Structural and functional studies of the human mitotic kinesins MPP1 and KifC1, kinase DYRK2, and antibody A33 Fab and their potential as targets for development of cancer chemotherapy drugs

Jiazhi Tang

A thesis submitted in accordance with the requirements of UCL School of Pharmacy for the degree of Doctor of Philosophy

30th June 2022

UCL SCHOOL OF PHARMACY
29-39 Brunswick Square
London WC1N 1AX

Plagiarism statement

This thesis describes research carried out at the University College London School of Pharmacy between November 2017 and February 2022 under the supervision of Prof. Frank Kozielski. I, Jiazhi Tang, hereby confirm that the work presented in this thesis is my own. Any ideas, quotations, and paraphrasing from other people's work and publications have been appropriately referenced. I have not violated the UCL School of Pharmacy's policy on plagiarism.

Signature: Date:

Abstract

The work described in this thesis focuses on the early-phase development of anti-cancer treatments based on small therapeutic molecules and therapeutic antibodies, which target key factors, including Kinesin Family Member C1 (KifC1), M-Phase Phosphoprotein 1 (MPP1), Dual Specificity Tyrosine Phosphorylation Regulated Kinase 2 (DYRK2), and human glycoprotein A33 antigen, during the formation and development of cancer.

The first part of the thesis covers investigation of two human mitotic kinesins, KifC1 (also called human spleen, embryo and testes protein, HSET; kinesin-14 family) and MPP1 (kineisn-6 family), which are pivotal in cell division, with an emphasis on early-phase fragment-based drug discovery. Several small-molecule hits were identified and characterised using various assays. Attempts for acquiring the protein crystals of these kinesins are also described.

The second part of the thesis describes the crystallisation and structure determination of a well-developed immunotherapy molecule, human A33 Fab, which targets a key factor during colorectal cancer development. Protein crystals of the human A33 Fab were obtained, and the structure of A33 Fab was determined. A comprehensive analysis of A33 Fab structure was conducted, which formed the basis of research conducted by our collaboration partners.

The last part of the thesis covers a mechanistic investigation of selective inhibitors of human DYRK2, a cancer-related kinase. DYRK2 plays a pivotal role in the development of triple-negative breast cancer (TNBC). The crystal

structures of DYRK2 in complex with three distinct inhibitors were determined. Ligand-protein interactions were analysed to explain the affinity and selectivity of the inhibitors, which may provide guidance for further drug development.

Impact Statement

In this PhD project, different cancer drug targets and immunotherapies were investigated. During the course of this project, several MPP1 inhibitor scaffolds were investigated, a SAXS-based MPP1 model of the motor domain was generated, and an approach for acquiring the MPP1 crystal structure using crystallisation helper was established; furthermore, two A33 Fab crystal structures and three DYRK2-inhibitor crystal structures were acquired. Together, these data and results contribute to enhancing the understanding of cancer drug targets and may foster development of drugs based on these targets. The financial impact includes cost reduction in lead identification and optimisation as well as antibody stability research. A paper on human A33 Fab has been published, and several other publications are under preparation.

Lists of manuscripts

Co-author publication

Cheng Zhang, Nuria Codina, **Jiazhi Tang**, Haoran Yu, Nesrine Chakroun, Frank Kozielski, Paul A. Dalby; Comparison of the pH- and thermally-induced fluctuations of a therapeutic antibody Fab fragment by molecular dynamics simulation. *Computational and Structural Biotechnology Journal.* 2021. **19**: p. 2726-2741. (IF = 7.271)

First author manuscript

Jiazhi Tang, Cheng Zhang, Paul Dalby, Frank Kozielski; The structure of the humanised A33 Fab variant, an immunotherapy candidate for colorectal cancer. *bioRxiv 2022.06.21.497004; doi: https://doi.org/10.1101/2022.06.21.497004*

Acknowledgments

I would like to thank my father Jin Tang, mother Jin Sun, grandfather Zhangyun Tang, and grandmother Guifang Zhang, who offer me endless support and faith. The selfless support of my family has allowed me to present this thesis.

I am grateful to my supervisor Prof. Frank Kozielski, who devoted a lot of time and patience to my research projects. I hope that he gets more time for the academic research of his interest.

Furthermore, I thank Dr. Nikos Pinotsis, Mr. Helal, and Prof. Geoffery Wells for providing support for dealing with technical issues in crystallography and chemical analysis.

I would like to thank my friends and colleagues Jiaqi Lou, Shumeng Ma, Danni Dong, Qian Wang, Shymaa Damfo, and Aisha Alnami who make me proud to be a part of this group.

Finally, I offer heartful thanks to my sponsor, the Chinese Scholarship Council, who believed in my capabilities and offered me generous financial support.

List of Abbreviations

AMPPNP	Adenylyl-imidodiphosphate
ATP	Adenosine triphosphate
CDRs	Complementarity determining regions
CLK	Cdc-like kinase
CRO	Contract research organisation
CPMG	Carr-Purcell-Meiboom-Gill
Cryo-SEM	Cryo-scanning electron microscopy
DH	DYRK homology box
DMSO	Dimethyl sulfoxide
DYRKs	Dual-specificity yak-related kinases
EDTA	Ethylene diamine tetraacetic acid
EGTA	Ethylene glycol tetraacetic acid
eLBOW	Electronic ligand builder and optimisation workbench
FBDD	Fragment-based drug development
FHA	Forkhead-associated domain
FWRs	Framework regions
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HSET	Human spleen, embryo and testes protein
HSF1	Heat shock factor protein 1
IPTG	Isopropyl-D-1-thiogalactopyranoside
KIFs	Kinesin superfamily proteins
LB	Lysogeny broth
LDH	Lactate dehydrogenase

MD	Motor domain
MKLP2	Mitotic Kinesin-Like Protein 2
MM	Multiple myeloma
MPP1	M-phase phosphoprotein 1
MST	Microscale Thermophoresis
MTA	Microtubule-targeting agents
MTs	Microtubules
MTOC	Microtubule-organising centre
NADH	Nicotinamide adenine dinucleotide
NCD	Non-Claret disjunctional
NIS	Neck-linker initial sequences
PDAs	Pseudo-dyad axes
PDB	Protein data bank
PEP	Phosphoenolpyruvate
PH	Pleckstrin homology
PK	Pyruvate kinase
PMSF	Phenylmethylsulfonyl fluoride
PPI	Protein-protein interaction
PX	Phox homology domain
SAXS	Small-angle X-ray scattering
SEC	Size-exclusion chromatography
SMIs	Small molecule inhibitors
SRB	Sulforhodamine B
STD	Saturation-transfer difference
SUMO	Small ubiquitin-like modifier

SDS–PAGE Sodium dodecyl sulphate–polyacrylamide gel electrophoresis

TB Terrific broth

TNBC Triple-negative breast cancer

Contents

1. Introduction

Cancer refers to a group of diseases characterised by the uncontrolled growth of cells in a specific part of the body, with the potential to invade or spread to other parts of the body. In 2020, nearly 10 million cancer-related deaths were reported, making it a leading cause of deaths worldwide [1]. By 2040, this figure is expected to double, with the greatest increase in low- and mid-income countries. Cancer causes approximately 30% of all deaths resulting from non-communicable diseases among adults aged 30–69 years [2]. Lung cancer (accounts for 11.6% of all cases) and female breast cancer (accounts for 11.6% cases) are the most frequently diagnosed cancers, followed by colorectal cancers (10.2%). In terms of mortality, lung cancer is the leading cause of death (18.4% of all deaths), followed by colorectal (9.2%) and stomach cancers (8.2%).

The work described in this thesis focuses on the early phase development of anti-cancer treatments. Small fragments targeting the kinesins KifC1 and MPP1, which are potential cancer drug targets, were screened, characterised, and developed. Furthermore, crystal structures of DYRK2, a key factor in the development of TNBC, in complex with three potent inhibitors were solved for SAR analyses and further development of inhibitors. Finally, the crystal structure of human A33 Fab, a well-developed immunotherapy molecule for colorectal cancer treatment, was solved as a keystone of stability and dynamic analysis for engineering stable protein therapeutics under clinical conditions.

1.1 Kinesins

Kinesin superfamily proteins (KIFs) primarily function as molecular motors using the microtubule (MT) system, which transports various cargos, including

cellular proteins, macromolecules, and organelles [3]. Kinesins convert the chemical energy of adenosine triphosphate (ATP) into motility by driving internal conformational changes [4]. They travel along MTs in both the minus-end and plus-end directions to fulfil their crucial roles in intracellular transport and cell division.

The first kinesin was purified and identified in 1985 as a unique motile protein separated from the cytoplasm of the giant axon of the squid [5, 6]. This protein was named kinesin-1 (also known as conventional kinesin, Kif5B). In the following decades, an increasing number of kinesins with highly conserved motor domains were identified [7]. To date, more than 650 KIFs have been identified in eukaryotes. However, kinesins seem to be absent in bacteria and archaea [8]. Currently, mammalian kinesins can be classified into 16 subfamilies, viz. kinesin-1 to kinesin-14B, by phylogenetic analysis of the motor domain (Figure 1-1) [9]. The human genome harbours 44 different kinesins, and several of these play crucial roles in the development of various diseases [10].

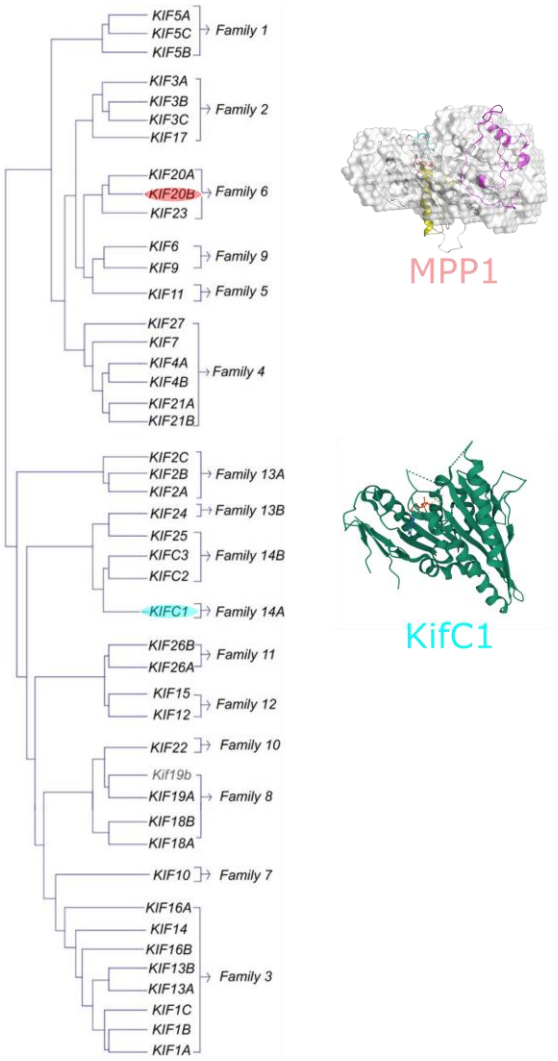


Figure 1-1. KIFs can be classified into 16 families by phylogenetic analysis of their motor domains. Figure adapted from Kalantari, S. et. al [9].

1.1.1 Structural characteristics of kinesins

A typical kinesin comprises three main domains—(1) the motor domain (MD), which converts chemical energy from ATP into motility via conformational changes with the help of the small neck linker region attached to the motor domain; (2) the stalk or coiled-coil domain, which links the motor and tail domains and mediates the oligomerisation of various kinesins; (3) and the tail domain, which recognises and binds to various cargos directly or indirectly through adaptor or scaffold proteins (Figure 1-2 B) [10].

Generally, KIFs can be grouped into three types based on the location of their motor domains — N-terminal kinesins, which have their motor domain in the N-terminal region; M-kinesins, which have a motor domain in the middle of the polypeptide chain; and C-terminal kinesins, which have their motor domain in the carboxyl-terminal region [3]. The N-terminal and C-terminal kinesins drive MT plus- and minus-end motilities, respectively, whereas the M-kinesins depolymerise MTs (Figure 1-2). Aside from conventional structures, KIFs from different families may harbour additional gene-specific domains. For example, KIF1A and KIF1B have a pleckstrin homology (PH) domain, which is involved in intracellular signalling [11, 12]. KIF13B harbours a conserved CAP-Gly, a Gly-rich domain of ~42 residues, which is found in some cytoskeleton-associated proteins [13]. KIF16B has a Phox homology (PX) domain, which is involved in membrane interaction [14], while KIF21A possesses a WD40 repeat domain, which is a short structural motif terminating in a tryptophan-aspartic acid (W-D) dipeptide and is involved in cargo binding [15].

The overall structure of the kinesin motor domain comprises an eight-stranded mixed β-sheet surrounded by three α-helices on each side (Figure 1-3). This arrangement shows some similarities with the arrangement in myosins [16]. To hydrolyse ATP and convert the resulting chemical energy into motility, the motor domain harbours a highly conserved GxxxxGKT motif (Walker motif), namely the P-loop, which interacts with the α- and β-phosphates of bound ADP [17, 18].

Other highly conserved motifs in the motor domain include the switch I (SSRSH) and switch II regions (DLAGSE), which experience continuous conformational changes during the ATP hydrolysis cycle [19]. These switch regions translate

local changes from the motor to larger conformational changes that are responsible for motility [16].

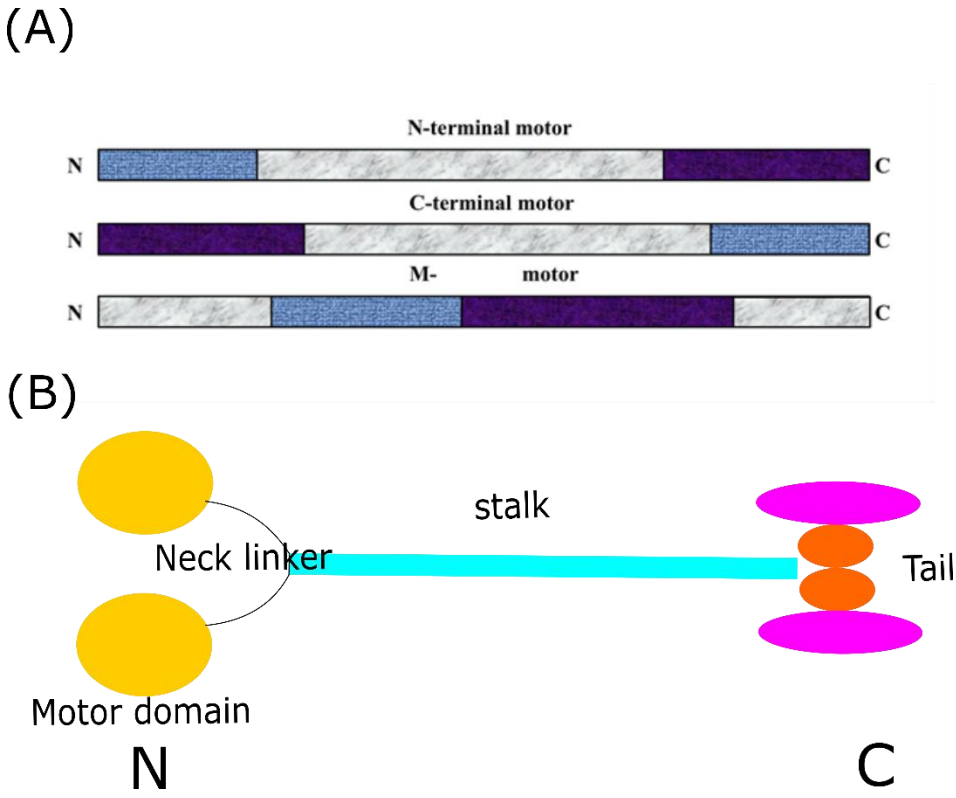


Figure 1-2. (A) Three-domain structure of KIFs. The motor (blue), coiled-coil (grey), and tail (purple) domains are marked. N-terminal kinesins have their motor domain in the N-terminal region, M-type kinesins have a motor domain in the middle of the polypeptide chain, and C-terminal kinesins have their motor domain in the C-terminal region. (B) Schematic of a typical N-terminal motor kinesin dimer. Some kinesins bind directly to cargos through their globular tails (red) as homodimers, whereas others, for example, Kif5B, bind to their cargos as heterotetramers with the assistance of kinesin light chains (purple).

The MT binding site lies on the opposite side of the P-loop. L7/L8, L11, and L12/α5 primarily form the MT binding site that communicates with MTs via electrostatic interactions [20].

The crystal structures of kinesins deposited in the Protein Data Bank (PDB) demonstrate the conserved nature of the motor domain irrespective of its position. A general layout of crucial features of kinesin motor domains is displayed in Figure 1-3 A, B.

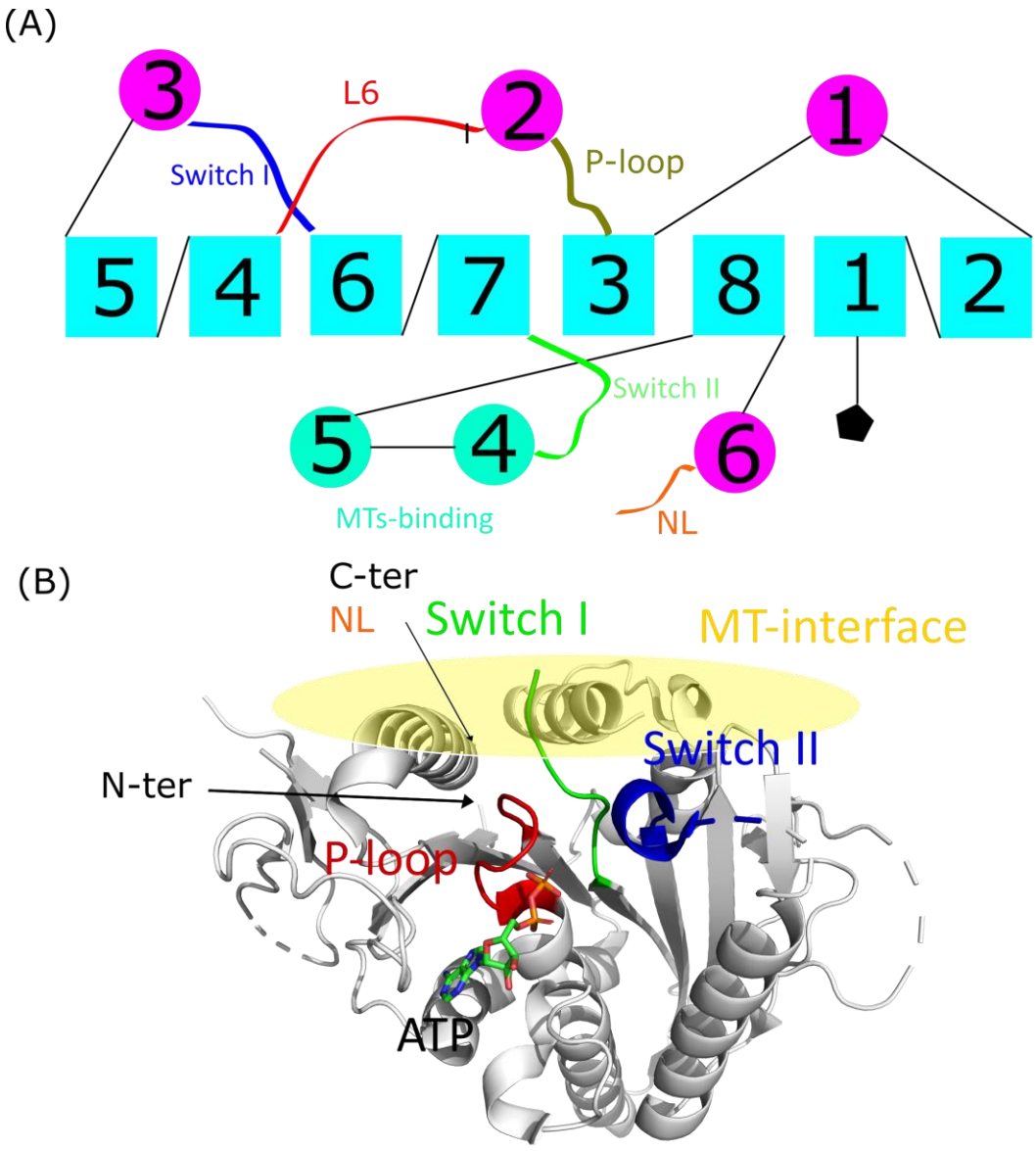


Figure 1-3. (A) Schematic of the major secondary structural elements of the kinesin motor domain showing the P-loop, switch I, switch II, loop-6, MT-binding regions, and neck linker regions (NL). The α-helixes are shown as circles, whereas the β-strands are shown as squares. (B) KifC1 motor domain (PDB entry: 5WDH) marked with the major secondary structural elements, including the P-loop (red), switch I (green), switch II (blue), MT-binding interface (light yellow), N-terminus, C-terminus, and neck linker region (black arrows).

Among conventional kinesins, Kinesin-1, the neck linker that links the motor domain and the stalk, is a short polypeptide that plays a crucial role in movement of the motor along MTs. When kinesin binds MTs, the neck linker is in an 'undocked' state. ADP is released, while ATP binds to the motor. ATP is hydrolysed, which drives the neck linker into a 'docked' state with the assistance of switch I and switch II [21]. Although several kinesins, such as NCD (Non-Claret Disjunctional) and kinesin-14 motors, do not have the neck linker structure, they possess similar structures with the same functions.

The stalk following the neck linker region is responsible for the oligomerisation of kinesins. Homodimerisation is necessary for the functioning of certain kinesins [22]. Furthermore, the stalk domain regulates motor domain activity by facilitating interactions between the motor and tail domains [23]. The typical structure of the stalk domain is a coiled-coil formed because of the presence of heptad repeats [24].

The tail domain links with the stalk domain, which is responsible for the binding of cargos directly or indirectly with the help of adapter/scaffold proteins [25]. In some kinesins, such as kinesin-1 and MCAK, the tail domain also interacts with the motor domain and forms a closed conformation to restrict unwanted ATP use when not activated [26, 27]. The tail domains share low similarity among different kinesins, which explains the specificity of diverse cargos, including vesicles, proteins, and organelles. Furthermore, some tail domains are capable of binding MTs in a nucleotide-independent manner, which induces the formation of large protofilament sheets, suggesting the potential role of tail domains in MT bundling [28].

1.1.2 Functions of Kinesins

Kinesins have diverse functions and can be generally grouped into mitotic kinesins, which are the focus of this research project, and kinesins involved in intracellular transport.

Kinesins are involved in intracellular transport. They play crucial roles in the generation and normal functioning of neurons through participating in the transport of vesicles, protein complexes, and membranous organelles [10]. As a vital protein involved in synapse functioning, kinesin-1 has been implicated in Alzheimer's, Huntington, and Parkinson's diseases [17]. Furthermore, kinesins are necessary for the transport of organelles in somatic cells are responsible for routing the endoplasmic reticulum and Golgi apparatus to their respective destinations in cells [29]. Moreover, kinesins can regulate transport inside cilia and flagella [30, 31].

So far, more than 16 human kinesins have been reported to function in various stages of mitosis by regulating spindle assembly and maintenance, chromosome segregation, mitotic checkpoint control, and cytokinesis [9] (Figure 1-4). Some kinesins, such as Eg5, only function at one particular stage, whereas others, such as members of the KIF4 family, which are implicated in chromosome alignment, anaphase, and cytokinesis, participate in multiple stages of mitosis [32, 33]. Furthermore, several kinesins have redundant functions. For example, in clinical tests of chemotherapeutic molecules targeting Eg5, tumour cells obtained drug resistance [34]. Subsequent investigations indicated that tumour cells use a different endogenous kinesin, Kif15, during bipolar spindle formation to compensate for the activity loss of Eg5 and rescue mitosis [35]. Moreover, several

kinesins play crucial roles in both mitosis and neuronal development, which indicates the complexity of the system [8].

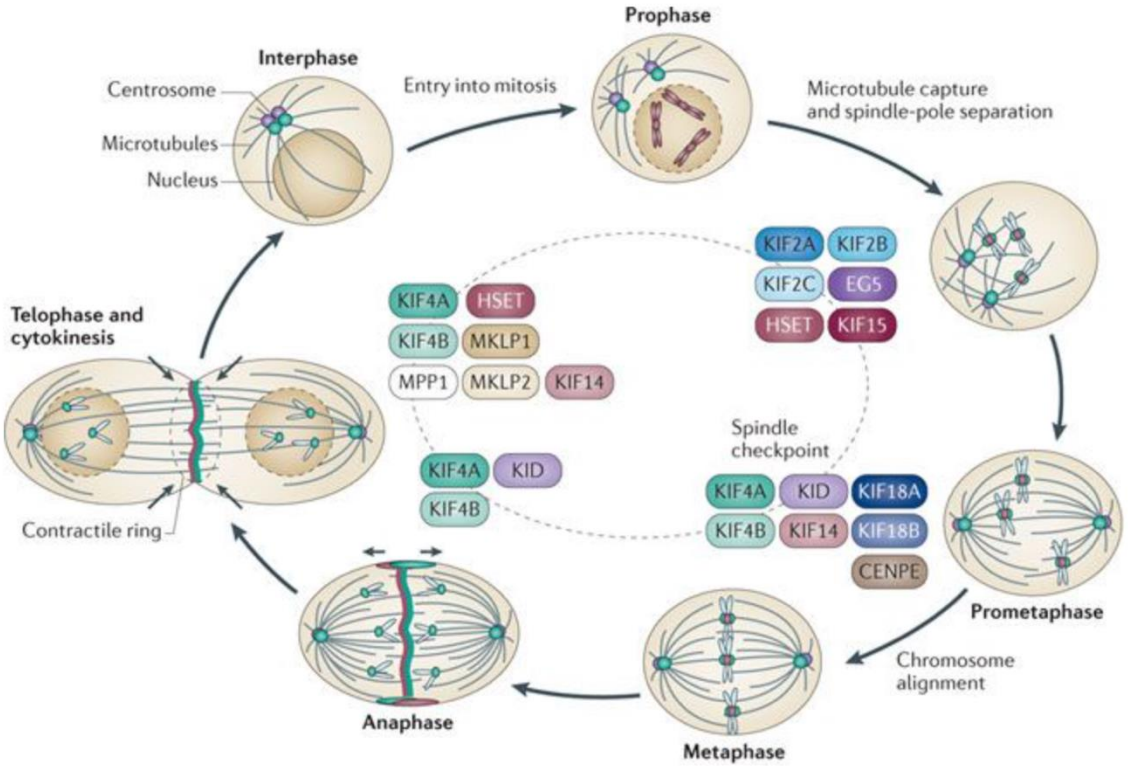


Figure 1-4. Functions of human kinesins involved in different stages of mitosis and cytokinesis. Figure adapted from Rath and Kozielski 2012 [8].

1.1.3 KifC1

1.1.3.1 Structure and functions of KifC1

KifC1, also known as Human Spleen, Embryo and Testes Protein (HSET) or kinesin-like protein, belongs to the kinesin-14 family. Kinesin-14 family members are C-terminal motors that cross-link with MTs during spindle assembly and move towards the MT minus-end [36].

Human KifC1 comprises 673 amino acid residues with three major domains, namely an N-terminal tail domain, a continuous coiled-coil stalk domain, and a C-terminal globular motor domain (Figure 1-5). As a C-terminal kinesin, KifC1 moves towards the minus-end of MTs with a slow gliding rate of approximately 5 μm/min in fluorescent microtubule gliding assays, which is much slower than the motility rate of conventional kinesins (30–48 μm/min) [37].

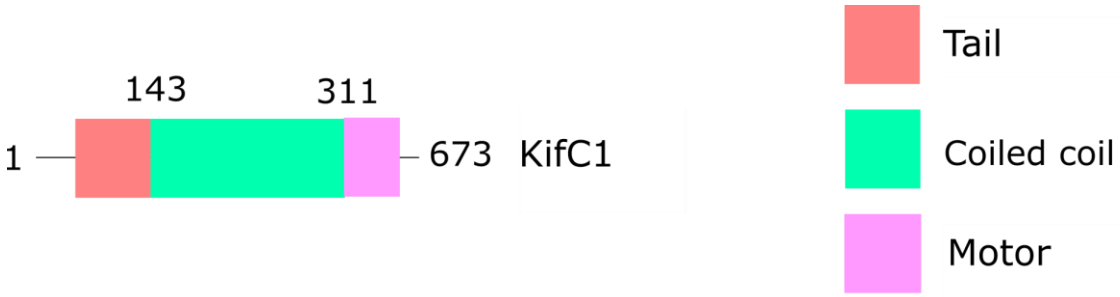


Figure 1-5. Bar diagram of Homo sapiens KifC1, depicting its three-domain structure.

The KifC1 motor domain structure has been determined previously (PDB Entry: 5WDH). Structural analysis indicates that the allosteric L5/α2/α3 inhibitor binding pocket identified in Eg5 (kinesin-5) and CENP-E (kinesin-7) has a reduced volume and is not accessible to current KifC1 inhibitors, including AZ82.

The L5/α2/α3 inhibitor binding pocket is frequently utilised by Eg5 and CENP-E inhibitors such as ispinesib and GSK923295, respectively [38, 39]. However, when rendered as space-filling residues, the KIFC1 L5/α2/α3 pocket is enclosed and not visible when viewed from the top (Figure 1-6 B, C), although a small opening is observed on the lower right side [40]. Recently, a new kinesin inhibitor-binding site composed of helices α4 and α6 was identified in Eg5. An Eg5 inhibitor, BI8, has been reported to bind to this pocket (Figure 1-6 D) [41]. The α4/α6 binding pocket also exists in KifC1 and is spacious enough to harbour potential KifC1 inhibitors (Figure 1-6 E). Unfortunately, as most KifC1 inhibitors bind to the KifC1-MT complex rather than the bare motor domain, no structural data are currently available for any KifC1-inhibitor complex.

In normal cells, KifC1 participates in the establishment and maintenance of mammalian meiotic and mitotic spindles and spindle pole focusing [37]. Furthermore, it plays a crucial role in the formation of MT asters [42]. Moreover, KifC1 transports bare double-stranded DNA, which indicates it potential as a key transporter in gene therapy. KifC1 is highly expressed in the testis, spleen, and ovary and is responsible for acrosome biogenesis and nuclear reshaping, an important procedure in formation of the acrosome, which is located in the front part of the sperm nucleus and is crucial for sperm viability during spermatogenesis [43, 44]. Reduced KifC1 expression is a key factor in a rare male infertility disease characterised by defective acrosome formation and failure in the elongation of sperm heads [44].

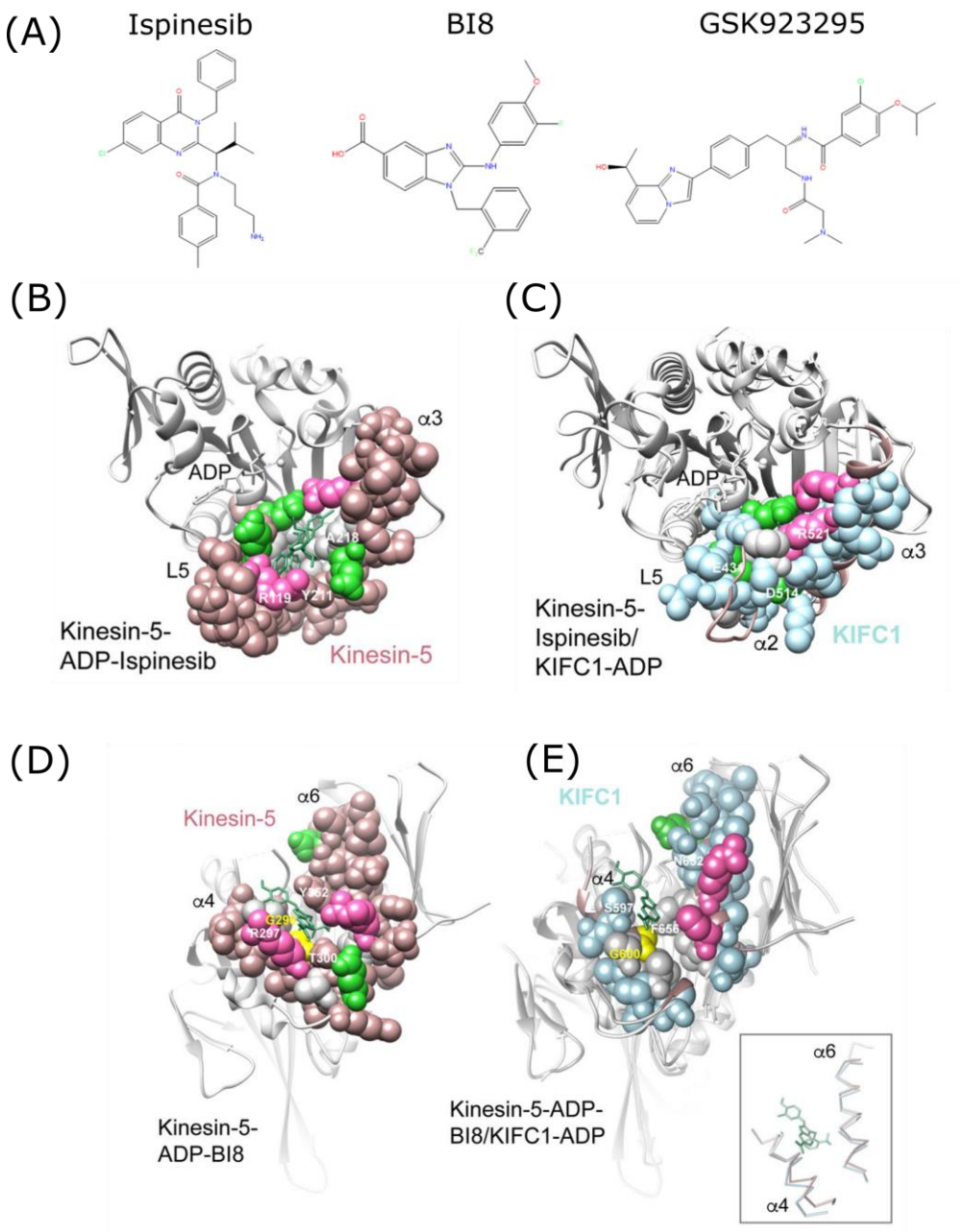


Figure 1-6. (A) Chemical structures of Ispinesib, BI8, and GSK923295. (B) Space-filled L5/α2/α3 pocket in Eg5, bound with ispinesib. Viewed from the top. Hydrophobic residues, grey; basic residues, hot pink; acidic residues, green. (C) Space-filled alignment of KifC1 with Eg5 at the L5/α2/α3 site. (D) Space-filled α4/α6 pocket in Eg5, bound with BI8. (E) Space-filled alignment of KifC1 with Eg5 at the α4/α6 site. Images adapted from Park et al. [40].

1.1.3.2 KifC1 as an anti-cancer drug target

During the prometaphase, the nuclear envelope breaks down when the two centrosomes begin to separate, resulting in the formation of a bipolar spindle. The centrosomes function as the primary microtubule-organising centre (MTOC) throughout mitosis. In normal cells, only two centrosomes emerge before the prometaphase. However, many solid and haematological cancers examined to date have been reported to have supernumerary (more than two) centrosomes [45, 46].

Extra centrosomes may contribute to tumorigenesis by promoting genetic instability through multipolar mitoses or by disrupting cell polarity and asymmetric cell division [45]. However, extra centrosomes can also result in multipolar mitoses and missegregation of genetic data, which finally causes aneuploidy. Although moderate levels of genetic instability contribute to tumorigenesis, the high levels of genetic instability caused by multipolar mitoses are detrimental to cancer cells and eventually cause apoptosis [47].

To avoid multipolar mitoses, several mechanisms, including extrusion of extra centrosomes, inactivation of extra centrosomes, asymmetric segregation, and centrosome clustering [48], are utilised by cancer cells, of which centrosome clustering is the most common and important one. Cancer cells coalesce their extra centrosomes into two groups to form a pseudo-bipolar spindle to avoid multipolar mitoses (Figure 1-7).

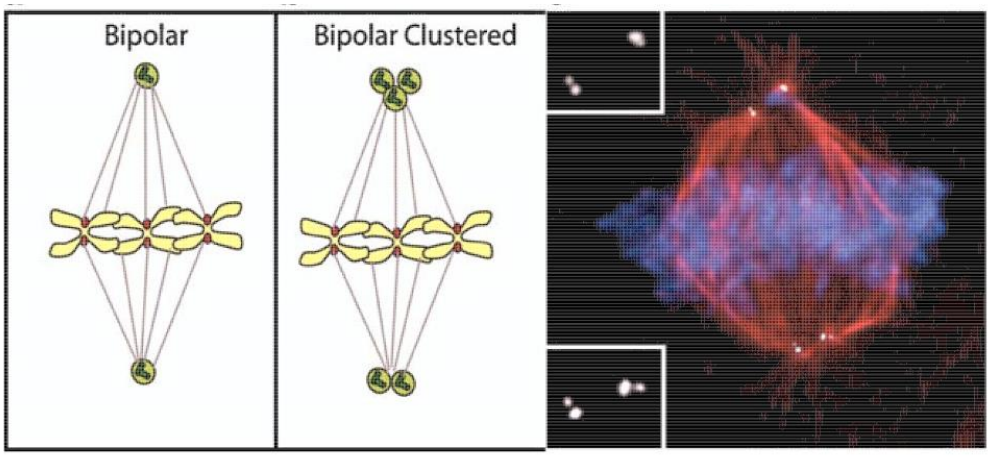


Figure 1-7. Cancer cells cluster extra centrosomes to form a pseudo-bipolar spindle to avoid multipolar mitoses. (left) Normal bipolar spindle. (middle) A pseudo-bipolar spindle with three clustered centrosomes. (right) Image from a CaCo-2 colon cancer cell showing the clustering of supernumerary centrosomes. Cells were stained for centrioles (centrin, green), microtubules (α-tubulin, red) and DNA (blue). Figure adapted from [48].

Various cancer cells with extra centrosomes seem to rely on KifC1 to cluster centrosomes for their survival [49]. KifC1 is reported to bundle MTs at the spindle poles and cluster multiple centrosomes into pseudo-bipolar spindles using their minus-end direct motility. Prolonged prometaphase and metaphase are observed when KifC1 maintains the bipolarity of cancer cells [37].

Although KifC1 is indispensable in some cancer cell lines and functions in reproductive cells, it is not necessary for most somatic cells, which makes it an appealing chemotherapeutic target [48]. To date, three inhibitors of KifC1, namely AZ82, CW069, and SR31527, have been reported (Figure 1-8). However, the exact action mechanisms of these inhibitors remain unknown owing to the difficulties in obtaining structural information. Meanwhile, current inhibitors do not have optimal efficacy for further clinical investigation (*KifC1 discussion section 4.4.3*) [50, 51]. The work described in this thesis focuses on fragment-

based drug development to obtain drug candidates. A set of active fragments was identified via fragment screening, followed by an initial SAR analysis.

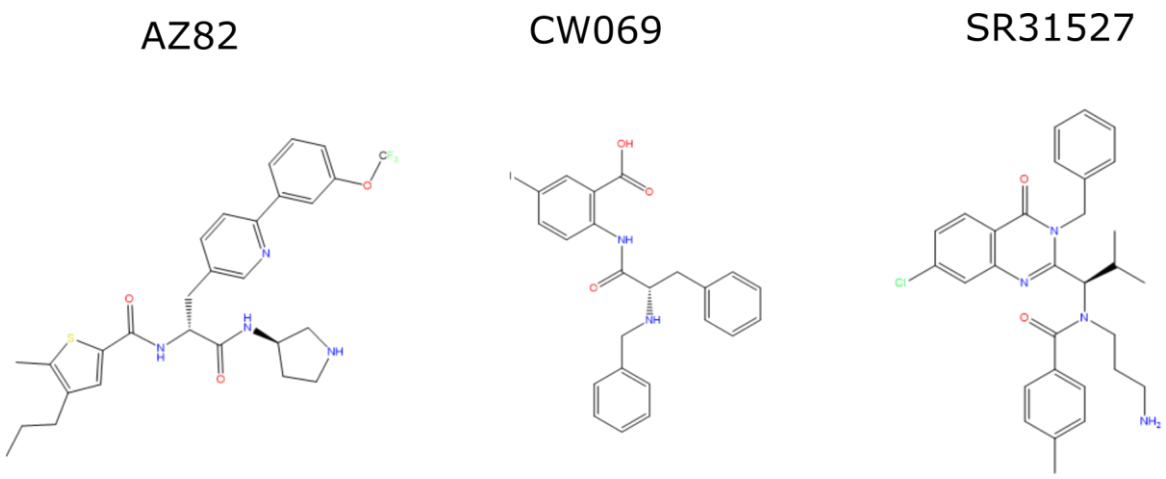


Figure 1-8. Chemical structures of the currently known KifC1 inhibitors.

1.1.4 MPP1

1.1.4.1 Structure and functions of MPP1

Human M-phase phosphoprotein 1 (MPP1), also known as KIF20B, KRMP1 (Kinesin-related M-phase phosphoprotein 1), and MPHOSPH1, is an N-terminal motor protein with 1853 residues, belonging to the kinesin-6 family, and features a large loop L6 insertion in the motor domain of approximately 100 residues [8]. To date, only one X-ray structure of a kinesin-6 motor domain is available, which is named Zen4 (PDB entry: 5X3E) [52], a kinesin from *Caenorhabditis elegans*. Furthermore, four MT-bound *Bos taurus* mitotic kinesin-like protein 2 (MKLP2) structures have been solved using Cryo-EM (MKLP2 with ADP, 5.8 Å, PDB entry: 5ND2; MKLP2 alone, 6.1 Å, PDB entry: 5ND3; MKLP2 with ADP.AIFx, 4.4 Å, PDB entry: 5ND4; and MKLP2 with Adenylyl-imidodiphosphate (AMP-PNP), 7.9 Å, PDB entry: 5ND7) [53].

Three human kinesins, viz. mitotic kinesin-like protein 1 (MKLP1), mitotic kinesin-like protein 2 (MKLP2), and M-phase phosphoprotein 1 (MPP1), belong to the kinesin-6 family (Figure 1-9). All these proteins are N-terminal motors with plus-end motility. Kinesin-6 members are involved in the metaphase-to-anaphase transition, which leads to physical division of the cell into two daughter cells via cytokinesis [54-56].

The difficulties in acquiring kinesin-6 protein crystals are attributed to the high flexibility of the L6 insertion. Based on secondary structure predictions, loop L6 is mostly disordered. Another structural feature of the kinesin-6 family is an extension of 54 amino acids preceding the conserved motor domain. This

extension is also predicted to be structurally disordered, and its function remains unknown.

Following the motor domain is the α-helical coiled-coil stalk, which is connected to the motor domain via the neck linker region. The stalk of MPP1 is significantly longer than that of other kinesins and spans from the 500th residue to the 1520th residue (Figure 1-9) [54]. It is also substantially longer than the stalk region of other kinesin-6 members, including MKLP-1 and MKLP-2. The stalk region mediates MPP1 dimerisation, which is pivotal for the physiological function of MPP1 [57].

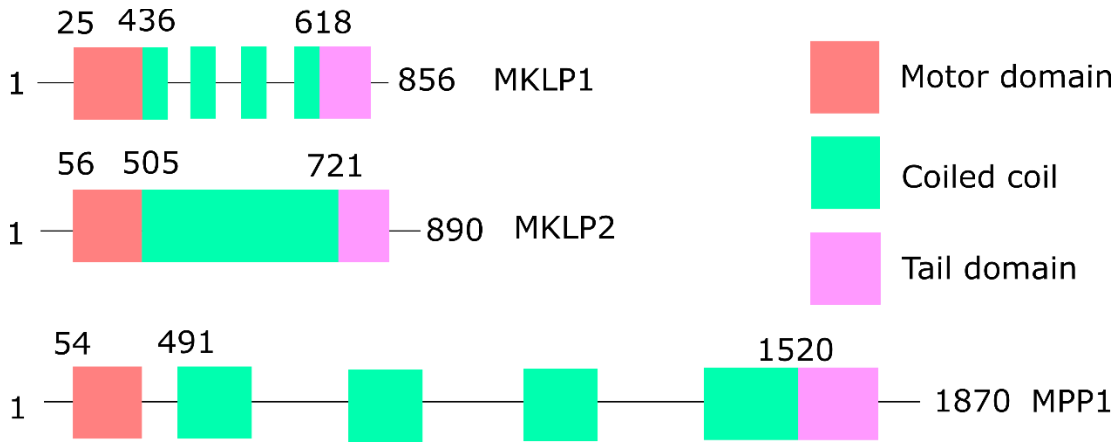


Figure 1-9. Bar diagram of three Homo sapiens kinesin-6 members—MKLP1, MKLP2, and MPP1.

The tail region of MPP1 comprises approximately 200 residues. One significant difference between the tail region of MPP1 and other kinesin-6 proteins is the lack of a leucine zipper, a motif composed of α-helices with periodic repetition of leucine residues [58].

MPP1 has been reported to be regulated *in vivo* by posttranslational modification through phosphorylation of the C-terminal tail domain by cdc2 kinase. This phosphorylation activates MPP1 and induces its binding with mitotic peptidyl-prolyl isomerase (Pin1) via the Pin1 WW domain [57]. Furthermore, casein kinase II is thought to phosphorylate the motor domain of MPP1. However, the importance of these interactions has not been fully investigated.

1.1.4.2 Clinical relevance of MPP1

MPP1 plays a crucial role in cytokinesis. During mitosis, MPP1 disperses throughout the cytoplasm in the metaphase and subsequently concentrates on the midbody [54]. Experiments in HCT116 colon carcinoma cells indicate that MPP1 depletion using RNAi leads to defects in cytokinesis, eventually causing apoptotic cell death through two mechanisms—1. the midbody of RNAi-treated cells survives without abscission and the two physically connected daughter cells undergo apoptosis; 2. the midbody regresses, causing the emergence of binucleated cells that undergo apoptosis [59]. A subsequent study demonstrated that only specific cancer cell lines, such as bladder cancer and large B cell lymphoma, highly rely on MPP1 for cytokinesis [60]. In contrast, no significant defects in cytokinesis are observed upon MPP1 depletion in HeLa cells [61]. Thus, MPP1 is considered a potential target for drug development against bladder cancer.

In a previous study, our group reported depsidones from lichens as natural product inhibitors of MPP1 [59]. One of the inhibitors showed an IC_{50} value of 9.9 ± 1.1 μM. In this thesis, initial SAR-based inhibitor development will be discussed based on these previous findings. Meanwhile, alternative approaches,

such as SAXS and αRep-helper-co-crystallisation, to acquire structural information on MPP1 will also be discussed.

1.1.4.3 Crystallisation helper, an alternative approach to obtain MPP1 crystals

Owing to the flexibility and disorder of the loop L6 and N-terminal regions, none of the kinesin-6 family proteins from *Homo sapiens* have been crystallised neither have their structures been solved. Thus, an alternative approach is proposed to crystallise the MPP1 motor domain.

Alpha repeat proteins (αReps) are a library of artificial proteins developed from HEAT repeats involved in protein-protein interactions [62]. αRep libraries contain a wide range of αReps with various combinations of side chains in hypervariable positions as well as different numbers of repeats. In general, each artificial αRep comprises an N-cap, a C-cap, and an internal domain consisting of random repeats of αRep internal motifs. Each repeat carries five highly randomised amino acid positions, which form a hypervariable surface that ensures specific recognition of the target protein (Figure 1-10) [62].

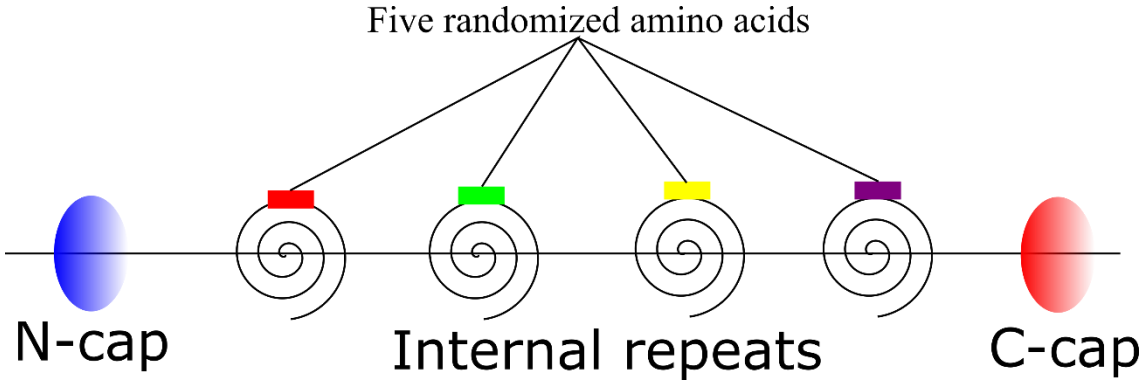


Figure 1-10. Schematic showing αRep structures.

αReps can bind specifically to the target proteins, which are then selected via phage display and phage-ELISA. The selected αReps act as crystallisation helper proteins and bind non-covalently to the target protein, thus enhancing the chances of crystallising the protein complex [63]. We worked in cooperation with the developer of αReps, Prof. Philippe Minard (Institut de Biologie Intégrative de la Cellule, Université Paris Sud, France), to generate optimal helper proteins for human MPP1 and attempted to crystallise the MPP1-aRep complexes.

1.2 Human A33 Fab

1.2.1 Colorectal cancer and current treatments

Colorectal cancer (CRC) is a common cancer type worldwide; it has a high metastasis rate and is the second leading cause of cancer-related deaths [64]. Approximately one to two million new cases of CRC are diagnosed every year, making it the third most common cancer and the fourth most common cause of cancer-related deaths, leading to 700,000 deaths per year, exceeded only by lung, liver, and stomach cancers. By gender, CRC is the second most common cancer in women (9.2%) and the third most common in men (10%) [65]. Worldwide, the probability of suffering from colorectal cancer is about 4–5%. Approximately 20% of newly diagnosed patients are identified with metastases in the liver, lungs, lymph nodes, peritoneum, or soft tissues [66]. Recently, this percentage has dropped significantly owing to the screening methods such as faecal occult blood tests, colonoscopy, and colonography, which have increased the possibility of diagnosing CRC at earlier stages [67].

However, disease control for patients with advanced-stage CRC remains challenging, and intensive treatments such as irinotecan or oxaliplatin-based

chemotherapies, signalling inhibitors, and antibodies are required [68]. The most widely accepted guidelines indicate surgical resection of the metastases with supporting chemotherapy comprising a combination of cytotoxic and targeted biological agents [65]. Along with traditional chemotherapy, immunotherapies, including monoclonal antibodies or proteins targeting vascular endothelial growth factor (VEGF) and epidermal growth receptor (EGFR), have been developed and reported to improve the outcome of CRC [69]. The monoclonal antibody bevacizumab is the most commonly used anti-VEGF agent, which targets circulating VEGF-A. Furthermore, a recombinant fusion protein aflibercept, which blocks VEGF-A, VEGF-B, and placental growth factors, has been developed. These immunotherapies represent optimal treatments when used in combination with cytotoxic agents. In contrast, anti-EGFR treatment can only be used under limited conditions, either as a single agent or in combination with cytotoxic molecules. The most important anti-EGFR agents include monoclonal antibodies such as cetoximab and panitumumab [65].

1.2.2 Human A33 antigen and development of A33 Fab

Human A33 antigen, an M_r 43,000 glycoprotein, has been used as an immunotherapy target. The functional significance of the A33 antigen remains unknown; however, it has been shown to mediate colonic mucosal repair in an animal model of colitis [70]. A33 antigen expression is restricted to the epithelia of the lower gastrointestinal tract and to the carcinoma lesions originating from rectal and colonic mucosa [71]. The A33 antigens expressed in metastatic colorectal cancers share 95% similarity [72]. As A33 expression is uniform across both disease stages and degree of histological differentiation and the antigen is not detectably secreted or shed into the bloodstream, the A33 antigen is considered an optimal immunotherapy target for CRC [70].

The original monoclonal antibody (mAb) against A33, originating from murine IgG2a, has undergone a set of preclinical analyses followed by a series of phase I clinical trials. However, haematological toxicity triggered by the human anti-mouse antibody (HAMA) response was reported as the major limiting factor [73]. To reduce the toxicity as well as extend the half-life of mAb A33, a chimeric A33 antibody was developed by combining the variable region of mAb A33, which is responsible for antigen binding, with the constant region of the human antibody LAY. The chimeric A33 antibody has 75% sequence identity with human IgG1 [73]. Although the human immune system reacts with a mild response to chimeric A33, the human anti-chimeric antibody (HACA) response is still triggered. Failure of the chimeric antibody implied that further humanisation of the chimeric A33 antibody was needed to reduce the immune response. Thus, a fully humanised A33 Fab was developed by grafting only the CDRs from the mAb, which are responsible for antigen binding, onto the variable region framework of the human antibody [74].

Several A33 Fab variants have been developed by Union Chimique Belge Pharma (UCB) for treating colorectal cancer. These were tested as therapeutic antibodies either in their isolated forms, or chemically cross-linked using trimaleimide, which produces a trivalent Fab fragment that can bind antigens with greater avidity, to further increase druggability [73]. However, none of these particular variants progressed to the market because of toxicity or stability issues. Although several manuscripts have been published on the conformational flexibility and kinetic stability of the A33 Fab mutant H/C226S (named as A33 Fab throughout this article) [75, 76], no crystal structure of any A33 Fab variant is available to date. This PhD thesis reports the structure of the fully humanised A33 Fab H/C226S. In the H/C226S variant, cysteine at the 226th position of the heavy chain is mutated to serine for eliminating intermolecular dimerisation [75]. The

A33 Fab structure is either used for structural analysis or as a key model for protein stability investigation. This work has been published as 'Comparison of the pH- and thermally-induced fluctuations of a therapeutic antibody Fab fragment by molecular dynamics simulation' [77]. The manuscript on the structural analysis of A33 Fab is currently in preparation and will be published in future.

1.3 Dual specificity tyrosine-phosphorylation-regulated kinase 2 (DYRK2)

1.3.1 Triple-negative breast carcinoma (TNBC) and treatments under development

Triple-negative breast cancer (TNBC) refers to a phenotypically diverse disease comprising different subtypes with distinct behaviours and responses to therapy [78]. Approximately 15–20% of breast tumours are reported to be triple-negative, with around 7,500 women in the UK being diagnosed each year (data from Cancer Research UK). As a heterogeneous group of tumours, TNBC is defined by the absence of oestrogen (ER), progesterone (PR), and Her2 receptors. As most basal-like cancers are also triple-negative breast cancers and most triple-negative breast cancers (approximately 80%) are basal-like breast cancers, TNBC is often used as a surrogate for identifying the aggressive basal breast cancer subtype [78, 79]. Compared with other breast cancer subtypes, TNBC has the highest mortality rates mainly because of its aggressive behaviour and lack of treatments. According to clinical research, patients with TNBC have an increased likelihood of distant recurrence and higher mortality within five years of diagnosis [80]. The five-year disease-free survival rate is approximately 61% for women diagnosed with TNBC [81].

The therapeutic options for TNBC are limited owing to the lack of response to ER or HER2 targeted treatments [82, 83]. The current NICE guidelines advise that patients diagnosed with early TNBC be treated with a combination of surgery, chemotherapy (anthracyclines and/or taxanes), and radiotherapy [32]. Despite this aggressive approach to treatment, recurrence rates remain high with low survival rates.

Currently, there are three main classes of TNBC treatments under development; these include immunotherapy, androgen receptor (AR)-based therapies, and immune checkpoint inhibitors. The AR-targeted therapies show the greatest potency for luminal AR-like (LAR) and AR-positive TNBC. However, both LAR and AR-positive types account for less than 20% of all TNBCs, indicating that AR-based therapies are likely to benefit only a small subset of patients. Furthermore, a recent phase II clinical trial evaluating the anti-AR agent enzalutamide suggested that further biomarker development is needed to identify the patients that are most likely to benefit from this type of intervention because AR immunohistochemical (IHC) analysis does not correlate with the treatment response [84]. In contrast, most trials of immune checkpoint inhibitors with chemotherapy are still in phase I or II. Although the results are encouraging, with one trial reporting response rates of more than 30%, this approach is likely to be suitable for only a subset of patients, and more reliable biomarkers are required to ensure correct patient selection similar to that for AR-targeted therapies [85]. Thus, new therapeutic targets and biomarkers that offer new mechanistic interventions for TNBC are urgently needed, along with patient stratification approaches to predict likely responders to the new emerging treatments.

1.3.2 Cytoprotective stress response pathway and clinical relevance of DYRK2

In contrast to normal cells, cancer cells are constitutively under cellular stress (i.e., DNA damage/replication stress, proteotoxic stress, mitotic stress, metabolic stress, and oxidative stress). Sustained activation of cytoprotective stress response pathways is therefore pivotal to cancer cell survival [86]. The essential nature of these pathways in cancer cells makes them attractive targets for modulating stress responses and impairing cancer cell survival. To overcome proteotoxic stress inherent to malignant transformation, cancer cells induce a range of adaptive mechanisms, with the master transcription factor heat-shock factor 1 (HSF1)-orchestrated response taking centre stage [87]. Adding to the current knowledge of proteotoxic stress, our collaboration partner Prof. Simon Mackay (Strathclyde University, Glasgow, Scotland) has recently disclosed new evidence demonstrating DYRK2 as an important positive upstream regulator of Heat-Shock Factor 1 (HSF1) (to be published).

As most solid human tumours are aneuploidic, tumour cells are under inherent proteotoxic stress induced by toxic, unfolded proteins. The proteotoxic response pathway helps counter the proteotoxic stress by promoting the proper folding and/or degradation of proteins. The transcription factor HSF1 (Heat-Shock Factor 1), which activates heat shock proteins (HSPs) that function as molecular chaperones, along with the proteasome, which turns over unfolded proteins, plays a key role in the proteotoxic response pathway. HSF1 is activated, translocates to the nucleus, and initiates the transcription of HSPs to guard the proteome against misfolding and aggregation, which further protects cancer cells against proteotoxic stress. The importance of HSF1 to the viability of cancer cells has been demonstrated by the reduced susceptibility of Hsf1-knockout mice to

tumour formation driven either by Ras or p53 mutations or by chemical carcinogens [87]. In addition to playing a pivotal role in countering proteotoxic stress, HSF1 also regulates genes involved in proliferation, survival, and the responses to drugs, which highlights its central role in both cancer progression and chemoresistance [88].

With regards to breast cancer, elevated HSF1 levels are associated with a poor prognosis [89]. Selective therapeutic interventions targeting this pathway are viable because HSF1 depletion strongly reduces the proliferation and viability of breast cancer cells while negligibly affecting normal breast epithelial cells [90]. Meanwhile, proteasome addiction has been identified as a vulnerability of TNBC, suggesting the importance of maintaining protein homeostasis for TNBC cell survival [91]. Accordingly, TNBC cells tend to respond with enhanced sensitivity to modulation of the HSF1 pathway. Although pharmacological and genetic evidence supports the fact that sensitising tumour cells to proteotoxic stress strongly suppresses tumour progression [92], developing a direct HSF1 inhibitor has proved to be challenging. A more practical approach is to target upstream regulatory pathways of HSF1 or HSF1-mediated transcription. Dr. Laureano de la Vega (University of Dundee, Scotland) has disclosed that DYRK2 is an important upstream positive regulator of HSF1. Thus, pharmacological inhibition of this kinase offers a new strategy for suppressing this pro-survival pathway, particularly in TNBC (to be published).

1.3.3 DYRK family and DYRK2 functions

DYRK2 belongs to the dual-specificity tyrosine-phosphorylation-regulated (dual-specificity yak-related) kinase family (DYRKs) and can phosphorylate both

Ser/Thr and Tyr substrates [93]. To date, seven DYRK family members, viz. DYRK1A, DYRK1B, DYRK1C, DYRK2, DYRK3, DYRK4A, and DYRK4B, have been identified. DYRKs rapidly autophosphorylate a critical tyrosine in their conserved activation loop YxY during folding. After autoactivation, they merely function as serine/threonine kinases toward their substrates [94].

DYRK2 acts as both a tumour promoter and suppressor in various cancers. In non-small cell lung cancer, DYRK2 overexpression is associated with better survival, which implies its tumour suppressive role [95]. Furthermore, DYRK2 overexpression in breast cancer and pulmonary adenocarcinoma is reported to be correlated with a higher survival rate [96]. Recent findings indicate that a low expression of DYRK2 is correlated with a poor prognosis in colorectal cancer. However, the significantly increased expression of DYRK2 in oesophageal and lung adenocarcinoma and gastric stromal tumours suggests its potential as an oncogene [93, 97, 98].

Notably, several studies indicate that DYRK2 can promote tumour growth in TNBC [99]. Nevertheless, the underlying mechanisms mediating its function in TNBC are not well characterised and require more extensive investigations.

The *in vivo* and *in vitro* studies conducted by Dr. Laureano de la Vega have demonstrated functional interaction between DYRK2 and HSF1 in TNBC cells. These studies indicated that proteotoxic stress stabilises DYRK2, which in turn phosphorylates and activates HSF1. Furthermore, DYRK2 depletion/inhibition impairs HSF1 activity and sensitises TNBC cell lines to proteotoxic stress. Furthermore, DYRK2 depletion reduces TNBC growth in mouse xenografts in

an HSF1-dependent manner. Finally, high protein levels of DYRK2 are associated with poor prognostic factors in clinical breast cancer samples.

Depending on the connection between DYRK2 and HSF1, selective DYRK inhibitors targeting the DYRK2-HSF1 axis could represent new opportunities to improve TNBC treatment. Targeting DYRK2 using small molecules is hypothesised to validate this target as a potential new strategy to treat patients with TNBC, either as a mono- or combined therapy.

1.3.4 DYRK2 inhibitors under development

As there are seven DYRKs, a highly selective inhibitor that does not affect other DYRK family members needs to be developed for DYRK2. Whether the participation of each kinase in different pathways will have common or opposite outcomes remains unpredictable. In particular, the role of DYRK1A in tumour development remains unknown to date; similarly to those for DYRK2, both tumour suppressor and promoter activities have been reported for DYRK1A. Moreover, many normal cell types rely on DYRK1A for survival, which implies that inhibition of DYRK1A in healthy tissues should be avoided. Leucettines were one of the earlier groups of compounds developed as dual inhibitors of DYRK1A, DYRK2, and Cdc-Like Kinases (CLKs), but no compounds from this series showed any selectivity for DYRK2 over DYRK1A [100]. EHT5372 is the most potent DYRK2 inhibitor reported to date. However, the selectivity of EHT5372 is problematic, as it is even more potent against DYRK1A/1B. EHT5372 has been reported to inhibit DYRK1A-induced Tau phosphorylation in cell lines and target the G0/G1 transition, which suppresses cancer cells to some extent. However, considering its poor DYRK selectivity, it is impossible to

specify a specific isoform associated with its functional outputs [101]. Therefore, a potent and selective DYRK2 inhibitor is needed to delineate the roles of these different isoforms and validate DYRK2 as a target.

2. Hypothesis and aims

2.1 Kinesins KifC1 and MPP1

Since the discovery of kinesins 36 years ago, more than 650 kinesin members have been identified to date. Among the kinesin superfamily members, the cancer-related roles of Eg5 and CENP-E have been extensively investigated. To date, a phase I clinical trial of the CENP-E inhibitor GSK923295 as refractory cancer treatment and multiple phase II clinical trials of Eg5 inhibitors, such as filanesib, to treat multiple myeloma have been completed. Although various studies have been conducted on the cancer-related roles of KifC1 and MPP1 (and several other kinesins), no KifC1 or MPP1 inhibitors have proceeded to clinical trials. Thus, the hypothesis for MPP1 and KifC1 in this thesis is that these two kinesins may be potential targets for drug development in cancer chemotherapy.

The aims of these projects were as follows:

1. To investigate analogue hits that specifically inhibit KifC1 based on the initial fragment hits identified via NMR.
2. To characterise the KifC1 hits using biochemical ATPase assays and biophysical microscale thermophoresis (MST) and thermal shift assays (TSA).
3. To determine the low-resolution structure of the MPP1 motor domain in solution using Small Angle X-ray Scattering (SAXS). The resulting structural information will give us a glimpse of the overall shape of MPP1.
4. To obtain MPP1 crystals and solve the MPP1 structure using X-ray diffraction.

2.2 DYRK2

Recent findings have revealed that DYRK2 plays a crucial role in TNBC development. However, to date, no selective DYRK2 inhibitor has been developed. Therefore, selective DYRK2 inhibitors are hypothesised to represent a potential treatment for TNBC.

The aims of this project are as follows:

1. To obtain DYRK2-inhibitor crystals and solve their structures using X-ray diffraction to contribute to the development of inhibitors selective for DYRK2 over other DYRK family members.
2. To elucidate the mechanism underlying the selectivity of DYRK2 inhibitors using structural analysis.

2.3 Human A33 Fab

The human A33 transmembrane glycoprotein is a potential tumour-associated antigen expressed in 95% of primary and metastatic colorectal cancers [102]. Various therapeutic A33 antibodies have progressed to clinical trials. However, compared with that of small chemotherapeutic molecules, successful development of therapeutic antibodies critically depends on achieving stability under a wide range of conditions. The hypothesis for the A33 Fab project is that by obtaining the crystal structure of human A33 Fab and performing stability analyses, therapeutic A33 antibody variants that are stable within clinical conditions can be developed.

The aims of this project are as follows:

1. To obtain the human A33 Fab crystal structure.
2. To perform stability analyses based on the solved structure [77].

3. Materials and methods

3.1 Materials

The human KifC1 expression clone was provided by Dr. Catherine Tham. The human MPP1 expression clones were obtained from Dr. Sandeep Talapatra. The human $DYRK2_{73-478}$ expression clone was purchased from Genescript (Piscataway, New Jersey, United States). The αRep expression clones were provided by Dr. Minard Philippe (Institut de Biologie Intégrative de la Cellule, Université Paris Sud). The αRep G8-$MPP1_{54-491}$ fusion protein expression clones (short and long linkers) were purchased from Genescript. *E. coli* strain W3110 containing plasmid pTTOD A33 IGS2 was provided by UCB (Slough, UK). Purified A33 Fab was provided by Dr. Cheng Zhang from the UCL Bioengineering Department. BL21 (DE3), BL21 (DE3) pLysS, and BL21-Codon Plus (DE3) RIPL competent cells were purchased from Merck Millipore (Burlington, Massachusetts, United States). The Qiagen HiSpeed Plasmid Midi Kit and cis-Repressed pQE Kan Vector Set (pQE-80l-Kan, pQE-81l-Kan, and pQE-82l-Kan) were obtained from Qiagen (Hilden, Germany). Kanamycin sulphate was purchased from Bio Basic Canada Inc. Propylene glycol, lysogeny broth (LB) medium, and LB-agar medium were purchased from MP Biomedicals (Irvine, California, United States). Terrific broth (TB) medium was obtained from Melford. 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) was purchased from ACROS ORGANICS (Geel, Belgium). Tris (hydroxymethyl) aminomethane, imidazole, $MgCl_2$, glycerol, dimethyl sulfoxide (DMSO), ethylene glycol tetraacetic acid (EGTA), streptomycin sulphate, NaCl, pyruvate kinase (PK), lactate dehydrogenase (LDH), chloramphenicol, ethylene diamine tetraacetic acid (EDTA), phosphoenolpyruvate (PEP), nicotinamide adenine dinucleotide (NADH), KCl, isopropyl-D-1-thiogalactopyranoside (IPTG), ammonium citrate tribasic, Sigma AST protease inhibitor cocktail tablets, BIS-TRIS propane, and Amicon® Ultra 10k, 30k, and 50k centrifuge filters were

purchased from Sigma Aldrich (St. Louis, Missouri, United States). Ethanol, NaOH, sodium citrate dehydrate, and isopropanol were obtained from Fisher Scientific. Ampicillin sodium salt and citrinin were bought from Cayman Chemical (Ann Arbor, Michigan, United States). The ÄKTA start protein purification system, 5 ml HisTrap FF crude columns, and the HiLoad 16/600 Superdex 200 pg size exclusion column were purchased from GE Healthcare (Chicago, Illinois, United States). NuPAGE® MES SDS Running Buffer (20×), SeeBlue® Plus2 Pre-stained Protein Standard, SimpleBlue™ SafeStain, NuPAGE® Sample Reducing Agent (10×), NuPAGE® LDS Sample Buffer (4×), and NuPAGE®Novex 4–12% Bis-Tris Protein Gels (1.0 mm 10 well) were obtained from Life Technologies. Quick Start™ Bradford 1× Dye Reagent was obtained from Bio-Rad (Hercules, California, United States). Ninety-six-well half-area microplates (white, μclear base, medium binding) were purchased from Greiner Bio-One (Kremsmünster, Austria). SnakeSkin® Dialysis Tubing was purchased from Thermo Scientific (Waltham, Massachusetts, United States). Twenty-four-well Linbro plates, the PEG/Ion™ commercial screen set, Index HT commercial screen set, Crystal Screen HT commercial screen set, and Low Ionic Strength Screen Kit were bought from Hampton Research (Aliso Viejo, California, United States). Circular coverslips (22 mm) were obtained from Jena Bioscience (Jena, Germany). PACT premier™, JCSG-plus™, PGA screen™, Morpheus® commercial screen, and Rubic Buffer Screen sets were obtained from Molecular Dimension (Altamonte Springs, Florida, United States). Tubulin (> 99% pure) from porcine brain and taxol were purchased from Cytoskeleton Inc (Denver, Colorado, United States). The Tecan Sunrise microplate reader was obtained from TECAN (Männedorf, Switzerland). Ascochitine analogues were obtained from Sai life Sciences. Chromenone-3-carboxylic acids and quinolone-3-carboxylic acid analogues were obtained from Dr. Geoff Wells (University College London, London), and three new depsidone analogues were obtained from Dr. Joelle Prunet (Glasgow University, Scotland). The Tycho NT.6 and

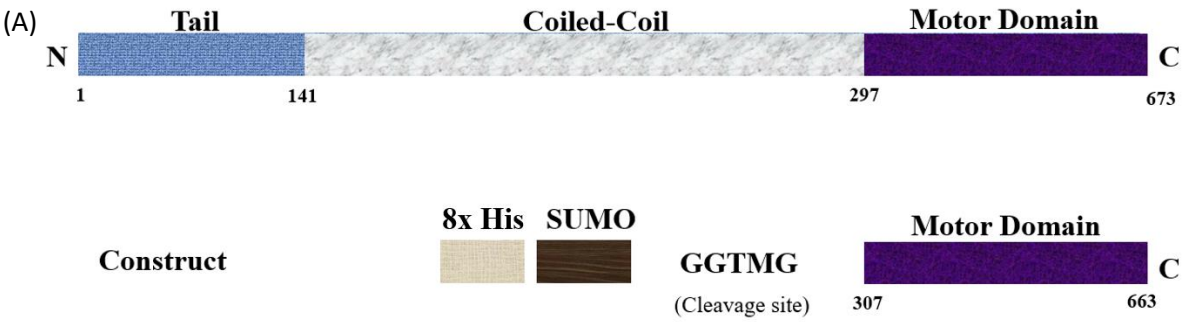
Monolith NT.115 instruments, His-Tag Labeling Kit RED-tris-NTA 2nd Generation, and Monolith premium capillaries were purchased from Nanotemper Technologies (München, Germany). Amicon concentrators (30 kDa) were purchased from Generon Ltd. (Berkshire, UK). Dual Thickness Microloops and UniPucks were obtained from MiTeGen (Ithaca, New York, United States).

3.2 KifC1

3.2.1 Preparation of the KifC1 expression vector

The expression clone containing the coding sequence optimised for expression in *E. coli*, for the amino acid sequence of KifC1$_{307-663}$ inserted into the ppSUMO-2 vector, was purchased from Genscript (Figure 3-1).

(A)



(B)

MGSSHHHHHHHHSSGLVPRGSHMASMSDSEVNQEAKPEVKPEVKP
ETHINLKVSDGSSEIFFKIKKTTPLRRLMEAFAKRQGKEMDSLRFLY
DGIRIQADQTPEDLDMEDNDIIEAHREQIGG

TMGLKGNIRVFCRVRPVLPGEPTPPPGLLLFPSGPGGPSDPPTRLSLS
RSDERRGTLSGAPAPPTRHDFSFDRVFPPGSGQDEVFEEIAMLVQSA
LDGYPVCIFAYGQTGSGKTFTMEGGPGGDPQLEGLIPRALRHLFSVA
QELSGQGWTYSFVASYVEIYNETVRDLLATGTRKGQGGECEIRRAG
PGSEELTVTNARYVPVSCEKEVDALLHLARQNRAVARTAQNERSSR
SHSVFQLQISGEHSSRGLQCGAPLSLVDLAGSERLDPGLALGPGERE
RLRETQAINSSLSTLGLVIMALSNKESHVPYRNSKLTYLLQNSLGGS
AKMLMFVNISPLEENVSESLNSLRFASKVNQC

Figure 3-1. (A) Bar diagram of full-length human KifC1 and the KifC1 motor domain used for expression. (B) Protein sequence of the KiC1 expression clone. A SUMO tag with an 8x His-tag is located at the C-terminus of KifC1$_{307-663}$. The SUMO tag is highlighted in blue, and the Ulp1 cleavage site is marked in red. The estimated pI value before and after cleavage is 6.01 and 6.33, respectively. The MW before and after cleavage is 52.8 and 38.8 kDa, respectively.

The purchased plasmid was then amplified in DH5α *E. coli* cells and extracted using the Qiagen HiSpeed plasmid Midi kit.

3.2.2 Protein expression and purification

The expression plasmid was transformed into competent BL21 (DE3) *E. coli* cells and incubated overnight at 37 °C on an agar plate supplemented with 50 μg/ml kanamycin.

To express human KifC1, a single colony was picked from the plate and added to 40 ml TB medium supplemented with 50 μg/ml kanamycin. The small culture was then incubated overnight at 37 °C in a shaker, following which it was separately added to six 2 L Erlenmeyer flasks, each containing 1 L TB medium supplemented with 50 μg/ml kanamycin, and incubated in a shaker at 37 °C for 16 h. After incubation, the resulting large cultures were induced with 0.5 mM IPTG and incubated at 20 °C for 24 h. On the next day, the culture was harvested via centrifugation at 4000 rpm (Avanti® J-E centrifuge from Beckman Coulter, rotor: JLA 16.250) and 4 °C for 20 min. The cell pellets were then resuspended in 200 ml Buffer A (50 mM Tris pH 7.8, 300 mM NaCl, 10 mM imidazole and 10% glycerol), containing 1 mM phenylmethylsulphonyl fluoride (PMSF), and frozen in liquid nitrogen. Finally, the resuspended pellets were stored at -80 °C.

3.2.3 Purification of KifC1 for ATPase assays and crystallisation

KifC1 pellets were thawed and sonicated for 10 rounds on ice, each sonication round with 30 s on and 30 s off at 16 μm amplitude, using the Soniprep 150

sonicator. The lysate was centrifuged at 20000 rpm and 4 °C for 90 min (Avanti® J-E centrifuge from Beckman Coulter, rotor: JLA 25.50); the supernatant was then collected and loaded onto a 5 mL HisTrap FF crude column pre-equilibrated with Buffer A. The column was then washed with 50 column volumes (CVs) of Buffer B (50 mM Tris pH 7.8, 300 mM NaCl, 30 mM imidazole, and 5% glycerol). The protein bound to the column was then eluted isocratically using Buffer C (50 mM Tris pH 7.8, 300 mM NaCl, 250 mM imidazole, 1 mM $MgCl_2$, and 5% glycerol) and collected. The collected protein fractions were then dialysed with 1 mM Mg^{2+}ATP, 3 mM DTT, and Ubiquitin-like-specific protease 1 (Ulp1) (1 mg Ulp1 for 100 mg protein) in 1 L Buffer D (50 mM Tris pH 7.8, 300 mM NaCl, 1 mM $MgCl_2$ and 5% glycerol) overnight at 4 °C using SnakeSkin® Dialysis membrane, 10 kDa.

The uncleaved protein and protease were removed using a second Ni-NTA affinity column. The cleaved protein present in the flow-through fractions was qualitatively tested using the Bradford reagent and further verified via SDS-PAGE. Pure fractions were combined. The pooled sample was concentrated to 5 ml with an Amicon Ultra Centrifugal Filter (cut-off, 30 kDa).

The concentrated protein sample was finally subjected to size-exclusion chromatography on a HiLoad 16/600 Superdex 200 pg size exclusive column (GE Healthcare) equilibrated in Gel-filtration Buffer (50 mM Tris pH 7.8, 150 mM NaCl, 1 mM $MgCl_2$, 1 mM DTT, and 5% glycerol). Fractions containing the protein of interest were pooled and concentrated on Amicon Ultra Centrifugal Filters (30 kDa) to a final concentration of approximately 3.8 mg/ml for ATPase assays and approximately 16 mg/ml for crystallisation, aliquoted, frozen in liquid nitrogen, and stored at -80 °C.

3.2.4 Buffer tests using thermal shift assays

The Tycho NT.6 instrument (Nanotemper) was used to identify the best buffer for *in vitro* experiments on KifC1. The fluorescence of intrinsic tryptophan and tyrosine residues was detected at both 350 and 330 nm as a 30 °C/minute temperature ramp was applied from 35–95°C to identify the inflection temperature (T_i) representing the unfolding transition in the structural integrity of a protein.

$KifC1_{307-663}$ (8.4 mg/ml) in Gel-filtration Buffer was diluted 20-fold with buffers from the Rubic buffer Screen Set (Molecular Dimension) and subjected to thermal shift measurement. All experiments were conducted in triplicate, and T_i values were calculated using the Tycho NT.6 software.

3.2.5 Purification of $KifC1_{307-663}$ for microscale thermophoresis (MST) measurement and microtubule (MT)-stimulated ATPase assays

The expression clone described in the previous section 3.2.1 was used to transform competent BL21 (DE3) Gold cells. The expression methods were the same as those described in the previous section, except that Buffer A was replaced with 120 mM potassium phosphate monobasic pH 7.5, 150 mM NaCl, 10 mM imidazole, and 5% glycerol.

KifC1 pellets were thawed and sonicated for 10 rounds on ice, each sonication round with 30 s on and 30 s off at 16 μm amplitude, using the Soniprep 150 sonicator. The lysate was centrifuged at 20000 rpm and 4 °C for 90 min (Avanti®

J-E centrifuge from Beckman Coulter, rotor: JLA 25.50). The supernatant was collected and loaded onto a 5 mL HisTrap FF crude column pre-equilibrated with Buffer A. The column was then washed with 50 CVs of Buffer B (120 mM potassium phosphate monobasic pH 7.5, 150 mM NaCl, 30 mM imidazole, and 5% glycerol). Protein bound to the column was then eluted isocratically with Buffer C (120 mM potassium phosphate monobasic pH 7.5, 150 mM NaCl, 300 mM imidazole, and 5% glycerol) and pooled. The collected sample was concentrated to 5 ml using a 30 kDa Amicon Ultra Centrifugal Filter.

The concentrated protein sample was finally subjected to size-exclusion chromatography on a HiLoad 16/600 Superdex 200 pg size exclusion column (GE Healthcare) equilibrated with Gel-filtration Buffer (120 mM potassium phosphate monobasic pH 7.5, 150 mM NaCl, 1 mM $MgCl_2$, 1 mM DTT, and 5% glycerol). Fractions containing the protein of interest were pooled and concentrated on 30 kDa Amicon Ultra Centrifugal Filters to a final concentration of approximately 2 mg/ml for MST measurement and approximately 3.8 mg/ml for ATPase assays.

3.2.6 Steady-state ATPase assays

Steady-state basal and MT-stimulated ATPase rates were measured using the pyruvate kinase/lactate dehydrogenase linked assay [103]. Kinesins convert ATP to ADP and inorganic phosphate. PEP is then converted to pyruvate by pyruvate kinase (PK) in the presence of ADP. Pyruvate is converted to lactate by LDH. In this process, NADH is converted to NAD+. As NADH absorbs UV light at 340 nm, while NAD+ strongly absorbs at 260 nm, the reduction in A_{340} is proportional to the ATP consumption rate, which represents kinesin activity (Figure 3-2).

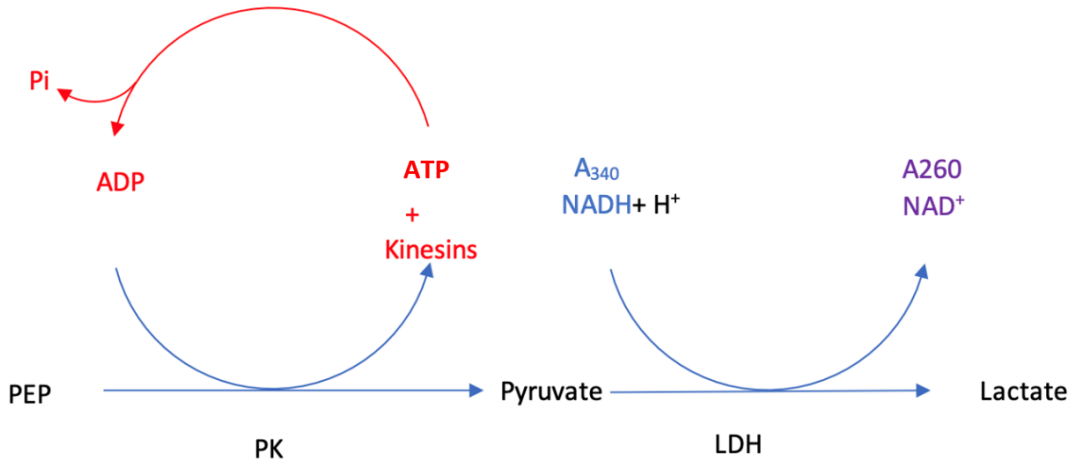


Figure 3-2. Schematic representation of reactions involved in the ATPase assay used to measure the ATPase activities of kinesins.

3.2.7 ATPase characterisation of KifC1$_{307-663}$

The amount of KifC1 was optimised at approximately 2.73 μM for salt dependence and ATPase activity characterisation and at approximately 0.364 μM for MT-stimulated ATPase activity measurement. The enzymatic activity of KifC1$_{307-663}$ was measured in the presence of NaCl or KCl to examine the effect of ion strength. ATPase activity was measured, in triplicate, using a salt concentration series of 0, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, and 1100 mM. The basal ATPase activity of KifC1$_{307-663}$ was then characterised in the presence of 350 mM NaCl. ATPase activity was measured, in triplicate, using Mg^{2+}ATP in a gradient concentration series. The MT-dependence of KifC1$_{307-663}$ was investigated using MT concentrations ranging from 0–10 μM, supplemented with 1 mM Mg^{2+}ATP.

3.2.8 Composition of ATPase buffers

A25 Buffer: 25 mM ACES/KOH, pH 6.9, 2 mM magnesium acetate, 2 mM potassium EGTA, 0.1 mM potassium EDTA, and 1 mM 2-mercaptoethanol.

ATPase Buffer: 1 mM Mg^{2+}ATP, 2 mM PEP, 0.25 mM NADH, 3–10 μg/ml PK, and 3 μg/ml LDH, scaled to the required volume with A25 Buffer. The prepared ATPase buffer was aliquoted, frozen in liquid nitrogen, and stored at -80 °C [104].

3.2.9 Measurement of kinetic data

For each measurement, a concentration series of either Mg^{2+}ATP, MTs, salt, or inhibitors was used. The amount of KifC1 was optimised at approximately 2.7 μM for basal and approximately 144 nM for MT-stimulated ATPase assays. The mixture of ATPase Buffer together with increasing concentrations of inhibitor/MTs/salt/ATP was added to each well. The enzyme was added just before the start of each experiment. All data were obtained in triplicate.

Absorbance was read at 340 nm and 25 °C for 30 min using a Tecan Sunrise reader. The rate was calculated using the following formula:

$$\textbf{rate}\ (s^{-1}) = \frac{-\left(\frac{\Delta A}{t}\right) \times 2 \times 10^6}{\varepsilon_{NADH} \times 60 \times c}$$

where (ΔA/t) is the absorbance reduction at 340 nm per min, ε_{NADH} represents the molar extinction coefficient of NADH (6220 ΔA/mol/cm), and c is the molar

concentration of the tested kinesin. The Michaelis–Menten curve was fitted using Kaleidagraph 4.0.3.

3.2.10 Measurement of ATPase activity of KifC1 fragments

Twenty fragment hits of $KifC1_{307-663}$ were identified with fragment screening via NMR using saturation-transfer difference (STD), Carr-Purcell-Meiboom-Gill (CPMG), and Waterlogsy NMR, which were facilitated by the Monash University. Ten of the fragments were then subjected to basal ATPase assays for verification. The fragments were dissolved in DMSO and transferred to 96-well plates in gradient dilution (Table 3-1). Characterisation of the ATPase activity of KifC1 fragments was carried out with 2.7 μM $KifC1_{307-663}$, in triplicate, using 96-well half-area μclear microplates (Greiner Bio-One).

Table 3-1. Fragment stock concentrations and final ATPase activity measurement

Well number	**1**	**2**	**3**	**4**	**5**	**6**
Stock concentration (mM)	100.00	65.00	42.25	27.46	17.85	11.60
Final concentration in measurement (mM)	2.00	1.30	0.85	0.55	0.36	0.23
Well number	**7**	**8**	**9**	**10**	**11**	**12**
Stock concentration (mM)	7.54	4.90	3.19	2.07	1.35	0.00
Final concentration in measurement (mM)	0.15	0.10	0.06	0.04	0.03	0.00

The inhibition curves of fragments 4 and 9 were fitted using an adapted Hill equation as follows:

$$v/v_0 = F_{min} + (F_{max} - F_{min}) /(1 + ([I] / IC_{50})^{\wedge}h) \quad [105]$$

where v is the reaction velocity at different fragment concentrations; v_0 is the control velocity in the absence of the inhibitor; h represents the Hill coefficient (steepness of the curve); *[I]* indicates fragment concentration; IC_{50} indicates the median inhibitory concentration; F_{min} indicates the minimum fractional activity, and F_{max} indicates the maximum fractional activity. The data were fitted and analysed using Kaleidagraph 4.0.3.

The binding specificities of fragments 4 and 9 were tested by measuring the basal inhibition of another kinesin, Eg5, using the same protocol as described previously.

The fragment hits investigated via ATPase and MST experiments (see next sections) were further verified via the measurement of MT-stimulated ATPase activity. Tubulin powder (Cytoskeleton Inc.) was dissolved in G-PEM Buffer (80 mM PIPES pH 6.9, 2 mM $MgCl_2$ and 0.5 mM EGTA) at 100 μM. To polymerise tubulin, 100 μM taxol (Cytoskeleton Inc.) was added and incubated overnight at 37 °C. The resulting MTs were added to ATPase Buffer containing 2 μM taxol, before the measurements, to a final concentration of 4 μM.

The ATPase measurement protocol has been described in previous sections, and 144 nM $KifC1_{307\text{-}663}$ was optimised for the MT-stimulated ATPase assays. A two-fold dilution series was prepared for the fragments (Table 3-2). Only a single measurement was performed for each fragment owing to the high cost of MTs.

Table 3-2. Fragment concentration during MT-stimulated ATPase activity measurement

Well number	**1**	**2**	**3**	**4**	**5**	**6**	**7**
Stock concentration (mM)	100.00	50.00	25.00	12.50	6.25	3.13	0.00
Final concentration in measurement (mM)	2.00	1.00	0.50	0.25	0.13	0.06	0.00

3.2.11 Thermal shift assays in the presence of potential KifC1 fragments

Nine fragments from the previous section were characterised using thermal shift assays (TSA) on the Tycho NT.6 instrument. $KifC1_{307\text{-}663}$ was expressed and purified as described previously. $KifC1_{307\text{-}663}$ was finally concentrated to 8.4 mg/ml in the Assay Buffer (potassium phosphate monobasic pH 7.0, 150 mM NaCl, 5% glycerol, 1 mM DTT, and 1 mM Mg^{2+}ATP). The unfolding curves and inflection temperatures (T_i) were calculated using Tycho NT.6 software.

The nine fragments were measured at 0.625, 1.25, 2.5, and 5.0 mM in triplicate in the presence of 8 μM KifC1 using the Assay Buffer. Simultaneously, 1.25%, 2.5%, 5.0%, and 10% DMSO were measured separately with 8 μM KifC1 as controls. Finally, 8 μM KifC1 alone was measured as a blank.

3.2.12 MST measurement of potential KifC1 fragments

The NMR fragment hits were then measured via MST using the Monolith NT.115 system. A His-Tag Labeling Kit RED-tris-NTA 2nd Generation was used to label $KifC1_{307\text{-}663}$ with an 8× His-tag. The ratio of labelling dye to $KifC1_{307\text{-}663}$ was optimised at 1:3. The labelled protein stock solution was prepared by mixing 100

nM labelling dye with 300 nM KifC1$_{307-663}$ at a 1:1 ratio and incubating on ice for 50 min. After incubation, the protein stock was further diluted to 20 nM with the Assay Buffer (sodium phosphate monobasic pH 7.0, 150 mM NaCl, 5% glycerol, 1 mM DTT, 1 mM Mg^{2+}ATP, and 0.05% F-127).

Two-fold serial dilutions of the fragments were prepared (Table 3-3) using the Assay Buffer. For each MST measurement, the serially diluted fragment stock was transferred to PCR tubes. The labelled protein stock was then mixed with the fragment solution to a final concentration of 10 nM and incubated for 20 min on ice. Monolith NT.115 series capillaries (Cat. MO-K022) were used to load the solutions. The excitation power was set at 20% and MST power was set at 'Medium'. The temperature was set at 25 °C throughout the measurements.

Table 3-3. Fragment concentrations used during MST measurements

Tube	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Stock concentration (mM)	20.0000	10.0000	5.0000	2.5000	1.2500	0.6250	0.3125	0.1563
Final concentration (mM)	10.0000	5.0000	2.5000	1.2500	0.6250	0.3125	0.1563	0.0781
Tube	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**
Stock concentration (mM)	0.0781	0.0391	0.0195	0.0098	0.0049	0.0024	0.0012	0
Final concentration (mM)	0.0391	0.0196	0.0098	0.0049	0.0024	0.0012	0.0006	0

The MO.Control Software automatically performs several checks, including variations in fluorescence intensity, adsorption to surfaces, sample aggregation, and photobleaching rate changes. If the data pass all these quality checks, the

MO.Control Software performs an initial analysis to determine whether a K_d value of the interaction can be determined. If the K_d of the fragment could be determined, more comprehensive analyses were carried out using the MO.Affinity Analysis software. The analysis time was set to 15 s. The triplicate data were then fitted with the K_d model to obtain dose-response curves that could be normalised to ΔFNorm.

After MST measurement of the initial NMR hits, nineteen KifC1 analogues purchased from eMolecules were measured. The analogues were first subjected to the binding check. A target-only group, composed of 10 nM $KifC1_{307-663}$ in Assay Buffer, and a complex group, composed of 10 nM $KifC1_{307-663}$ with 10 mM analogue, were measured to identify whether the analogue showed binding affinity. Nine analogues were identified as potential hits from the binding check and further measured using MST to identify K_d values using the same protocol.

3.2.13 Crystallisation trials

Crystallisation trials of $KifC1_{307-663}$ were set up. The Mosquito® crystallisation robot was used to set up sitting drops using five commercial screens, viz. PACT premier™, PEG/Ion™, JCSG-plus™, PGA screen™, and Morpheus®. Each drop was set up by mixing 90 nl each of KifC1 and reservoir solution with 1 mM Mg^{2+}ATP. The commercial screens were set up at both 4 and 18 °C. Crystallisation drops were prepared at both 28 and 16 mg/ml.

The published crystallisation conditions for KifC1 were also tested [40]. The published reservoir contains 3.5 M sodium chloride, 0.1 M bis-tris propane, and

5% glycerol, pH 7.5. The drop was set up at 18 °C by mixing 16 mg/ml KifC1 with the reservoir at a 1:1 ratio. Based on the published reports, a set of varying conditions was screened (Supplementary materials 7.5). Two drops were set up by mixing 16 mg/ml and 28 mg/ml KifC1 with reservoirs at a 1:1 ratio in each well. The crystallisation screens were repeated at both 4 °C and 18 °C.

3.3 Human MPP1

3.3.1 Preparation of three MPP1 expression vectors

Three MPP1 expression clones, codon-optimised for expression in *E. coli*, were designed and then synthesized by Genscript. $MPP1_{1\text{-}491}$, $MPP1_{54\text{-}491}$, and $MPP1_{54\text{-}491\ \Delta L6}$ (loop L6 deleted) sequences were inserted into the ppSUMO-2 expression vector.

3.3.2 Transformation and expression of MPP1 constructs

All three MPP1 constructs were expressed using the same protocol. The expression plasmid was cloned into competent BL21 (DE3) pLysS Gold cells and incubated overnight at 37 °C on an agar plate supplemented with 50 μg/ml kanamycin and 34 μg/ml chloramphenicol.

To express MPP1, a single colony was picked from the plate and added to 40 ml TB medium supplemented with 50 μg/ml kanamycin and 34 μg/ml chloramphenicol. A small culture was then incubated overnight at 37 °C in a shaker. Afterward, the small culture was separately added into twelve 2 L

Erlenmeyer flasks, each containing 1 L TB medium supplemented with 50 μg/ml kanamycin, and incubated in a shaker at 37 °C for 16 h. After incubation, the large culture was induced with 0.5 mM IPTG and incubated at 20 °C for 24 h. On the next day, the culture was harvested via centrifugation at 4000 rpm (Avanti® J-E centrifuge from Beckman Coulter, rotor: JLA 16.250) and 4 °C for 20 min. The cell pellets were then resuspended in 200 ml Lysis Buffer (described in the next section) and frozen in liquid nitrogen. Finally, the resuspended pellets were stored at -80 °C.

3.3.3 Purification of $MPP1_{1-491}$, $MPP1_{54-491}$, and $MPP1_{54-491\ \Delta L6}$ without the His-tag

The following buffers were used during the purification process:

$MPP1_{1-491}$

Buffer A (Lysis Buffer): 50 mM Tris pH 8.1, 300 mM NaCl, 10 mM imidazole, and 5% glycerol.

Buffer B (Elution Buffer): 50 mM Tris pH 8.1, 300 mM NaCl, 300 mM imidazole, and 5% glycerol.

Buffer C (Dialysis Buffer): 50 mM Tris pH 8.1, 300 mM NaCl, 1 mM Mg^{2+}ATP, 3 mM DTT (fresh), and 5% glycerol.

Buffer D (Gel filtration Buffer): 50 mM Tris pH 8.1, 200 mM NaCl, 1 mM DTT (fresh), and 5% glycerol.

$MPP1_{54-491}$ and $MPP1_{54-491\ \Delta L6}$

Buffer A (Lysis Buffer): 50 mM PIPES pH 6.5, 300 mM NaCl, 10 mM imidazole, and 5% glycerol.

Buffer B (Elution Buffer): 50 mM PIPES pH 6.5, 300 mM NaCl, 300 mM imidazole, and 5% glycerol.

Buffer C (Dialysis Buffer): 50 mM PIPES pH 6.5, 300 mM NaCl, 1 mM Mg^{2+}ATP, 3 mM DTT (fresh), and 5% glycerol.

Buffer D (Gel filtration Buffer): 50 mM PIPES pH 6.5, 200 mM NaCl, 1 mM DTT (fresh), and 5% glycerol.

MPP1 pellets were thawed and supplemented with 1 mM PMSF and 2 μg/ml DNase I. The mixture was then sonicated for 16 cycles on ice, each sonication round with 30 s on and 30 s off at 16 μm amplitude, using the Soniprep 150 sonicator. The lysate was subsequently centrifuged at 20000 rpm and 4 °C for 90 min (Avanti® J-E centrifuge from Beckman Coulter, rotor: JLA 25.50). The supernatant was collected and loaded onto a 5 ml HisTrap FF crude column pre-equilibrated with Buffer A using the ÄKTA start protein purification system. The column was then washed with 100 CVs of a mixture of Buffer A and Buffer B containing 30 mM imidazole. Protein bound to the column was then eluted with 100% Buffer B. Pure fractions were collected and dialysed with Ulp1 (1 mg Ulp1: 50 mg protein) overnight at 4 °C using Buffer C.

A second nickel-affinity column purification step was then applied. The dialysed protein was loaded onto a 5 mL HisTrap FF crude column pre-equilibrated with Buffer C, followed by washing with 20 CVs of Wash Buffer (a mix of Buffer A and Buffer B containing 30 mM Imidazole). The flow-through and washing flow-through were collected in fractions of 2 ml. The fractions were qualitatively tested

using the Bradford reagent and further verified via SDS-PAGE. The MPP1-containing fractions were pooled and concentrated to 5 ml with an Amicon Ultra Centrifugal Filter (30 kDa).

A final gel filtration step was applied using the ÄKTA start protein purification system, equipped with a HiLoad 16/600 Superdex 200 pg size exclusion column, using Buffer D. Fractions containing the protein of interest were pooled and concentrated using an Amicon Ultra Centrifugal Filter (30 kDa) to a final concentration of approximately 2 mg/ml for the MST and ATPase assays and approximately 18 mg/ml for the crystallisation trials. Protein concentration was determined using the Bradford reagent, supplemented with 1 mM Mg^{2+}ATP, and divided into 100 µl aliquots. The aliquots were frozen in liquid nitrogen and stored at -80 °C.

3.3.4 Purification of $MPP1_{1-491}$, $MPP1_{54-491}$, and $MPP1_{54-491\ \Delta L6}$ with the His-tag

The following buffers were used during the purification process:

$MPP1_{1-491}$

Buffer A (Lysis Buffer): 50 mM HEPES pH 7.8, 300 mM NaCl, 10 mM imidazole, and 5% glycerol.

Buffer B (Elution Buffer): 50 mM HEPES pH 7.8, 300 mM NaCl, 300 mM imidazole, and 5% glycerol.

Buffer C (Gel filtration Buffer): 50 mM HEPES pH 7.8, 200 mM NaCl, and 1 mM DTT (freshly prepared).

$MPP1_{54-491}$ and $MPP1_{54-491\ \Delta L6}$

Buffer A (Lysis Buffer): 50 mM PIPES pH 6.5, 300 mM NaCl, 10 mM imidazole, and 5% glycerol.

Buffer B (Elution Buffer): 50 mM PIPES pH 6.5, 300 mM NaCl, 300 mM imidazole, and 5% glycerol.

Buffer D (Gel filtration Buffer): 50 mM PIPES pH 6.5, 200 mM NaCl, 1 mM DTT (freshly prepared), and 5% glycerol.

The protocol was the same as described in the previous section, except that the dialysis process was excluded.

3.3.5 Crystallisation trials

Crystallisation trials of $MPP1_{1-491}$ without the His-tag were set up. The Mosquito® crystallisation robot was used to set up sitting drops using five commercial screens, viz. PACT premier™, PEG/Ion™, JCSG-plus™, PGA screen™, and Morpheus®. Each drop was set up by mixing 90 nl $MPP1_{1-491}$ and 90 nl reservoir solution. The commercial screens were set up at both 4 °C and 18 °C. The crystallisation drops were prepared at 15 mg/ml.

3.3.6 Screening of Alpha-Rep (αRep) crystallisation helper proteins

The purified $MPP1_{1-491}$ with a His-tag was sent to Prof. Minard Philippe (Institut de Biologie Intégrative de la Cellule, Université Paris Sud, France) as a target for screening Alpha-Rep (αRep) crystallisation helper proteins. During the αRep screening process, the target protein $MPP1_{1-491}$ was immobilised on immuno plates and subjected to three rounds of panning with αRep libraries using phage display. The randomly picked αRep clones from the phage display were further verified and analysed via phage-ELISA. The clone hits with significantly positive phage-ELISA signals were sequenced and expressed in bacteria as isolated proteins for further ELISA verification [63]. Finally, the plasmids of αRep hits were sent to us.

3.3.7 Transformation and expression of αRep proteins

The expression plasmids were cloned into competent BL21 (DE3) Gold cells, which were then plated on agar supplemented with 100 μg/ml ampicillin and incubated overnight at 37 °C. To express αReps, a single colony was picked from the plate and added to 40 ml TB medium supplemented with 100 μg/ml ampicillin. The small culture was then incubated overnight at 37 °C in a shaker, following which it was added to 1 L TB medium supplemented with 100 μg/ml ampicillin and incubated at 37 °C for 8 h in a shaker. After incubation, the large culture was induced with 0.5 mM IPTG and incubated at 37 °C for 5 h. The culture was harvested via centrifugation at 4000 rpm (Avanti® J-E centrifuge from Beckman Coulter, rotor: JLA 16.250) and 4 °C for 20 min. The cell pellets were then resuspended with 20 ml Lysis Buffer (described in the next section)

and frozen in liquid nitrogen. Finally, the resuspended pellets were stored at -80 °C.

3.3.8 Purification of αReps

Lysis Buffer: 20 mM Tris pH 7.5, 200 mM NaCl, and one tablet of sigma AST protease inhibitor cocktail for every 100 ml.

Buffer A: 100 mM Tris pH 7.5, 150 mM NaCl, and 1 mM EDTA.

Buffer B: 100 mM Tris pH 7.5, 150 mM NaCl, 1 mM EDTA, and 2.5 mM d-desthiobiotine.

The αRep pellets were thawed and sonicated for 16 rounds at 4 °C on ice, each sonication round with 30 s on and 30 s off at 16 μm amplitude, using the Soniprep 150 sonicator. The lysate was then centrifuged at 20000 rpm for 30 min at 4 °C (Avanti® J-E centrifuge from Beckman Coulter, rotor: JLA 25.50). The supernatant was collected and loaded onto a 5 ml Strep Trap HP column pre-equilibrated with Lysis Buffer using the ÄKTA start protein purification system. The column was then washed with 10 CVs of Buffer A. Protein bound to the column was then eluted with 100% Buffer B.

The eluted protein was then subjected to gel filtration chromatography using a HiLoad 16/600 Superdex 200 pg size exclusion column. Lysis Buffer was used to pre-equilibrate the column and elute the protein. The eluate was collected in 4 ml fractions at 4 °C. The pooled sample was then concentrated to approximately 10 mg/ml with an Amicon Ultra Centrifugal Filter (10 kDa) and qualitatively verified using SDS-PAGE. The sample was quantified using the Bradford

reagent, divided into 100 µl aliquots of 15 mg/ml, frozen in liquid nitrogen, and stored at -80 °C.

3.3.9 MST measurement of the protein-protein interactions (PPI) between αReps and MPP1 constructs

To verify the PPI between αReps and MPP1 constructs, MST measurement was conducted using the Monolith NT.115 instrument. The three MPP1 proteins, $MPP1_{1-491}$, $MPP1_{54-491}$, and $MPP1_{54-491\ \Delta L6}$, were labelled with the His-Tag Labeling Kit RED-tris-NTA 2nd Generation. The ratio of labelling dye to the proteins was optimised at 1:2. The proteins were mixed with the labelling dye and incubated on ice for 50 min. After incubation, the protein stocks were further diluted to 20 nM with the assay buffers (same as the gel filtration buffers in the previous section).

Two-fold serial dilutions of αReps were prepared using the αReps lysis buffer, with the highest concentration of 45 µM. For each MST measurement, the serially diluted αRep stock was transferred to PCR tubes. The labelled MPP1 proteins were then mixed with the αRep stock to obtain a final concentration of 10 nM and incubated on ice for 20 min. The solutions were loaded in Monolith NT.115 series capillaries. The excitation power was set at 20%, and the MST power was set at 'Medium' (40%). The temperature was set at 25 °C throughout the measurement. The curve fitting process and K_d calculations were automatically conducted by the MO.Control software as described in the previous section on KifC1. The αReps with the highest binding affinities were selected for subsequent co-crystallisation trials.

3.3.10 Co-crystallisation trials of the $MPP1_{1-491}$-αRep complexes

$MPP1_{1-491}$ crystallisation trials were set up in the presence of two αReps, named E4 and G8, which acted as crystallisation helpers. $MPP1_{1-491}$ was concentrated to 18 mg/ml (311 μM) and mixed with 8.8 mg/ml (374 μM) E4 at a 1:1 ratio. The final concentrations were 150 and 187 μM for $MPP1_{1-491}$ and E4, respectively. The Mosquito® nanodrop crystallisation robot was used to set up sitting drops with seven commercial screens, namely PACT premier™, PEG/Ion™, JCSG-plus™, PGA screen™, Morpheus®, Index HT, and Crystal Screen HT. Each drop was set up by mixing 90 nl each of the $MPP1_{1-491}$-E4 mixture and reservoir solution. Screens were set up at both 4 °C and 18 °C. Initial hits were found in the condition Index F1 (0.2 M L-Proline, 0.1 M HEPES pH 7.5, and 10% w/v PEG3350) at 18 °C.

The same screening procedure was then performed using αRep G8. Initial hits were found in the condition Morpheus® A2 (0.06 M Calcium chloride dihydrate, 0.06 M magnesium chloride hexahydrate, 0.1 M Imidazole, 0.1 M MES monohydrate pH 6.5, 6% w/v PEG20000, and 12% v/v PEG methyl ether 500) at 4 °C.

To acquire optimal crystals for X-ray diffraction experiments, crystallisation conditions that yielded crystals in nanodrops were repeated using 24-well Linbro Plates. The ratios of protein mixtures to reservoir ratios were set at 1:2, 1:1, and 2:1. Sitting drops and hanging drops were set up in separate Linbro plates.

3.3.11 Data collection and processing, structure determination, and refinement

X-ray diffraction data were collected on the Diamond Light Source (DLS, Harwell, UK) beamline I04-1 for αRep E4 and on the European Synchrotron Radiation Facility (ESRF, Grenoble France) beamline ID30A-3 for αRep G8. Data were indexed and integrated with *iMOSFLM* [106]. Data reduction and scaling were accomplished using *SCALA* [107] for αRep G4 and *aimless* [108] for αRep G8 within the *CCP4* suite of programs [108]. The structures were solved using *phenix.phaser* [109] with the αRep iiiA5 molecule as the reference model (PDB entry 6GX7) for αRep G4 and the αRep iE5 molecule as the reference model (PDB entry 6GWC) for αRep G8. Electron-density and difference-density maps, all σ_A-weighted, were inspected, and the model was improved using *Coot* [110].

3.3.12 Construction of $MPP1_{54-491}$- G8 fusion proteins

Two fusion protein expression constructs were designed by linking αRep G8 to the N-terminus of $MPP1_{54-491}$ with a short linker of 10 residues (SGGGGSGGGG) or a long linker of 32 residues (GSAAGSGGASGGGGSGGGGSAGSAAGSGG) [111]. The two codon-optimised constructs were subcloned into the pQE-81L-Kan vector by Genscript.

3.3.13 Expression and purification of G8-$MPP1_{54-491}$ fusion protein (short)

G8-$MPP1_{54-491}$ fusion protein with the short linker was selected for further crystallisation experiments. The purification buffers are described below:

MPP1$_{54-491}$- G8 fusion protein

Buffer A (Lysis Buffer): 50 mM PIPES pH 6.8, 200 mM NaCl, 10 mM imidazole, and 5% glycerol.

Buffer B (Wash Buffer): 50 mM PIPES pH 6.8, 200 mM NaCl, 30 mM imidazole, and 5% glycerol.

Buffer C (Elute Buffer): 50 mM PIPES pH 6.8, 200 mM NaCl, 250 mM imidazole, and 5% glycerol.

Buffer D (Gel filtration Buffer): 50 mM PIPES pH 6.8, 200 mM NaCl, 1 Mm DTT, and 5% glycerol.

Based on small-scale results, MPP1$_{54-491}$- G8 (short) was cloned into competent One Shot™ BL21 Star™ (DE3) Chemically Competent *E. coli* and incubated overnight at 37 °C on an agar plate supplemented with 50 μg/ml kanamycin and 34 μg/ml chloramphenicol.

To express the fusion protein, a single colony was picked from the plate and added to 40 ml TB medium supplemented with 50 μg/ml kanamycin and 34 μg/ml chloramphenicol. A small culture was then incubated overnight at 37 °C in a shaker. Afterward, the small culture was separately added into six 2 L Erlenmeyer flasks, each containing 1 L TB medium supplemented with 50 μg/ml kanamycin, and incubated in a shaker at 37 °C for 16 h. After incubation, the large culture was induced with 0.5 mM IPTG and incubated at 20 °C for 24 h. On the next day, the culture was harvested via centrifugation at 4000 rpm (Avanti® J-E centrifuge from Beckman Coulter, rotor: JLA 16.250) and 4 °C for 20 min. The cell pellets

were then resuspended in 100 ml Lysis Buffer and frozen in liquid nitrogen. Finally, the resuspended pellets were stored at -80 °C.

Fusion protein pellets were thawed and supplemented with 1 mM PMSF. The mixture was then sonicated for 16 cycles on ice, each sonication round with 30 s on and 30 s off at 16 μm amplitude, using the Soniprep 150 sonicator. The lysate was subsequently centrifuged at 20000 rpm and 4 °C for 90 min (Avanti® J-E centrifuge from Beckman Coulter, rotor: JLA 25.50). The supernatant was collected and loaded onto a 5 ml HisTrap FF crude column pre-equilibrated with Buffer A using the ÄKTA start protein purification system. The column was then washed with 50 CVs of Buffer B . Protein bound to the column was then eluted with 100% Buffer C. Pure fractions were collected. The fractions were qualitatively tested using the Bradford reagent and further verified via SDS-PAGE. The fusion protein-containing fractions were pooled and concentrated to 5 ml with an Amicon Ultra Centrifugal Filter (30 kDa).

A final gel filtration step was applied using the ÄKTA start protein purification system, equipped with a HiLoad 16/600 Superdex 200 pg size exclusion column, using Buffer D. Fractions containing the protein of interest were pooled and concentrated using an Amicon Ultra Centrifugal Filter (30 kDa) to a final concentration of approximately 20 mg/ml for the crystallisation trials. Protein concentration was determined using the Bradford reagent, supplemented with 1 mM Mg^{2+}ATP.

3.3.14 Crystallisation trials of G8-MPP1$_{54-491}$ fusion protein (short)

Purified G8-MPP1$_{54-491}$ fusion protein (short) was then subjected to crystallisation trials. The Mosquito® crystallisation robot was used to set up sitting drops using

five commercial screens, viz. PACT premier™, PGA screen™, Index HT, Crystal Screen HT and Morpheus®. Each drop was set up by mixing 80 nl fusion protein and 80 nl reservoir solution. The commercial screens were set up at both 4 °C and 18 °C. The crystallisation drops were prepared at 20 mg/ml.

3.3.15 Measurement of ATPase activity of MPP1 analogues

MPP1 inhibitors, obtained from Sai Life Sciences Ltd, Dr. Joelle Prunet, and Dr. Geoffrey Wells, were analysed using ATPase assays. A commercially available analogue, citrinin, was purchased from Cayman Chemical. A set of 38 inhibitors was synthesised by Mr. Helal Helal from Dr. Geoffrey Wells' group, based on the SAR analysis of ascochitine. Diluted fragment stocks were prepared in 96-well plates by serial dilution (Table 3-4). All small molecules were characterised using ATPase assays in the presence of 710 nM $MPP1_{1\text{-}491}$ construct. ATPase assays were carried out by mixing 96 µl ATPase buffer with 2 µl $MPP1_{1\text{-}491}$ construct ($MPP1_{1\text{-}491}$, $MPP1_{54\text{-}491}$, and $MPP1_{54\text{-}491\ \Delta L6}$) and 2 µl small molecule stocks in 96-well half-area µclear plates. All measurements were conducted in triplicate. Absorbance was read at 340 nm and 25 °C for 30 min using a Tecan Sunrise photometer.

Table 3-4. Small molecule concentrations in stocks and basal MPP1 ATPase activity measurement

Well number	**1**	**2**	**3**	**4**	**5**	**6**
Stock concentration (mM)	10	5	2.5	1.25	0.63	0.31
Final concentration (µM)	200	100	50	25	12.5	6.25
Well number	**7**	**8**	**9**	**10**	**11**	**12**
Stock concentration (mM)	0.16	0.08	0.04	0.02	0.01	0.00
Final concentration (µM)	3.13	1.56	0.78	0.39	0.20	0.00

3.3.16 Measurement of MT-stimulated ATPase activities of MPP1 inhibitors

Small molecules **62**, **89**, **91**, **99**, and **100** were further analysed via the measurement of MT-stimulated ATPase activity. Tubulin was polymerised into MTs before the measurements, as described in the previous section on KifC1. The final concentration of MTs was optimised at 500 nM. For each measurement, a series of fragment concentrations was used. The mixture of ATPase buffer along with the fragments in decreasing concentrations was added to each well. MPP1 was added just before the start of the experiment at a final concentration of 66 nM (Table 3-5).

Table 3-5. Fragment concentrations in stocks and MT-stimulated MPP1 ATPase activity measurement

Well number	**1**	**2**	**3**	**4**	**5**	**6**	**7**
Stock concentration (mM)	10	5	2.5	1.25	0.63	0.31	0
Concentration in measurement (μM)	200	100	50	25	12.5	6.25	0

3.3.17 SAXS measurement of MPP1 constructs

The three MPP1 constructs were purified as previously described and measured on the SEC-SAXS beamline, EMBL, Hamburg. The samples were subjected to size-exclusion chromatography (SEC) before SAXS measurement to ensure high purity and removal of potential aggregates. The compositions of SEC running buffers were the same as those of the gel filtration buffers of the constructs. Data were processed using software within the *ATSAS 3.0 suite* [112]. The *R(g)*, distance distribution, and estimated MW were calculated using *Primus* [113].

Three distinct MPP1 homology models representing $MPP1_{1-491}$, $MPP1_{54-491}$, and $MPP1_{54-491\ \Delta L6}$ were generated by Dr. Nikos Pinotsis (Protein Crystallography and Biophysics, Birkbeck College, London, England) based on the KifC1 and MKLP2 models (PDB entries 5WDH and 5ND2). *Crysol* and *Sreflex* were used to analyse the fit of the homology models to the experimental data [112]. The *ab initio* bead models of three distinct MPP1 constructs were generated using *Dammif* and further integrated and improved in *Dammin* [114, 115]. The MPP1 homology models were finally superimposed onto the *ab initio* bead models using *SASpy* [116].

3.4 Human A33 Fab

3.4.1 Sample preparation

Purified A33 Fab (22 mg/ml in water) was provided by Dr. Cheng Zhang from the UCL Bioengineering Department. The crystallisation stock was prepared by diluting A33 Fab with PIPES buffer (10× stock of 50 mM PIPES pH 7.0 and 100 mM NaCl) and MilliQ water to final concentrations of 10, 15, and 20 mg/ml.

3.4.2 Crystallisation of human A33 Fab

The Low Ionic Strength Screen Kit (Hampton Research) was used to screen A33 Fab for crystallisation conditions. Three hanging drops were set up in each well of 24-well Linbro plates for 10, 15, and 20 mg/ml of A33 Fab. Then, 4 µl A33 Fab solution was mixed with 2 µl buffer reagent and 5 µl precipitant reagent and equilibrated against 1 mL of 24% w/v PEG3350 reservoir, as described in the

Low Ionic Strength Screen handbook (Hampton Research). Two screening experiments were conducted at either 4 or 18 °C using the same protocol. Next, 5 µl cryoprotectant solution containing 1.2× drop composition reagents and 30% glycerol was pipetted on the hanging drops containing A33 Fab crystals. Needle-like A33 Fab crystal clusters were broken manually into small pieces using a steel wire and then fished out with MicroMount cryo-loops. The A33 Fab crystals were immediately frozen in liquid nitrogen and stored in Universal Pucks (MiTeGen).

3.4.3 Data processing and structure determination

X-ray diffraction data were collected on a DLS beamline I04 for both crystal forms at 100 K. *iMosflm* was used to index and integrate the diffraction data [106]. The subsequent data reduction scaling was accomplished using *SCALA* [107] within the CCP4 suite of programs [108]. Molecular replacement (MR) was performed using *phenix.phaser* [109]. For the MR of the triclinic structure, the model comprised a heavy chain from the humanised RK35 antibody [117] (PDB entry 5F3H) and a light chain from anti-ErbB2 Fab2C4 [118] (PDB entry 1L7I). The following hexagonal structure used the refined triclinic model as a model. The triclinic structure contains two copies of Fab complexes, each consisting of one light chain and one heavy chain, while the hexagonal structure contains only one copy of the complex. The structures from MR were then refined using *phenix.refine* [109]. The calculation of R_{free} used 10% of the data. Electron-density and difference-density maps, all σ_A-weighted, were inspected, and the model was improved using *Coot* [110].

The elbow angle of the Fab is the angle between the pseudo-twofold axes between the variable and constant domains. It is calculated using the Pymol (Version

2.3.2) elbow_angle.py script, setting the V_L domain as residues 1–109 and the V_H domain as residues 1–117. The elbow angles were calculated for the *P*1 and $P6_5$ crystal structures, the free and binding states of Certolizumab, along with the trajectories from molecular dynamics based on full-residue *P*1 and $P6_5$ structures.

3.5 Human DYRK2

3.5.1 Expression and purification of $DYRK2_{73-478}$

The expression clone containing the coding sequence for the amino acid sequence of human $DYRK2_{73-478}$ (residues 73–478, NCBI gi number 4503427), inserted into the ppSUMO-2 vector, was purchased from Genscript. The $DYRK2_{73-478}$ plasmid was cloned into competent BL21-Codon Plus (DE3) RIPL cells for expression. The bacterial culture was grown at 37 °C in TB medium, supplemented with 100 mg/L kanamycin and 100 mg/L streptomycin, to an A_{600} of approximately 0.8 and induced overnight with 1 mM IPTG at 20 °C. The harvested cells were resuspended in buffer A (50 mM HEPES pH 7.5, 500 mM NaCl, 5 mM imidazole, 5% glycerol, and 0.5 mM tris(2-carboxyethyl)phosphine (TCEP)) and then subjected to sonication for 16 cycles on ice with 1 mM PMSF added, each sonication cycle with 30 s on and 30 s off at 16 μm amplitude, using the Soniprep 150 sonicator. The clear lysate was centrifuged at 20000 rpm and 4 °C for 1 h (Avanti® J-E centrifuge from Beckman Coulter, rotor: JLA 25.50). The supernatant was collected and loaded onto a 5 mL FF crude column pre-equilibrated with Buffer A and then washed with 50 CVs of buffer A containing 40 mM imidazole. $DYRK2_{73-478}$ was then eluted with buffer B (50 mM HEPES pH 7.5, 500 mM NaCl, 250 mM imidazole, 5% glycerol, and 0.5 mM TCEP). The eluted protein was subjected to protease cleavage using Ulp1 in Buffer C (50

mM HEPES pH 7.5, 500 mM NaCl, 5% glycerol, 1 mM MgATP, and 3 mM DTT) overnight at 4 °C with SnakeSkin Dialysis membrane, 10 kDa. The dialysed protein was loaded onto a pre-equilibrated 5 mL HisTrap FF crude column, followed by washing with 20 CVs of Buffer A. The flow-through was collected in 2 ml fractions. The fractions were qualitatively tested using the Bradford reagent and further verified via SDS-PAGE. Pure fractions were combined. The pooled sample was concentrated to 5 ml with an Amicon Ultra Centrifugal Filter, 30 kDa. The cleaved protein was finally subjected to size-exclusion chromatography on a HiLoad 16/600 Superdex 200 pg column (GE Healthcare) equilibrated with buffer D (25 mM HEPES pH 7.5, 500 mM NaCl, and 5 mM DTT). The final purity of the purified $DYRK2_{73-478}$ was verified using SDS PAGE. The protein was then concentrated to 10 mg/ml with an Amicon ultrafiltration device (Millipore), aliquoted, frozen in liquid nitrogen, and stored at -80 °C.

3.5.2 Crystallisation of the $DYRK2_{73-478}$-inhibitor complexes

The purified $DYRK2_{73-478}$ was incubated with four compounds, namely LMB017, LMB035, CI641, and CI709 (Table 3-7), each at 1 mM for 1 h at 18 °C. A set of crystallisation conditions extracted from recent publications was screened for co-crystallisation of $DYRK2_{73-478}$ with different compounds [99-101, 119]. Sitting drops were set up by mixing 2 µl protein with 1 µl reservoir solution. The crystals appeared after three weeks. The crystallisation conditions are summarised in Table 3-6.

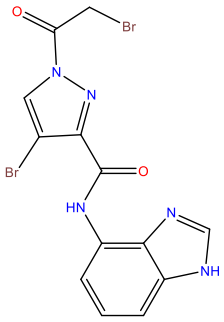
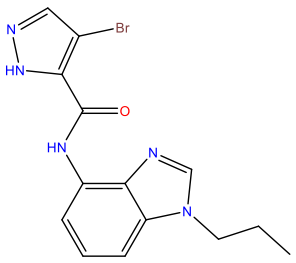
The crystallisation experiments were repeated because of the low-resolution of the diffraction data. The crystals were further dehydrated by transferring sitting

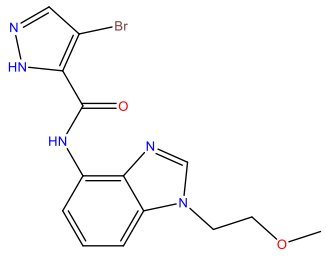
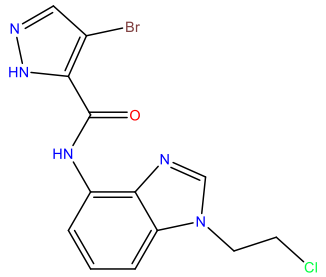
bridges that held the drops from original wells into new reservoir wells containing the same precipitants at 2× concentrations. However, the resolution of the dehydrated crystals did not increase further.

Table 3-6. Crystallisation conditions for the DYRK2-inhibitor complexes

Complex	Buffer system	Additive	Precipitant
DYRK2$_{73-478}$-LMB035	0.1 M citric acid pH 5.0	0.2 M NaCl	20% PEG3350
DYRK2$_{73-478}$-LMB017	0.1 M citric acid pH 5.5	0.2 M NaCl	15% PEG3350
DYRK2$_{73-478}$-CI709	0.1 M citric acid pH 5.5	0.2 M NaCl	15% PEG3350
DYRK2$_{73-478}$-CI641	0.1 M citric acid pH 5.0	0.2 M NaCl	15% PEG3350

Table 3-7. Selective DYRK2$_{73-478}$ inhibitors obtained from Prof. Simon Mackay. Molecular weight, MW; hydrogen-bond donors, HBD; hydrogen-bond acceptors, HBA; polar surface area, PSA.

Name	*Chemical structure*	*MW [DA]*	*cLogP*	*Number of HBA/HBD*	*PSA [Å^2]*
					
CI641	$C_{13}H_9Br_2N_5O_2$	427.056	0	5/2	86.16
					
CI709	$C_{14}H_{14}BrN_5O$	348.204	1.6615	5/2	69.09

					
LMB017	$C_{14}H_{14}BrN_5O_2$	364.203	0.6817	6/2	78.32
					
LMB035	$C_{13}H_{11}BrClN_5O$	368.619	1.5467	5/2	69.09

3.5.3 Data processing and structure determination

X-ray diffraction data for DYRK2$_{73-478}$-LMB035 and DYRK2$_{73-478}$-LMB017 were collected using the ESRF beamline ID30A-1. DYRK2$_{73-478}$-CI709 diffraction data were recorded on beamline ID23-2. Data were indexed and integrated with *XDS* [120]. Data reduction and scaling were performed using *Aimless* within the *CCP4* suite of programs [108]. The structures were solved using *phenix.phaser* [109] with one DYRK2$_{73-478}$ molecule as a search model (PDB entry 3K2L). Calculation of R_{free} used 5% of the data. Electron-density and difference-density maps, all σ_A-weighted, were inspected, and the models were improved using *Coot* [110]. The coordinates and the cif dictionary for the inhibitor molecules were calculated using *electronic Ligand Builder and Optimisation Workbench* (*eLBOW)* within the *phenix* suite of programs.

Further symmetry analyses were carried out using *Xtriage* within the *phenix* suite of programs [109]. Indexing, data reduction, and scaling procedures were

conducted using the same protocols but in the space group $C12_1$ space. The refinements were carried out using *Refmac* within the *CCP4i2* suite of programs with twinning laws applied [108].

4. Results and discussion

4.1 Human MPP1

4.1.1 Efforts to obtain the crystal structure of the human MPP1 motor domain

MPP1 plays a key role in cytokinesis and is reported to be a potential target for drug development against bladder cancer [60]. To date, the crystal structure of MPP1 has not been reported; the structures of MKLP1 and MKLP2 are also unknown. Thus, a valid structure of the MPP1 motor domain is needed for understanding its function and developing MPP1-specific inhibitors. We did not acquire any crystals from crystallisation attempts using the $MPP1_{1\text{-}491}$ construct (*Materials and methods 3.3.5*). Meanwhile, a postdoctoral researcher in our laboratory, Dr. Sandeep Talapatra, conducted crystallisation trials on human, mouse, and rat MPP1 constructs. However, only disordered crystals that did not diffract were acquired from those attempts. Hence, we referred to a new method for co-crystallising the MPP1 motor domain with αRep crystallisation helper proteins [63, 111].

The purities of the proteins in this section are indicated in Figure 4-1. $MPP1_{1\text{-}491}$, $MPP1_{54\text{-}491}$ ΔL6, αRep E7, and αRep H12 were purified to > 95% purity. Meanwhile, $MPP1_{54\text{-}491}$, αRep E4, αRep G4, αRep G5, and αRep G8 were purified to > 80% purity.

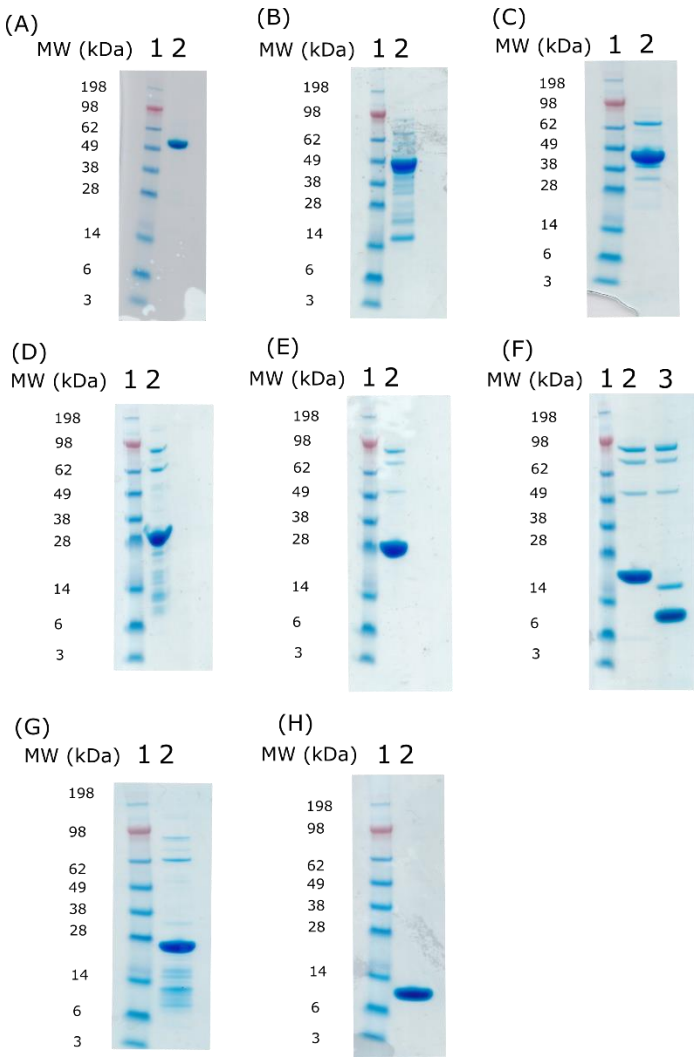


Figure 4-1. SDS-PAGE images of purified proteins. Column 1 represents the protein markers (SeeBlue™ Plus2, Thermo Fisher). (A) Column 2: $MPP1_{1-491}$. The protein band corresponds to the MW of $MPP1_{1-491}$ (56.1 kDa) with a purity of 95%. (B) Column 2: $MPP1_{54-491}$. The protein band corresponds to the MW of $MPP1_{54-491}$ (50.0 kDa) with a purity of approximately 80%. (C) Column 2: $MPP1_{54-491\,\Delta L-6}$. The protein band corresponds to the MW of $MPP1_{54-491\,\Delta L-6}$ (41.4 kDa) with a purity of 85%. (D) Column 2: αRep E4. The protein band corresponds to the MW of αRep E4 (30.3 kDa) with a purity of 80%. (E) Column 2: αRep E7. The protein band corresponds to the MW of αRep E7 (30.2 kDa) with a purity of 80%. (F) Column 2: αRep G4. The protein band corresponds to the MW of αRep G4 (23.5 kDa) with a purity of 85%; Column 3: αRep G5. The protein band corresponds to the MW of αRep G5 (13.3 kDa) with a purity of 80%. (G) Column 2: αRep G8. The protein band corresponds to the MW of αRep G8 (26.6 kDa) with a purity of 85%. (H) Column 2: αRep H12. The protein band corresponds to the MW of αRep H12 (13.4 kDa) with a purity of 95%.

4.1.1.1 Initial screening of αRep binders

$MPP1_{1-491}$ with a His-tag was purified and sent to Dr. Minard Philippe as a target for the high-throughput screening of αRep crystallisation helpers. Eight hits were identified via phage-display and six of these were sequenced and sent back to us. Each artificial αRep is composed of an N-cap, a C-cap, and an internal domain comprising random repeats of αRep internal motifs. Each repeat carries five highly randomised amino acid positions, which form a hypervariable surface, ensuring specific recognition of the target protein. The plasmids of the hits were sent back to us after one year. The molecular weights and pI values of the αRep hits are indicated in Table 4-1, and the protein sequences are listed in Figure 4-2. The αRep hits were expressed on a large scale and purified. The final purities of the αReps used for MST measurement and the crystallisation experiments are indicated in Figure 4-1.

Table 4-1. Molecular weights and pI values of the αRep hits identified via phage display by Dr. Minard Philippe

Name	***MW (kDa)***	***pI***
E4	30.3	7.65
E7	30.2	5.03
G4	23.5	5.64
G5	13.3	8.36
G8	26.6	7.71
H12	13.4	7.09

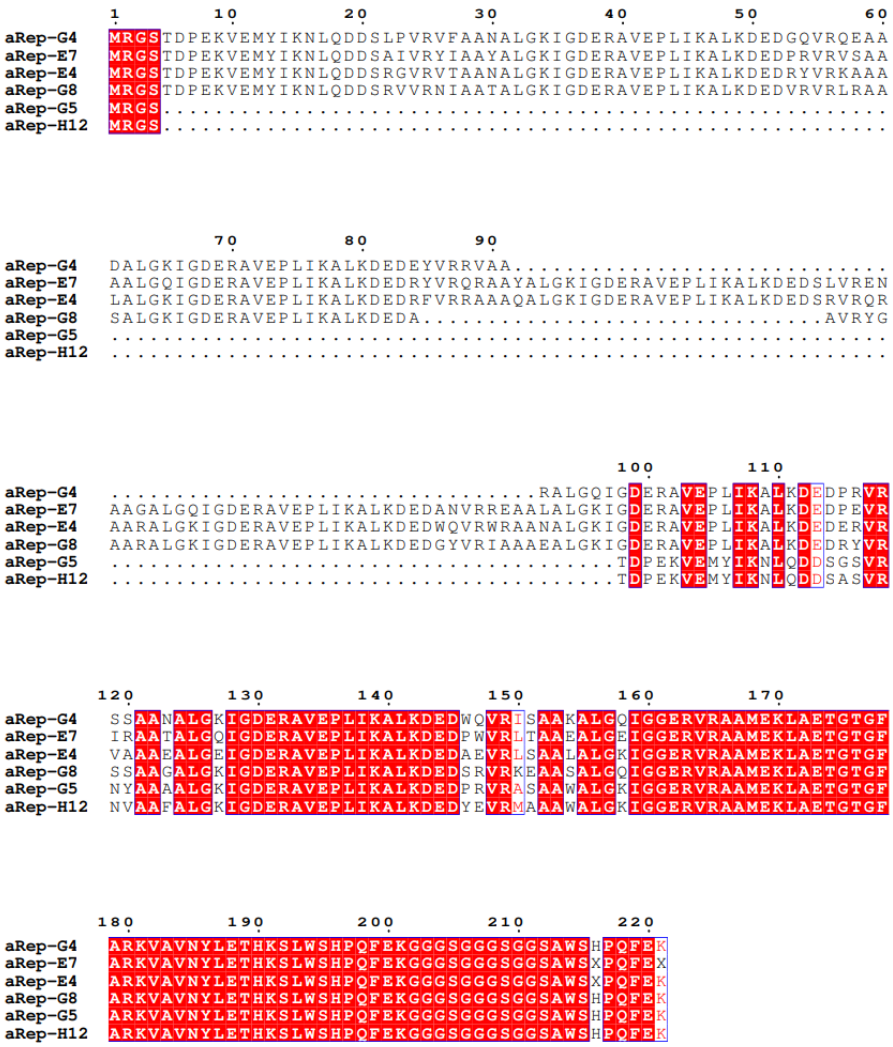


Figure 4-2. Alignment of the αRep hit protein sequences. Six hits with different lengths and repeats of internal motifs were identified via phage-display.

4.1.1.2 MST measurement of the protein-protein interactions (PPI) between αReps and MPP1

To select the optimal αReps for subsequent co-crystallisation trials, the affinities of αReps to three distinct MPP1 constructs were measured using MST (Figure 4-3). According to the MST measurements, αRep H12 showed no binding affinity to any of the three MPP1 constructs. Meanwhile, αRep G5 showed weak binding affinities to $MPP1_{1\text{-}491}$ (K_d = 19.6 μM) and $MPP1_{54\text{-}491}$ (K_d = 42.4 μM) but no affinity to $MPP1_{54\text{-}491}$ ΔL6.

The other four αReps, E4, E7, G4, and G8, displayed decent affinities to all the three MPP1 constructs (Table 4-2). Among the αReps, E4 and G8 showed the highest affinities to all MPP1 constructs. The K_d values of αRep E4 binding to $MPP1_{1-491}$, $MPP1_{54-491}$, and $MPP1_{54-491\ \Delta L-6}$ were 0.69 ± 0.14 μM, 1.30 ± 1.01 μM, and 0.92 ± 0.33 μM, respectively. The K_d values of αRep G8 binding to $MPP1_{1-491}$, $MPP1_{54-491}$, and $MPP1_{54-491\ \Delta L-6}$ were 0.60 ± 0.00 μM, 1.02 ± 0.43 μM, and 0.43 ± 0.15 μM, respectively. Subsequently, the E4 and G8 αReps were selected as crystallisation helpers for the subsequent co-crystallisation trials.

*Table 4-2. K_d values of αReps binding to MPP1 constructs measured using microscale thermophoresis (MST). αRep G5 showed the weakest binding affinities to all the three constructs. In contrast, αReps E4 and G8 demonstrated optimal binding affinities to all the three constructs and were selected for the subsequent co-crystallisation trials. *As only a single measurement was conducted for each test, some K_d confidences could not be identified. *n.b., no detectable binding affinity was observed in the measurement.*

αRep names	***$MPP1_{1-491}$***	***$MPP1_{54-491}$***	***$MPP1_{54-491,\Delta L-6}$***
		MST K_d [μM]	
E4	0.69 ± 0.14	1.30 ± 1.01	0.92 ± 0.33
E7	1.08 ± 0.38	10.77 ± 5.39	10.21 ± 6.57
G4	0.17 ± null*	4.92 ± 4.23	2.45 ± 1.22
G5	19.64 ± null*	42.41 ± 13.40	n.b.
G8	0.60 ± null*	1.02 ± 0.43	0.43 ± 0.15
H12	n.b.	n.b.	n.b.

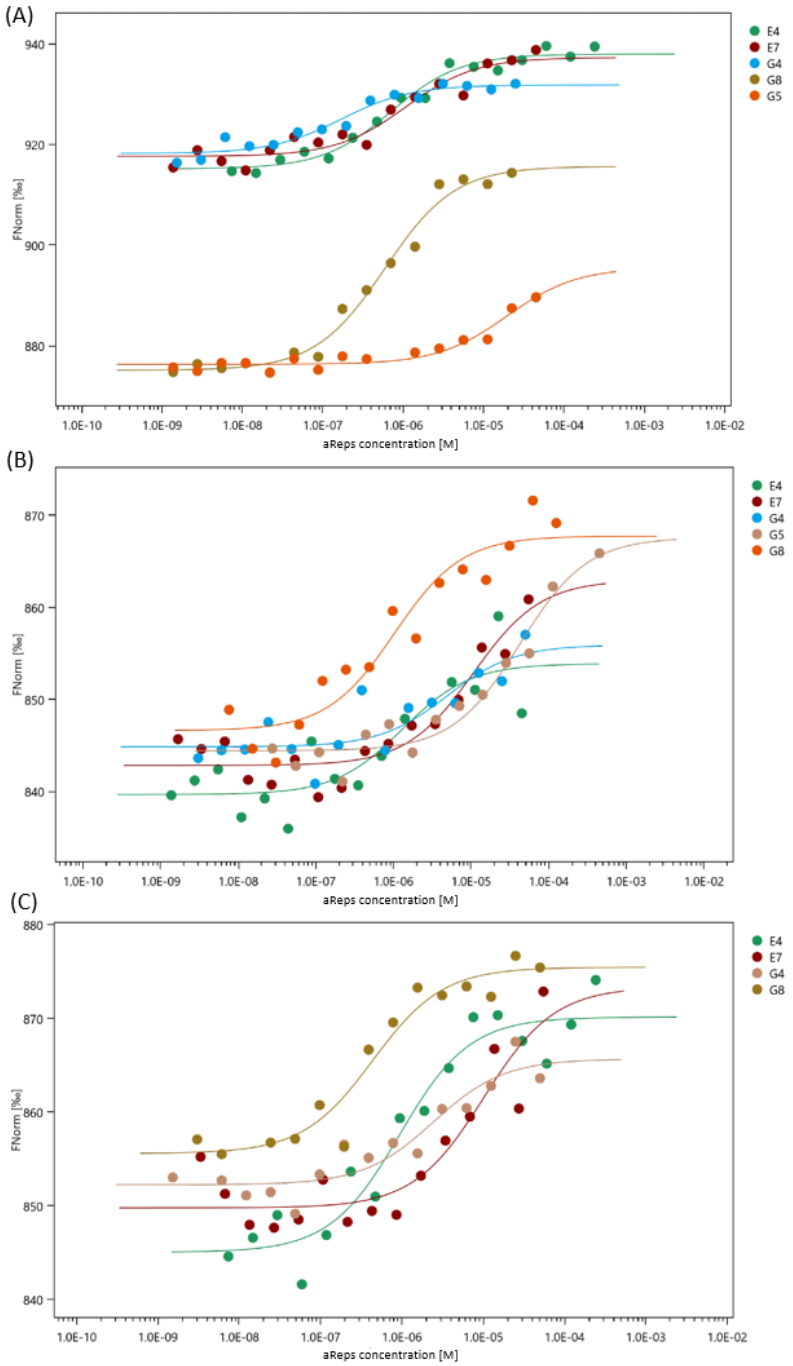


Figure 4-3. Microscale thermophoresis (MST) measurement of the binding affinities of αReps to MPP1 constructs. Concentration–response curves display changes in the ratio between fluorescence after 5 s MST on time and that before activation of the IR laser under different αReps concentrations. (A) Five αReps showed detectable binding affinities to $MPP1_{1\text{-}491}$. The curves were fitted with the K_d model within the MO.Affinity Analysis software. (B) The same five αReps hits showed detectable binding affinities to $MPP1_{54\text{-}491}$. (C) Four αReps showed detectable binding affinities to $MPP1_{54\text{-}491\ \Delta L\text{-}6}$. αRep G5 showed no detectable binding affinity to $MPP1_{54\text{-}491\ \Delta L\text{-}6}$, although it did bind to $MPP1_{1\text{-}491}$ and $MPP1_{54\text{-}491}$.

4.1.1.3 Co-crystallisation trials of MPP1$_{1-491}$ with αReps

The initial crystallisation experiments were carried out by setting up sitting drops with commercial screens, including PACT premier™, PEG/Ion™, JCSG-plus™, PGA screen™, Morpheus®, Index HT, and Crystal Screen HT. For αRep E4, an initial hit was found in the Index F1 well (0.2 M L-Proline, 0.1 M HEPES pH 7.5, 10% w/v PEG3350) at 18 °C. For αRep G8, initial hits were found in the Morpheus® A2 well (0.06 M Calcium chloride dihydrate, 0.06 M Magnesium chloride hexahydrate, 0.1 M Imidazole, 0.1 M MES monohydrate pH 6.5, 6% w/v PEG20000, 12% v/v PEG methyl ether 500) at 4 °C. Conditions that yielded small crystals were repeated in 24-well Linbro plates to generate crystals large enough for X-ray diffraction experiments. Both E4-MPP1$_{1-491}$ and G8-MPP1$_{1-491}$ co-crystallisation trials generated crystals that yielded solvable X-ray diffraction data (Figure 4-4). However, the solved structures contained only αReps, indicating that the crystals were only αRep crystals. The crystallographic statistics are presented in Table 4-3.

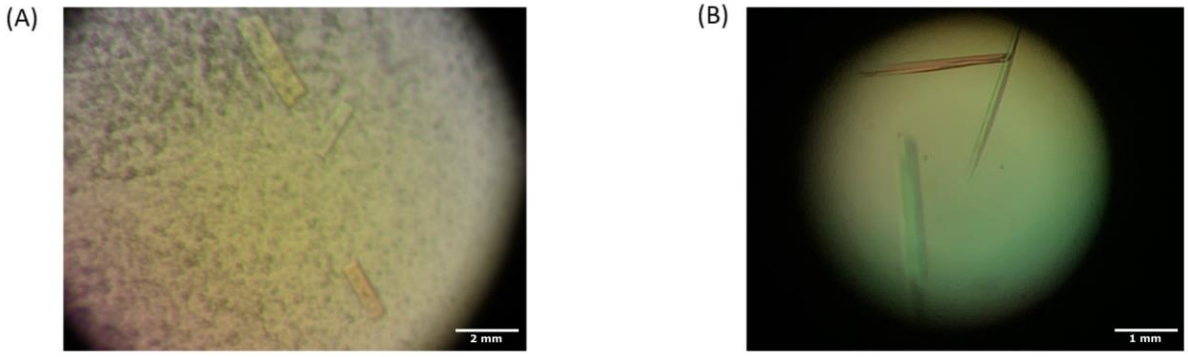


Figure 4-4. (A) αRep E4 crystals and (B) αRep G8 crystals.

*Table 4-3. Data collection and structure refinement statistics for the αReps E4 and G8. Values for the outer shell are given in parentheses. *Water molecules were not identified owing to low resolution.*

Crystal name	*αRep E4*	*αRep G8*
Data collection		
Wavelength (Å)	0.9119	0.9677
Space group	$P12_11$	$P2_12_12_1$
a, b, c (Å)	60.58, 46.14, 60.76	57.61, 96.65, 149.89
α, β, γ (°)	90, 112.58, 90	90, 90, 90
Resolution range (Å)	27.2 - 2.2 (2.28 - 2.2)	45.67 - 3.2 (3.315 - 3.2)
No. of reflections	512811 (48735)	358025 (35811)
Unique reflections	15839 (1532)	13722 (1353)
Multiplicity	32.4 (31.3)	26.1 (25.8)
Completeness (%)	98.14 (97.15)	92.08 (96.06)
⟨I/σ(I)⟩	81.88 (29.44)	14.53 (1.47)
R_{meas} *(%)*	57.15 (63.83)	86.27 (301.2)
Refinement		
Overall Wilson B ($Å^2$)	17.63	83.23
R.m.s. deviations		
Bond lengths (Å)	0.010	0.009
Bond angles (°)	1.29	1.22
No. of atoms		
Protein non-hydrogen atoms	2081	4448
Water	188	None*
B factors ($Å^2$)	19.96	101.43
Protein	19.21	101.43
Water	27.54	N/A*
R_{work}/R_{free} *(%)*	20.0 / 25.3	22.1 / 29.1
Ramachandran		
Preferred, allowed, outliers (%)	98.8, 0.4, 0.8	90.8, 7.8, 1.4

The resolution of the E4 structure was 2.2 Å. The outer-shell *I/σ(I)* value, which was 29.44, indicated that the resolution had not reached the limit. However, as

our purpose was to determine the co-crystal structure of E4 with $MPP1_{1-491}$, data collection and analysis were not improved further. The E4 structure demonstrated a typical αRep fold composed of an N-terminal cap, a C-terminal cap, and an αRep internal domain (Figure 4-5). The internal domain comprised six repeats of αRep internal motifs. Each repeat carried five highly randomised amino acid positions, which formed a hypervariable surface, ensuring specific recognition of the target protein (*Introduction 1.1.4.3*, Figure 1-10) [62].

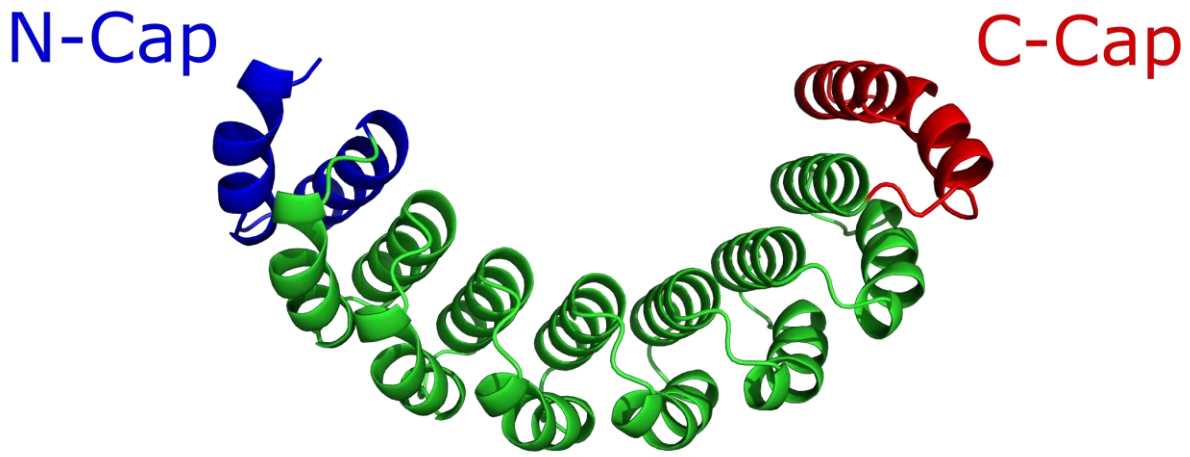


Figure 4-5. Crystal structure of αRep E4, comprising an N-terminal cap (blue), a C-terminal cap (red), and six repeats of αRep internal motifs (green).

The resolution of the solved G8 structure was 3.2 Å, which was insufficient to identify water molecules in the structure. However, the general structure of αRep G8 was obvious (Figure 4-6). There were three molecules in the asymmetric unit (AU). The crystal structure of αRep G8 comprised an N-terminal cap, a C-terminal cap, and five repeats of αRep internal motifs.

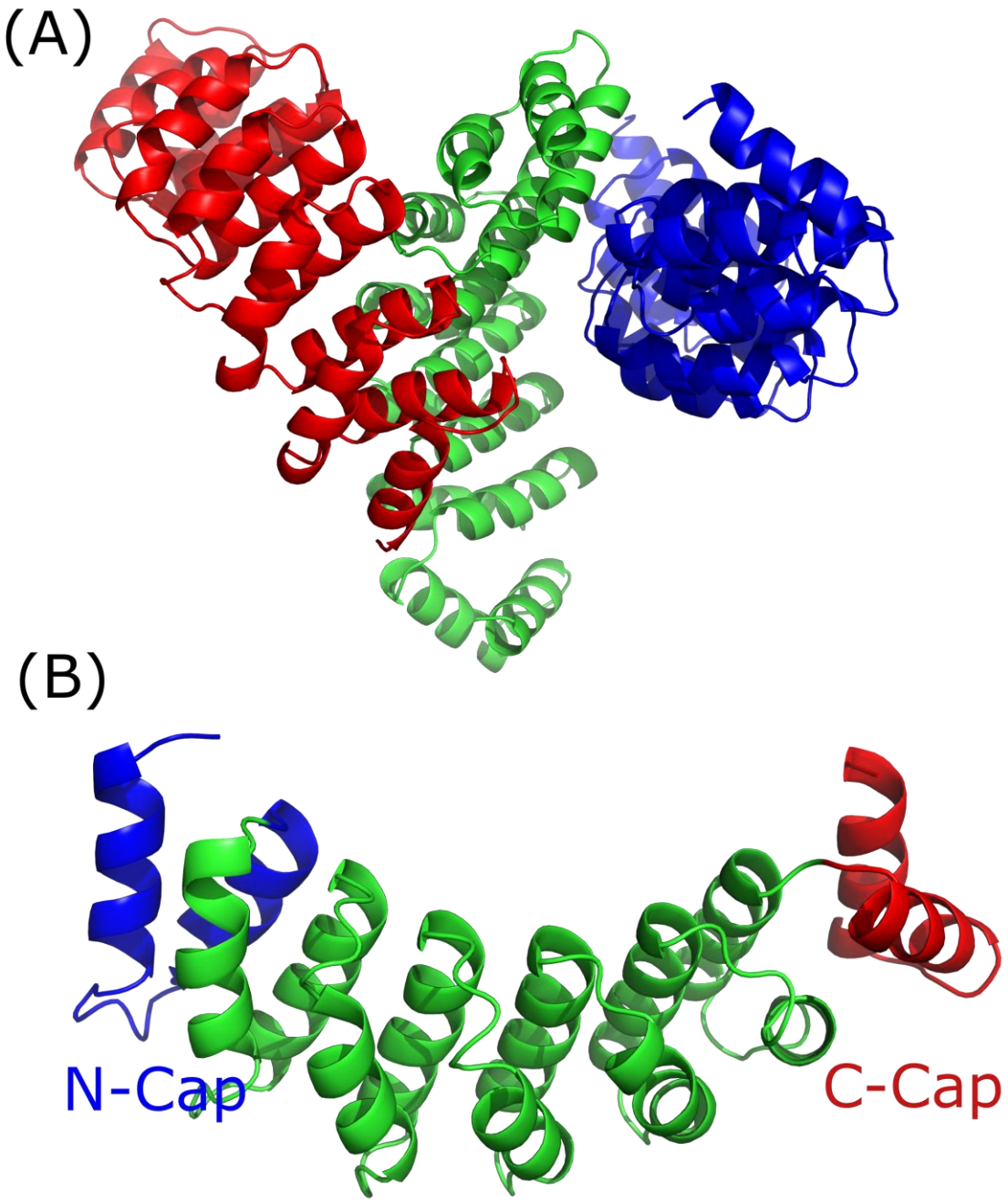


Figure 4-6. (A) Three copies of αRep G8 form the asymmetric unit. (B). Crystal structure of αRep G8, comprising an N-terminal cap (blue), a C-terminal cap (red), and five repeats of αRep internal motifs (green).

The co-crystallisation trials and structure determination indicated that the αReps did not form a complex with $MPP1_{1\text{-}491}$ and did not co-crystallise with $MPP1_{1\text{-}491}$. Further experiments were conducted by increasing the ratios of $MPP1_{1\text{-}491}$ to αRep

to 2:1 and 3:1 when setting up drops. However, only αReps were found in the AU.

4.1.1.4 Designing αRep-MPP1 fusion proteins

The co-crystallisation trials indicated that covalent complexes between αReps and MPP1 may be needed to yield complex crystals. An αRep-Kinase (YabT) fusion protein structure has been reported previously (PDB entry 6G4J) [111]. αReps not only help crystallise the target protein but also increase the expression yields in *E. coli*. The αRep was linked to the N-terminus of the target protein with either a short (SGGGGSGGGG) or long linker (GSAGSAAGS GGASGGGGSGGGGSAGSAAGSGG). As described in the published manuscript, both short and long fusion proteins were crystallised and their structures were determined. The short fusion protein formed a stable dimer in solution, whereas the long fusion protein formed an equilibrium between monomers and dimers, which implies that the short linker impairs the formation of an intramolecular interaction between the αRep and target domains of the fusion protein, thus favouring the formation of intermolecular interactions through dimerisation.

Based on the published structure and our previous MST and crystallisation results, two fusion protein constructs were designed by linking αRep G8 to $MPP1_{54-491}$ with the reported short and long linkers. A short version of $MPP1_{54-491}$ was selected to reduce the molecular weight of the fusion protein (Figure 4-7). The molecular weights of the short and long fusion proteins were ~75 and ~77 kDa, respectively.

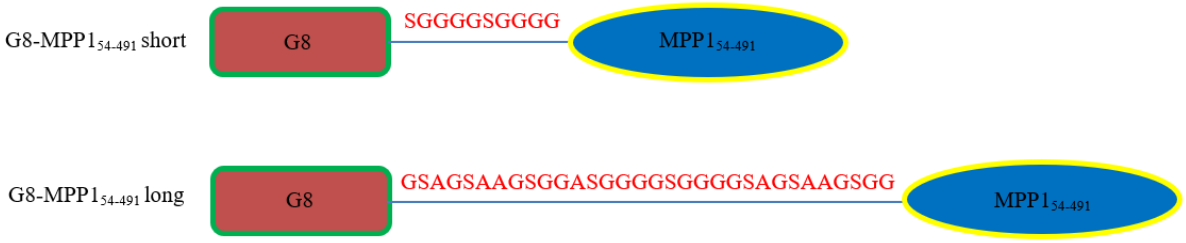


Figure 4-7. G8-MPP1$_{54-491}$ fusion protein constructs. The fusion protein constructs were designed by linking αRep G8 to the N-terminus of MPP1$_{54-491}$ with the reported short and long linkers.

4.1.1.5 Crystallisation trials of G8-MPP1$_{54-491}$ fusion protein (short)

G8-MPP1$_{54-491}$ fusion protein with the short linker was selected for further crystallisation trials as short constructs are favoured by the *E.coli* expression system. The G8-MPP1$_{54-491}$ (short) protein was purified with an 80% purity (Figure 4-8 A) and subjected to crystallisation trials (*Materials and methods 3.3.14*) on 23rd June 2022. We have acquired several crystals (Figure 4-8 B) and are now sending them for measuremens.

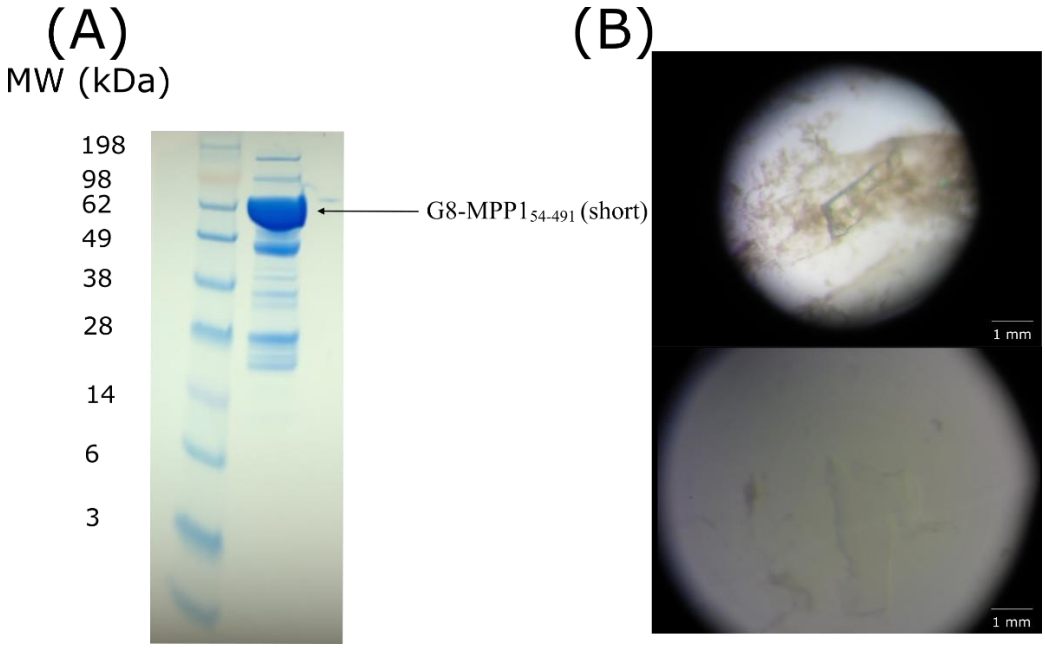


Figure 4-8. (A) G8-MPP1$_{54-491}$ (short) protein was purified with an 80% purity (75.2 kDa). (B) Several crystals have been acquired and sent for measurements.

4.1.2 Pharmacophore-based development of ascochitine analogues as MPP1 inhibitors

4.1.2.1 Characterisation of $MPP1_{1-491}$ in the presence of ascochitine and citrinin

A previous member of our group identified ascochitine (Compound **A9b**) as a potential inhibitor of MPP1 (Figure 4-9 A, B). Here, ascochitine and citrinin were characterised in the presence of $MPP1_{1-491}$ based on the inhibition of basal ATPase activity. Results indicated that ascochitine reduced the basal ATPase activity of $MPP1_{1-491}$ by 85%, with an IC50 value of 19.8 ± 3.3 μM. In contrast, citrinin, a commercially available analogue of ascochitine, did not inhibit $MPP1_{1-491}$ (Figure 4-9 C). The inhibition data indicate that despite a similar overall structure to that of ascochitine, the four minor changes in citrinin completely abolish the inhibition of MPP1.

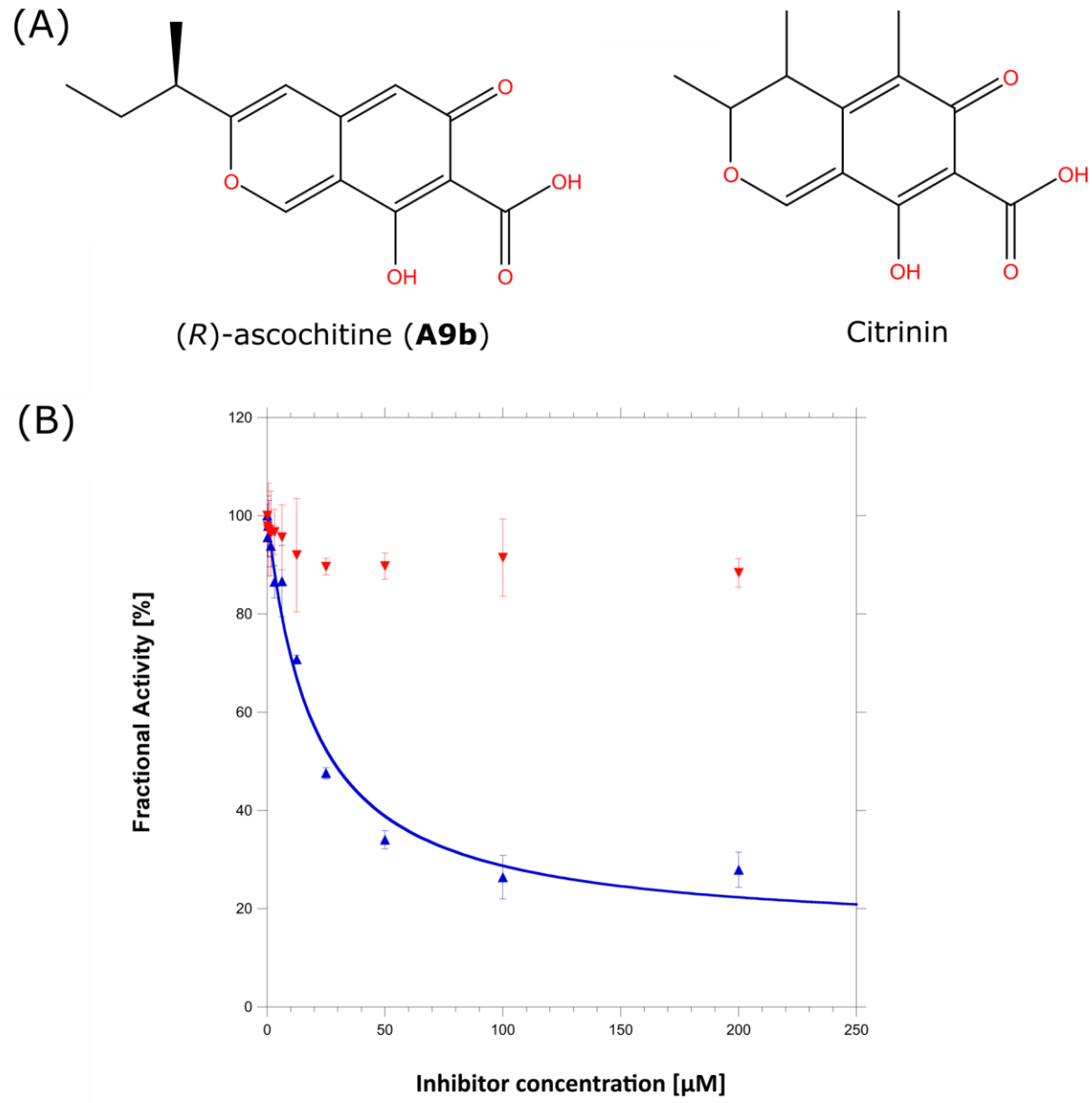


Figure 4-9. (A) Chemical structure of (R)-ascochitine. (B) Chemical structure of citrinin. (C) Concentration-response plot of ascochitine (blue) and citrinin (red) using basal ATPase activity. Ascochitine reduced the activity of $MPP1_{1\text{-}491}$ by 85%, whereas citrinin showed no inhibition.

4.1.2.2 Determination of the importance of the unique loop L6 region for inhibitor binding

Inhibition of the basal and MT-stimulated MPP1 ATPase activities was determined in the presence of ascochitine and two additional control compounds, depsidone analogues **D1** (norlobaridone) and **D2** (physodic acid) (Table 4-4, Table 4-10) [59]. The two depsidone analogues were characterised by a previous group member in the presence of two distinct MPP1 constructs, $MPP1_{57–491}$ and $MPP1_{2–477}$. **D2** inhibits the basal MPP1 ATPase activity with an IC_{50} of 18.7 ± 2.6 μM and MT-stimulated ATPase activity with an IC_{50} of 24.4 ± 3.5 μM, which indicates that it is a weak inhibitor of MPP1. Meanwhile, the other analogue **D2** is more active than **D1**, inhibiting the basal and MT-stimulated ATPase activities with IC_{50} values of 9.9 ± 1.1 μM and 5.8 ± 0.9 μM, respectively [59]. Both basal and MT-stimulated MPP1 ATPase assays were established in the 96-well format in half-area plates and were used to determine the IC_{50} values of the inhibitor analogues. The IC_{50} value of ascochitine for inhibition of the $MPP1_{1-491}$ construct was 16 μM; values of 9.5 μM and 20.1 μM were obtained for the $MPP1_{54-491}$ and $MPP1_{54-491}$ ΔL-6 constructs, respectively (Figure 4-10 A). Similar results were obtained for the inhibition of MT-stimulated MPP1 ATPase activity (Figure 4-10 D), although the IC50 values were systematically better in the presence (the physiologically more relevant condition) than in the absence of MTs (e.g., $MPP1_{1-491}$ construct: $IC_{50,basal}$ = 16 μM, $IC_{50,MTs}$ = 11.3 μM).

*Table 4-4. Inhibition of basal and MT-stimulated ATPase activities in three MPP1 constructs used in this study. n.i.: no inhibition; n.d.: not determined. MIA: maximal inhibition acquired. Depsidone analogues **D1** and **D2** represent a different chemical scaffold derived from another more complex natural product and were used as controls. The data for $MPP1_{57-491}$ and $MPP1_{2-477}$ are acquired from the study by Talapatra et al. [59].*

MPP1 constructs	***Inhibition of basal MPP1 ATPase activity, IC_{50} [μM]*** ***MIA [%]***			***Inhibition of MT-stimulated MPP1 ATPase activity, IC_{50} [μM]*** ***MIA [%]***		
	ascochitine	***D1***	***D2***	***ascochitine***	***D1***	***D2***
$MPP1_{1-491}$	16.0 ± 1.8 (75)	11.6 ± 2.1 (75)	13.6 ± 2.8 (70)	11.3 ± 1.1 (75)	6.3 ± 0.3 (85)	3.8 ± 0.6 (75)
$MPP1_{54-491}$	9.5 ± 0.5 (78)	10.6 ± 1.3 (85)	14.8 ± 3.0 (75)	8.6 ± 1.0 (75)	7.3 ± 1.2 (80)	7.1 ± 1.6 (65)
$MPP1_{54-491\ \Delta L-6}$	20.1 ± 3.8 (75)	37.7 ± 6.7 (70)	13.6 ± 1.9 (60)	4.5 ± 0.9 (60)	9.9 ± 1.6 (70)	5.8 ± 0.6 (60)
$MPP1_{57-491}$ [59]	17. 6 ± 2.0 (80)	9.9 ± 1.1	18.7 ± 2.6	n.d.	5.8 ± 0.9	24.4 ± 3.5
$MPP1_{2-477}$ [59]	n.d.	10.4 ± 1.7	10.8 ± 1.7	n.d.	4.6 ± 0.8	7.6 ± 0.7

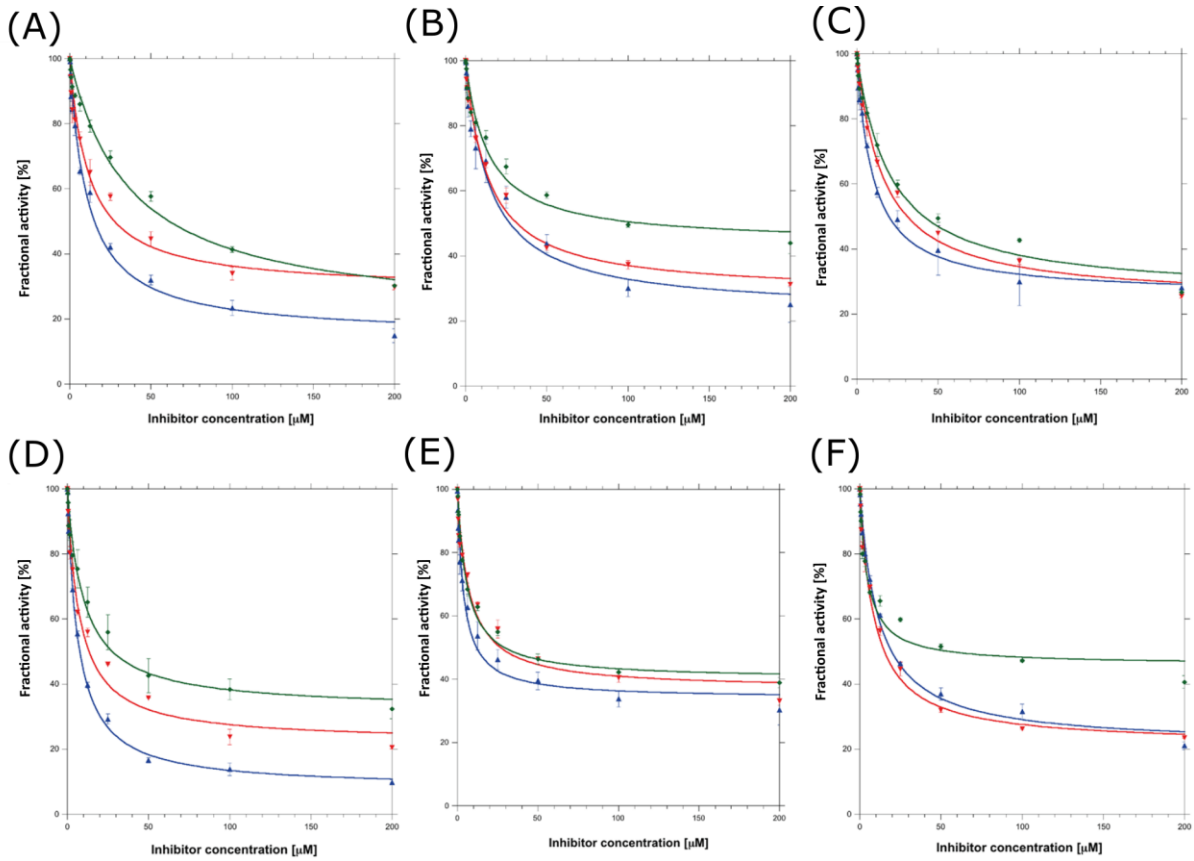


*Figure 4-10. Concentration-response plots of MPP1 inhibitors. (A) Basal ATPase characterisation of ascochitine in the presence of three distinct MPP1 constructs. (B) Basal ATPase characterisation of depsidone analogue **D1** in the presence of three distinct MPP1 constructs. (C) Basal ATPase characterisation of depsidone analogue **D2** in the presence of three distinct MPP1 constructs. (D) MT-stimulated ATPase characterisation of ascochitine in the presence of three distinct MPP1 constructs. (E) MT-stimulated ATPase characterisation of depsidone analogue **D1** in the presence of three distinct MPP1 constructs. (F) MT-stimulated ATPase characterisation of depsidone analogue **D2** in the presence of three distinct MPP1 constructs. $MPP1_{1-491}$ (red), $MPP1_{54-491}$ (blue), and $MPP1_{54-491\,\Delta L-6}$ (green).*

Interestingly, there was no significant difference in IC_{50} values between the MPP1 constructs with and without the loop L6 region, indicating that this unique and characteristic region of MPP1 is not involved in binding with ascochitine and depsidones. In summary, inhibition assays were set up with the help of control

compounds. These miniaturised assays can now be applied for determining the IC_{50} values of novel inhibitor analogues.

4.1.2.3 Synthesis of inhibitor analogues targeting human MPP1

The synthesis of novel MPP1 inhibitor analogues was outsourced to the company Sai (Figure 4-11). In total, 15 inhibitor analogues were ordered. Sai successfully synthesised 13 of the 15 inhibitor analogues, including seven intermediates (**A4a** to **A4g**) and six ascochitine analogues (**A5b, A5c, A5d, A5e, A5i, and A10**). Synthesis of the three remaining analogues (**A5f, A5g,** and **A5h**) failed at the last cyclisation step of the synthetic route (Table 4-5). All 12 analogues were provided along with their characterisation results obtained via H-NMR, C-NMR, and high-resolution mass spectrometry.

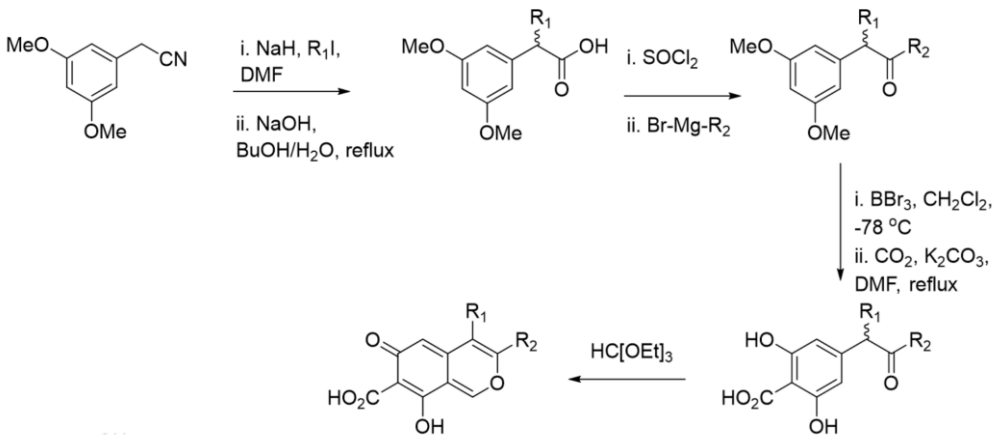
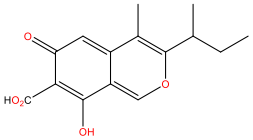
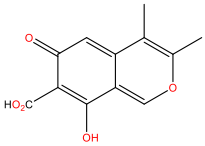
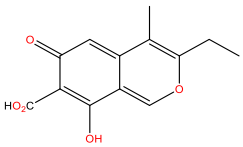
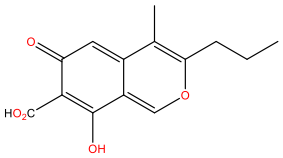
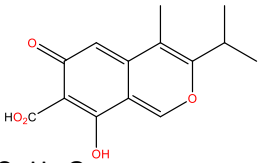
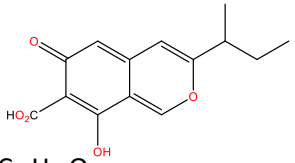
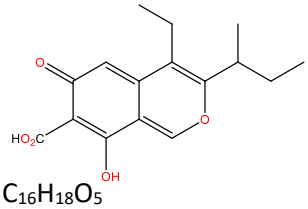
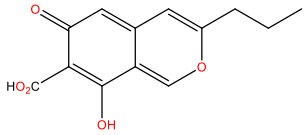
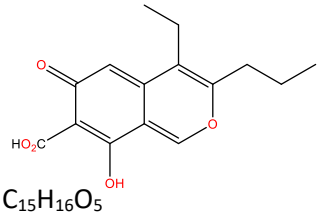
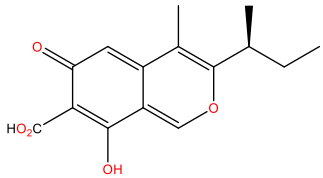
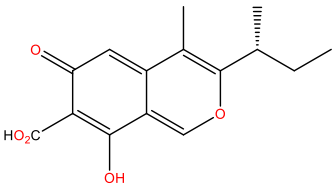
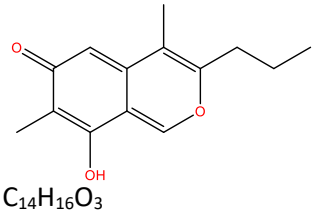


Figure 4-11. Synthetic route for MPP1 inhibitor analogues. The synthesis was challenging and was outsourced to the company Sai.

*Table 4-5. Chemical properties of MPP1 analogues ordered from Sai and the ascochitine controls. *1 Intermediates **A4a** to **A4g** are not included, as they showed no inhibition in the ATPase assay. *2 The synthesis of analogues **A5e** to **A5g** failed at the last cyclisation step of the synthetic route. *3**A9a** and **A9b** are ascochitine controls representing different enantiomers, whereas **A5a** is a mixture of enantiomers. Molecular weight, MW; hydrogen-bond donors, HBD; hydrogen-bond acceptors, HBA; polar surface area, PSA.*

Name	***Structure***	***MW [Da]***	***Number of HBA/ HBD***	***cLogP***	***PSA [Å²]***
****3A5a*** ***(+/-) ascichitine***	 $C_{15}H_{16}O_5$	276.3	4/2	2.31	83.8
A5b	 $C_{12}H_{10}O_5$	234.2	4/2	1.22	83.8
A5c	 $C_{13}H_{12}O_5$	248.2	4/2	1.47	83.8
A5d	 $C_{14}H_{14}O_5$	262.3	4/2	2.03	83.8
A5e	 $C_{14}H_{14}O_5$	262.3	4/2	1.76	83.8
****2A5f***	 $C_{14}H_{14}O_5$	262.3	4/2	2.18	83.8

[2]A5g	 $C_{16}H_{18}O_5$	290.3	4/2	2.78	83.8
[2]A5h	 $C_{13}H_{12}O_5$	248.2	4/2	1.90	83.8
A5i	 $C_{15}H_{16}O_5$	276.3	4/2	2.50	83.8
[3]A9a ***(S) ascochitine***	 $C_{15}H_{16}O_5$	276.3	4/2	2.31	83.8
[3]A9b ***(R) ascochitine***	 $C_{15}H_{16}O_5$	276.3	4/2	2.31	83.8
A10	 $C_{14}H_{16}O_3$	232.3	3/1	1.75	46.5

4.1.2.4 SAR study of novel MPP1 analogues

An SAR study, using the first batch of 12 analogues, was conducted by measuring the inhibition of the basal ATPase activities of the $MPP1_{1\text{-}491}$, $MPP1_{54\text{-}491}$, and $MPP1_{54\text{-}491\ \Delta L\text{-}6}$ constructs. The results are summarised in Table 3. All intermediate analogues (**A4a** to **A4g**) were inactive, indicating that the two-ring system is essential for inhibition (data not shown). This signifies that the final and difficult to achieve cyclisation step, which is very challenging, cannot be avoided when obtaining inhibitor analogues with inhibitory activity. The analogue with the methyl substituent (**A5b**) is a weak inhibitor of $MPP1_{54\text{-}491}$ (IC_{50} = 53 μM), whereas the analogue with the ethyl substituent (**A5c**) has an IC_{50} value of 11.7 μM. The most active analogue (**A5d**) of this series with an n-propyl group has an IC_{50} of 4.2 μM and is slightly more active than the iso-propyl containing analogue (**A5e**; IC_{50} = 5.6 μM). Both analogues are approximately 2-folds more active than the parent compound ascochitine (**A5a**), with an IC_{50} of 9.1 μM. The same trend could be observed for the $MPP1_{1\text{-}491}$ construct (Table 4-6). The results for the $MPP1_{54\text{-}491\ \Delta L\text{-}6}$ construct are quite consistent; the activity of four out of five analogues dropped by 2- to 3-folds, indicating that this unique region (loop L6), which is only present in kinesin-6 family members, may be involved in some aspect of inhibitor binding.

The most active compounds **A5d** and **A5e** were further characterised using the MT-stimulated ATPase assay (Table 4-7). Interestingly, the inhibitors were systematically more potent in the presence of MTs, indicating that the compounds probably bind to an allosteric pocket rather than the ATP-binding site or MT-binding interface.

In summary, the results from the first batch of inhibitor analogues showed that all tested intermediates were inactive, indicating that the two-ring system is essential for inhibition. Therefore, further synthesis and testing of synthetic intermediates is unnecessary. As the n-propyl substituent is more active than the sec-butyl analogue, separation of enantiomeric mixtures during subsequent analogue synthesis or development of an enantio-selective synthesis is unnecessary, thus shortening the synthesis process of novel inhibitor analogues.

*Table 4-6. Initial SAR study measuring the inhibition of the basal MPP1 ATPase activity for the first batch of MPP1 inhibitor analogues. MIA: maximum inhibition attained in percentage. *[1]The control ascochitine (**A5a**) is a racemic mixture. *[2]The synthesis of this analogue is still in progress. *[3]analogues **A9a** and **A9b** represent the separated enantiomers of ascochitine. *[4]Citrinin is a close natural product analogue with a slightly modified core scaffold. *[5]The inhibitory activity of the remaining analogues will be measured once compound **A5** has been synthesised and delivered. *[6]Synthesis of these analogues carrying the butyl substituent failed but was subsequently performed using the n-propyl group (analogues **A5h** and **A5i**). *[7]To be measured (tbm). *[8]No inhibition (n.i.).*

Compound No.	***MPP1$_{1-491}$ construct IC_{50} [µM] MIA [%]***	***Short MPP1$_{54-491}$ construct IC_{50} [µM] MIA [%]***	***Short MPP1$_{54-491\ \Delta L6}$ construct IC_{50} [µM] MIA [%]***
****[1]ascochitine Control A5a***	16.0 ± 1.8 (75)	9.1 ± 0.6 (75)	20.1 ± 3.8 (75)
A4a	n.i. (10)	n.i. (10)	n.d.
A4b	n.i. (10)	n.i. (5)	n.d.
A4c	n.i. (5)	n.i. (5)	n.d.
A4d	n.i. (5)	n.i. (5)	n.d.
A4e	n.i. (10)	n.i. (10)	n.d.
A4f	n.i.	n.i.	n.d.

	(10)	(10)	
A4g	n.i. (20)	n.i. (15)	n.d.
A5b	68.7 ±12.5 (59)	53.2 ± 3.1 (60)	17.2 ± 3.6 (50)
A5c	26.8 ± 6.5 (62)	11.7 ± 1.7 (65)	31.7 ± 0.6 (50)
A5d	**6.6 ± 0.3 (62)**	**4.2 ± 0.6 (75)**	**12.4 ± 1.7 (58)**
A5e	11.9 ± 2.0 (61)	5.6 ± 0.6 (70)	13.6 ± 1.2 (72)
[*6]A5f	Synthesis failed		
[*6]A5g	Synthesis failed		
A5h	Synthesis failed		
A5i	23.3 ± 4.8	[*7]tbm	[*7]tbm
[*3]A9a (ascochitine-S)	14.9 ± 3.4	[*7]tbm	[*7]tbm
[*3]A9b (ascochitine-R)	19.8 ± 3.3	[*7]tbm	[*7]tbm
[*2]A5	The synthesis of this analogue is still in progress.		
A10	*[8]n.i.	[*7]tbm	[*7]tbm
Citrinin	*[8]n.i.	[*7]tbm	[*7]tbm

*Table 4-7. Characterisation of **A5d** and **A5e** using the MT-stimulated ATPase assay. The inhibitors are systematically more potent in the presence of MTs, which resemble conditions closer to physiological conditions.*

Compound No	***$MPP1_{1-491}$ IC_{50} [µM] MIA (%)***	***$MPP1_{54-491}$ IC_{50} [µM] MIA (%)***
	MT-stimulated	*MT-stimulated*
A5d	5.5 ± 0.6 (86)	3.4 ± 0.6 (75)
A5e	10.3 ± 1.0 (90)	4.7 ± 0.7 (80)
Ascochitine (A5a)	16.0 ± 1.8	9.5 ± 0.5
Control	(75)	(78)

4.1.2.5 Determination of the drug-like properties of potent inhibitor analogues

A range of drug-like property assays was outsourced to Cyprotex (Alderley Park, Cheshire, United Kingdom), a contract research organisation (CRO), to profile ascochitine and the most potent analogues.

The results for analogues **A5d** and **A5e** are summarised in Table 4-8. Notably, the most active inhibitor analogues had an average molecular weight of 262.3 Da, which is very low, providing the option for further chemical optimisation by adding new functional groups, if necessary, without unfavourably increasing the MW beyond the Lipinski limit (500 Da). Compound **A5d**, the most potent analogue tested so far, showed good solubility (> 100 μM) at all the three pH values tested. For **A5e**, solubility was pH-dependent, but good solubility was obtained at neutral pH. Lipophilicity (logP or logD) was not experimentally determined at this stage, but the calculated cLogP values of 3.75 and 3.62 indicated that both analogues were relatively balanced compounds. P450 inhibition was experimentally tested using five substrates. With these, both analogues showed only moderate to low inhibition of the CYP2C9 isoform, with IC_{50} values of 23.0 ± 0.8 μM and 10.8 ± 0.6 μM. Importantly, the isoforms CYP2D6 and CYP3A4, which together metabolise 80% of all drugs, were not inhibited by these two analogues. Furthermore, we tested the analogues in human plasma protein binding assays. The test compounds were not detectable in the assays in either the buffer or plasma compartment, indicating the instability of these analogues in human plasma. Although this suggests that the analogues contain hydrolytically labile functional groups, activity may vary considerably (with a 2- to 20-fold difference) depending greatly on the batch or vendor. This was repeated and confirmed. To determine *in vitro* permeability, a bidirectional

Caco-2 cell layer assay was employed. In this assay, **A5d** showed good to moderate permeability, whereas **A5e** was notably less permeable, with low to moderate permeability. Analogue **A5d** showed no indication of efflux, whereas **A5e** had an efflux ratio of 4.6, which indicated that this compound may be a substrate for efflux transporters such as Pgp, BCRP, or MRP2. We also tested the metabolic stability of the analogues in liver microsomes. Both the analogues showed high clearance in both human and mouse microsomal stability assays, with half-lives between 7.6 and 16.6 min, although analogue **A5e** was slightly more stable than **A5d**. A significant turnover of the compounds was also observed in the absence of NADPH, indicating that the metabolism is partly mediated by non-NADPH-dependent enzymes. Finally, we tested whether our analogues display hERG liabilities. However, no inhibition was observed in the concentration range investigated.

In summary, the potency of the inhibitor scaffold as well as its metabolic and plasma stability need to be improved while maintaining its solubility, cell permeability, and absence of other liabilities such as CYP or hERG inhibition.

Table 4-8. Determination of various drug-like properties of two selected MPP1 inhibitor analogues

Assay / Analogues	***A5d***	***A5e***
MW [Da]	262.3	262.3
Turbidimetric solubility Calculated Mid-range [μM] (pH 2.0, 6.0, 7.4)	> 100; > 100; > 100	20; 65; > 100
clogP	3.75	3.62
Cytochrome P450 inhibition [μM]		
2C19	> 25	> 25
1A	> 25	> 25
2D6	> 25	> 25
2C9	23.0 ± 0.8	10.8 ± 0.6
	> 25	> 25

3A4		
Plasma protein binding [$fu_{100\%}$] [% recovery]	Test compound not detectable in the assay, suggesting instability of the test compound in human plasma	Test compound not detectable in the assay, suggesting instability of the test compound in human plasma
Stability in human plasma $t_{1/2}$ [min]	9.3	14.4
Chemical stability $t_{1/2}$ [min]	Predominantly stable	Predominantly stable
Caco-2 permeability		
Direction A2B		
Mean P_{app} [$10^{-6}cms^{-1}$]	18.7 ± 11.0	3.7 ± 0.3
Mean recovery [%]	49.8	31.1
Direction B2A		
Mean P_{app} [$10^{-6}cms^{-1}$]	13.9 ± 1.0	17.0 ± 1.2
Mean recovery [%]	60.2	63.0
Efflux ratio [mean P_{app} B2A / mean P_{app} A2B]	0.7	4.6
Microsomal stability		
Human		
Cl_{int} [μL/min/mg protein]	177 ± 32.1	89.8 ± 32.8
$t_{1/2}$ [min]	7.8	15.4
Mouse		
Cl_{int} [μL/min/mg protein]	157 ± 26.4	83.6 ± 12.1
$t_{1/2}$ [min]	8.8	16.6
hERG [μM]	> 25	> 25

4.1.2.6 Determination of GI_{50} values using cell proliferation assays, overexpression of MPP1 in tumour cell lines, and investigation of drug combinations

Cell proliferation experiments were performed by a post-doctoral researcher in our laboratory, Dr. Ahmed Ahmed. The data are discussed here for a comprehensive view on MPP1 analogues. Ahmed performed proliferation assays and determined the GI_{50} values for the inhibitor analogues in bladder tumour cell lines, investigated the overexpression in bladder cancer tumour cell lines, and

examined the potential additive or synergistic effects of MPP1 inhibitors with other kinesin-targeting inhibitors.

GI_{50} values of Eg5 inhibitor STLC and potential MPP1 inhibitors

Compound	RT4	UM-UC-03	HeLa
STLC	0.849 ± 0.17	1.169 ± 0.27	1.23 ± 0.14
Ascochitine	142.10 ± 32.32	96.38 ± 3.50	130.43 ± 5.35
5e	85.27 ± 6.84	56.75 ± 5.12	64.95 ± 2.04
5d	83.54 ± 1.09	71.19 ± 12.0	71.16 ± 8.84

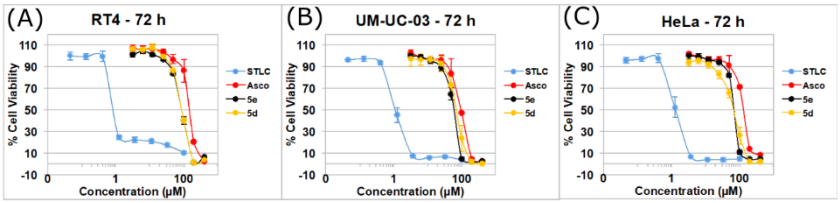


*Figure 4-12. Concentration-response curves of MPP1 inhibitors in bladder (UM-UC-03 and RT4) and cervical (HeLa) cancer cell lines. The inhibitor STLC (Eg5 inhibitor) was used as a control. Briefly, 4000 cells/well of RT4 and 1000 cells/well of UM-UC-03 and HeLa cells were seeded in triplicate in 96-well plates and incubated for 24 h. The cells were then treated with eight different concentrations of STLC, ascochitine, or compounds **A5e** and **A5d** for 72 h. After treatment, the cells were fixed with 10% TCA and cell viability was measured using the Sulforhodamine B (SRB) assay. Data represent the mean ± SEM of values from four independent experiments.*

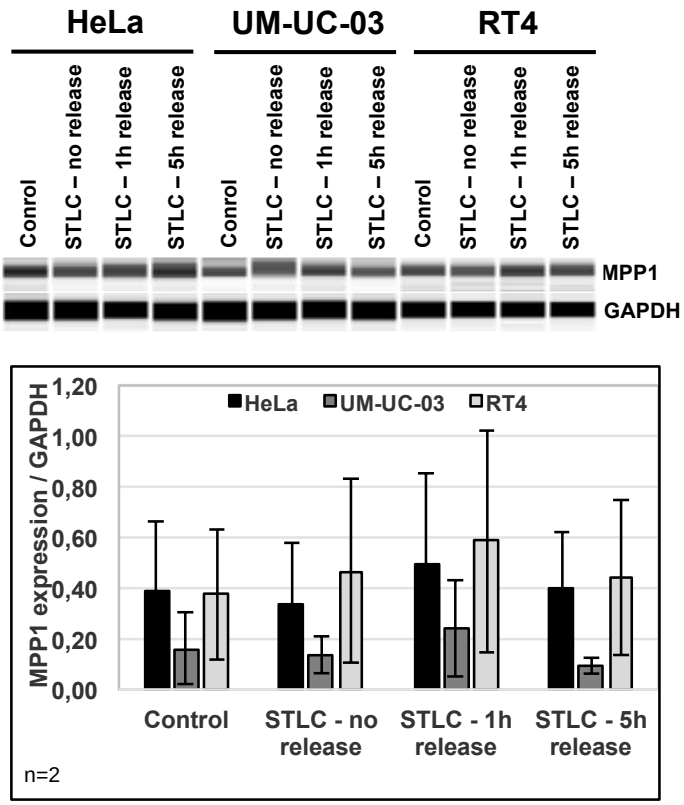


Figure 4-13. Summary of MPP1 protein expression analysis in three cancer cell lines. Protein expression levels of MPP1 in bladder (UM-UC-03 and RT4) and cervical (HeLa) cancer cell lines. Cells were seeded in 6-well plates and incubated for 24 h for attachment. The cells were then blocked in mitosis with 10 µM STLC for 18 h. Three conditions were tested—1) no release, 2) 1 h release, and 3) 5 h release; an untreated control was also included. Cells were collected and lysed with RIPA lysis buffer supplemented with protease and phosphatase inhibitors. Western blot analyses were performed to detect MPP1 protein levels according to the manufacturer's instructions (ProteinSimple). Protein bands were quantified using the software Compass for SW, and the values were normalised to those of the loading control GAPDH. Data represent the mean ± SEM of values from two independent experiments.

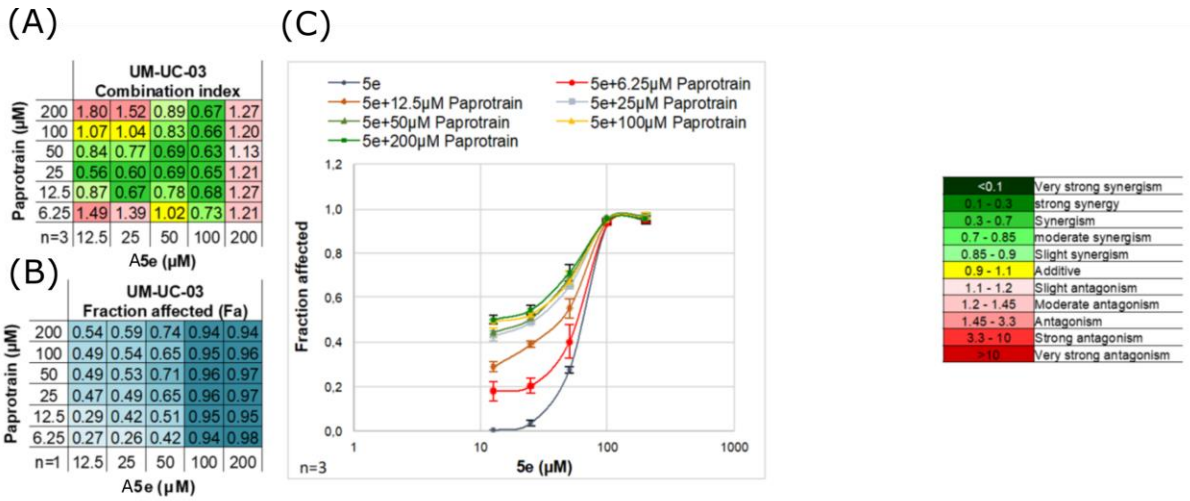


*Figure 4-14. Synergy analysis between **A5e** and paprotrain in bladder cancer (UM-UC-03) cells. (A) Combination index (CI) values of the two-drug combination, CI < 1, CI = 1, and CI > 1 indicate synergism (green), additive effects (yellow), and antagonism (red), respectively. (B) Fraction affected (Fa) by each drug combination using the Sulforhodamine B (SRB) cytotoxicity assay. (C) Dose-response curves for **A5e** alone and in combination with paprotrain. UM-UC-03 cells (1000 cells/well) were seeded in a 96-well plate and incubated overnight. Cells were then treated with the two-drug combination or a single drug for 72 h. A 6 × 5 matrix format was used to assess all two-drug combinations using a range of concentrations at, above, and below the GI_{50} of each compound. After treatment, cells were fixed with 10% TCA and viability was measured using the SRB assay. The combination indices (CI) were calculated using the CalcuSyn software based on the Chou-Talalay non-constant ratio method.*

Proliferation assays were conducted using two bladder cancer cell lines and HeLa cells. There were no significant differences in the GI_{50} values for all three tumour cell lines (Figure 4-12). The low inhibitory activity of MPP1 analogues was significantly less pronounced in cell-based assays compared to that *in vitro*, possibly due to their weak metabolic stability (see *drug-like properties 4.1.2.5*). Furthermore, no significant differences were observed in protein expression levels between bladder cancer tumour cell lines and HeLa cells, which appears to contradict the published data, indicating that MPP1 is overexpressed in bladder cancers (Figure 4-13) [60]. However, additional analysis is needed in various

tumour cell lines to confirm this result. We also compared the inhibitory effects of MPP1 inhibitors alone and in combination with paprotrain, an inhibitor targeting the kinesin-6 family member MKLP-2 [121]. The results indicated weak synergistic effects when using both inhibitors targeting different aspects of cytokinesis.

4.1.2.7 Metabolite identification study of MPP1 inhibitor analogue A5d

Metabolite profiling was performed on the MPP1 analogue **A5d** in human plasma samples. In summary, a total of three metabolites have been found and reported (Figure 4-15). Structural elucidation has been provided for all three metabolites, M1, M2, and M3. Surprisingly, oxidation initially occurred at the n-propyl and methyl substituents but not in the ring system. The second metabolite displayed further oxidation, whereas the third metabolite showed two oxidation reactions and hydration of the ring system. This is surprising, as we mainly identified the hydroxyl, ketone, and carboxylic acid substituents as potential liabilities with respect to metabolism. Unfortunately, although high-resolution mass spectrometry techniques were used, it could not be clarified whether hydroxylation, epoxidation, or oxide formation occurred or at which alkyl substituent this occurred. Furthermore, it was unclear how the ring system was hydrated in the third metabolite.

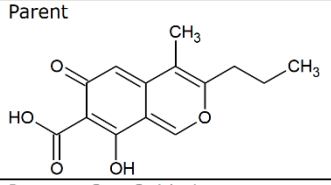
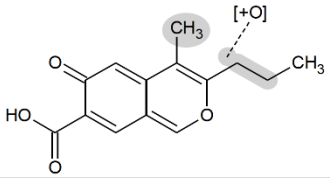
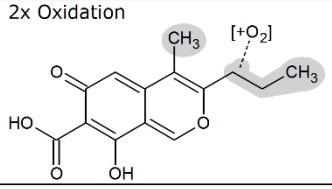
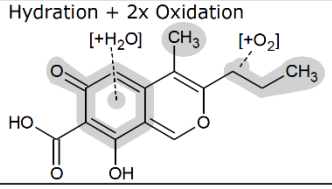
Proposed Assignment		Human
Parent	Parent 	✓
M1	Parent - O + Oxidation 	✓
M2	2x Oxidation 	✓
M3	Hydration + 2x Oxidation 	✓

*Figure 4-15. Summary of the identified **A5d** metabolites in human plasma samples.*

4.1.2.8 Development of additional synthetically accessible MPP1 analogues

The synthesis of novel ascochitine-based inhibitors is challenging, with up to 11 steps required for certain analogues (see *synthesis of inhibitor analogues 4.1.2.3*). The final cyclisation step is particularly challenging and failed for some interesting analogues. Therefore, to speed up the generation of novel inhibitor analogues, it is essential to simplify and shorten their synthesis by modifying the pharma.cophore that is capable of inhibiting MPP1.

A pharmacophore-based search was conducted by Mr. Helal Abdulghani Helal, a PhD student in Dr. Geoffrey Wells' group (Department of Pharmaceutical and

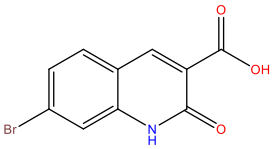
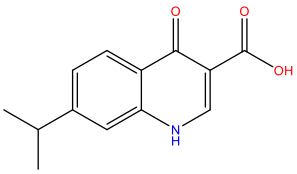
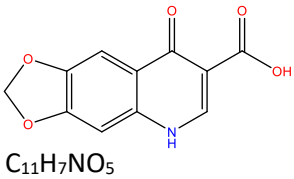
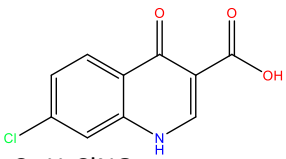
Biological Chemistry, UCL School of Pharmacy). The search suggested that the core structures of chromenone-3-carboxylic acid and quinolone-3-carboxylic acid would be isosteres of the less synthetically accessible ascochitine, with a significantly reduced number of synthesis steps, i.e., 2–3 (Mr. Helal Helal, personal communication) (Figure 4-16 A).

Analogues were synthesised based on the chromenone-3-carboxylic acid and quinolone-3-carboxylic acid core structures. Finally, 24 analogues were measured using the basal ATPase assay. However, most of them showed very weak or no inhibition of $MPP1_{1\text{-}491}$. Nine analogues showed atypical basal ATPase curves, which were further measured using the MT-stimulated ATPase assay. Finally, the four analogues, **A84, A91, A99,** and **A100,** were identified using the MT-stimulated ATPase assay (Figure 4-16). Analogue **A84** reduced the MT-stimulated ATPase activity of $MPP1_{1\text{-}491}$ by up to 98%, with an IC_{50} value of 42.5 ± 17.6 μM. Meanwhile, analogue **A91** reduced the MT-stimulated ATPase activity of $MPP1_{1\text{-}491}$ by up to 84%, with an IC_{50} value of 59.7 ± 16.1 μM. The other two hits, analogues **A99** and **A100**, showed inhibition of $MPP1_{1\text{-}491}$, with IC_{50} values of 58.2 ± 16.1 μM and 33.3 ± 15.8 μM, respectively. However, the inhibition curves of **A99** and **A100** were slightly atypical, suggesting that their IC_{50} values were higher than the measured values and more data points should be measured at higher concentrations to accurately identify their IC_{50} values (Figure 4-16 B).

All four hits identified in the ATPase assay shared a quinolone-3-carboxylic acid core scaffold, whereas analogues based on the chromenone-3-carboxylic acid scaffold did not cause inhibition of $MPP1_{1\text{-}491}$. Meanwhile, analogue **A84**, which was synthesised based on the quinolone-3-carboxylic acid scaffold, was a

stronger inhibitor compared with the 4(1H)-quinolinone derivatives. In terms of the 4(1H)-quinolinone derivatives, any replacements of the 8-isopropyl substituent would sharply reduce the IC_{50} values of the analogues.

Table 4-9. MPP1 analogues identified using the MT-stimulated ATPase assay. The IC_{50} values were measured in the presence of $MPP1_{1-491}$. The analogues share a core quinolone-3-carboxylic acid scaffold, which is an isostere of the less synthetically accessible ascochitine. Molecular weight, MW; hydrogen-bond donors, HBD; hydrogen-bond acceptors, HBA; polar surface area, PSA.

Compound name	***Structure***	***MW [Da]***	***Number of HBA/ HBD***	***cLogP***	***PSA [Å²]***	***IC_{50} [μM]***
A84	 $C_{10}H_6BrNO_3$	268.1	2/2	1.40	66.4	42.5 ± 17.6
A91	 $C_{13}H_{13}NO_3$	231.3	3/2	1.65	66.4	59.7 ± 16.1
A99	 $C_{11}H_7NO_5$	233.2	5/2	0.20	84.9	58.2 ± 16.1
A100	 $C_{10}H_6ClNO_3$	223.6	3/2	0.98	66.4	33.3 ± 15.8

In summary, we identified four MPP1 analogues that inhibited the $MPP1_{1\text{-}491}$ construct with a more synthetically accessible quinolone-3-carboxylic acid scaffold compared with that of ascochitine. However, the quinolone analogues were 5- to 8-fold less active than the ascochitine analogue **A5d** (5.5 ± 0.6 μM) in MT-stimulated ATPase assays. The quinolone analogues did not show inhibition of $MPP1_{1\text{-}491}$ in the basal ATPase assay. We thus conclude that quinolone analogues are not potent MPP1 inhibitors despite their shorter synthetic route.

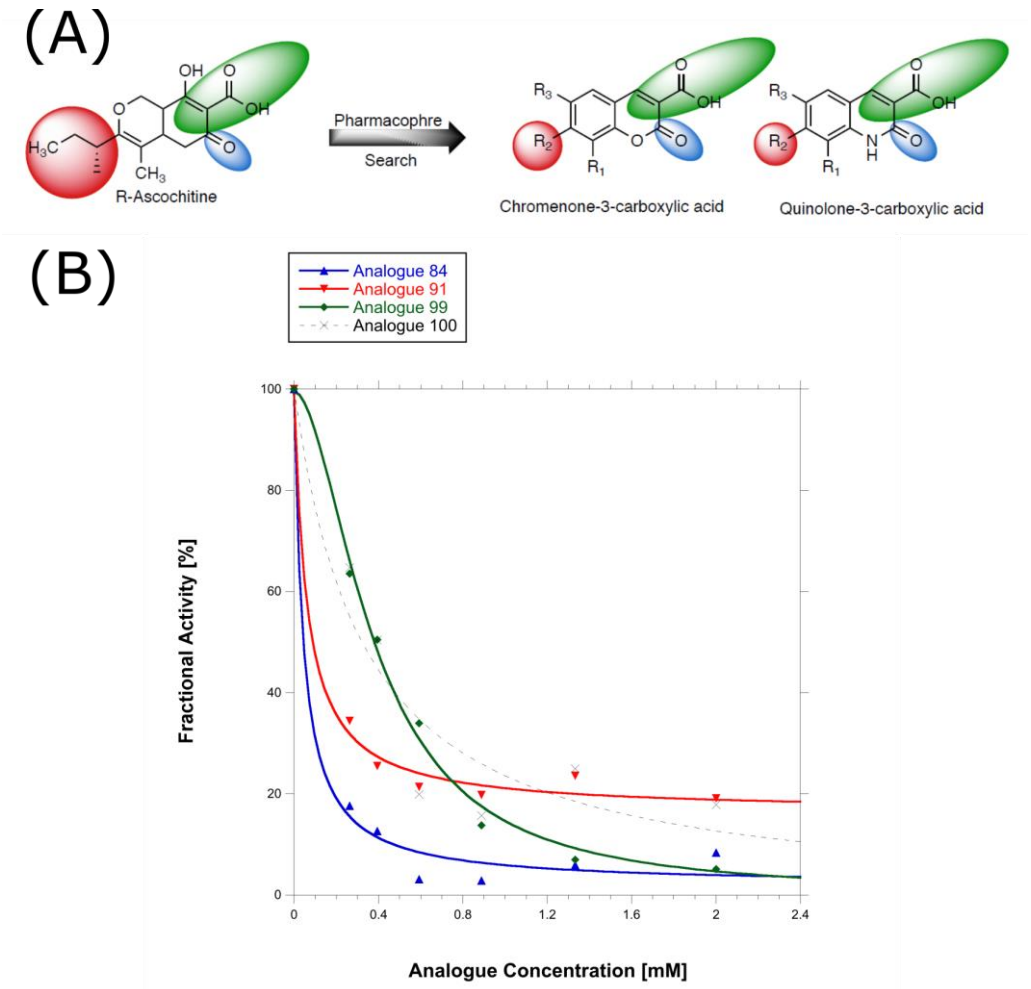


*Figure 4-16. (A) Pharmacophore modelling suggests that chromenone-3-carboxylic acids and quinolone-3-carboxylic acid core structures are isosteres of ascochitine. (B) Inhibition of the MT-stimulated ATPase activity by analogues **A84** (blue), **A91** (red), **A99** (green), and **A100** (black dot) in the presence of $MPP1_{1\text{-}491}$. Analogues **A84** and **A91** show better inhibition of $MPP1_{1\text{-}491}$ than analogues **A99** and **A100**.*

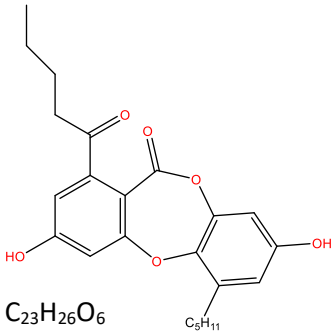
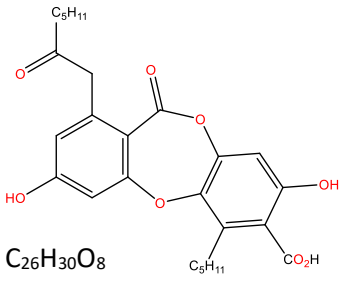
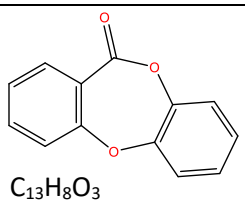
4.1.3 Development of depsidone analogues as MPP1 inhibitors

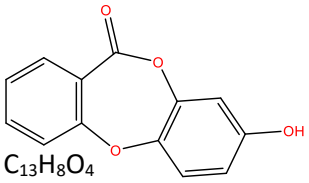
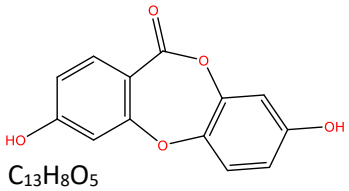
A previous member of our group, Dr. Sandeep Talapatra, identified two depsidone analogues, norlobaridone (**D1**) and physodic acid (**D2**) (Table 4-10), as MPP1 inhibitors by screening several National Cancer Institute (NCI) libraries using inhibition of basal ATPase activity [59]. These two inhibitors were further characterised in the presence of five MPP1 constructs, including $MPP1_{57–491}$, $MPP1_{2–477}$, $MPP1_{1–491}$, $MPP1_{54–491}$, and $MPP1_{54\text{-}491\ \Delta L\text{-}6}$, in both basal and MT-stimulated ATPase assays (see *Results 4.1.2.2*, Figure 4-10, Table 4-4). The IC_{50} values measured using the ATPase assays indicated that both inhibitors **D1** and **D2** are low micromolar inhibitors of MPP1 and that **D2** is slightly more active than **D1**. The type of inhibition by **D2** was then characterised, which suggested that **D2** is an uncompetitive inhibitor with respect to ATP under basal conditions and is a mixed inhibitor with respect to MTs [59]. As depsidone inhibitors do not compete with ATP or MTs, MPP1 must contain at least one allosteric inhibitor-binding pocket, as observed in other kinesins, including Eg5, CENPE, and KifC1. **D1** and **D2** were then subjected to proliferation assays using HCT116 (colon), BxPC3 (pancreatic), K562 (leukaemia), and J82 and UM-UC-3 (both bladder cancer) cells. Both **D1** and **D2** displayed a measurable effect in the panel of cell lines tested, with EC_{50} values ranging between 31.8 and 48.3 μM [59].

The drug-like properties of inhibitors **D1** and **D2** were then profiled [59]. Both **D1** and **D2** showed unfavourable cLogP values, which exceeded Lipinski's limit (< 5). **D2** did not show satisfactory results in both the PAMPA and Caco-2 assays, whereas **D1** showed decent permeability in PAMPA assays ($P_{app} > 10 \times 10^{-6}$ cm/s). The Caco-2 efflux ratio of **D1** was 0.5 (< 2), indicating no drug efflux. The final test of both compounds in human and mouse microsomal stability assays showed high clearance. **D1** showed a half-life of only 3.5 and 3.1 min in human

and mouse microsomes, respectively. In parallel, the half-lives of **D2** were 4.5 and 22.5 min in human and mouse microsomes, respectively. The high clearance and low half-life of both inhibitors indicated that the minimal pharmacophore should be determined to simplify the chemical structure. Thus, three more depsidone analogues, **D3**, **D4,** and **D5**, with minimal scaffolds were synthesised by our collaboration partner Dr. Joëlle Prunet, University Glasgow, Scotland (Table 4-10).

Table 4-10. Chemical structures and properties of depsidone analogues. Molecular weight, MW; hydrogen-bond donors, HBD; hydrogen-bond acceptors, HBA; polar surface area, PSA.

Compound number/ Commercial name	***Structure***	***MW [Da]***	***Number of HBA/ HBD***	***cLog P***	***PSA [Å²]***
D1 (norlobaridone)	 $C_{23}H_{26}O_6$	398.5	5/2	5.20	93.1
D2 (physodic acid)	 $C_{26}H_{30}O_8$	470.5	6/3	5.21	130.4
D3 (depsidone)	 $C_{13}H_8O_3$	212.2	2/0	3.02	35.5

D4 ***(CD13)***	 $C_{13}H_8O_4$	228.2	3/1	2.63	55.8
D5 ***(bis-Desmethylcorynesidone A)***	 $C_{13}H_8O_5$	244.2	4/2	2.24	76.0

Compound **D3**, **D4** and **D5** were characterised using basal ATPase assay in the presence of $MPP1_{1\text{-}491}$ (Figure 4-17) but showed no inhibition. The conclusion is that the 12-pentanoyl and 4-pentyl substitutes are crucial for the MPP1 inhibition.

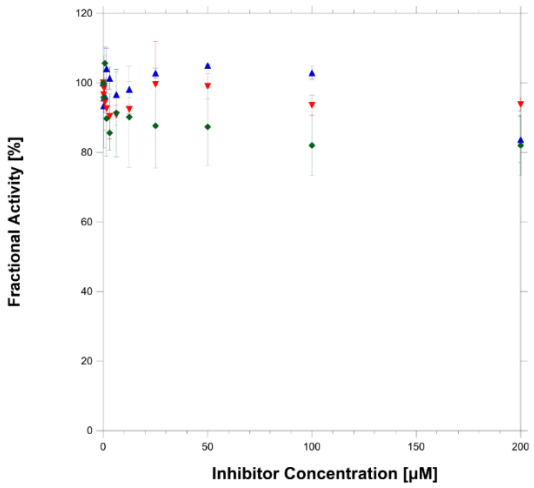


*Figure 4-17. Compound **D3** (blue), **D4** (red) and **D5** (green) were characterised using a basal ATPase assay in the presence of $MPP1_{1\text{-}491}$. No inhibition was observed in all three measurements.*

4.1.4 Low-resolution models of human MPP1 generated from SAXS data

The SAXS data of three distinct MPP1 constructs, $MPP1_{1\text{-}491}$, $MPP1_{54\text{-}491}$, and $MPP1_{54\text{-}491\ \Delta L\text{-}6}$, were measured at EMBL Hamburg. The datasets were analysed using *ATSAS*. The *R(g)* (Guinier) values of $MPP1_{1\text{-}491}$, $MPP1_{54\text{-}491}$, and $MPP1_{54\text{-}491\ \Delta L\text{-}6}$ were 2.78 ± 0.01, 2.54 ± 0.01, and 2.82 ± 0.01 nm, respectively. *P(r)* plots were then generated using *Primus* to analyse the data (Figure 4-18). The *P(r)* plots were sensitive to any conformational changes caused by intramolecular interactions. According to the *P(r)* plots, the $MPP1_{54\text{-}491}$ construct was more compact than the other two constructs. Furthermore, the D_{max} values indicated that the $MPP1_{54\text{-}491}$ monomer was the smallest particle in solution. In contrast, the general shift of $MPP1_{54\text{-}491\ \Delta L\text{-}6}$, the smallest construct, to a larger distance further confirmed the formation of artificial dimers in solution. The calculated *R(g)* values from the *P(r)* plots were 2.84, 2.58, and 2.82 nm for $MPP1_{1\text{-}491}$, $MPP1_{54\text{-}491}$, and $MPP1_{54\text{-}491\ \Delta L\text{-}6}$, respectively, in agreement with the Guinier *R(g)*s.

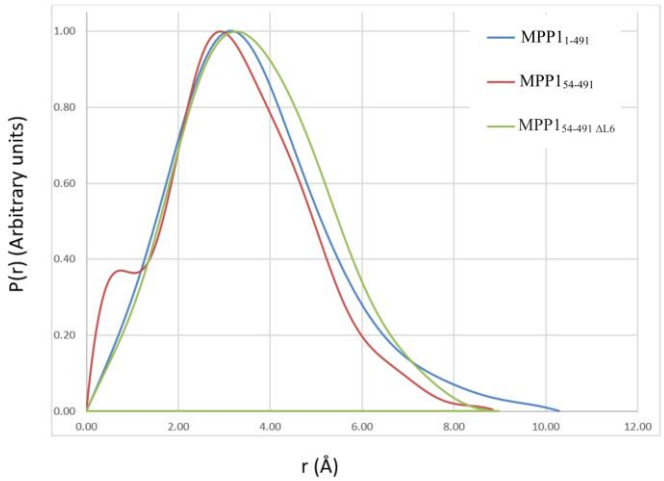


Figure 4-18. P(r) plots of $MPP1_{1\text{-}491}$, $MPP1_{54\text{-}491}$, and $MPP1_{54\text{-}491\ \Delta L\text{-}6}$. The P(r) plots were generated using ATSAS and represent distances between all possible pairs of atoms within the molecule.

Three distinct MPP1 homology models representing $MPP1_{1\text{-}491}$, $MPP1_{54\text{-}491}$, and $MPP1_{54\text{-}491\ \Delta L6}$ were generated by Dr. Nikos Pinotsis (Protein Crystallography and Biophysics, Birkbeck College, London, England) based on KifC1 and MKLP2 models (PDB entries 5WDH and 5ND2). The quality of the homology models was analysed using *Crysol* and *Sreflex*, which compare the SAXS scatter from macromolecules in solution with that in the given models. Among the three distinct MPP1 homology models, $MPP1_{1\text{-}491}$ showed the best fit with the SAXS scatter curve, with a *Crysol* x^2 value of 3.1.

The *ab initio* bead models of the three distinct MPP1 constructs were then generated using *Dammif* and further improved using *Dammin* in the ATSAS suit. The final x^2 values of the $MPP1_{1\text{-}491}$, $MPP1_{54\text{-}491}$, and $MPP1_{54\text{-}491\ \Delta L6}$ dummy models were 1.06, 1.13, and 1.11, respectively. These low x^2 values indicated optimal fits of the dummy models to the SAXS scatter curves. The *ab initio* models of the $MPP1_{1\text{-}491}$ and $MPP1_{54\text{-}491}$ constructs showed similar shapes, whereas the $MPP1_{54\text{-}491\ \Delta L6}$ bead model was larger and significantly different in shape.

Finally, the MPP1 homology models were superimposed onto the *ab initio* bead models using *SASpy* [116]. The normalised spatial discrepancy (NSD) for $MPP1_{54\text{-}491}$ is 1.5, indicating that the homology model fits properly into the *ab initio* model. The NSD value for $MPP1_{1\text{-}491}$ is 2.1, which is slightly higher because of the missing densities of the N-terminal residues (residues 1 to 53). In contrast, $MPP1_{54\text{-}491\ \Delta L6}$ forms artificial dimers in solution, as indicated by the *R(g)* and *P(r)* plots and the shape of the *ab initio* model. The homology $MPP1_{54\text{-}491\ \Delta L6}$ model was subjected to *ClusPro* analysis for protein-protein docking [122]. The dimer models generated using *ClusPro* were then fitted onto the *ab initio* bead model

using *SASpy*. However, even the best fit had an NSD value of 2.5, indicating a large discrepancy between the homology and *ab initio* models. The $MPP1_{54\text{-}491\ \Delta L6}$ homology model may not represent the correct conformation in dimers.

The $MPP1_{1\text{-}491}$ homology model showed a typical kinesin fold comprising an eight-stranded antiparallel β-sheet surrounded by three major helices on either side (Figure 4-19 A, C). Its N-terminal region is composed of 54 residues (residues 1 to 54), but its function remains unknown. The N-terminal region of $MPP1_{1\text{-}491}$ was missing in the *ab initio* model generated from SAXS data, indicating that it is mostly disordered. The $MPP1_{54\text{-}491}$ model is similar to the $MPP1_{1\text{-}491}$ model but without the first 54 residues. In these two models, the large loop L6 insertion emerges from the α2 helix, interacts with α1, α2, α3, and L8, and is finally connected with β4. The large insertion forms intramolecular interactions with multiple secondary structure elements and may function in stabilising the structure.

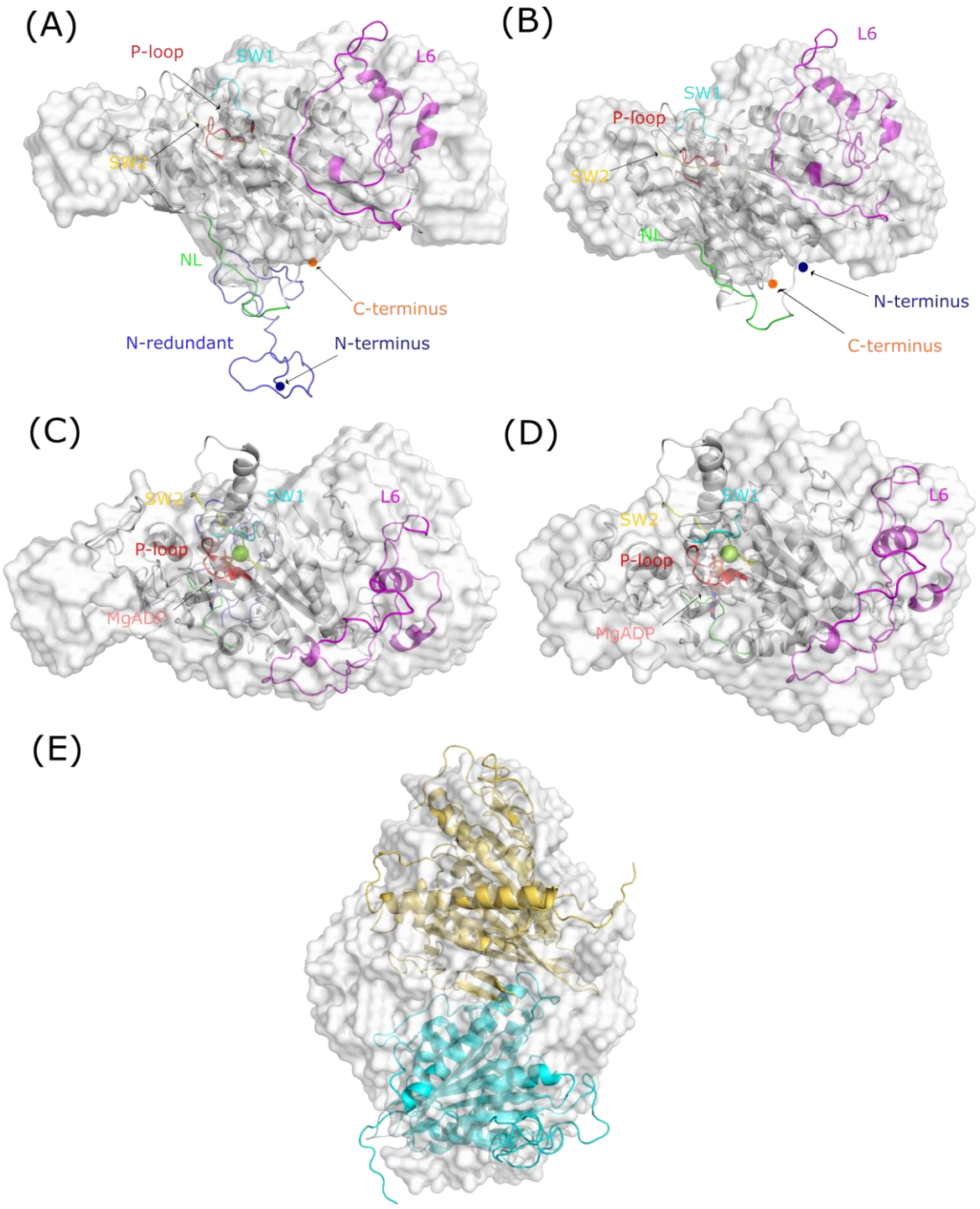


Figure 4-19. MPP1 homology models (dark grey) superimposed onto their ab initio models (transparent white). (A) The MPP1$_{1-491}$ homology model fit onto the ab initio model with an NSD value of 2.1. The large normalised spatial discrepancy (NSD) is caused by the missing densities of the N-terminal residues 1-53 (blue). The structural features are marked as follows: P-loop (red), switch I (cyan), switch II (yellow), loop L6 (magenta), neck linker (green), N-

terminus (dark blue), and C-terminus (orange). (B) The MPP1$_{54-491}$ homology model fit properly onto the ab initio model, with an NSD value of 1.5. (C) MPP1$_{1-491}$ model from the top view. (D) MPP1$_{54-491}$ model from the top view. (E) The MPP1$_{54-491\ \Delta L6}$ ab initio model is larger and significantly different from that of MPP1$_{1-491}$ and MPP1$_{54-491}$, indicating the formation of MPP1$_{54-491\ \Delta L6}$ dimers. However, the dimer models generated using ClusPro have NSD values larger than 2.5, indicating substandard fits. Two identical MPP1$_{54-491\ \Delta L6}$ molecules are marked in yellow and cyan.

The MT binding site of MKLP-2, another kinesin-6 family member, has been described by Atherton *et al.* [53]. They solved the structure of the MT-bound MKLP-2 (*Bos taurus* origin) motor domain in the presence of ADP at a resolution of 5.8 Å using Cryo-EM (PDB entry 5ND2). The MT binding region of MKLP2 is composed of L2, L7, β5/L8, α4/L11 (switch II), α5, α6, and L12 domains (Figure 4-20 A). The MT-proximal domain of MPP1 homology models shares a high similarity with MKLP2, indicating a similar interaction footprint between MPP1 and MTs. Alignment of the MPP1$_{1-491}$ homology model with the MT-MKLP-2 model shows that aside from the similar footprint, the redundant N-terminus of MPP1 forms a distinct domain that displays additional interactions with β-tubulin within the MT protofilament. This assumption is also supported by the biochemical data previously published by our group [59]. In MT-stimulated ATPase assays, the $K_{0.5,\ MT}$ of MPP1$_{2-477}$ was 102.0 ± 3.2 μM, which was slightly lower than the $K_{0.5,\ MT}$ of MPP1$_{57-491}$ (112.0 ± 1.0 μM), indicating that the MPP1 construct with the N-terminus has a higher affinity for MTs [59].

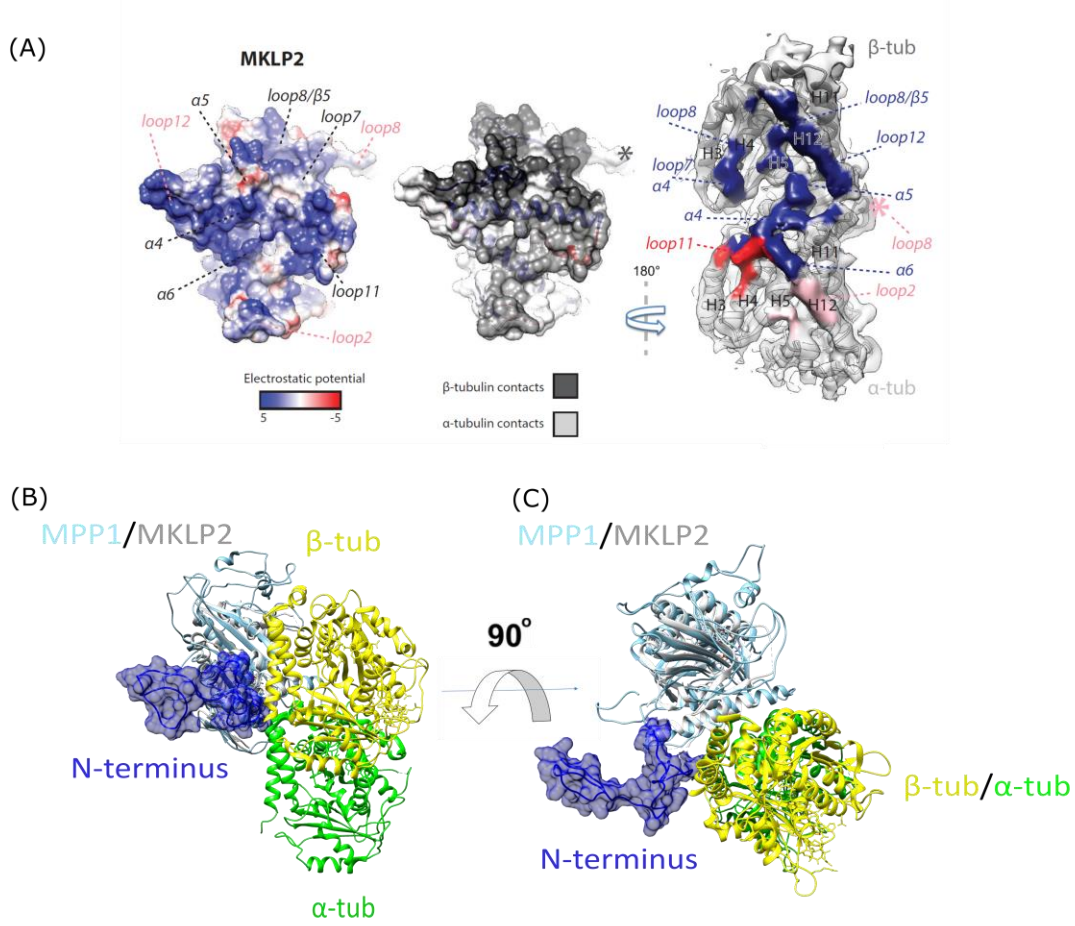


Figure 4-20. (A) MT-MKLP2 footprint at the binding interfaces on both MKLP2 and α/β tubulins (PDB entry 5ND2). The MT binding region of MKLP-2 includes L2, L7, β5/L8, α4/L11 (switch II), α5, α6, and L12. Figure adapted from Atherton et al. [53]. (B) Alignment of MPP1$_{1-491}$ homology model (light blue) to the MT-MKLP2 model (grey). The MPP1$_{1-491}$ homology model shows a high similarity with the MKLP2 model. Aside from similar interactions between the kinesin and MTs, the N-terminus of MPP1 (blue) forms a distinct domain that interacts with β-tubulin (yellow). The contour of the N-terminus was calculated using Chimera. (C) Top view of the alignment of the MPP1$_{1-491}$ homology model with the MT-MKLP2 model.

In the MPP1$_{1-491}$ and MPP1$_{54-491}$ *ab initio* models, the density of loop L6 forms a distinct domain emerging from core β-sheets and facing the MT plus end. Although the density of loop L6 spreads adjacent to the MT-binding region, loop L6 itself does not appear to contact the MT surface (Figure 4-21 C). Atherton *et al.* reported that loop L6 of MKLP-2 connects the core β-sheet and β5a/b of the

tubulin-binding subdomain, which is an effective part of the Switch I/II subdomain (Figure 4-21 A). The density of loop L6 rotates by a small angle during transition from the MKLP-2 MT-NN state (MT binding, absence of nucleotide, PDB entry 5ND7) to the MKLP-2 MT-ADP.AlFx state (MT binding, ATP analogue binding, PDB entry 5ND4) [53] (Figure 4-21 B). In all three MKLP2 models (absence of nucleotide, ADP binding, and ATP analogue binding), a 'sausage' density exists at the N-terminus of loop L6. A 4-turn helix can be fitted into this density (Figure 4-21 A, B). Interestingly, this helix is also observed in the X-ray model of Zen4 (PDB entry 5X3E) [52], an MKLP-1 homologue, and the $MPP1_{1\text{-}491}$ and $MPP1_{54\text{-}491}$ homology models, although the sequence alignment of kinesin-6 members indicates a very low similarity with the loop L6 region (Figure 4-22). The 4-turn αL6 helix within loop L6 seems to be conserved throughout the kinesin-6 family (Figure 4-21 D). In contrast, the C-terminus of loop L6 is mostly disordered. The electron density of the C-terminus of loop L6 is missing in both X-ray (Zen4) and Cryo-EM (MKLP-2) structures [52, 53].

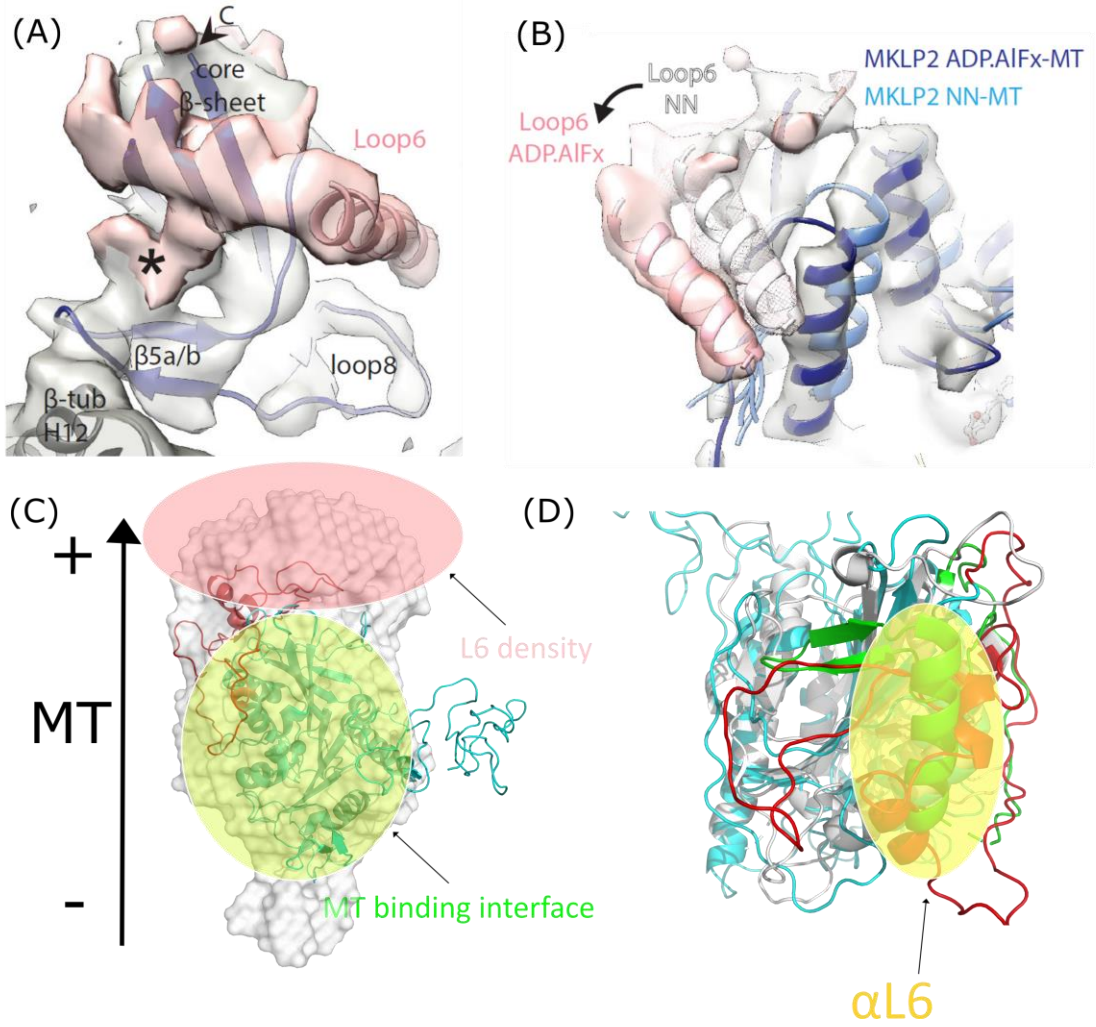


Figure 4-21. (A) Loop L6 of MKLP-2 connects the core β-sheet and β5a/b of the tubulin-binding subdomain. A 'sausage' density exists at the N-terminus of loop L6, which can be fitted with a 4-turn helix (B) Loop 6 is an effective part of Switch I/II subdomain, which rotates during ATP binding. Figure adapted from Atherton et al. [53]. (C) The density of loop L6 in MPP1 homology models forms a distinct domain emerging from the core β-sheets and facing the MT plus end, adjacent to the MT-binding region. (D) Alignment of the $MPP1_{1-491}$ homology model (cyan) with Zen4 (grey). Loops L6 of these models are marked in red ($MPP1_{1-491}$) and green (Zen4). An α-helix at the N-terminus of L6 (αL6) is observed in all three kinesin-6 members.

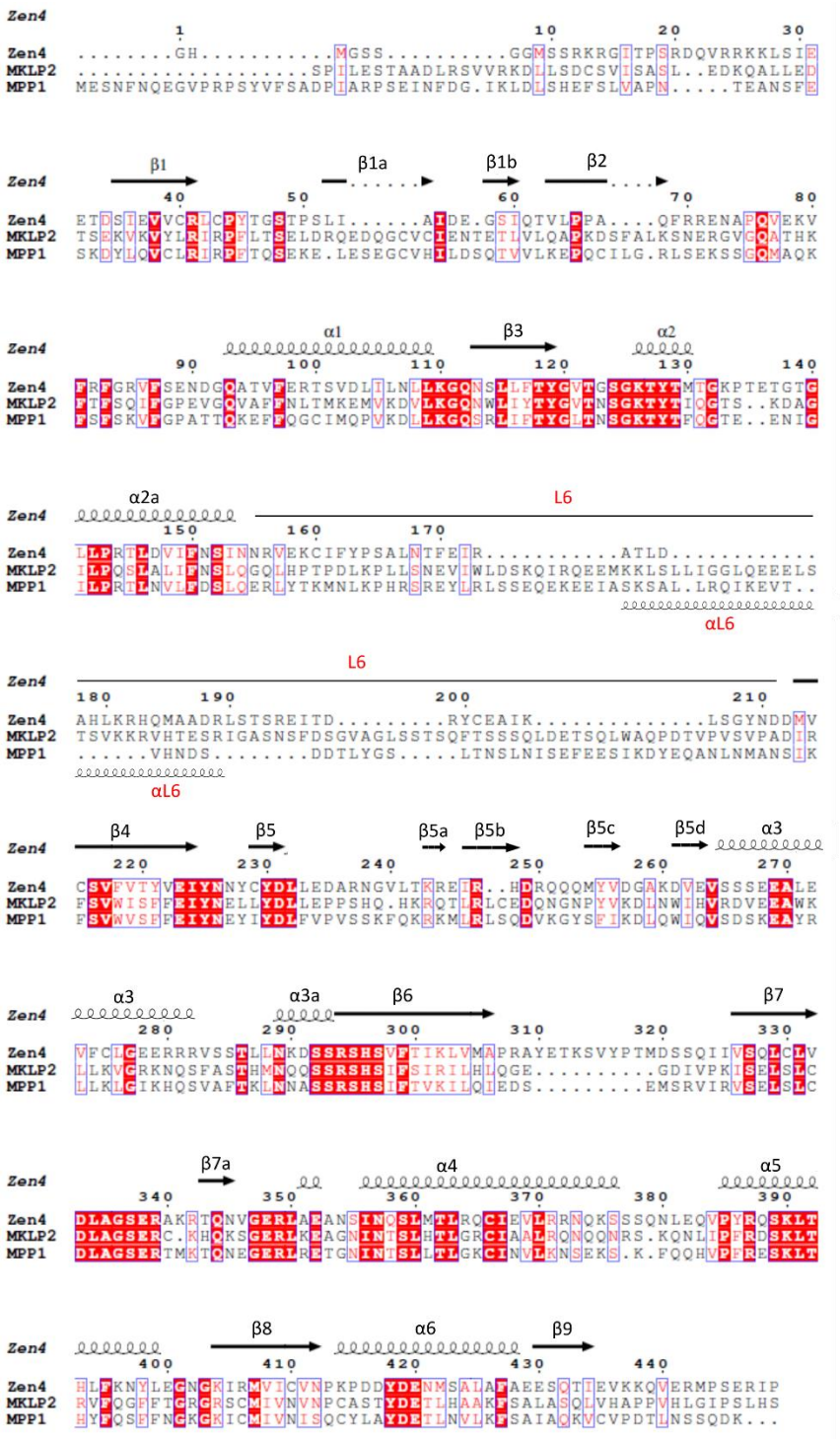


Figure 4-22. Structural sequence alignment of kinesin-6 family member motor domains. Although Zen4, MKLP-2, and MPP1 share relatively low sequence similarities, especially around the loop L6 region, a conventional helix αL6 (red) is observed at the N-terminus of loop L6 in all three kinesins.

Unfortunately, the *ab initio* models generated from SAXS data only represent the general shape of MPP1 constructs in solution and cannot provide insights into the structural and mechanochemical details of the MPP1 motor domain. The *ab initio* model as well as the homology models provide an opportunity to observe the MPP1 motor domain. A more detailed X-ray model is still needed to investigate the conformational rearrangements at the catalytic and neck linker site, which is crucial for understanding MPP1 mobility and the development of MPP1 inhibitors.

4.1.5 Discussion

4.1.5.1 Challenges in obtaining the crystal structure of human MPP1 motor domain

MPP1 is an interesting therapeutic target owing to its exclusive role in cytokinesis and apparent clinical relevance in bladder cancer and hepatocellular carcinoma [60, 123-125].

Currently, there are five PDB entries for the kinesin-6 motor domain, which include the MT-bound MKLP-2 motor domain in the presence of Mg^{2+}ADP (*B. taurus*, Cryo-EM, 5.80 Å, PDB entry 5ND2), the MT-bound MKLP-2 motor domain without nucleotide (*B. taurus*, Cryo-EM, 6.10 Å, PDB entry 5ND3), the MT-bound MKLP-2 motor domain in the presence of ADP.AlFx (*B. taurus*, Cryo-EM, 4.4 Å, PDB entry 5ND4), the MT-bound MKLP-2 motor domain in the presence of the slowly hydrolysable adenylyl-imidodiphosphate (AMPPNP) (*B. taurus*, Cryo-EM, 7.9 Å, PDB entry 5ND7), and Zen4 in the apo state (*C. elegans*, X-ray diffraction, 2.6 Å, PDB entry 5X3E) [52, 53]. These publications

report a significant divergence between kinesin-6 members and conventional kinesin-1, one of the most investigated kinesins with respect to mechanochemistry. In conventional kinesin-1, conformation of the catalytic site and core β-sheets is significantly different in the ADP-binding and nucleotide-free states. In the ADP-binding state, switch I transforms into an ordered α helix. A critical salt bridge between switch I (Arg203, SSRSH) and switch II (Glu236, DLAGSE) unlatches, and switch II adopts an 'open' conformation to release ADP from the binding site. Furthermore, helix α4 extends by two more turns and rotates away from the core β-sheets for the neck-linker docking [113]. In a nucleotide-free state, switch I 'melts' into a disordered loop and forms a salt bridge with switch II in preparation for ATP binding. The neck-linker docking site is also blocked owing to the rearrangement of α4. However, in the MKLP-2 structures, no significant differences were observed in the MT-ADP and MT-NN (nucleotide-free) models at the catalytic site. Switch I, switch II, and the onset of helix α4 are generally disordered and flexible throughout the ADP-NN transition [39]. Moreover, in kinesin-1, the neck-linker docks to the core β-sheet upon ATP binding, thereby causing kinesin to step forward. In contrast, docking of the neck-linker is not visible in the MKLP-2 MT-ADP.AlFx (non-hydrolysable ATP analogue) model. Unfortunately, none of the kinesin-6 motor domains from *H. sapiens* have been determined yet, and alignment of the MPP1 motor domain sequence with that of MKLP-2 and Zen4 shows a low sequence similarity (43.8%) (Figure 4-22). Thus, acquiring crystals of human MPP1 is rather challenging.

We conducted SAXS measurements for three distinct MPP1 constructs, $MPP1_{1–491}$, $MPP1_{54–491}$, and $MPP1_{54\text{-}491\ \Delta L\text{-}6}$, and determined the nanoparticle size distribution (*R(g)*) as well as resolved the size (*Dmax*) and shape of MPP1 in solution. Three *ab initio* models were generated from the SAXS data and aligned

with MPP1 homology models. The *ab initio* models indicated that $MPP1_{1–491}$ and $MPP1_{54–491}$ are monomers in solution, whereas $MPP1_{54\text{-}491\ \Delta L\text{-}6}$ unexpectedly forms dimers. The $MPP1_{54–491}$ homology model fits best into the *ab initio* model, with a x^2 value of 1.5. The x^2 value of the $MPP1_{1–491}$ homology model is 2.1 owing to missing density at the N-terminus, suggesting that this N-terminal region (54 residues) is disordered. The models of the $MPP1_{54\text{-}491\ \Delta L\text{-}6}$ dimer were predicted using *ClusPro* [122], but none of them fit properly into the *ab initio* model, indicating that the $MPP1_{54\text{-}491\ \Delta L\text{-}6}$ homology model does not represent the correct dimer conformation.

Alignment of the $MPP1_{1\text{-}491}$ homology model with the MKLP-2 MT-ADP model indicated that the redundant N-terminus of $MPP1_{1\text{-}491}$ forms a distinct domain adjacent to the MT-binding interface. The N-terminus may interact with the MT protofilament. The published $K_{0.5,\ MT}$ value of the short $MPP1_{2\text{-}477}$ construct is slightly lower than that of the long $MPP1_{57\text{-}491}$ construct, which further supports this assumption [59].

Moreover, alignment of the MPP1 homology models with MKLP-2 and the MKLP-1 orthologue Zen4 models indicates the existence of an α-helix (αL6) at the N-terminus of loop L6 in all kinesin-6 members despite significant differences in their protein sequences. The helix αL6 moves away from the catalytic core of MKLP-2 during ATP binding, suggesting that it is an effective part of the Switch I/II subdomain [53].

Although SAXS data indicate the general shape of the MPP1 structure, a crystal structure of MPP1 is still needed for mechanochemical analysis and drug development, particularly structure-based drug design. Unfortunately, our efforts

to acquire MPP1 crystals failed. Dr. Sandeep Talapatra also tried a reductive methylation approach, which could promote crystallisation via improved crystal packing; however, this approach also failed [58].

The key to acquiring MPP1 crystals is to find a method for stabilising MPP1, especially the highly disordered loop L6 region. Thus, a co-crystallisation strategy with αRep proteins was chosen. αReps function as molecular chaperones, which bind to and stabilise the target. αReps are artificially synthesised proteins that mimic heat-shock proteins (HSP), which bind specifically to the target proteins during multiple cellular processes. αReps bind to the target protein and stabilise it, thereby increasing the possibility of generating stable complexes. This strategy has proven effective, as seven different structures of αReps/target complexes with unrelated target structures have been solved and three of them have been obtained for proteins that previously resisted all crystallisation efforts (FNE and Octarellin) [111, 126-129]. Remarkably, when Van der Waals forces alone cannot form solid interactions between the αRep and the target, artificial peptide linkers can be introduced to enhance these interactions and foster the crystallisation process [111].

The αRep binders for MPP1 motor domain constructs were selected via phage display and further ranked by affinity. Afterwards, initial co-crystallisation trials were conducted. However, only crystals of αReps were acquired. Apparently, MPP1 only functioned as a precipitant, which pushed the αReps concentration to supersaturation phase. A possible solution is to covalently link αReps to the N-terminus of MPP1 using a linker [111]. Chevrel *et al.* claim that by linking αReps to the N-terminus of the target protein, the expression yield of the target protein can be increased [63].

Thus, two fusion protein constructs, αRep G8-MPP1$_{54-491}$ (short) and αRep G8-MPP1$_{54-491}$ (long), were designed. The short fusion protein was anticipated to form dimers. As the linker is relatively short, the αRep domain of one molecule could bind to the MPP1 domain of another molecule through intermolecular interactions. In contrast, the long fusion protein was anticipated to form monomers. The linker was long enough to provide adequate flexibility for flipping the αRep domain toward the MPP1 domain and form intramolecular interactions [111] (Figure 4-23).

The G8-MPP1$_{54-491}$ (short) fusion protein has been purified and subjected to crystallisation trials (23rd June 2022). We have acquired several crystals and now preparing to send them for measurements.

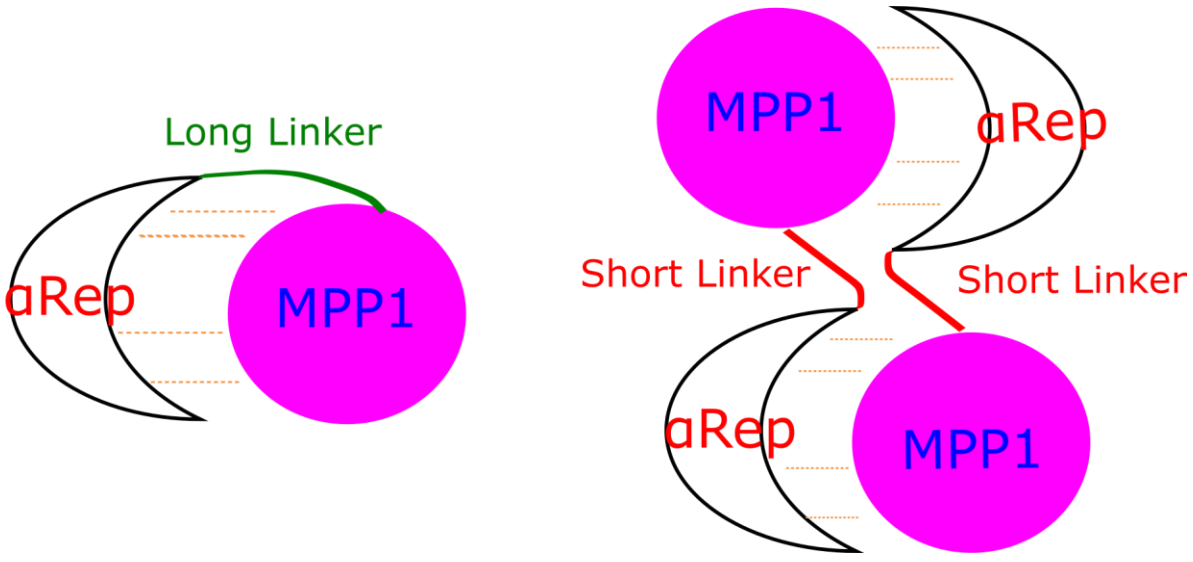


Figure 4-23. Schematic illustrating potential intramolecular and intermolecular interactions within/between the αRep/MPP1 fusion proteins. αRep is shown as a crescent shape; MPP1 is shown as a purple sphere; the short linker is coloured in red; and the long linker is coloured in green.

4.1.5.2 Development of inhibitors against human MPP1

MPP1 was first identified by screening a subset of proteins specifically phosphorylated at the G_2/M transition using the MPM2 monoclonal antibody [130]. The MPP1-depleted cells showed cytokinesis failure. Either the midbody regressed with the appearance of a binucleated cell or the midbody persisted and the two ill-separated daughter cells finally underwent apoptosis [123]. MPP1 has a specific pattern of localisation and expression during the cell cycle. During the interphase, MPP1 is mostly located in the nucleus and is well expressed in G_2. MPP1 is diffused throughout the cytoplasm during the metaphase and subsequently associates with the central spindle and the midbody at the end of mitosis [123].

Although the detailed mechanism of how MPP1 functions during cytokinesis remains unknown, several protein regulators have been found to interact with MPP1 during the cell cycle. Transition from the G_2 to M phase is triggered by the nuclear import of cdc2 kinase, which then phosphorylates a set of mitotic proteins on the Ser/Thr-Pro motif, including MPP1. Peptidyl-prolyl isomerase (Pin1) then preferentially binds to the phosphorylated Ser/Thr-Pro motif through the N-terminal WW domain, which comprises two highly conserved tryptophan amino acids separated by ~20 residues, and catalyses the peptidyl-prolyl isomerisation of MPP1 through the C-terminal isomerase domain [57]. In bladder cancer cell lines, both MPP1 and protein regulator of cytokinesis 1 (PRC1) are upregulated. MPP1 interacts with PRC1 and forms a complex, which is likely to play a crucial role in bladder carcinogenesis [60]. Moreover, MPP1 is stabilised by forkhead box protein O38 (FOXO38), a poorly characterised F-box protein. FOXO38 is further regulated by ubiquitin specific protease 7 (USP7). Depletion of USP7

significantly increases the percentage of multinucleated cells, which can be rescued via the upregulation of MPP1 or FBXO38 [131].

MPP1 plays roles in some tumours, including hepatocellular [124, 125], bladder [60], colorectal [132], breast [133], renal cell [134], and pancreatic cancers [135], and is considered a candidate therapeutic target in hepatocellular carcinoma and bladder cancer. To date, only two MPP1 inhibitors, depsidone analogues norlobaridone (**D1**) and physodic acid (**D2**), have been reported [59]. Both **D1** and **D2** are low micromolar inhibitors of MPP1. Proliferation assays showed that **D1** and **D2** displayed a measurable inhibitory effect in a cell line panel, including HCT116 (colon), BxPC3 (pancreatic), K562 (leukaemia), as well as J82 and UM-UC-3 cells (both bladder cancer), with EC_{50} values ranging between 31.8 and 48.3 μM [59]. However, both compounds showed high clearance and low half-lives in human and mouse microsomal stability assays, indicating a need for further improvement of the scaffold to overcome these issues.

In this study, we investigated and improved the two ascochitine and depsidone-based MPP1 inhibitors. Ascochitine is an MPP1 hit identified by Dr. Sandeep Talapatra. In the basal ATPase assay, ascochitine reduced the activity of $MPP1_{1-491}$ to 85%, with an IC_{50} value of 19.8 ± 3.3 μM. Interestingly, a commercially available chemically related ascochitine-like analogue, citrinin, caused no inhibition of $MPP1_{1-491}$, implying that despite a similar overall structure, the four minor changes in citrinin completely abolished MPP1 inhibition. To further identify whether the loop L6 region is involved in the binding of ascochitine and depsidone analogues, three distinct MPP1 constructs, $MPP1_{1-491}$, $MPP1_{54-491}$, and $MPP1_{54-491\ \Delta L-6}$, were characterised using both basal and MT-stimulated ATPase assays in the presence of ascochitine and depsidone analogues, **D1** and **D2**. Finally, no significant difference in IC_{50} values was observed between the MPP1

constructs with and without the loop L6 region, indicating that the featured loop L6 is not involved in the binding of ascochitine and depsidones.

Fifteen MPP1 inhibitor analogues were ordered from the company Sai. Owing to major difficulties in synthesising these analogues, especially in the last cyclisation step of the long synthetic route, only thirteen analogues could be finally synthesised, including seven intermediates (**A4a** to **A4g**) and six ascochitine analogues (**A5b, A5c, A5d, A5e, A5i,** and **A10**). The synthesised analogues were then subjected to basal ATPase activity measurement. None of the intermediates (**A4a** to **A4g**) caused the inhibition of MPP1, indicating that the two-ring core scaffold is essential for inhibition. The most active analogues, **A5d** and **A5e,** had IC_{50} values of 4.2 ± 0.6 and 5.6 ± 0.6 μM, respectively; both were approximately 2-fold more active than the parent compound ascochitine.

Analogues **A5d** and **A5e** were subjected to drug-like property tests. Both showed chemical properties that met Lipinski's rules (MW < 500 Da, number of hydrogen-bond donors (HBD) < 5, number of hydrogen-bond acceptors (HBA) < 10, and cLogP < 5) and acceptable permeability in the Caco-2 cell layer assay [136]. In the PAMPA assays, **A5d** showed no indication of efflux, whereas **A5e** was identified as a potential substrate for efflux transporters [137]. Meanwhile, they showed no significant inhibition of P450 enzymes and hERG, indicating a low metabolic or cardiac risk. However, both **A5d** and **A5e** showed high clearance and low half-lives in human and mouse microsomal stability assays, indicating that they are highly unstable under physiological conditions. This conclusion was also supported by plasma protein binding assays, wherein neither **A5d** nor **A5e** was detectable in the buffer or plasma compartment.

The cell proliferation experiments were performed using analogues **A5d** and **A5e**. Compared with that in *in vitro* assays, the analogues demonstrated significantly lower inhibitory effects, probably because of their instability under physiological conditions. Contrary to the results published in previous reports, no significant differences were observed in MPP1 expression levels between bladder cancer cell lines and HeLa cells [60]. Moreover, only weak synergistic effects were observed when using **A5e** with paprotrain. These results indicate weak synergistic effects when using both inhibitors targeting different aspects of cytokinesis. A recent study reported that reducing MPP1 expression using shRNAs sensitises hepatocellular carcinoma cells to MT-targeting agents (MTAs), including taxol, epothilone B, and vincristine, by blocking cytokinesis [124]. This suggests that further proliferation experiments could be performed to investigate the synergistic effects of MPP1 inhibitors with MTAs, but more potent MPP1 inhibitors are needed first.

Three metabolites were identified during subsequent metabolite profiling, with oxidation initially occurring at the n-propyl and methyl substituents rather than in the ring system.

As described previously, ascochitine analogues are difficult to synthesise through a multistep procedure comprising up to 11 steps. To speed up the syntheses of MPP1 analogues, a pharmacophore-based search was conducted, suggesting that the core structures of chromenone-3-carboxylic acid and quinolone-3-carboxylic acid would be more synthetically available isosteres of ascochitine. The synthetic route could be significantly reduced to 2 or 3 steps. However, the chromenone-3-carboxylic acid and quinolone-3-carboxylic acid analogues showed no inhibitory effects in the basal ATPase assays. Inhibition of the MPP1 constructs could only

be observed in MT-stimulated ATPase assays, implying that these analogues only inhibit MPP1 in an MT-bound state. Moreover, the chromenone-3-carboxylic acid and quinolone-3-carboxylic acid analogues were five to eight-fold less active than the ascochitine analogue **A5d**, indicating that they are not suitable for further development.

The other set of three analogues was developed based on the depsidone scaffold. Depsidones **D1** and **D2** have been investigated and reported by Talapatra *et al* [59]. We obtained three additional depsidone analogues for initial SAR analysis from Dr. Joëlle Prunet. However, all of them caused no inhibition of $MPP1_{1\text{-}491}$ in the basal ATPase assay, which indicated that the 12-pentanoyl and 4-pentyl substituents are crucial for inhibitory activity.

In summary, we investigated analogues for two MPP1 inhibitors, among which analogues **A5d** and **A5e** were found to be the most active compounds. Although these compounds demonstrated good inhibitory activities in ATPase assays, it may be challenging to develop these into potential drug-like molecules because of several drug-like property liabilities, including high metabolic clearance and low half-life. However, they could be developed into tool compounds to study the function of MPP1 in tumour cell lines. The significantly low inhibitory activities of MPP1 analogues in cell-based assays compared with those *in vitro* also support this conclusion. Further investigations could be conducted into the synergistic effects of **A5d** and **A5e** combined with MTAs. Meanwhile, a crystal structure of the MPP1-ligand complex could help elucidate the interactions and foster inhibitor development.

4.2 Humanised A33 Fab

4.2.1 Determination of two A33 Fab structures

The crystallisation conditions for humanised A33 Fab were identified using the low ionic strength screen kit (Hampton Research). Crystals grew in two different space groups. Triclinic crystals (Space group P1, 2.3 Å resolution) grew from 4 µl of a 20 mg/ml protein solution mixed with 2 µl of 50 mM glycine (pH 9.0) and 5 µl of 16% w/v PEG3350 at 18 °C. Hexagonal crystals (Space group $P6_5$, 2.2 Å resolution) grew from 4 µl of a 20 mg/ml protein solution mixed with 2 µl of 50 mM citric acid (pH 4.0) and 5 µl of 8% w/v PEG3350 at 18 °C. The crystallographic statistics are presented in Table 4-11.

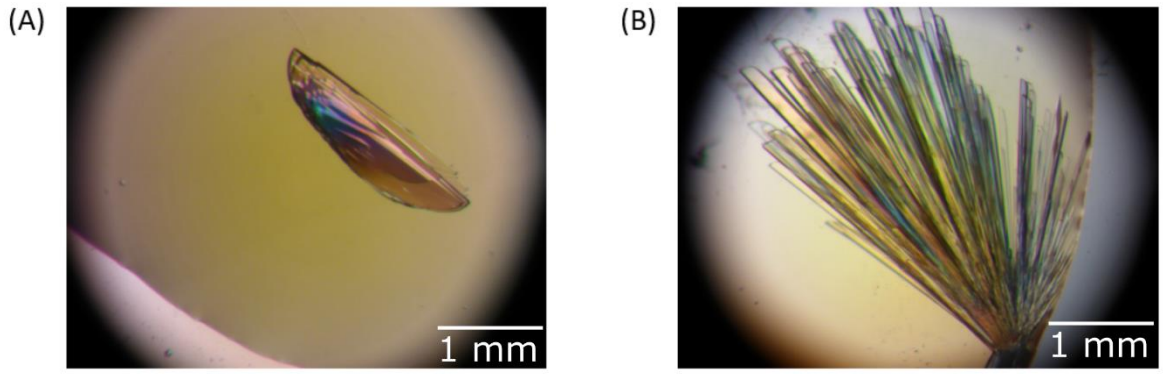


Figure 4-24. Fab crystals in (A) the triclinic space group and (B) hexagonal space group.

Table 4-11. Data collection and structure refinement statistics for A33 Fab in the space groups P1 and $P6_5$. Values for the outer shell are indicated in parentheses.

Crystal Form	*Triclinic*	*Hexagonal*
Data collection		
Wavelength (Å)	0.9763	0.91188
Space group	*P*1	*P*6_5
a, b, c (Å)	37.67, 69.73, 89.85	93.83, 93.83, 201.74
α, β, γ (°)	99.71, 101.43, 105.66	90, 90, 120
Resolution range (Å)	85.59-2.30 (2.38–2.30)	75.37-2.20 (2.28–2.20)
No. of reflections	106776 (7689)	202208 (26490)
Unique reflections	35070 (2710)	27308 (2676)
Multiplicity	3.0 (2.8)	7.4 (6.8)
Completeness (%)	94.0 (72.9)	98.4 (90.3)
⟨I/σ(I)⟩	11.0 (5.7)	12.1 (2.1)
R_{merge} *(%)*	5.6 (14.5)	9.6 (107.8)
R_{meas} *(%)*	6.8 (17.7)	10.3 (116.7)
Refinement		
Overall Wilson B ($Å^2$)	31.8	36.7
R.m.s. deviations		
Bond lengths (Å)	0.012	0.011
Bond angles (°)	1.41	1.29
No. of atoms	6612	3338
Protein non-hydrogen atoms	6608	3342
Water molecules	267	166
B factor ($Å^2$)	34.6	54.8
Protein	34.7	55.1
Water	32.1	48.1
R_{work}/R_{free} *(%)*	19.4/27.1	20.5/25.6
Ramachandran		
Preferred, allowed, outliers (%)	94.4, 5.0, 0.6	93.1, 5.2, 1.7

As a humanised Fab fragment, A33 Fab comprises a γ heavy chain and a κ light chain. Each chain contains a variable domain (V_L and V_H) and constant domain

(C_L and C_H1). The variable framework sequence is derived from the human antibody LAY and is substituted with murine CDRs, whereas the constant domain is fully derived from mammalian cells [73]. Slight differences were identified in the elbow angles of the solved structures. The elbow angles, which are defined as the intersection angles of two pseudo-dyad axes (PDAs) between the variable domain and constant domain, of structures solved in the *P*1 and $P6_5$ space groups were 155.9° and 146.3°, respectively. The variation in elbow angles indicates the flexibility of the switch region [138].

A33 Fab showed a canonical β-sandwich Ig fold within four domains (V_L, V_H, C_L, and C_H1). Each domain had two layers of β-sheets—an inner and an outer β-sheet. One canonical disulphide bridge was identified in each domain between the β-sheet layers (between Cys_{22}-Cys_{96} in V_H, Cys_{23}-Cys_{88} in V_L, Cys_{144}-Cys_{200} in C_H, and Cys_{134}-Cys_{194} in C_L), which contributed to the stability of the Fab. An inter-chain disulphide bridge, Cys_{214}_C_L-Cys_{220}_C_H, was not resolved owing to the missing loop in the hinge region of the heavy chain. Seven inter-chain hydrogen bonds were identified between the heavy and light chains; these were V_HPro102-V_LTry36, V_HPhe103-V_LThr46, V_HTrp106-V_LThr46, V_HGln39-V_LGln38, V_HGln39-V_LGln38, C_HPro170-C_LSer162, and C_HPro126-C_LSer121. In addition to hydrophobic interactions, there are five inter-chain hydrogen bonds in the variable regions and only two in the constant domains, illustrating tighter contacts within variable regions (Figure 4-25).

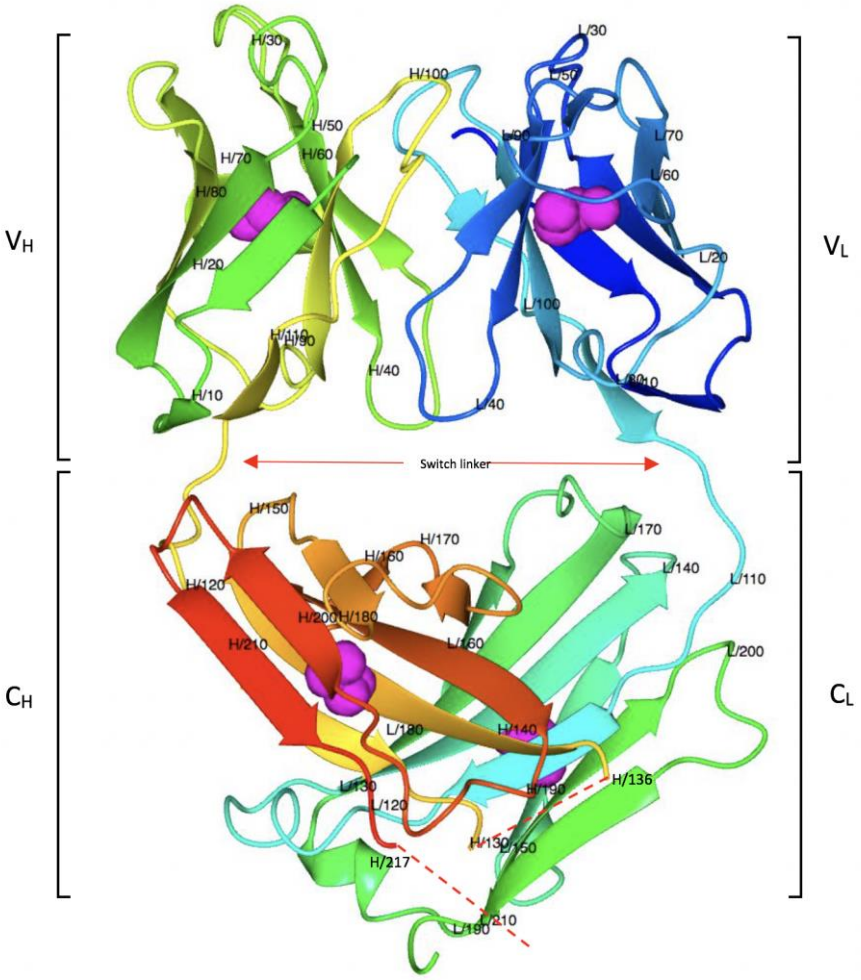


Figure 4-25. The overall structure of A33 Fab. Residue numbers are labelled at every ten amino acids. A33 Fab shows a canonical β-sandwich Ig fold within four domains (VL, VH, CL, and CH1), and each domain contains a disulphide bridge between the inner and outer β-sheets (magenta sphere). The missing loop regions are marked with a red dash.

4.2.2 Comparison of A33 Fab structures among different space groups

Triclinic crystals (*P1*) (PDB entry 7NFA) had two copies of A33 Fab per asymmetric unit (AU), which diffracted to 2.3 Å resolution, whereas the hexagonal crystals ($P6_5$) (PDB entry 7NC0) had only one copy per AU and diffracted to 2.2 Å. The overall structures in both crystal forms were similar, with slight differences （Figure 4-26 B). The main difference between the two

structures was the elbow angle. Different compact patterns affected the intersection angle between the variable and constant regions. The elbow angles shifted from 155.9° (triclinic form) to 146.3° (hexagonal form), indicating the flexibility of the switch region. However, the variable and constant regions shared a considerable structural similarity. The RMSD calculated between the C^{α} atoms of matched residues at the 3D superposition of the two structures was 1.33 Å. In contrast, the RMSD decreased to only 0.5 and 0.76 Å for separately superposed variable- and constant regions, respectively.

A closer insight into the linkers between the variable and constant regions revealed conformational differences in heavy chain switch peptides between the two models. The side chains of heavy chain Ser117 and Ser119 had different orientations and formed different hydrogen bonds with adjacent residues. Furthermore, two unique hydrogen bonds were formed within Ala118 and its adjacent residues in the hexagonal model. In contrast, the linker in the light chain was fixed by a set of hydrogen bonds, which conferred it less flexibility (Figure 4-26 B, C).

Two loops were missing in both the structural models, which indicated high flexibility in these regions. In the C_H domain, the missing loop comprised four residues (SKST, $_{heavy}$Ser132 to $_{heavy}$Thr135) in the $P6_5$ model and six residues (SKSTSG, $_{heavy}$Ser132 to $_{heavy}$Gly137) in the P1 model. The hydrophilicity of the missing loop could facilitate interactions with solvents, thereby contributing to the loop flexibility. The other missing loop was located at the C-terminus of the heavy chain, which is also the hinge region of A33 Fab. The missing loop in the hinge region contained 11 residues (KSCDKTHTSAA, $_{heavy}$Lys218 to $_{heavy}$Ala228).

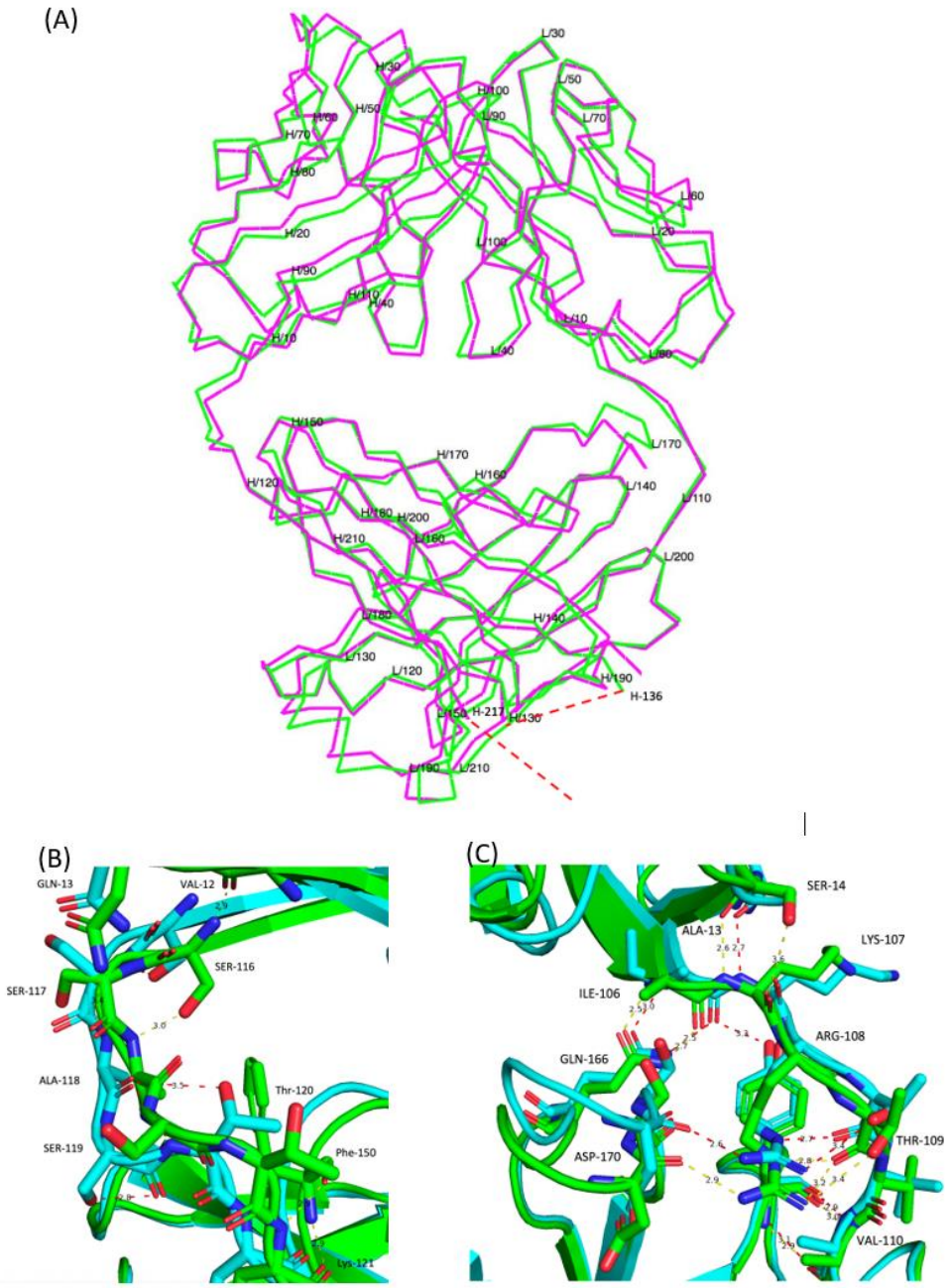


Figure 4-26. (A) Superposed Cα trace of A33 Fab structures in space groups P1 (magenta) and P65 (green). The missing insertions (heavySer132 to heavyThr135 and heavyLys218 to heavyAla228) are marked with red dash lines. Residues are labelled every 10 amino acids. The two structures share a high similarity but display differences in the switch region. (B) Superposed heavy chain switch region. The triclinic model is coloured in green, while the hexagonal model is coloured in cyan. Significant conformational differences can be viewed from heavySer117 to heavySer119, while the light chain switch peptides are fixed by a set of hydrogen bonds. (C) Superposed light chain switch region.

4.2.3 Elbow angles for the *P*1 and *P*6_5 structures obtained from simulations

As shown in Figure 4-27 A, B, the elbow angles in the P1 and $P6_5$ structures were initialised at 151° and 148.7°, respectively. In the P1 structure, this angle remained at around 150° for the first 15 ns before decreasing to 144° at 17 ns onwards. The P65 structure briefly dipped to 138° at 6 ns but returned to 145° from 15 ns onwards. From 20 ns onwards, the elbow angles in both the P1 and $P6_5$ structures fluctuated slightly in the range of 137°–146°. This suggested that at these simulation conditions (pH 7.0), the elbow-angle in both structures equilibrated rapidly and converged on the elbow angle observed in the $P6_5$ crystal form. Figure 3 shows the distribution of elbow angles for the two structures throughout their respective 100 ns simulations. Compared with the P65 structure, the P1 structure had a marginally high frequency at 135–140° and >150°. Nevertheless, the difference between the two structures was negligible.

In summary, the elbow angle simulation implies that the $P6_5$ structure resembles the average conformation sampled in the simulation more closely than the P1 structure, and therefore, it better represents the structure of A33 Fab likely to be observed in solution.

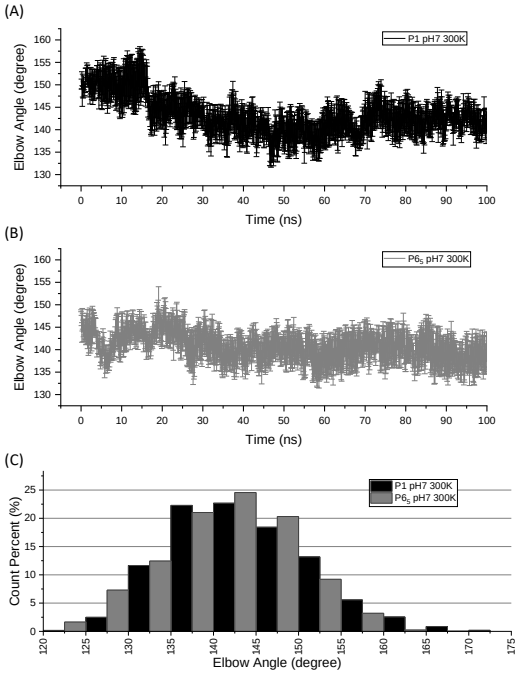


Figure 4-27. Elbow angles for the P1 and P65 structures obtained from simulations. The average elbow angles are shown for P1 (A) and P6$_5$ (B) structures during the 100 ns simulation at pH 7.0 and 300 K. Error bars represent standard errors of the mean (SEM). (C) Histogram for elbow angle distribution in the P1 and P6$_5$ trajectories obtained during the simulations. In all images, P1 and P6$_5$ are indicated in black and grey, respectively.

4.2.4 Humanised complementarity determining regions (CDRs)

A33 Fab is a protein with a highly humanised variable region, composed of a human variable framework from the LAY antibody and murine CDRs, to minimise the formation of human anti-chimeric antibodies (HACAs) [73]. The three murine CDRs responsible for antigen binding are located at the 'top' site of Fab structure (Figure 4-28).

Certolizumab, a peer therapeutic antibody, and A33 Fab have been developed from the same humanisation scaffold [139]. To analyse the structural differences between these, we aligned A33 Fab with Certolizumab for comparison. The constant regions of A33 Fab and certolizumab share a high similarity. The

calculated RMSDs for the C_H and C_L regions between A33 Fab and Certolizumab are 0.65 and 0.49 Å, respectively (missing insertions were excluded from the calculations). Despite large variations between their CDRs, their framework regions (FWRs) remained highly similar (Figure 4-28 B), indicating the rigidity of FWRs.

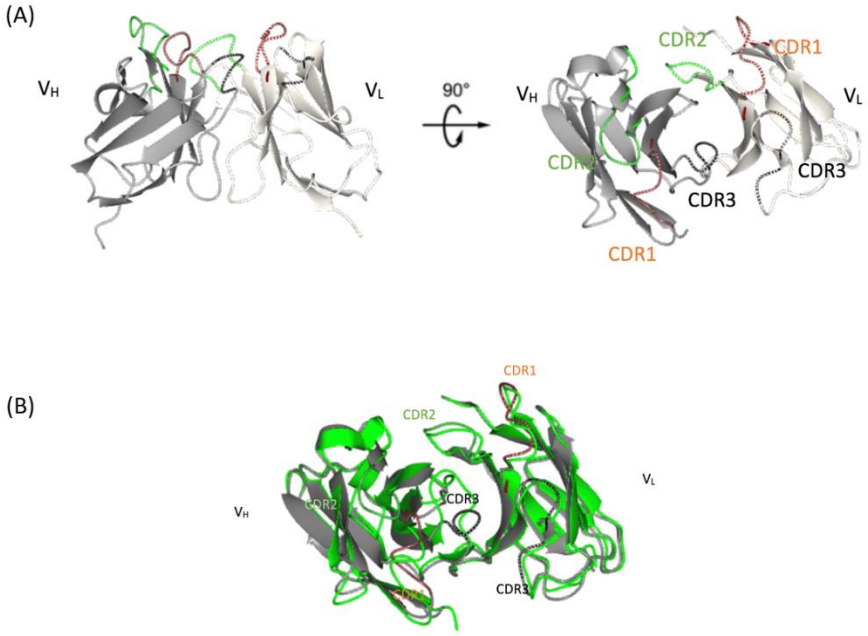


Figure 4-28. CDRs of A33 Fab. (A) CDR1 (brown), CDR2 (green), and CDR3 (black) originating from monoclonal murine A33 Fab are responsible for antigen binding. (B) Superposition of A33 FAB (grey) with certolizumab (green).

4.2.5 Discussion

Over the past decades, significant efforts have been undertaken for the development of therapeutic antibodies against CRC, targeting human antigen A33. Various A33 Fab variants have been developed, and multiple clinical trials have been undertaken [71, 140]. However, the normal gut localisation and intrinsic stability issues with these variants resulted in continuous failure in clinical trials [140, 141]. To date, the structure of A33 Fab or its variants has

never been reported. For the first time, the structure of an A33 Fab mutant, H/C226S, has been elucidated in two crystal forms.

The structural analysis was conducted on A33 Fab in two different crystal forms. The A33 Fab structures demonstrated the intrinsic rigidity of four separate regions (V_L, C_L, C_{H1}, C_L) and the flexibility of the switch region. The noticeable RMSD of the structures between the two crystal forms was identified because of differences in elbow angles between the two pseudo-dyad axes. The elbow angles varied from 156° (triclinic form) to 145° (hexagonal form) because of the different compact patterns and interactions with solvents. In contrast, the canonical β-sandwich Ig folds contributed to the rigidity of both variable and constant regions, leading to the high similarity among separate domains in different crystal forms. Subsequent simulations showed that the elbow angles of both structures were prone to equilibration and converged on an average of 145°, implying that the $P6_5$ structure is the most likely representation of the average solution structure of A33 Fab at pH 7.0.

To investigate the similarity between the elbow angles connecting the variable and constant domains, molecular dynamics analyses were performed on the full-residue homology models of *P1*, $P6_5$, and certolizumab. Principal Component Analysis (PCA) was performed by Dr. Cheng Zhang to capture the major collective motions throughout the simulation (SI). The PC1 motion demonstrated elbow bending, whereas all the three structures shared overlapped distribution when their trajectories were projected onto the PC1 eigenvectors. This implies that the three structures, despite minor sequence differences, demonstrate similar dynamics in the elbow bending movement.

The hypothesis was that the major differences between A33 Fab and Certolizumab are observed in the variable region. A33 Fab was developed from the same humanisation scaffold as certolizumab, possessing a similar variable region framework and the same constant region protein sequence. Interestingly, the variable regions of these two Fabs shared greater similarities than expected (RMSD 0.91 Å). Notwithstanding the different protein sequences, only the V_H CDR3 and V_H CDR2 of both Fabs varied significantly, whereas the remaining CDRs showed similar conformations, implying the key roles of V_H CDR3 and V_H CDR2 in the target identification process. Although their sequences were different, their FWRs were similar, implying that FWRs offer good support to CDRs while being neutral to scaffolding for antibody structural integrity. The other main difference was the elbow angle. As published before, the elbow angle of certolizumab changes by 9° during antigen binding, from 138° (free) to 129° (binding) [139]. In contrast, the elbow angles of *P*1 and $P6_5$ structures of A33 Fab were 156° and 145°, respectively. Certolizumab and Fab A33 possess the same protein sequence in the switch and constant regions, whereas the elbow angles differ, implying the flexibility in the switch region. The flexibility of the elbow angle is reported to be vital for the binding affinity of Fabs to antigens, as it optimises antibody conformational dynamics and adaptation to antigen structure [142, 143]. In conclusion, the scaffold of A33 Fab is optimal to support the CDRs targeting different antigens and has the potency to be applied to other therapeutic candidates from an engineering perspective.

4.3 Human DYRK2

4.3.1 SAR-based development of DYRK2 inhibitors

An initial SAR study of DYRK2 inhibitors in the absence of structural information was performed by our collaboration partner Prof. Simon Mackay (Strathclyde University, Scotland). The structure of one of the DYRK2 inhibitors was delineated into three regions, namely ring A, ring B, and the CONH core scaffold. A set of compounds was then synthesised and tested using biochemical assays. The results indicated that only compounds with a 4-substituted pyrazole ring in ring A had moderate activities. In terms of ring B, only the benzimidazole ring showed good activity, whereas any substitution of the indole, indazole, quinoxaline, and 8-quinoline rings reduced activity. In conclusion, a substituted pyrazole ring and a substituted benzimidazole ring are essential for maximal activity (Figure 4-29). Based on this scaffold, the effects of substituents at the R_1, R_2, and R_3 positions, which are pyrazole 4-substituents, benzimidazole 1-substituents, and benzimidazole 2-substituents, respectively, were further investigated.

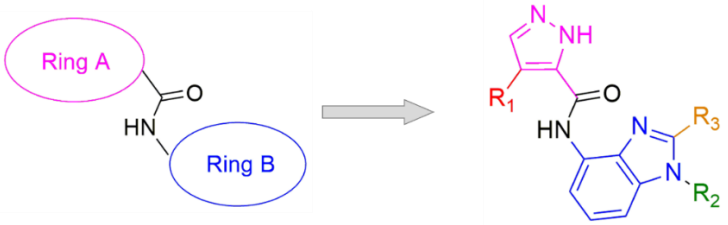


Figure 4-29. General scaffold of a DYRK2 inhibitor.

In terms of R_1, the 4-bromo substituent was more favourable than the 4-chloro and 4-iodo substituents. Meanwhile, any large groups at the R_2 position could

abrogate potency. Complete removal of the benzimidazole 1-substituent was well tolerated. Interestingly, smaller lipophilic substituents showed better activity, with the chloroethyl substituent being the most active (inhibitor LMB035). Finally, any replacement of the H atom at position 2 of the benzimidazole ring (R_3) abrogated activity, indicating that the benzimidazole ring should remain unsubstituted at this position.

The SAR map summarising these findings from our collaboration partner is shown in Figure 4-30. Four inhibitors based on these findings were sent to us and subjected to subsequent crystallisation experiments (Table 4-12).

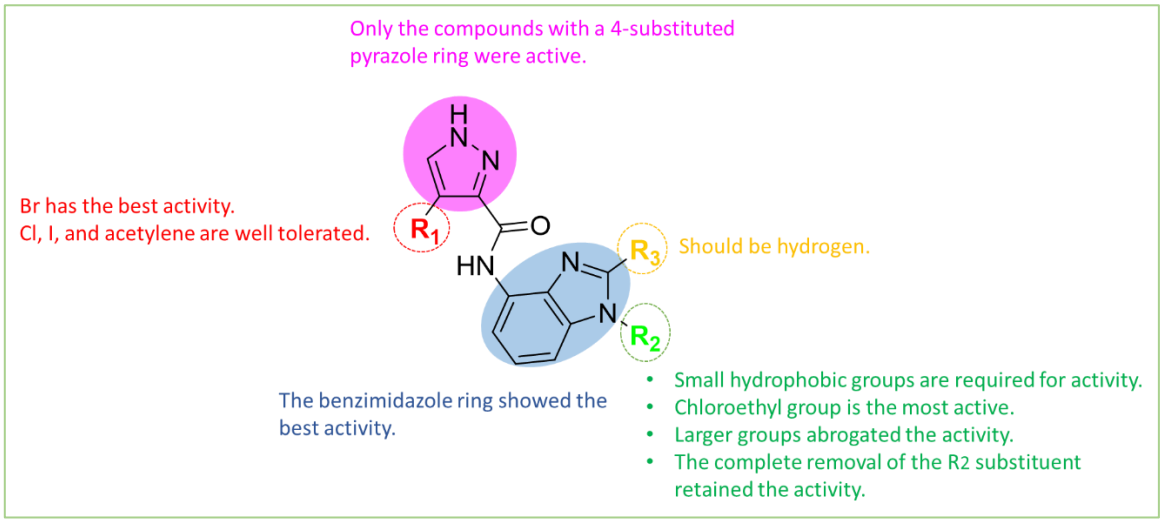
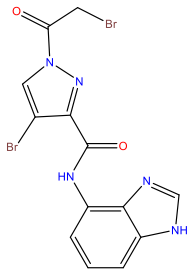
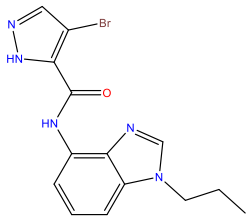
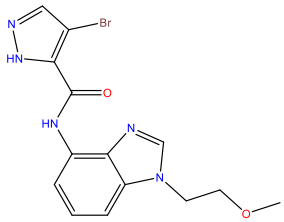
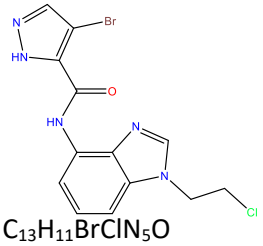


Figure 4-30. SAR summary of the benzimidazol-7-yl-pyrazole-3-carboxamide DYRK2 inhibitors

*Table 4-12. Chemical structures and selected calculated properties of the inhibitors subjected to co-crystallisation experiments. The inhibitors were developed based on SAR investigations by our collaboration partner Prof. Simon MacKay. The chemical properties of the inhibitors obey the Lipinski rules of 5.*n.d. Not identified.*

Name [Number]	***Structure***	***MW [Da]***	***cLogP***	***Number of HBA/HBD***	***PSA [Å²]***	***K_i [nM]***
CI641 (DK1)	 $C_{13}H_9Br_2N_5O_2$	427.1	0.00	5/2	86.2	n.d.
CI709 (DK2)	 $C_{14}H_{14}BrN_5O$	348.2	1.66	5/2	69.1	210
LMB017 (DK3)	 $C_{14}H_{14}BrN_5O_2$	364.2	0.68	6/2	78.3	890
LMB035 (DK4)	 $C_{13}H_{11}BrClN_5O$	368.6	1.55	5/2	69.1	19

4.3.2 Crystallisation and data processing of DYRK2$_{73-478}$

The co-crystallisation trials of DYRK2$_{73-478}$ in the presence of four different compounds were conducted separately. The purity of DYRK2$_{73-478}$ used for crystallisation trials is shown in Figure 4-31. DYRK2$_{73-478}$-inhibitor crystals appeared after three weeks. However, the first crystals only diffracted to ~8 Å. With further optimisation, the resolution was increased to ~3.0 Å. The DYRK2$_{73-478}$-CI641 crystals diffracted to 3.5 Å. Thus, this dataset was initially ignored and processed after the other three datasets. Meanwhile, the DYRK2$_{73-478}$-LMB035 crystals diffracted to 2.5 Å, whereas the DYRK2$_{73-478}$-LMB017 and DYRK2$_{73-478}$-CI709 crystals diffracted to 3.0 Å (Figure 4-32). The DYRK2$_{73-478}$-LMB035, DYRK2$_{73-478}$-LMB017 and DYRK2$_{73-478}$-CI709 structures were first solved in space group C222$_1$ using XDS. Among the solved structures, the DYRK2$_{73-478}$-LMB035 model showed the best R_{free} value of 29.7%. In contrast, the R_{free} values of the DYRK2$_{73-478}$-LMB017 and DYRK2$_{73-478}$-CI709 models were 33.6% and 34.4%, respectively. The crystallographic statistics are given in Table 4-13.

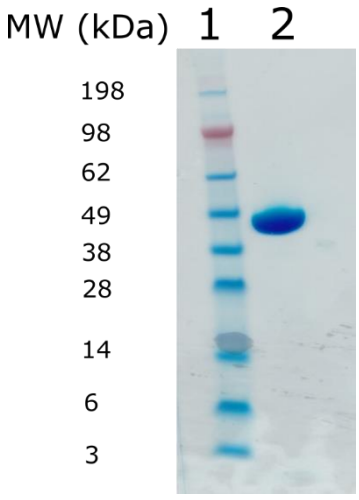


Figure 4-31. SDS-PAGE of purified DYRK2. Lane 1, protein markers (See blue plus2, Thermo Fisher). Lane 2, purified DYRK2. The protein band corresponds to the MW of DYRK2$_{73-478}$ (46.8 kDa) with a purity of > 95%.

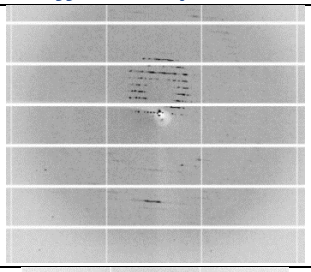
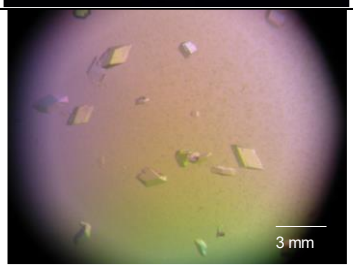
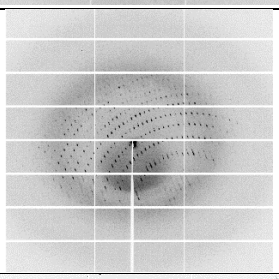
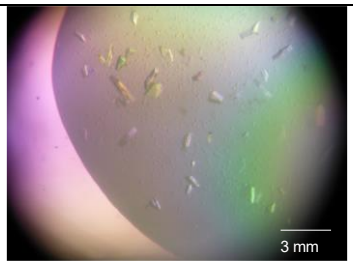
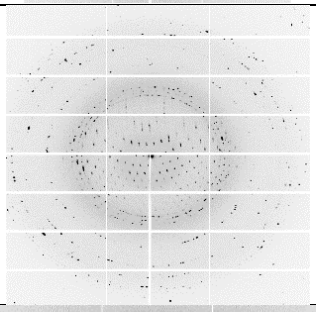
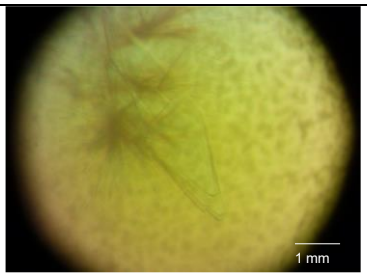
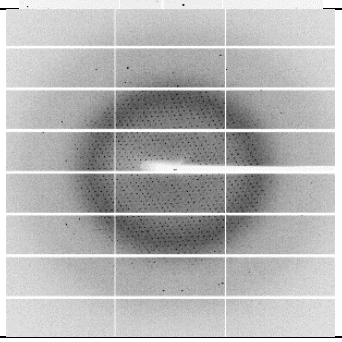
Crystal Name	Crystal image	Diffraction pattern
$DYRK2_{73-478}$-CI641 (DK1)		
$DYRK2_{73-478}$-LMB035 (DK2)		
$DYRK2_{73-478}$-LMB017 (DK3)		
$DYRK2_{73-478}$-CI709 (DK4)		

Figure 4-32. Crystal images and diffraction patterns of the $DYRK2_{73-478}$-inhibitor complexes

Table 4-13. Data collection and refinement statistics for the DYRK2$_{73-478}$-inhibitor complexes processed in space group C222$_1$. Values for the outer shell are given in parentheses.

Crystal Name	**DYRK2$_{73-478}$-LMB035**	**DYRK2$_{73-478}$-LMB017**	**DYRK2$_{73-478}$-CI709**
Data collection			
Wavelength (Å)	0.9655	0.9655	0.9655
Space group	*C222$_1$*	*C222$_1$*	*C222$_1$*
a, b, c (Å)	60.60, 130.01, 282.98	60.78, 129.75, 281.66	60.02, 128.50, 279.11
α, β, γ (°)	90, 90, 90	90, 90, 90	90, 90, 90
Resolution range (Å)	47.87–2.5 (2.589–2.5)	47.48–3.0 (3.107–3.0)	47.27–3.0 (3.18–3.0)
No. of reflections	238409 (21443)	104829 (9089)	93211 (14520)
Unique reflections	39217 (3853)	21135 (1984)	20109 (3203)
Multiplicity	6.1 (5.6)	5.0 (4.6)	4.6 (4.5)
Completeness (%)	99.81 (99.69)	92.43 (88.01)	90.9 (91.5)
⟨I/σ(I)⟩	9.66 (1.75)	6.47 (1.85)	5.6 (1.7)
R_{meas} (%)	13.6 (69.9)	24.9 (73.6)	25.8 (111.5)
Refinement			
Overall Wilson B (Å^2)	49.8	42.2	61.0
Bond lengths (Å)	0.013	0.011	0.011
Bond angles (°)	1.54	1.51	1.60
Protein non-hydrogen atoms	6181	6177	6120
Solvent	135	20	42
B factors (Å^2)	66.9	37.3	55.2
Protein	67.1	37.4	55.1
Water	60.3	27.1	66.36
R_{work}/R_{free} (%)	23.3/29.65	25.2/33.60	27.2/34.4
Ramachandran			
Preferred, allowed, outliers (%)	89.0, 9.2, 1.9	81.3, 15.6, 3.1	71.4, 18.6, 10.0

However, further analyses using Xtriage suggested another solution to solve the structure of $DYRK2_{73-478}$-inhibitor crystals in space group $C12_1$ with twinning laws applied. The datasets were thus reprocessed in this space group. The R_{free} values decreased to 28.3% for the $DYRK2_{73-478}$-LMB035 structure and to 30.5% for the $DYRK2_{73-478}$-LMB017 structure. Refmac identified the twinning law of the $DYRK2_{73-478}$-LMB035 dataset as H, -K, -L and the twinning law of the $DYRK2_{73-478}$-LMB017 dataset as -H, -K, L. Unfortunately, as the $DYRK2_{73-478}$-CI709 dataset was only collected covering 100°, it showed a low overall completeness of 55% when indexed in the $C12_1$ space group, and the complex structure could not be determined.

Finally, the $DYRK2_{73-478}$-CI641 dataset, which has a resolution of 3.5 Å, was processed. $DYRK2_{73-478}$-CI641 was indexed via XDS in the $C12_1$ space group, with an R_{merge} of 16%, an I/σ of 3.8 (1.2), and a completeness of 79%. We then forced XDS to index the dataset in space group $C222_1$. However, XDS could only index 1/3 of the spots in $C222_1$ and the R_{merge} was > 40%, indicating that it is not the solution. The merged dataset in space group $C12_1$ was subjected to molecular replacement using the solved $DYRK2_{73-478}$-LMB035 structure as the reference model and refined using *Refmac* with the twinning law (-H, -K, L). Unfortunately, owing to the low resolution of the dataset, the density of water molecules and inhibitor CI641 was mostly missing, and the conformation of the inhibitor could not be properly identified (Figure 4-33). *Refmac* could not perform any further refinements even if we forced the inhibitor into the binding pocket; hence, we decided to stop here.

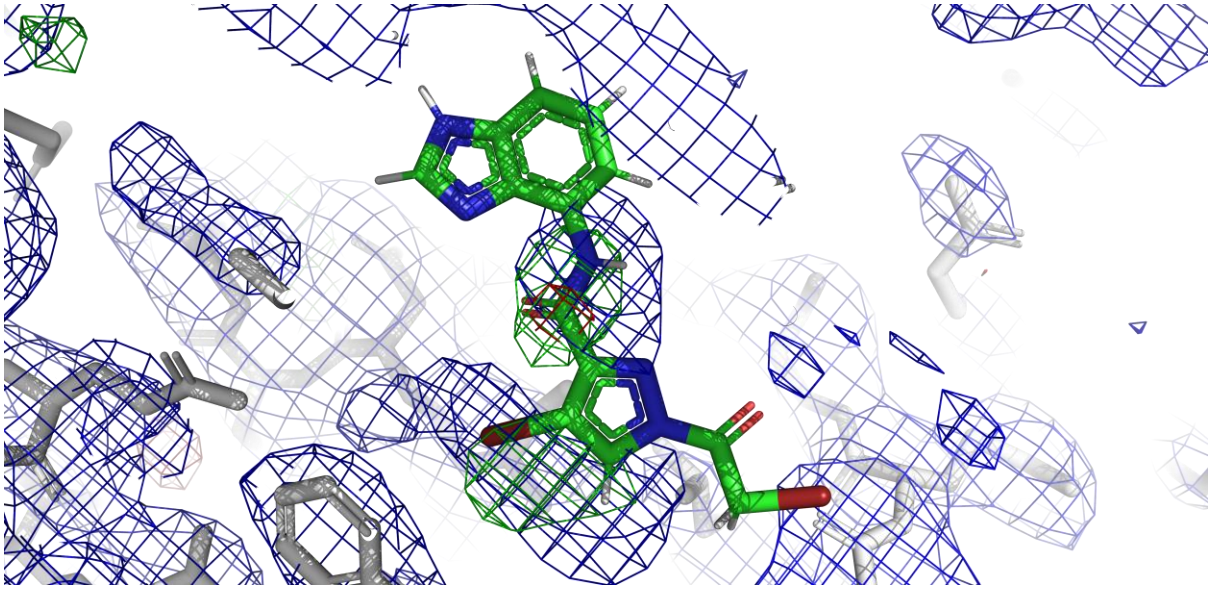


Figure 4-33. Owing to the low resolution of the DYRK2$_{73-478}$-CI641 dataset, the electron density of the CI641 inhibitor is mostly missing. The conformation of CI641 could not be properly identified. DYRK2$_{73-478}$ is coloured in grey, CI641 is coloured by elements (carbon: green, nitrogen: blue, oxygen: red, bromine: brown), the 2Fo-Fc map (contoured at 1σ) is coloured in blue, and the Fo-Fc map (contoured at 3σ) is coloured in green (positive) and red (negative).

Table 4-14. Data collection and refinement statistics for the DYRK2$_{73-478}$ complexes in space group C12$_1$. Values for the outer shell are given in parentheses.

Crystal Name	**DYRK2$_{73-478}$-LMB035**	**DYRK2$_{73-478}$-LMB017**	**DYRK2$_{73-478}$-CI641**
Data collection			
Wavelength (Å)	0.9655	0.9655	0.9686
Space group	*C12$_1$*	*C12$_1$*	*C12$_1$*
a, b, c (Å)	60.77, 130.32, 287.08	60.82, 129.35, 280.73	60.89, 129.16, 284.99
α, β, γ (°)	90.00, 90.11, 90.00	90.00, 90.45, 90.00	90.00, 91.55, 90.00
Resolution range (Å)	47.87–2.5 (2.589–2.5)	47.48–3.0 (3.107–3.0)	48.14–3.5 (3.780–3.5)
No. of reflections	246926 (22637)	108978 (9844)	64604 (10058)
Unique reflections	75828 (7327)	40629 (3910)	22378 (4001)
Multiplicity	3.3 (3.1)	2.7 (2.5)	2.9 (2.5)
Completeness (%)	98.1 (95.6)	93.2 (89.9)	79.8 (69.1)
⟨I/σ(I)⟩	6.8 (1.3)	7.0 (1.9)	3.8 (1.2)
R_{meas} (%)	12.9 (62.3)	18.1 (67.8)	21.8 (127.2)
Refinement			
Bond lengths (Å)	0.016	0.015	0.0173
Bond angles (°)	2.22	2.01	1.95
No. of atoms			
Protein non-hydrogen atoms	12402	12403	12156
Solvent	50	350	none
B factors (Å^2)	38.9	26.3	43.02
Protein	39.0	26.5	42.16
Water	35.8	18.9	none
R_{work}/R_{free} (%)	25.4/28.3	22.7/30.5	33.0/38.2
Ramachandran			
Preferred, allowed, outliers (%)	92.1, 7.3, 0.66	83.8, 14.6, 1.6	82.4, 15.1, 2.7

We further investigated whether $C222_1$ or $C12_1$ is the correct solution for the datasets. As the datasets were collected from crystals with similar shapes and grown under similar crystallisation conditions, we assumed that they were all in either the $C222_1$ or $C12_1$ space group.

The electron density maps of inhibitor binding sites were analysed to identify the correct space group. In the $DYRK2_{73-478}$-LMB035 models of both $C222_1$ and $C12_1$, we found that two LMB035 molecules bound $DYRK2_{73-478}$ at distinct binding pockets. One LMB035 bound to the ATP binding pocket, which is frequently utilised by type-1 kinase inhibitors. The second LMB035 molecule bound to a pocket formed by αD, αE, and a loop between β7 and β8 (Figure 4-35 A). With respect to the αD/αE binding pocket, negative density was observed in the $C222_1$ model, whereas the density looked tidy in the $C12_1$ model. Furthermore, the density of LMB035 in the $C222_1$ model 'merged' with the other LMB035 from an asymmetric molecule in the adjacent AU. The inhibitors were too close to each other in this model. In contrast, the densities of these two inhibitors were well separated in the $C12_1$ model (Figure 4-34 A, B, C, D).

In the $DYRK2_{73-478}$-LMB017 model in space group $C222_1$, LMB017 covalently bound to Ile294, which contradicted the biochemical data. In contrast, LMB017 was well separated from Ile294 in space group $C12_1$. A similar issue was observed for the $DYRK2_{73-478}$-CI709 complex in $C222_1$ (Figure 4-34 G, H).

Finally, the side chain of Phe160, which is located at the G-loop (P-loop) and is vital for DYRK2 selectivity (*Results 4.3.4.2*), was missing in all the $C222_1$ models, whereas its electron density could be identified in the $C12_1$ models (Figure 4-34). Based on these observations, we conclude that $C12_1$ is the

correction solution, whereas $C222_1$ most likely overestimates the crystal symmetry.

In summary, the $DYRK2_{73-478}$-LMB035 and $DYRK2_{73-478}$-LMB017 structures were solved in space group $C12_1$. We excluded the possibility of the $C222_1$ space group based on electron density analyses. The resolution of the $DYRK2_{73-478}$-CI641 structure was too low to observe the density of the inhibitor. Furthermore, the $DYRK2_{73-478}$-CI709 dataset was only collected at 100° and could not be properly indexed in space group $C12_1$. Thus, in the following chapters, we will mainly discuss the $DYRK2_{73-478}$-LMB035 and $DYRK2_{73-478}$-LMB017 complexes.

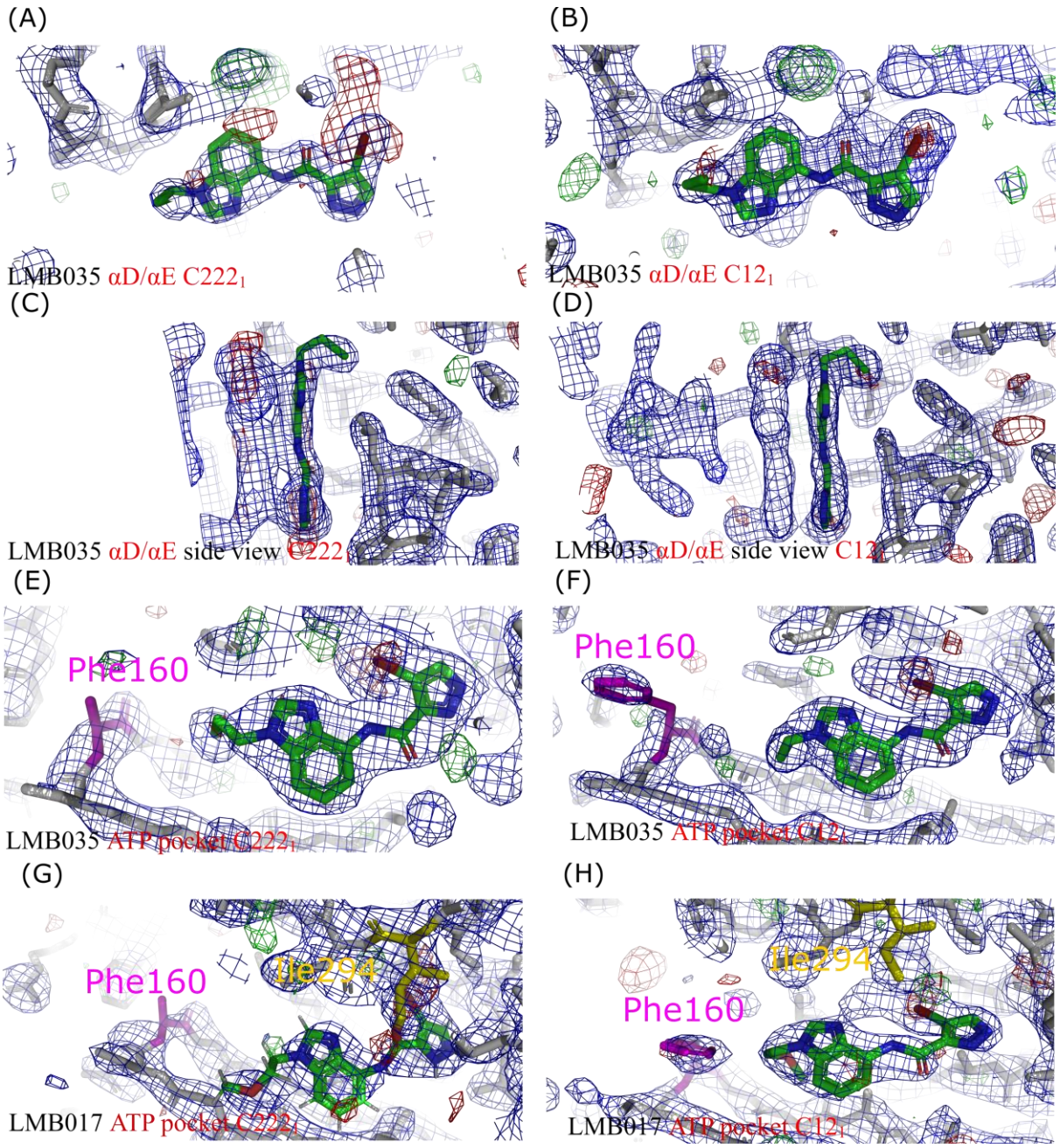


Figure 4-34. Density maps of the binding pockets of DYRK2$_{73-478}$-LMB035, DYRK2$_{73-478}$-LMB017, and DYRK2$_{73-478}$-CI709. (A), (B) αD/αE binding pockets of DYRK2$_{73-478}$-LMB035 C222$_1$ and C12$_1$ models. The C222$_1$ model shows negative electron density (red), whereas the C12$_1$ model looks tidy. (C), (D) LMB035s from two AUs are located too close to each other, and their density is 'merged' together. In contrast, LMB035s from two AUs are well separated in the C12$_1$ model. (E), (F) The density of Phe160 is missing in the C222$_1$ DYRK2$_{73-478}$-LMB035 model, whereas it is visible in the C12$_1$ model. (G), (H) LMB017 covalently binds to Ile294 in the C222$_1$ model. In contrast, the densities of LMB017 and Ile294 are separated in the C12$_1$ model.

4.3.3 Binding analysis of the inhibitors LMB035 and LMB017

The interactions between DYRK2 and inhibitors were analysed.

In both $DYRK2_{73-478}$-LMB035 and $DYRK2_{73-478}$-LMB017 complexes, each AU contained four copies of the complex. Interestingly, in the 'outer' molecules of $DYRK2_{73-478}$-LMB035 AU (molecules A, D), LMB035 bound twice to the same $DYRK2_{73-478}$ molecule in two distinct binding pockets (Figure 4-35 C). One LMB035 bound to the ATP binding pocket, while the other bound to a pocket consisting of αD, αE, and a loop between β7 and β8 (Figure 4-35 A). The LMB035 at αD/αE pocket also interacted with another LMB035 from the adjacent AU by pi stacking (Figure 4-34 D). In contrast, only one LMB017 was observed in each of the four DYRK2 of the AU (Figure 4-35 B). Although there were some uncertainties in the $DYRK2_{73-478}$-CI709 and $DYRK2_{73-478}$-CI641 models, only one CI709/CI641 was confirmed to bind to each DYRK2.

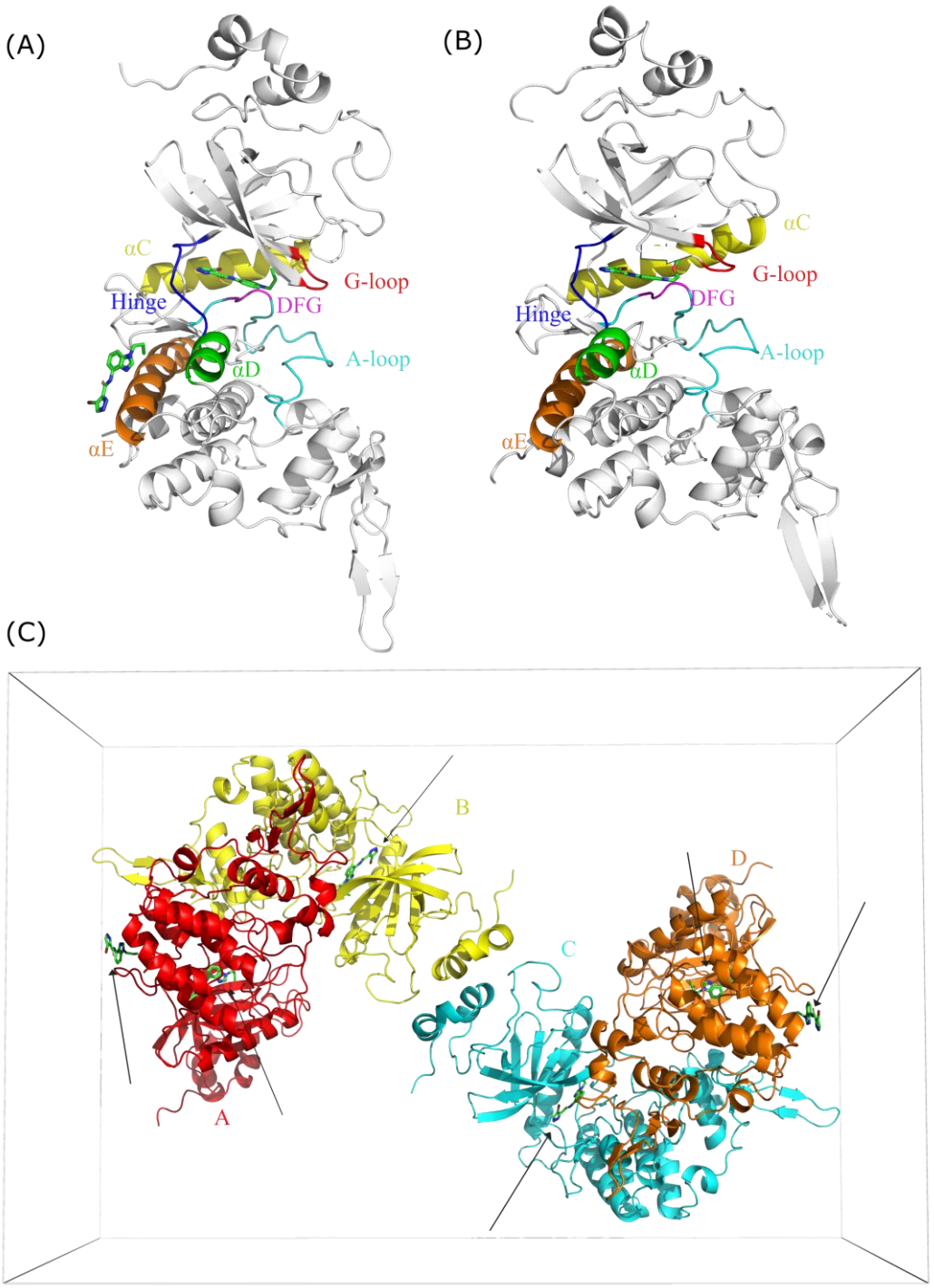


Figure 4-35. (A) Crystal structure of the DYRK2$_{73-478}$-LMB035 complex. LMB035 is coloured by elements, αC is coloured in yellow, αD is coloured in green, αE is coloured in orange, the G-loop is coloured in red, the A-loop is coloured in cyan, the Hinge is coloured in blue, and the DFG motif is coloured in magenta. (B) Crystal structure of the DYRK2$_{73-478}$-LMB017 complex. (C) Four molecules in the DYRK2$_{73-478}$-LMB035 AU. Among the 'outer' molecules within the AU (molecules A, D), LMB035 binds to DYRK2$_{73-478}$ at two distinct binding pockets (black arrows). Among the 'inner' molecules (B, C), only one inhibitor binds to each DYRK2$_{73-478}$.

In the ATP-binding pocket, the N2 of the pyrazole ring of LMB035 forms a hydrogen bond with Leu231 of the hinge and the N1 forms another H-bond with a buried water molecule (Figure 4-36 A). Furthermore, LMB035 forms numerous hydrophobic interactions with residues Ala176, Leu230, and Leu282. The 4-bromo substituent of pyrazole forms three hydrophobic interactions with Ala176, Phe228, and Ile294. Interestingly, Br also forms a halogen bond with a water molecule deeply buried in the ATP-binding pocket, which in turn is H-bonded with the carboxylate of Glu193 and the nitrogen of Asp295. Meanwhile, the benzimidazole ring forms hydrophobic interactions with Val163 and Ile155. The terminal chloro substituent points upward within a small pocket, herein G-pocket, formed by residues Phe160, Gly161, Gln162, and Val163 of the G-loop (ATP-binding loop).

A similar binding conformation can be observed in the $DYRK2_{73-478}$-LMB017 complex. The hydrophobic interactions essentially remain the same (Figure 4-35 Figure 4-36) The hydrogen bond between the N2 of the pyrazole ring and Leu231 can still be observed, while the N1-water hydrogen bond cannot be confirmed. An additional hydrogen bond can be observed between the methoxyethyl substituent and Lys178. The featured halogen bond between the 4-bromo substituent and the deeply buried water can still be observed.

As discussed in the SAR-based inhibitor development section, halo-substituted pyrazoles were the most active compounds within the designed series. The structure observations further indicated that the halogen bond between the inhibitor and the deeply buried water molecule is pivotal for effective ligand binding. Halogen bonds are reported to be significant in stabilising the receptor-ligand binding and mediating recognition [144]. Furthermore, halogen bonds are strong enough to compete with standard hydrogen bonds [145].

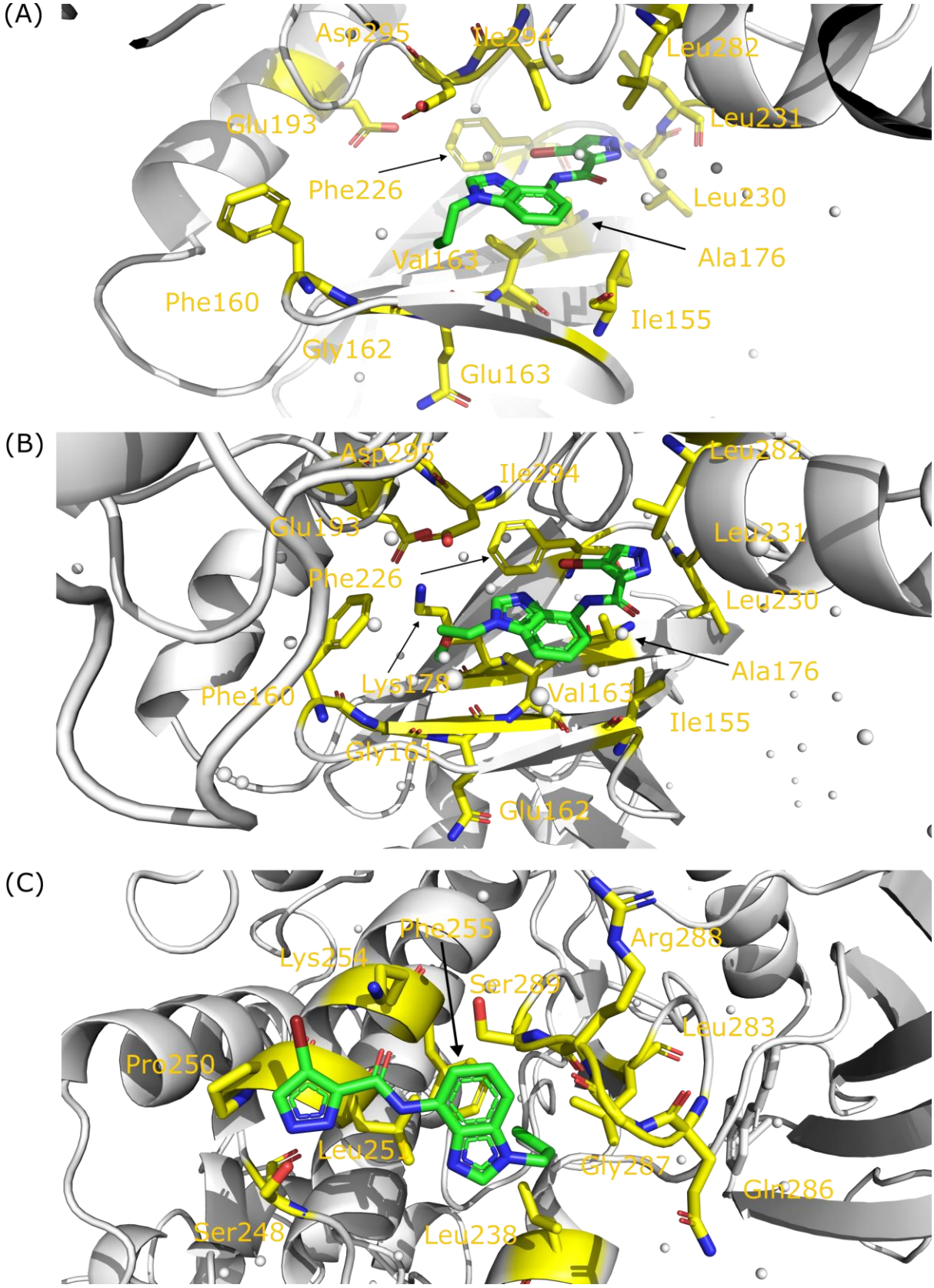


Figure 4-36. (A) LMB035 in an ATP-binding pocket. (B) LMB017 in an ATP-binding pocket. (C) LMB035 in the αD/αE pocket. Water molecules are shown as grey spheres.

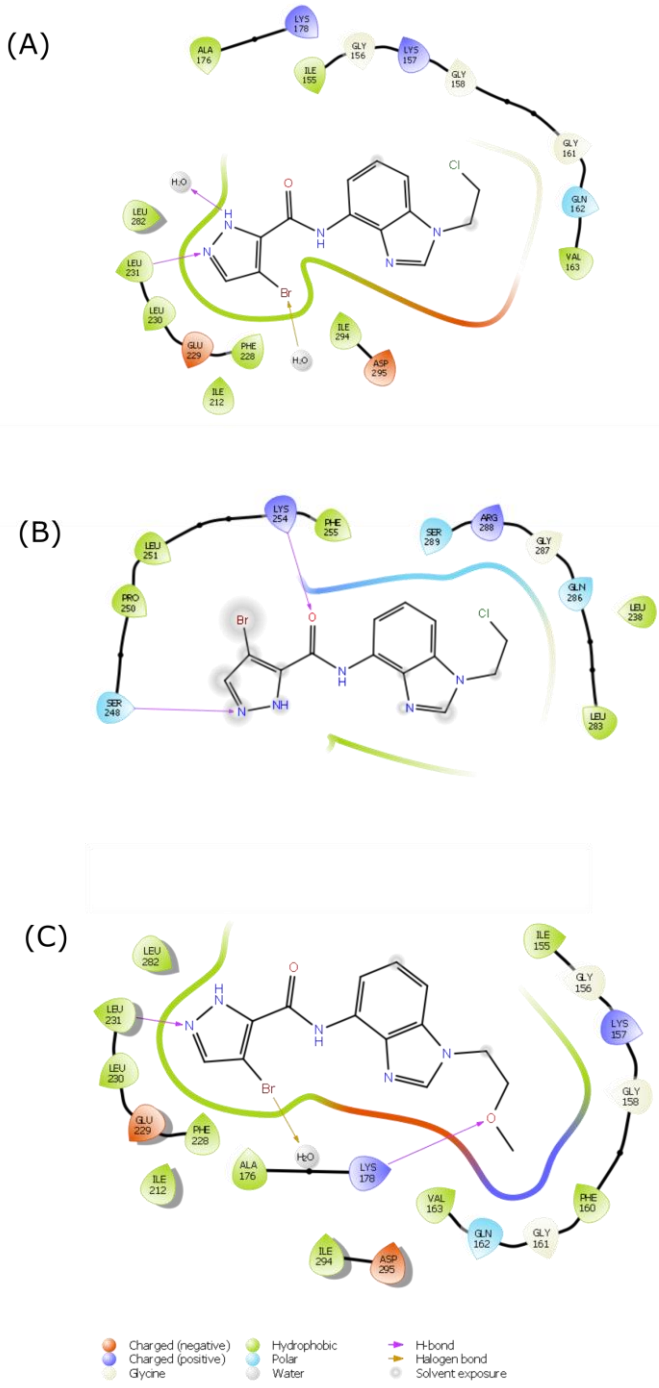


Figure 4-37. 2D interaction maps. (A) LMB035 binds to the ATP-binding pocket. Two hydrogen bonds, along with a halogen bond between bromine and water, can be observed at this site. (B) LMB035 also binds to the αD/αE binding site. Two hydrogen bonds can be observed here. The chloroethyl substituent is twisted with chlorine pointing towards the hydrophilic loop composed of Gln286, Gly287, Arg288, and Ser289. (C) LMB017 binds to the ATP-binding site. Two hydrogen bonds along with the Br-water halogen bond can be observed.

Structural analysis of LMB035 and LMB017 indicated that effective DYRK2 type-1 inhibitors, which bind to the ATP-binding site of the kinase in its active form (DFG-in), require four essential binding elements, namely, H-bonding with the hinge, halogen bonding with the buried water molecule, hydrophobic interactions with benzimidazole mainly via Val163, and finally, a critical shape and size limit for the benzimidazole-2 substituent to fit the G-pocket (Figure 4-38).

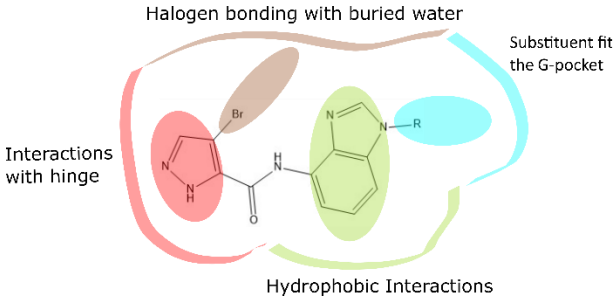


Figure 4-38. Summary of essential binding elements required for effective DYRK2 type-1 inhibitors.

During inhibitor development, LMB035 demonstrated better potency than other candidates (K_i = 19 nM). The explanation is that LMB035 not only binds to the ATP-binding pocket but also to a novel allosteric pocket, αD/αE, located at the C-lobe of DYRK2 (Figure 4-35 A, C, Figure 4-37 B). The N1 of the pyrazole ring forms a hydrogen bond with Ser248, whereas the amide carbonyl forms a hydrogen bond with Lys254. Hydrophobic interactions can be observed between the inhibitor and residues Pro250, Leu251, Phe255, Leu239, and Leu283. The chloroethyl substituent fits into a small hydrophilic pocket composed of Gln286, Gly287, Arg288, Ser289, Phe255, and Leu283 (Figure 4-36 C, Figure 4-37 B). The surface analysis implies that LMB017 has a methoxyethyl substituent that is too large to fit the small pocket and thus cannot fit into this binding site. In

conclusion, the high potency of LMB035 is attributed to the utilisation of both the ATP-binding and αD/αE pockets.

4.3.4 Discussion

4.3.4.1 Comparison of solved structures with available DYRK complex structures from the PDB

Kinases experience phosphorylation as an essential maturation step in the activation process. Previous studies have reported that DYRKs, including DYRK2 and DYRK1A, rapidly autoactivate during folding via phosphorylation on the second tyrosine residue of the conserved activation loop YxY motif (Tyr321 in DYRK1A and Tyr309 in DYRK2) [119]. This phosphorylation event occurs in cis, while DYRKs are still bound to the ribosome. The DYRKs subsequently lose tyrosine phosphorylation function and retain only the S/T phosphorylation ability [146].

Another pivotal conformational change during the maturation step is the switch from the DFG-out (inactive) to DFG-in (active) state. The DFG (Asp-Phe-Gly) motif is a common regulatory motif in protein kinases [147]. The DFG phenylalanine packs into a hydrophobic pocket between one residue from the N-lobe and one residue from the C-lobe, creating a hydrophobic regulatory spine. This packing interaction is referred to as the “DFG-in” conformation. The Asp of the DFG (Asp295 in DYRK2) coordinates the magnesium during ATP hydrolysis in this conformation. In contrast, inactivation occurs when phenylalanine (Phe296 in DYRK2) moves out of the hydrophobic pocket, disrupting the orientation of the DFG aspartate and in some cases, sterically blocking the ATP-binding site,

which are referred to as “DFG-out” conformations. The Asp of the DFG is no longer able to coordinate magnesium in this state [148]. The vast majority of ATP-competitive inhibitors (type-1 and 1.5 inhibitors) bind to active conformations (DFG-in) [147]. In contrast, the ‘off’ conformation of the ‘gatekeeper’, a highly conventional Phe residue located at the hinge region (Phe228 in DYRK2), and the ‘DFG-out’ conformation are thought to be pivotal for the binding of type-2 inhibitors (inhibitors utilising both ATP-binding and the adjacent allosteric binding site) [148].

Currently, there are 17 DYRK2-inhibitor complex structures available in the PDB. The structures from PDB were aligned with the DYRK2-LMB017 and DYRK2-LMB035 structures. During alignments, only the DFG-in conformation was observed (Figure 4-39). Simultaneously, 29 available DYRK1A structures were also aligned. The DFG motifs of DYRK1A structures are also in their active conformations. In conclusion, all DRYKs adopt the DFG-in conformation (active) wherein access to the close-up allosteric pocket is blocked. As a result, no DYRK inhibitor could be a type-2 inhibitor.

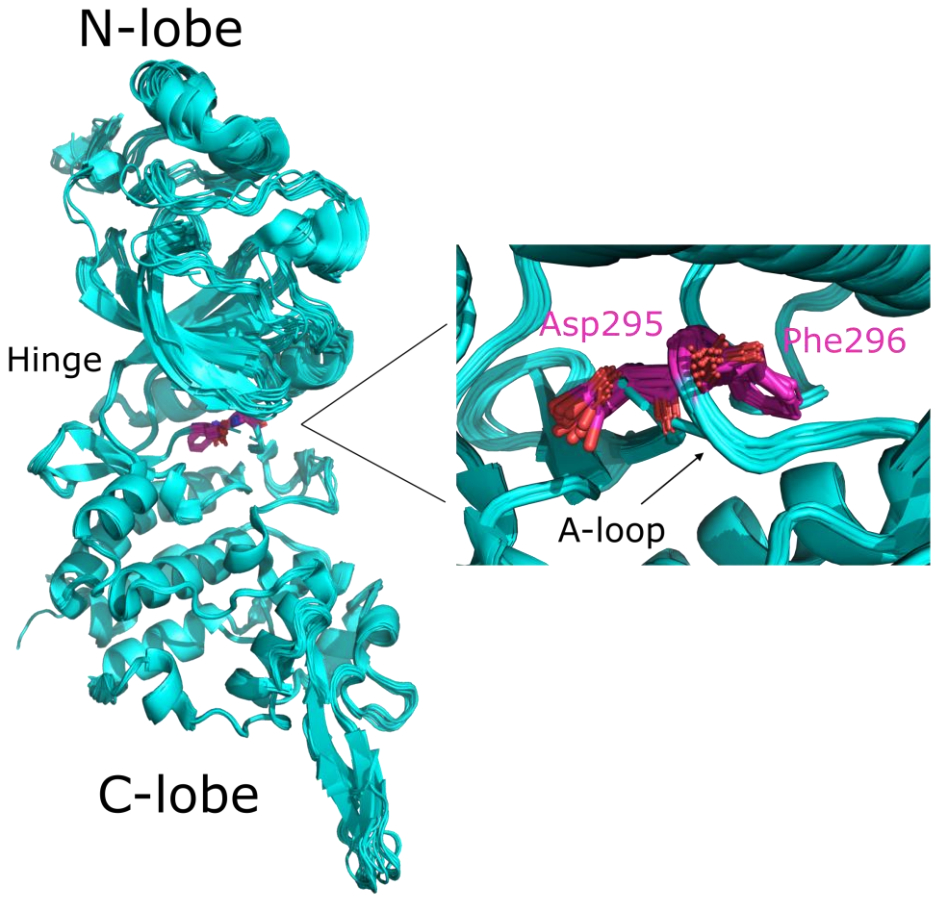


Figure 4-39. Alignments of the available DYRK2 structures from PDB with DYRK2-LMB017 and DYRK2-LMB035 structures. During alignment, only the DFG-in conformation was observed.

4.3.4.2 Analysis of the DYRK2 selectivity of LMB035 and LMB017

Among the 29 DYRK1A-inhibitor complexes deposited to PDB, 4YLL had the highest resolution (1.4 Å). The $DYRK2_{73-478}$-LMB035 and $DYRK2_{73-478}$-LMB017 complexes were thus aligned and superposed with the 4YLL structure for selectivity analysis using the Align and Superimpose Proteins protocol within *Discovery Studio* (DS) [149]. With respect to the kinase active site, three residues of DYRK2, namely Ile212, Leu230, and Ile294, are replaced with Val222, Met240, and Val306 in DYRK1A, respectively (Figure 4-40 A). This difference

would make the active site of DYRK2 a bit smaller, with a less accessible pocket volume according to DS.

A pivotal difference between the DYRK1A and DYRK2 inhibitor binding site is the Phe at the G-loop (Phe160 of DYRK2 and Phe170 in DYRK1A) (Figure 4-40 A, B). By superposing the ATP-binding sites of DYRK1A (4YLL) and DYRK2, the clash between the chloroethyl substituent (LMB035) and the Phe170 of DYRK1A is observed. The Phe170 of DYRK1A adopts an ‘in’ conformation, which completely blocks the G-pocket and subsequently clashes with any substituent pointing toward the G-pocket. In contrast, the ‘out’ conformation of Phe160 in DYRK2 cannot block the G-pocket completely to allow proper binding of LMB035 (Figure 4-40 B).

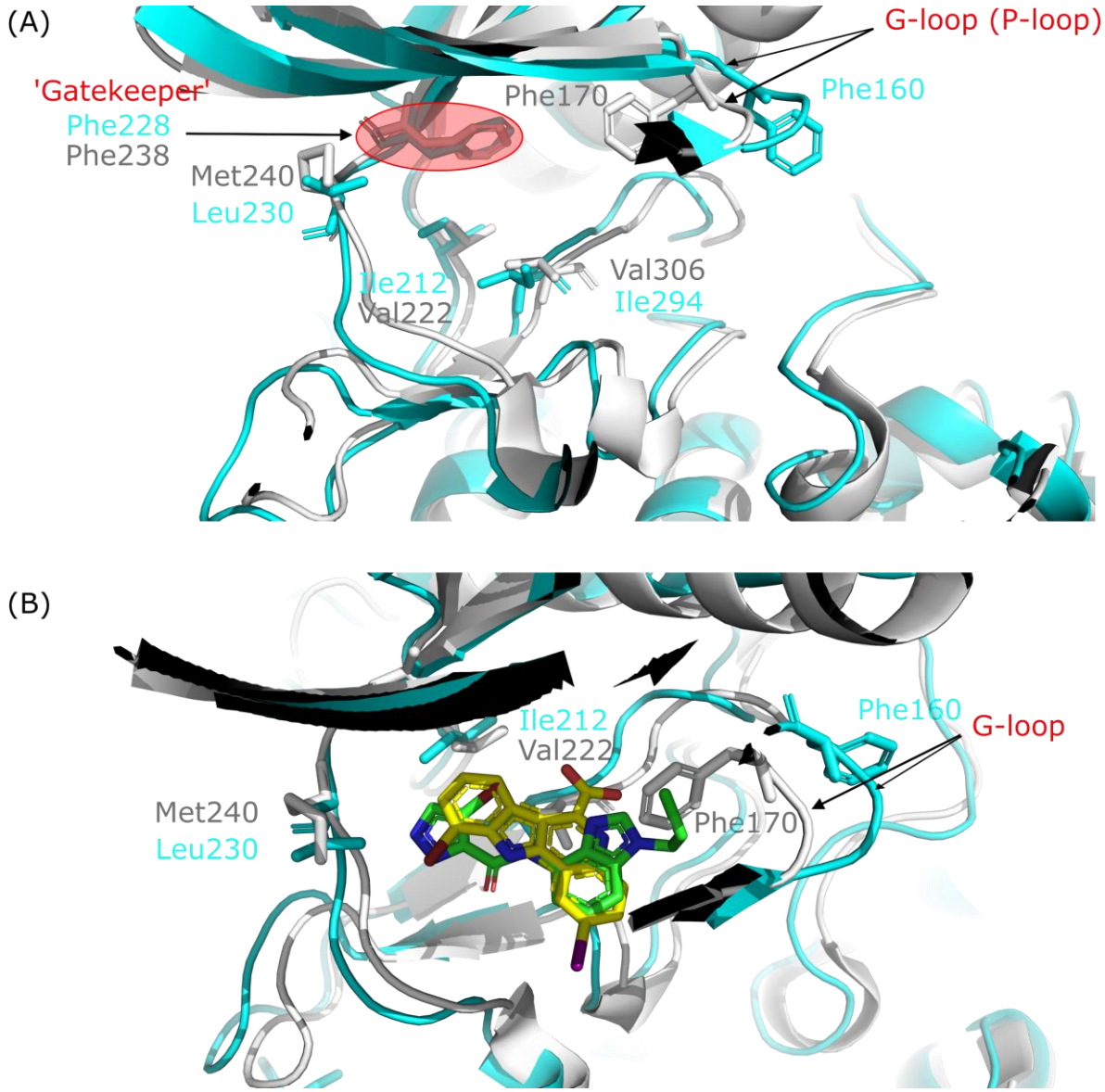


Figure 4-40. ATP-binding pocket of DYRK2$_{73-478}$-LMB035 (cyan) with DYRK1A (PDB entry 4YLL) (grey). (A) The inhibitors are not shown. Three residues of DYRK2, namely Ile212, Leu230, and Ile294, are replaced with Val222, Met240, and Val306 in DYRK1A. Furthermore, the conformations of Phe160 (DYRK2) and Phe170 (DYRK1A) are significantly different. (B) Top view of the binding pockets. DYRK1A inhibitor 4E3 from PDB entry 4YLL is marked in yellow, while LMB035 is marked in green. The Phe170 of DYRK1A adopts an 'in' conformation, which would clash with LMB035 or any substituent pointing toward the G-pocket.

A docking simulation of LMB035 into DYRK1A (4YLL) was carried out in DS by Prof. Simon Mackay. The docking score of LMB035 into the active site of

DYRK1A enzyme was much lower than its respective score calculated against DYRK2. Meanwhile, the free binding energy score showed a positive value, indicating unfavourable binding of LMB035 into DYRK1A, in perfect agreement with the measured biological activities where LMB035 is inactive against DYRK1A. In the docking model, LMB035 did not form a halogen bond with the deep water molecule. Superimposition of $DYRK2_{73-478}$-LMB035 and DYRK1A (4YLL) revealed that the distance between the Br of LMB035 and the two equivalent deep water molecules was different. In the crystal structure of DYRK2-LMB035, the distance was 2.5 Å, while in the superimposed DYRK1A-LMB035 virtual complex, it was 5.0 Å, which is beyond the effective interaction distance. A comparison between the binding pockets of DYRK1A and DYRK2 indicated that the active site in DYRK2 is generally more hydrophobic and tighter (smaller) than that in DYRK1A. Furthermore, the docking model indicated that the halo-pyrazole ring of LMB035 would be flipped during the binding into DYRK1A, probably because of the 'in' conformation of Phe170 (Figure 4-41). Based on the docking simulations, we hypothesised that the halogen bond against the deep water molecule as well as the conformational change in Phe160 (Phe170 in DYRK1A) is pivotal for the selectivity of DYRK2-specific inhibitors over DYRK1A.

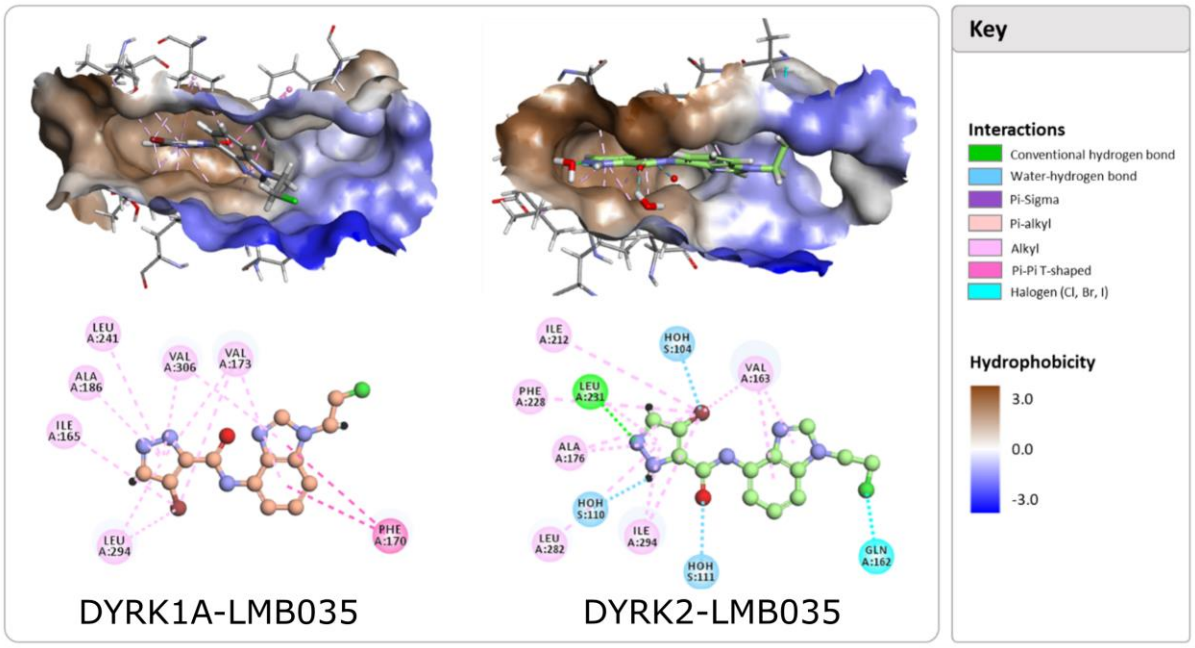


Figure 4-41. Binding pocket of the DYRK2$_{73-478}$-LMB035 complex compared to that of the DYRK1A-LMB035 docking model. Figure created by Prof. Simon Mackay.

Further analysis was conducted by aligning all 29 DYRK1A and 19 DYRK2 structures available from the PDB with those of the DYRK2$_{73-478}$-LMB035 and DYRK2$_{73-478}$-LMB017 complexes. The Phe170 in 27 of the 29 DYRK1A crystal structures adopted the 'in' conformation, while in the other two, the Phe170 adopted the 'out' conformation similar to that of the DYRK2$_{73-478}$-LMB035 complex. Interestingly, the Phe160 in the DYRK2$_{73-478}$-LMB017 model also adopted a unique 'half-in' conformation (Figure 4-42 A). Compared with those in DYRK1A structure, the β2 strand moves closer to the active pocket and the connected G-loop moves away from the pocket. Simultaneously, Phe160 rotates slightly outside but does not flip into the 'out' position. These conformational changes together make space for the ether substituent of LMB017 and avoid clashes (Figure 4-42 A).

Within 17 other available PDB entries, there are two predominant conformations of G-loop/Phe160. In these conformations, the Phe160 is either too close to LMB035/LMB017 or directly clashes with LMB035/LMB017 (Figure 4-42 B). In $DYRK2_{73-478}$-LMB035 and $DYRK2_{73-478}$-LMB017 structures, G-loop/Phe160 adopts unique conformations different from all other deposits. During the binding process, LMB035/LMB017 pushes the G-loop and Phe160 away, and the β2 strand moves closer to the active pocket. Unfortunately, we cannot prove whether the same conformational changes occur in DYRK1A.

Although conformational changes at the G-loop/Phe160 are crucial for binding of the DYRK2-selective inhibitors LMB035 and LMB017, we cannot prove the assumption introduced from docking simulation that the unique conformation of Phe160 is key to their selectivity. The involvement of the G-loop/Phe170 motif of DYRK1A in restraining the binding of LMB035/LMB017 also remains unknown.

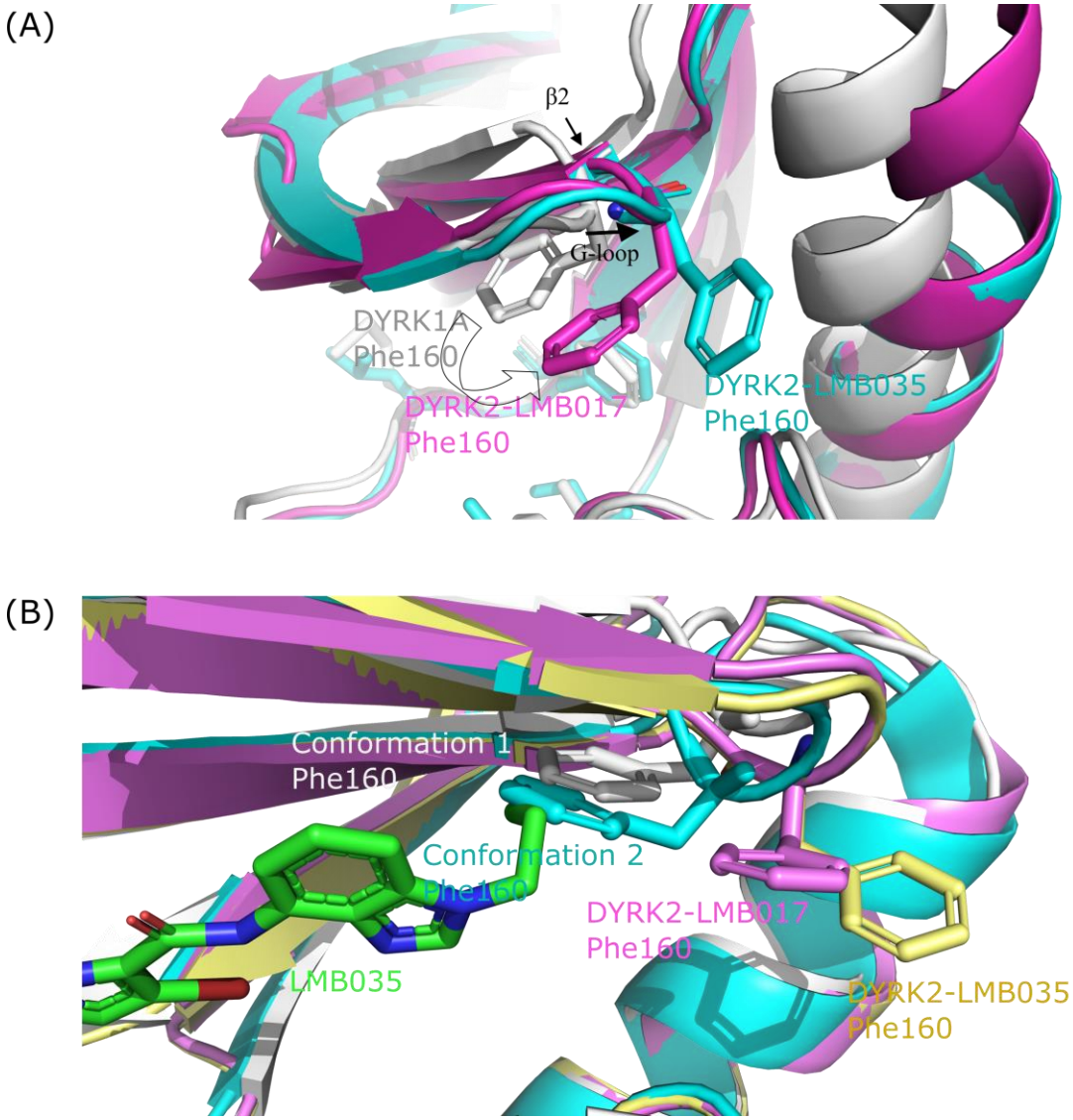


Figure 4-42. (A) Alignment of DYRK1A-4E3 (grey), DYRK2$_{73-478}$-LMB035 (cyan), and DYRK2$_{73-478}$-LMB017 (magenta) structures. Inhibitors are not shown. Compared with those in DYRK1A structure, the β2 strand of DYRK2$_{73-478}$-LMB017 moves closer to the active pocket (black arrow) and the connected G-loop moves away from the pocket (black arrow). Simultaneously, Phe160 rotates slightly outside but does not flip into the 'out' position (transparent arrow). (B) Two major conformations of G-loop/Phe160 exist among the other 17 DYRK2 deposits. These predominant conformations clash directly or indirectly with LMB035 (green) or LMB017 (not shown).

4.4 Developing KifC1 inhibitors

4.4.1 Construct properties and purification of KifC1$_{307-663}$

Human KifC1 is a member of the kinesin-14 family and consists of 673 amino acids. It contains three major domains—an N-terminal tail (residues 1 to 138), an internal coiled-coil domain (residues 141 to 297), and a C-terminal motor domain (307-663), which houses the nucleotide-binding pocket and the MT-interacting region. The expression clone KifC1$_{307-663}$ covers the motor domain of KifC1.

The KifC1 motor domain with bound Mg^{2+}ADP has been reported (PDB entry 5WDH). Alignment of the protein sequences of the published KifC1 motor domain (residues 307-663) with the KifC1$_{307-663}$ showed a very high identity between the two constructs. However, a single mutation T368P was accidentally introduced in the published KifC1 motor domain (personal information provided by Prof. Heewon Park), which may be vital for the formation of crystals (Figure 4-43). Furthermore, as stated in the publication [40], the protein was crystallised in the presence of a His-tag (Figure 4-43).

The calculated MW of KifC1$_{307-663}$ is 38.8 kDa. SDS-PAGE was performed using 12 µg purified protein after size-exclusive chromatography to verify protein purity and MW. The protein band located between the 32 and 46 kDa bands was estimated to be ~38 kDa, corresponding to the estimated MW of KifC1$_{307-663.}$ The purity was ~95% according to the SDS-PAGE result (Figure 4-44).

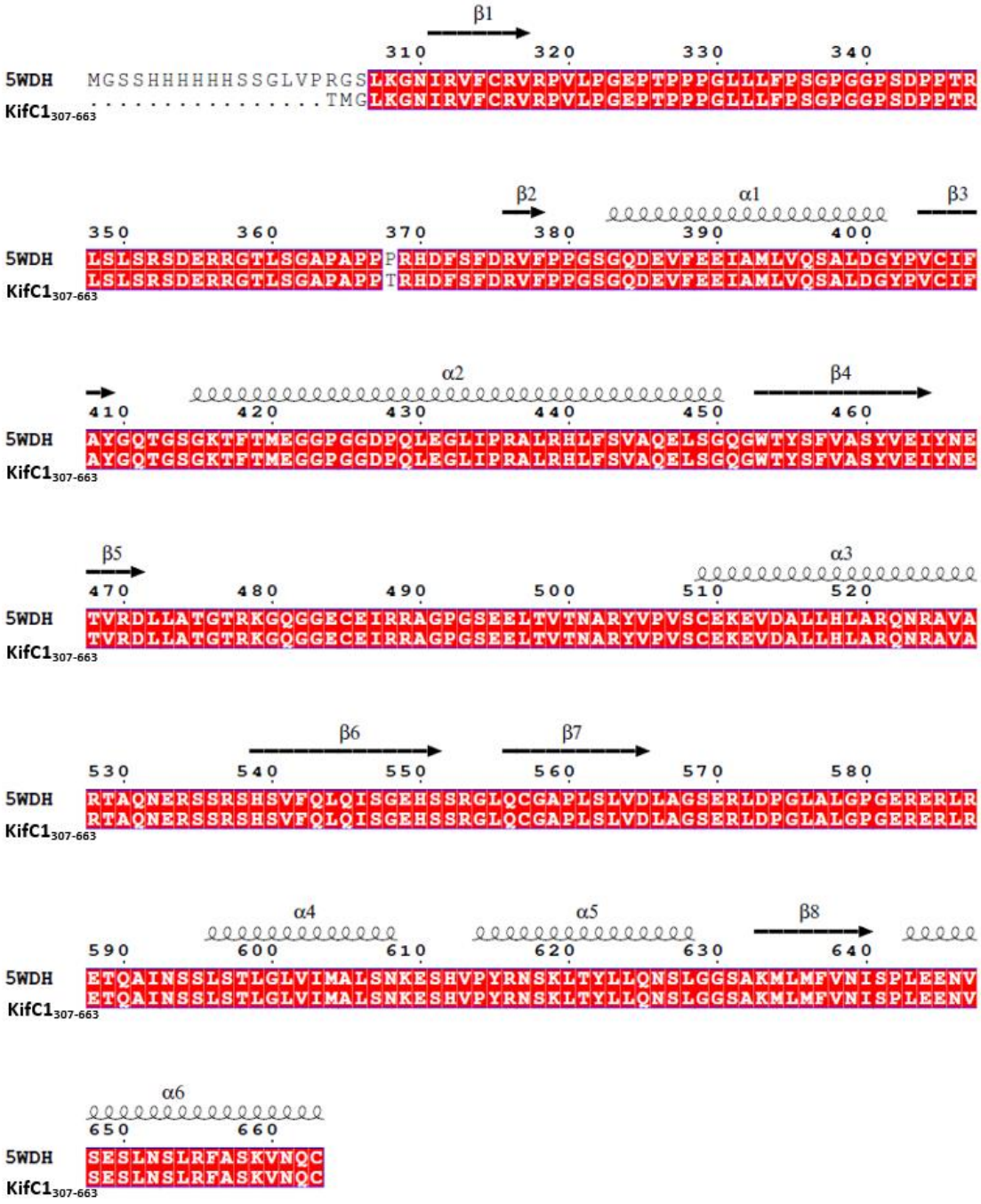


Figure 4-43. Alignment of the published KifC1 motor domain (PDB entry 5WDH) with the KifC1$_{307-663}$ construct. A single mutation T368P was introduced in the published KifC1 motor domain. Furthermore, a 6x His-tag and various additional residues were attached to the N-terminus of the 5WDH sequence, and these were not cleaved during purification.

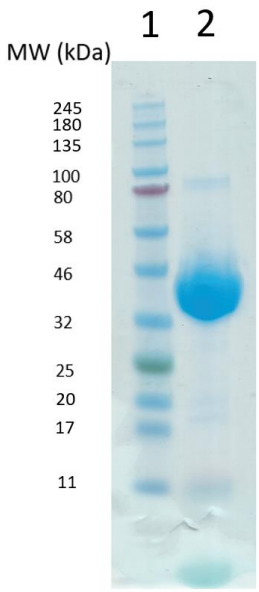


Figure 4-44. The protein band with 95% purity represents KifC1$_{307-663}$, with an MW of ~38 kDa. Lane 1: Protein marker. Lane 2: Purified KifC1$_{307-663}$.

4.4.2 Kinetic analysis and characterisation of KifC1$_{307-663}$

4.4.2.1 Salt dependence of basal ATPase activity

The enzymatic activity of KifC1$_{307-663}$ was measured in the presence of KCl and NaCl. With an increase in KCl concentration, the rate of KifC1$_{307-663}$ increased from 0.004 to 0.0157 s^{-1} and peaked at 800 mM KCl. When KCl concentration was increased further, a slight reduction was observed in the enzymatic activity. The basal KifC1$_{307-663}$ ATPase activity increased and peaked at 0.01 s^{-1} when NaCl concentration reached 300 mM. A decreasing trend was observed with an increasing NaCl concentration. In summary, 300 mM NaCl or 800 mM KCl was optimal for measuring the ATPase activity of KifC1$_{307-663}$. In further experiments, 300 mM NaCl was used in the basal ATPase assays (Figure 4-45 A).

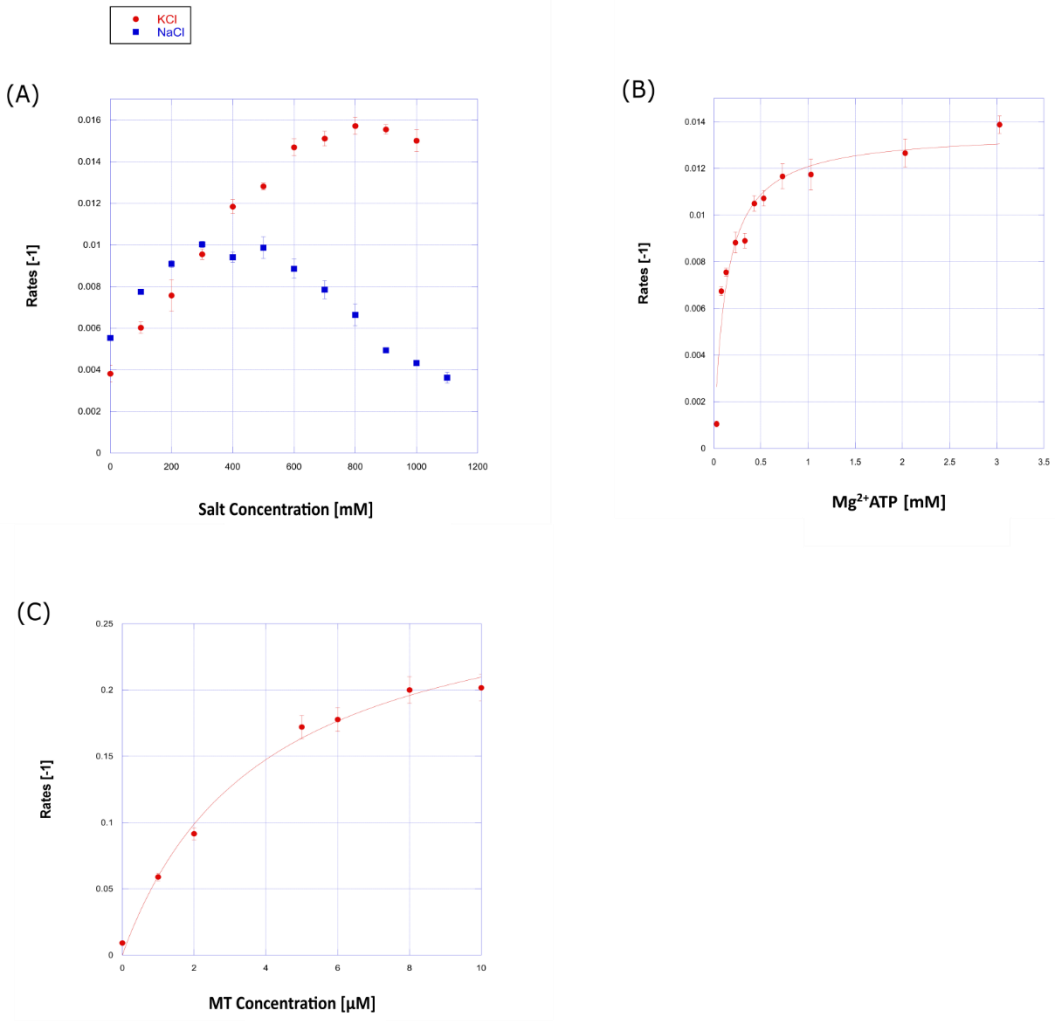


Figure 4-45. (A) Comparison of KCl (red)/NaCl (blue)-dependent basal ATPase activity of KifC1$_{307-663}$. KifC1$_{307-663}$ ATPase activity peaked at 800 mM KCl or 300 mM NaCl. (B) Basal ATPase activity of KifC1$_{307-663}$. The K_{cat} and K_M values were 0.0136 ± 0.0007 s^{-1} and 0.1244 ± 0.0228 µM, respectively. (C) Characterisation of KifC1 MT-stimulated ATPase activity. The MT-stimulated ATPase activity of KifC1$_{307-663}$ was investigated in the presence of increasing concentrations of MTs, ranging from 0–10 µM, supplemented with 1 mM Mg^{2+}ATP.

4.4.2.2 Characterisation of basal ATPase activity of KifC1$_{307-663}$

Basal KifC1$_{307-663}$ activity was characterised in the presence of increasing concentrations of Mg^{2+}ATP and the presence of 300 mM NaCl (Figure 4-45 B).

The K_M value was 0.1244 ± 0.0228 μM. Meanwhile, the k_{cat} value of KifC1$_{307-663}$ was 0.0136 ± 0.0007 s^{-1}. KifC1$_{307-663}$ showed a turnover of ~0.014 ATP molecules/s.

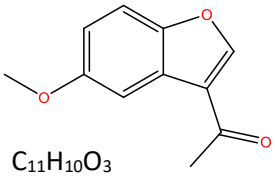
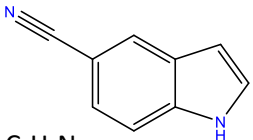
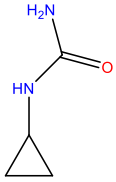
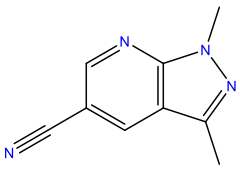
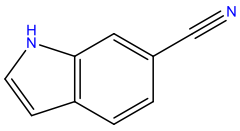
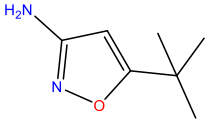
4.4.2.3 MT-stimulated ATPase activity

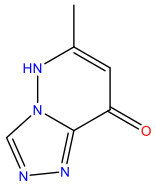
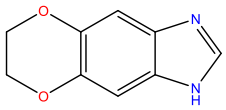
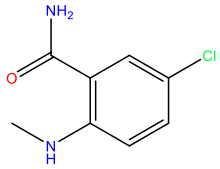
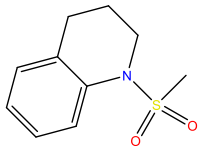
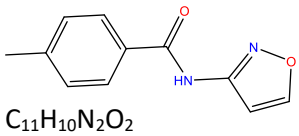
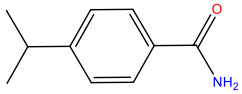
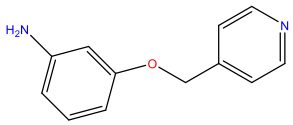
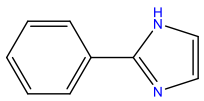
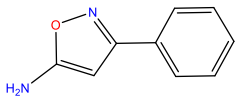
The MT-stimulated KifC1$_{307-663}$ ATPase activity was characterised in the presence of increasing concentrations of MTs, from 0–10 μM, supplemented with 1 mM Mg^{2+}ATP (Figure 4-45 C). The k_{cat} value of KifC1$_{307-663}$ was 0.802 ± 0.054 s^{-1}, and the $K_{0.5,MT}$ value was 3.908 ± 0.662 μM. The k_{cat} value of KifC1$_{307-663}$ at maximum activation in the presence of MTs was increased 80-fold compared with that of the basal activity.

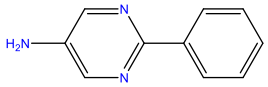
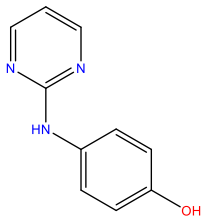
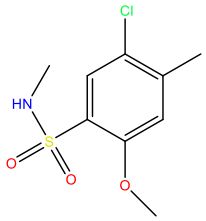
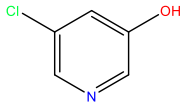
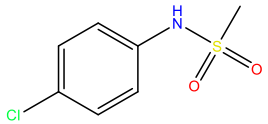
4.4.2.4 Evaluation of the KifC1 fragments identified via NMR

Our collaboration partner from Monash University identified 20 fragment hits of KifC1 using NMR screens (Table 4-15). Among these 20 hits, 10 fragments, (fragments **1**–**10**) identified via STD or more than one NMR methods, were selected for further validation.

Table 4-15. Fragments identified by initial NMR screens, including STD, CPMG, and Water Logsy. The number, commercial name (if applicable), chemical structure, molecular formulae, molecular weight, cLogP, hydrogen-bond donors (HBD), and hydrogen-bond acceptors (HBA), LogS, polar surface area (PSA), and NMR methods applied are summarised.

Number/ Commercial name	***Chemical structure***	***MW [Da]***	***cLogP***	***Number of HBA/ HBD***	***PSA [Å²]***	***LogS [mol/L]***	***NMR hit***
1 FCB102830 (Fluorochem)	 $C_{11}H_{10}O_3$	190.2	2.29	3/0	35.53	-2.17	CPMGLogsy
2 EN300-97267 (Enamine BB)	 $C_9H_6N_2$	142.2	1.96	2/1	35.82	-2.39	STD CPMG Logsy
3 EN300-13158 (Enamine BB)	 $C_4H_8N_2O$	100.1	-0.72	1/2	55.12	-0.13	STD
4 EN300-52293 (Enamine BB)	 $C_9H_8N_4$	172.2	0.36	4/0	51.75	-2.07	STD
5 EN300-100999 (Enamine BB)	 $C_9H_6N_2$	142.2	1.96	2/1	35.82	-2.38	STD CPMGLogsy
6 EN300-67494 (Enamine BB)	 $C_7H_{12}N_2O$	140.2	1.37	3/1	47.61	-1.49	STD

7 BBV-39790414 (Enamine BB)	 $C_6H_6N_4O$	150.1	-2.41	5/1	57.06	-1.61	STD
8 STL368348 (Vitas M Labs)	 $C_9H_8N_2O_2$	176.2	2.06	4/1	42.85	-1.79	STD
9 BTB01822SC (MOLPORT)	 $C_8H_9ClN_2O$	184.6	1.77	2/1	55.12	-2.48	STD CPMGLogsy
10 Synthesised by Mr. Helal	 $C_{10}H_{13}NO_2S$	211.3	1.68	2/0	37.38	-3.04	STD Logsy
11	 $C_{11}H_{10}N_2O_2$	202.2	1.93	3/1	50.69	-3.06	CPMG
12	 $C_{10}H_{13}NO$	163.2	2.08	1/1	43.09	-2.47	CPMG
13	 $C_{12}H_{12}N_2O$	200.2	1.28	3/1	47.61	-2.40	CPMG
14	 $C_9H_8N_2$	144.2	2.07	2/1	24.39	-2.30	CPMG
15	 $C_9H_8N_2O$	160.2	1.87	3/1	47.61	-2.75	CPMG

16	 $C_{10}H_9N_3$	171.2	1.77	3/1	50.74	-1.85	CPMG
17	 $C_{10}H_9N_3O$	187.2	1.66	4/2	56.98	-2.12	CPMG
18	 $C_9H_{12}ClNO_3S$	249.7	2.13	3/1	55.4	-2.73	Logsy
19	 C_5H_4ClNO	129.5	1.77	2/1	32.59	-1.07	Logsy
20	 $C_7H_8ClNO_2S$	205.6	1.86	2/1	46.17	-2.85	Logsy

4.4.2.5 Thermal shift assay (TSA)

In total, 96 buffer conditions from the Rubic buffer screen (Molecular Dimension) were tested in duplicate to determine the optimal buffer conditions for stabilising the KifC1$_{307\text{-}663}$ motor protein. Among various buffer compositions and ionic strengths, the potassium phosphate buffer (100 mM potassium phosphate pH 7.0 and 250 mM NaCl) showed the highest T_i value, which was 53.4 °C (Supplementary materials 7.2) (Figure 4-46). Considering the biophysical properties of KifC1$_{307\text{-}663}$, a modified potassium phosphate buffer (50 mM

potassium phosphate pH 7.0, 150 mM NaCl, 5% glycerol and 1 mM DTT) was employed for future experiments, including TSA, ATPase assays, and MST.

Fragments **1**–**9** were subjected to TSA measurements using the modified potassium phosphate buffer (50 mM potassium phosphate pH 7.0, 150 mM NaCl, 5% glycerol, and 1 mM DTT) as the assay buffer. A control group composed of $KifC1_{307\text{-}663}$ dissolved in 1.25%, 2.5%, 5.0%, and 10% DMSO, which are compatible with the final DMSO concentrations in the fragment measurements, was set up to investigate the solvent effects. In the control group, as DMSO concentration increased, the T_i values decreased, implying a destabilising effect of DMSO on $KifC1_{307\text{-}663}$ (Figure 4-46). The fragments should, in theory, bind to and stabilise the protein, which would increase the T_i value. However, the TSA measurements of $KifC1_{307\text{-}663}$ in the presence of these fragments showed reduced T_i values. The T_i values of the fragments were less than those of the DMSO blank groups (Table 4-16) (*Supplementary materials 7.3*). TSA was thus considered unsuitable for analysing KifC1 fragments under the present conditions. However, it may be employed later, for more potent fragment analogues.

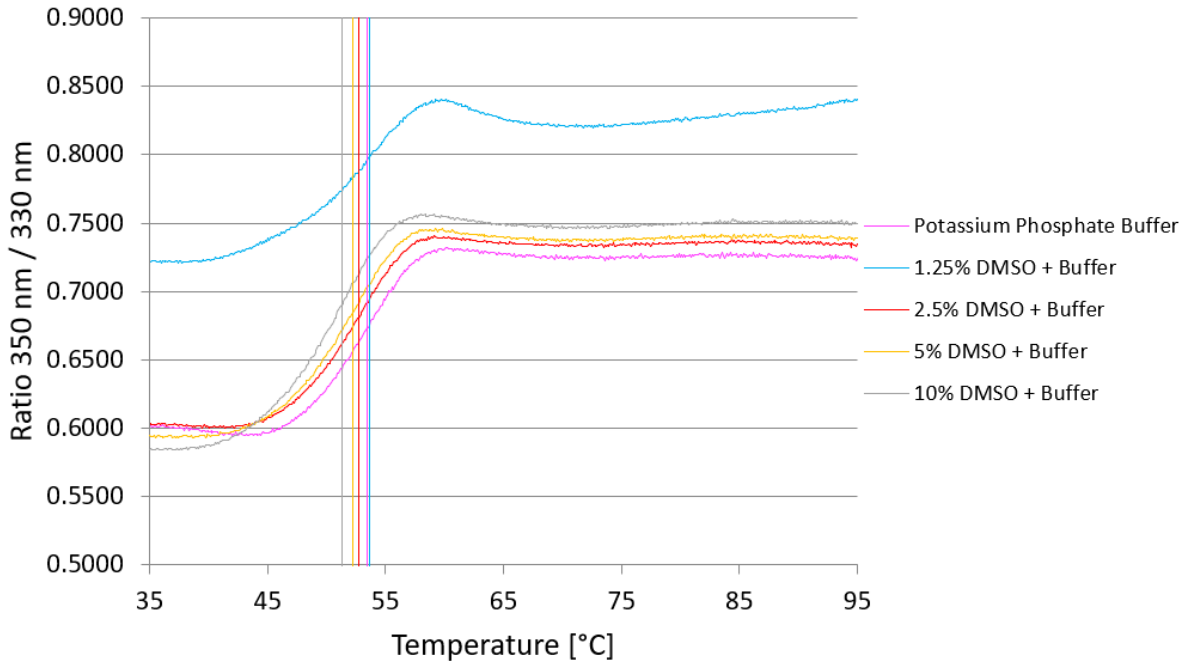


Figure 4-46. Thermal shift assay (TSA) curves averaged from duplicate measurements. TSA measurements indicate that the potassium phosphate buffer (100 mM potassium phosphate pH 7.0, 250 mM NaCl) was optimal for KifC1$_{307-663}$. The T_i value of KifC1$_{307-663}$ measured in the potassium phosphate buffer is 53.4°C, which is the highest among all buffers included in the Rubic buffer screen kit. As the fragments were dissolved in DMSO before measurement, the effects of DMSO on KifC1$_{307-663}$ were investigated by measuring KifC1$_{307-663}$ in a DMSO gradient. The measured Ti values decreased with an increasing DMSO concentration, indicating the destabilising effect of DMSO on KifC1$_{307-663}$.

*Table 4-16. Thermal shift assay (TSA) measurements of KifC1 targeting fragments. *Fragment **8** could not be measured because of intrinsic fluorescence.*

Measurements	***T_i (5 mM) [°C]***	***T_i (2.5 mM) [°C]***	***T_i (1.25 mM) [°C]***	***T_i (0.625 mM) [°C]***
KifC1 control without DMSO	53.6 ± 0.25			
control (10%, 5%, 2.5%, 1.25% DMSO)	51.3 ± 0.14	52.2 ± 0.15	52.7 ± 0.17	53.6 ± 0.18
Fragment 1	50.7 ± 0.42	51.3 ± 0.25	52.2 ± 0.23	53.1 ± 0.09
Fragment 2	39.7 ± 0.07	39.6 ± 0.22	40.6 ± 0.10	41.9 ± 0.19
Fragment 3	50.7 ± 0.35	52.2 ± 0.11	atypical curve	53.3 ± 0.15
Fragment 4	41.3 ± 0.19	41.8 ± 0.13	42.2 ± 0.14	43.1 ± 0.14
Fragment 5	atypical curve	atypical curve	atypical curve	atypical curve
Fragment 6	atypical curve	50.0 ± 0.09	51.8 ± 0.18	53.0 ± 0.41
Fragment 7	atypical curve	atypical curve	atypical curve	atypical curve
Fragment 8	None*	None*	None*	None*
Fragment 9	atypical curve	atypical curve	52.3 ± 0.31	52.9 ± 0.25

4.4.2.6 Measurement of ATPase activity of KifC1 fragments

To further validate and determine the binding affinities, basal ATPase activity measurements of $KifC1_{307\text{-}663}$ were conducted in the presence of the fragments.

Fragment **1** was not subjected to ATPase activity measurement owing to its limited availability. Among fragments **2**–**10**, which were subjected to ATPase activity measurement, fragments **4** and **9** showed detectable inhibition of $KifC1_{307\text{-}663}$ activity, whereas the remaining compounds demonstrated very weak or no binding (*Supplementary materials 7.4*). However, the inhibition curves of fragments **4** and **9** showed a sigmoidal shape rather than the typical nonlinear regression curves. Copeland *et al*. have reported that the concentration-response relationships of molecules often reflect non-ideal behaviours because of stoichiometric binding or other biophysical factors [105], which eventually result in the generation of sigmoidal curves (see *Discussion 4.4.3*). Based on the research by Copeland *et al*., an adapted Hill equation was applied to fit the curves (see *Materials and Methods 3.2.10*). An additional Hill coefficient was included in the equation to describe the steepness of the curve. The IC_{50} values for fragments **4** and **9** were 0.99 ± 0.04 mM and 0.91 ± 0.02 mM, respectively. Fragment **4** reduced the basal ATPase activity of $KifC1_{307\text{-}663}$ by 48% at most whereas fragment **9** reduced the basal ATPase activity by 84% (Figure 4-47 B). Notably, the curves for fragments **4** and **9** had high Hill coefficients of 7.6 and 6.2, respectively, indicating the stoichiometric binding of multiple ligands to the same target.

Fragments **4** and **9** were further measured against another kinesin, human Eg5, to test binding specificity. The sigmoidal curves could be fitted with basal ATPase

data, which indicated the non-specific binding of fragments **4** and **9**. The IC_{50} values of fragments **4** and **9** were 0.85 ± 0.12 mM and 1.09 ± 0.02 mM, respectively. Fragment **4** reduced the basal ATPase activity of Eg5 by 33%, whereas fragment **9** reduced it by 79% (Figure 4-47 C). The Hill coefficients for fragments **4** and **9** were 9.3 and 7.8, respectively. The binding specificity test indicated that fragments **4** and **9** bound non-specifically to kinesins.

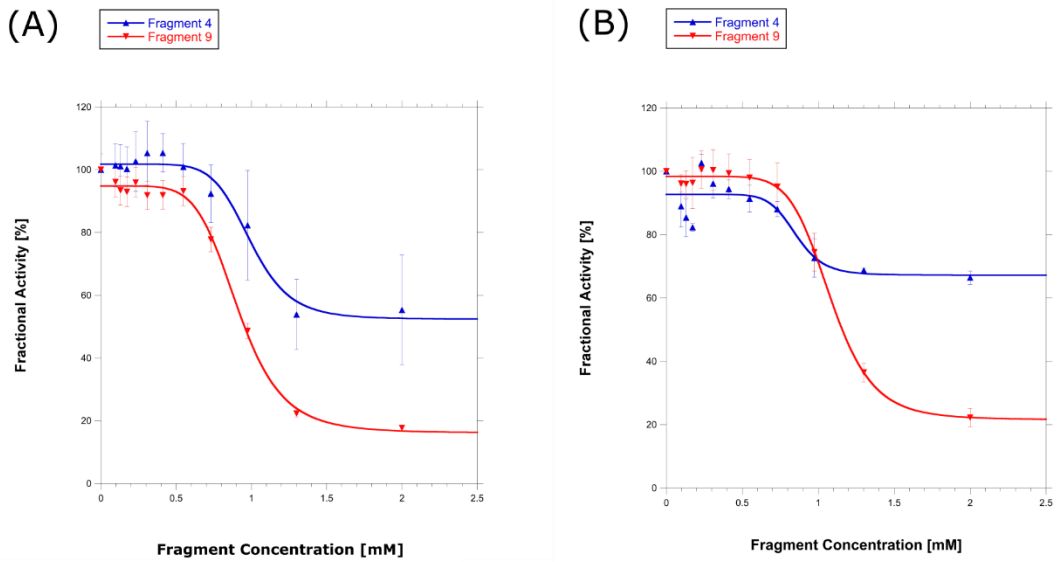


*Figure 4-47. Concentration-response curves of fragments 4 and 9 in basal ATPase assays. (A) Fragments **4** (blue) and **9** (red) reduced the activity of $KifC1_{307-663}$ by 48% and 84% at most, with IC_{50} values of 0.85 ± 0.12 mM and 1.09 ± 0.02 mM, respectively. (B) Fragments **4** (blue) and **9** (red) reduced the basal ATPase activity of Eg5 by 33% and 79%, with IC_{50} values of 0.85 ± 0.12 mM and 1.09 ± 0.02 mM, respectively, indicating no selectivity between KifC1 and Eg5.*

Nineteen KifC1 analogues based on fragments **1** to **9** were purchased from a commercial supplier, eMolecules. Meanwhile, nineteen additional KifC1 analogues were synthesised by Mr. Helal Helal. These thirty-eight analogues were subjected to further basal ATPase activity measurement (*Supplementary material section 7.1*). However, most analogues showed poor solubility in the assay buffer, causing the measurements of triplicate data to show large error bars

at high concentrations. Thus, data points with large error bars at high concentrations were excluded when plotting the figures.

Six analogues of fragment **10** showed either inhibition or activation of $KifC1_{307-663}$ during basal ATPase activity measurement (*Supplementary material section* Table 7-1). However, analogues **10A**, **10C**, **10D,** and **10E** acted as either weak antagonists or agonists, and their IC_{50} values could not be determined appropriately from the weak data. Meanwhile, owing to poor solubility, the data points at high concentrations were missing for **10H** and **10I**, and their IC_{50} values could not be identified either (Figure 4-48).

Fragments **2** and **5** are isomers but did not show inhibition of basal ATPase activity. However, six analogues of these two fragments demonstrated detectable activities in the ATPase assay, acting as agonists (Table 4-17). Analogues **2I, 2L, 2M, 5G, 5J,** and **5K** increased the basal ATPase activity of $KifC1_{307-663}$, with IC_{50} values of 0.17 ± 0.01 mM, 0.42 ± 0.01 mM, 2.45 ± 7.10 mM, 0.24 ± 0.01 mM, 0.41 ± 0.04 mM, and 1.63 ± 4.02 mM, and Hill coefficients of 4.30, 5.37, 0.93, 3.16, 5.71, and 0.90, respectively (Figure 4-49). The relatively large data confidence indicated that the curves for analogues **2M** and **5K** are atypical, indicating that their IC_{50} values are questionable. However, the rising trend of $KifC1_{307-663}$ ATPase activity was confirmed in these measurements. The ATPase activity data of these analogues indicated that adding 1-sulfonyl substituents to the 1H-indole-5-carbonitrile and 1H-indole-6-carbonitrile (fragments **2** and **5**) scaffolds increased the binding affinities of analogues **2I** and **5G**. However, large sulfonyl substituents, for example, benzenesulphonyl (**2M** and **5K**), reduced the binding affinity.

The other analogues from either eMolecules or Mr. Helal Helal did not show detectable inhibition or activation effects on the basal ATPase activity of KifC1$_{307\text{-}663}$ and thus are not discussed here.

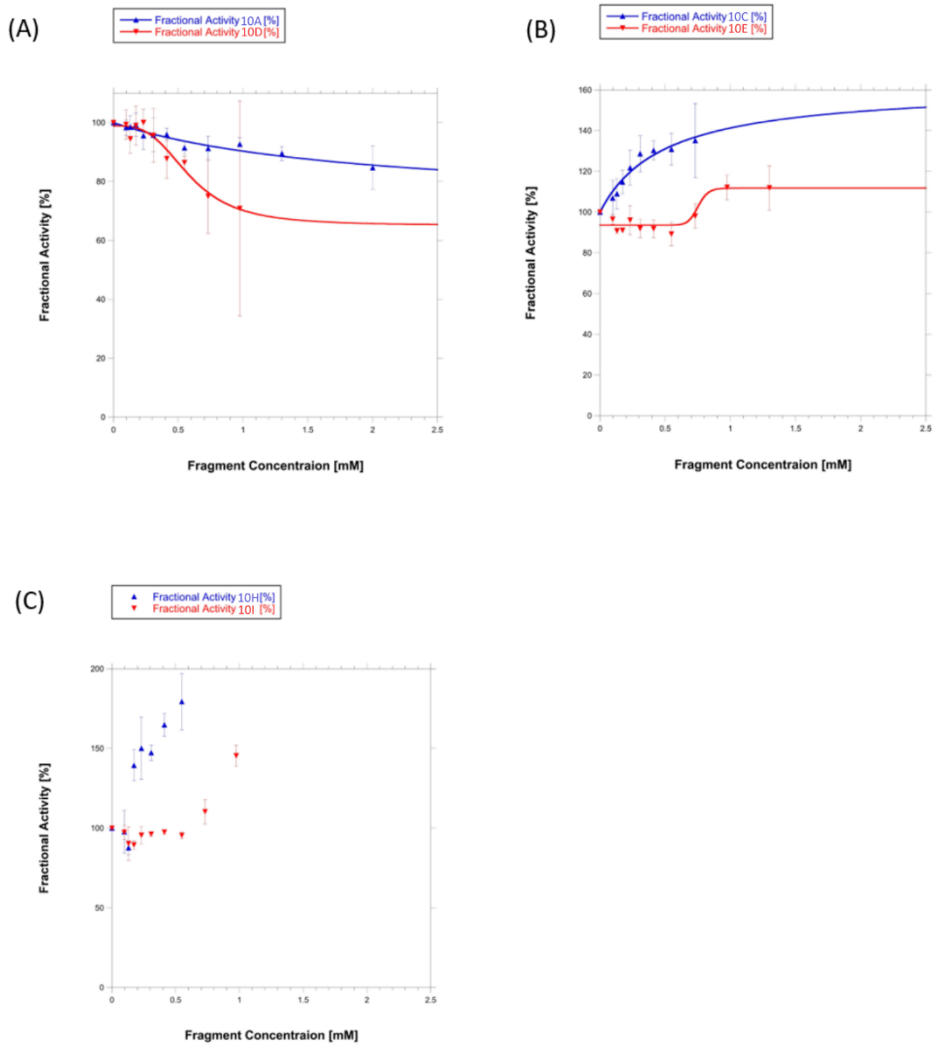


*Figure 4-48. Basal ATPase activities in the presence of analogues of fragment hit **10**. (A), (B) Analogues **10A**, **10C**, **10D**, and **10E** show either weak inhibition or weak activation effects on KifC1$_{307\text{-}663}$. The IC$_{50}$ values cannot be properly identified from the weak data. (C) Analogues **10H** and **10I** act as agonists. These data are more solid than the previous data. However, the IC$_{50}$ values could not be identified owing to poor solubility. Chemical structures of these analogues are shown in Supplementary material section 7.1.*

In summary, several analogue hits were identified using basal ATPase activity measurements. However, these hits did not show typical nonlinear regression inhibition curves that are normally observed during the measurement of ATPase activities of other kinesins. Fragments **4** and **9** inhibited $KifC1_{307\text{-}663}$ in a sigmoidal pattern, whereas analogues of fragments **2** and **5** activated the $KifC1_{307\text{-}663}$ basal ATPase activity, together implying that the ATPase assay is not yet suitable for measuring the weak-binding affinities of fragments at such high concentrations. These results will be further analysed in the discussion section.

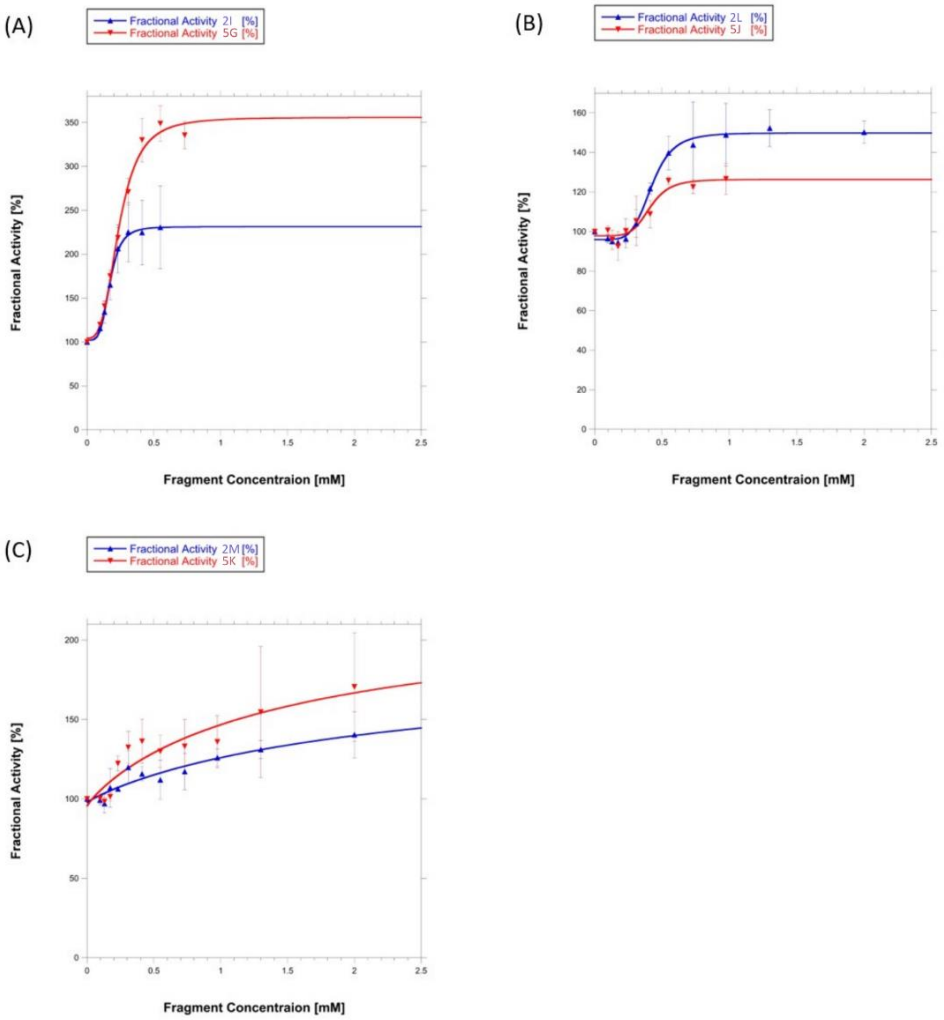
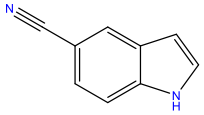
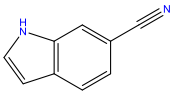
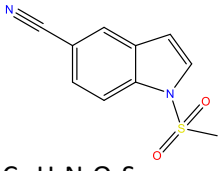
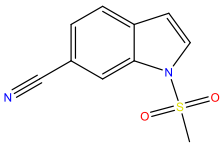
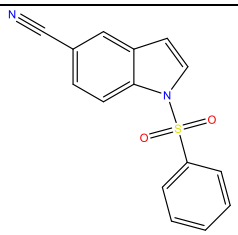
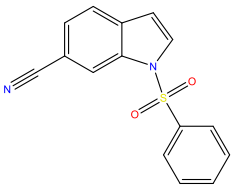
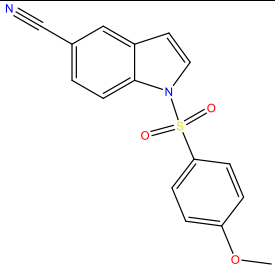
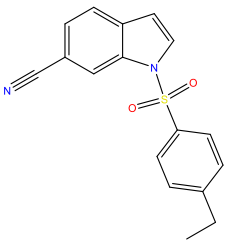


*Figure 4-49. Basal ATPase activities of fragment analogues **2** and **5**. Six analogues demonstrated activation effects on $KifC1_{307\text{-}663}$. **2I**, **2L**, and **2M** are isomers of **5G**, **5J**, and **5K**, and thus separated into three panels (A), (B) and (C) for comparison. (A), (B) Analogues **2I**, **2L**, **5G**, and **5J** increased the basal ATPase activity of $KifC1_{307\text{-}663}$ in a sigmoidal pattern. (C) The ATPase curves of **2M** and **5K** are atypical, indicating that their IC_{50} values are questionable.*

*Table 4-17. IC_{50} values of fragment analogues **2** and **5** measured using basal ATPase assays. Fragments **2**, **2I**, **2L**, and **2M** are isomers of fragments **5**, **5G**, **5J**, and **5K**, respectively. **2I, 2L, 2M, 5G, 5J**, and **5K** acted as agonists in ATPase activity measurement. *Not determined because of atypical curves.*

Number/ Commercial name	***Structure***	***IC_{50} (mM)***	***Number/ Commercial name***	***Structure***	***IC_{50} (mM)***
2 EN300-97267 (Enamine BB)	 $C_9H_6N_2$	none	***5*** EN300-100999 (Enamine BB)	 $C_9H_6N_2$	none
2I	 $C_{10}H_8N_2O_2S$	0.17 ± 0.01	***5G***	 $C_{10}H_8N_2O_2S$	0.24 ± 0.01
2L	 $C_{15}H_{10}N_2O_2S$	0.42 ± 0.01	***5J***	 $C_{15}H_{10}N_2O_2S$	0.41 ± 0.04
2M	 $C_{16}H_{12}N_2O_3S$	n.d.*	***5K***	 $C_{17}H_{14}N_2O_2S$	n.d.*

4.4.2.7 MST measurement of KifC1 fragments

TSA, a biophysical assay, was used to examine the KifC1 fragments. However, the results indicated that TSA was unsuitable for analysing weak-binding KifC1

fragments (Table 4-16). Here, another biophysical assay, MST, was applied to determine the binding affinities of the fragment hits and their analogues.

Among the ten original hits identified using NMR, fragments **1** and **10** showed severe adsorption/ligand-induced photobleaching changes in MST measurements, implying that they cannot be measured using MST; they were thus excluded from further measurements. Among fragments **2** to **9**, which were subjected to MST measurements, only fragment **8** showed a detectable binding affinity (Figure 4-50 A, B). The K_d value was 1.34 ± 0.43 mM, with a response amplitude of 20.9 and a signal-to-noise ratio of 12.4. In contrast, the other fragments had signal-to-noise ratios lower than 5 and, therefore, could not be identified as hits (Figure 4-50 C). The MST response was normalised to F_{Norm} using the MO.Affinity Analysis Software, and the curves were fitted using the K_d model.

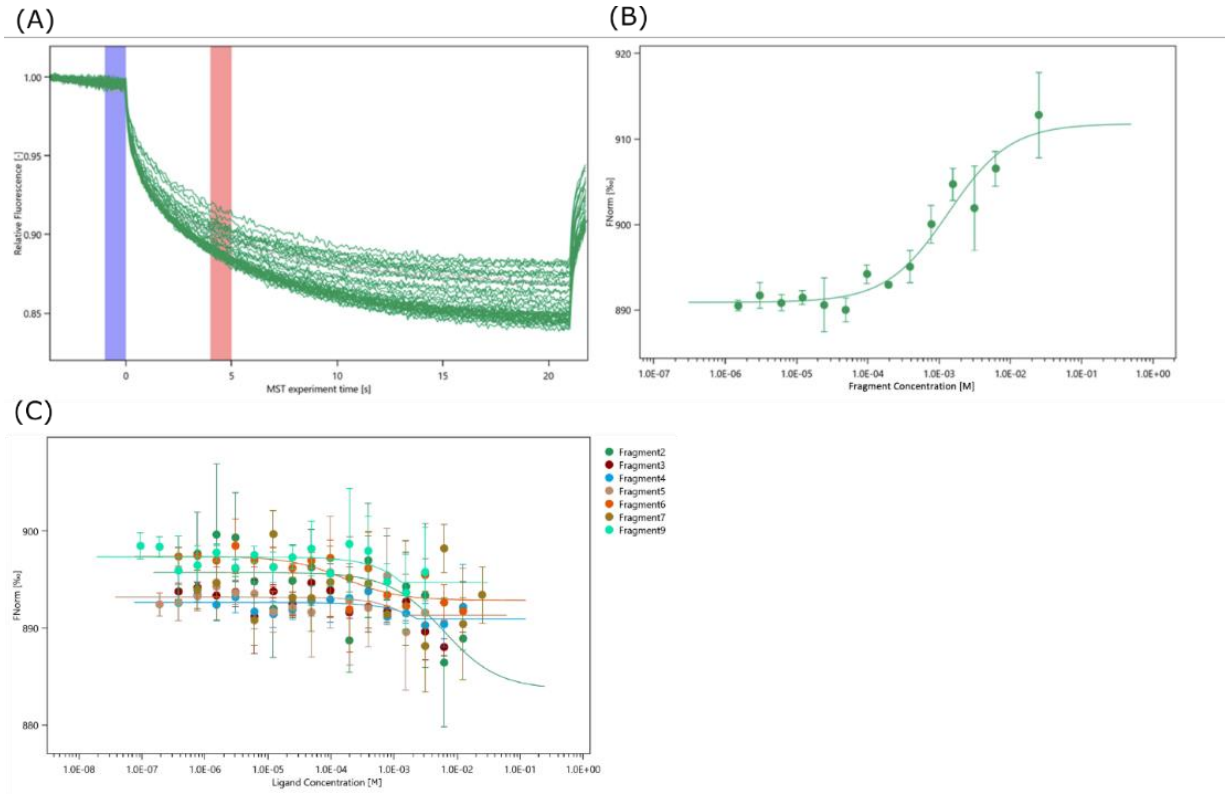


*Figure 4-50. Fragment **8** was identified as a hit using microscale thermophoresis (MST). (A) MST trace of fragment **8** showed changes in fluorescence upon activation and deactivation of the IR laser and demonstrated no aggregation or ligand-induced photobleaching rate changes. The relative fluorescence was read at 5 s. (B) Dose-response curve of fragment **8**. The MST response was normalised to FNorm and resulted in a K_d value of 1.34 ± 0.43 µM, with a response amplitude of 20.9 and a signal-to-noise ratio of 12.4. (C) The other fragments had signal-to-noise ratios lower than 5 and could therefore not be identified as hits.*

Aside from the original NMR hits, nineteen fragment analogues ordered from eMolecules were subjected to MST measurement. The analogues were first subjected to a binding check. Nine analogues (**2A, 2C, 2B, 2G, 5A, 5B, 6B, 8C,** and **9A**) (*Supplementary materials 7.1*) were identified as potential hits from the binding check and subjected to triplicate measurement of K_d values. During MST measurement, analogue **8C** was excluded because of ligand-induced photobleaching rate changes. Finally, two analogues, **2C** and **6B**, demonstrated decent binding, with K_d values of 23.06 ± 12.57 and 62.45 ± 12.52 µM,

respectively (Figure 4-51 B, D). The response amplitudes were 3.4 and 3.9, while the signal-to-noise ratios were 8.4 and 16.7 for **2C** and **6B**, respectively. The other analogues had signal-to-noise ratios lower than 5 and, therefore, could not be identified as hits (Figure 4-51 E). However, fragments **2C** and **6B** may likely be false-positive hits. Their K_d values are too good for fragments, and their K_d confidences are too large. Furthermore, the MST response amplitudes of fragments **2C** and **6B** are too low compared with those of fragment **8** (response amplitude 20.9). In summary, we did not identify any valid fragment hit from eMolecules based on MST measurements.

In summary, fragment **8** was identified as a hit via MST, with a K_d of 1.34 ± 0.43 mM, and will be subjected to further investigations in the future (Table 4-18).

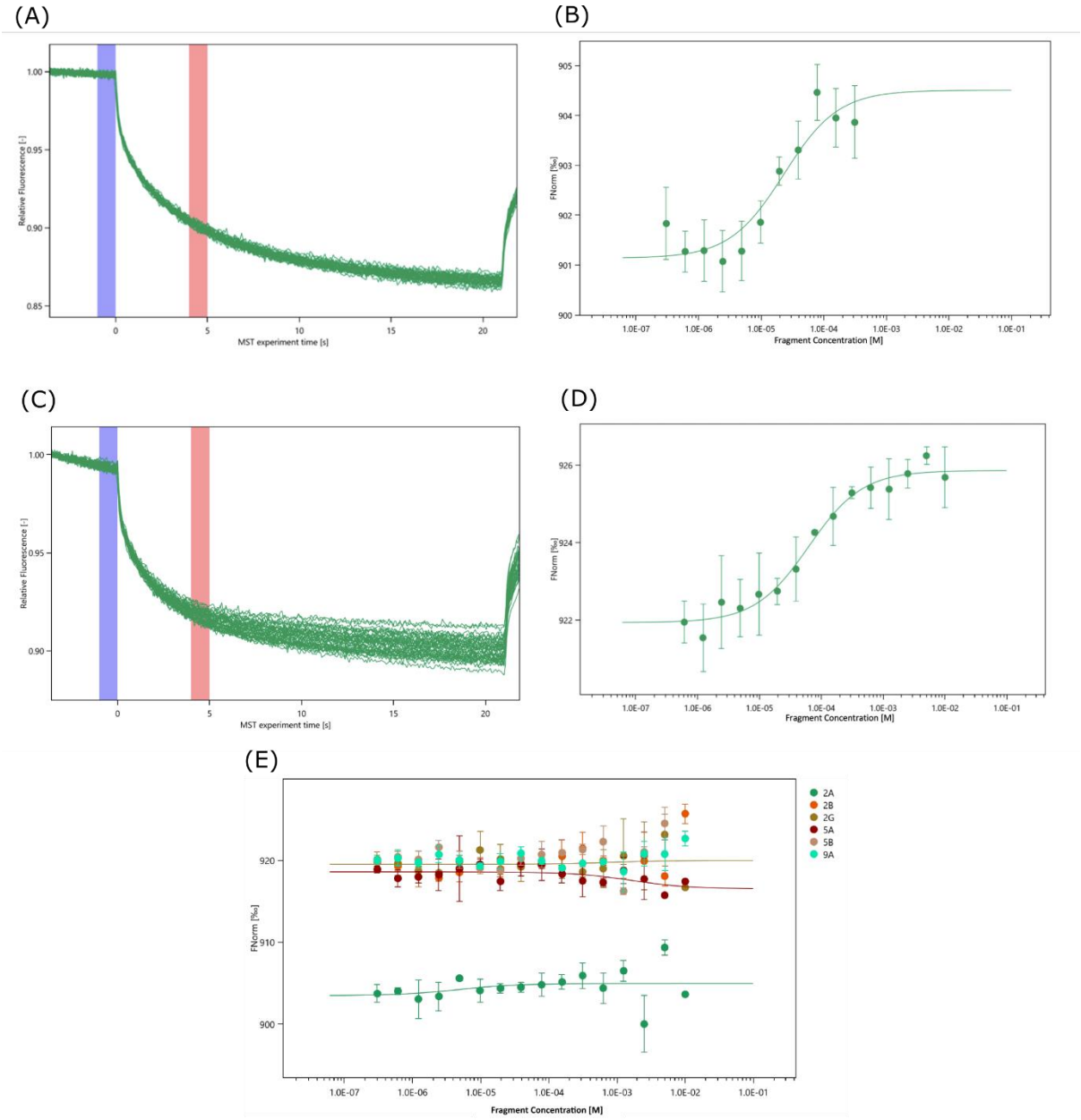
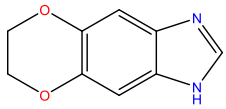
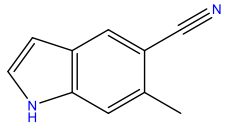
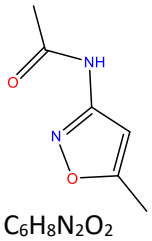


*Figure 4-51. MST measurements of KifC1 fragment analogues. (A) MST trace of **2C** showed no aggregation or ligand-induced photobleaching rate changes. The relative fluorescence was read at 5 s. (B) Dose-response curve of **2C**. The MST response indicated a K_d value of 23.06 ± 12.57 µM, with a response amplitude of 3.4 and a signal-to-noise ratio of 8.4. (C) The MST trace of **6B** demonstrated no aggregation or ligand-induced photobleaching rate changes. The relative fluorescence was read at 5 s. (D) Dose-response result of **6B**. The MST response indicated a K_d value of 62.45 ± 12.52 µM, with a response amplitude of 3.9 and a signal-to-noise ratio of 16.7. (E) The other fragments had signal-to-noise ratios lower than 5 and could not be identified as hits.*

*Table 4-18. Three hits were identified using MST. Fragment **8** showed a reasonable affinity and large response amplitude as well as signal-to-noise ratio, whereas fragments **2C** and **6B** were false-positive hits.*

Number/ Commercial name	***Structure***	***K_d (mM)***	***Response amplitude***	***Signal-to-noise ratio***
8 STL368348 (Vitas M Labs)	 $C_9H_8N_2O_2$	1.34 ± 0.43	20.9	12.4
2C PBLY8010 (Pharmablock)	 $C_{10}H_8N_2$	False-positive	3.4	8.4
6B 7113213850 (Otava)	 $C_6H_8N_2O_2$	False-positive	3.9	16.7

4.4.2.8 Attempts for crystallisation of KifC1$_{307\text{-}663}$

Several attempts were made to obtain KifC1$_{307\text{-}663}$ crystals. Various commercially available crystallisation screens were tested. The published crystallisation condition for the KifC1 mutant (PDB entry 5WDH) was also evaluated (*Supplementary materials 7.5*). However, all crystallisation trials were unsuccessful.

4.4.3 Discussion

KifC1 is a kinesin that is indispensable in some cancer cell lines and reproductive cells but is not necessary for most somatic cells [150]. Thus, inhibitors targeting KifC1 can probably selectively target cancer cells with minimal effects on somatic cells.

To date, three KifC1 inhibitors, AZ82, CW069, and SR31527, have been developed. However, none of them has progressed to clinical trials. AZ82 has been reported to specifically bind the KifC1-MT complex but not to KifC1 or MT alone. It binds to the KifC1-MT complex in an ATP-competitive manner and causes mitotic catastrophe in BT-549 breast cancer cells with amplified centrosomes [50]. Although AZ82 inhibits the MT-stimulated KifC1 ATPase activity, with an IC_{50} of 0.3 μM, it also demonstrates non-specific cytotoxic effects at concentrations above 4 μM [50]. Meanwhile, it is challenging to solve the KifC1-AZ82 complex for structure-based design as AZ82 binds to the KifC1-MT complex but not to KifC1 alone. These challenges have prevented further development of AZ82 as a potential cancer treatment. SR31527 directly binds to KifC1 without involving MTs. This binding has a K_d value of 25.4 nM, as measured using the Biolayer Interferometry (BLI) assays [151]. However, ST31527 shows severe off-target effects and kills almost all LL47 cells (human lung fibroblast line) at a 100 μM inhibitor concentration [151], which has led to the suspension of further developments. CW069 is reported to cause multipolar mitoses in N1E-115, BT549, and MDA-MB-231 breast cancer cells, with limited effects on MCF-7 cancer cells without extra centrosomes [51]. Furthermore, CW069 does not cause neutropenia in primary human bone marrow cells, a side effect that is commonly observed when inhibiting human Eg5 or CENP-E. However, among the three known KifC1 inhibitors, CW069 has the lowest

efficacy, with an IC_{50} value of 75 ± 20 μM. The absence of selectivity of CW069 to KifC1 over Eg5 is also an issue [51].

The work in this thesis aimed to develop novel inhibitors of KifC1. A previous student from our group tried to identify KifC1 hits using high-throughput screening (HTS). However, no hit was identified from the Maybridge Hitfinder library (Thermo Fisher), which contains more than 15,000 small molecules. Thus, identifying KifC1 hits has proven extremely challenging. AZ82 was developed by screening more than 800,000 compounds, followed by subsequent SAR analysis [152]. SR31527 was identified from a library containing 30,000 compounds [151], and CW069 was developed after exploring millions of compounds *in silico* followed by convergent syntheses [51]. Considering the difficulties in identifying KifC1 hits from HTS, the FBDD approach was chosen.

Twenty hits were identified via NMR fragment screening, and ten of these were selected for further analyses. These ten fragments were then subjected to TSA measurement. However, none of the fragments bonded to and stabilised the KifC1 motor domain. The measured T_i values in the presence of the fragments decreased compared with those in the control group, suggesting that TSA is not suitable for measuring the weak-binding KifC1 fragments.

The initial fragment hits were then subjected to basal ATPase activity measurement. Interestingly, in contrast to typical non-linear regression inhibition curves, which are frequently observed during the measurement of ATPase activities of other kinesins, fragments **4** and **9** showed perfectly fitted sigmoidal curves. Copeland *et al*. have reported that the concentration-response relationships of molecules often reflect non-ideal behaviours because of

stoichiometric binding or other biophysical factors [105]. In stoichiometric binding, especially situations where 1-to-1 binding does not properly describe the inhibition mechanism, a Hill equation: $100\%\text{inhibition} = \frac{100}{1+(\frac{IC50}{[I]})\text{^h}}$ is used to describe the concentration-response relationships, where h refers to the Hill coefficient, which describes the steepness of the curves. To fit the measured KifC1 inhibition curves more properly, the Hill equation was further modified to $v/v0 = F_{min} + \frac{Fmax - Fmin}{1 + (\frac{[I]}{IC50})\text{^h}}$ with two additional factors, F_{min} and F_{max}, introduced to describe the minimum and maximum fractional activities observed in the ATPase assay.

The Hill coefficients of fragments **4** and **9** were significantly larger than 1, indicating the stoichiometric binding of multiple ligands to the same target. Aside from the possibility of stoichiometric binding, the sigmoidal curve pattern can also be explained using the biophysical changes induced by protein denaturants or promiscuous inhibitors [105]. Such compounds do not behave properly by specifically interacting with a defined binding pocket on the enzyme molecule and are therefore, intractable as drug leads. These compounds can form micelles and inhibit enzyme function. Micelles show a very abrupt concentration-response plot, reflecting not the response of the enzyme to inhibition but rather the critical micellar concentration (CMC) of the compound. Detergents, chaotropic agents, aprotic and nonpolar solvents (e.g., DMSO, acetonitrile), and other nonspecific enzyme denaturants also display high Hill coefficients when titrated in enzyme assays. None of these inhibition mechanisms is tractable from a pharmacological perspective. Therefore, a high Hill slope in the concentration-response plot for a compound should cause some scepticism [105]. However, fragments **4** and **9** were less possible to cause inhibitions through biophysical changes for the following reasons: (1) promiscuous inhibitors should completely sabotage enzyme activity

with a F_{min} value close to zero. However, the enzymatic activities of $KifC1_{307-663}$ in the presence of fragments **4** and **9** decreased and reached plateaus of 52.4% and 16.2%, respectively; (2) Sigmoidal curves with high Hill slopes could also be observed in some analogues (**2I, 2L**, *etc.*) but in an increasing manner. Further, enzyme activities cannot be boosted by micelle formation.

In summary, fragments **4** and **9** showed inhibitory effects in the basal ATPase assay. However, as they showed sigmoidal inhibition curves, their inhibitory effects remain questionable until further investigation.

In Eg5 and CENP-E, the L5/α2/α3 pocket is frequently utilised as the inhibitor binding site. A recent paper reported a second allosteric binding pocket in KifC1, formed by helices α4 and α6 (Figure 4-52) [40]. Alignment of the α4/α6 binding pockets between KifC1 and Eg5 demonstrated significant differences in volume and the distribution of surface hydrophobicity and charges (*Introduction section 1.1.3.1*, Figure 1-6 B, C). The charged Arg521 of helix α3 interacts with other residues, stabilises loop L5, and maintains the closed conformation of the potential binding pocket, disfavouring the binding of molecules (especially AZ82) at this site (*Introduction section 1.1.3.1*, Figure 1-6 C). Although the author claimed that AZ82 can not bind to this site, relatively small fragments still have a chance to slip into the pocket as it is not completely closed. Thus, some fragments and analogues identified in our works can be assumed to bind to both L5/α2/α3 and α4/α6 pockets and synergistically inhibit or activate the enzymatic activity of $KifC1_{307-663}$.

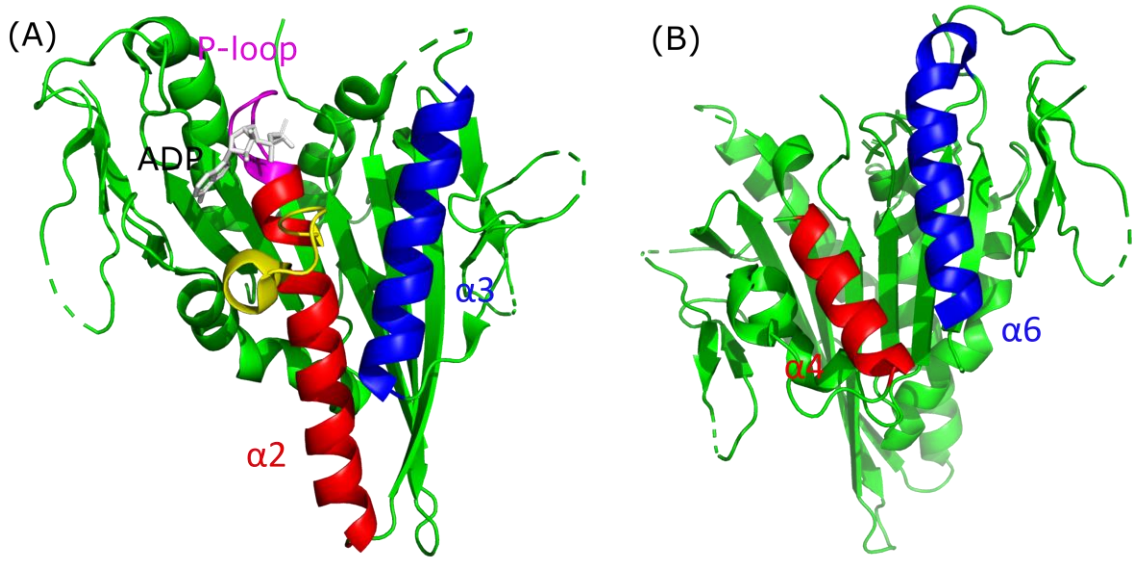


Figure 4-52. (A) L5/α2/α3 pocket in KifC1. The L5/α2/α3 pocket is frequently utilised for developing Eg5 and CENP-E inhibitors but is too small to harbour the KifC1 inhibitor AZ82. The p-loop is coloured in magenta, ADP is coloured in pale white, helix α2 is coloured in red, loop L5 is coloured in yellow, and helix α2 is coloured in blue. (B) α4/α6 pocket in KifC1. The α4/α6 pocket most likely harbours AZ82. Helix α4 is coloured in red, and helix α6 is shaded in blue.

Another interesting observation in the ATPase activity measurement of KifC1$_{307-663}$ was that some fragments and analogues functioned as agonists rather than antagonists. The work of a previous Ph.D. student from our group showed that a well-developed inhibitor, AZ82, also activates the enzymatic activity of KifC1$_{307-663}$ in the basal ATPase assay [153]. The inhibition could only be observed in MT-stimulated ATPase assays. It remains unclear why KifC1 inhibitors function as agonists in the basal ATPase assay. To date, there have been no KifC1-ligand complex structures available in the PDB. Our attempts to repeat the crystallisation process of the binary KifC1-ADP motor domain based on the published conditions failed. The author of the KifC1 crystal structure, Prof. Hewwon Park, told us personally that the mutation (T368P) in the KifC1 sequence could be pivotal to the crystallisation process, in addition to the significant difference in the N-terminal region.

In addition to the biochemical assay, MST was conducted to validate the potential fragment hits for $KifC1_{307\text{-}663}$. In terms of MST measurements, fragment **8** was considered a reasonable starting point for subsequent fragment-based drug discovery.

There are three strategies for improving fragment hits in FBDD, namely fragment growing, fragment hopping/merging, and fragment linking. However, the two latter approaches highly rely on the crystal structures of protein-ligand complexes, which are not available during the development of KifC1 inhibitors. The development of KifC1 fragments can only be conducted by fragment growing based on SAR analysis. In our study, analogues of the initial NMR hits were either purchased from chemical suppliers or synthesised by our collaboration partner. Several hits were identified (Table 4-17 and Table 4-18). For future development, additional analogues could be synthesised and tested. However, it is more important to crystallise the human KifC1 motor domain and redetermine its structure for applying a structure-based design.

5. Summary and future directions

5.1 Developing MPP1 inhibitors based on ascochitine and depsidone scaffolds

Recently, MPP1 has gained attraction as a potential cancer drug target owing to its crucial roles in hepatocellular [124, 125], bladder [60], colorectal [132], breast [133], renal cell [134], and pancreatic cancers [135]. MPP1 interacts with a set of cell cycle regulators, including Pin1 [57], PRC1 [60], FOXO38, and USP7 [131], directly or indirectly (*Human MPP1, Discussion 4.1.5.2*). However, its detailed mechanism in regulating cytokinesis remains unclear. To date, only two MPP1 inhibitors, norlobaridone (**D1**) and physodic acid (**D2**), have been reported by a previous member of our group [59]. Unfortunately, both **D1** and **D2** have stability issues, which necessitate examination of the effects of various substituents on the activity and stability of the scaffold.

We identified ascochitine as a novel chemical scaffold for developing MPP1 inhibitors, whereas one of its analogues, citrinin, showed no inhibition of MPP1. A set of ascochitine analogues as well as several intermediates were synthesised. The following ATPase activity measurement indicated that the two-ring system is vital for inhibitory activity and identified two improved hits, **A5d** and **A5e**. The drug-like properties of **A5d** and **A5e** were then profiled. Both analogues showed favourable chemical properties, which meet Lipinski's rules and show good CYP and hERG liabilities. However, the analogues had stability issues, as indicated by high clearance and low half-lives in human and mouse microsomal stability assays. Cell proliferation experiments indicated weak inhibitory effects of **A5d** and **A5e** in several cancer cell lines and weak synergistic effects with paprotrain,

an MKLP-2 inhibitor [121]. In summary, they were not found to be suitable for further development.

The difficulties in synthesising ascochitine analogues hinder the identification of more potent MPP1 analogues. Thus, we referred to Dr. Geoffrey Wells for the synthesis of analogues with fewer synthetic steps. The new analogues are isosteres to ascochitine, but are synthetically more available. The following ATPase activity measurements indicated that the analogues based on new chemical scaffolds did not inhibit the basal ATPase activity of MPP1. Instead, they only inhibited the MT-stimulated MPP1 ATPase activity. The new analogues also showed weaker inhibitory effects than those of ascochitine analogues, 5 to 8-folds less active than **A5d**. In summary, the ascochitine isosteres do not have much potential as MPP1 inhibitors.

Parallel work was also conducted on developing new MPP1 analogues based on the depsidone scaffold. Compound **D3**, **D4**, and **D5** with minimised depsidone scaffolds were characterised in the presence of $MPP1_{1\text{-}491}$, but showed no inhibition. We thus concluded that the 12-pentanoyl and 4-pentyl substitutes are crucial for the MPP1 inhibition.

To foster the development of MPP1 inhibitors, an MPP1 crystal structure is needed. In extended crystallisation experiments, we failed to acquire MPP1 crystals. Therefore, we tried to co-crystallise the MPP1 motor domain with αRep proteins, which increased the chances of obtaining crystals. This strategy has been reported to be effective for proteins resisting crystallising [111, 126-129]. Although we acquired several crystals, they were subsequently found to only be αReps rather than the αRep-MPP1 complex. Thus, we ordered vectors of the

αRep-MPP1 fusion protein, which are composed of an αRep covalently linked with MPP1. We have set up crystallisation trials and acquired several crystals recently and are now waiting for the measurements. Meanwhile, three MPP1 motor domain constructs were subjected to SAXS measurement and some structural information was acquired.

In summary, we identified two ascochitine analogues, **A5d** and **A5e**, showing low micromolar inhibition of MPP1. However, their drug-like properties indicated that they are not suitable for further development. The development of potential MPP1 inhibitors is thus currently halted until further knowledge of the crystal structure.

5.2 Crystal structure of human A33 Fab as a cornerstone for molecular dynamic simulations and further drug development

Human A33 antigen is a potential drug target for colorectal cancer [70, 71, 73]. Extended investigations have been performed to develop A33 antibodies as an immunotherapy and to humanise the variable region of murine mAb to reduce the immune response [73, 74]. However, the normal gut localisation and intrinsic stability issues in the variants caused continuous failure in clinical trials [140, 141]. To date, the structure of A33 Fab or its variants has never been reported. To provide a solid model for stability analyses, we disclosed the structure of an A33 Fab mutant, H/C226S, in two crystal forms.

During structural analysis, we found that the four separate regions of A33 Fab (V_L, C_L, C_{H1}, and C_L) are intrinsically rigid and that flexibility lies within the

switch region. The elbow angles of the models in two different crystal forms varied from 156° (triclinic form) to 145° (hexagonal form).

Further comparison of A33 Fab with certolizumab, an immunotherapeutic molecule developed from the same humanisation scaffold as A33 Fab, showed that in the variable regions, FWRs offer good support to CDRs while being neutral to scaffolding for antibody structural integrity. The scaffold of the A33 Fab is optimal to support CDRs targeting different antigens and could be applied to other therapeutic candidates from an engineering perspective. Furthermore, our partner has performed stability analyses based on this model, which identified the early stages of unfolding and stability-limiting regions of Fab A33 and the V_H and C_L domains as interesting future targets for engineering stability to both pH- and thermal stresses simultaneously [77].

5.3 Developing selective DYRK2 inhibitors

DYRK2 is considered a potential TNBC drug target; however, its mechanism in promoting TNBC was unknown [99]. Recently, *in vivo* and *in vitro* studies were conducted by our collaboration partner Dr. Laureano de la Vega, who disclosed the following mechanism. The proteotoxic stress in TNBC stabilises DYRK2, which, in turn, phosphorylates and activates heat shock factor protein 1 (HSF1), which further triggers a downstream mechanism to guard the cancer proteome against misfolding and aggregation (to be published). DYRK2 intervention using small molecules can validate this target as a new potential approach to treat patients with TNBC, either as a mono- or combined therapy.

Several inhibitors have been developed by our collaboration partner Prof. Simon Mackay (Strathclyde University, Scotland) using SAR-based drug development. The inhibitors with the highest efficacy and best selectivity were sent to us for co-crystallisation with DYRK2. During data processing, we found that the DYRK2-inhibitor datasets could be processed in both the $C222_1$ and $C12_1$ space groups. Further comparison between the densities and solved models indicated that $C12_1$ would be the correct solution. Unfortunately, we could not identify the correct conformation of CI641 in the crystal structure owing to a low resolution of 3.5 Å. The $DYRK2_{73-478}$-CI709 dataset could not be processed properly because of low completeness, a result of incomplete data collection. Thus, only two crystal structures were determined.

By analysing the interactions between the inhibitors LMB035/LMB017 and DYRK2, we summarised that selective binding to DYRK2 active site requires four essential binding elements of the inhibitor, namely H-bonding with the hinge, halogen bonding with the buried water molecule, hydrophobic interactions with benzimidazole mainly via Val163, and a critical shape and size limit for the benzimidazole-2 substituent to fit the G-pocket. Interestingly, we found that LMB035 binds twice to DYRK2 in two distinct pockets. A second binding pocket was identified between helices αD and αE, which may explain the higher binding affinity of LMB035 than that of other analogues.

Finally, experiments to elucidate the selectivity of LMB035/LMB017 to DYRK2 over DYRK1A via docking simulation and binding pocket comparison were conducted by Prof. Simon Mackay. The docking simulation indicated that the following two elements are vital for selectivity: 1. halogen bonding with the buried water molecule and 2. access to the G-pocket, which is regulated by Phe160.

We further aligned the $DYRK2_{73-478}$-LMB035/$DYRK2_{73-478}$-LMB017 models with other available DYRK1A and DYRK2 structures to investigate their ATP-active sites. Interestingly, the G-loop/Phe160 adopted unique conformations that have never been reported before. LMB035 and LMB017 seemed to push the G-loop away during the binding process, and the β2 strand moved closer to the binding pocket. The Phe160 in the LMB017 model completely flipped to the other side. It remains unknown whether the same conformational change occurs in DYRK1A. If the β2/G-loop/Phe170 in DYRK1A is rigid, it will severely clash with LMB035 and LMB017 and reject their binding, which may finally result in their selectivity.

5.4 Dilemma and future directions in developing KifC1 inhibitors

Halting the cell cycle in mitosis and interfering with its normal progression is one of the most successful anti-cancer strategies. Over the past decades, various studies have investigated mitotic kinesins for their regulatory roles during the cell cycle. Among mitotic kinesins, Eg5 is the most comprehensively investigated drug target. Several Eg5-targeted drugs have entered phase I, II, and III clinical trials either as monotherapies or in combination with other drugs. Unfortunately, most of them have failed owing to low efficacy or adverse side effects. For example, AZD4877, a urothelial cancer drug, failed phase II trials due to low efficacy [154], whereas MK-0731, a treatment for solid tumours, produced grade 4 toxicities of myelosuppression [155]. The most successful Eg5 inhibitor to date is filanesib, which has successfully passed phase II trials with encouraging results [156] and entered phase III trials against multiple myeloma [34]. Apart from those of Eg5, recent studies have revealed the cancer-related roles of other kinesins, including KifC1 and MPP1.

KifC1 is utilised by some cancer cell lines to cluster supernumerary centrosomes and form pseudo-bipolar spindles, which subsequently avoid multipolar mitoses and apoptosis [45-47]. Although KifC1 is indispensable in some cancer cell lines and reproductive cells, it is not necessary for most somatic cells, which makes it an appealing chemotherapeutic target [48]. A limited number of KifC1-targeting inhibitors have been developed and tested *in vitro*. However, none of them have progressed to clinical trials because of low efficacy or cytotoxic effects (*Developing KifC1 inhibitors, Discussion 4.4.3*). Thus, KifC1 inhibitors with novel chemical scaffolds are needed to improve potency and reduce cytotoxicity.

In this thesis, several KifC1 fragments have been identified and characterised using different biophysical methods. As the KifC1-ligand complex structures were impossible to acquire, we tried improving the KifC1 fragment hits through fragment growth using SAR-by-catalogue and synthesis in collaboration with Dr. Geoff Well's group. However, neither biochemical assays such as ATPase activity assays nor biophysical assays, including TSA and MST, generated solid results that would allow an initial SAR study. In ATPase assays, fragments **4** and **9** showed inhibitory effects in the presence of $KifC1_{307\text{-}663}$. However, the inhibitory effect remains questionable because the inhibition curves had a sigmoidal shape (*Developing KifC1 inhibitors, Discussion 4.4.3*). Meanwhile, MST measurement led to the identification of three hits, fragments **2C**, **6B**, and **8**, whereas the binding of fragments **4** and **9** to $KifC1_{307\text{-}663}$ was not confirmed. In MST measurement, fragment **8** showed the highest affinity, with a K_d of 1.34 ± 0.43 mM. In summary, developing KifC1 inhibitors by growing fragments without structural data is very challenging. SAR-by-catalogue development of KifC1 inhibitors cannot be conducted using the current tools, possibly owing to a lack of sufficient commercially available fragment analogues. Therefore, a

KifC1-inhibitor crystal structure is needed before the project can continue. Future work should include ordering an identical expression construct to the published KifC1 structure and attempt reproducing crystallisation [40].

6. Conclusions

Over the past decades, various novel cancer therapies have been developed. These therapies are applied either alone or in combination with other drugs. Examples of cancer therapies can be classified into several groups: (1) conventional chemotherapeutic molecules (cytostatic drugs), (2) small molecule inhibitors (SMIs), (3) therapeutic mAbs including antitumor mAbs, anti-angiogenesis mAbs, and checkpoint inhibitor mAbs, chimeric antigen-specific receptor (CAR)-transfected T-cells (CAR-T cells), (5) antitumor vaccines, and (6) oncolytic viruses (OVs) [157]. In this thesis, both SMIs and mAb approaches have been investigated.

We first investigated small molecule inhibitors targeting key proteins such as human MPP1, KifC1, and DYRK2, which are involved in oncogenic signal transduction [37, 134, 135, 158]. Compared with cytostatic drugs interfering with cell proliferation, SMIs are targeted drugs that can specifically target cancer cells but spare normal cells, hence having high potency and low toxicity [159]. The SMIs development projects in this thesis were conducted following the rational drug discovery cycle, which represents the process of finding new medications based on the knowledge of a biological target [160].

In the MPP1 project, a recently discovered natural product lead compound, ascochitine, was first characterised using ATPase assays. In the absence of structural information for human MPP1, a set of ascochitine analogues was synthesised and their inhibitory effects were characterised against MPP1. Finally, two potent analogues, **A5d** and **A5e**, with better IC_{50} values compared to the lead compound were subjected to cell-based assays and drug-like property analyses.

However, these analyses indicated that the two compounds were not sufficiently cell permeable and displayed liabilities and were therefore not suitable for further development. We also attempted to crystallise MPP1 for subsequent structure determination using X-ray crystallography for structure-based drug design, which did not succeed. In the future, we could try to co-crystallise **A5d** or **A5e** with MPP1, as the compound may stabilise MPP1 and increase the chance to obtain crystals. In parallel, several depsidone analogues, a recently reported natural product MPP1 inhibitor were synthesised and characterised, with the conclusion that the 12-pentanoyl and 4-pentyl substitutes are crucial for their activity. Due to these multiple challenges, the drug discovery project on human MPP1 will be terminated and the results will be published.

In the DYRK2 project, our partners conducted an HT screen followed by hit optimisation. They subsequently summarised a SAR map for DYRK2 inhibitors and selected four analogues for subsequent crystallisation experiments. We then co-crystallised Dyrk2 in the presence of these analogues and determined three complex structures. The protein-ligand interactions were analysed and the pharmacophore features essential for DYRK2 binding were described. Additionally, we identified a novel allosteric binding pocket in DYRK2, which should be verified in the futre by additional biophysical experiments. Afterwards, binding pocket comparisons and docking simulations were conducted by Prof. Simon Mackay, with the conclusion that halogen bonding with the buried water molecule and access to the G-pocket are vital for selectivity. The DYRK inhibitors previously reported were generally non-specific [100, 101]. Here we reported the first selective DYRK2 inhibitors and elucidated the selectivity mechanism for future structure-based drug design. Meanwhile, the *in vivo* and *in vitro* studies conducted by our partner Dr. Laureano de la Vega demonstrated that DYRK2 is an important positive upstream regulator in the cytoprotective stress

response pathway. In conclusion, the selective DYRK2 inhibitors may contribute to the development of potential TNBC SMIs.

In the KifC1 project, twenty fragment hits were identified using complementary NMR screens. Due to the absence of solid structural data, we attempted to improve the KifC1 fragment hits through fragment growthing using a SAR-by-catalogue strategy. We either ordered analogues from chemical retailers or synthesised them in collaboration with Dr. Geoff Well's group. Although three different biophysical and biochemical assays (ATPase activity, TSA and MST) were employed to characterise the NMR hits and their analogues, no compound was found suitable for further development. Moreover, the results from different assays did not agree with each other. Several attempts to develop KifC1 inhibitors (AZ82, CW069 and SR31527) have been reported in the absence of structural data [51, 151, 152]. However, none of these lead to KifC1 inhibitors with potential for clinical studies as described in *Discussion 4.4.3*. In conclusion, developing KifC1 inhibitors by growing fragments without structural data is very challenging. KifC1-fragment crystal structures are needed before the project can continue. In the future, we could attempt to co-crystallise KifC1 with SR31527, which has been reported to directly bind to KifC1 rather than to the KifC1-MT complex [151].

Therapeutic mAbs also represent successful cancer treatments. According to the FDA, there are approximately 80 distinct therapeutic mAbs in clinical use, including 30 mAbs for the treatment of cancer [161]. In this thesis, we reported the crystal structure of human A33 Fab, a potential therapeutic antibody, in two distinct space groups. A33 Fab is a humanised Fab that was developed using CDR grafting technology but failed to reach the market because of toxicity and stability issues [71, 162]. Subsequently, the A33 Fab crystal structure has been used for *in*

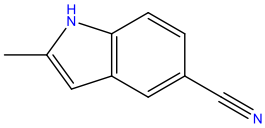
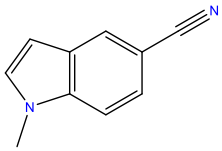
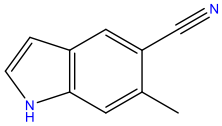
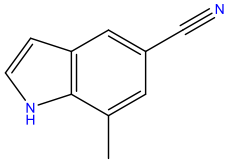
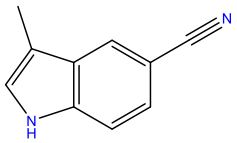
silico protein stability investigations [77, 163]. We are currently preparing a manuscript on the human A33 Fab based on the crystal structures and relevant SAXS solution data together with our collaboration partners. The crystal structure of A33 Fab combined with biophysical data may contribute to epitope-directed mAb design in the future.

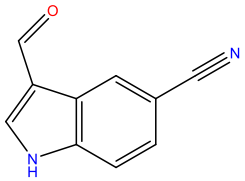
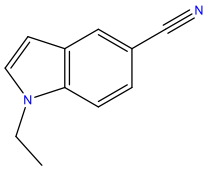
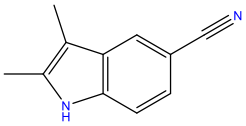
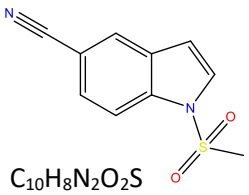
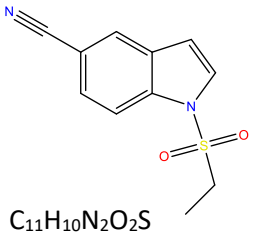
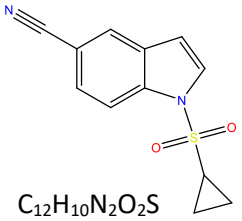
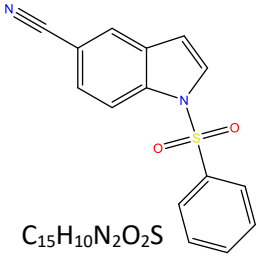
Despite their severe side effects, chemothereutic drugs remain the main treatment option for cancer. Examples, still in use, are melphalan, chlorambucil, cyclophosphamide, ifosfamid, and many others [157]. Compared with chemotherapy, targeted therapy including SMIs and mAbs have higher tumor specificity and lower toxicity. The first part of this thesis describes the SMIs projects of MPP1, DYRK2, and KifC1, representing the modern aspect of rational drug design. By December 2020, 89 SMIs have been approved by the FDA and the National Medical Products Administration (NMPA) of China [159]. SMIs have advantages in pharmacokinetic (PK) properties, costs, patient compliance, and drug storage and transportation. The second part of this thesis is the A33 Fab project, which described another modern aspect of cancer drug development, focusing on the tumor-bearing host organism and its immune system. Compared with SMIs, mAbs exert only mild side effects and are usually well tolerated [164]. In conclusion, all four projects in this thesis represent parts of the modern cancer drug design process and together contribute to a better understanding of these potential targets in the future.

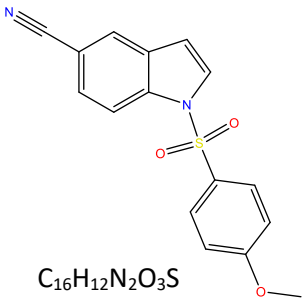
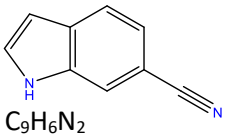
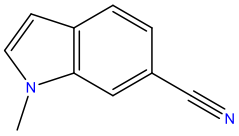
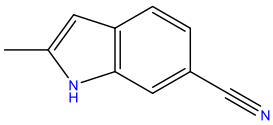
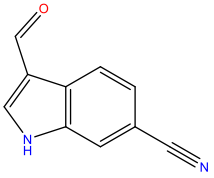
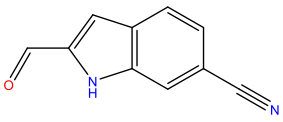
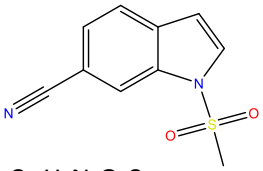
7. Supplementary Material Section

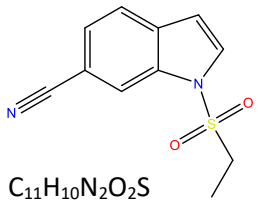
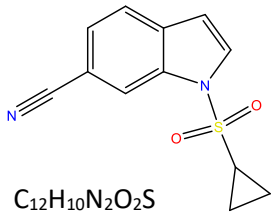
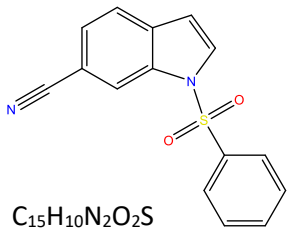
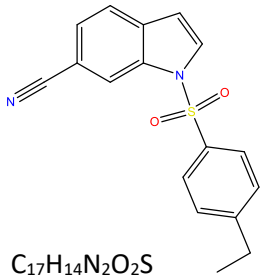
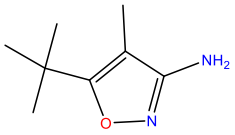
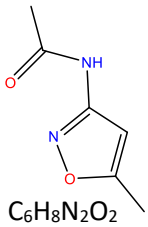
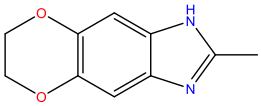
7.1 KifC1 analogues

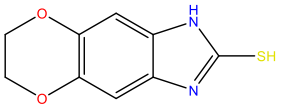
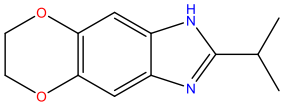
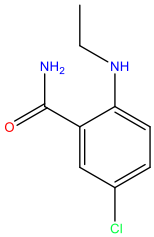
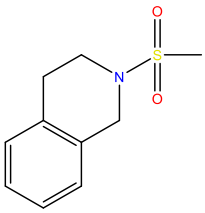
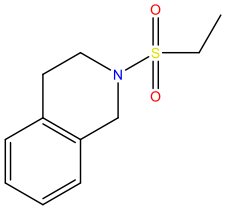
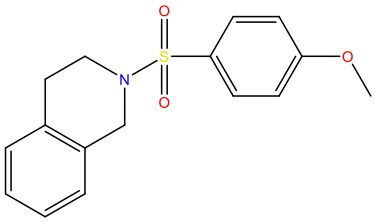
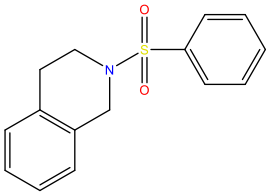
*Table 7-1. KifC1 fragment analogues purchased from commercial suppliers or synthesised by Mr. Helal based on the structures of the NMR hits. *Synthesised by Mr. Helal. Molecular weight, MW; hydrogen-bond donors, HBD; hydrogen-bond acceptors, HBA; polar surface area, PSA.*

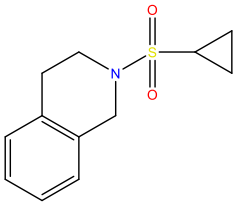
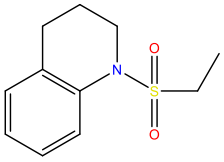
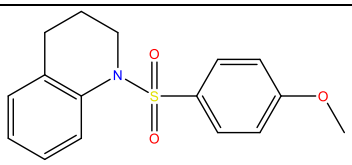
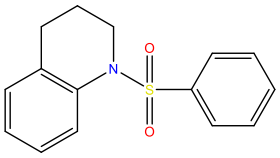
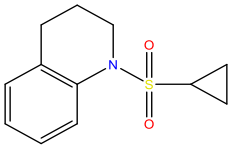
Number/ Commercial name	***Structure***	***MW (Da)***	***cLogP***	***Number of HBA/ HBD***	***PSA (Å²)***
2A CS-0102115 (ChemScene)	 $C_{10}H_8N_2$	156.2	1.94	2/1	35.82
2B LN00005444 (WuXi AppTec)	 $C_{10}H_8N_2$	156.2	1.84	2/0	27.03
2C PBLY8010 (Pharmablock)	 $C_{10}H_8N_2$	156.2	2.10	2/1	35.82
2D SL-05640 (Sinova)	 $C_{10}H_8N_2$	156.2	2.10	2/1	35.82
2E CS-0103492 (ChemScene)	 $C_{10}H_8N_2$	156.2	2.10	2/1	35.82

2F A176749 (Ambeed)	 $C_{10}H_6N_2O$	170.2	1.36	3/1	52.89
2G 4033033 (ChemBridge)	 $C_{11}H_{10}N_2$	170.2	2.18	2/0	27.03
2H EN300-49912 (Enamine)	 $C_{11}H_{10}N_2$	170.2	2.43	2/1	35.82
2I*	 $C_{10}H_8N_2O_2S$	220.2	0.87	3/0	61.17
2J*	 $C_{11}H_{10}N_2O_2S$	234.2	1.38	3/0	61.17
2K*	 $C_{12}H_{10}N_2O_2S$	246.3	1.51	3/0	61.17
2L*	 $C_{15}H_{10}N_2O_2S$	282.3	3.06	3/0	61.17

2M*	 $C_{16}H_{12}N_2O_3S$	312.3	2.93	4/0	70.4
5A AX146297 (Aldlab Chemicals)	 $C_9H_6N_2$	142.2	1.61	2/1	35.82
5B A103498 (Ambeed)	 $C_{10}H_8N_2$	156.2	1.84	2/0	27.03
5C EN300-4289085 (Enamine)	 $C_{10}H_8N_2$	156.2	1.94	2/1	35.82
5D A095037 (Aldlab Chemicals)	 $C_{10}H_6N_2O$	170.2	1.36	3/1	52.89
5F AC749692 (AbovChem)	 $C_{10}H_6N_2O$	170.2	1.41	3/1	52.89
5G*	 $C_{10}H_8N_2O_2S$	220.3	0.87	3/0	61.17

5H*	 $C_{11}H_{10}N_2O_2S$	234.3	1.38	3/0	61.17
5I*	 $C_{12}H_{10}N_2O_2S$	246.3	1.51	3/0	61.17
5J*	 $C_{15}H_{10}N_2O_2S$	282.3	3.06	3/0	61.17
5K*	 $C_{17}H_{14}N_2O_2S$	310.4	3.96	3/0	61.17
6A 33746 (AstaTech)	 $C_8H_{14}N_2O$	154.2	2.37	3/1	47.61
6B 7113213850 (Otava)	 $C_6H_8N_2O_2$	140.1	0.17	3/1	50.69
8A EN300-06637 (Enamine)	 $C_{10}H_{10}N_2O_2$	190.2	0.98	4/1	42.85

8B BBL010809 (Vitas M Labs)	 $C_9H_8N_2O_2S$	208.2	1.48	4/2	42.85
8C BBL026495 (Vitas M Labs)	 $C_{12}H_{14}N_2O_2$	218.3	2.04	4/1	42.85
9A CTBB-005833 (Chemtellect Inc)	 $C_9H_{11}ClN_2O$	198.7	1.33	2/2	55.12
10A*	 $C_{10}H_{13}NO_2S$	211.3	0.84	2/0	37.38
10B*	 $C_{11}H_{15}NO_2S$	225.3	1.35	2/0	37.38
10C*	 $C_{16}H_{17}NO_3S$	303.4	2.90	3/0	46.61
10D*	 $C_{15}H_{15}NO_2S$	273.4	3.02	2/0	37.38

					
10E*	$C_{12}H_{15}NO_2S$	237.3	1.48	2/0	37.38
					
10F*	$C_{11}H_{15}NO_2S$	225.3	0.00	2/0	37.38
					
10G*	$C_{16}H_{17}NO_3S$	303.4	2.61	3/0	46.61
					
10H*	$C_{15}H_{15}NO_2S$	273.4	2.48	2/0	37.38
					
10I*	$C_{12}H_{15}NO_2S$	237.3	0.00	2/0	37.38

7.2 Buffer screening results of KifC1$_{307-663}$ using the thermal shift assay (TSA)

*Table 7-2. T_i values of KifC1$_{307-663}$ measured using thermal shift assays (TSA) with the Rubic Buffer Screen. *Measurements with significantly atypical curves where T_i values cannot be properly calculated are marked as '0.0'.*

Buffer	Ti [°C]	pH	Salt (sodium chloride) [M]
Water	51.9 ± 0.1	7.0	0
100 mM Citrate	0.0	4.0	0
100 mM Na acetate	42.5 ± 0.9	5.0	0
100 mM Citrate	0.0	5.0	0
100 mM MES	46.8 ± 0.1	6.0	0
100 mM K phosphate	50.1 ± 0.1	6.0	0
100 mM Citrate	55.1 ± 0.1	6.0	0
100 mM Bis-Tris	41.7 ± 0.5	6.5	0
100 mM MES	55.6 ± 0.1	6.5	0
100 mM Na2 phosphate	0.0	7.0	0
100 mM K phosphate	57.2 ± 0.1	7.0	0
100 mM HEPES	0.0	7.0	0
100 mM MOPS	51.1 ± 0.1	7.0	0
100 mM Am acetate	49.6 ± 0.1	7.3	0
100 mM Tris-HCl	48.8 ± 0.1	7.5	0
100 mM Na2 phosphate	53.6 ± 0.3	7.5	0
100 mM Imidazole	0.0	7.5	0
100 mM HEPES	53.8 ± 0.2	8.0	0
100 mM Tris-HCl	52.1 ± 0.1	8.0	0
100 mM Tricine	54.2 ± 0.1	8.0	0
100 mM BICINE	54.3 ± 0.2	8.0	0
100 mM BICINE	54.4 ± 0.2	8.0	0
100 mM Tris-HCl	53.1 ± 0.1	8.5	0

100 mM CHES	49.7 ± 0.1	9.0	0
Water	52.4 ± 0.1	7.0	0.298
100 mM Citrate	0.0	4.0	0.298
100 mM Na acetate	0.0	4.5	0.298
100 mM Citrate	0.0	5.0	0.298
100 mM MES	46.0 ± 0.3	6.0	0.298
100 mM K phosphate	48.1 ± 0.5	6.0	0.298
100 mM Citrate	51.8 ± 0.1	6.0	0.298
100 mM Bis-Tris	49.0 ± 0.1	6.5	0.298
100 mM MES	48.5 ± 0.1	6.5	0.298
100 mM Na2 phosphate	53.4 ± 0.2	7.0	0.298
100 mM K phosphate	53.4 ± 0.3	7.0	0.298
100 mM HEPES	51.7 ± 0.1	7.0	0.298
100 mM MOPS	51.4 ± 0.1	7.0	0.298
100 mM Am acetate	49.5 ± 0.3	7.3	0.298
100 mM Tris-HCl	51.2 ± 0.3	7.5	0.298
100 mM Na2 phosphate	53.1 ± 0.2	7.5	0.298
100 mM Imidazole	0.0	7.5	0.298
100 mM HEPES	53.2 ± 0.1	8.0	0.298
100 mM Tris-HCl	52.3 ± 0.1	8.0	0.298
100 mM Tricine	53.6 ± 0.2	8.0	0.298
100 mM BICINE	53.6 ± 0.4	8.0	0.298
100 mM BICINE	53.6 ± 0.1	8.5	0.298
100 mM Tris-HCl	0.0	8.5	0.298
100 mM CHES	50.4 ± 0.1	9.0	0.298
100 mM SPG	0.0	4.0	0
100 mM SPG	0.0	4.5	0
100 mM SPG	44.2 ± 0.3	5.0	0
100 mM SPG	48.4 ± 0.2	5.5	0
100 mM SPG	51.2 ± 0.1	6.0	0
100 mM SPG	52.2 ± 0.1	6.5	0

100 mM SPG	53.3 ± 0.1	7.0	0
100 mM SPG	54.4 ± 0.1	7.5	0
100 mM SPG	54.2 ± 0.1	8.0	0
100 mM SPG	54.2 ± 0.1	8.5	0
100 mM SPG	54.4 ± 0.1	9.0	0
100 mM SPG	53.5 ± 0.1	10.0	0
20 mM HEPES	51.8 ± 0.2	7.5	0
50 mM HEPES	52.1 ± 0.1	7.5	0
125 mM HEPES	52.4 ± 0.1	7.5	0
250 mM HEPES	53.0 ± 0.3	7.5	0
20 mM Na2 phosphate	53.5 ± 0.1	7.5	0
50 mM Na2 phosphate	53.5 ± 0.1	7.5	0
125 mM Na2 phosphate	53.6 ± 0.1	7.5	0
250 mM Na2 phosphate	54.3 ± 0.1	7.5	0
20 mM Tris-HCl	51.9 ± 0.1	8.0	0
50 mM Tris-HCl	52.3 ± 0.1	8.0	0
125 mM Tris-HCl	52.5 ± 0.1	8.0	0
250 mM Tris-HC	50.9 ± 0.1	8.0	0
50 mM HEPES	52.6 ± 0.2	7.5	0.06
50 mM HEPES	53.0 ± 0.1	7.5	0.149
50 mM HEPES	52.5 ± 0.1	7.5	0.298
50 mM HEPES	52.8 ± 0.2	7.5	0.595
50 mM HEPES	53.8 ± 0.1	7.5	0.893
50 mM HEPES	54.0 ± 0.1	7.5	1.19
50 mM Tris-HCl	52.6 ± 0.1	8.0	0.06
50 mM Tris-HCl	52.7 ± 0.1	8.0	0.149
50 mM Tris-HCl	52.5 ± 0.1	8.0	0.298
50 mM Tris-HCl	52.5 ± 0.3	8.0	0.595
50 mM Tris-HCl	53.9 ± 0.1	8.0	0.893
50 mM Tris-HCl	55.0 ± 0.1	8.0	1.19
50 mM MES /Bis-Tris	48.7 ± 0.1	6.0	0

50 mM MES /Imidazole	0.0	6.5	0
50 mM Bis-Tris /PIPES	49.7 ± 0.2	6.5	0
50 mM MOPS /Bis-Tris propane	0.0	7.0	0
50 mM Na phosphate /Citrate	52.2 ± 0.5	7.5	0
50 mM MOPS /Na HEPES	0.0	7.5	0
0.1 M BICINE /Tris	53.3 ± 0.5	8.5	0
50 mM Imidazole	0.0	7.5	0.119
125 mM Imidazole	0.0	7.5	0.119
250 mM Imidazole	0.0	7.5	0.119
350 mM Imidazole	0.0	7.5	0.119
500 mM Imidazole	0.0	7.5	0.119

7.3 Thermal shift assay (TSA) measurements of KifC1 fragments

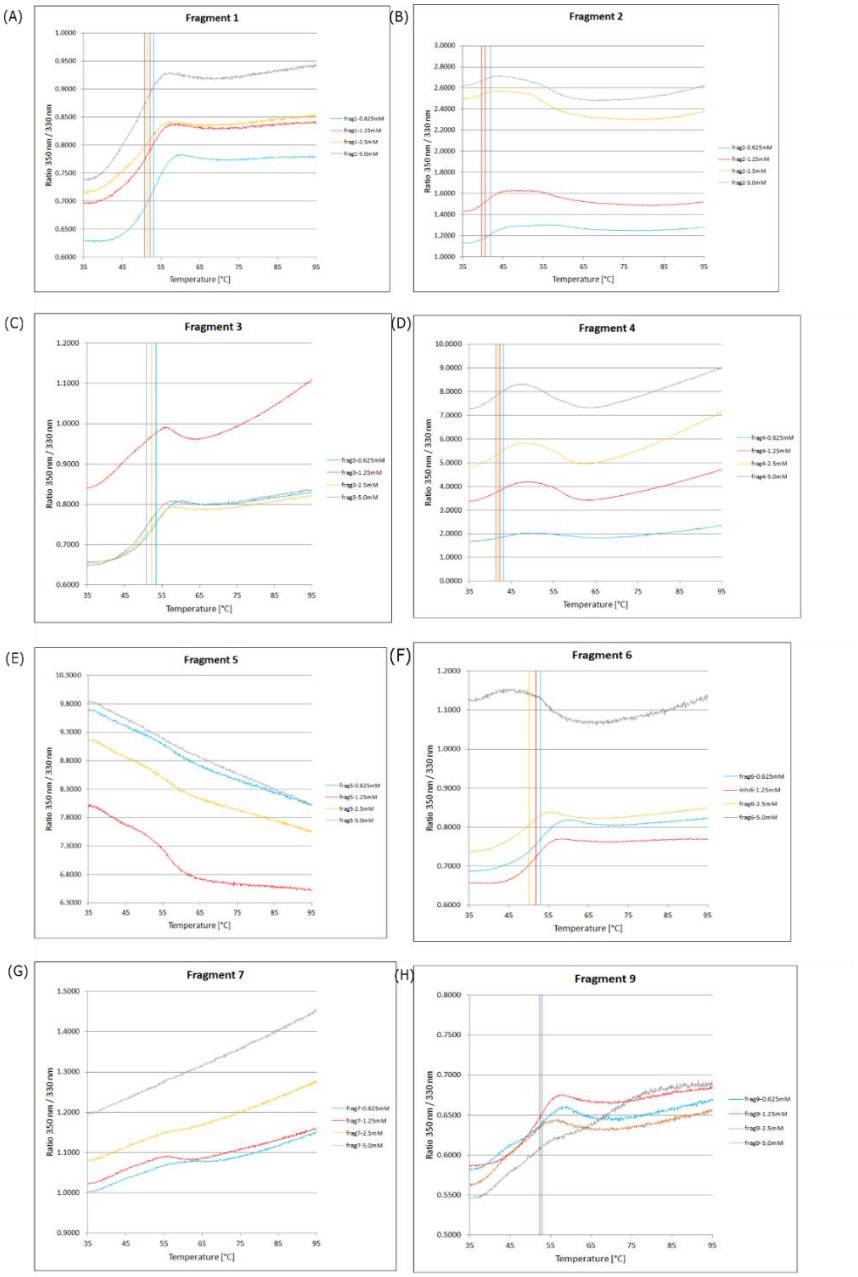


Figure 7-1. Thermoshift measurements of KifC1 fragments. In theory, the fragments should stabilise the protein and increase the Ti values. However, reduced Ti values and atypical curves were identified in presence of the fragments, indicating that the thermal shift assay (TSA) is not suitable for measuring KifC1 fragments.

7.4 Inhibition of KifC1 fragments measured using basal ATPase assays

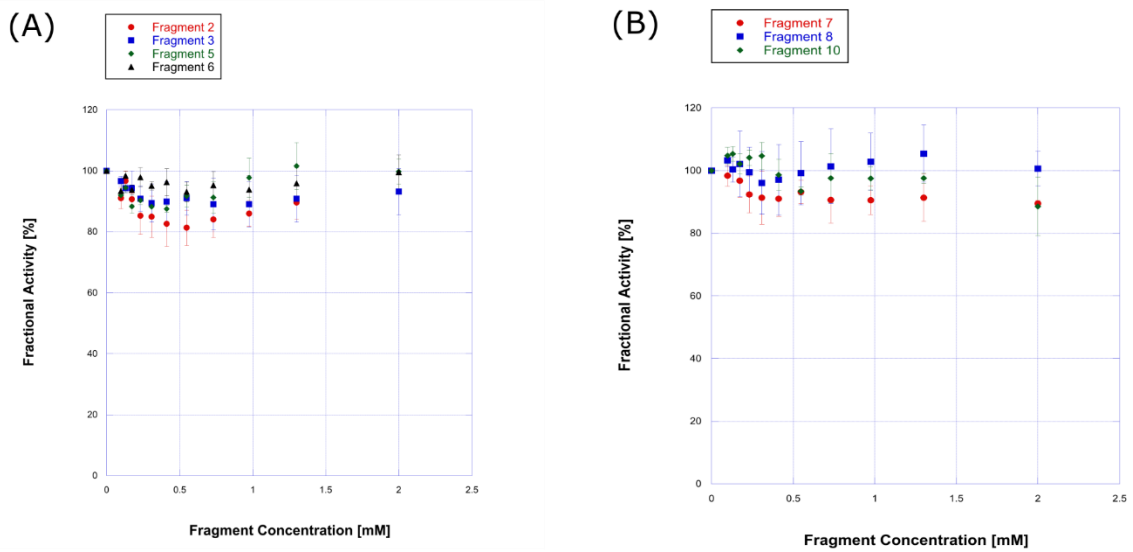


Figure 7-2. Nine fragments were measured using basal ATPase assays in the presence of KifC1$_{307-663}$. Fragments 4 and 9 were identified as hits and have been discussed individually in the results section 4.4.2.6. The other seven fragments displayed here showed no inhibition in the assays.

7.5 Crystallisation screening of KifC1$_{307-663}$

Table 7-3. Crystallisation conditions tested based on the published conditions. A set of pH values and precipitant concentrations were prepared according to the published conditions. The crystallisation plates were set up at both 4 °C and 18 °C. Two drops were set up by mixing 16 and 28 mg/ml KifC1 with reservoir solutions at a 1:1 ratio in each well (with 1 mM MgATP added).

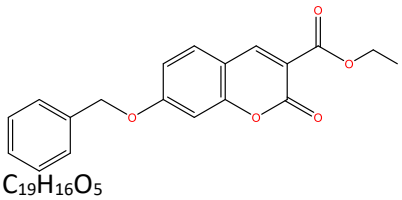
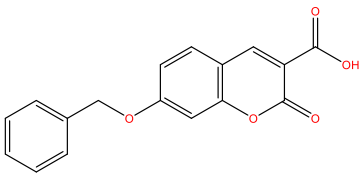
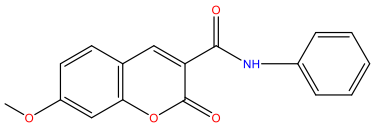
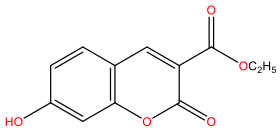
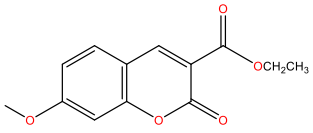
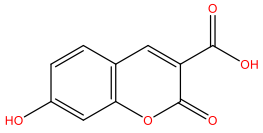
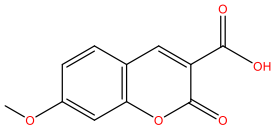
Number	***Buffer***	***Precipitant***	***Additive***
1	0.1 M Bis-Tris Propane, pH 6.6	2.6 M NaCl	5% Glycerol
2	0.1 M Bis-Tris Propane, pH 6.9	2.6 M NaCl	5% Glycerol
3	0.1 M Bis-Tris Propane, pH 7.2	2.6 M NaCl	5% Glycerol
4	0.1 M Bis-Tris Propane, pH 7.5	2.6 M NaCl	5% Glycerol
5	0.1 M Bis-Tris Propane, pH 7.8	2.6 M NaCl	5% Glycerol
6	0.1 M Bis-Tris Propane, pH 8.0	2.6 M NaCl	5% Glycerol
7	0.1 M Bis-Tris Propane, pH 6.6	2.7 M NaCl	5% Glycerol
8	0.1 M Bis-Tris Propane, pH 6.9	2.7 M NaCl	5% Glycerol
9	0.1 M Bis-Tris Propane, pH 7.2	2.7 M NaCl	5% Glycerol
10	0.1 M Bis-Tris Propane, pH 7.5	2.7 M NaCl	5% Glycerol
11	0.1 M Bis-Tris Propane, pH 7.8	2.7 M NaCl	5% Glycerol
12	0.1 M Bis-Tris Propane, pH 8.0	2.7 M NaCl	5% Glycerol
13	0.1 M Bis-Tris Propane, pH 6.6	2.8 M NaCl	5% Glycerol
14	0.1 M Bis-Tris Propane, pH 6.9	2.8 M NaCl	5% Glycerol
15	0.1 M Bis-Tris Propane, pH 7.2	2.8 M NaCl	5% Glycerol
16	0.1 M Bis-Tris Propane, pH 7.5	2.8 M NaCl	5% Glycerol
17	0.1 M Bis-Tris Propane, pH 7.8	2.8 M NaCl	5% Glycerol
18	0.1 M Bis-Tris Propane, pH 8.0	2.8 M NaCl	5% Glycerol
19	0.1 M Bis-Tris Propane, pH 6.6	2.9 M NaCl	5% Glycerol
20	0.1 M Bis-Tris Propane, pH 6.9	2.9 M NaCl	5% Glycerol
21	0.1 M Bis-Tris Propane, pH 7.2	2.9 M NaCl	5% Glycerol
22	0.1 M Bis-Tris Propane, pH 7.5	2.9 M NaCl	5% Glycerol

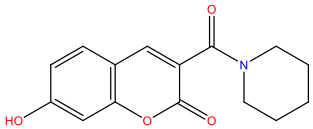
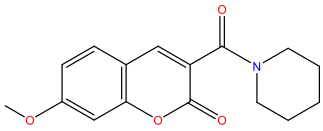
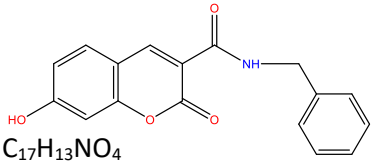
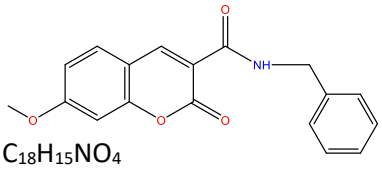
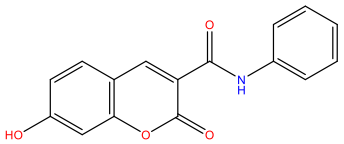
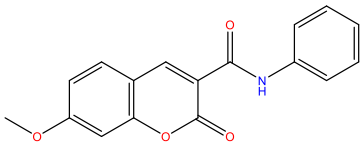
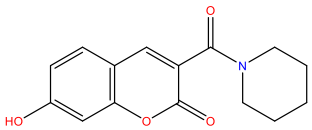
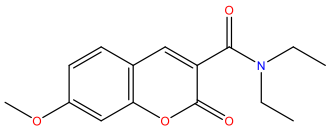
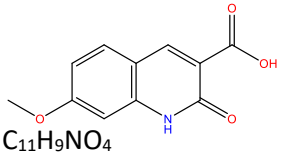
23	0.1 M Bis-Tris Propane, pH 7.8	2.9 M NaCl	5% Glycerol
24	0.1 M Bis-Tris Propane, pH 8.0	2.9 M NaCl	5% Glycerol
25	0.1 M Bis-Tris Propane, pH 6.6	3.0 M NaCl	5% Glycerol
26	0.1 M Bis-Tris Propane, pH 6.9	3.0 M NaCl	5% Glycerol
27	0.1 M Bis-Tris Propane, pH 7.2	3.0 M NaCl	5% Glycerol
28	0.1 M Bis-Tris Propane, pH 7.5	3.0 M NaCl	5% Glycerol
29	0.1 M Bis-Tris Propane, pH 7.8	3.0 M NaCl	5% Glycerol
30	0.1 M Bis-Tris Propane, pH 8.0	3.0 M NaCl	5% Glycerol
31	0.1 M Bis-Tris Propane, pH 6.6	3.1 M NaCl	5% Glycerol
32	0.1 M Bis-Tris Propane, pH 6.9	3.1 M NaCl	5% Glycerol
33	0.1 M Bis-Tris Propane, pH 7.2	3.1 M NaCl	5% Glycerol
34	0.1 M Bis-Tris Propane, pH 7.5	3.1 M NaCl	5% Glycerol
35	0.1 M Bis-Tris Propane, pH 7.8	3.1 M NaCl	5% Glycerol
36	0.1 M Bis-Tris Propane, pH 8.0	3.1 M NaCl	5% Glycerol
37	0.1 M Bis-Tris Propane, pH 6.6	3.2 M NaCl	5% Glycerol
38	0.1 M Bis-Tris Propane, pH 6.9	3.2 M NaCl	5% Glycerol
39	0.1 M Bis-Tris Propane, pH 7.2	3.2 M NaCl	5% Glycerol
40	0.1 M Bis-Tris Propane, pH 7.5	3.2 M NaCl	5% Glycerol
41	0.1 M Bis-Tris Propane, pH 7.8	3.2 M NaCl	5% Glycerol
42	0.1 M Bis-Tris Propane, pH 8.0	3.2 M NaCl	5% Glycerol
43	0.1 M Bis-Tris Propane, pH 6.6	3.3 M NaCl	5% Glycerol
44	0.1 M Bis-Tris Propane, pH 6.9	3.3 M NaCl	5% Glycerol
45	0.1 M Bis-Tris Propane, pH 7.2	3.3 M NaCl	5% Glycerol
46	0.1 M Bis-Tris Propane, pH 7.5	3.3 M NaCl	5% Glycerol
47	0.1 M Bis-Tris Propane, pH 7.8	3.3 M NaCl	5% Glycerol
48	0.1 M Bis-Tris Propane, pH 8.0	3.3 M NaCl	5% Glycerol
49	0.1 M Bis-Tris Propane, pH 6.6	3.4 M NaCl	5% Glycerol
50	0.1 M Bis-Tris Propane, pH 6.9	3.4 M NaCl	5% Glycerol
51	0.1 M Bis-Tris Propane, pH 7.2	3.4 M NaCl	5% Glycerol
52	0.1 M Bis-Tris Propane, pH 7.5	3.4 M NaCl	5% Glycerol
53	0.1 M Bis-Tris Propane, pH 7.8	3.4 M NaCl	5% Glycerol

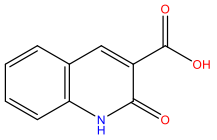
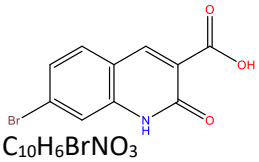
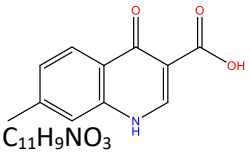
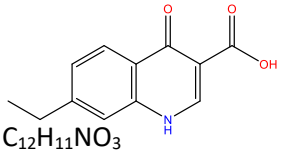
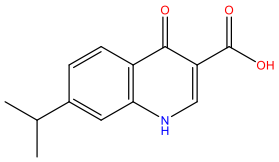
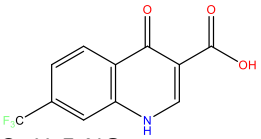
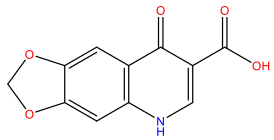
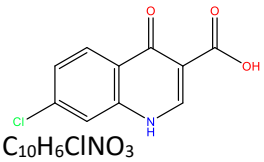
54	0.1 M Bis-Tris Propane, pH 8.0	3.4 M NaCl	5% Glycerol
55	0.1 M Bis-Tris Propane, pH 6.6	3.5 M NaCl	5% Glycerol
56	0.1 M Bis-Tris Propane, pH 6.9	3.5 M NaCl	5% Glycerol
57	0.1 M Bis-Tris Propane, pH 7.2	3.5 M NaCl	5% Glycerol
58	0.1 M Bis-Tris Propane, pH 7.5	3.5 M NaCl	5% Glycerol
59	0.1 M Bis-Tris Propane, pH 7.8	3.5 M NaCl	5% Glycerol
60	0.1 M Bis-Tris Propane, pH 8.0	3.5 M NaCl	5% Glycerol
61	0.1 M Bis-Tris Propane, pH 6.6	3.6 M NaCl	5% Glycerol
62	0.1 M Bis-Tris Propane, pH 6.9	3.6 M NaCl	5% Glycerol
63	0.1 M Bis-Tris Propane, pH 7.2	3.6 M NaCl	5% Glycerol
64	0.1 M Bis-Tris Propane, pH 7.5	3.6 M NaCl	5% Glycerol
65	0.1 M Bis-Tris Propane, pH 7.8	3.6 M NaCl	5% Glycerol
66	0.1 M Bis-Tris Propane, pH 8.0	3.6 M NaCl	5% Glycerol

7.6 MPP1 analogues synthesised by Mr. Helal.

Table 7-4. MPP1 analogues synthesised by Mr. Helal using chromenone-3-carboxylic acids and quinolone-3-carboxylic acid core structures.

Compound name	Structure	MW (Da)	Number of HBA/ HBD	cLogP	PSA ($Å^2$)
A13	 $C_{19}H_{16}O_5$	324.3	3/0	3.45	61.8
A14	 $C_{17}H_{12}O_5$	296.3	3/1	2.85	72.8
A17	 $C_{17}H_{13}NO_4$	295.3	3/1	2.37	64.6
A40	 $C_{12}H_{10}O_5$	234.2	3/1	1.46	72.8
A41	 $C_{13}H_{12}O_5$	248.2	3/0	1.72	61.8
A42	 $C_{10}H_6O_5$	206.2	3/2	0.86	83.8
A43	 $C_{11}H_8O_5$	220.2	3/1	1.12	72.8

					
A44	$C_{15}H_{15}NO_4$	273.3	3/1	1.41	66.8
					
A45	$C_{16}H_{17}NO_4$	287.3	3/0	1.67	55.8
					
A46	$C_{17}H_{13}NO_4$	295.3	3/2	2.17	75.6
					
A47	$C_{18}H_{15}NO_4$	309.3	3/1	2.44	64.6
					
A48	$C_{16}H_{11}NO_4$	281.3	3/2	2.10	75.6
					
A49	$C_{17}H_{13}NO_4$	295.3	3/1	2.37	64.6
					
A50	$C_{15}H_{15}NO_4$	273.3	3/1	1.41	66.8
					
A51	$C_{15}H_{17}NO_4$	275.3	3/0	1.62	55.8
					
A62	$C_{11}H_9NO_4$	219.2	3/2	0.44	75.6

A67	 $C_{10}H_7NO_3$	189.2	2/2	0.57	66.4
A84	 $C_{10}H_6BrNO_3$	268.1	2/2	1.40	66.4
A87	 $C_{11}H_9NO_3$	203.2	3/2	0.90	66.4
A89	 $C_{12}H_{11}NO_3$	217.2	3/2	1.32	66.4
A91	 $C_{13}H_{13}NO_3$	231.3	3/2	1.65	66.4
A94	 $C_{11}H_6F_3NO_3$	257.2	6/2	1.34	66.4
A99	 $C_{11}H_7NO_5$	233.2	5/2	0.20	84.9
A100	 $C_{10}H_6ClNO_3$	223.6	3/2	0.98	66.4

7.7 Basal ATPase activity characterisation of the MPP1 analogues synthesised by Mr. Helal

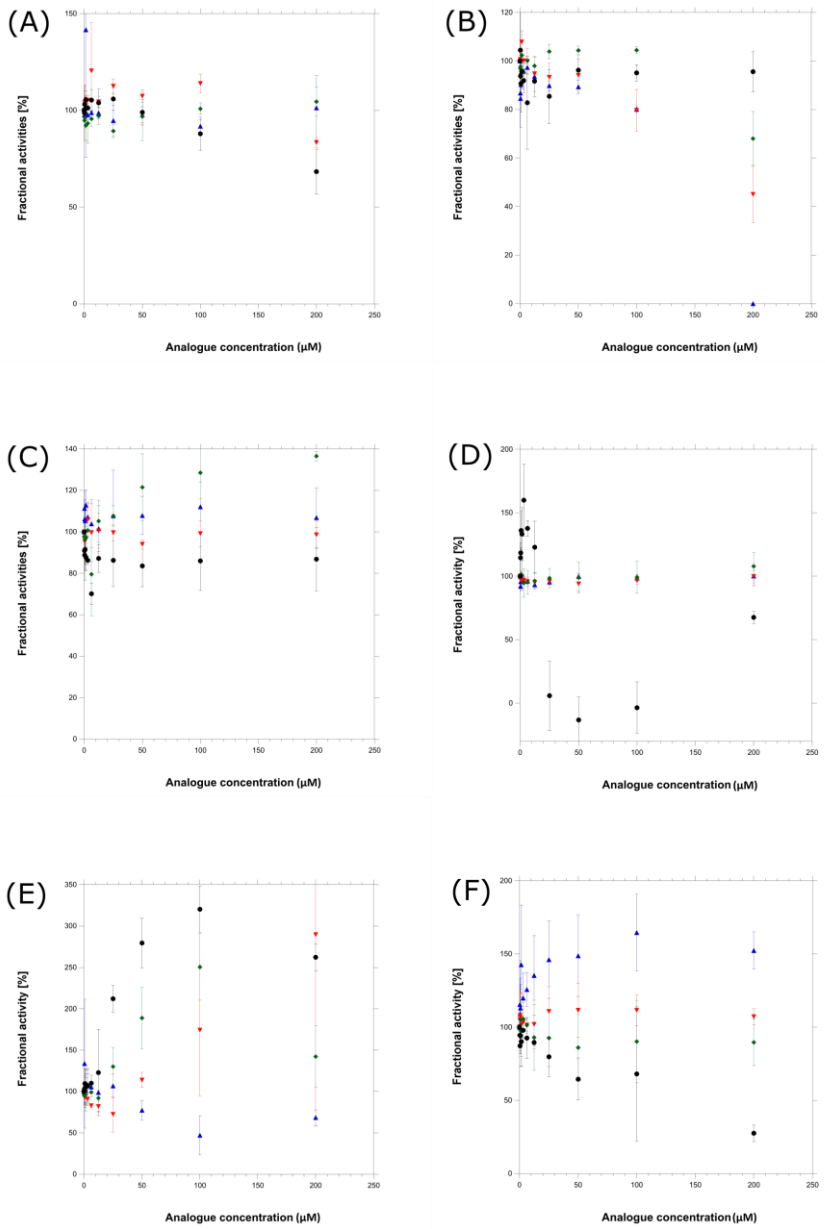


*Figure 7-3. Basal ATPase activity characterisation of the MPP1 analogues synthesised by Mr. Helal. (A) Analogues **A13** (blue), **A14** (red), **A17** (green), and **A40** (black). (B) Analogues **A41** (blue), **A42** (red), **A43** (green), and **A44** (black). (C) Analogues **A45** (blue), **A46** (red), **A47** (green), and **A48** (black). (D) Analogues **A49** (blue), **A50** (red), **A51** (green), and **A62** (black). (E) Analogues **A67** (blue), **A84** (red), **A87** (green), and **A89** (black). (F) Analogues **A91** (blue), **A94** (red), **A99** (green), and **A100** (black). All the analogues showed either no inhibition or atypical binding curves in the basal ATPase assay.*

8. Bibliography

1. Ferlay, J., et al., *Global cancer observatory: cancer today.* 2018: p. 1-6.
2. Organization, W.H., *WHO report on cancer: setting priorities, investing wisely and providing care for all.* 2020.
3. Hirokawa, N. and Y. Tanaka, *Kinesin superfamily proteins (KIFs): Various functions and their relevance for important phenomena in life and diseases.* Experimental cell research, 2015. **334**(1): p. 16-25.
4. Kolomeisky, A.B. and M.E.J.A.R.P.C. Fisher, *Molecular motors: a theorist's perspective.* 2007. **58**: p. 675-695.
5. Vale, R.D., T.S. Reese, and M.P.J.C. Sheetz, *Identification of a novel force-generating protein, kinesin, involved in microtubule-based motility.* 1985. **42**(1): p. 39-50.
6. Bloom, G.S., et al., *Native structure and physical properties of bovine brain kinesin and identification of the ATP-binding subunit polypeptide.* 1988. **27**(9): p. 3409-3416.
7. Kolomeisky, A.B., *Motor proteins and molecular motors: how to operate machines at the nanoscale.* Journal of Physics: Condensed Matter, 2013. **25**(46): p. 463101.
8. Rath, O. and F. Kozielski, *Kinesins and cancer.* Nature reviews cancer, 2012. **12**(8): p. 527.
9. Kalantari, S. and I.J.J.o.m.g. Filges, *'Kinesinopathies': emerging role of the kinesin family member genes in birth defects.* 2020. **57**(12): p. 797-807.
10. Hirokawa, N., et al., *Kinesin superfamily motor proteins and intracellular transport.* Nature reviews Molecular cell biology, 2009. **10**(10): p. 682.
11. Hummel, J.J. and C.C.J.J.o.C.B. Hoogenraad, *Specific KIF1A–adaptor interactions control selective cargo recognition.* 2021. **220**(10): p. e202105011.
12. Nangaku, M., et al., *KIF1B, a novel microtubule plus end-directed monomeric motor protein for transport of mitochondria.* 1994. **79**(7): p. 1209-1220.
13. Steinmetz, M.O. and A.J.T.i.b.s. Akhmanova, *Capturing protein tails by CAP-Gly domains.* 2008. **33**(11): p. 535-545.
14. Pyrpassopoulos, S., H. Shuman, and E.M.J.M.b.o.t.c. Ostap, *Adhesion force and attachment lifetime of the KIF16B-PX domain interaction with lipid membranes.* 2017. **28**(23): p. 3315-3322.
15. Lee, K.-H., et al., *KIF21A-mediated axonal transport and selective endocytosis underlie the polarized targeting of NCKX2.* 2012. **32**(12): p. 4102-4117.
16. Kull, F.J., et al., *Crystal structure of the kinesin motor domain reveals a structural similarity to myosin.* Nature, 1996. **380**(6574): p. 550.
17. Endow, S.A., F.J. Kull, and H. Liu, *Kinesins at a glance.* J Cell Sci, 2010. **123**(20): p. 3420-3424.
18. Kull, F.J., R.D. Vale, and R.J. Fletterick, *The case for a common ancestor: kinesin and myosin motor proteins and G proteins.* Journal of Muscle Research & Cell Motility, 1998. **19**(8): p. 877-886.
19. Hirose, K., et al., *Large conformational changes in a kinesin motor catalyzed by interaction with microtubules.* Molecular cell, 2006. **23**(6): p. 913-923.
20. Woehlke, G., et al., *Microtubule interaction site of the kinesin motor.* Cell, 1997. **90**(2): p. 207-216.
21. Turner, J., et al., *Crystal structure of the mitotic spindle kinesin Eg5 reveals a novel conformation of the neck-linker.* Journal of Biological Chemistry, 2001. **276**(27): p. 25496-25502.
22. De Cuevas, M., T. Tao, and L. Goldstein, *Evidence that the stalk of Drosophila kinesin heavy chain is an alpha-helical coiled coil.* The Journal of cell biology, 1992. **116**(4): p. 957-965.
23. Grummt, M., et al., *Importance of a flexible hinge near the motor domain in kinesin-driven motility.* The EMBO journal, 1998. **17**(19): p. 5536-5542.

24. Verhey, K.J. and J.W. Hammond, *Traffic control: regulation of kinesin motors.* Nature reviews Molecular cell biology, 2009. **10**(11): p. 765.
25. Verhey, K.J., et al., *Cargo of kinesin identified as JIP scaffolding proteins and associated signaling molecules.* 2001. **152**(5): p. 959-970.
26. Hackney, D.D. and M.F. Stock, *Kinesin's IAK tail domain inhibits initial microtubule-stimulated ADP release.* Nature Cell Biology, 2000. **2**(5): p. 257.
27. Talapatra, S.K., B. Harker, and J.P.J.E. Welburn, *The C-terminal region of the motor protein MCAK controls its structure and activity through a conformational switch.* 2015. **4**: p. e06421.
28. Wendt, T., et al., *A structural analysis of the interaction between ncd tail and tubulin protofilaments.* Journal of molecular biology, 2003. **333**(3): p. 541-552.
29. Hirokawa, N., *Organelle transport along microtubules—the role of KIFs.* Trends in cell biology, 1996. **6**(4): p. 135-141.
30. Cole, D.G., *Intraflagellar transport: keeping the motors coordinated.* Current biology, 2005. **15**(19): p. R798-R801.
31. Hao, L., et al., *Analysis of intraflagellar transport in C. elegans sensory cilia*, in *Methods in cell biology*. 2009, Elsevier. p. 235-266.
32. !!! INVALID CITATION !!! {}.
33. Zhu, C. and W.J.P.o.t.N.A.o.S. Jiang, *Cell cycle-dependent translocation of PRC1 on the spindle by Kif4 is essential for midzone formation and cytokinesis.* 2005. **102**(2): p. 343-348.
34. Garcia-Saez, I. and D.A.J.B.P. Skoufias, *Eg5 targeting agents: From new anti-mitotic based inhibitor discovery to cancer therapy and resistance.* 2021. **184**: p. 114364.
35. Milic, B., et al., *KIF15 nanomechanics and kinesin inhibitors, with implications for cancer chemotherapeutics.* Proceedings of the National Academy of Sciences, 2018. **115**(20): p. E4613-E4622.
36. Chandra, R., et al., *Structural and functional domains of the Drosophila ncd microtubule motor protein.* Journal of Biological Chemistry, 1993. **268**(12): p. 9005-9013.
37. Pannu, V., P.C. Rida, and R. Aneja, *The human Kinesin-14 motor KifC1/HSET is an attractive anti-cancer drug target*, in *Kinesins and Cancer*. 2015, Springer. p. 101-116.
38. Ulaganathan, V., et al., *Structural insights into a unique inhibitor binding pocket in kinesin spindle protein.* 2013. **135**(6): p. 2263-2272.
39. Wood, K.W., et al., *Antitumor activity of an allosteric inhibitor of centromere-associated protein-E.* 2010. **107**(13): p. 5839-5844.
40. Park, H.-W., et al., *Structural basis of small molecule ATPase inhibition of a human mitotic kinesin motor protein.* Scientific reports, 2017. **7**(1): p. 15121.
41. Myers, S.M. and I.J.F.m.c. Collins, *Recent findings and future directions for interpolar mitotic kinesin inhibitors in cancer therapy.* 2016. **8**(4): p. 463-489.
42. Norris, S.R., et al., *Microtubule minus-end aster organization is driven by processive HSET-tubulin clusters.* bioRxiv, 2018: p. 278481.
43. Zhang, Y. and A.O. Sperry, *Comparative analysis of two C-terminal kinesin motor proteins: KIFC1 and KIFC5A.* Cell motility and the cytoskeleton, 2004. **58**(4): p. 213-230.
44. Zhi, E., et al., *Decreased expression of KIFC1 in human testes with globozoospermic defects.* Genes, 2016. **7**(10): p. 75.
45. Krämer, A., K. Neben, and A.D. Ho, *Centrosome aberrations in hematological malignancies.* Cell biology international, 2005. **29**(5): p. 375-383.
46. Nigg, E.A., *Centrosome aberrations: cause or consequence of cancer progression?* Nature Reviews Cancer, 2002. **2**(11): p. 815.
47. Weaver, B.A., et al., *Aneuploidy acts both oncogenically and as a tumor suppressor.* Cancer cell, 2007. **11**(1): p. 25-36.
48. Godinho, S.A., M. Kwon, and D. Pellman, *Centrosomes and cancer: how cancer cells divide with too many centrosomes.* Cancer and Metastasis Reviews, 2009. **28**(1-2): p. 85-98.

49. Kwon, M., et al., *Mechanisms to suppress multipolar divisions in cancer cells with extra centrosomes.* Genes & development, 2008. **22**(16): p. 000-000.
50. Wu, J., et al., *Discovery and mechanistic study of a small molecule inhibitor for motor protein KIFC1.* ACS chemical biology, 2013. **8**(10): p. 2201-2208.
51. Watts, C.A., et al., *Design, synthesis, and biological evaluation of an allosteric inhibitor of HSET that targets cancer cells with supernumerary centrosomes.* Chemistry & biology, 2013. **20**(11): p. 1399-1410.
52. Guan, R., et al., *Crystal structure of Zen4 in the apo state reveals a missing conformation of kinesin.* 2017. **8**(1): p. 1-9.
53. Atherton, J., et al., *The divergent mitotic kinesin MKLP2 exhibits atypical structure and mechanochemistry.* 2017. **6**: p. e27793.
54. Abaza, A., et al., *M-phase phosphoprotein 1 is a human plus-end directed kinesin-related protien required for cytokinesis.* 2003.
55. Hill, E., M. Clarke, and F.A.J.T.E.j. Barr, *The Rab6-binding kinesin, Rab6-KIFL, is required for cytokinesis.* 2000. **19**(21): p. 5711-5719.
56. Hardin, J., et al., *Zygotic loss of ZEN-4/MKLP1 results in disruption of epidermal morphogenesis in the C. elegans embryo.* 2008. **237**(3): p. 830-836.
57. Kamimoto, T., et al., *Identification of a novel kinesin-related protein, KRMP1, as a target for mitotic peptidyl-prolyl isomerase Pin1.* 2001. **276**(40): p. 37520-37528.
58. Talapatra, S.K., *Mechanistic investigation of small molecule inhibitors of kinesin-5 and kinesin-6 family members in cancer drug development*. 2012, University of Glasgow.
59. Talapatra, S.K., et al., *Depsidones from lichens as natural product inhibitors of M-phase phosphoprotein 1, a human kinesin required for cytokinesis.* 2016. **79**(6): p. 1576-1585.
60. Kanehira, M., et al., *Oncogenic role of MPHOSPH1, a cancer-testis antigen specific to human bladder cancer.* 2007. **67**(7): p. 3276-3285.
61. Zhu, C., et al., *Functional analysis of human microtubule-based motor proteins, the kinesins and dyneins, in mitosis/cytokinesis using RNA interference.* 2005. **16**(7): p. 3187-3199.
62. Urvoas, A., et al., *Design, production and molecular structure of a new family of artificial alpha-helicoidal repeat proteins (αRep) based on thermostable HEAT-like repeats.* 2010. **404**(2): p. 307-327.
63. Chevrel, A., et al., *Alpha repeat proteins (αRep) as expression and crystallization helpers.* 2018. **201**(2): p. 88-99.
64. Bray, F., et al., *Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries.* CA: a cancer journal for clinicians, 2018. **68**(6): p. 394-424.
65. Mármol, I., et al., *Colorectal carcinoma: a general overview and future perspectives in colorectal cancer.* 2017. **18**(1): p. 197.
66. Riihimäki, M., et al., *Patterns of metastasis in colon and rectal cancer.* Scientific reports, 2016. **6**(1): p. 1-9.
67. Noguchi, T., G. Ritter, and H. Nishikawa, *Antibody-based therapy in colorectal cancer.* Immunotherapy, 2013. **5**(5): p. 533-545.
68. Van Loon, K. and A.P. Venook, *Adjuvant treatment of colon cancer: what is next?* Current opinion in oncology, 2011. **23**(4): p. 403-409.
69. Van Cutsem, E., et al., *Metastatic colorectal cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up.* 2014. **25**: p. iii1-iii9.
70. Infante, J.R., et al., *Safety, pharmacokinetics and pharmacodynamics of the anti-A33 fully-human monoclonal antibody, KRN330, in patients with advanced colorectal cancer.* 2013. **49**(6): p. 1169-1175.
71. Welt, S., et al., *Phase I study of anticolon cancer humanized antibody A33.* Clinical Cancer Research, 2003. **9**(4): p. 1338-1346.

72. GarinChesa, P., et al., *Organ-specific expression of the colon cancer antigen A33, a cell surface target for antibody-based therapy.* International journal of oncology, 1996. **9**(3): p. 465-471.
73. King, D., et al., *Preparation and preclinical evaluation of humanised A33 immunoconjugates for radioimmunotherapy.* British journal of cancer, 1995. **72**(6): p. 1364-1372.
74. King, D.J., J.R. Adair, and R.J. Owens, *Humanized antibodies directed against A33 antigen*. 2001, Google Patents.
75. Chakroun, N., et al., *Mapping the aggregation kinetics of a therapeutic antibody fragment.* Molecular pharmaceutics, 2016. **13**(2): p. 307-319.
76. Zhang, C., et al., *Computational design to reduce conformational flexibility and aggregation rates of an antibody fab fragment.* Molecular pharmaceutics, 2018. **15**(8): p. 3079-3092.
77. Zhang, C., et al., *Comparison of the pH-and thermally-induced fluctuations of a therapeutic antibody Fab fragment by molecular dynamics simulation.* 2021. **19**: p. 2726-2741.
78. Foulkes, W.D., I.E. Smith, and J.S.J.N.E.j.o.m. Reis-Filho, *Triple-negative breast cancer.* 2010. **363**(20): p. 1938-1948.
79. Urru, S.A.M., et al., *Clinical and pathological factors influencing survival in a large cohort of triple-negative breast cancer patients.* 2018. **18**(1): p. 1-11.
80. Dent, R., et al., *Triple-negative breast cancer: clinical features and patterns of recurrence.* 2007. **13**(15): p. 4429-4434.
81. Dawood, S., et al., *Impact of body mass index on survival outcome among women with early stage triple-negative breast cancer.* 2012. **12**(5): p. 364-372.
82. Lyons, T.G.J.C.t.o.i.o., *Targeted therapies for triple-negative breast cancer.* 2019. **20**(11): p. 1-13.
83. Yao, H., et al., *Triple-negative breast cancer: is there a treatment on the horizon?* 2017. **8**(1): p. 1913.
84. Traina, T.A., et al., *Results from a phase 2 study of enzalutamide (ENZA), an androgen receptor (AR) inhibitor, in advanced AR+ triple-negative breast cancer (TNBC)*. 2015, American Society of Clinical Oncology.
85. Katz, H. and M.J.M.o. Alsharedi, *Immunotherapy in triple-negative breast cancer.* 2018. **35**(1): p. 1-9.
86. Herr, I. and K.-M.J.B. Debatin, The Journal of the American Society of Hematology, *Cellular stress response and apoptosis in cancer therapy.* 2001. **98**(9): p. 2603-2614.
87. Li, D., et al., *A gain-of-function mutant p53–HSF1 feed forward circuit governs adaptation of cancer cells to proteotoxic stress.* 2014. **5**(4): p. e1194-e1194.
88. Whitesell, L. and S.J.E.o.o.t.t. Lindquist, *Inhibiting the transcription factor HSF1 as an anticancer strategy.* 2009. **13**(4): p. 469-478.
89. Santagata, S., et al., *High levels of nuclear heat-shock factor 1 (HSF1) are associated with poor prognosis in breast cancer.* 2011. **108**(45): p. 18378-18383.
90. Gabai, V.L., et al., *Heat shock transcription factor Hsf1 is involved in tumor progression via regulation of hypoxia-inducible factor 1 and RNA-binding protein HuR.* 2012. **32**(5): p. 929-940.
91. Petrocca, F., et al., *A genome-wide siRNA screen identifies proteasome addiction as a vulnerability of basal-like triple-negative breast cancer cells.* 2013. **24**(2): p. 182-196.
92. Luo, J., N.L. Solimini, and S.J.J.C. Elledge, *Principles of cancer therapy: oncogene and non-oncogene addiction.* 2009. **136**(5): p. 823-837.
93. Yan, H., et al., *Low expression of DYRK2 (Dual Specificity Tyrosine Phosphorylation Regulated Kinase 2) correlates with poor prognosis in colorectal cancer.* 2016. **11**(8): p. e0159954.
94. Yoshida, K.J.B.p., *Role for DYRK family kinases on regulation of apoptosis.* 2008. **76**(11): p. 1389-1394.
95. Yamashita, S.-I., et al., *DYRK2 expression may be a predictive marker for chemotherapy in non-small cell lung cancer.* 2009. **29**(7): p. 2753-2757.

96. Yamashita, S.-i., et al., *Expression of dual-specificity tyrosine-(Y)-phosphorylation–regulated kinase 2 (DYRK2) can be a favorable prognostic marker in pulmonary adenocarcinoma.* 2009. **138**(6): p. 1303-1308.
97. Miller, C.T., et al., *Gene amplification in esophageal adenocarcinomas and Barrett's with high-grade dysplasia.* 2003. **9**(13): p. 4819-4825.
98. Koon, N., et al., *Molecular targets for tumour progression in gastrointestinal stromal tumours.* 2004. **53**(2): p. 235-240.
99. Banerjee, S., et al., *Ancient drug curcumin impedes 26S proteasome activity by direct inhibition of dual-specificity tyrosine-regulated kinase 2.* 2018. **115**(32): p. 8155-8160.
100. Tahtouh, T., et al., *Selectivity, cocrystal structures, and neuroprotective properties of leucettines, a family of protein kinase inhibitors derived from the marine sponge alkaloid leucettamine B.* 2012. **55**(21): p. 9312-9330.
101. Chaikuad, A., et al., *An unusual binding model of the methyl 9-anilinothiazolo [5, 4-f] quinazoline-2-carbimidates (EHT 1610 and EHT 5372) confers high selectivity for dual-specificity tyrosine phosphorylation-regulated kinases.* 2016. **59**(22): p. 10315-10321.
102. Murer, P., L. Plüss, and D. Neri. *A novel human monoclonal antibody specific to the A33 glycoprotein recognizes colorectal cancer and inhibits metastasis*. in *MAbs*. 2020. Taylor & Francis.
103. Hackney, D.D. and W. Jiang, *Assays for Kinesin Microtubule-Stimulated AT Pase Activity*, in *Kinesin Protocols*. 2001, Springer. p. 65-71.
104. Vernos, I., *Kinesin Protocols*. Vol. 164. 2001: Springer.
105. Copeland, R.A., *Evaluation of enzyme inhibitors in drug discovery: a guide for medicinal chemists and pharmacologists*. 2013: John Wiley & Sons.
106. Battye, T.G.G., et al., *iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM.* Acta Crystallographica Section D: Biological Crystallography, 2011. **67**(4): p. 271-281.
107. Evans, P., *Scaling and assessment of data quality.* Acta Crystallographica Section D: Biological Crystallography, 2006. **62**(1): p. 72-82.
108. Collaborative, C.P., *The CCP4 suite: programs for protein crystallography.* Acta crystallographica. Section D, Biological crystallography, 1994. **50**(Pt 5): p. 760.
109. Liebschner, D., et al., *Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix.* Acta Crystallographica Section D: Structural Biology, 2019. **75**(10): p. 861-877.
110. Emsley, P., et al., *Features and development of Coot.* Acta Crystallographica Section D: Biological Crystallography, 2010. **66**(4): p. 486-501.
111. Shi, L., et al., *Structural analysis of the Hanks-type Protein Kinase YabT from Bacillus subtilis provides new insights in its DNA-Dependent Activation.* 2019. **9**: p. 3014.
112. Manalastas-Cantos, K., et al., *ATSAS 3.0: expanded functionality and new tools for small-angle scattering data analysis.* 2021. **54**(1).
113. Konarev, P.V., et al., *PRIMUS: a Windows PC-based system for small-angle scattering data analysis.* 2003. **36**(5): p. 1277-1282.
114. Franke, D. and D.I.J.J.o.a.c. Svergun, *DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering.* 2009. **42**(2): p. 342-346.
115. Svergun, D.I.J.B.j., *Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing.* 1999. **76**(6): p. 2879-2886.
116. Panjkovich, A. and D.I.J.B. Svergun, *SASpy: a PyMOL plugin for manipulation and refinement of hybrid models against small angle X-ray scattering data.* 2016. **32**(13): p. 2062-2064.
117. Apgar, J.R., et al. *Beyond CDR-grafting: structure-guided humanization of framework and CDR regions of an anti-myostatin antibody*. in *MAbs*. 2016. Taylor & Francis.

118. Vajdos, F.F., et al., *Comprehensive functional maps of the antigen-binding site of an anti-ErbB2 antibody obtained with shotgun scanning mutagenesis.* Journal of molecular biology, 2002. **320**(2): p. 415-428.
119. Soundararajan, M., et al., *Structures of Down syndrome kinases, DYRKs, reveal mechanisms of kinase activation and substrate recognition.* 2013. **21**(6): p. 986-996.
120. Kabsch, W.J.A.C.S.D.B.C., *Xds.* 2010. **66**(2): p. 125-132.
121. Labrière, C., et al., *New MKLP-2 inhibitors in the paprotrain series: Design, synthesis and biological evaluations.* 2016. **24**(4): p. 721-734.
122. Kozakov, D., et al., *The ClusPro web server for protein–protein docking.* 2017. **12**(2): p. 255-278.
123. Abaza, A., et al., *M phase phosphoprotein 1 is a human plus-end-directed kinesin-related protein required for cytokinesis.* 2003. **278**(30): p. 27844-27852.
124. Liu, X., et al., *Inhibition of kinesin family member 20B sensitizes hepatocellular carcinoma cell to microtubule‐targeting agents by blocking cytokinesis.* 2018. **109**(11): p. 3450-3460.
125. Liu, X., et al., *MPHOSPH1: a potential therapeutic target for hepatocellular carcinoma.* 2014. **74**(22): p. 6623-6634.
126. Guellouz, A., et al., *Selection of specific protein binders for pre-defined targets from an optimized library of artificial helicoidal repeat proteins (alphaRep).* 2013. **8**(8): p. e71512.
127. Ferrandez, Y., et al., *Amphipol-mediated screening of molecular orthoses specific for membrane protein targets.* 2014. **247**(9): p. 925-940.
128. Tiouajni, M., et al., *Structural and functional analysis of the fibronectin‐binding protein FNE from Streptococcus equi spp. equi.* 2014. **281**(24): p. 5513-5531.
129. Figueroa, M., et al., *The unexpected structure of the designed protein Octarellin V. 1 forms a challenge for protein structure prediction tools.* 2016. **195**(1): p. 19-30.
130. Matsumoto-Taniura, N., et al., *Identification of novel M phase phosphoproteins by expression cloning.* 1996. **7**(9): p. 1455-1469.
131. Georges, A., et al., *USP7 regulates cytokinesis through FBXO38 and KIF20B.* 2019. **9**(1): p. 1-16.
132. Lin, W.F., et al., *Pseudopod‐associated protein KIF20B promotes Gli1‐induced epithelial‐mesenchymal transition modulated by pseudopodial actin dynamic in human colorectal cancer.* 2018. **57**(7): p. 911-925.
133. Li, T.-F., et al., *Overexpression of kinesin superfamily members as prognostic biomarkers of breast cancer.* 2020. **20**(1): p. 1-16.
134. Li, G., et al., *KIF20B promotes the progression of clear cell renal cell carcinoma by stimulating cell proliferation.* 2019. **234**(9): p. 16517-16525.
135. Chen, J., et al., *KIF20B Promotes Cell Proliferation and May Be a Potential Therapeutic Target in Pancreatic Cancer.* 2021. **2021**.
136. Wang, Z., et al., *Determination of in vitro permeability of drug candidates through a Caco‐2 cell monolayer by liquid chromatography/tandem mass spectrometry.* 2000. **35**(1): p. 71-76.
137. Di, L., et al., *High throughput artificial membrane permeability assay for blood–brain barrier.* 2003. **38**(3): p. 223-232.
138. Stanfield, R.L., et al., *Antibody elbow angles are influenced by their light chain class.* Journal of molecular biology, 2006. **357**(5): p. 1566-1574.
139. Lee, J.U., et al., *Molecular basis for the neutralization of tumor necrosis factor α by certolizumab pegol in the treatment of inflammatory autoimmune diseases.* International journal of molecular sciences, 2017. **18**(1): p. 228.
140. O'Donoghue, J.A., et al., *124I-huA33 antibody uptake is driven by A33 antigen concentration in tissues from colorectal cancer patients imaged by immuno-PET.* Journal of Nuclear Medicine, 2011. **52**(12): p. 1878-1885.
141. Codina Castillo, N., *Stability and aggregation-prone conformations of an antibody fragment antigen-binding (Fab)*. 2019, UCL (University College London).

142. Henderson, R., et al., *Selection of immunoglobulin elbow region mutations impacts interdomain conformational flexibility in HIV-1 broadly neutralizing antibodies.* 2019. **10**(1): p. 1-14.
143. Lord, D.M., et al. *Structure-based engineering to restore high affinity binding of an isoform-selective anti-TGFβ1 antibody*. in *MAbs*. 2018. Taylor & Francis.
144. Auffinger, P., et al., *Halogen bonds in biological molecules.* 2004. **101**(48): p. 16789-16794.
145. Metrangolo, P. and G.J.C.a.E.j. Resnati, *Halogen bonding: a paradigm in supramolecular chemistry.* 2001. **7**(12): p. 2511-2519.
146. Lochhead, P.A., et al., *Activation-loop autophosphorylation is mediated by a novel transitional intermediate form of DYRKs.* 2005. **121**(6): p. 925-936.
147. Fabbro, D., S.W. Cowan‐Jacob, and H.J.B.j.o.p. Moebitz, *Ten things you should know about protein kinases: IUPHAR R eview 14.* 2015. **172**(11): p. 2675-2700.
148. Treiber, D.K., N.P.J.C. Shah, and biology, *Ins and outs of kinase DFG motifs.* 2013. **20**(6): p. 745-746.
149. Spassov, V.Z., L. Yan, and P.K.J.P.S. Flook, *The dominant role of side‐chain backbone interactions in structural realization of amino acid code. ChiRotor: A side‐chain prediction algorithm based on side‐chain backbone interactions.* 2007. **16**(3): p. 494-506.
150. Xiao, Y.-X. and W.-X.J.O. Yang, *KIFC1: a promising chemotherapy target for cancer treatment?* 2016. **7**(30): p. 48656.
151. Zhang, W., et al., *Discovery of a novel inhibitor of kinesin-like protein KIFC1.* 2016. **473**(8): p. 1027-1035.
152. Yang, B., et al., *Discovery of potent KIFC1 inhibitors using a method of integrated high-throughput synthesis and screening.* 2014. **57**(23): p. 9958-9970.
153. Tham, C.L., *Biochemical characterisation of human KifC1 and Eg5, two potential targets for drug development in cancer chemotherapy*. 2019, UCL (University College London).
154. Jones, R., et al., *Phase II study to assess the efficacy, safety and tolerability of the mitotic spindle kinesin inhibitor AZD4877 in patients with recurrent advanced urothelial cancer.* 2013. **31**(4): p. 1001-1007.
155. Holen, K., et al., *A phase I trial of MK-0731, a kinesin spindle protein (KSP) inhibitor, in patients with solid tumors.* 2012. **30**(3): p. 1088-1095.
156. Shah, J.J., et al., *A Phase 1 and 2 study of Filanesib alone and in combination with low‐dose dexamethasone in relapsed/refractory multiple myeloma.* 2017. **123**(23): p. 4617-4630.
157. Schirrmacher, V.J.I.j.o.o., *From chemotherapy to biological therapy: A review of novel concepts to reduce the side effects of systemic cancer treatment.* 2019. **54**(2): p. 407-419.
158. Correa-Sáez, A., et al., *Updating dual-specificity tyrosine-phosphorylation-regulated kinase 2 (DYRK2): Molecular basis, functions and role in diseases.* 2020. **77**: p. 4747-4763.
159. Zhong, L., et al., *Small molecules in targeted cancer therapy: Advances, challenges, and future perspectives.* 2021. **6**(1): p. 1-48.
160. Stromgaard, K., P. Krogsgaard-Larsen, and U. Madsen, *Textbook of drug design and discovery*. 2009: CRC press.
161. Lu, R.-M., et al., *Development of therapeutic antibodies for the treatment of diseases.* 2020. **27**(1): p. 1-30.
162. Kim, J.H. and H.J. Hong, *Humanization by CDR grafting and specificity-determining residue grafting*, in *Antibody Engineering*. 2012, Springer. p. 237-245.
163. Tang, J., et al., *The structure of the humanised A33 Fab C226S variant, an immunotherapy candidate for colorectal cancer.* 2022.
164. Zahavi, D. and L.J.A. Weiner, *Monoclonal antibodies in cancer therapy.* 2020. **9**(3): p. 34.