Meged, R., Berry, I., Jelloul, M.B., Sussman, J.L., Stuart, D.I., and Perrakis, A. 2005. Towards rationalization of crystallization screening for small- to medium-sized academic laboratories: The PACT/JCSG+ strategy. *Acta Crystallographica Section D*. **61** (10), 1426–1431.

113. Page, R., Grzechnik, S.K., Canaves, J.M., Spraggon, G., Kreusch, A., Kuhn, P., Stevens, R.C., and Lesley, S.A. 2003. Shotgun crystallization strategy for structural genomics: An optimized two-tiered crystallization screen against the Thermotoga maritima proteome. *Acta Crystallographica Section D*. **59** (6), 1028–1037.

114. Emsley, P., and Cowtan, K. 2004. Coot: Model-building tools for molecular graphics. *Acta Crystallographica Section D2*. **60** (12), 2126–2132.

115. Stone, A., Friedman, D.N., Worgall, S., Kushner, B.H., Wolden, S., Modak, S., LaQuaglia, M.P., Wu, X., Cheung, N.K., and Sklar, C.A. 2017. Long-term pulmonary outcomes in pediatric survivors of high-risk neuroblastoma. *Journal of Pediatric Hematology/Oncology*. **39** (7), 547–554.

116. Sang, C.Y., Qin, W.W., Zhang, X.J., Xu, Y., Ma, Y.Z., Wang, X.R., Hui, L., and Chen, S.W. 2019. Synthesis and identification of 2,4-bisanilinopyrimidines bearing 2,2,6,6-tetramethylpiperidine-N-oxyl as potential Aurora A inhibitors. *Bioorganic and Medicinal Chemistry*. **27** (1), 65–78.

117. Wang, W., Feng, X., Liu, H.X., Chen, S.W., and Hui, L. 2018. Synthesis and biological evaluation of 2,4-disubstituted phthalazinones as Aurora kinase inhibitors. *Bioorganic and Medicinal Chemistry*. **26** (12), 3217–3226.

118. Graff, J.N., Higano, C.S., Hahn, N.M., Taylor, M.H., Zhang, B., Zhou, X., Venkatakrishnan, K., Leonard, E.J., and Sarantopoulos, J. 2016. Open-label, multicenter, phase 1 study of alisertib (MLN8237), an Aurora A kinase inhibitor, with docetaxel in patients with solid tumors. *Cancer*. **122** (16), 2524–2533.

119. Lea, W.A., and Simeonov, A. 2011. Fluorescence polarization assays in small molecule screening. *Expert Opinion on Drug Discovery*. **6** (1), 17–32.

120. Hall, M.D., Yasgar, A., Peryea, T., Braisted, J.C., Jadhav, A., Simeonov, A., and Coussens, N.P. 2016. Fluorescence polarization assays in high-throughput screening and drug discovery: A review. *Methods and Applications in Fluorescence*. **4** (2), article no: 22001 [no pagination].

121. Schrödinger 2020. LigPrep Schrödinger Release 2020-3.

122. Friesner, R.A., Murphy, R.B., Repasky, M.P., Frye, L.L., Greenwood, J.R., Halgren, T.A., Sanschagrin, P.C., and Mainz, D.T. 2006. Extra precision glide: Docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes. *Journal of Medicinal Chemistry*. **49** (21), 6177–6196.

123. Hawkins, P.C.D., Skillman, A.G., and Nicholls, A. 2007. Comparison of shape-matching and docking as virtual screening tools. *Journal of Medicinal Chemistry*. **50** (1), 74–82.

124. GraphPad Software 2020. GraphPad Prism.

125. Ostermann, N., Ruedisser, S., Ehrhardt, C., Breitenstein, W., Marzinzik, A., Jacoby, E., Vangrevelinghe, E., Ottl, J., Klumpp, M., Hartwieg, J.C.D., Cumin, F., Hassiepen, U., Trappe, J., Sedrani, R., Geisse, S., Gerhartz, B., Richert, P., Francotte, E., et al. 2013. A novel class of oral direct renin inhibitors: Highly potent 3,5-disubstituted piperidines bearing a tricyclic P3 - P1 pharmacophore. *Journal of Medicinal Chemistry*. **56** (6), 2196–2206.

Chapter 9 Appendix

Appendix 1: Compound library based on MES structure
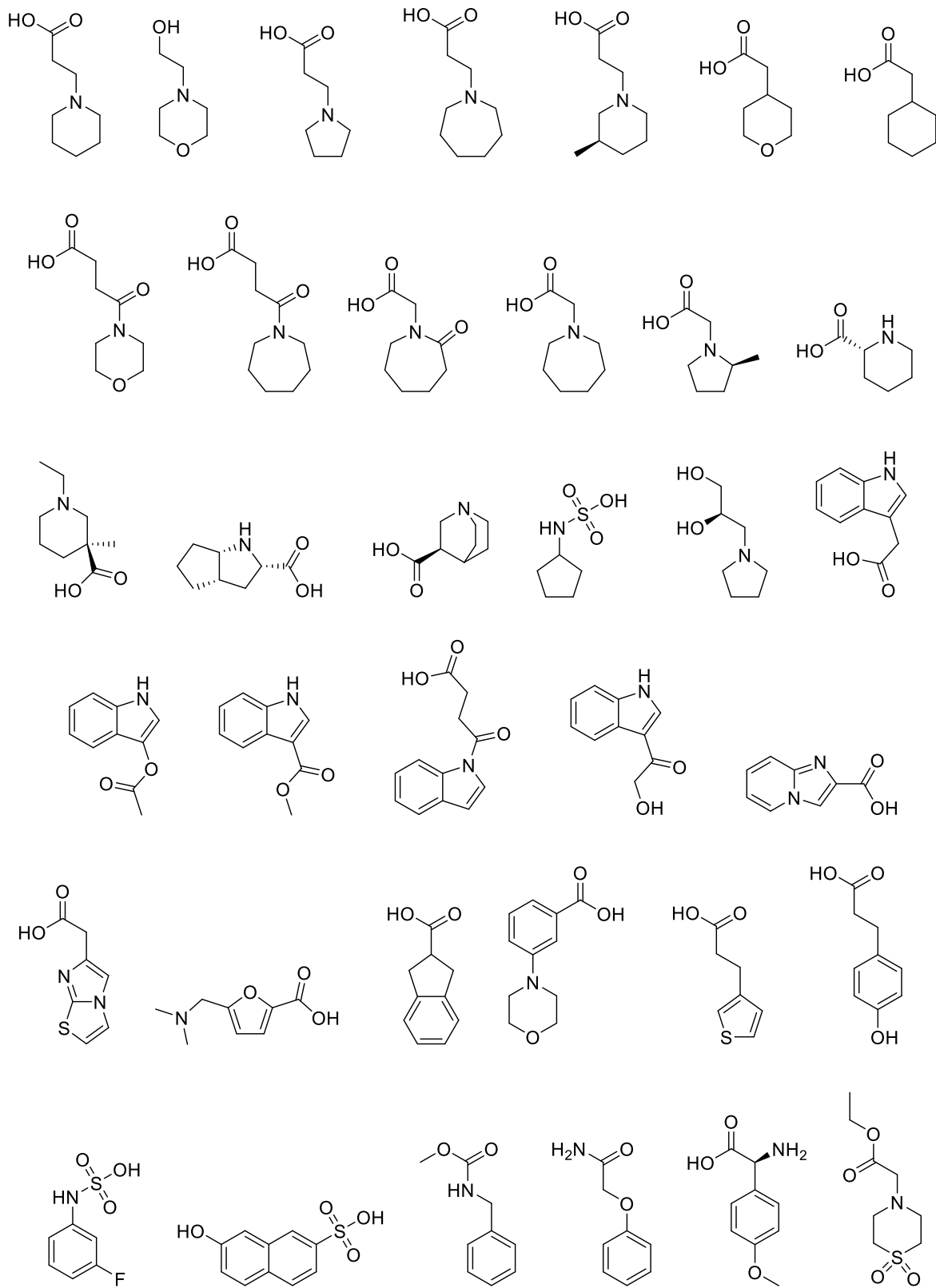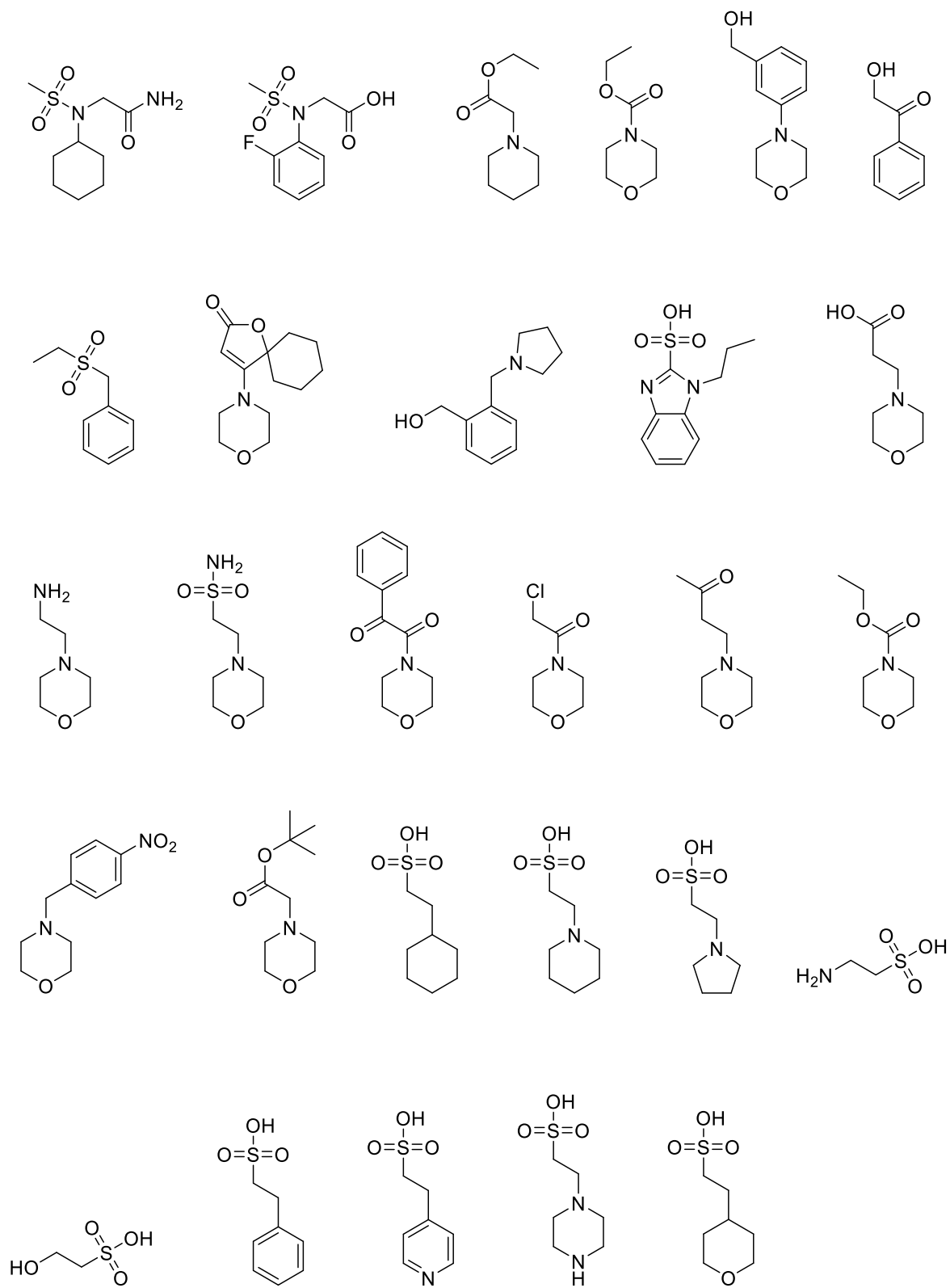
Appendix 2: Aurora A protein BLAST results

Table 9.1 Protein BLAST results with similarity to Aurora A

Protein	Accession Number	% Identity	E value	Total score
Aurora A	NP_001310234.1	Query (100%)	Query (0)	-
Aurora B	NP_001300881.1	71	7e-139	401
Aurora C	NP_001015878.1	68	1e-141	408
PLK4	BAB69958.1	38	2e-58	207
SNRK	AAH71567.1	36	3e-49	181
MARK3	pdb|2QNJ|A	35	3e-49	172
MARK4	pdb|5ES1|A	34	4e-49	172
Greatwall kinase	pdb|5LOH|A	35	4e-49	172
PDK1	pdb|3PWY|A	37	9e-49	170
ULK3	pdb|6FDY|U	35	1e-48	169
Ribosomal S6	AAH06106.3	39	1e-48	174
MARK2	pdb|5EAK|A	33	6e-47	166
PLK2	pdb|4I5M|A	31	9e-47	165
SIK3	XP_016872913.1	34	8e-47	175
DCLK2	XP_016863321.1	35	2e-46	172
PLK1	pdb|3THB|A	33	1e-44	160
DCLK3	NP_208382.1	35	1e-44	166
CAMK1	pdb|4FG9|A	32	2e-44	159

AKT2	pdb|1MRV|A	36	2e-43	157
AKT3	NP_001193658	33	4e-43	159
SGK1	pdb|3HDM|A	36	5e-43	157
SIK1	NP_775490.2	31	3e-42	160
SGK2	AAV38901.1	35	7e-41	151
BRSK1	NP_115806.1	30	2e-40	155
SNF1-like kinase	EAX09496.1	31	3e-40	155
CAMK4	pdb|2W4O|A	33	4e-40	148
DAPK3	pdb|1YRP|A	34	8e-40	145
ROCK1	pdb|3TV7|A	33	2e-38	145
SAD-A	AAS86443.1	30	4e-38	149
CAMK2	pdb|6CMJ|A	32	1e-37	141
DAPK1	pdb|1ZWS|A	32	1e-37	140
DAPK2	pdb|1Z9X|A	32	2e-37	140
MST4	AAH99843.1	32	7e-37	141
ROCK2	pdb|4L6Q|A	31	5e-36	139
Chk2	pdb|2CN5|A	33	5e-36	137
CDK3	NP_001249.1	30	9e-36	135
PKCα	BAU98542.1	32	7e-36	142
PKCβ	NP_997700.1	32	1e-35	141
BRSK2	NP_001243559.1	30	1e-35	142
CaMKII	AAD04566.1	31	2e-35	137

PKN2	AAI25200.1	33	5e-35	140
ULK2	pdb|6QAT|A	36	5e-35	133
MAST3	NP_055831.1	29	6e-35	140
PKCδ	NP_001341605.1	31	6e-35	139
MRCKβ	NP_006026.3	35	8e-35	140
MRCKα	AAI36334.1	33	8e-35	140
ULK1	pdb|4WNO|A	33	1e-34	132
MKNK1	pdb|2HW6|A	31	1e-34	132
MAST2	AAH06166.1	30	3e-34	137
CDK5	pdb|4AU8|A	30	7e-34	130
CDK2	pdb|5K4J|A	30	7e-34	130
Nek3	NP_002489.1	30	2e-33	133
Nek2	pdb|2W5A|A	32	5e-33	127
PKCι	pdb|5LIH|A	28	7e-33	129
MLCK	BAB21504.1	26	7e-33	134
SLK3	AAB82560.1	31	9e-33	130
PKCζ	AAA36488.1	28	3e-32	131
CDK1	pdb|4Y72|A	29	3e-32	126
PKN1	AAH40061.1	33	4e-32	131
PAK6	pdb|4KS7|A	30	7e-32	125
PKD1	ABE96833.1	32	8e-32	130
PKCμ	CAA53384.1	32	8e-32	130
STK10	pdb|6GTT|A	30	1e-31	124

DMPK	AAC14450.1	34	2e-31	128
MYLK2	NP_149109.1	30	3e-31	128
MAPKAPK3	NP_001230854.1	31	3e-31	125
Nek1	pdb|4APC|A	27	3e-31	124
PAK2	NP_002568.2	27	4e-31	127
STK4	pdb|3COM|A	29	4e-31	123
PSKH2	NP_149117.1	28	5e-31	124
Chk1	pdb|4FSY|A	29	7e-31	122
Nek7	pdb|5DE2|A	30	9e-31	122
MAP2K1	pdb|3SLS|A	28	2e-30	121
STK3	pdb|6AO5|A	29	3e-30	122
MAPKAPK2	pdb|2JBO|A	32	3e-30	121
PIM3	NP_001001852.2	30	6e-30	120
MEK3	AAB41729.1	32	1e-29	124
Nek6	AAG13417.1	27	2e-29	123
CASK	pdb|3MFR|A	28	3e-29	119
Nek8	NP_835464.1	31	4e-29	122
PNCK	AAH64422.1	34	1e-28	115
SPEG kinase	NP_005867.3	32	2e-28	121
PAK1	AAC50590.1	28	2e-28	119
PAK4	AAH02921.1	27	2e-28	118
PAK7	AAH24179.1	30	2e-28	120
PAK5	NP_065074.1	30	3e-28	120

CAMKV	AAH00497	26	3e-28	119
STK11	NP_000446.1	28	4e-28	117
MKNK2	pdb|2AC5|A	29	4e-28	115
Nek9	NP_149107.4	29	7e-28	119
PIM-2	pdb|2IWI|A	29	1e-27	114
MAP3K5	pdb|2CLQ|A	29	9e-28	114
PIM-1	pdb|4BZN|A	29	2e-27	113
LATS1	NP_001337268.1	30	3e-27	117
MAP3K8	NP_001231063.1	31	7e-27	114
MAP3K6	AAI09033.1	27	1e-26	114
CDK13	NP_003709.3	28	1e-26	115
TAOK1	NP_079418.1	31	1e-26	115
MAP3K2	NP_001358839.1	29	1e-26	115
TAOK3	pdb|6BDN|A	32	3e-26	110
MAP3K4	AAI43736.1	27	4e-26	114
CDK16	AAH06190.1	26	5e-26	112
TAOK2	NP_004774.1	32	1e-25	112
MAP3K3	pdb|5J5T|A	27	2e-25	109
LATS2	NP_055387.2	30	2e-25	111
PASK	NP_001239051.1	29	3e-25	111
OSR1	NP_005100.1	28	2e-24	107
TNIK	NP_001155034.1	29	1e-23	106

Appendix 3: Initial virtual screening hits at the salt bridge binding site taken forward for investigation

Compound ID	Stock concentration (mM)	Compound ID	Stock concentration (mM)
AC11871	100.00	AC16190	100.00
AC42061	110.62	CC70663	100.00
CC47846	109.93	CC68213	100.00
CC26813	144.43	CC52213	100.00
CC00713	127.53	CC00813	100.00
CC01963	110.72	AC25987	100.00
CC08909	125.42	CC02113	100.00
CC60513	100.00	SEW03804	100.00
CC61313	90.03	CC66813	100.00
CC61963	111.22	CC06363	100.00
CC66913	104.34	CC35563	100.00
CC71601	46.74	CC30613	100.00
MAY00096	173.83	CC13813	100.00
BTB10042	86.68	CC00913	100.00
CC42545	67.37	CC25513	100.00
CC75719	87.64	MO01167	100.00
MO08563	71.22	CC29313	100.00
AC39789	66.24	KM05166	100.00
AC39921	109.41	ML00416	100.00

AC12261	194.17	SP01447	100.00
AC10026	62.83	CC21913	100.00
AC29626	99.93	BTBG00035	100.00
AC33758	76.23	MO07746	100.00
AC34683	186.01	BTB13426	100.00
AC37311	65.37	TL00150	100.00
AC42043	73.12	BTB00733	100.00
AC42399	49.64	AC39858	100.00
CC49713	120.65	SEW05227	100.00
CC14713	193.68	CC24413	100.00
AW00189	67.52	BTB10558	100.00
BTB13068	82.04	KM10724	100.00
AC13217	115.21	AC10460	100.00
AC15193	144.53	AC36409	100.00
RH00879	123.93	RH00729	100.00
BTB13009	64.63	CC65813	100.00
BTB07041	137.07	MO01181	100.00
HTS09269	77.79	S01381	100.00
RH01534	124.64	CC56046	100.00
RH01844	126.20	HTS00626	100.00
RJC00373	89.00	CC13546	100.00
SEW04444	121.99	CC58701	100.00
BTB00070	88.97	AC44411	100.00

RJC04046	83.67	**AC33438**	100.00
SEW01483	135.56	**BTB10358**	100.00
HTS01520	134.03	**RH00001**	100.00
AC39841	172.63	**CC40996**	100.00
BTB14320	51.98	**TL00838**	100.00
DP01481	106.01	**BTB06289**	100.00
RF03300	106.77	**KM08985**	100.00
CC00413	130.26	**CC12296**	100.00
GK01607	99.26	**MO08374**	100.00
BTB02424	113.84	**BTB04350**	100.00
CC07114	100.00	**MO08607**	100.00
CC65063	100.00	**10M-523S**	100.00
CC61114	100.00	**AS-5557**	100.00
AC15782	100.00	**FS-2138**	100.00
CD10231	100.00	**PS-4320**	100.00
CD04945	104.71	**PS-3600**	100.00
DSHS01050	83.82	**PS-4439**	100.00
SB01866	165.5	**FS-1321**	100.00
CC29716	118.99	**PS-3424**	100.00
PS-3756	100.00	**1R-1190**	100.00
PS-3417	100.00	**AS-5528**	100.00
SS-3747	100.00		

Appendix 4: Crystallography values (provided by Dr Syed Ahanger)

Table 9.2 Crystal data for bound fragments

Compound	R work	R free	B factor
BTB10042	**20.2**	**23.1**	**60.0**
AC36409	**20.1**	**24.6**	**69.0**
SEW03804	**23.3**	**29.5**	**62.0**