Of the vulnerability of orphan proteins: The case study of the *Arabidopsis thaliana* p70 ribosomal S6 Kinase 2

Kestutis Lapenas

A thesis submitted for the degree of
Doctor of Philosophy in Biochemistry

School of Life Sciences

University of Essex

October 2023

Impact of COVID-19

As a result of COVID-19 lockdown experimental work was interrupted abruptly, and, unfortunately, this had major long-term repercussions, which could not be fully mitigated, on the progress and delivery of some parts of the project. Due to Health & Safety, it was not possible to carry out lab work between 17th March 2020 and 20th August 2020. After the 20th of August we worked with a 20% max capacity of the School of Life Sciences research building for 1 year.

Abstract

Plants have developed a diverse and multi-layered range of mechanisms to perceive, respond and adapt to changing environmental conditions, understanding of which is vital for global food production and security. One of the most conserved cellular machineries involved in stress signalling is the Target of Rapamycin (TOR) kinase, which acts as an “energy status” sensor, connecting nutrient and energy signalling to growth and homeostasis, under normal and stress conditions. The plant p70 ribosomal S6 kinase (AtS6K) family, composed of two members (AtS6K1 and AtS6K2), is a critical downstream effector of TOR involved in environmental stress response. Despite earlier studies which showed that AtS6K2, but not AtS6K1, is activated in response to cold and salt stresses, many key aspects of its regulation and biological role are poorly understood. Indeed, AtS6K2 is an orphan protein in search of a function.

Here, to gain structural information indispensable to increase our understanding of S6Ks activity, function and regulation, an extensive range of experiments aiming to express, purify and crystallise *Arabidopsis thaliana* S6K2 were attempted. A combination of sequence alignment, modelling and molecular dynamics simulation approaches were employed to predict the protein structure and dynamics, and guide design and optimisation of the expression constructs. A total of 33 different expression experiments, about 26 different purification condition tested, for the 23 constructs cloned, failed to produce samples of active protein amenable for structural studies. Specifically, KL_123_471_NC_C-His_N-T7; KL_123_471_NC_C-His; KD_123_414_NC_C-His_N-T7; KD_123_414_C_C-His_N-T7; and KD_123_414_C_N-GST constructs were expressed in soluble form, purified and used for crystallisation screening. Constructs for co-expression with λ-phosphatase were extremely promising for both crystallisation

(KDN_133_414_NC_C-His_λ and KDN_133_414_C_C-His_λ) and activity assays (FL_1_471_NC_N-His_λ both WT and mutant), however they need to be optimised for larger scale production. Despite the lack of experimental structural data, this case study represents one of the largest, compared to published work on recombinant plant proteins, systematic recombinant protein expression and purification campaign, which provides clear directions for future studies on AtS6K and, in general, plant AGC kinases *(Pearce et al.*, 2010; Rademacher and Offringa, 2012).

Acknowledgements

I would like to thank Dr Filippo Prischi and Dr Ulrike Bechtold for their support throughout my PhD. I could have never expected so much positivity and push towards common goals during the last four years. It was an eye-opening experience on how can one work in academia with so much passion and humane qualities. I felt a constant academic and physiological support while working in Essex, mainly thanks to my supervisors. My lab colleagues also helped me a greatly, namely Dr Marina Rozman and Dr Ryan Cronin, who I met before acquiring their PhDs and with whom I had many amazing memories in and out of the lab. I thank my family that was there to support and encourage. Deep thanks to Sima, my girlfriend who provided never ending energy supply through my last days and weeks of writing. Finally, I want to give special thanks to Prof Dr Karsten Niefind and Christine Tölzer, both of who initiated my crystallography dreams and made the it possible to aim for this PhD.

Contents

List of abbreviations

Å	Ångstrom
β-MEt	beta-mercaptoethanol
°C	degrees Celcius
µg	microgram
µL	microlitre
µM	micromolar
α	alpha
β	beta
λ	lambda
AMP-PNP	adenyline-imidodiphosphate
AtS6K1	*Arabidopsis thaliana* p70 ribosomal protein S6 kinase 1
AtS6K2	*Arabidopsis thaliana* p70 ribosomal protein S6 kinase 2
ATP	adenosine triphosphate
bp	base pairs
C	carbon
Ca	calcium
cm	centimetres
Co	cobalt
CR	conserved region
CV	Column Volume
Da	Dalton
ddH_2O	double deionised water
DMSO	Dimethyl Sulphoxide
DNA	Deoxyribonucleic Acid
DNase	Deoxyribonuclease
dsDNA	double-stranded Deoxyribonucleic Acid
EDTA	Ethylenediaminetetraacetic Acid

eIF	eukaryotic Initiation Factors
FL	Full Length
g	gram
GST	Glutathione-S-Transferase
h	hour
HM	Hydrophobic Motif
HP SP	High Performance Sulphopropyl
IMAC	Immobilised Metal-ion Affinity Chromatography
IPTG	Isopropyl-β-D-Thiogalactopyranoside
kb	kilobases
KD	Kinase Domain
kDa	kilo Dalton
LB	Luria broth
m	metre
M	molar
mg	milligram
Mg	magnesium
min	minute
mL	millilitres
mM	millimolar
Mn	manganese
mol	mole
MR	molecular replacement
mRNA	messenger Ribonucleic Acid
mTOR	mammalian Target of Rapamycin
mTORC1	mammalian Target of Rapamycin Complex 1
MW	molecular weight
N	nitrogen
Na	sodium
Ni	nickel

NLS	Nuclear Localisation Signal
nm	nanometre
NTA	Nitrilotriacetic Acid
O	oxygen
P	phosphorous
PCR	Polymerase Chain Reaction
PDB	Protein Data Bank
PDK1	3-phosphoinositide-dependent protein kinase 1
PEG	Polyethyleneglycol
PI3K	Phosphoinositide 3-Kinase
RNA	Ribonucleic Acid
RNAi	Ribonucleic Acid interference
rpm	revolutions per minute
rRNA	ribosomal Ribonucleic Acid
S	sulphur
s	second
S6K1	p70 ribosomal protein S6 kinase 1
S6K2	p70 ribosomal protein S6 kinase 2
SDS	Sodium Dodecyl Sulphate
SDS-PAGE	Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis
SEC	Size Exclusion Chromatography
T_m	melting temperature
TM	Turn Motif
Tris	tris(hydroxymethyl)aminomethane
UTR	Untranslated Region
UV	ultraviolet
V	Volts

List of Amino Acid and Nucleotide Abbreviations

Amino acids

Ala	A	Alanine
Arg	R	Arginine
Asn	N	Asparagine
Asp	D	Aspartic Acid
Ala	A	Alanine
Cys	C	Cysteine
Glu	E	Glutamic Acid
Gln	Q	Glutamine
His	H	Histidine
Ile	I	Isoleucine
Leu	L	Leucine
Lys	K	Lysine
Met	M	Methionine
Phe	F	Phenylalanine
Pro	P	Proline
Ser	S	Serine
Thr	T	Threonine
Trp	W	Tryptophan
Tyr	Y	Tyrosine
Val	V	Valine

Nucleotides

A	Adenine
C	Cytosine
G	Guanine
T	Thymine
U	Uracil

1. Introduction

The ability to maintain internal equilibrium in response to external stresses and changes is a distinctive feature of all living organisms. Proficiency in coping with stress conditions is especially important to plants that cannot relocate as a response to environmental cues. As such, sessile plants have evolved unique ways to adapt to stress (Verma, Nizam and Verma, 2013). Major crop yield-limiting factors such as extreme temperatures, drought, flooding, and pathogen infections perturb homeostasis and pose a growing challenge to crop plants worldwide (Alcázar *et al.*, 2006). Environmental stress causes energy deprivation due to inhibition of photosynthesis (Baena-González and Sheen, 2008). As a result, a variety of cellular processes are altered, including the activation of rapid protective mechanism (e.g.: antioxidants and osmolyte accumulation) and other metabolic reprogramming promoting stress tolerance often at the expense of plant growth (Baena-González, 2010; Bechtold and Field, 2018; Bechtold, Ferguson and Mullineaux, 2018). Since crop plants represent a major source of food throughout the globe, reduced harvest yield is a major challenge for current and future global food production (Gornall *et al.*, 2010).

Plants have evolved unique tightly regulated signaling pathways that detect environmental fluctuations and elicit an appropriate reaction (Dobrenel *et al.*, 2011). Reversible phosphorylation is one of the most widespread and well-researched mechanisms for the regulation of general signal transduction pathways as well as specific protein properties (Hunter, 1995). Phosphorylation is catalyzed by kinases while the reverse process called dephosphorylation is catalyzed by phosphatases. Protein kinases, regulatory enzymes that alter the properties of their substrates by attaching a phosphate group to serine, threonine or tyrosine amino acid residues, are not only the main players in

these signaling pathways but are also involved in the regulation of most cellular processes of all eukaryotic organisms (Endicott, Noble and Johnson, 2012). Kinases act as molecular switches that are capable of finely tuning gene expression depending on internal and external conditions, resulting in a broad array of phenotypes. Evolutionary studies have evidenced that kinases are highly conserved across eukaryotes, and while the human genome is predicted to code for approximately 500 kinases (Manning *et al.*, 2002), the kinome of plants is generally larger, ranging in size between 600 and 2500 members (Lehti-Shiu and Shiu, 2012). For example, around 1000 kinases can be identified in the Arabidopsis genome (Zulawski *et al.*, 2014). Eukaryotic plant kinases can be grouped into two paraphyletic groups: the membrane-located receptor-like kinases and the soluble cytosolic kinases. Both types of kinases catalyze the transfer of a phosphoryl group from ATP to the hydroxyl group of specific amino acid residues, mainly serine, threonine and tyrosine (Champion *et al.*, 2004). In addition, atypical kinases of prokaryotic origin and kinases containing nonstandard kinase domains (e.g. PDK1) have been identified in plants (Zulawski and Schulze, 2015).

In 1988, p70 ribosomal protein S6 kinases (S6Ks) were identified in animal models as kinases that phosphorylate the 40S ribosomal protein S6 (rpS6) (Jenö *et al.*, 1988; Nemenoff *et al.*, 1988) in response to growth factors and mitogens signalling (Gressner and Wool, 1974a, 1974b; Anjum and Blenis, 2008). It was not until 1994 that the plant S6Ks orthologues were identified in Arabidopsis (Zhang, Broome, *et al.*, 1994; Zhang, Lawton, *et al.*, 1994). Specifically, plant S6Ks were originally identified via amplification of human kinase homologues in Arabidopsis genomic DNA using two degenerate oligonucleotides coding for the amino acid sequence of a conserved motif in the catalytic domains of the human protein kinase C (PKC) and protein kinase A (PKA), which resulted in the isolation of two genes, AtS6K1 and AtS6K2 (also called atpk1/atpk6 and

atpk2/atpk19) (Zhang, Lawton, *et al.*, 1994). The head-to-tail tandem array organization on chromosome 3 and the conservation of intron-exon boundaries within the genomic sequences strongly suggest that the two genes were originated by an event of gene duplication (Zhang, Lawton, *et al.*, 1994; Turck *et al.*, 1998). The two AtS6K genes were not initially assigned as AGC kinases because of low sequence identity (42 – 53%) with kinase domains of human PKA, PKCα, and S6K1. The name "AGC" stands for the three founding members of this kinase group: Protein Kinase A (PKA), Protein Kinase G (PKG), and Protein Kinase C (PKC). An initial functional characterization study of AtS6K1 identified several ribosomal proteins as potential substrates (Zhang, Broome, *et al.*, 1994). The first *in vitro* evidence that AtS6K1/2 are rpS6 kinases came from the ectopic expression of AtS6K2 in human 293 cells and incubation of AtS6K2 isolated from plant cells with the 40 S subunit, which in both cases resulted in an increased phosphorylation of rpS6 (Turck *et al.*, 1998).

Due to the high sequence identity of AtS6K1 and AtS6K2 (88% overall and 96% for the catalytic kinase domain; Fig. 1.1), AtS6K1 has been considered the prototypical S6Ks and AtS6K2 has been largely neglected. However, a growing number of studies have indicated that two kinases are functionally different. For example, the mRNAs for AtS6K1 and AtS6K2 cycle with opposite phases, AtS6K1 peaking around dawn and AtS6K2 peaking in the afternoon (Mockler *et al.*, 2007). AtS6Ks have different cellular localization, with AtS6K1 predominantly localized in the cytoplasm, and AtS6K2 in the nucleus (Mahfouz *et al.*, 2006; Sun *et al.*, 2016). Furthermore, expression data from the Arabidopsis eFP Browser (Winter *et al.*, 2007) shows that AtS6K1 expression is induced in response to UV-B, oxidative, and genotoxic stresses and downregulated following osmotic stress specifically in shoot, while AtS6K2 is overexpressed in plant roots in response to salt treatment. While information about AtS6K isoform-specific roles is limited, initial

evidence suggest that AtS6Ks mediate different stress response pathways (Turck *et al.*, 2004; Mahfouz *et al.*, 2006; Rajamäki *et al.*, 2017). AtS6K2 is involved in abiotic stress responses (Mizoguchi *et al.*, 1995), however compared to AtS6K1 there is less information regarding AtTOR-AtS6K2 pathway and the only relevant structural data comes from crystallographic data of human S6K1 (Wang *et al.*, 2013).

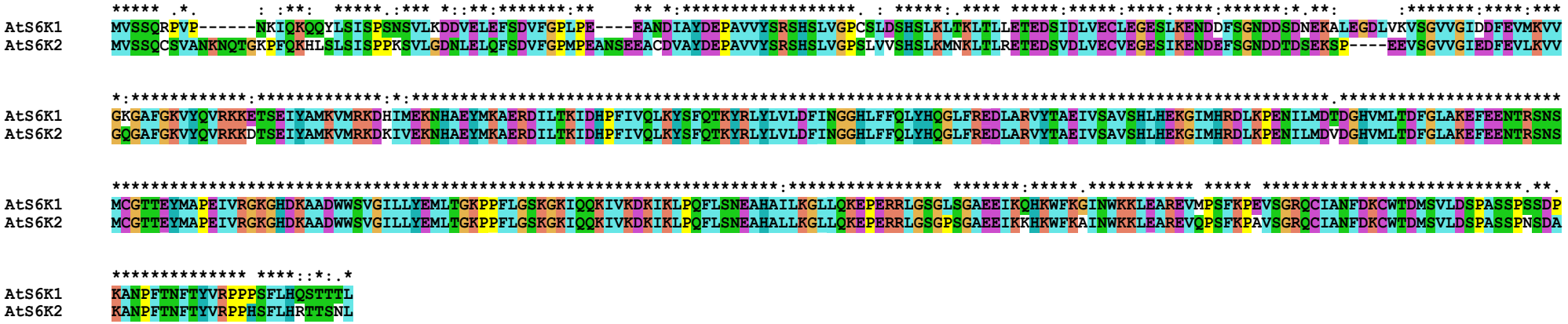


Fig. 1.1| AtS6K1 alignment with AtS6K2 (88% overall sequence identity). Amino acid residues were colour coded: light blue – hydrophobic; red – positively charged; magenta – negatively charged; green – polar; orange – glycines; yellow – prolines; cyan – aromatic; and white – unconserved.

1.1 Target of Rapamycin

Target of Rapamycin (TOR), is a Ser/Thr kinase that is viewed as the master regulator of growth and metabolism which directly or indirectly regulates transcription, translation, ribosome biogenesis, translocation of regulatory proteins, autophagy, and storage of reserve compounds (Dobrenel *et al.*, 2011). Nutrient availability activates plant TOR in contrast to Snf1-Related Protein Kinase 1 (SnRK1), which operates under low carbon conditions, promotes energy saving and counteracts TOR (Robaglia, Thomas and Meyer, 2012; Margalha, Confraria and Baena-González, 2019). SnRK1 is heterotrimeric kinase that regulates AtTOR by directly interacting and phosphorylating AtRAPTOR (Regulatory-Associated Protein of TOR), an important component of the AtTOR complex (Nukarinen *et al.*, 2016). AtS6Ks are downstream effectors of AtTOR and, similarly to most eukaryotes (Kozma and Thomas, 2002; Liao *et al.*, 2008), the S6Ks pathway regulates cell growth, cell proliferation, and stress response by modifying protein synthesis and ribosomal biogenesis (Turck *et al.*, 1998; Mahfouz *et al.*, 2006; Rhoads, Dinkova and Jagus, 2007).

TOR was discovered in 1991 via isolation of *S. cerevisiae* mutants resistant to rapamycin, that is a naturally secreted product of the soil bacteria that acts as an anti-microbial agent by inhibiting TOR (Heitman, Movva and Hall, 1991). In plants, TOR is essential to lipid and starch metabolism, cell cycle progression, cell wall modifications and senescence in addition to cell proliferation in shoot and root meristems (Caldana *et al.*, 2013; Li *et al.*, 2017). In *Arabidopsis*, the only copy of At*TOR* gene is located in the lower arm of chromosome 1. Inactivation of the gene results in embryonic arrest, implying AtTOR importance for the early development of *Arabidopsis*. Homozygous *AtTOR* knockout mutants would not evolve further than the dermatogen stage while cellular divisions still occurred, proposing that disruption of *AtTOR* is not detrimental for the cell-

division machinery (Menand *et al.*, 2002). AtTOR was found to be expressed in proliferating cells and tissues, i.e. primary meristem, embryos, and endosperms, while it was absent in differentiated cells (Menand *et al.*, 2002). In contrast, AtTOR mRNA is expressed almost equally in all plant tissues suggesting post-transcriptional regulation of AtTOR (Robaglia *et al.*, 2004). Translational regulation is often linked to distinct features of untranslated mRNA sequences, which are free from restraints of encoding proteins and may be able to influence any stage of translation (Leppek, Das and Barna, 2018). Importantly, highly conserved in *Arabidopsis*, uORF (upstream open reading frame) sequence was identified upstream of the AtTOR coding region providing insights on possible post-transcriptional regulation mechanism (Robaglia *et al.*, 2004). AtTOR is localised in both nucleus and cytoplasm, as shown by transient AtTOR-GFP expression in onion (*Allium cepa*) epidermal cells. An RPRK motif (amino acids 2077-2080) in the kinase domain is necessary for the nuclear localization of AtTOR (Ren et al., 2011). TOR regulatory function is maintained through interaction with various partners (Jacinto and Hall, 2003). In mammals, two structurally and functionally diverse high molecular weight protein complexes use mTOR as their kinase activity core: mTORC1 and mTORC2. However, an increasing volume of evidence has shown that TORC1 is the only TOR complex in plants (Maegawa *et al.*, 2015). Similarly to mammals, the plant TORC1 complex is composed of TOR, RAPTOR and LST8 (Rexin *et al.*, 2015). Research on Arabidopsis RAPTOR (AtRAPTOR) suggests that the same region, i.e. HEAT repeats, is responsible for RAPTOR binding to TOR in plants and humans (Mahfouz *et al.*, 2006). Furthermore, both *Arabidopsis* and mammalian LST8 interact with the kinase domain of TOR (Shi, Wu and Sheen, 2018). RICTOR, a component of mTORC2, has not been found in plant species or the absence of the subunit has not been defined (Xiong *et al.*, 2017a). Growth-related TOR pathway output mechanisms appear to have a higher degree of

conservation across eukaryotic groups than upstream inputs. In fact, conservation of TOR structural elements across eukarya implies a well-established functionality that is backed by a substantial amount of biochemical, genetic and pharmacological data (Rexin *et al.*, 2015).

1.2 Structure and activation of AtS6Ks

The AtS6Ks is a serine/threonine protein kinase family which belong to the AGC kinase group (Zhang, Lawton, *et al.*, 1994). The AtS6Ks consist of a) a non-conserved N-terminal regulatory domain, containing the TOR signalling (TOS) motif; b) a catalytic kinase domain with an activation loop; c) a regulatory AGC C-terminal region, containing the hydrophobic motif (HM) and the Turn motif (TM) (Fig. 1.2A) (Turck *et al.*, 1998; Bögre *et al.*, 2003). Although 3D structures of plant S6Ks or of any other plant AGC kinase domain are not available, general kinase fold components can be identified using sequence analysis (Endicott, Noble and Johnson, 2012). Approximately 256 residues long, the kinase domains of AtS6Ks consists of an N-lobe, mainly of beta-sheet structure with an α-helix (i.e. C-helix), a predominantly alpha-helical C-lobe, and the ATP binding site sandwiched between the two lobes (Endicott, Noble and Johnson, 2012b; Rademacher and Offringa, 2012). The activation segment (DFG-X12-SMCGTTEYMAPE), containing of the DFG motif, required for binding with the Mg^{2+} which orients the ATP for ɣ-phosphate transfer, the activation loop (SMCGTTEY) and the APE motif, that mediates binding of substrates, are located in the C-lobe (Fig. 1.2A) (Endicott, Noble and Johnson, 2012; Rademacher and Offringa, 2012). Additionally, ATP binding is stabilized by interaction with the conserved glycine-rich loop (GXGXXG) within the N-lobe (Fig. 1.2A) (Endicott, Noble and Johnson, 2012; Rademacher and Offringa, 2012). Sequence analyses of AtS6Ks revealed two putative NLS (Nuclear Localization Sequence) motifs located at both ends of

the kinase domain in AtS6K2, as opposed to a single NLS in AtS6K1, situated at the N–terminus of its kinase domain (Fig. 1.2A) (Mahfouz *et al.*, 2006). Similarly, expression of AtS6K2-GFP fusion protein in BY2 (*Nicotiana tabacum*) cells and AtS6K1-GFP protein in both BY2 and Arabidopsis cells, demonstrated nuclear and cytoplasmic localization of AtS6K1 while AtS6K2 was mostly confined within the nucleus (Mahfouz *et al.*, 2006; Henriques *et al.*, 2010). These results underline the functional diversity of the two S6K isoforms in plants.

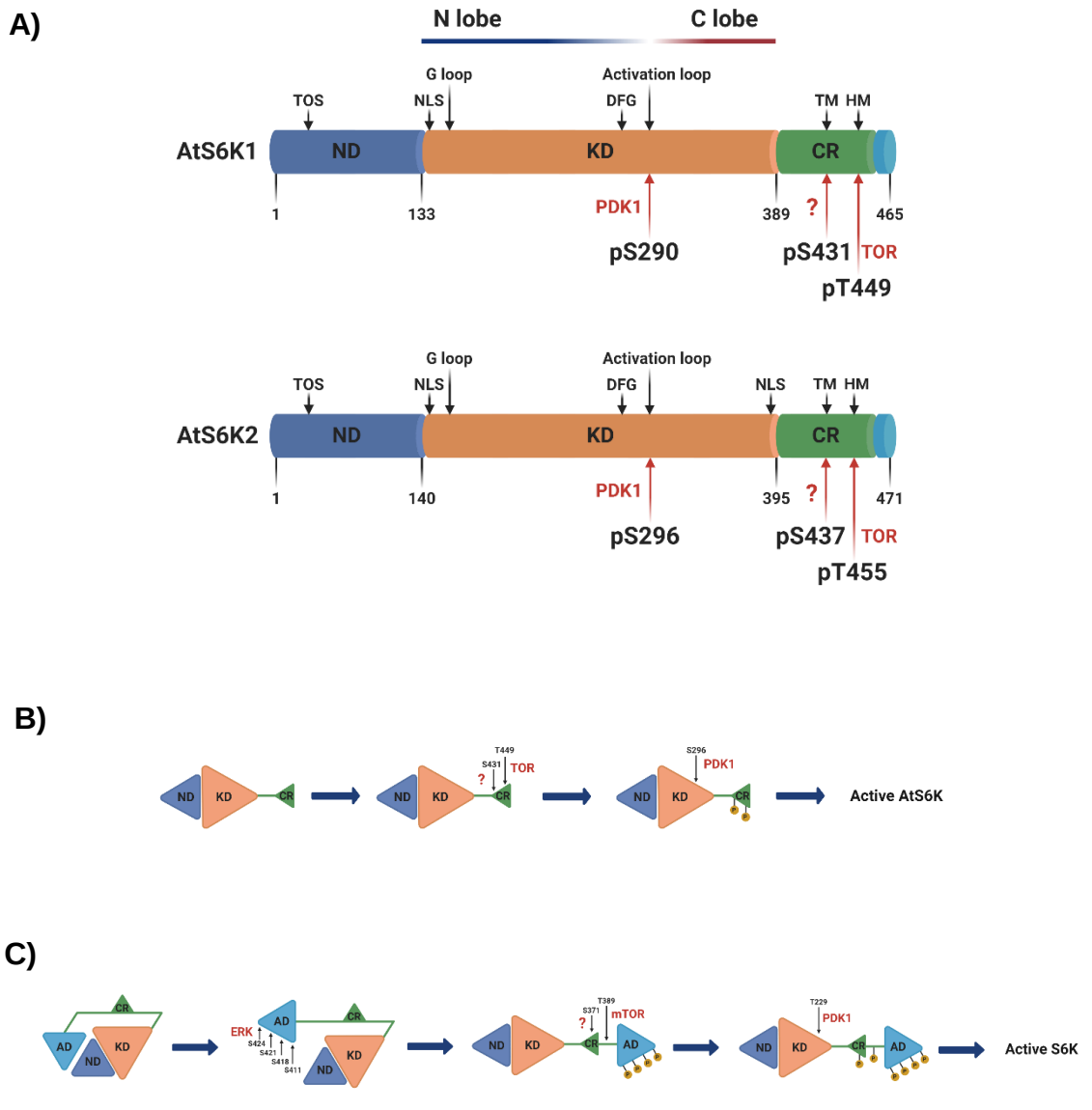


Figure 1.2| Schematic representation of AtS6Ks architecture and activation. **(A)** Comparison of AtA6K1 and AtS6K2 primary structures colored by domain. The nonconserved N-terminal domain (ND) is depicted in blue; the AGC kinase domain in orange (KD); the AGC C-terminal region in green (CR); and the nonconserved C-terminal domain in light blue. Motifs and phosphorylated residues are highlighted in the structure with black and red arrows respectively. TOS, TOR signaling motif; TM, turn motif; HM, hydrophobic motif; NLS, nuclear localization sequence; G loop, glycine rich loop. **(B)** Stepwise models of activation of Arabidopsis and **(C)** human S6Ks. The residue numbers correspond to Arabidopsis and human S6K1. AD, human S6Ks autoinhibitory domain, missing in the plant proteins, created with BioRender.com.

The activity of plant S6Ks can be modulated by several factors that trigger or directly affect TOR signalling, which includes auxin and other phytohormones (Turck *et al.*, 2004; Schepetilnikov *et al.*, 2013; Zhang *et al.*, 2016), insulin or insulin-like growth factors (Jiménez, Beltrán-Peña and Ortíz-López, 1999; Garcia Florez *et al.*, 2001; Reyes de la Cruz, Aguilar and Sánchez de Jiménez, 2004; Dinkova *et al.*, 2007; Garrocho-Villegas, Aguilar C and Sánchez de Jiménez, 2013), nutrients (Zhang *et al.*, 2016; Van Leene *et al.*, 2019), a range of stress factors (Williams *et al.*, 2003; Mahfouz *et al.*, 2006), light signals (Turkina, Klang Årstrand and Vener, 2011; Chen *et al.*, 2018; Enganti *et al.*, 2018) and also several synthetic compounds designed to target the TOR-S6K pathway (Xiong and Sheen, 2012; Schepetilnikov *et al.*, 2013; Xiong *et al.*, 2017a). The activation of AtS6Ks is initiated by the binding of Arabidopsis TOR Complex 1 (AtTORC1) to the TOS motif, which, differently from the canonical TOS motif found in other eukaryotes, is composed of a stretch of 44 amino acid with a 12 amino acid core element (Son *et al.*, 2017a). AtTORC1 is composed of 3 proteins with distinct roles; 1) AtTOR, a large (~250 KDa) evolutionary conserved serine/threonine protein kinase belonging to the phosphatidylinositol 3-kinase-related kinase (PIKK) family, 2) Regulatory-Associated Protein of TOR (AtRAPTOR), which influence the activity and substrate specificity of TOR, and 3) a small Lethal with Sec Thirteen 8 (AtLST8) protein, involved in TOR-mediated signalling processes (Anderson, Veit and Hanson, 2005; Moreau *et al.*, 2012; Rexin *et al.*, 2015). Following direct interaction between AtS6Ks and AtRAPTOR1 (Mahfouz *et al.*, 2006), AtTOR phosphorylates T449 and T455 within the HM (FTNFTpYVRP) located in the AGC C-terminal region of AtS6K1 and AtS6K2 respectively (Fig. 1.2B) (Xiong and Sheen, 2012). This creates a docking site, called PDK1 interacting fragment (PIF) (Biondi *et al.*, 2000; Frödin *et al.*, 2002), recognised by Arabidopsis 3-Phosphoinositide-Dependent Kinase 1 (AtPDK1), which directly binds AtS6Ks and phosphorylates Ser-290

and Ser-296 on the activation loop (SpMCGTEEY) of AtS6K1 and AtS6K2, respectively (Mahfouz *et al.*, 2006). In many AGC-kinases the phosphorylated PIF has also been shown to enhance kinase activity by folding back on to the N-lobe of the kinase domain48, but no studies have investigated this mechanism in plant S6Ks.

Phosphorylation of the activation loop is a well characterised conserved priming mechanism that enables kinases to phosphorylate their substrates (Rademacher and Offringa, 2012). Specifically, in a resting state the activation loop is usually disordered, preventing the binding of both Mg^{2+} and substrates (Nolen, Taylor and Ghosh, 2004). Phosphorylation of the activation loop promotes formation of a network of hydrogen bonds with the C-helix and the activation segment, resulting in an “open” structure conformation allowing binding and phosphorylation of a wide range of substrates (Huse and Kuriyan, 2002). Interestingly, other post-translational modifications (PTMs) have been reported in human S6Ks but, due to limited studies on plants, no information about the presence and roles of additional PTMs on AtS6Ks is available (Wang *et al.*, 2008; Gwalter, Wang and Gout, 2009; Fenton *et al.*, 2010). Despite functional conservation with the human S6Ks (Turck *et al.*, 1998), the activation mechanism of plant and human proteins is significantly different. The human S6Ks are activated by ERK, which phosphorylates the S6Ks C-terminal autoinhibitory domain, which is missing in the plant proteins (Pardo and Seckl, 2013). Only upon release of this autoinhibitory segment, are TORC1 and PDK1 able to phosphorylate S6Ks (Fig. 1.2C) (Pardo and Seckl, 2013). This would suggest that, while the kinase domain has an evolutionary conserved function (Turck *et al.*, 1998), the N- and C-terminal of the protein may provide species-specific regulation and substrate selection (Rademacher and Offringa, 2012). The Kozaki group (Yaguchi and Kozaki, 2018a; Yaguchi, Ikeya and Kozaki, 2020a) interrogated the role of Arabidopsis, rice, maize, tomato, and soybean S6Ks phosphorylations and, although they were not able to fully

characterise plants proteins activation due to shortfalls of the experimental approach adopted (i.e. the use of a heterologous system, the not quantified ectopic expression of plants proteins in ypk3Δ yeast cells, and the lack of protein-specific antibodies), they were able to confirm the presence of differences in pant S6Ks regulation compared to humans.

1.3 AtS6K Substrates

Conserved amongst the AGC kinase family, the recognition sequence phosphorylated by S6Ks is RXRXXS/T (Tavares *et al.*, 2015). However, knowledge of the consensus sequence provides limited information about AtS6Ks' substrates, since analysis of basophilic kinases consensus sequences, which includes the majority of known Ser/Thr protein kinases, reveals high level of similarities (Pinna and Ruzzene, 1996). In fact, AGC kinases can redundantly phosphorylate overlapping sites on substrates, further complicating the assignment of a specific substrate to a kinase within a signaling pathway (Manning and Cantley, 2007; Mendoza, Er and Blenis, 2011). Data on plant kinase promiscuity are extremely limited, but studies on human kinases showed that several S6Ks substrates can also be targeted by p90 ribosomal S6 kinase (RSK), Protein Kinase B (Akt), Protein Kinase A (PKA) (Manning and Cantley, 2007; Mendoza, Er and Blenis, 2011). Importantly, while kinases preferentially phosphorylate residues within the consensus motif, recent structural and biochemical studies have shown that substrate recognition can be driven by docking interactions with regions distant from the phosphorylation and the kinase catalytic sites. It has also been shown that temporally and spatially regulated complex interactions are required for substrate phosphorylation *in vivo* (Miller and Turk, 2018). It is not known yet if plant S6K-isoform specific features and, most importantly, substrates recognition are driven by the non-conserved N-termini.

The most well characterised substrate and primary target of S6Ks is the ribosomal protein S6 (rpS6). In Arabidopsis, two rpS6 have been identified, rp6SA and rpS6B, which are functionally equivalent (Chang *et al.*, 2005; Creff, Sormani and Desnos, 2010). rpS6 is located in the mRNA/tRNA binding site of the 40S subunit and is the only protein in the 40s ribosomal subunit to be phosphorylated in a controlled way (Mahfouz *et al.*, 2006a). rpS6 phosphorylation increases cap-dependent translation in humans (Roux *et al.*, 2007), but its role in plants is largely unknown (Enganti *et al.*, 2018). Studies on maize have shown that auxin stimulation enhanced rpS6 phosphorylation with a concomitant recruitment of 5′TOP-like mRNAs into polysomes (Levy *et al.*, 1991; Beltrán-Peña *et al.*, 2002). This would suggest that, at least in maize tissues, rpS6 phosphorylation regulates translation of specific proteins downstream of auxin signalling (Beltrán-Peña *et al.*, 2002). Plant rpS6 is most commonly phosphorylated on two residues S237 and S240, which have been detected in Arabidopsis (Enganti *et al.*, 2018) and maize (Williams *et al.*, 2003). The level of phosphorylation on these residues changes during the day and night. Turkina et al. (Turkina, Klang Årstrand and Vener, 2011) carried out an extensive mass spectrometry study on Arabidopsis cytosolic rpS6 and identified a novel phosphorylation site on S231. A higher phosphorylation level was detected during the daytime and, specifically, the day/night phosphorylation ratio on S231, S237 and S240 was 2.2, 4.2 and 1.8, respectively. The authors speculated that the higher phosphorylation level of rpS6 is at least in part responsible for the higher protein synthesis during the light period (Turkina, Klang Årstrand and Vener, 2011). The number of rpS6 phosphorylations is different in humans. In fact, S6Ks phosphorylate rpS6 on S235 (structurally equivalent to S240 in Arabidopsis), S236, S240 and S244, while RSKs (absent in plants) phosphorylate only two (S235 and S236) (Roux *et al.*, 2007). Despite these differences, earlier experiments by Turck et al.(Turck *et al.*, 1998) showed that human 293 cells treated with rapamycin had

reduced rpS6 phosphorylation, which was rescued by transient expression of AtS6K2. Human rpS6 phosphorylation was detected via ^{32}P incorporation, so information about the residues phosphorylated by AtS6K2 is not available. However, this experiment showed that AtS6Ks are functional homologues of human S6Ks (Roux *et al.*, 2007).

Studies focusing specifically on S6K2 are scarce, in both human and plants (Pardo and Seckl, 2013). Xiong et al. (Xiong *et al.*, 2017a) showed that AtS6K2, but not AtS6K1, regulates photoautotrophic growth downstream of AtTOR. In order to map the network of AtTOR-AtS6K2 interaction, the authors used a yeast two-hybrid screening and identified Brassinosteriod Insensitive 2 (BIN2) as a key binding partner of AtS6K2 (Fig. 1.3). BIN2, one of ten GSK3-like kinases in Arabidopsis, is responsible for blocking the transduction of brassinosteroids (BRs) signal by phosphorylating and deactivating the transcription factors BES1 and BZR1 (Peng *et al.*, 2008; Li *et al.*, 2020). BIN2 is a homologue of human GSK3β and, since GSK3β is phosphorylated by human S6Ks, Xiong et al. (Xiong *et al.*, 2017a) carried out kinase assays and co-immunoprecipitation experiments to study if a similar regulation was also present in plants. The authors demonstrated that BIN2 directly binds AtS6K2, but not AtS6K1, and is phosphorylated on S187 and S203 by AtS6K2 in an AtTOR dependent way. This data suggests that AtTOR-AtS6K2-BIN2 pathway negatively regulates photoautotrophic growth (Xiong *et al.*, 2017a).

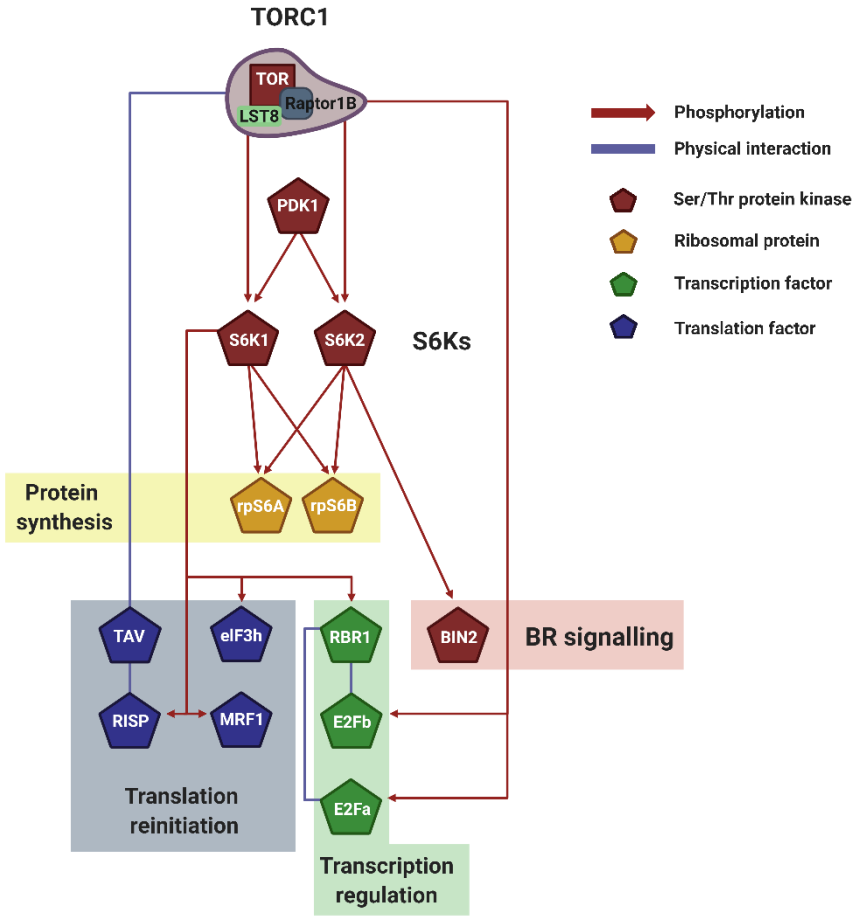


Figure 1.3| Visualization of plant S6Ks signaling network. Schematic representation of the downstream TORC1/S6Ks signaling events, which control transcription, translation, and ribosomal biogenesis. The network legend is shown on the right. Created with BioRender.com.

AtS6Ks play a major role in plant reinitiation of translation (Schepetilnikov *et al.*, 2011, 2013; Lee *et al.*, 2017). Leet et al. (Lee *et al.*, 2017) recently demonstrated that increase in cellular energy causes phosphorylation of Arabidopsis MA3 Domain-Containing Translation Regulatory Factor 1 (MRF1) downstream of AtTOR-AtS6Ks (Fig. 1.3). MRF is a family of translation regulatory factors composed of 4 isoforms, MRF1-4, which are transcriptionally and functionally regulated by TOR. Rapid phosphorylation of MRF1 under light and glucose conditions was seen to positively correlate with MRF1 association with the ribosome and eIF4A-1. In this complex, MRF1 may play a role in helping eIF4A-1 catalyse the unwinding of mRNA at the 5′-UTR to facilitate ribosome scanning (Lee *et al.*, 2017). It was concluded that the TOR-S6Ks-MRF1 pathway has a potential role in rapidly rebooting translation when the environment is favourable for growth (Lee *et al.*, 2017). Furthermore, association of AtS6K1 with the translation reinitiation-promoting factor eIF3 non-core subunit h (eIF3h) is triggered by the phytohormone auxin downstream of AtTOR (Fig. 1.3) (Schepetilnikov *et al.*, 2013). In a resting state, AtS6K1 is inactive and bound to polysomes. Upon auxin treatment, AtTORC1 associates with the polysome, phosphorylates AtS6K1, which is then able to phosphorylate eIF3h, activating the translation factor. Active eIF3h increases ribosomal loading of uORF-containing mRNAs, promoting, among others, translation of auxin response factors (ARFs) and basic zipper transcription factors (bZIPs) (Kim *et al.*, 2004; Schepetilnikov *et al.*, 2013). AtS6K1 activation by auxin-triggered AtTOR contributes to reinitiating the translation of mRNAs containing uORFs in the 5’-untranslated regions (Schepetilnikov *et al.*, 2013). A specific case of reinitiation is induced by TAV, the Cauliflower mosaic virus (CaMV) reinitiation factor (Schepetilnikov *et al.*, 2011). TAV binds and activates AtTOR, resulting in polysome association and AtS6K1 phosphorylation. AtS6K1 in turn phosphorylates the reinitiation-supporting protein (RISP) on S267. TAV then binds phosphorylated RISP, keeping it

associated to the polysome. In this way, after termination of translation, ribosomes are equipped with the necessary machinery to reinitiate translation of the downstream ORFs (Thiébeauld *et al.*, 2009, p. 3; Schepetilnikov *et al.*, 2011).

AtS6K1 is able to repress cell proliferation under nutrient-limiting conditions by directly binding and activating Retinoblastoma-Related 1 (AtRBR or AtRBR1; Fig. 1.3) (Henriques *et al.*, 2010). AtRBR is a highly conserved protein that inhibits the activity of the transcription factors E2Fs (E2FA/B/C) (van den Heuvel and Dyson, 2008). E2F transcription factors are targets of universal CYC-CDK-RBR cascade that initiates cell cycle (Francis, 2007). E2Fa and E2Fb phosphorylations activate S-phase genes, which are pivotal to cell proliferation in shoot and root apexes (Li *et al.*, 2017). Co-immunoprecepitation experiments have shown that AtS6K1 can bind directly with AtRBR (Henriques *et al.*, 2010). Although experimental evidences are available only for AtS6K1, both AtS6K1 and AtS6K2 contains the LxCxE-like motif on the N-term domain required for AtRBR binding. Interestingly, E2FB was also detected in co-immunoprecepitation experiment and its relative abundance was higher in presence of AtS6K1. This could suggest that AtS6K1 is able to associate with the AtRBR-E2Fs complex. Although AtRBR might be phosphorylated by AtS6K1 upon binding, no experimental data have linked a specific phosphorylation on AtRBR to AtS6K1. The binding of AtS6K1 results in the nuclear localisation of AtRBR-E2Fs and, conversely, RNAi silencing of AtS6K1 increases cytosol amount of AtRBR. Importantly, silencing of AtS6K1 or reduction of AtS6K1/2 (s6k1s6k2/++ hemizygous mutants) caused an increase in chromosome numbers, which was linked to the downregulation of AtRBR and concomitant increase in E2Fs activity. AtS6Ks can negatively regulate cell division through RBR1–E2F transcriptional switch and are important to chromosome stability as their deregulation can alter the incident rate of polyploidization (Henriques *et al.*, 2010).

1.4 S6Ks in growth and development

Plant growth is a function of a series of cell divisions and expansion which occur in specialized regions known as meristems. The meristems, shoot apical meristem (SAM) and root apical meristem, are made up of continuously dividing and growing cells which later form specialised tissues during cell expansion and differentiation. S6Ks have been linked to plant growth and development especially in meristems and regions of active cell proliferation. For example, expression of S6K1 was observed in both apical meristems and fast-growing organs like lateral root tips and reproductive tissues of Arabidopsis (Zhang, Lawton, *et al.*, 1994; Tzeng *et al.*, 2009). In Arabidopsis protoplasts silencing of AtS6Ks led to an increase in cell number and cell cycle regulators, suggesting a negative effect of AtS6Ks on cell proliferation (Henriques *et al.*, 2010). S6Ks involvement in regulating cell size was also highlighted when ectopic expression of Lily S6K1 (LS6K1) in Arabidopsis resulted in flowers with short petals and stamens due to a significant inhibition of cell expansion (Tzeng *et al.*, 2009). This implies, at least in Arabidopsis, that S6Ks are involved in the transition from cell division to cell expansion. In agreement with this, disruption of both AtS6K1 and 2 led to halted embryo development and a reduction in epidermal leaf size as a result of inadequate cell expansion, rather than cell division, demonstrating S6Ks role in plant growth and development (Henriques *et al.*, 2010).

The phosphorylation status of plant S6Ks during periods of accelerated growth also links S6Ks to growth responses. For example, phosphorylated maize S6K (ZmS6K) isolated from germinating maize axes positively correlated with germination time, suggesting that plant S6Ks activity is developmentally regulated (Reyes de la Cruz *et al.,* 2004; Dinkova *et al.*, 2007). In addition, S6Ks phosphorylation is equally observed when plant growth is stimulated exogenously with hormones, sugars or growth factors, linking growth to S6Ks activity (Garcia Florez *et al.*, 2001; Xiong and Sheen, 2012; Schepetilnikov

et al., 2013; Van Leene *et al.*, 2019). Cyclodipeptides (CDPs) produced by several bacteria species (plant-growth-promoting-rhizobacteria), known to alter plant root and shoot architecture by altering hormonal responses (Kimura *et al.*, 2005; Ortiz-Castro *et al.*, 2011; González *et al.*, 2017; Ortiz-Castro *et al.,* 2019), have been linked with S6Ks phosphorylation. Indeed, Corona-Sánchez *et al.* (2019) showed that the increase in plant growth by *Pseudomonas aeruginosa* CDPs promoted S6K phosphorylation and activation.

It is suggested that the phosphorylation of RPS6 by S6Ks downstream of the TOR pathway is needed to promote proteins synthesis required for growth processes (Ren *et al.*, 2013). Therefore, the phosphorylation status of rpS6 is used to monitor S6Ks activity since S6Ks are the only known proteins to phosphorylate rpS6 in plants. In fact, increased growth of hormone-stimulated maize axes were positively correlated with an enhanced rpS6 phosphorylation, which was blocked by the application of rapamycin, an inhibitor of the TOR-S6K signalling pathway (Reyes de la Cruz *et al.*, 2004; Dinkova *et al.*, 2007). A similar result was observed with germinating maize callus, suggesting that functional TOR-S6K activity is essential for rpS6 phosphorylation during plant growth (Garrocho-Villegas *et al.*, 2013)

Growth inhibition also induce changes in S6K activity. For example, treatment of 10-day old maize seedlings with high doses of oligogalacturonides (OGs), a plant growth regulator which inhibits coleoptile growth, induced changes in ZmS6K activity in a similar way to rapamycin-treated seedling (Peña-Uribe *et al.*, 2012). Varying degrees of rpS6 phosphorylation was also observed in Arabidopsis leaves at different levels of photosynthetic capacity, suggesting the S6Ks could be modulated by photosynthesis via TOR (Boex-Fontvieille *et al.*, 2013). Indeed, disrupted mutants of rice S6K1 (OsS6K1) displayed defective chloroplasts and reduced photosynthesis, limiting the carbon source important for plant growth, ultimately resulting in reduced growth (Sun *et al.*, 2016). This

produced a phenotype akin to that induced by TOR-inactivated mutants, which was shown to cause downregulation of nuclear genes which code for many chloroplastic proteins (Dobrenel *et al.*, 2016b; Xiong *et al.*, 2017). Owing to TOR pathway being the central hub for integrating signals from the environment to regulate plant growth and development (Deprost *et al.*, 2007; Dobrenel, Caldana, *et al.*, 2016; Zhang *et al.*, 2016; Ryabova *et al.*, 2019), loss of function mutations on members of the TORC1 complex have also been reported to have varying developmental phenotypes in plants (Menand *et al.*, 2002; Anderson, Veit and Hanson, 2005; Deprost *et al.*, 2005; Moreau *et al.*, 2012). The same is true downstream of TORC1 with S6Ks and rpS6 genetic mutants exhibiting varying developmental phenotypes, including smaller leaves, increased trichome branching, decreased root growth and arrested embryo development (Henriques *et al.*, 2010; Kim *et al.*, 2014; Xiong *et al.*, 2017). This clearly demonstrates that modulation of the TOR-S6K pathway and its associated components is crucial for plant growth and development, especially in unfavourable conditions.

1.5. S6Ks and stress

In a fluctuating environment, the biggest challenge to plant survival is the ability to adapt to stressful conditions. The impact of stress on plant growth is well documented and involves a complex co-ordination of physiological processes to maintain homeostasis (Kliebenstein, 2016; Bechtold and Field, 2018; Bechtold *et al.*, 2018). Factors that modulate the TOR pathway in turn impact on the activity of S6Ks thereby affecting growth. One such factor is abiotic stress. Mahfouz *et al.*, (2006) showed that the activity of AtS6K1 was significantly reduced under osmotic stress in a TOR dependent manner, suggesting that the TOR-S6K1 pathway is modulated by osmotic stress. Conversely, transgenic seeds constitutively expressing AtS6K1 were hypersensitive to osmotic stress (Mahfouz *et al.*,

2006), suggesting that plants cope with stressful situations by actively reducing S6K activity via the TOR pathway, thereby limiting growth. In support of this hypothesis, a significant drop in mitotic index, a measure of cell proliferation, during heat stress was accompanied by a complete de-phosphorylation of rpS6 in tomato cell cultures (Scharf and Nover, 1982). Considering growth arrest being one of the major consequence of heat stress in plants (Mittler *et al.*, 2012; Vile *et al.*, 2012; Bita and Gerats, 2013; Albihlal *et al.*, 2018), the increase in mitotic index prior to rpS6 re-phosphorylation was observed during recovery, implicating S6K in stress-related growth arrest (Scharf and Nover, 1982). Although S6K regulation during heat stress was relatively unknown at the time, its involvement became prominent when Turck *et al.* (1998) showed that AtS6K2 was unable to phosphorylate mammalian rpS6 at high temperatures, due to kinase inactivation.

Other types of stresses which affect growth have also been linked with S6Ks activity, mostly by measuring the phosphorylation status of rpS6. Williams *et al.* (2003) reported a reduction in rpS6 phosphorylation in response to oxygen deprivation and a subsequent increase following re-oxygenation in maize root tips. A similar de-phosphorylation of rpS6 was also observed in response to heat stress (Williams *et al.*, 2003; Enganti *et al.*, 2018). In contrast, cold stress stimulated an accumulation of phosphorylated rpS6 (Williams *et al.*, 2003; Enganti *et al.*, 2018), while salt stress had no significant effect (Williams *et al.*, 2003). Although these experiments do not provide direct evidences of endogenous S6Ks activity under stress conditions, due to protein levels falling below the detection limits of western blots (Turck *et al.*, 1998, 2004), there is a positive correlation (Meyuhas, 2015).

1.6 Hypothesis and Aims

Hypothesis – AtS6Ks are key kinase effectors of TOR pathways. Based on similarities with the human orthologue (Pearce, Komander and Alessi, 2010) and *in vivo* plant data (Obomighie *et al.*, 2021, and unpublished), the hypothesis is that AtS6K2 phosphorylation/activation and not overexpression enhances *A. thaliana* resistance to cold and increased salinity in soil stresses. HM phosphorylation by TOR (T455) and subsequent activation loop phosphorylation by PDK1 (S296) activates AtS6K2 and promotes plant stress response.

Aims – The overall goal of this Ph.D. project was to ultimately understand how AtS6K2 modulates cellular responses to changes in the external environment in the model plant Arabidopsis. However, no structural data is available for any plant AGC kinase or even human S6K2. As such, my goal was to carry out a structural and functional characterization of AtS6K2, to rationalize the isoform specific roles of S6Ks in plants, and to ultimately increase our understanding of S6Ks' functions and regulation mechanisms. X-ray crystallography to determine the 3D structures of AtS6K2 in different activation states (wild-type – inactive; T455E_S296E mutant – active) and activity assays were planned. In addition, computational studies were coupled with experimental work to help increase our understanding of AtS6K2 stability and solubility.

2. Methods

2.1 Bioinformatics

2.1.1 Sequence alignment, structure prediction and modelling

The amino acid sequence for AtS6K2 was retrieved from UniProt (ID: Q39030), while the human S6K1 sequences used for crystallisation studies were retrieved from the PDB (ID: 3A62 and 4L43 for the KD and KL constructs, respectively). Sequences were aligned using the Clustal Omega (https://www.ebi.ac.uk/Tools/msa/clustalo/) multiple-sequence alignment algorithm (Sievers *et al.*, 2011), which also provided sequence similarity and conservation.

Secondary structure and disordered regions prediction analyses were performed using the PSIPRED server (http://bioinf.cs.ucl.ac.uk/psipred/) (Jones, 1999). To overcome the intrinsic limitations of the modelling process for proteins with no high sequence similarity homologue with solved 3D structure, the AtS6K2 models were generated using I-TASSER (https://zhanggroup.org/I-TASSER/) (Yang *et al.*, 2015) and AlphaFold (https://alphafold.ebi.ac.uk/) (Jumper *et al.*, 2021). The resulting modelled 3D structures were visualized utilizing PyMOL (Schrödinger, LLC, 2015).

2.1.2. Preparation for Molecular Dynamics simulations

The human S6K1 (hk1) 3D structure and FASTA sequence were retrieved from the Protein Data Bank (PDB code: 4L3J) (Wang *et al.*, 2013) and the UNIPROT database (entry P23443), respectively. The Preparation Wizard implemented in Maestro graphical interface (Madhavi Sastry *et al.*, 2013) was used to resolve high energy intramolecular steric clashes and to add missing amino acid side chains. The co-crystallized ligand and zinc ion were removed to obtain the protein apo form. The hydrogen atoms of the protein were added with PROPKA at pH 7 ensuring accurate ionization states of amino acid residues. The PROPKA server uses its own internal hydrogen placement algorithm and predicts the protonation states at a given pH of ionizable amino acids by assessing their pKa values (Søndergaard *et al.*, 2011). The human S6K2 (hk2), *Drosophila Melanogaster* S6K (dk), *Strongylocentrotus purpuratus* S6K (sk), *Caenorhabditis elegans* S6K (ck), *Arabidopsis thaliana* S6K1 (ak1), *Arabidopsis thaliana* S6K2 (ak2), and *Schizosaccharomyces pombe* (pk) sequences were provided by Dr J. Paps (University of Bristol) (Guijarro-Clarke, Holland and Paps, 2020). Their 3D structures were predicted with a homology modelling approach performed by Schrödinger Suites 2019-2 (Jacobson *et al.*, 2004), using hk1 as template. Loops were refined after each model generation, using a Generalized Born model in aqueous solvent without any constraints. No pre-defined coordination was used to place the Zn^{2+}, Mn^{2+} and Li^{+} ions in the models. Each S6K model was aligned with the human S6K1 3D structure (PDB code: 4L3J) (Wang *et al.*, 2013) that includes the Zn^{2+} ion. The Zn^{2+}, Mn^{2+} and Li^{+} ions were placed in the exact position of Zn^{2+} ion from the human S6K1 structure and coordinated/assumed their coordination during the system equilibration. Residues involved in the coordination were not conserved.

2.1.3. Production and analysis of Molecular Dynamics trajectories

Systems preparation for molecular dynamics (MD) was carried out with tleap from ambertools22 (Case *et al.*, 2022). ATP and two Mg^{2+} ions were inserted in the models by a 3D structural superimposition of the X-ray structure of cAMP-dependent protein kinase catalytic subunit alpha (PKA) (PDB ID: 1Q24) with the S6K models. Ff14SB, gaff force fields and 12-6-4 Lennard Jones potentials were used to parametrize proteins, ligands, and Zn^{2+}, Mn^{2+}, Li^{+}, and Mg^{2+} ions, respectively. All biological systems were inserted in a cubic box filled with water molecules (TIP3P) and Na^{+}, Cl^{-} counter ions to neutralize the system. Periodic boundary conditions were employed, and long-range electrostatics was evaluated through the particle-mesh Ewald (PME) algorithm using a real-space cut-off of 12 Å and a grid spacing of 1 Å per grid point in each dimension. The 12 Å cut-off allows for reasonably accurate short-range interactions while shifting the computationally intensive part of the long-range interactions to the reciprocal space calculations performed by the PME algorithm. Using a relatively short 12 Å real-space cut-off helps avoid artefacts associated with the truncation of electrostatic interactions. A longer cut-off could introduce unphysical interactions and may require additional corrections to account for these artefacts. Additional parameters for ATP and phosphorylated amino acids were added (Meagher, Redman and Carlson, 2003; Craft and Legge, 2005). The systems were initially minimized in three steps: i) only hydrogen atoms; ii) ATP (if present) and side chains; iii) the whole system. Then, an annealing molecular dynamics protocol heated the system in two steps: i) from 0 to 100 K (in 50 ps), and ii) from 100 to 300 K (in a 150 ps). Finally, the system was equilibrated at 300 K for 50 ps with constraints of 5 kcal/mol on the backbone for the first 500 ps, while the pressure was kept constant at 1 atm for 1 ns with a Monte-Carlo barostat. The cut-off for van der Waals interactions was set to 9.0 Å, to account for the majority of significant interactions at a reasonable computational cost and to avoid

artefacts across periodic boundaries. The collision frequency coefficient was set to 2 ps^{-1}. The time step was set to 2 fs and the pressure relaxation time to 3 ps. A total of 10,000 frames were collected for an MD run of 500 ns. Root Mean Square Deviation (RMSD), Radius of Gyration (Rg) and hydrogen bonding analyses of MD trajectories were carried out with CPPTRAJ (Case *et al.*, 2022). CPPTRAJ was also used to compute the hydrogen bond lifetime between water molecules and ions and/or residues along the entire MD run.

2.2. Production of protein expression vectors

DNA corresponding to AtS6K2 gene inserts was amplified via PCR using full-length wild-type AtS6K2 or kinase domain constructs (residues 123 to 414) as templates, which were codon-optimized for bacterial expression. Five different AtS6K2 inserts were used in cloning: (i) FL encoding the full length protein (spanning residues 1 to 471), (ii) KL encoding the Kinase domain with the C-terminal region (residues 123 to 471), (iii) KD encoding for the Kinase Domain (residues 123 to 414), (iv) KN encoding the optimised (Section 3.1.6) Kinase Domain sequence with the C-terminal region (residues 136 to 460), and (v) KDN encoding the optimised (Section 3.1.7) Kinase Domain (residues 133 to 414). FL, KL and KD sequences were initially cloned into a PETM-14 vector for expression of recombinant proteins with a cleavable (HRV 3C PreScission protease) N-terminal hexahistidine tag (FL_1_471_C_N-His, KL_123_471_C_N-His and KD_123_414_C_N-His). KL and KD inserts were also cloned into the pGEX-6P-1 vector in order to produce protein with cleavable (HRV 3C PreScission protease) glutathione S-transferase (GST) tag on the N-terminus (KL_123_471_C_N-GST and KD_123_414_C_N-GST). Similarly, FL, KD and KL constructs were cloned into a pET-21a(+) vector to produce C-terminal hexahistidine tagged proteins with an uncleavable tag (FL_1_471_NC_C-His_N-T7, KL_123_471_NC_C-His_N-T7 and KD_123_414_NC_C-His_N-T7).

In order to generate expression constructs of AtS6K2 with cleavable His-tag on the C-terminus, pET-21a(+) was modified into pET-21(3C) vector by introducing an HRV 3C PreScission protease cleavage site (Sections 2.2.7 and 3.1.4.2). pET-21(3C) was digested with BamHI/XhoI and FL, KD, KL genes were cloned creating recombinant AtS6K2 variants with cleavable His-tag on the C-terminus (FL_1_471_C_C-His_N-T7, KL_123_471_C_C-His_N-T7 and KD_123_414_C_C-His_N-T7). The constructs also included an uncleavable T7 tag (MASMTGGQQMGRGS) on the N-terminus. To remove

the N-terminal T7 tag, a new KL gene insert was synthesised via PCR using primers with NheI/XhoI restriction sites and previously generated FL_1_471_C_C-His_N-T7 as template. Digested KL insert was cloned into both pET-21a(+) and pET-21(3C) vectors creating the KL_123_471_NC_C-His construct that contains uncleavable His-tag and no T7 tag and the KL_123_471_C_C-His construct that contains cleavable His-tag and no T7 tag. KN_136_460_C_C-His insert already included the HRV 3C PreScission protease cleavage site on the C-terminus, and thus was cloned into the unmodified pET-21a(+) vector. Lastly, two KDN inserts were generated via PCR: (i) KDN with cleavable C-terminal His-tag (HRV 3C PreScission protease) and (ii) KDN with uncleavable C-terminal His-tag. Both were cloned into the pETDuet-1 vector, that already contained λ-phosphatase gene, generating co-expression constructs (KDN_133_414_NC_C-His_λ and KDN_133_414_C_C-His_λ). All the constructs that were used for AtS6K2 expression and purification are summarised in the Table 2.1.

Table 2.1| Master Table of AtS6K2 Expression Experiments

Name`	Construct size	Vector	Tag	HRV 3C site	T7 tag	MW/ cleaved	Restriction sites	Media	Expressed	Soluble	WT or mutant	pI/cleaved
FL_1_471_C_N-His	1-471	pETM-14	N-His	Yes	No	55.20580/ 53.43787	NcoI/NotI	AIM	No	–	WT	6,28/6.18
KL_123_471_C_N-His	123-471	pETM-14	N-His	Yes	No	41.96115/ 40.19322	NcoI/NotI	AIM	Yes	No	WT	8,68/8.81
								LB	Yes	Not tested		
KD_123_414_C_N-His	123-414	pETM-14	N-His	Yes	No	35.67717/ 33.90924	NcoI/NotI	LB	No	–	WT	8.63/8.81
KL_123_471_C_N-GST	123-471	pGEX-6P-1	GST	Yes	No	66.61694/ 40.20423	BamHI/XhoI	AIM	No	–	WT	7.63/8.93
								LB	Yes	Yes		
KD_123_414_C_N-GST	123-414	pGEX-6P-1	GST	Yes	No	60.33296/ 33.92025	BamHI/XhoI	LB	Yes	Not tested	WT	7.28/8.95
								AIM	Yes	Yes		
FL_1_471_NC_C-His_N-T7	1-471	pET-21a(+)	C-His	No	Yes	55.48312	BamHI/XhoI	AIM	Yes	Yes	WT	6.43
								LB	Yes	Not tested		
FL_1_471_NC_C-His_N-T7_dm	1-471	pET-21a(+)	C-His	No	Yes	55.55317	BamHI/XhoI	LB	Yes	Yes	T455E+ S296E	6.27
								AIM	Yes	Yes		
KL_123_471_NC_C-His_N-T7	123-471	pET-21a(+)	C-His	No	Yes	42.23847	BamHI/XhoI	AIM	Yes	Yes	WT	8.93
								LB	Yes	Yes		
KL_123_471_NC_C-His_ N-T7_sm	123-471	pET-21a(+)	C-His	No	Yes	42.26648	BamHI/XhoI	LB	Yes	Not tested	T455E	8.81
KL_123_471_NC_C-His_ N-T7_dm	123-471	pET-21a(+)	C-His	No	Yes	42.30852	BamHI/XhoI	LB	Yes	Not tested	T455E+ S296E	8.67
KD_123_414_NC_C-His_N-T7	123-414	pET-21a(+)	C-His	No	Yes	35.95449	BamHI/XhoI	LB	Yes	Not tested	WT	8.94
								AIM	Yes	Yes		
FL_1_471_C_C-His_N-T7	1-471	pET-21(3C)	C-His	Yes	Yes	56.60845/ 55.39014	BamHI/XhoI	AIM	Yes	No	WT	6.35/6.18
KL_123_471_C_C-His_N-T7	123-471	pET-21(3C)	C-His	Yes	Yes	43.36380/ 42.14550	BamHI/XhoI	AIM	Yes	No	WT	8.81/8.81
KD_123_414_C_C-His_N-T7	123-414	pET-21(3C)	C-His	Yes	Yes	37.07982/ 35.86152	BamHI/XhoI	LB	Yes	Not tested	WT	8.81/8.80
								AIM	Yes	Yes		
KL_123_471_C_C-His	123-471	pET-21(3C)	C-His	Yes	No	42.27258/ 41.05427	NheI/XhoI	AIM	Yes	No	WT	8.67/8.67
KL_123_471_NC_C-His	123-471	pET-21a(+)	C-His	No	No	41.14724	NheI/XhoI	LB	Yes	Yes	WT	8.81
								AIM	Yes	Yes		
KL_123_471_NC_C-His_sm	123-471	pET-21a(+)	C-His	No	No	41.17525	NheI/XhoI	LB	Yes	Not tested	T455E	8.67
KL_123_471_NC_C-His_dm	123-471	pET-21a(+)	C-His	No	No	41.21729	NheI/XhoI	AIM	Yes	No	T455E+ S296E	8.48
KN_136_460_C_C-His	136-460	pET-21a(+)	C-His	Yes	No	39.56262/ 38.34333	NheI/XhoI	LB	Yes	No	T455E+ S296E	8.68/8.82
FL_1_471_NC_N-His_λ	1-471	pETDuet-1	N-His	No	No	54.22267	Commerciall y acquired	LB	Yes	Yes	WT	6.43
FL_1_471_NC_N-His_λ_dm	1-471	pETDuet-1	N-His	No	No	54.29272	Commerciall y acquired	LB	Yes	Yes	T455E+ S296E	6.27
KDN_133_414_NC_C-His_λ	133-414	pETDuet-1	C-His	No	No	33.61701	NcoI/NotI	LB	Yes	Yes	WT	9.16
KDN_133_414_C_C-His_λ	133-414	pETDuet-1	C-His	Yes	No	34.74235/ 33.52404	NcoI/NotI	LB	Yes	Yes	WT	9.06/9.06

2.2.1. Gene and vector amplification

Chemically competent DH5α *E. coli* cells (Invitrogen) were transformed with an appropriate vector (Table 2.2). 100-300 µL of transformation reaction mixture was transfered to 10 mL LB-broth and grown overnight at 37°C, 180 rpm. The cells were harvested by centrifugation at 3739 g for 10 minutes at 4°C and DNA extracted using a GeneJET MiniPrep Kit according to the manufacturer's instructions.

To amplify the AtS6K2 genes and to insert phospho-mimetic mutations (serine/threonine residues mutated to a glutamic acid residue), PCR was performed using a TC-412 thermal cycler (Techne^{TM}) (Tables 2.3 and 2.4). Primers used in PCR are listed in the Table A1. Reaction products were analysed via 2% agarose gel electrophoresis (Section 2.2.5) and DNA was extracted using the NucleoSpin Gel and PCR Clean-up kit according to manufacturer's instructions.

Table 2.2| Bacterial transformation

Stage	Conditions	Length, comments
Thaw aliquot of competent cells (50 µL)	On ice	Gently mix when thawed
Add 1/5 µL of DNA	On ice	5 µL when using ligation reaction mixture
Incubation	On ice	30 min
Heat shock	42°C	45 s in a water bath
Recovery	On ice	2-5 min
Add LB media	Room temperature	200 µL for LB-agar plates and 500 µL for liquid cultures
Expand culture	37°C at 180-220 rpm	1 h

Table 2.3| PCR reaction components

Substance	Quantity (gene amplification)	Quantity (mutagenesis)
Template DNA	100 – 200 ng	15 – 20 ng
5x Phusion GC/HF buffer	1x	1x
dNTPs 10 mM	0.2 mM	0.2 mM
F/R primer	125 ng	75 ng
Phusion DNA pol.	2 U	2 U
DMSO	10%	Up to 6%*
H_2O	Make up to 50 µL	Make up to 30 µL

*Mutagenesis reactions were performed both with 6% DMSO and without it.

Table 2.4| PCR reaction steps

Step	Temperature (°C)	Length, min	Number of cycles
Initial denaturation	98	5^{P}, 2^{M}	1
Denaturation	98	1	25^{P}, 16^{M}
Annealing	55-60	1	25^{P}, 16^{M}
Elongation	72^{P}, 68^{M}	1^{P}, 12^{M}	25^{P}, 16^{M}
Final elongation	72	10, 0^{M}	1
Hold	4^{P}, 12^{M}	∞	

^{P} – gene amplification, ^{M} – mutagenesis.
Length of mutagenesis elongation (in minutes) is equal to 2x vector size in kb. The rest of conditions were adapted from QuikChange II Site-Directed Mutagenesis Kit (Agilent™).

2.2.2. Double digestion of DNA

Extracted insert DNA and empty vectors were digested with appropriate pairs of restriction endonucleases (Tables 2.1). The digestion reaction mixture was incubated at 37°C for either 1-3 hours or overnight. When using ThermoFisher restriction enzymes, the vector digestion mixture was also supplemented with 10U of Alkaline phosphatase (ThermoFisher Scientific) to prevent re-ligation. If enzymes from New England Biolabs (NEB) were used, an additional dephosphorylation step was performed. The digested vector was run through a 1% agarose gel (Section 2.2.5) and extracted using a Gel and PCR Clean-up Kit. Digested insert DNA was purified using the same kit via ‘PCR clean-up’ manufacturer’s protocol.

Table 2.5| Components of DNA double digestion reaction

Substance	Quantity
DNA	Up to 35 µL (7.5 µg)
Restriction enzymes	10 U of each
10X buffer	1/2X*
H_20	Make up to final volume of 50 µL
Alkaline phosphatase	10 U**

*2X Tango buffer was used for digestion with BamHI/XhoI, while reaction with NheI/XhoI was performed in 1X Tango buffer. NcoI/NotI reaction was performed in 1X NEBuffer r3.1
**AP only present in vector digestion mixture with ThermoFisher enzymes.

2.2.3 DNA ligation

Ligation reactions were performed by mixing 100 ng of digested, dephosphorylated vector DNA and 100-400 ng of digested insert DNA followed by incubation with T4 DNA Ligase (ThermoFisher Scientific) for 16-72h at room temperature (20°C) (Table 2.6).

Table 2.6| Components of DNA ligation reaction

Substance	Quantity
Vector DNA	100 ng
Insert DNA	100 – 400 ng
T4 ligase buffer (10X)	1X
T4 DNA ligase	5U
H_20	Make up to final volume of 20 - 30 µL*

Excess of digested insert DNA was used in order to increase the probability of a successful ligation reaction.
*Final volume of the reaction was adjusted according the concentration of digested DNA after the clean-up or gel extraction (Section 2.2.2)

2.2.4. Construct verification

5 µL of the reaction solution was used to transform DH5α competent cells (Table 2.2) that were plated onto LB-agar plates containing an appropriate antibiotic (Table 2.7) and grown overnight at 37°C. Selected colonies were expanded in 10 mL of LB media at 37°C, 180 rpm overnight. Plasmid DNA was extracted and purified using a GeneJET MiniPrep Kit and verified by automated Sanger sequencing (Eurofins). AtS6K2 constructs were then used for expression or amplified according to section 2.2.1.

Table 2.7| Antibiotic concentrations

Antibiotic	Final concentration	Expression constructs
Ampicillin	100 µg/mL	C-Histag, N-GST
Kanamycin	50 µg/mL	N-Histag
Spectinomycin	50 µg/mL	λ-phosphatase co-expression

2.2.5 Agarose gel electrophoresis

1-2% (w/v) agarose gels were prepared by dissolving agarose in in 1xTBE buffer. Samples were prepared by adding 5 µL GelRed™ Nucleic Acid Gel Stain (1:10000 dilution) and appropriate volume of agarose gel loading dye (6X) and run at 100V for 40-60 minutes against GeneRuler 1 kb DNA Ladder (Thermo Scientific). Gel analysis was performed under 470nm wavelength light in a Syngene LED Blue Light Transilluminator.

2.2.6. Site-directed mutagenesis

Mutagenesis was performed using a QuikChange II Site-Directed Mutagenesis Kit (Agilent Technologies) according to the manufacturer's instructions. Two pairs of primers for S296E and T455E phosphomimetic mutations were designed using an online Primer Design tool from Agilent (Table A1). PCR was used to synthesise the expression vector (Tables 2.3 and 2.4). Reaction products were then incubated with DpnI (10 U) restriction endonuclease at 37°C overnight to remove parental DNA. The remaining DNA was purified using the Gel and PCR Clean-up Kit. DH5α competent cells were transformed with the digested product (Table 2.2) and plated onto LB-agar plates containing appropriate antibiotic (Table 2.7). Selected colonies were grown in 10 mL of LB media overnight and the plasmids purified using a Thermo Scientific GeneJET Plasmid Purification Kit. Samples were then verified by Sanger sequencing.

2.2.7. pET-21(3C) plasmid production

A pair of complementary primers containing HRV 3C PreScission protease cleavage site (Table A1) were annealed together by heating the reaction mixture to 95°C and cooling it down to room temperature. Oligonucleotide duplex and pET-21a(+) vector was digested with XhoI at 37°C for 3 hours and ligated together using T4 DNA ligase at 20°C for 20 hours (Table 2.6). Orientation of the cleavage site insert was verified by Sanger sequencing. The XhoI restriction site at the 3' of the insert was mutated by site-directed mutagenesis (Tables 2.3, 2.4 and S1) and the sequence of pET-21(3C) confirmed by Sanger sequencing (Figure 2.1).

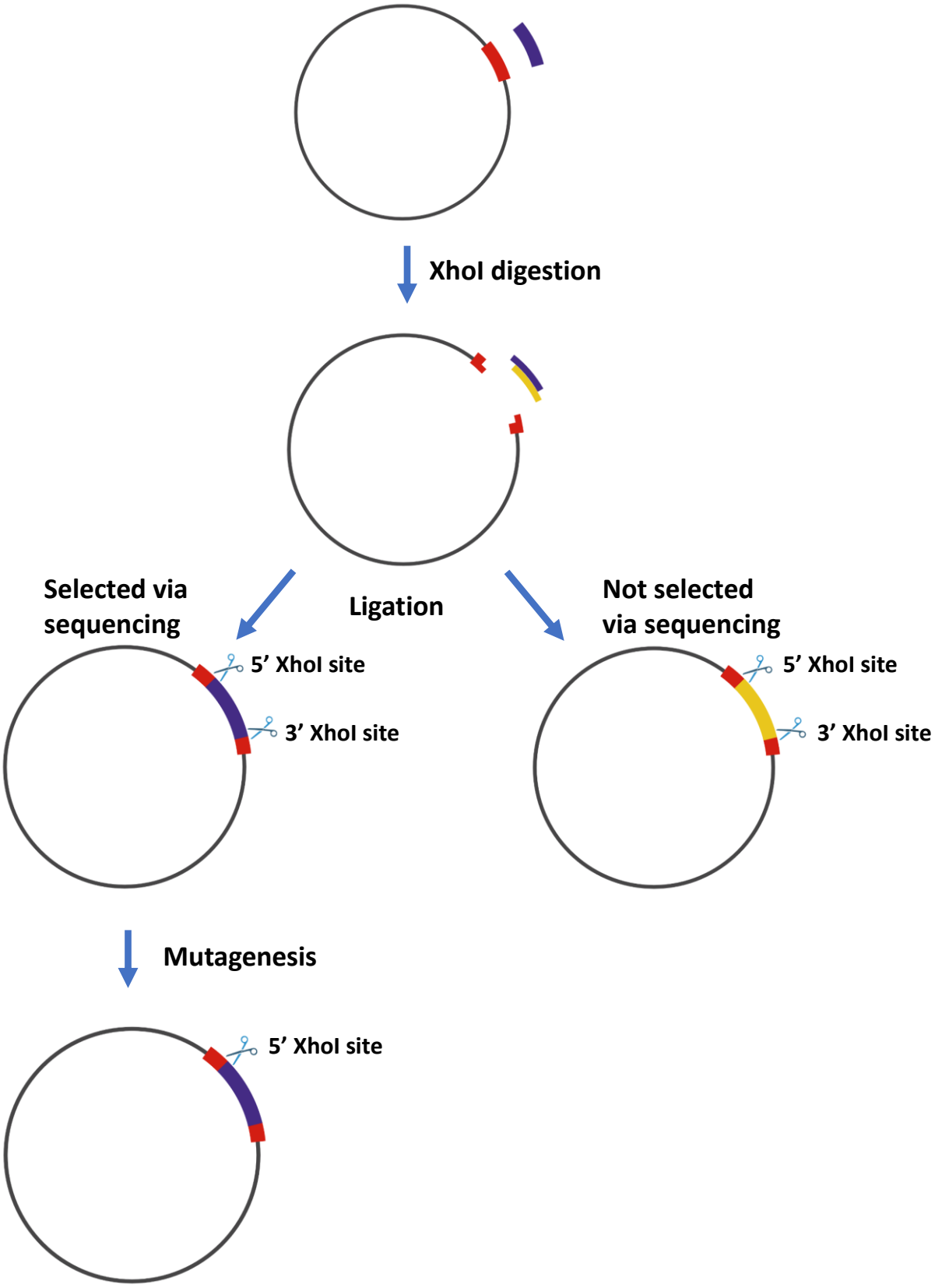


Figure 2.1| Cloning of the pET-21(3C) vector. Schematic representation of the insertion of the HRV 3C PreScission protease cleavage site (in blue) into a pET-21(a) vector.

2.3. Protein expression

2.3.1. Test expression

50 mL LB media pre-cultures were prepared in 250 mL Erlenmeyer Flasks and sterilised by autoclaving. *E. coli* expression strains i.e. One Shot BL21(DE3), Rosetta (DE3) and SHuffle T7, were transformed with appropriate construct (Table 2.2) and transferred to 50 ml LB supplemented with antibiotics (Table 2.7). Cells were grown until absorbance at 600 nm (OD_{600}) reached 0.5, 0.5 mL of pre-induced cell culture was collected and protein expression induced by addition of IPTG to 250 µM final concentration. Cells were incubated with shaking at 37°C, 180 rpm for 3 hours. 0.5 mL of cell culture was collected and protein expression analysed using SDS-PAGE (Section 2.4.5).

2.3.2. Large scale expression

BL21(DE3) OneShot (Invitrogen) competent cells were used for expression of AtS6K2 constructs. Pre-cultures were grown in 10 mL or 50 mL LB overnight at 37°C, 180 rpm. 2L Erlenmeyer flasks containing 1L of Auto-induction Media - AIM (FORMEDIUM) or LB media (MELFORD) were inoculated with 1% (v/v) of pre-culture. For expression in AIM, the cells were grown at 37°C, 180 rpm for 1-2 hours. The temperature was lowered to 18-20°C and expression continued overnight. For expression in LB, competent cells were grown at 37°C, 180 rpm until OD_{600} reached value of 0.5. The temperature was lowered to 18°C and protein expression was induced by 250 µM IPTG. Cell growth was continued for 16-20 hours. Cells were harvested by centrifuging at 4°C, 3739x*g* for 15 minutes, and the resulting cell pellet was collected and either stored at -80°C or lysed for purification.

2.4 Protein purification

AtS6K2 for structural studies was purified by a three-stage protocol: 1) Immobilized metal ion affinity chromatography (IMAC); 2) Ion exchange chromatography (IEC); 3) Size exclusion chromatography (SEC). AtS6K2 for activity measurements was purified by either IMAC and SEC or by IMAC only. The compositions of purification buffers for particular constructs are shown in tables A2-A12. β-mercaptoethanol, Deoxyribonuclease I from Bovine Pancreas (DNase I) and protease inhibitors (Table A13) were added to buffers immediately before use.

2.4.1. Cell lysis

Harvested cell pellets were resuspended in 2-6 mL of lysis buffer (Tables A2-A12) per gram of cell pellet and incubated on ice for 30-60 minutes with gentle stirring to resuspend the cells. The cells were then lysed using a high pressure homogeniser (Avestin Emulsiflex C3) by passing the solution through two times at a pressure of 15000 psi. The total cell lysate was centrifuged at 32931x*g* for 40-60 minutes at 4°C and the supernatant was collected for further purification.

2.4.2. Immobilised Metal Ion Affinity Chromatography

The supernatant was applied to an equilibrated affinity chromatography column (HisTrap FF crude 5 mL or HisTrap excel 1/5 mL or GSTrap 4B 5 mL) using a Peristaltic Pump P-1 (Cytiva). Wash buffers (5 mL/min) and elution buffer (Tables A2-A12) containing 300 mM imidazole or 10 mM L-glutathione (1-2 mL/min) were applied by using AKTÅ chromatography systems (PrimePlus, Pure or Purifier). Column elution fractions were collected and overnight dialysis at 4°C was performed in IEC buffer containing a low NaCl concentration (50-100 mM) to increase binding to the IEC column. Alternatively, elution fractions were concentrated to approximately 20% of initial volume at 3739x*g*, 4°C using a Macrosep Advance Centrifugal Device (Pall) with a molecular weight cut-off (MWCO) of 10 kDa and diluted back to the original volume using the IEC buffer A (Tables A2-A12). If needed, Human Rhinovirus 3C protease (PreScission protease) and/or λ-phosphatase was added to elution fractions and incubated at 4°C overnight. HRV 3C PreScission protease was a gift from Dr F. Prischi. Presence of target protein was verified via SDS-PAGE.

2.4.3. Ion Exchange Chromatography

After reducing the NaCl concentration, the IMAC elution was applied to a 5 mL High Performance Sulphopropyl (HiTrap™ HP SP) cation exchange column using a Peristaltic Pump P-1 at a flow rate of 1 mL/min. Further purification steps were performed using an AKTÅ chromatography system. The column was washed with 5 CV of 5% Buffer B and eluted using a linear gradient of NaCL or Li_2SO_4 over 16-20 CV at 2 mL/min (Tables A2-A12). Collected fractions were analysed with SDS-PAGE.

2.4.4. Size Exclusion Chromatography

Most SEC experiments were carried out using a HiLoad 16/600 Superdex 75 pg column (V=120 mL, GE Healthcare). HiLoad 16/600 Superdex 200 pg was also used for purification of KL_123_471_NC_C-His (Section 3.1.4.5.4). Analytical SEC was performed with a TSKgel G3000SW$_{XL}$ HPLC column. Purification runs were carried out using the AKTÅ systems at 1 mL/min (0.7 mL/min for analytical SEC). The columns were equilibrated with 1 CV of SEC buffer (Tables A2-A12) while the collected IE/IMAC fractions were concentrated in a filter unit with MWCO 10 kDa and loaded into the injection loop. Sample purity was verified using SDS-PAGE. Pure samples were concentrated to 2 mg/mL, flash-frozen in liquid nitorgen to prevent the formation of ice crystals and stored at -80°C.

2.4.5. SDS-PAGE

Protein samples were heated with 2x Laemmli sample buffer for 10-30 mins at 80-90°C prior loading to the gel. Electrophoresis was performed at 200V for 40-50 minutes until the lower fraction of the dye reached the bottom of the gel. 1) If using the Coomassie dye, the dye was applied directly on the gel and incubated for 1 hour on a shaker (Stuart See-Saw Rocker SSL4). Once the dye had been removed, the gel was rinsed with water a few times and destained in 10% methanol and acetic acid solution. 2) If using the SimplyBlue SafeStain dye, the gel was firstly rinsed thoroughly in water and boiled for 1-2 minutes with the dye. The gel was incubated with the dye for approximately 15 minutes and destained in water for at least 15 minutes.

2.4.6. Western-Blot

Protein samples were run on a SDS-PAGE and transferred onto a polyvinylidene fluoride membrane (GVS) that had been previously activated with methanol. Membrane pores were then blocked with PBS-T containing 5% milk for 30-60 minutes and washed with PBS-T buffer 3 times for 5 min per wash. Primary antibody in 5% milk was applied by rotating the membrane overnight (16-20 hours) at 4°C (Table A14). Membrane was then washed with PBS-T 3 times for 5 min per wash, before incubation with secondary antibody in PBS-T for 1 hour at room temperature (Table A14). The membrane was then washed with PBS-T 3 times for 5 min per wash, before applying 500 µL Luminata Forte Western Horseradish peroxidase substrate. Western blot membrane was visualised in a Fusion Fx (Vilber) machine.

2.4.7. Mass spectrometry

Protein samples were run on a SDS-PAGE. The gel was transferred to a clean petri dish in a biological safety cabinet to avoid contamination and destained in water. Appropriate bands were cut out using a sterile scalpel and transferred to 1.5 mL Eppendorfs in biological safety cabinet. The tubes were either stored at -20°C or sent straight for MS analysis.

Samples were digested by trypsin and analysed by the staff of the Functional Genomics and Proteomics Laboratories, the University of Birmingham (KL_123_471_NC_C-His_N-T7 construct) or the Yusuf Hamied Department of Chemistry, the University of Cambridge (KL_123_471_NC_C-His construct).

2.5. Kinase activity assay

The measurement of the activity of the AtS6K2 constructs was performed using the HTRF KinEASE-Serine / Threonine Kinase assay (HTRF, Cisbio Bioassays). In this assay, the phosphorylation of the substrate (peptides S1, S2 and S3 with unknown proprietary sequence recognised by any kinase) by AtS6K2 is detected by a phospho-specific monoclonal antibody labelled with Eu^{3+}-Cryptate. Peptides are all biotinylated and are recognized by a streptavidin-XL665 conjugate. In presence of phosphorylation, the Eu^{3+}-Cryptate from the antibody is in close proximity with the XL665-streptavidin, resulting in a fluorescence resonance energy transfer (FRET) signal at 665 nm. This system also incorporates its own internal control in that Eu^{3+}-Cryptate produces an emission peak at 620 nm. The ratio of fluorescence at 665 to 620 can be used to calculate the concentration of Eu3+-Cryptate-XL665 complexes. A CLARIOstar *Plus* (BMG LABTECH) plate reader was used to measure emission at 665 nm and 620 nm. Plate reader parameters were set up according to manufacturer's recommendations for the HTRF KinEASE assay. The protein activity was then determined by plotting the 665/620 ratio versus time. The assays were performed in AtS6K2 storage buffers: (i) KL_123_471_NC_C-His_N-T7 – 25 mM Tris-HCl buffer (pH 7.5) containing 250 mM NaCl, 5 mM β-mercaptoethanol, 5% (v/v) glycerol, 200 µM $MgCl_2$, 100 µM ATP; (ii) FL_1_471_NC_C-His_N-T7 - 25 mM Tris-HCl buffer (pH 8.0 for wild type StS6K2 and pH 7.5 for T455E, S296E mutant) containing 150 mM NaCl, 5 mM β-mercaptoethanol, 5% (v/v) glycerol, 10 mM $MgCl_2$, 2 mM ATP. AtS6K2 KL_123_471_NC_C-His_N-T7 concentrations from 2.7 nM to 0.47 µM were used. Full length AtS6K2 was assayed at 5 µM (both wild type and mutant), while the mutant was also tested at 10 µM and 20 µM.

2.6. Protein crystallization

Purified protein samples (>90-95% homogeneity) were screened for crystallisation conditions using sparse matrix kits (Table A15) that allow rapid and systematic testing of large numbers of crystallisation solutions. Conditions where crystal growth was detected were further optimised by a large volume crystallisation that increases quality and reproducibility of crystal growth.

2.6.1 Crystallization conditions' screening

Crystallisation screening trials were set up in 96 well sitting drop plates (Hampton Research). Protein samples were supplemented with 1 mM adenylyl-imidodiphosphate (AMP-PNP) and 2 mM $MgCl_2$ prior to concentrating. The protein concentrations of specific constructs are shown in tables A16-A18. Droplets of different protein to reservoir ratio (1:1 and 1:2) were dispensed using a Crystal Gryphon liquid handling robot. Plates were immediately sealed with Crystal Clear Sealing Tape (Hampton Research) to avoid sample evaporation and incubated at 4°C, 16°C or 18°C (Tables A16-A18).

2.6.2. Hanging Drop Vapour Diffusion Crystallisation

Hanging drop vapour diffusion crystallisation experiments were set up in 24-well plates (Hampton Research). 1 µL of 10 mg/mL AtS6K2 KD_123_414_NC_C-His_N-T7 protein was mixed with 1 µL of reservoir solution on a coverslip and was sealed with 1 mL reservoir. Similarly to the sitting drop experiments, protein solution was supplemented with 1 mM AMP-PNP and 2 mM $MgCl_2$. All the equipment (plate, buffers, coverslip, pipette tips) was pre cooled to 4°C and the drops prepared in a cold room environment (4°C). PACT C11 condition (Section 3.1.5.2) containing HEPES pH=7.0 was expanded by varying calcium chloride dihydrate (0.18M, 0.2M, 0.22M) and PEG 6000 (16%, 18%, 20%, 22%) concentrations. In addition, an ammonium sulphate screen was set up in a similar manner.

3. AtS6K2 expression, purification and crystallisation

3.1. Results

The catalytic domain of protein kinases is highly conserved and consists of two lobes, the N-lobe and the C-lobe. Each lobe contains key structural elements that are essential for the catalytic activity of a kinase. The most notable of which are the C-helix and the β4 strand located in the N-lobe and the HRD and DFG motifs in the C-lobe. These elements orientate ATP and substrate molecules for an effective transfer of the γ-phosphate (Meharena *et al.*, 2013).

Multiple AGC kinases, including *Arabidopsis* S6K2 (AtS6K2), are activated by step-wise phosphorylations of two highly conserved motifs. Firstly, the hydrophobic motif (HM), located in the non-catalytic AGC C-terminal region, is phosphorylated. Then the activation loop in the kinase domain is phosphorylated, activating the kinase. The activation of several AGC kinases also includes the phosphorylation of the turn motif (TM), which is required for kinase stability and may also protect the HM site from dephosphorylation (Pearce, Komander and Alessi, 2010). Activation of plant S6Ks includes phosphorylation of all three motifs. However, phosphorylation of only two out of the three sites is necessary to activate the kinases (Yaguchi, Ikeya and Kozaki, 2020).

Initial evidence suggest that in plants, like in humans (Pardo and Seckl, 2013), AtS6K1 and AtS6K2 have different roles in plant homeostasis, despite a high level of sequence conservation (Pardo and Seckl, 2013; Obomighie *et al.*, 2021). In order to rationalise the different cellular functions of the S6K isoforms, high quality structural data are essential. This is currently very challenging (if not impossible), as only the structure of

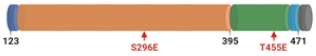
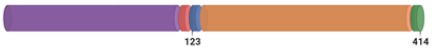
the human S6K1 has been solved (Wang *et al.*, 2013). Although AlphaFold predicts the catalytic domain of AtS6K2 with a similar architecture to human AGC kinase (Endicott, Noble and Johnson, 2012), this structure needs to be confirmed, as no structural data is available for any AGC kinase in plants. Unfortunately, the model doesn't provide any information for the non-conserved N-terminal domain and C-terminal region, suggested to have a key role in isoform-specific substrate selection and activation, respectively (Pardo and Seckl, 2013; Obomighie *et al.*, 2021).

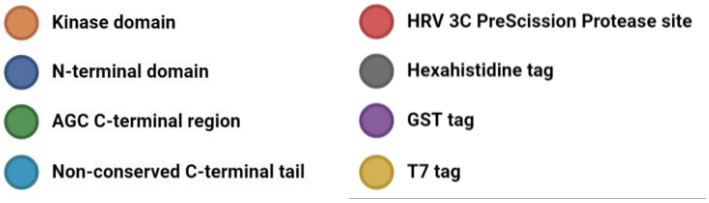
To perform structural and functional characterization of AtS6K2 a series of experiments were carried out: (I) purification and activity assay of full-length AtS6K2; (II) purification and crystallization of the AtS6K2 kinase domain with C-terminal region; (III) purification and crystallization of the AtS6K2 kinase domain; (IV) molecular dynamics (MD) simulations with principal component analysis (PCA) of the S6K kinase domain.

3.1.1. An overview of AtS6K2 constructs employed

No expression and purification protocol is available for plant (nor human) S6K2. In order to produce soluble AtS6K2 for crystallisation studies, several factors were taken into account, including the construct length, type of a tag and position of a tag, vector used for recombinant expression, expression temperature, host strain, expression media and duration of expression (Tables 2.1 and 3.1). Results chapters have been organised in chronological order and not according to construct type.

Table 3.1| Summary of AtS6K2 constructs. The 23 tested protein expression vectors have been grouped into types according to AtS6K2 residues covered: full-length (1-471), kinase domain-long (123-471), kinase domain (123-414), kinase new (136-460) and kinase domain-new (133-414). Residue numbers (based on UniProt ID:Q39030 domains' definition) and, where present, phosphomimetic mutations (S296E, T455E) are shown below the constructs' graphical representations.

Construct type	Name	Section	Graphical depiction
Full-length	FL_1_471_C_N-His	3.1.3.3	
	FL_1_471_NC_C-His_N-T7	4.1.1	
	FL_1_471_C_C-His_N-T7	3.1.4.2	
	FL_1_471_NC_N-His_λ	4.1.2	
Kinase domain-long	KL_123_471_C_N-His	3.1.3.1 3.1.4.5.1	
	KL_123_471_C_N-GST	3.1.2.1 3.1.4.5.2	
	KL_123_471_NC_C-His_N-T7	3.1.4.1 3.1.4.4	
	KL_123_471_C_C-His_N-T7	3.1.4.2	
	KL_123_471_C_C-His	3.1.4.3	
	KL_123_471_NC_C-His	3.1.4.5	
Kinase domain	KD_123_414_C_N-His	3.1.3.2	
	KD_123_414_C_N-GST	3.1.5.1	
	KD_123_414_NC_C-His_N-T7	3.1.5.2	
	KD_123_414_C_C-His_N-T7	3.1.5.3	
Kinase New	KN_136_460_C_C-His	3.1.6	
Kinase domain-New	KDN_133_414_NC_C-His_λ	3.1.7.1	
	KDN_133_414_C_C-His_λ	3.1.7.2	



3.1.2. GST tagged AtS6K2

As a starting point for recombinant protein expression, Glutathione S-transferase (GST) tagged constructs were designed, as GST has been demonstrated to increase the yield of soluble proteins from a variety of eukaryotic organisms, including plants (Smith and Johnson, 1988; Shih *et al.*, 2002). The pGEX-6P-1 vector, which has an N-terminal GST-tag and a HRV 3C PreScission protease cleavage site, was selected for the expression of the KL and KD constructs, hereafter KL_123_471_C_N-GST and KD_123_414_C_N-GST (Table 3.1). All the constructs were codon optimised for expression in *E. coli.* The KL sequence was designed based on the sequence alignment of the crystallised human S6K1-HM construct, that includes the kinase domain plus the hydrophobic motif (encompassing residues 52–394, PDB ID 4L46) (Wang *et al.*, 2013), with the full length AtS6K2 (Fig. 3.1). Despite the relatively low sequence similarity between human S6K1 and AtS6K2 (59%), particularly evident for the AGC C-term region, key conserved kinase motifs (i.e., Y/HRD, DFG, APE, S on the activation loop and T on the hydrophobic motif) were correctly aligned which allowed the design of the AtS6K2 KL and KD constructs based on the corresponding crystallised human S6K1 constructs.

In the designing of the purification conditions, several different protocols used for the purification of evolutionary related kinases, i.e., S6K1, RSK1-NTKD, RSK2-NTKD, RSK4-NTKD, were compared. The solubility of a protein in water-based solutions is dependent on its isoelectric point (pI). A protein will be the least soluble in a buffer with a pH that is equal to its pI (Arakawa and Timasheff, 1985). Accordingly, the ProtParam tool (Gasteiger *et al.*, 2005) was used to calculate the pI and choose the buffer pH that was at least 0.5 pH units different from the pI. Different salts have been used for different proteins, but, due to the lack of a rationale in different publications for using specific salts (Section 5), NaCl was used in all purification steps, unless specified. Indeed, several papers have shown that

ribosomal kinases have low solubility in low salt conditions (Malakhova *et al.*, 2009; Chrysostomou *et al.*, 2021). A stepwise approach based on the RSK4-NTKD purification protocol (Chrysostomou *et al.*, 2021) was adopted, which includes 3 steps: immobilized metal ion affinity chromatography (IMAC), ion exchange chromatography (IEC) and size exclusion chromatography (SEC).

Initially, the focus was on the purification of the KL constructs that would provide the most valuable structural information, as it would allow insights to the role of the AGC C-term region in the kinase domain activation and regulation.

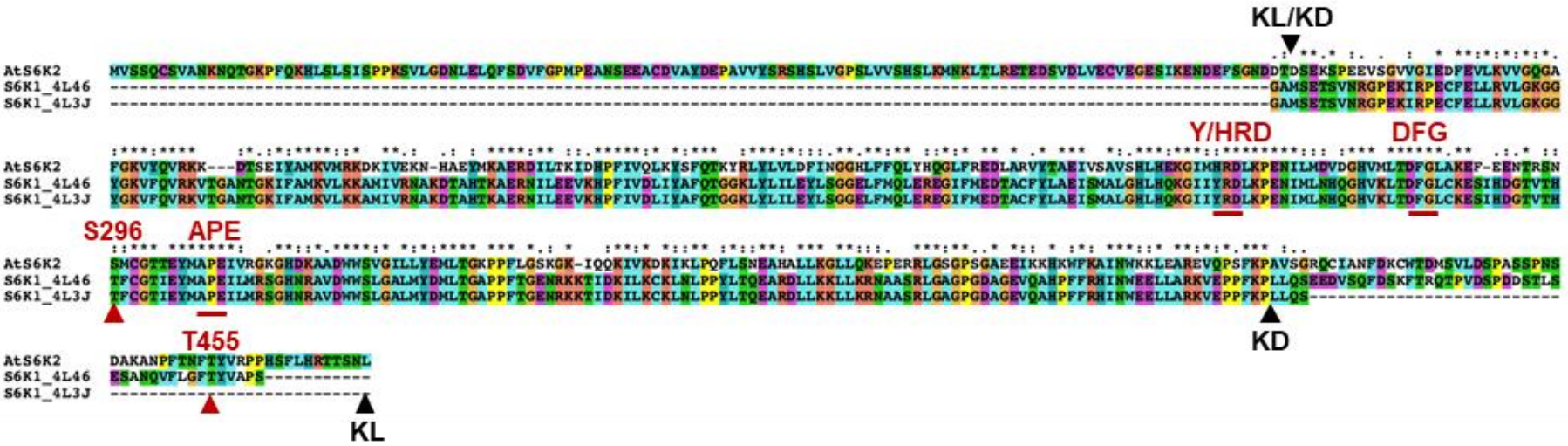


Fig. 3.1| Sequence alignment of two crystallised human S6K1 constructs and AtS6K2. Start/End of the Kinase Domain with AGC C-term region have been highlighted with black arrows; AtS6K2 residues S296 and T455 with red arrows; Y/HRD, DFG and APE motifs with red lines.

3.1.2.1. AtS6K2 KL_123_471_C_N-GST

To begin with, a small scale test expression of the N-terminal GST-tagged AtS6K2 KL (KL_123_471_C_N-GST) (MW=66.6 kDa) in LB media was carried out using three different *E. coli* strains: BL21(DE3), Rosetta (DE3) and K12 SHuffle. The BL21(DE3) strain has been constructed for inducible high-level expression of recombinant proteins (Studier and Moffatt, 1986). Rosetta(DE3) strain is a BL21 derivative that contains human codons rarely used in *E. coli* and has been shown to increase expression yield and protein purity (Tegel *et al.*, 2010). K12 SHuffle cells overexpress the disulphide bond isomerase DsbC that helps correct formation of disulphide bonds and prevents misfolding of recombinantly expressed AtS6K2 (Ren, Ke and Berkmen, 2016). SDS-PAGE analysis of expression samples after induction from all three strains contain a strong double band between 60 and 70 kDa, suggesting successful expression of KL_123_471_C_N-GST (Fig. 3.2A).

The expression of the KL_123_471_C_N-GST using BL21(DE3) and Auto-Induction media (AIM) (Studier, 2005, 2014) was then scaled up. A weaker band between 60-70 kDa is present in the cell lysate, the supernatant recovered from the centrifuged cell lysate and the flow-through of the GST affinity column, however absent from elution, suggesting that KL_123_471_C_N-GST is not expressed or insoluble and the band observed in the SDS-PAGE at 60-70 kDa is a contaminant (Fig. 3.2B). This suggests that the GST-tag may not be suitable for KL expression, hence an N-terminal hexahistidine tag was tested.

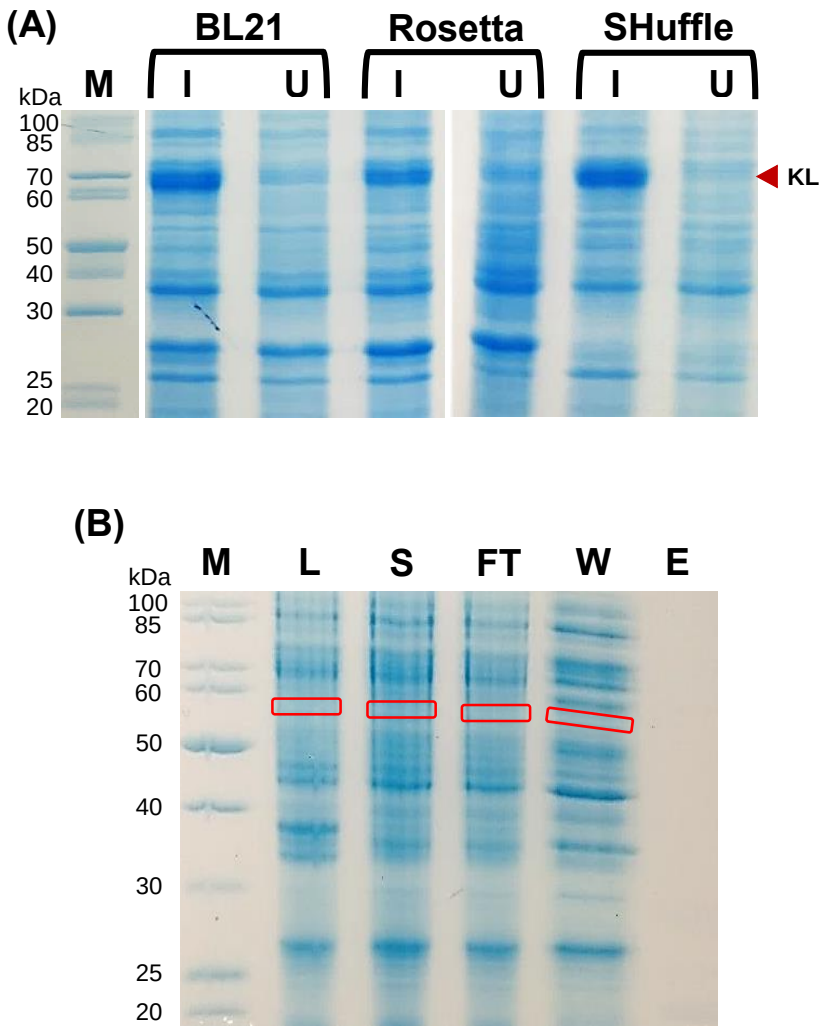


Figure 3.2| Test expression and GST affinity purification of AtS6K2 KL_123_471_C_N-GST. **(A)** Small scale test expression in different cell lines, *E. coli* BL21, Rosetta and Shuffle, showing IPTG induced (I) and uninduced (U) samples. The apparent weight of AtS6K2 KL_123_471_C_N-GST (MW=66.6 kDa) is highlighted with a red arrow on the side of the gel. **(B)** Samples from different affinity purification steps were analysed using SDS-PAGE. M – Protein ladder; L – Cell lysate; S – Supernatant of centrifuged cell lysate; FT – GST column Flow-through; W – GST column wash; E – Elution.

3.1.3. N-His tagged AtS6K2

To understand if the lack of reproducibility between test and large scale expressions of KL_123_471_C_N-GST was linked to the use of the pGEX-6P-1, AtS6K2 was cloned into a pETM-14 vector which has an N-terminal hexa-histidine tag (His-tag) and a HRV 3C PreScission protease cleavage site. Full-length (FL) AtS6K2, KL and KD sequences were inserted into the pETM-14 generating FL_1_471_C_N-His, KL_123_471_C_N-His and KD_123_414_C_N-His (Table 3.1).

3.1.3.1. AtS6K2 KL_123_471_C_N-His

Similarly to GST-tagged AtS6K2, I focused my initial efforts on the KL construct. Expression of KL_123_471_C_N-His (MW=42 kDa) was carried out in LB and AIM media. In both cases it was possible to detect a band between 40 and 50 kDa, compatible with KL_123_471_C_N-His, in the cell lysate and the soluble fraction, suggesting successful expression of soluble protein (Fig. 3.3A). However, when expressed again in the AIM media, the protein band was absent from the elution fraction of IMAC (Fig. 3.3B). To investigate if KL_123_471_C_N-His is misfolded, preventing affinity binding to the resin, a western blot analysis was carried out (Fig. 3.3C). This revealed that the protein is only present in the lysate, confirming that the protein is expressed but the construct used is insoluble. Although the KL_123_471_C_N-His is present in the supernatant when expressed in LB (Fig. 3.3A), there is also a similarly strong *E. coli* contaminant band slightly above it. It is likely that KL_123_471_C_N-His expressed in LB is equally as insoluble as KL_123_471_C_N-His expressed in AIM. Hence, instead of testing KL_123_471_C_N-His expression and solubility in LB media, I focused on the constructs with the C-terminal His-tag, as it has been shown previously that moving the His-tag from the N-term to the C-term can improve solubility by preventing potential disruption of the protein folding caused by the N-terminal His-tag (Eschenfeldt *et al.*, 2010). In addition, a C-terminal His-tag has been demonstrated to increase protein production levels and yield of purified protein compared to the N-terminal His-tag (Woestenenk, Hammarström, Van Den Berg, *et al.*, 2004).

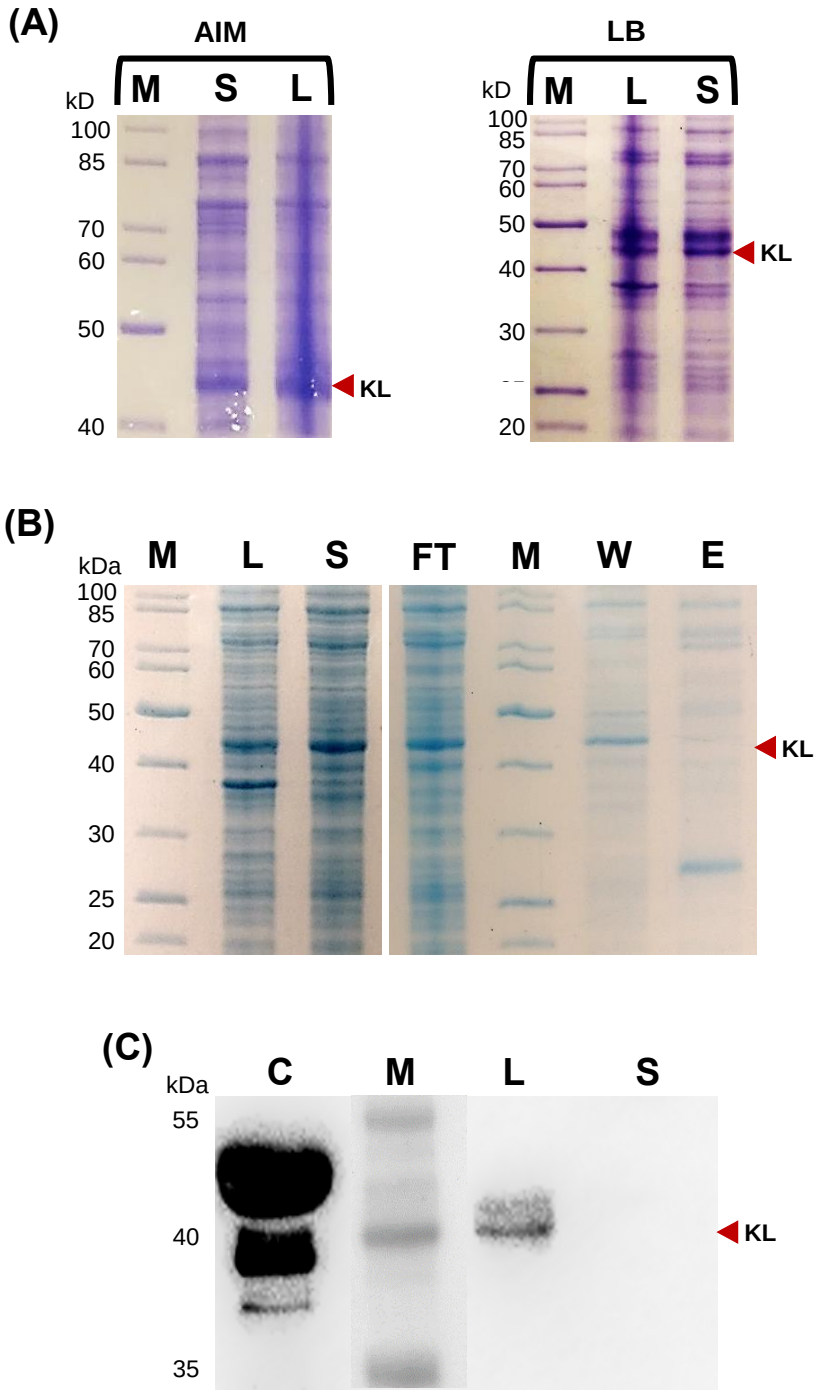


Figure 3.3| Expression and purification of AtS6K2 KL_123_471_C_N-His. **(A)** Small scale test expression in BL21 grown in either auto-induction media (AIM) or LB, showing cell lysate (L) and Supernatant of centrifuged cell lysate (S). AtS6K2 KL_123_471_C_N-His (42 kDa) is highlighted with a red arrow on the side of the gel. **(B)** Samples from different IMAC purification steps were analysed using SDS-PAGE. L – Cell lysate; S – Supernatant of centrifuged cell lysate; FT – HisTrap column flow-through; W – HisTrap column wash; E – Elution. **(C)** Expression in BL21 grown in AIM, showing the western-blot of cell lysate (L) and supernatant of centrifuged cell lysate (S). A positive HisTag control (C, N-hexahistidine-tagged YB-1 protein, received from F. Prischi) was included. M – Protein ladder.

3.1.3.2. AtS6K2 KD_123_414_C_N-His

To check if the AGC C-terminal region may be preventing expression of soluble AtS6K2, which may be likely due to the high content of hydrophobic residues in the AGC C-term, the catalytic domain construct, KD_123_414_C_N-His (35.7 kDa), was tested (Table 3.1). A small scale test expression was performed in LB media. Samples prior and after IPTG induction were taken but no overexpression band was present in both, suggesting that KD_123_414_C_N-His was not expressed (Fig. 3.4).

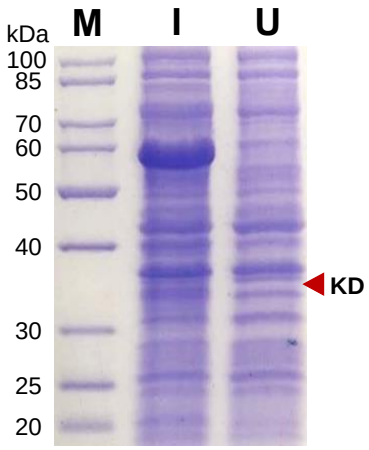


Figure 3.4| Test expression of AtS6K2 KD_123_414_C_N-His. Small scale test expression of AtS6K2 KD_123_414_C_N-His (MW=35.7 kDa) in BL21 grown in LB showing IPTG induced (I) and uninduced (U) samples. M – Protein ladder

3.1.3.3. AtS6K2 FL_1_471_C_N-His

AtS6K2 FL has a long N-term region which is intrinsically unfolded and is likely to affect protein solubility. Despite this, to test if the design of the KD and KL constructs affected the protein solubility, the full length AtS6K2 (FL_1_471_C_N-His) expression was tested (Table 3.1). FL_1_471_C_N-His (55.2 kDa) expression in AIM was not successful and no band of expected size was detected in any of the IMAC fractions (Fig. 3.5).

Taken together the results from the affinity purifications of the N-GST and N-His tagged fusion proteins suggest that the tags used don't improve protein solubility. Considering that the location of the His-tag might affect expression and solubility of the recombinant protein (Dan, Balachandran and Lin, 2009), the tag was moved to the C-terminus.

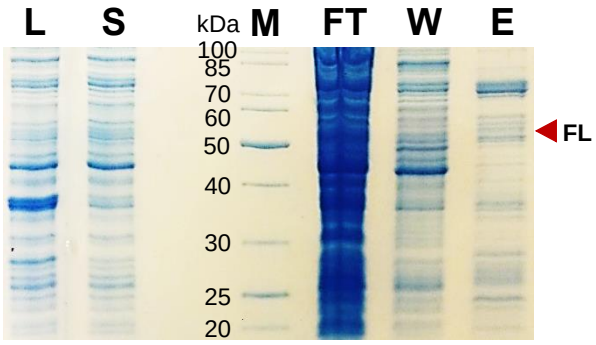


Figure 3.5| IMAC purification of AtS6K2 FL_1_471_C_N-His. Samples from different IMAC purification steps of AtS6K2 FL_1_471_C_N-His (55.2 kDa) were analysed using SDS-PAGE. M – Protein ladder; L – Cell lysate; S – Supernatant of centrifuged cell lysate; FT – HisTrap column flow-through; W – HisTrap column wash; E – Elution.

3.1.4. C-His tagged AtS6K2

To change the position of the affinity purification tag, KL, KD and FL were cloned into a pET-21a vector, to generate the KL_123_471_NC_C-His_N-T7, KD_123_414_NC_C-His_N-T7 and FL_1_471_NC_C-His_N-T7 constructs possessing an N-terminal T7 epitope tag (Enomoto, Chen and Berman, 1998) and a C-terminal His-tag. Unfortunately, the T7-tag is a relatively long flexible tag which antagonises the strategy to remove tag from the protein N-terminus. The constructs were then re-cloned to remove the T7-tag, and produce the KL_123_471_NC_C-His construct. In parallel, the pET-21a vector was modified by inserting the HRV 3C PreScission protease cleavage site at the C-terminus, creating the pET-21(3C) vector. Similarly, pET-21(3C) was used to clone the KL_123_471_C_C-His_N-T7, KD_123_414_C_C-His_N-T7 and FL_1_471_C_C-His_N-T7 constructs that contain an N-terminal T7-tag, a C-terminal His-tag and a HRV 3C PreScission protease cleavage site between the construct and the His-tag. KL was also re-cloned without the T7-tag resulting in the KL_123_471_C_C-His construct (Table 3.1).

3.1.4.1. AtS6K2 KL_123_471_NC_C-His_N-T7 expression, purification and activity assay

Similarly to N-terminal GST and His-tag constructs (Sections 3.1.2 and 3.1.3.1), I started from AtS6K2 KL_123_471_NC_C-His_N-T7 (42.2 kDa; Table 3.1), which was successfully expressed in BL21(DE3) cells using both LB and AIM media. A strong overexpression band of 47 kDa was detected in the cell lysate, the soluble fraction and the IMAC column eluate suggesting that KL_123_471_NC_C-His_N-T7 successfully bound to the HisTrap column and eluted at 300 mM imidazole (Table A5; Fig 3.6A). The elution fraction contained a high amount of KL_123_471_NC_C-His_N-T7 with some contaminants. The AtS6K2 band in the SDS-PAGE gel was running approximately 5 kDa

higher than the expected size of 42.2 kDa. This may have been caused by the buffer pH and/or the presence of salts, which can affect protein migration. The protein sample was further purified using cation exchange chromatography. A high intensity peak appeared between 350 and 400 mM NaCl (Fig. A1). Ion exchange chromatography (IEC) fractions that contained KL_123_471_NC_C-His_N-T7 (Fig. 3.6B) were pooled together, concentrated and then loaded onto a gel filtration column (HiLoad 16/600 Superdex 75 pg). Three peaks were visible in the chromarogram (Fig. 3.6D): the void containing unfolded protein and/or aggregates, and two partially overlapped peaks. AtS6K2 KL_123_471_NC_C-His_N-T7 was present in the last peak around 62.5 mL (Fig. 3.6C). Fractions containing the protein were pooled together, concentrated to 2 mg/mL, flash-frozen using liquid nitrogen and stored at -80°C.

In addition to wild type KL_123_471_NC_C-His_N-T7, I generated the T455 and S296 phosphomimetic mutants by replacing the S/T residues with glutamic acid to produce single and double mutants of AtS6K2 KL_123_471_NC_C-His_N-T7, aiming to produce constitutively active AtS6K2 kinase (Yaguchi, Ikeya and Kozaki, 2020). Specifically, T455E allows the binding of the HM to the phosphate-binding pocket on the N-lobe of the kinase and promotes the folding of the AGC C-terminus domain, while the S296E mutation should induce interaction between the E296 on the activation loop and the HRD motif in the kinase domain; this should cause flexible-ordered conformational transition, stabilising the kinase and improving its solubility (Endicott *et al.*, 2012; Nishi *et al.*, 2014; Chrysostomou *et al.*, 2021). Small scale test expression experiments demonstrated that both single and double (42.3 kDa) KL_123_471_NC_C-His_N-T7 mutants expressed in LB media and strong bands (~40 kDa) in the cell lysates were visible (Fig. 3.7). Unfortunately, similarly to all KL constructs containing phosphomimetic mutations, KL_123_471_NC_C-His_N-T7 mutants showed limited solubility (Section 3.1.4.5.3).

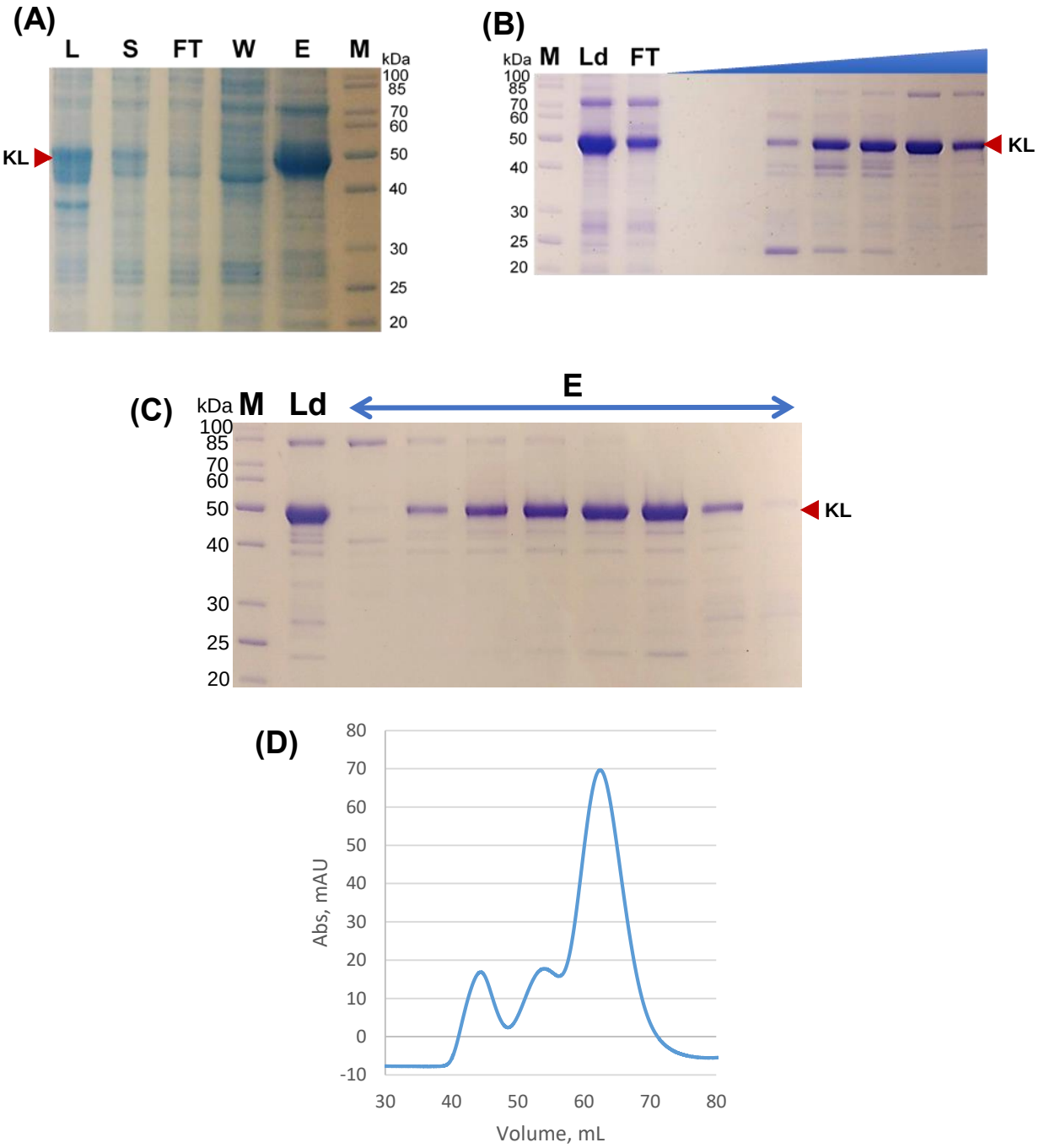


Figure 3.6| Purification of AtS6K2 KL_123_471_NC_C-His_N-T7. **(A)** Samples of AtS6K2 KL_123_471_NC_C-His_N-T7 (42.2 kDa) from different IMAC purification steps. M – Protein ladder; L – Cell lysate; S – Supernatant of centrifuged cell lysate; FT – HisTrap column Flow-through; W – HisTrap column wash; E – Elution. **(B)** Samples from ion exchange chromatography (IEC). Blue triangle represents the NaCl gradient, and samples eluting between 300 mM and 400 mM NaCl were loaded on the SDS-PAGE. Ld – sample loaded on the cation exchange column; FT – cation exchange column flow-through. **(C)** Samples from size exclusion chromatography (SEC, HiLoad 16/600 Superdex 75 pg column, V=120 mL). Ld – sample loaded on the SEC column. E – samples from consecutive 2 ml fractions eluting in the range 55-70 ml were loaded on the SDS-PAGE. AtS6K2 KL_123_471_NC_C-His_N-T7 is highlighted with red arrows on the sides of the gels. **(D)** Chromatogram of SEC, showing the AtS6K2 KL_123_471_NC_C-His_N-T7 elution peak.

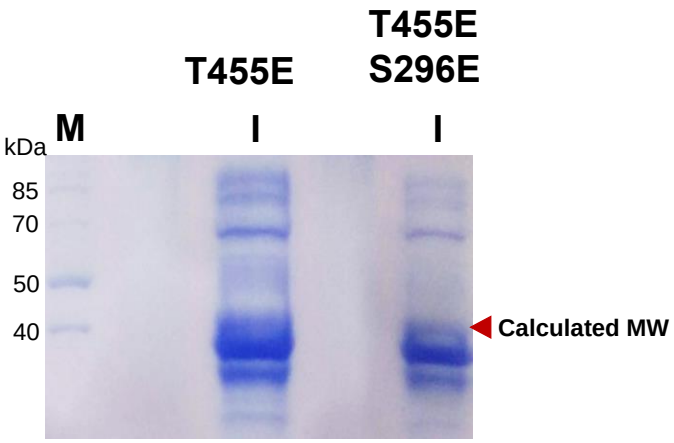


Figure 3.7| Test expression of AtS6K2 KL_123_471_NC_C-His_N-T7 T455E and T455E_S296E mutants. Small scale test expression of AtS6K2 KL_123_471_NC_C-His_N-T7 mutants (MW=42.3 kDa) in BL21 grown in LB showing IPTG induced (I) samples. Calculated size (ProtParam) of AtS6K2 KL_123_471_NC_C-His_N-T7 mutants is highlighted with a red arrow on the side of the gel.

To confirm that the purified wild-type AtS6K2 KL_123_471_NC_C-His_N-T7 has no enzymatic activity (differently from the mutants I tried to produce), the HTRF KinEASE STK kit was used (Section 2.5). The kit employs three universal peptidic substrates conjugated with biotin that are recognised and phosphorylated by the target kinase. Detection reagents include a phospho-specific monoclonal antibody labeled with Eu^{3+}-cryptate (donor fluorophore) that recognises the phosphorylated substrate and an allophycocyanin acceptor fluorophore XL665 conjugated with streptavidin. Only when the substrate is phosphorylated the antibody can bind, placing donor and acceptor molecules in close proximity. Excitation at 337 nm induce Eu^{3+}-cryptate emission at 620 nm, which in turn excites the XL665 acceptor, that emits light at 665 nm over millisecond scale, eliminating short-lived background fluorescence (Degorce *et al.*, 2009). Generated time-resolved fluorescence energy transfer (TR-FRET) signal is proportional to the phosphorylation level and is quantified by acceptor and donor (665 nm/620 nm) emission ratio. Three independent assays were carried out with a range of KL_123_471_NC_C-His_N-T7 concentrations (2.7 nM - 0.47 µM) and with a constant concentration (2 µM) of the three available substrates. The kinase was supplemented with saturating concentrations of ATP and $MgCl_2$ and incubated for 30 minutes at either 37°C or room temperature (20°C), and as expected no activity was detected. RSK4 C-terminal kinase domain was used as a positive control (gift from Dr F. Prischi). More work on phosphomimetic mutants was carried out using FL constructs (Section 4.1).

3.1.4.2. AtS6K2 KL_123_471_C_C-His_N-T7 and FL_1_471_C_C-His_N-T7

Any tag, independent of its size, might disrupt a protein's biological activity, crystallisation or influence its behaviour in some other way (Waugh, 2005). Although an uncleavable C-terminal His tag may not prevent the crystallization of AtS6K2, it usually increases the B-factor of the atoms in the crystal lattice (anharmonic vibrations of atoms around their equilibrium positions), making crystallization more challenging (Carson *et al.*, 2007). Thereby, the pET-21a vector was modified by inserting an HRV 3C PreScission protease cleavage site before the His-tag and generated the pET-21(3C) plasmid (Fig. 2.1). Firstly, pET-21a was cut open with XhoI restriction endonuclease. Single stranded oligonucleotides that contain the HRV 3C PreScission protease cleavage site (Table A1) were annealed together to form a duplex with blunt ends, which was digested with XhoI to create sticky ends. Then the digested vector was ligated with the oligoduplex and sequenced to confirm the correct positioning of the protease cleavage site, as due to identical sticky ends at both 5' and 3' the oligoduplex can ligate in opposing ways. This resulted in a plasmid with two XhoI digestion sites: one on both 5' and 3' of the inserted protease cleavage site. Finally, a single nucleotide in the the XhoI cleavage site 3' of the inserted protease cleavage sequence (CTCGAG to CTCCAG) was mutated using site-specific mutagenesis (Table A1) and the resulting plasmid was sequenced to confirm successful production of pET-21(3C). This was necessary as XhoI is an indispensable restriction enzymes used for cloning of many AtS6K2 constructs.

In parallel to the expression of KL_123_471_NC_C-His_N-T7, the KL_123_471_C_C-His_N-T7 and FL_1_471_C_C-His_N-T7 constructs that contain the N-terminal T7-tag and a cleavable C-terminal His-tag were tested (Table 3.1). Both KL_123_471_C_C-His_N-T7 (43.4 kDa) and FL_1_471_C_C-His_N-T7 (56.6 kDa) were successfully expressed in AIM as strong overexpression bands were present in their cell

lysates ≈40 kDa and ≈60 kDa respectively. Unfortunately, the bands were hardly visible in the soluble fraction (Fig. 3.8A and 3.8C) suggesting that both constructs were expressed in inclusion bodies. This was supported by SDS-PAGE analysis of affinity chromatography samples of KL_123_471_C_C-His_N-T7 which revealed that the weak band is not AtS6K2 as it was missing from the column elution (Fig. 3.8B). Cell lysis in reducing conditions with 5 mM β-mercaptoethanol (Table A19) did not improve solubility of the protein.

3.1.4.3. AtS6K2 KL_123_471_C_C-His

To check if the N-terminal T7-tag affects the solubility of KL_123_471_C_C-His_N-T7, a KL_123_471_C_C-His construct with a cleavable His-tag but without the T7 tag was tested (Table 3.1). Identically to KL_123_471_C_C-His_N-T7, KL_123_471_C_C-His (42.3 kDa) was expressed in AIM which resulted in insoluble protein. Cells were lysed in both reducing and non-reducing conditions (Table A19) and a strong band just above 40 kDa was only present in the cell lysate. The band was absent from the elution fraction of IMAC (Fig. 3.9). All the attempts to purify AtS6K2 KL construct with a cleaved His-tag were unsuccessful, despite the fact that sequencing did not show any issues with the plasmid.

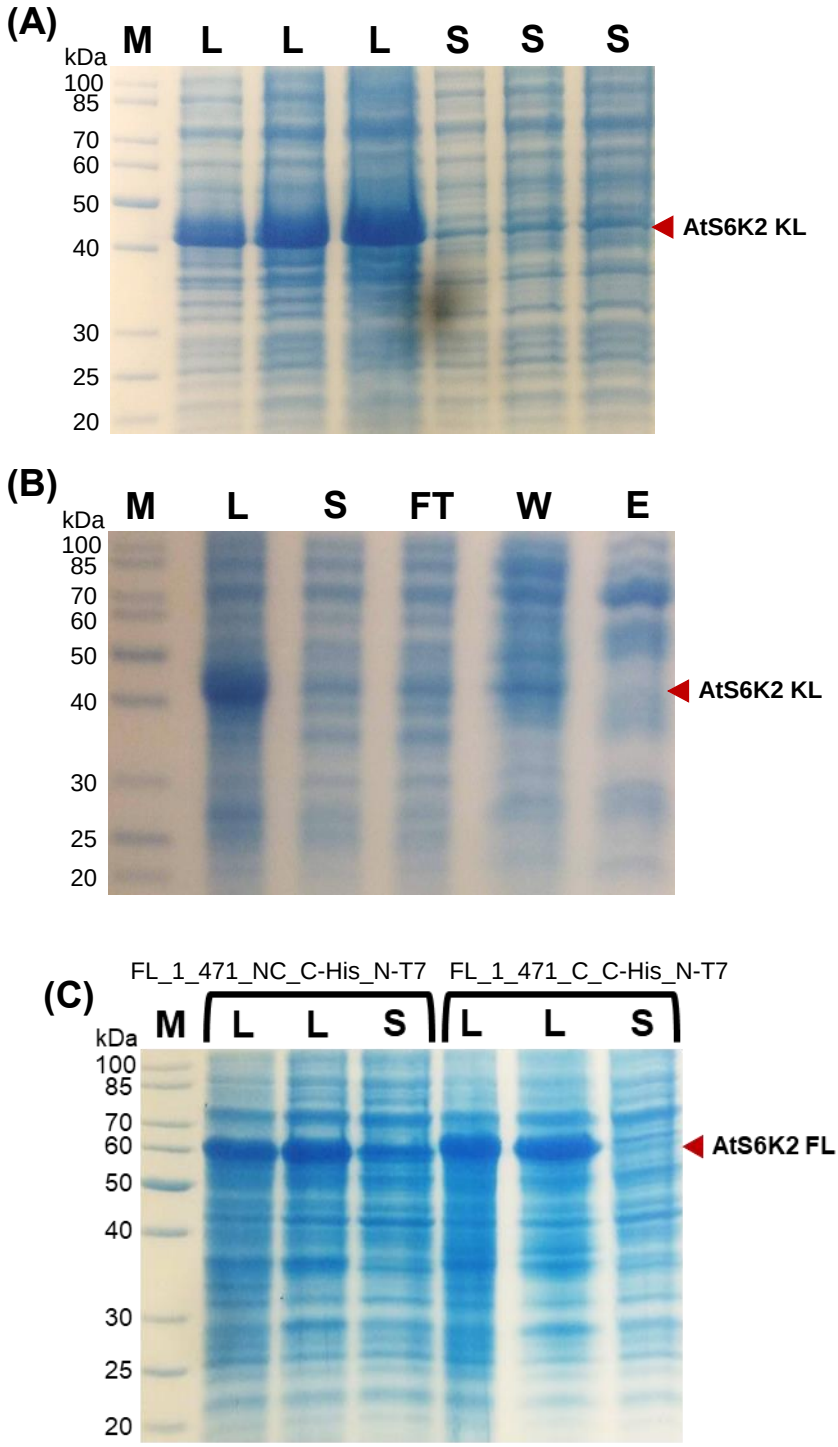


Figure 3.8| Expression and solubility of AtS6K2 KL_123_471_C_C-His_N-T7, FL_1_471_C_C-His_N-T7 and FL_1_471_NC_C-His_N-T7. **(A)** AtS6K2 KL_123_471_C_C-His_N-T7 expression in BL21 grown in AIM, showing increasing amounts (from left to right) of cell lysate (L) and Supernatant of centrifuged cell lysate (S). **(B)** Samples from different IMAC purification steps of AtS6K2 KL_123_471_C_C-His_N-T7 were analysed using SDS-PAGE. M – Protein ladder; L – Cell lysate; S – Supernatant of centrifuged cell lysate; FT – HisTrap column flow-through; W – HisTrap column wash; E – Elution. **(C)** AtS6K2 FL_1_471_C_C-His_N-T7 and FL_1_471_NC_C-His_N-T7 expression in BL21 grown in AIM, showing cell lysate (L) and Supernatant of centrifuged cell lysate (S). AtS6K2 KL_123_471_C_C-His_N-T7 (43.4 kDa), FL_1_471_C_C-His_N-T7 (56.6 kDa) and FL_1_471_NC_C-His_N-T7 (55.5 kDa) are highlighted with red arrows on the sides of gels.

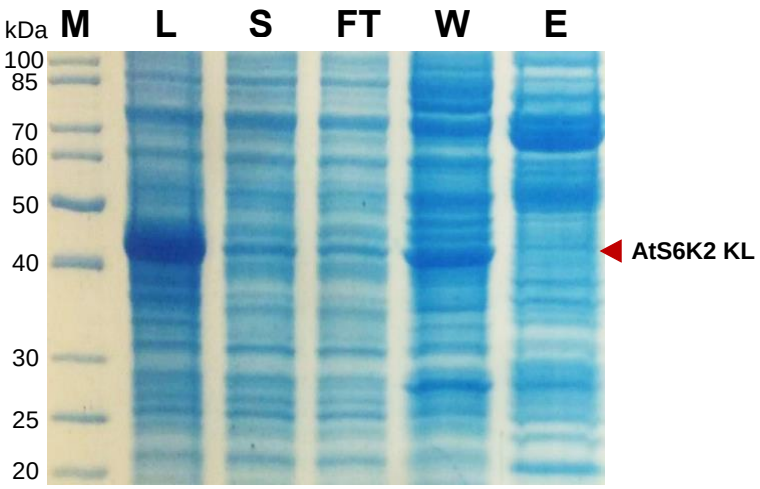


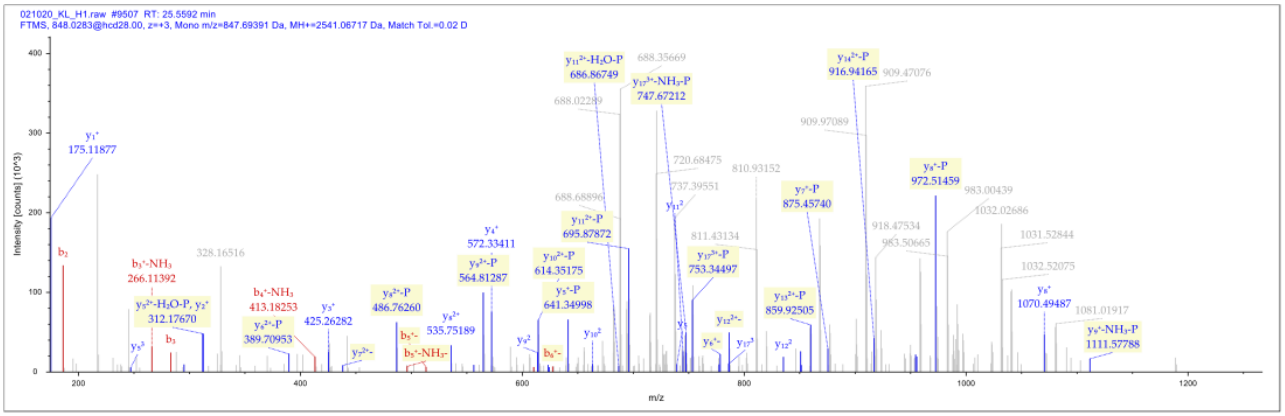
Figure 3.9| IMAC purification of AtS6K2 KL_123_471_C_C-His. Samples from different IMAC purification steps were analysed using SDS-PAGE. AtS6K2 KL_123_471_C_C-His (42.3 kDa) is highlighted with a red arrow on the side of the gel. M – Protein ladder; L – Cell lysate; S – Supernatant of centrifuged cell lysate; FT – HisTrap column flow-through; W – HisTrap column wash; E – Elution.

3.1.4.4. AtS6K2 KL_123_471_NC_C-His_N-T7 crystallisation and sequence validation

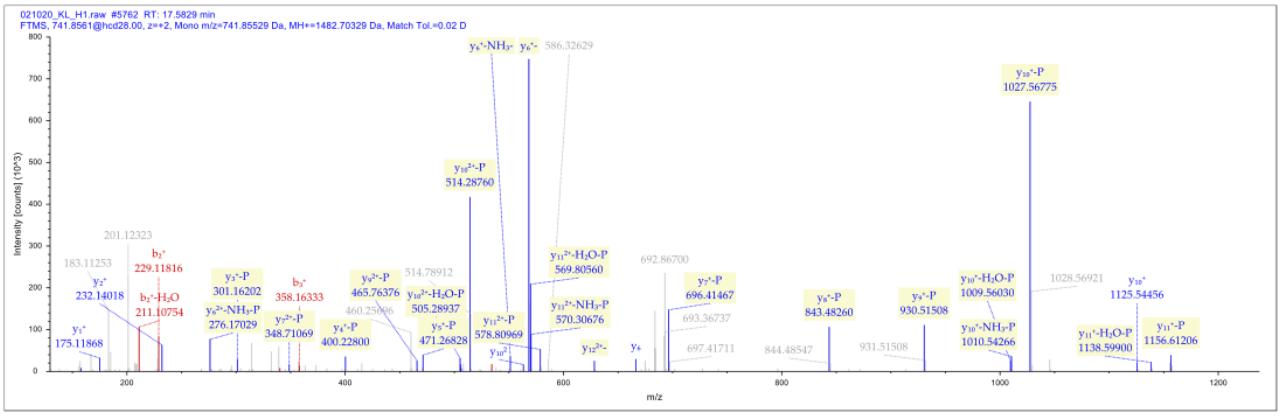
Success of protein crystallisation depends on finding a solution which is supersaturated in the biomolecule without disrupting its native structure by varying pH, temperature, protein concentration and precipitating agents (McPherson, 2017). Crystallisation experiments of the purified wild-type AtS6K2 KL_123_471_NC_C-His_N-T7 were based on crystallisation protocols of other AGC kinases (Ikuta *et al.*, 2007; Malakhova *et al.*, 2009; Wang *et al.*, 2013). KL_123_471_NC_C-His_N-T7 was concentrated to 10 mg/ml in the presence of 1 mM AMP-PNP and 2 mM $MgCl_2$ and screened for crystallisation conditions using a variety of commercially available sparse matrix screening kits (Table A15). A Gryphon liquid handling robot was used to set up crystallisation drops with either 1:1 or 1:2 volume ratios of protein to reservoir and plates were then incubated at 16°C. No crystals were obtained in any of the conditions tested and a vast majority of drops had no precipitation ($\gtrapprox$ 90% clear). This could suggest that the actual protein concentration used is too low. Thus, I concentrated KL_123_471_NC_C-His_N-T7 to 20 mg/ml and set up plates as before. The 20 mg/ml plates were incubated at 18°C (as this should at least increase precipitation) (Table A16) but surprisingly there was no evident increase in precipitation nor crystallisation.

Failure to crystallise the protein, coupled with excessive issues with protein solubility, urged a verification of the constructs produced. Samples of KL_123_471_NC_C-His_N-T7 (the only soluble constructs successfully purified) were analysed using mass-spectrometry (MS). Unexpectedly, three phosphorylations on residues S417, T455 and S462 were detected (Fig. 3.10). Of these only one has a known biological relevance, T455, which is phosphorylated by AtTOR during the protein activation in the cell. The other two phosphorylations have not been described before in AtS6K2 or human S6K2. Importantly, considering that S417 is predicted to be located above the catalytic region of the protein

(geometry not compatible with γ-phosphate transfer), and no other AGC-kinase has been reported to have phosphorylated residues in this position, these three phosphorylations are likely artefacts caused by the over-expression of a recombinant kinase in *E. coli*. As such, it is likely that the purification of KL_123_471_NC_C-His_N-T7 resulted in a heterogeneous protein sample, containing the recombinant protein with either 0, 1, 2 or 3 phosphorylated residues. Previous studies on active recombinant kinases have shown that phosphorylation and hyper-phosphorylation during recombinant protein expression may affect protein folding and prevent crystallisation and/or may induce non-native structural arrangements (Ali *et al.*, 2011; Shrestha *et al.*, 2012). This was an important result, as presence of these potentially random phosphorylation sites rationalised the lack of reproducibility observed in different expression tests and the failure to produce soluble samples. As such, no further tests on KL_123_471_NC_C-His_N-T7 were pursued and different strategies (i.e., changes in construct length and purification conditions) to obtain homogeneous soluble samples were employed.



Sequence: ANPFTNFTYVRPPHSFLHR (T455 and S462)



Sequence: EVQPSFKPAVSGR (S417)

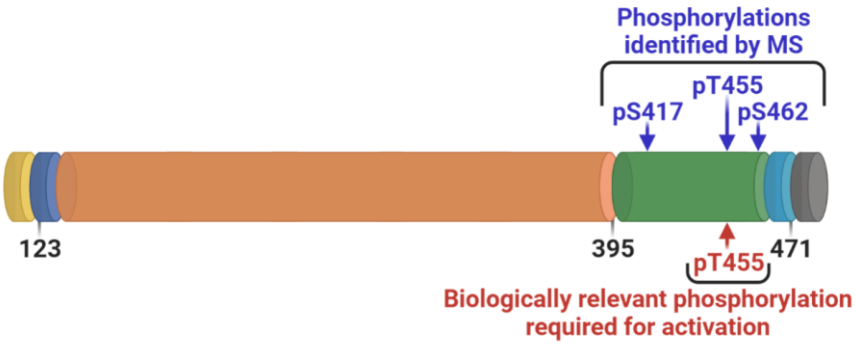


Figure 3.10| Mass Spectrometry spectra and schematic representation of KL_123_471_NC_C-His_N-T7 showing presence of 3 phosphorylations. AtS6K2 domains and motifs are coloured accordingly: yellow – T7 tag; blue – N-term domain; orange – kinase domain; green – AGC C-term region; cyan – non-conserved C-term tail; black – His-tag.

3.1.4.5. AtS6K2 KL_123_471_NC_C-His

An ulterior obstacle for protein crystallisation may have been caused by the T7-tag at the N-terminus of the recombinant protein and it may influence the native structure of AtS6K2. The KL sequence was re-cloned to generate the AtS6K2 KL_123_471_NC_C-His that has only the C-terminal His-tag and lacks the T7-tag (Table 3.1). KL_123_471_NC_C-His (41.1 kDa) was expressed in both LB and AIM and successfully purified by IMAC at pH 7.5 (Fig. 3.11A). Similarly to KL_123_471_NC_C-His_N-T7, KL_123_471_NC_C-His is likely to be randomly phosphorylated in several positions and form a heterogeneous solution. In order to produce homogenous AtS6K2 solution suitable for structural studies, the purified protein must be dephosphorylated. Unfortunately, treatment of the IMAC elution fraction with λ-phosphatase (26.2 kDa), using previously published protocols (Ali *et al.*, 2011; Prischi *et al.*, 2014), caused rapid protein aggregation (Fig. 3.11B), probably resulting from a destabilisation of a non-native protein fold. All IEC attempts were unsuccessful and dephosphorylated KL_123_471_NC_C-His was not produced.

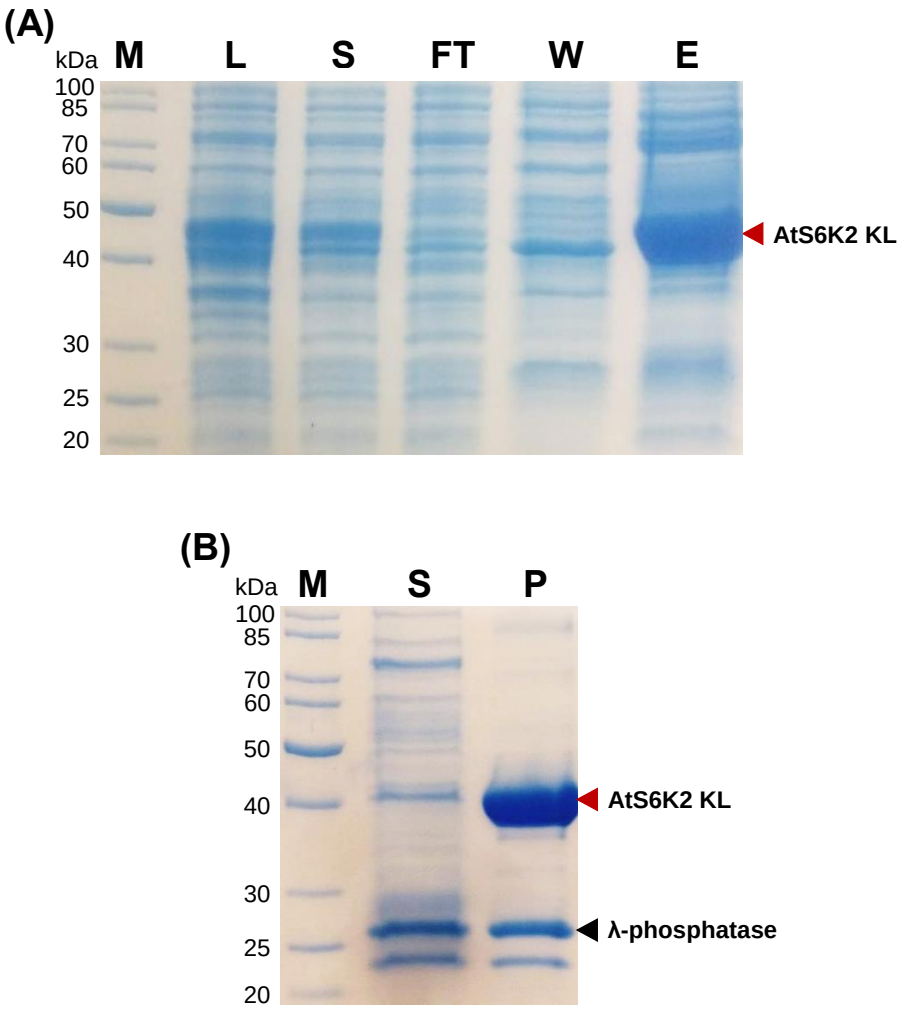


Figure 3.11| IMAC purification of AtS6K2 KL_123_471_NC_C-His. **(A)** Samples from different IMAC purification steps were analysed using SDS-PAGE. M – Protein ladder; L – Cell lysate; S – Supernatant of centrifuged cell lysate; FT – HisTrap column flow-through; W – HisTrap column wash; E – Elution. **(B)** Elution samples treated ON with λ-phosphatase were centrifuged and the supernatant (S) and precipitated fraction (P) were analysed by the SDS-PAGE. AtS6K2 KL_123_471_NC_C-His (41.1 kDa) is highlighted with red arrows on the sides of gels, while λ-phosphatase is highlighted with a black arrow.

3.1.4.5.1. AtS6K2 KL_123_471_NC_C-His and KL_123_471_C_N-His co-expression with λ-phosphatase

Instead of treating purified sample with λ-phosphatase a different dephosphorylation strategy was employed. It was previously shown that λ-phosphatase co-expression improved protein solubility and produced high yields of the human kinases Src and Abl expressed in bacteria (Seeliger *et al.*, 2005). To co-express the kinase constructs with λ-phosphatase, Albanese *et al.*, (2018) customised protein expression vectors from the pET-13 and -2 series, specifically the pET-13S-A expressing λ-phosphatase and pET-2B-T expressing kinases. Thus, I used pET-13S-A for λ-phosphatase expression but pET plasmids to express AtS6K2 were different from this study (pET-21a and pETM-14). Co-expression of KL_123_471_NC_C-His (41.1 kDa) with λ-phosphatase (25.2 kDa) in AIM followed by IMAC, revealed that only the λ-phosphatase was expressed (Fig. 3.12A). In addition, I tested whether co-expression with λ-phosphatase may improve solubility of KL_123_471_C_N-His which was previously shown to be insoluble (Section 3.1.3.1; Fig. 3.3). The result of co-expression in LB media and subsequent IMAC purification was similar to KL_123_471_NC_C-His as KL_123_471_C_N-His was not expressed (Fig. 3.12B). A common issue in co-expression using different plasmids is linked to the origin of replication, as similar origins tend to compete for the DNA polymerase resulting in the preferential replication of one plasmid at the expense of the other. However, both pET-21a and pETM-14 include a pBR322 origin of replication (ori) which is similar to the ColE1 ori present in pET-2B-T. pBR322/ColE1 ori is compatible with the CDF (CloDF13) ori and should be suitable for effective co-expression in *E. coli*. As such, it is unclear why the co-expression failed. The only information provided by Albanese *et al.*, (2018) that could potentially explain the failure of the failed co-expression is that the pET-13S-A has been engineered to be co-transformed with pET-2B-T for optimal expression (the presence of the 13S vector enhances expression of the gene in the 2A-T plasmid).

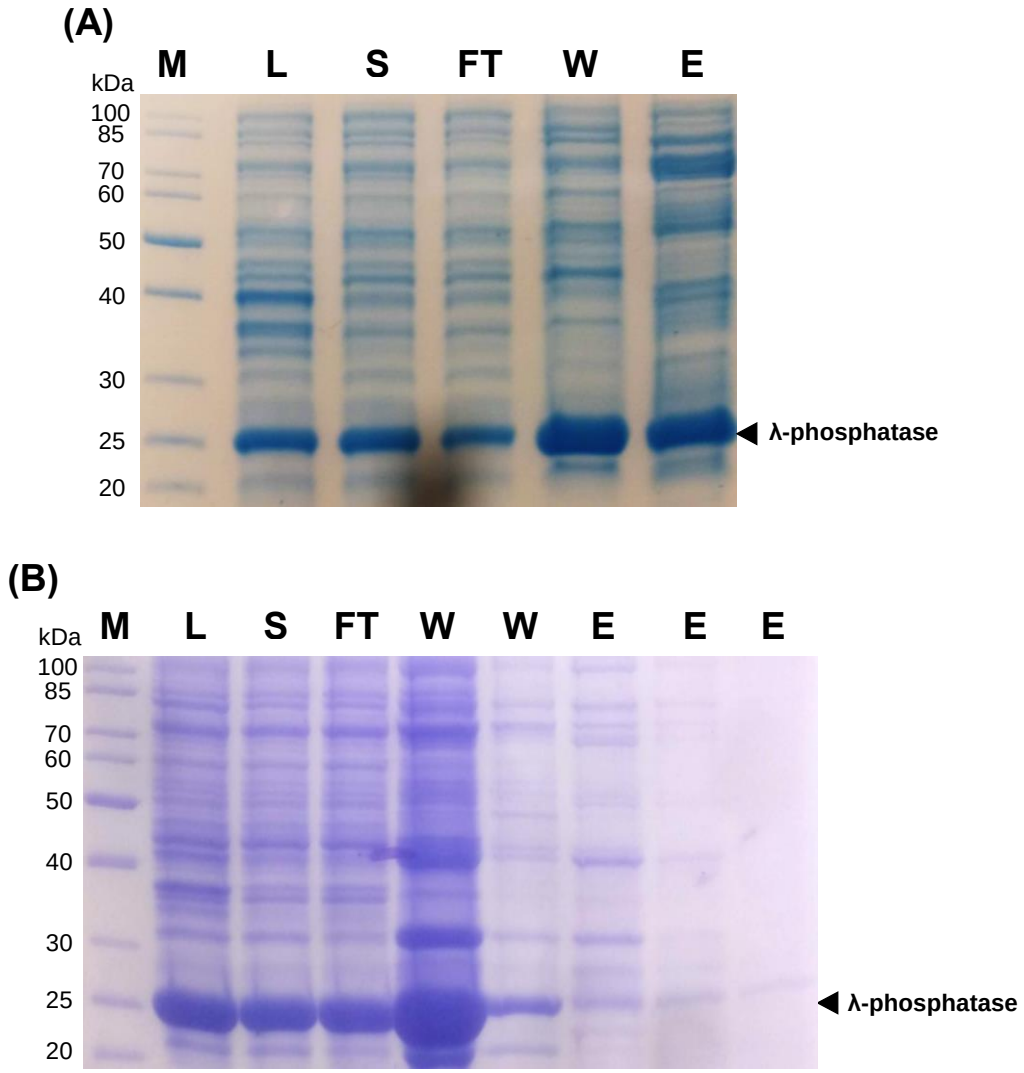


Figure 3.12| Co-expression of AtS6K2 KL_123_471_NC_C-His and KL_123_471_C_N-His with λ-phosphatase. Samples from different IMAC purification steps for **(A)** AtS6K2 KL_123_471_NC_C-His (41.1 kDa) co-expression in BL21 grown in AIM, and **(B)** AtS6K2 KL_123_471_C_N-His (42 kDa) co-expression in BL21 grown in LB. M – Protein ladder; L – Cell lysate; S – Supernatant of centrifuged cell lysate; FT – HisTrap column flow-through; W – HisTrap column wash; E – Elution. λ-phosphatase is highlighted with black arrows on the sides of gels.

3.1.4.5.2. AtS6K2 KL_123_471_C_N-GST co-expression with λ-phosphatase

In addition to the available pET constructs, I attempted to co-express λ-phosphatase with the KL_123_471_C_N-GST. Small scale test co-expression of KL_123_471_C_N-GST and λ-phosphatase indicated overexpression of both proteins. Samples taken after induction had strong bands at ≈60 kDa for KL_123_471_C_N-GST (66.6 kDa) and ≈25 kDa for λ-phosphatase (25.2 kDa) (Fig. 3.13A). I expressed the proteins in LB at a large scale and purified by IMAC. Cell lysate exhibited similar pattern to the small scale expression. Despite successful co-expression, only a very low amount of soluble protein was recovered after IMAC (Fig. 3.13B), which, unfortunately, was completely lost during ion exchange chromatography (either during concentration or low salt buffer exchange). The low expression of soluble AtS6K2 may also be linked to the use of plasmid not designed or optimised for co-expression.

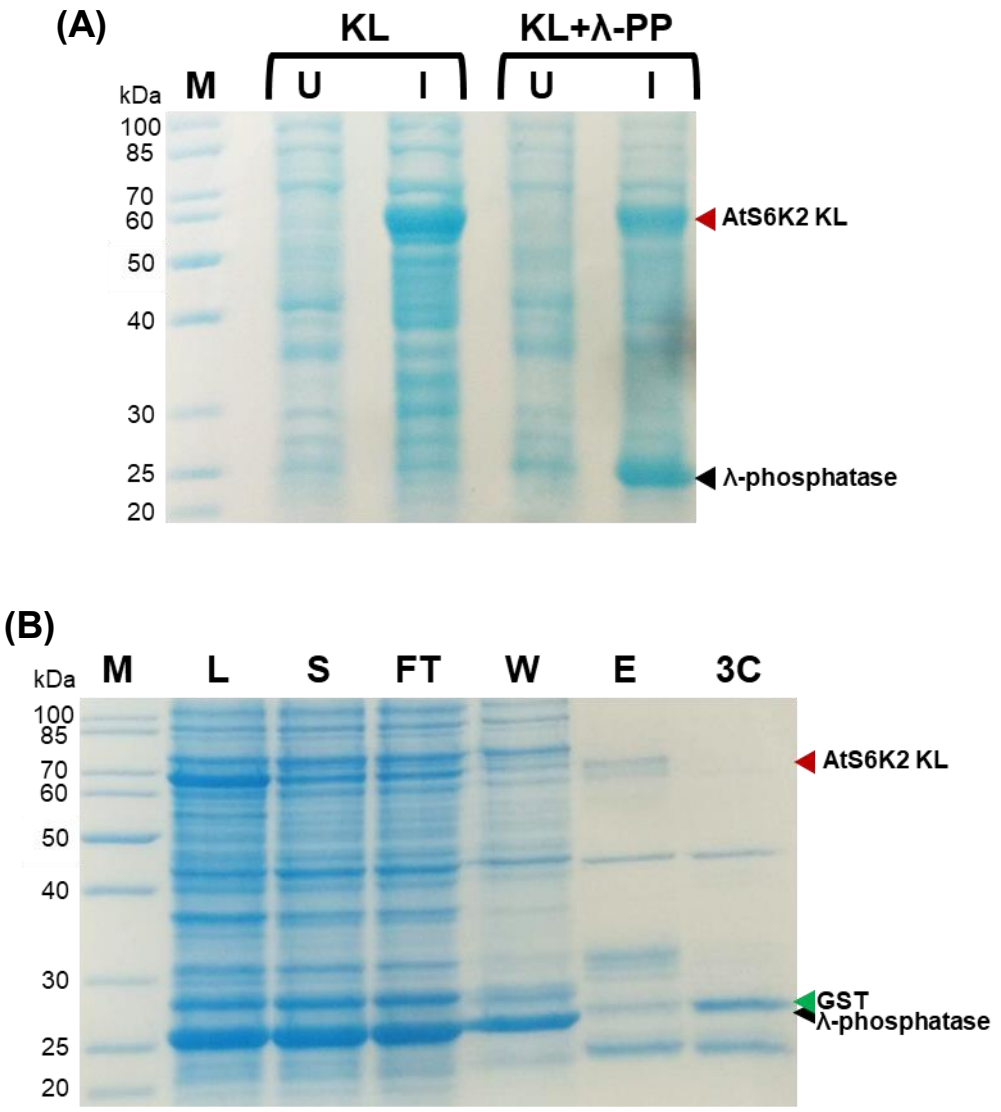


Figure 3.13| Co-expression of AtS6K2 KL_123_471_C_N-GST with λ-phosphatase. **(A)** Small scale test expression (using BL21 grown in LB) of KL_123_471_C_N-GST alone and co-expressed with λ-phosphatase showing IPTG induced (I) and uninduced (U) samples. **(B)** Samples from different IMAC purification steps of AtS6K2 KL_123_471_C_N-GST were analysed using SDS-PAGE. M – Protein ladder; L – Cell lysate; S – Supernatant of centrifuged cell lysate; FT – HisTrap column flow-through; W – HisTrap column wash; E – Elution. AtS6K2 KL_123_471_C_N-GST (66.6 kDa) is highlighted with red arrows on the sides of gels, λ-phosphatase (25.2 kDa) is highlighted with black arrows and GST (26 kDa) is highlighted with a green arrow.

3.1.4.5.3. AtS6K2 KL_123_471_NC_C-His production

In parallel to the co-expression experiment, I did more optimisation studies on the promising KL_123_471_NC_C-His (41.2 kDa) construct. To rule out the influence of pH in the solubility of phosphorylated-dephosphorylated proteins, due to the potential reduction in protein pI that is linked to hetereogenous phosphorylations (from 8.8 to about 7.3), KL_123_471_NC_C-His was also purified at pH 6.5, but the protein was insoluble in these conditions. Changing the pH to 8.0 and dephosphorylating after IEC proved to be the key to successful protein production (Table A6). Similarly to KL_123_471_NC_C-His_N-T7, a three-step purification was performed. IMAC results were analogous to those at pH=7.5 (Fig. 3.11A). The pooled fractions from a HisTrap column elution was then purified by cation exchange chromatography which resulted in a peak at 330 mM NaCl (Fig. 3.14A and 3.14B). After incubation with λ-phosphatase, the IEC fractions were loaded onto a gel filtration column (HiLoad 16/600 Superdex 75 pg). Three peaks were present in the chromatogram: a single peak at approximately 58 mL containing AtS6K2 with a small shoulder at the start; two overlapping beaks (at approximately 73 mL and 80 mL) containing the lambda phosphatase and a contaminant (Fig. 3.15C). Fractions containing AtS6K2 KL_123_471_NC_C-His were pooled together, concentrated to 2 mg/mL, flash-frozen using liquid nitrogen and stored at -80°C. Initial results indicated that SEC successfully removed the λ-phosphatase together with other contaminants and highly purified protein was produced (Fig. 3.15A).

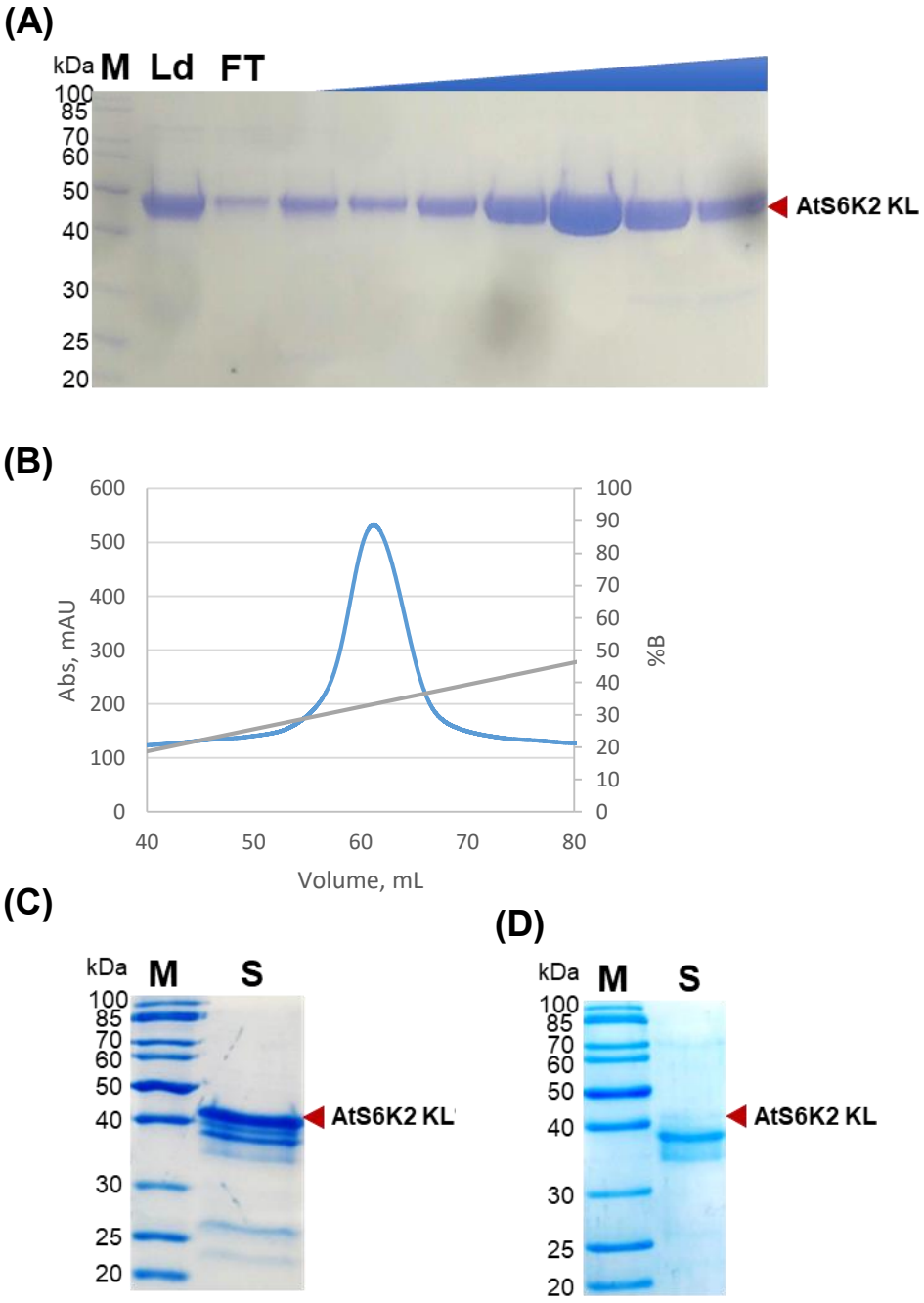


Figure 3.14| IEC Purification of AtS6K2 KL_123_471_NC_C-His. **(A)** Samples from ion exchange chromatography (IEC). Ld – sample loaded on the cation exchange column; FT – cation exchange column flow-through. Blue triangle represents the NaCl gradient, and samples eluting between 280 mM and 350 mM NaCl were loaded on an SDS-PAGE. AtS6K2 KL_123_471_NC_C-His (41.1 kDa) is highlighted with a red arrow on the side of the gel. **(B)** Chromatogram of IEC. **(C)** Purified AtS6K2 KL_123_471_NC_C-His sample (S) after SEC. AtS6K2 KL_123_471_NC_C-His is highlighted with a red arrow on the side of the gel. **(D)** Purified AtS6K2 KL_123_471_NC_C-His sample (S) after 7-day storage at 4°C. The Position of degraded AtS6K2 KL_123_471_NC_C-His band is highlighted with a red arrow on the side of the gel.

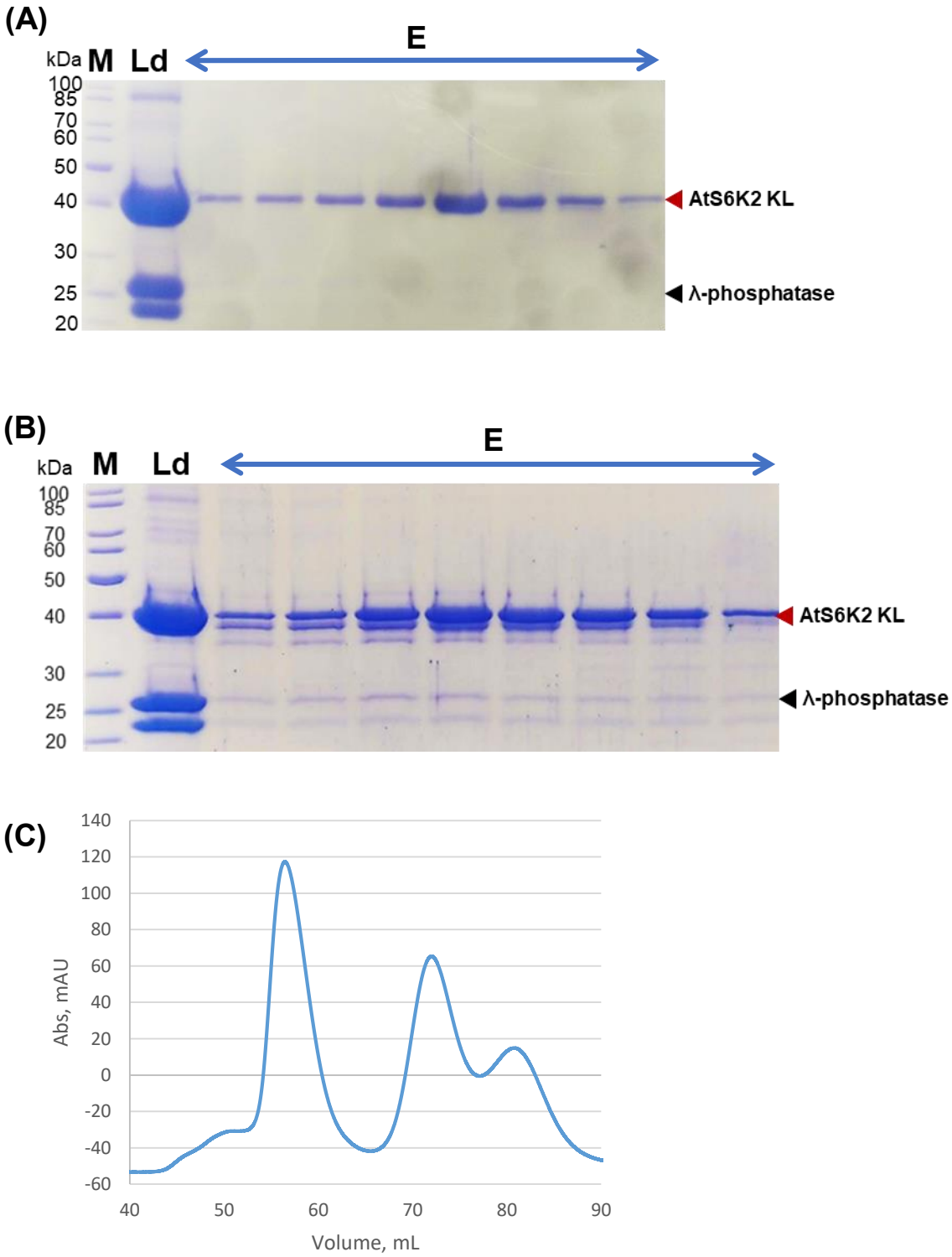


Figure 3.15| SEC purification of AtS6K2 KL_123_471_NC_C-His following enzymatic dephosphorylation. **(A-B)** Elution samples from two different SEC attempts with HiLoad 16/600 Superdex 75 pg column, V=120 mL. M – Protein ladder; Ld – sample loaded on the SEC column. E- samples eluting 50-65 mL were loaded on the SDS-PAGE. AtS6K2 KL_123_471_NC_C-His (41.1 kDa) is highlighted with red arrows on the sides of the gels and λ-phosphatase with black arrows. **(C)** Chromatogram of SEC.

Furthermore, produced KL_123_471_NC_C-His T455E single phosphomimetic mutant and S296E T455E double phosphomimetic mutant. Small scale test expression of KL_123_471_NC_C-His T455E was successful as there was a band of ≈40 kDa in the cell lysate (Fig. 3.16A). Double KL_123_471_NC_C-His mutant was expressed in AIM and lysed with and without 5 mM β- mercaptoethanol at pH 7.5 and 8.0. The protein was insoluble in any of the conditions tested as the ≈40 kDa band was absent from the supernatant of centrifuged cell lysate (Fig. 3.16B). No further experiments with KL_123_471_NC_C-His mutants were performed.

3.1.4.5.4. AtS6K2 KL_123_471_NC_C-His crystallisation and sequence validation

A pre-crystallization test was used to determine the appropriate protein concentration for crystallization screening, which was found to be 8-10 mg/mL. Initially, purified KL_123_471_NC_C-His was concentrated to a final concentration of 10 mg/ml in presence of 1 mM AMP-PNP and 2 mM $MgCL_2$. Unfortunately, the protein precipitated while being dispersed with the Gryphon robot at room temperature (20°C) and only one plate was set up (Table A17). Because KL_123_471_NC_C-His was unstable at room temperature, I lowered the protein concentration to 8 mg/ml to reduce precipitation rate. I also aliquoted the protein into small batches and kept the main sample on ice. This allowed me to set up multiple screening plates, which were incubated at 18°C. Amorphous precipitation formed in most conditions (>90%), and I decided to lower the protein concentration to 7 mg/mL. To reduce the possible precipitation due to temperature, I pre-cooled the screening plates on ice and incubated them at 4°C. This did not reduce the amount of precipitation that formed after 14 days suggesting that further optimisation of crystallisation conditions was required. However, additional purification attempts revealed that SEC elution samples contained a double band of ≈40 kDa (Fig. 3.15B). This may be an aggregate or a

contaminant eluting in the void volume, which partially overlapped with the KL_123_471_NC_C-His elution peak. To increase separation between void volume and KL_123_471_NC_C-His elution, I used an S200 SEC column (HiLoad 16/600 Superdex 200 pg, V=120 mL) instead of S75, which did not remove the contamination suggesting that the band in the SDS-PAGE is not a contamination eluting in the void volume, but it is likely an event of protein degradation. Indeed, storage of the protein at 4°C for 7 days resulted in no visible precipitation but the KL_123_471_NC_C-His upper band (≈41 kDa) almost completely disappeared, while the lower one had increased correspondingly in intensity (Fig. 3.14C and 3.14D), which confirmed the hypothesis that the 2 bands are both KL_123_471_NC_C-His. This warranted verification of the produced protein by MS. Both bands that were present after 7 days of storage (Fig. 3.14D) were sent for MS analysis, which confirmed that the ≈38 kDa band is AtS6K2, which is cleaved after residue K447 (Fig. 3.17). Presence of such degradation pattern could induce protein aggregation and heterogeneity of the sample rationalising the lack of stability and unsuccessful crystallisation of KL_123_471_NC_C-His.

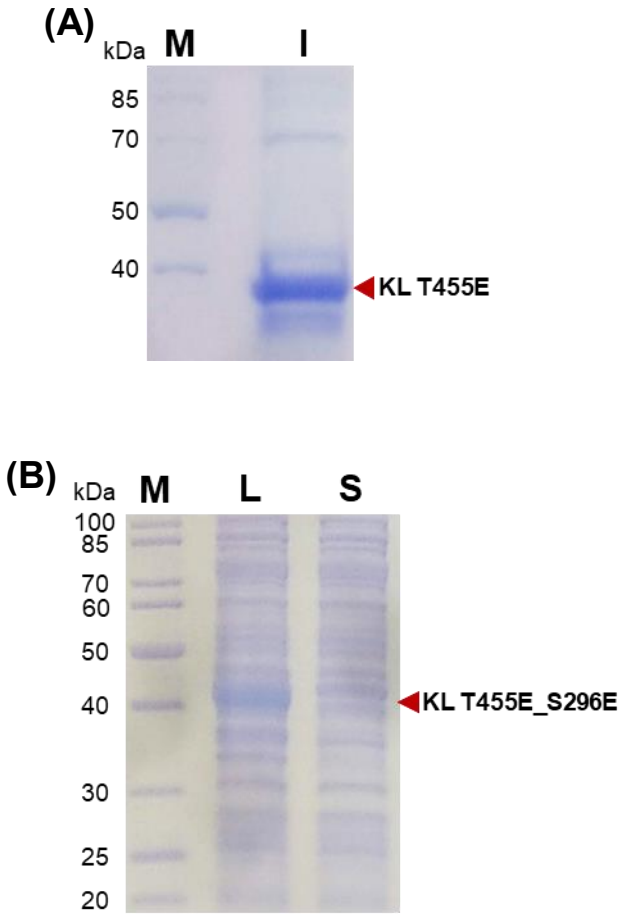


Figure 3.16| Expression of AtS6K2 KL_123_471_NC_C-His T455E and T455E_S296E mutants. **(A)** Small scale test expression (using BL21 grown in LB) of AtS6K2 KL_123_471_NC_C-His T455E mutant (MW=41.2 kDa) showing IPTG induced (I) sample. **(B)** AtS6K2 KL_123_471_NC_C-His T455E_S296E mutant (MW=41.2 kDa) expression (using BL21 grown in AIM), showing cell lysate (L) and supernatant of centrifuged cell lysate (S). AtS6K2 KL_123_471_NC_C-His mutants are highlighted with red arrows on the sides of gels.

(A)

Protein sequence coverage: 65%

Matched peptides shown in ***bold red***.

```
  1 DSEKSPEEVS GVVGIEDFEV LKVVGQGAFG KVYQVRKKDT SEIYAMKVMR
 51 KDKIVEKNHA EYMKAERDIL TKIDHPFIVQ LKYSFQTKYR LYLVLDFING
101 GHLFFQLYHQ GLFREDLARV YTAEIVSAVS HLHEKGIMHR DLKPENILMD
151 VDGHVMLTDF GLAKEFEENT RSNSMCGTTE YMAPEIVRGK GHDKAADWWS
201 VGILLYEMLT GKPPFLGSKG KIQQKIVKDK IKLPQFLSNE AHALLKGLLQ
251 KEPERRLGSG PSGAEEIKKH KWFKAINWKK LEAREVQPSF KPAVSGRQCI
301 ANFDKCWTDM SVLDSPASSP NSDAKANPFT NFTYVRPPHS FLHRTTSNL
```

(B)

Protein sequence coverage: 66%

Matched peptides shown in ***bold red***.

```
  1 DSEKSPEEVS GVVGIEDFEV LKVVGQGAFG KVYQVRKKDT SEIYAMKVMR
 51 KDKIVEKNHA EYMKAERDIL TKIDHPFIVQ LKYSFQTKYR LYLVLDFING
101 GHLFFQLYHQ GLFREDLARV YTAEIVSAVS HLHEKGIMHR DLKPENILMD
151 VDGHVMLTDF GLAKEFEENT RSNSMCGTTE YMAPEIVRGK GHDKAADWWS
201 VGILLYEMLT GKPPFLGSKG KIQQKIVKDK IKLPQFLSNE AHALLKGLLQ
251 KEPERRLGSG PSGAEEIKKH KWFKAINWKK LEAREVQPSF KPAVSGRQCI
301 ANFDKCWTDM SVLDSPASSP NSDAKANPFT NFTYVRPPHS FLHRTTSNL
```

Figure 3.17| Mass Spectrometry coverage analysis for AtS6K2 KL_123_471_NC_C-His, showing results for two repeat experiments **(A-B)**. Sequence coverage % is highlighted on top of the sequence. Peptides identified using mass spectrometry are mapped on the protein sequence and highlighted in red.

3.1.5. Purification of AtS6K2 KD constructs in the presence of Li_2S0_4

In an attempt to optimise the AtS6K2 purification protocol, I investigated the differences between my previous purification experiments and the ones that resulted in successful crystallisation and structure determination of S6K1, as there are no *in vitro* data for any S6K2. Initial buffer composition was based on Qin *et al.*, (2015), however other published purification buffers for human S6K1 contain Li_2SO_4, instead of NaCl (Sunami *et al.*, 2010; Wang *et al.*, 2013). As such, I re-optimised the lysis and purification buffers by completely substituting the NaCl with Li_2SO_4 for AtS6K2 KD_123_414_C_N-GST, KD_123_414_NC_C-His_N-T7 and KD_123_414_C_C-His_N-T7.

3.1.5.1. AtS6K2 KD_123_414_C_N-GST

Small scale test expression of KD_123_414_C_N-GST (60.3 kDa; Table 3.1) resulted in a strong overexpression band of ≈60kDa after IPTG induction (Fig. 3.18A). Large scale expression was performed in AIM and KD_123_414_C_N-GST purified according to the 3-step protocol. An overexpression band at ≈60 kDa was present in all the IMAC fractions while the GST (26 kDa) band was visible in the wash and elution fractions (Fig. 3.18B). GST column elution fractions were pooled together and incubated with λ-phosphatase and HRV 3C PreScission protease at 4°C to remove the GST tag and any phosphorylations. No precipitation formed during the overnight (≈20 hour) incubation and the sample was further purified by cation exchange column, eluting at 360 mM Li_2SO_4 with no visible contaminants (Fig. 3.18C and 3.18D). SEC purification resulted in a single peak at 62 mL that contained clean AtS6K2 KD_123_414_C_N-GST (Figs. 3.19A and 3.19B). To avoid protein precipitation due to freeze-thawing cycle, I screened for crystallisation conditions (Table A18) immediately after the protein was purified and concentrated without storing it at -80°C. This produced enough sample (10 mg/mL KD_123_414_C_N-GST in

presence of 1 mM AMP-PNP and 2 mM $MgCL_2$) to attempt crystallisation trials using commercial screens. Unfortunately, concentrated sample showed limited stability at room temperature (20°C), forming heavy precipitation while dispensed by the robot in the plates. However, the few resulting plates were incubated at 16°C. Surprisingly, crystals were detected in two conditions (Table 3.2): (i) 0.2M Calcium chloride dihydrate; 0.1M Sodium acetate pH=5.0; 20% w/v PEG6000 (Fig. 3.19D) and (ii) 0.2M Calcium chloride dihydrate; 0.1M HEPES pH=7.0; 20% w/v PEG6000 (Fig. 3.19C). Crystals were harvested and soaked in cryoprotectant (same as reservoir buffer with 50% glycerol). X-ray diffraction experiments were performed at Diamond Light Source (DLS), confirming that the crystals formed in both conditions were salt.

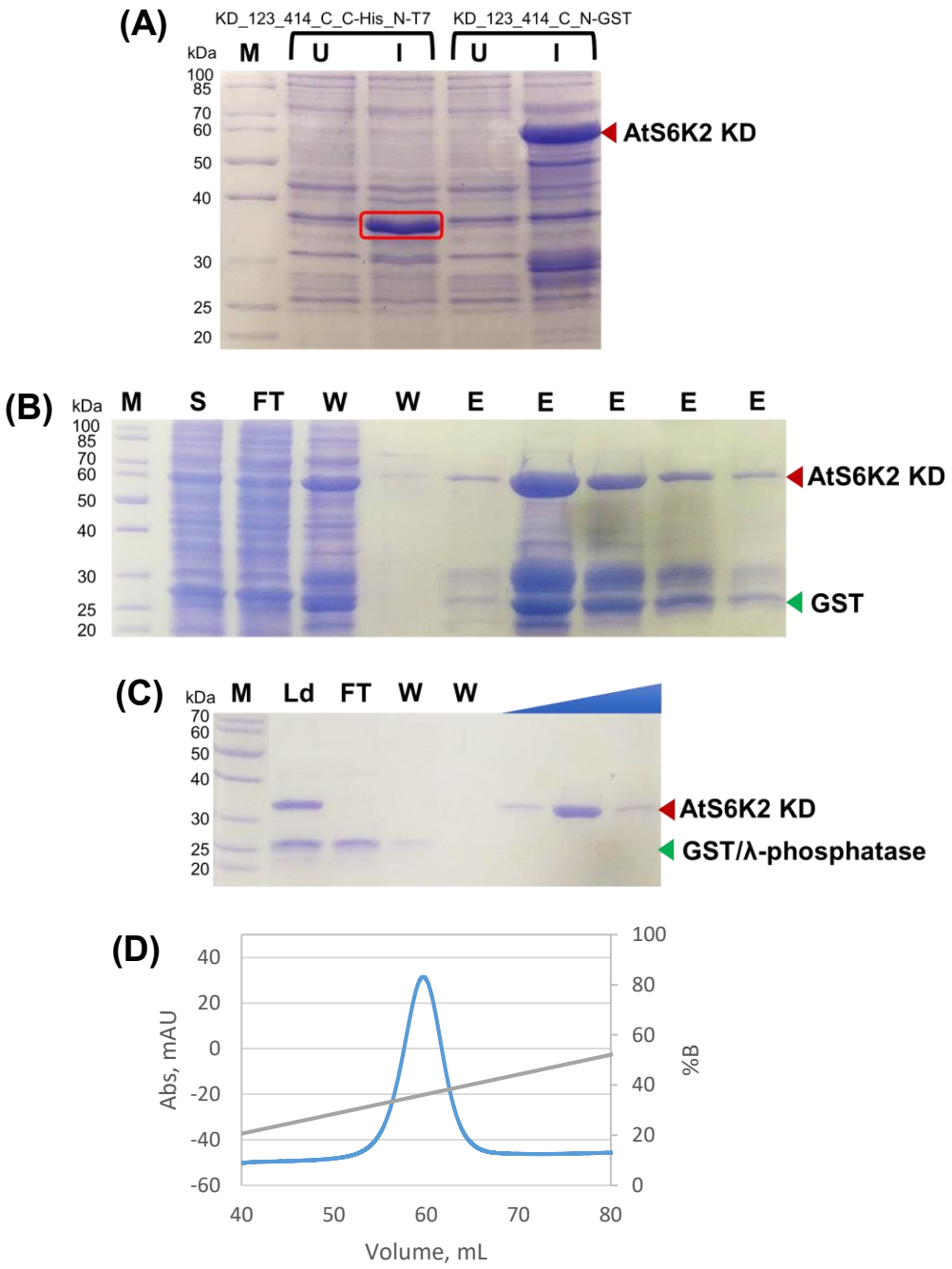


Figure 3.18| Expression and purification of AtS6K2 KD_123_414_C_N-GST. **(A)** Small scale test expression (using BL21 grown in LB) of AtS6K2 KD_123_414_C_N-GST (60.3/33.9 kDa) and KD_123_414_C_C-His_N-T7 (43.4 kDa) showing IPTG induced (I) and uninduced (U) samples. AtS6K2 KD_123_414_C_C-His_N-T7 is highlighted with red box. **(B)** Samples from different IMAC purification steps of AtS6K2 KD_123_414_C_N-GST. L – Cell lysate; S – Supernatant of centrifuged cell lysate; FT – GST column flow-through; W – GST column wash; E – Elution. **(C)** Samples from IEC of AtS6K2 KD_123_414_C_N-GST. Blue triangle represents the NaCl gradient, and samples eluting between 300 mM and 420 mM Li_2SO_4 were loaded on the SDS-PAGE. Ld – sample loaded on the cation exchange column; FT – cation exchange column flow-through; W – 50 mM Li_2SO_4 column wash. AtS6K2 KD_123_414_C_N-GST is highlighted with red arrows on the sides of the gels, while GST and λ-phosphatase with green arrows. **(D)** Chromatogram of the IEC. M – Protein ladder.

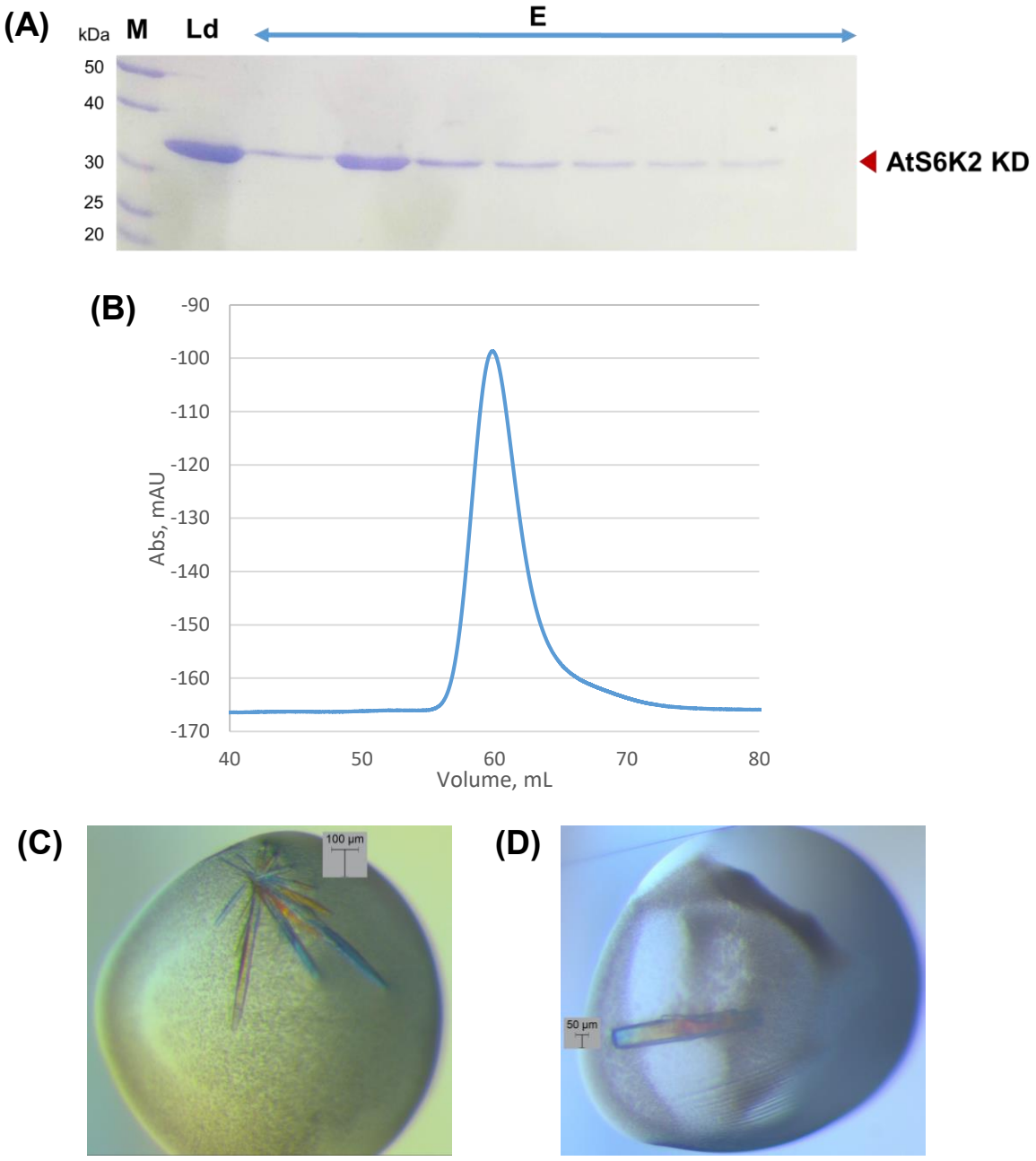


Figure 3.19| SEC purification and crystallisation of AtS6K2 KD_123_414_C_N-GST. **(A)** Elution Samples of AtS6K2 KD_123_414_C_N-GST (33.9 kDa) SEC with HiLoad 16/600 Superdex 75 pg column, V=120 mL. M – Protein ladder; Ld – sample loaded on the SEC column. E – samples eluting 56-73 mL were loaded on the SDS-PAGE. AtS6K2 KD_123_414_C_N-GST is highlighted with a red arrow on the side of the gel. **(B)** Chromatogram of SEC. **(C)** Crystals observed in PACT C11. **(D)** Crystals observed in PACT A11. Crystallisation conditions are set out in Table 3.2.

Table 3.2| Crystals in AtS6K2 KD constructs' screens

Screen, condition	Contents	Constructs
PACT A11	0.2 M Calcium chloride dihydrate; 0.1 M Sodium acetate 5.0; 20 % w/v PEG 6000	KD_123_414_C_N-GST; KD_123_414_NC_C-His_N-T7; KD_123_414_C_C-His_N-T7
PACT C11	0.2 M Calcium chloride dihydrate; 0.1 M HEPES 7.0; 20 % w/v PEG 6000	KD_123_414_C_N-GST; KD_123_414_NC_C-His_N-T7
PEG-Ion A7	0.2 M Calcium chloride dehydrate; 20% w/v Polyethylene glycol 3,350; pH=5.1	KD_123_414_NC_C-His_N-T7; KD_123_414_C_C-His_N-T7
PEG-Ion C4	0.2 M Calcium acetate hydrate; 20% w/v Polyethylene glycol 3,350; pH=7.5	KD_123_414_NC_C-His_N-T7
JCSG A4	0.02 M Calcium chloride dehydrate; 0.1 M Sodium acetate pH=4.6; 30% v/v MPD	KD_123_414_NC_C-His_N-T7

3.1.5.2. AtS6K2 KD_123_414_NC_C-His_N-T7

Next, I tested the KD construct with uncleavable C-term HisTag and N-term T7 tag (Table 3.1). Small scale test expression of KD_123_414_NC_C-His_N-T7 (36 kDa) showed that the protein is overexpressed in LB media (Fig. 3.20A). Expression was then scaled-up in AIM and the standart 3-step purification was performed. The pooled fractions from the HisTrap column elution were incubated with λ-phosphatase, forming a limited amount of precipitation (Fig. 3.20B). The Sample was then further purified using a cation exchange column. A peak appeared at 340 mM Li_2SO_4, that contained KD_123_414_NC_C-His_N-T7 with no visible contaminations (Fig. 3.20C and 3.20D). SEC resulted in a peak at 61.5 mL and fractions with high protein content (Fig. 3.21A and 3.21B) were concentrated to 10 mg/mL and screened for crystallisation conditions similarly to KD_123_414_C_N-GST. Enough sample was available to test four different screening 96-well plates (PACT, PGA, PEG/Ion Screen and JCSG-plus) and incubated at 16°C (Table A18). Crystals that lacked sharp edges appeared in a number of conditions after 24 hours (Table 3.2; Fig. 3.22).

Crystal growth in PACT C11 condition, that contained crystals with more defined and consistent shape, was optimised by varying PEG 6000 (16% - 22%) and $CaCl_2 \cdot 2H_20$ (0.18 M - 0.22 M) concentrations and setting up 24-well hanging drop vapor diffusion plate with 1µL protein + 1µL reservoir drops. To prevent possible precipitation due to temperature, the plate was set up at 4°C with pre-cooled buffers and equipment. Crystals formed after 40-hour incubation at 4°C in most conditions but the largest ones grew in 22% PEG6000 and 0.2/0.18 M $CaCl_2$. The condition with lower Ca^{2+} concentration was chosen for diffraction measurements with 20% glycerol used as cryoprotectant. Similarly to KD_123_414_C_N-GST, diffraction pattern confirmed that all of the conditions yielded salt crystals.

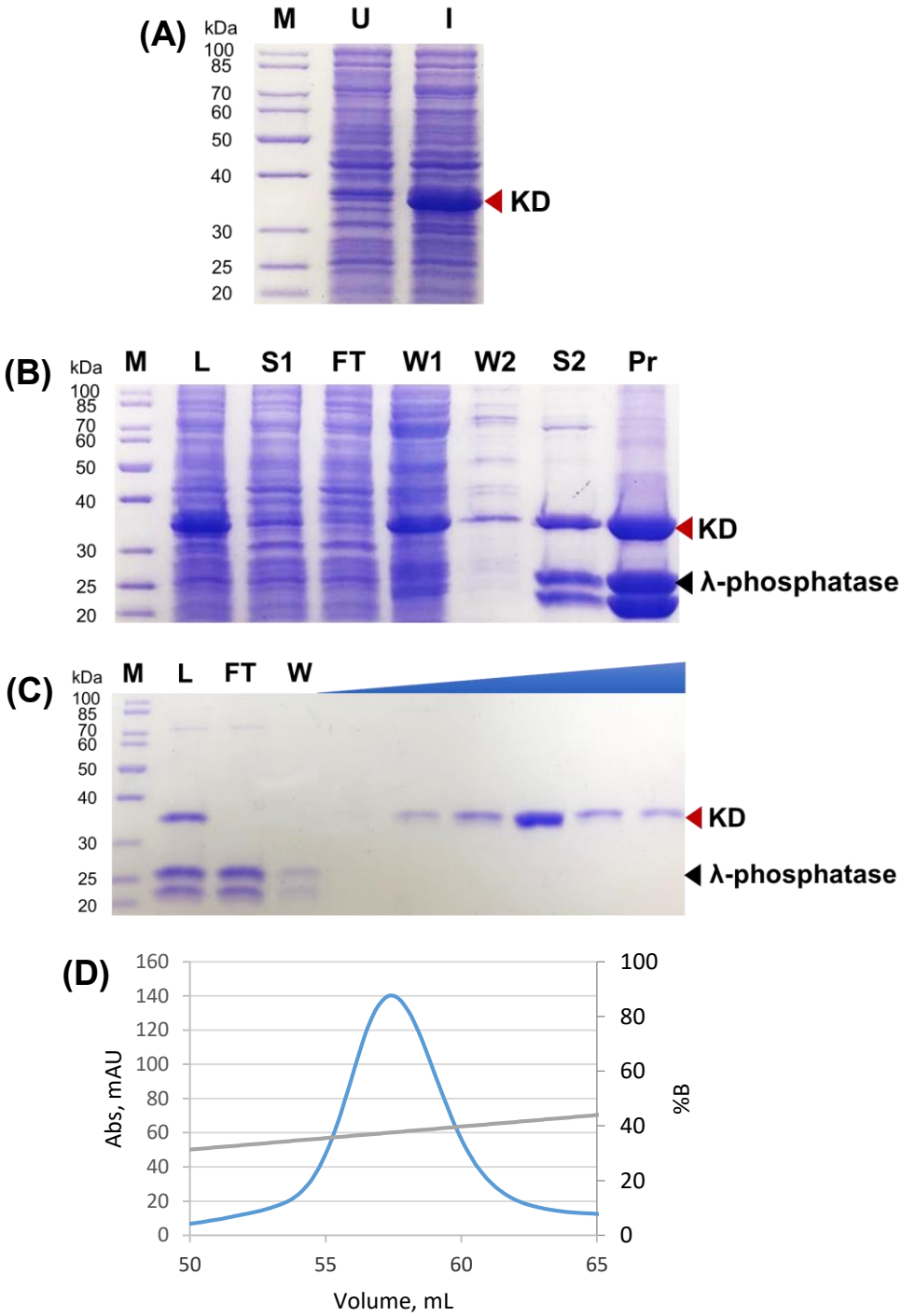


Figure 3.20| Expression and purification of AtS6K2 KD_123_414_NC_C-His_N-T7. **(A)** Small scale test expression (using BL21 grown in LB) of AtS6K2 KD_123_414_NC_C-His_N-T7 (36 kDa) showing IPTG induced (I) and uninduced (U) samples. **(B)** SDS-PAGE gel of samples from different IMAC purification steps of AtS6K2 KD_123_414_NC_C-His_N-T7. L – Cell lysate; S – Supernatant of centrifuged cell lysate; FT – HisTrap column flow-through; W – HisTrap column wash; E – Elution. **(C)** SDS-PAGE gel of samples from IEC of AtS6K2 KD_123_414_NC_C-His_N-T7. Blue triangle represents the NaCl gradient, and samples eluting between 330 mM and 420 mM Li_2SO_4 were loaded on the SDS-PAGE. Ld – sample loaded on the cation exchange column; FT – cation exchange column flow-through; W – 50 mM Li_2SO_4 column wash. AtS6K2 KD_123_414_NC_C-His_N-T7 is highlighted with red arrows on the sides of the gels and λ-phosphatase with black arrows. **(D)** Chromatogram of the cation exchange chromatography . M – Protein ladder.

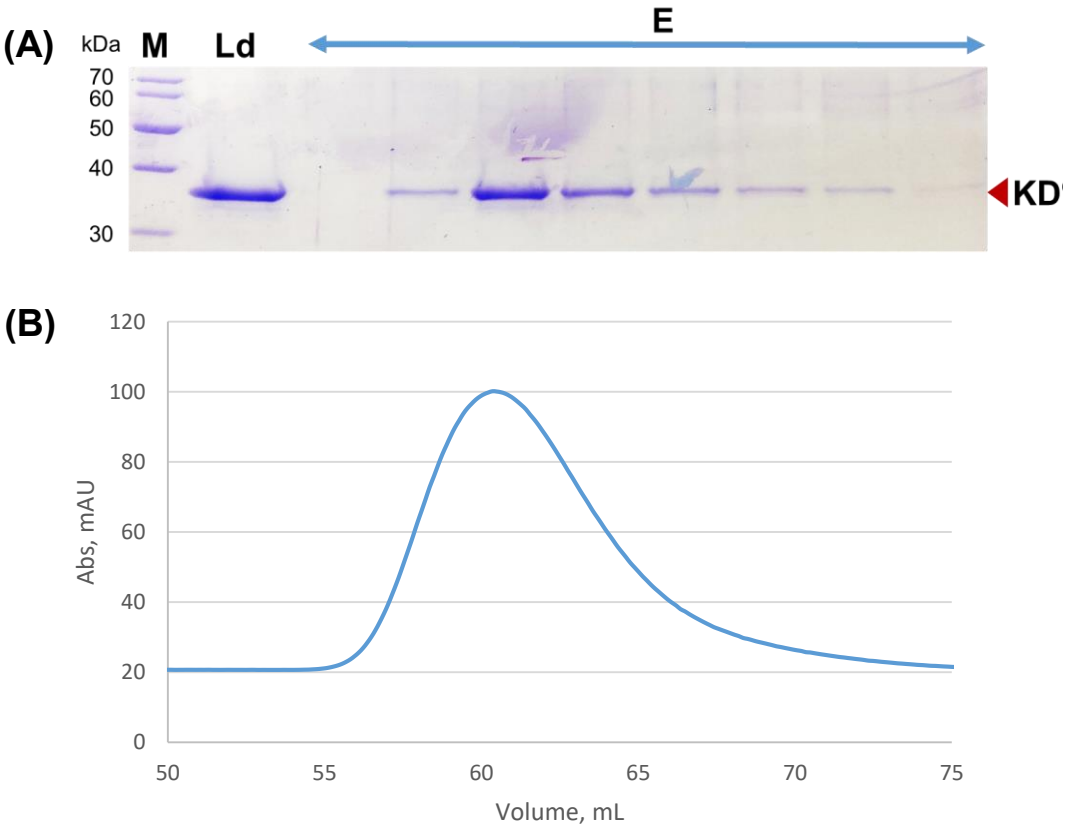


Figure 3.21| SEC purification of AtS6K2 KD_123_414_NC_C-His_N-T7. **(A)** Elution samples of AtS6K2 KD_123_414_NC_C-His_N-T7 (36 kDa) SEC with HiLoad 16/600 Superdex 75 pg column, V=120 mL. M – Protein ladder; Ld – sample loaded on the SEC column. E- samples eluting 55-70 mL were loaded on the SDS-PAGE. AtS6K2 KD_123_414_NC_C-His_N-T7 is highlighted with a red arrow on the side of the gel. **(C)** Chromatogram of SEC.

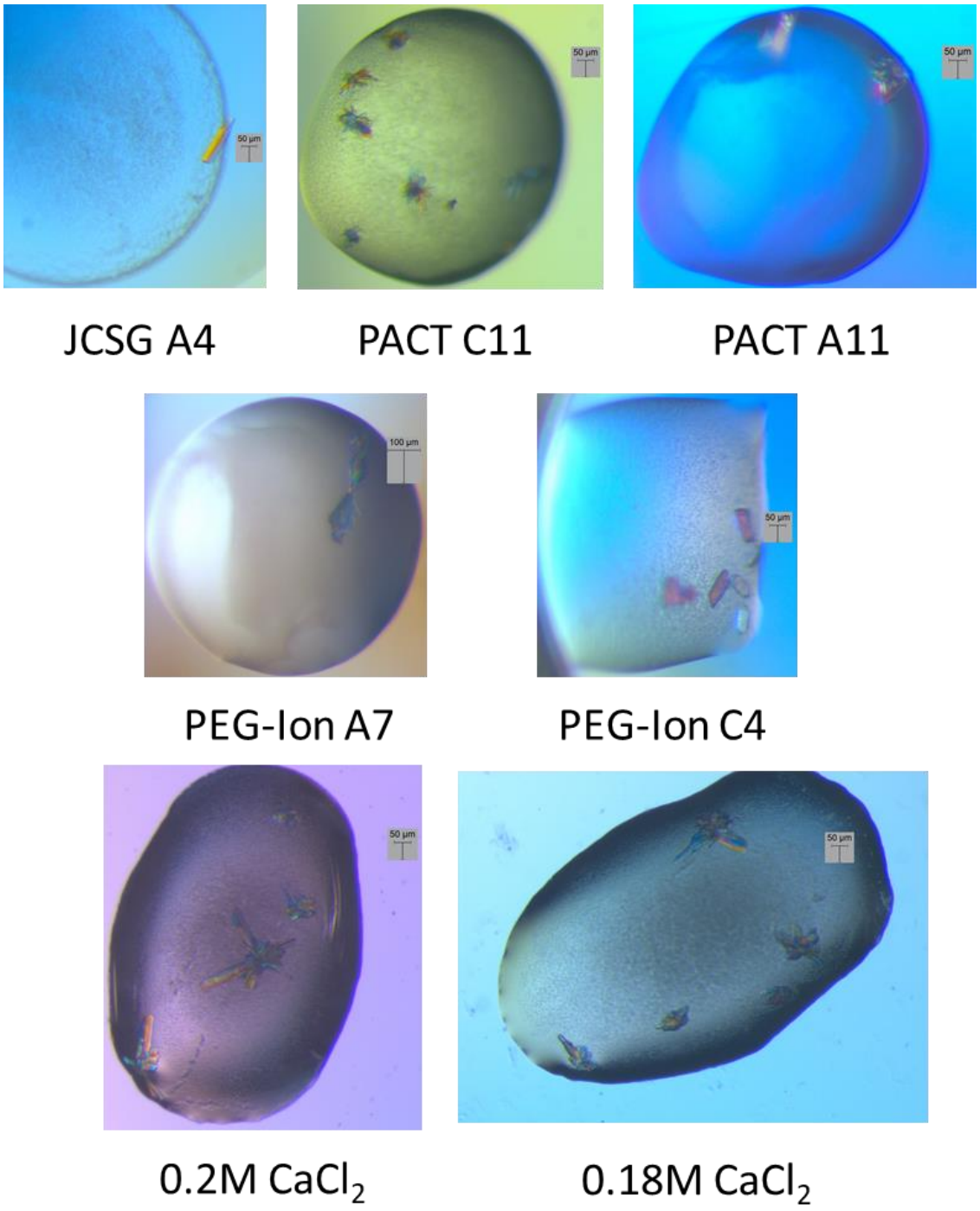


Figure 3.22| Crystallisation of AtS6K2 KD_123_414_NC_C-His_N-T7. Crystals observed in JCSG A4, PACT C11, PACT A11, PEG-Ion A7 and PEG-Ion C4. Relative crystallisation conditions are in Table 3.2. 0.2M and 0.18M $CaCl_2$ represent scaled-up PACT C11 condition containing 22% PEG6000 and 0.1M HEPES Ph=7.0.

3.1.5.3. AtS6K2 KD_123_414_C_C-His_N-T7

I then tested the KD_123_414_C_C-His_N-T7 (molecular weigths: 37 kDa – uncleaved and 35.9 kDa – cleaved) construct with cleavable C-term HisTag and N-term T7 tag (Table 3.1). The protein was successfully expressed in LB (small scale; Fig. 3.18A) and AIM (large scale). After successful IMAC (Fig. 3.23A), elution fraction was incubated with λ-phosphatase and HRV 3C PreScission protease (≈46 kDa) and, similarly to KD_123_414_NC_C-His_N-T7, some precipitation formed during overnight incubation (Fig. 3.23C). The recovered sample was purified via cation exchange chromatography, with the protein eluting at 370 mM Li_2SO_4, consistent with previous purification of the similar construct KD_123_414_NC_C-His_N-T7 (Fig. 3.23B and 3.23D). Collected IE fractions were further purified via SEC which resulted in a double band of apparent molecular weight of ≈35 kDa (Fig. 3.24A and 3.24B). The higher band was much more intense, suggesting that purity of the sample was sufficient for crystallisation (>90-95% homogeneity is usually enough) (Geerlof *et al.*, 2006). Crystallisation screens were set up similarly to KD_123_414_NC_C-His_N-T7, but the 96-well plates were pre-cooled and kept on ice. Screening plates were incubated at 4°C and clusters of crystals appeared after 20 hours (Fig. 24C and 24D) in two conditions: (i) a single cluster with rough edges (0.2 M Calcium chloride dihydrate; 20% w/v Polyethylene glycol 3350; pH=5.1) and (ii) multiple irregularly shaped crystal clusters (0.2 M Calcium chloride dihydrate; 0.1 M HEPES 7.0; 20 % w/v PEG 6000). Unfortunately, these conditions are similar to the ones seen for KD_123_414_NC_C-His_N-T7 and KD_123_414_C_N-GST, suggesting only salt crystals formed also for KD_123_414_C_C-His_N-T7 (Table 3.2).

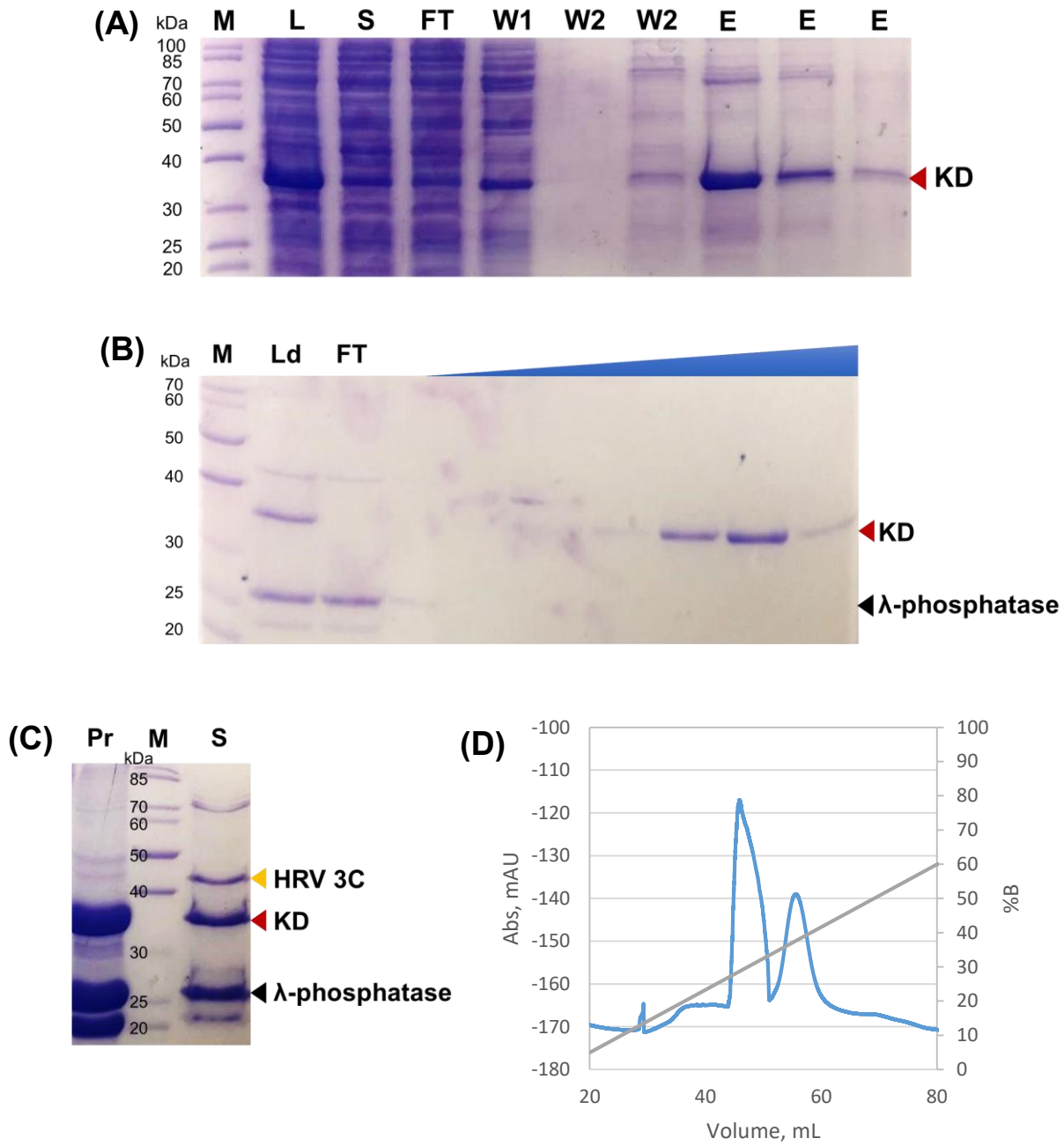


Figure 3.23| Purification of AtS6K2 KD_123_414_C_C-His_N-T7. **(A)** Samples of AtS6K2 KD_123_414_C_C-His_N-T7 (37.1/35.9 kDa) from different IMAC purification steps. M – Protein ladder; L – Cell lysate; S – Supernatant of centrifuged cell lysate; FT – HisTrap column Flow-through; W – HisTrap column wash; E – Elution. **(B)** Samples from ion exchange chromatography (IEC). Blue triangle represents the NaCl gradient, and samples eluting between 260 mM and 440 mM Li_2SO_4 were loaded on the SDS-PAGE. Ld – sample loaded on the cation exchange column; FT – cation exchange column flow-through. **(C)** IMAC elution samples treated ON with λ-phosphatase and HRV 3C PreScission protease were centrifuged and the supernatant (S) and precipitated fraction (Pr) were loaded on the SDS-PAGE. AtS6K2 KD_123_414_C_C-His_N-T7 is highlighted with red arrows on the sides of the gels, λ-phosphatase is highlighted with black arrows and HRV 3C PreScission protease with a yellow arrow. **(D)** Chromatogram of IEC, showing the AtS6K2 KD_123_414_C_C-His_N-T7 elution peak.

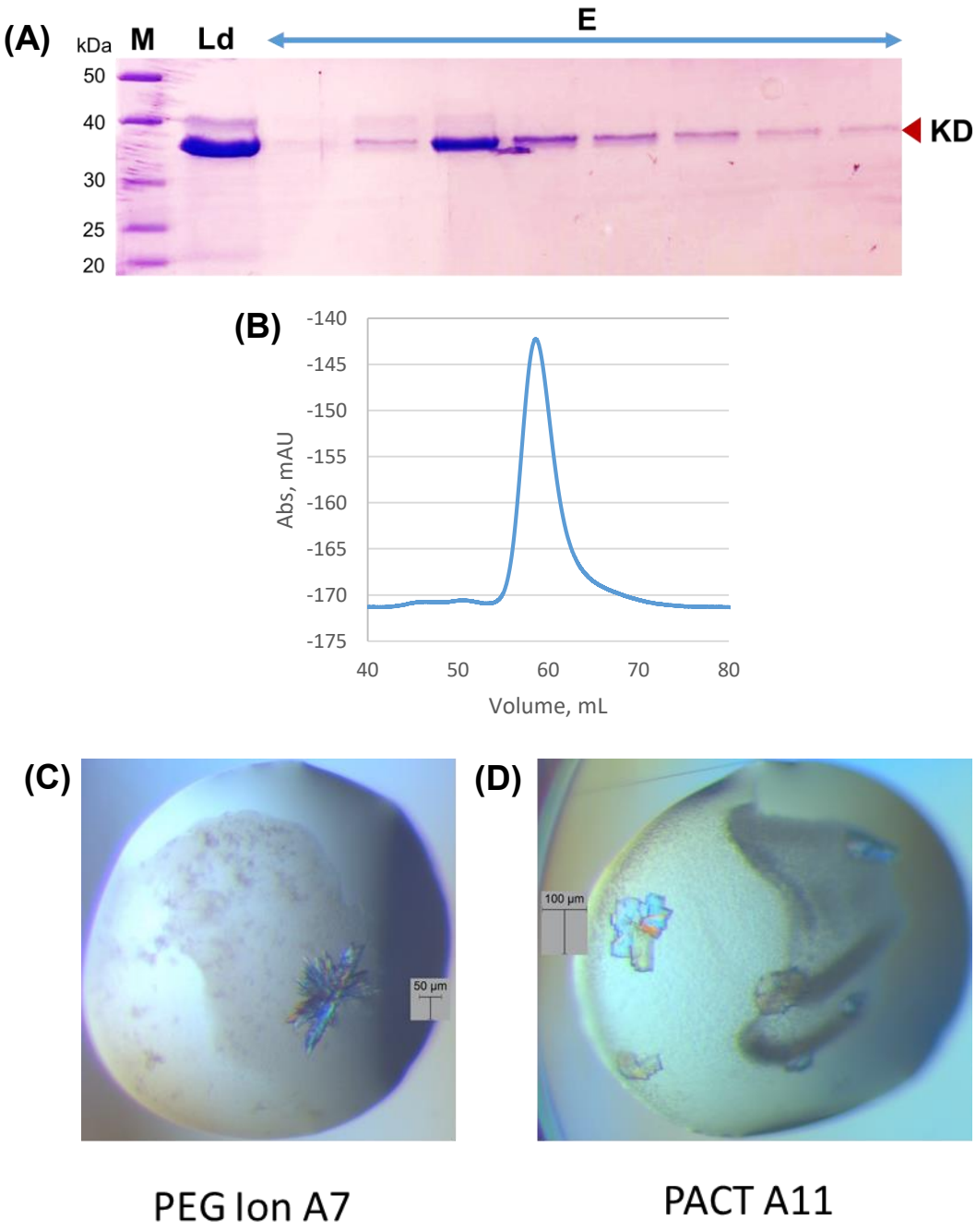
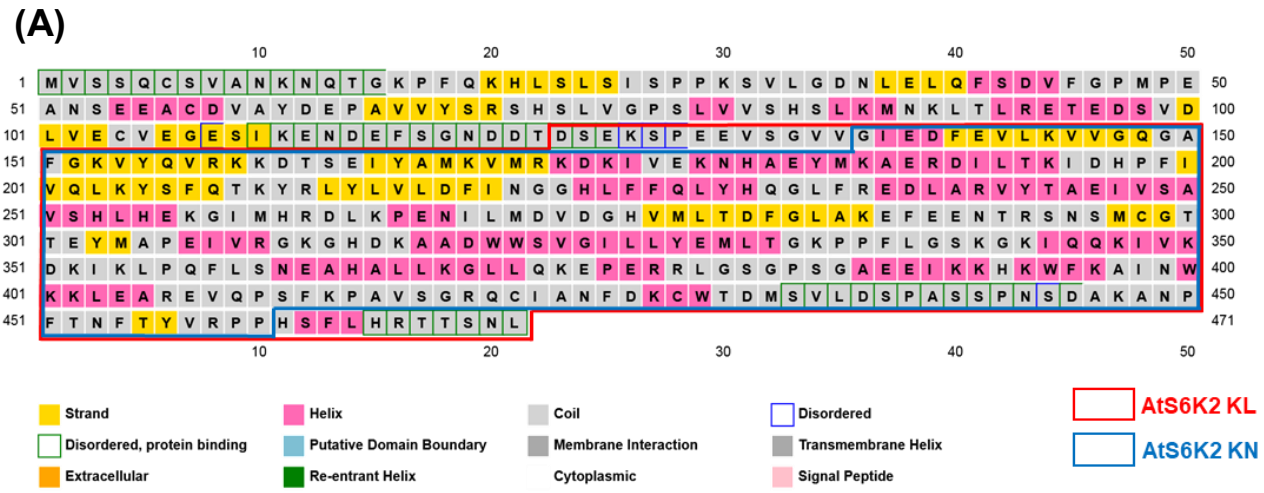


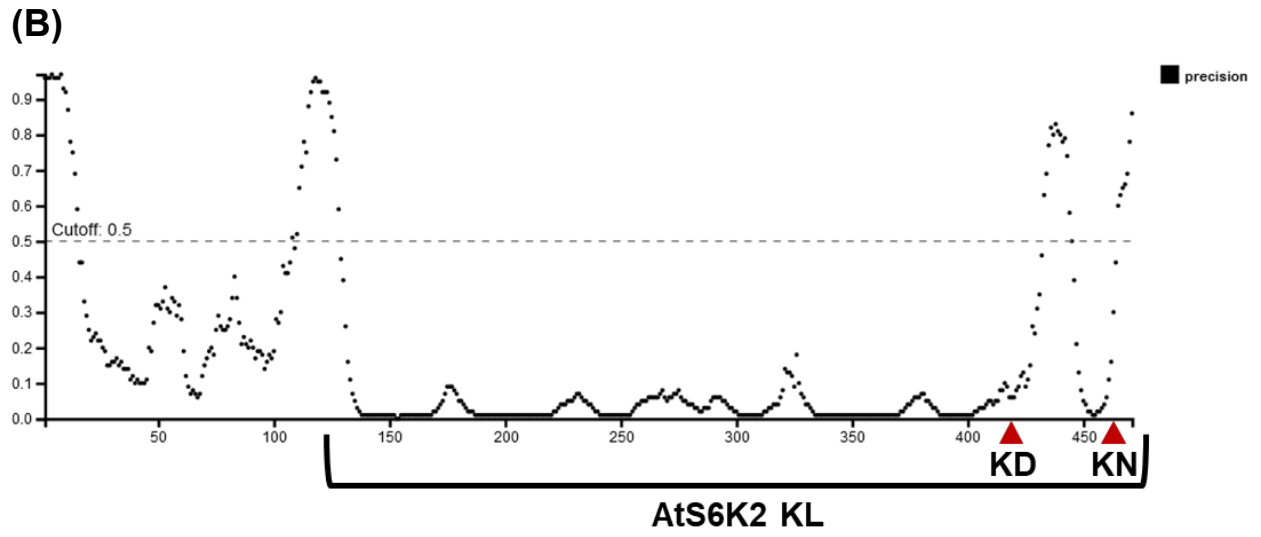
Figure 3.24| SEC purification and crystallisation of AtS6K2 KD_123_414_C_C-His_N-T7. **(A)** Elution Samples of AtS6K2 KD_123_414_C_C-His_N-T7 (35.9 kDa) SEC with HiLoad 16/600 Superdex 75 pg column, V=120 mL. M – Protein ladder; Ld – sample loaded on the SEC column. E – samples eluting 54-70 mL were loaded on the SDS-PAGE. AtS6K2 KD_123_414_C_C-His_N-T7 is highlighted with a red arrow on the side of the gel. **(B)** Chromatogram of SEC. **(C)** Crystals observed in PEG-Ion A7. **(D)** Crystals observed in PACT A11. Relative crystallisation conditions are in Table 3.2.

3.1.6. AtS6K2 KN_136_460_C_C-His

To further optimise AtS6K2 KD/KL constructs, making sure to include residues which are part of the folded domain and remove flexible and unstructured regions, computational modelling of AtS6K2 was performed to assess secondary and tertiary structures. AtS6K2 secondary structure elements (predicted with the PSIPRED online tool) (Buchan and Jones, 2019) corresponded to the well-known kinase topology (Fig. 3.25A). Importantly, the disorder plot showed very clearly that the N-terminal region is disordered (Fig. 3.25B), and likely to be an intrinsically unfolded domain important for protein interaction in line with the role of non-conserved regions in the human S6Ks (Obomighie *et al.*, 2021). Similarly, the C-terminal region is also predicted to be disordered (Fig. 3.25B). However, it is known that the AGC C-terminal domain is structured and binds on the N-lobe of the kinase upon phosphorylation of the HM motif – the reason why this region has been included in all my KL constructs. To further validate this data and design a new construct, I generated a AtS6K2 model using I-TASSER. The best model had a C-score of -1.03, which suggests a good reliability of the prediction, and a TM-score of 0.58, indicating a good overall topology (Zhang, 2008). The reliability of the model was further supported by the ligand binding prediction (docking simulation of ATP), which correctly identified the kinase active site and placed ATP in a pocket localized between the two lobes and which forms hydrogen bonds with the activation loop (where S296 is located). Superimposition of the model with the human S6K1 3D structure (PDB ID: 4L43) allowed to identify the N- and C-termini of the construct in a more reliable way than sequence alignment (Sections 2.1.1 and 3.1.2), especially for low sequence similarity proteins like the human and plant S6Ks (Fig. 3.26A). In agreement with the secondary structure prediction, the AtS6K2 construct encompassing G136-P460 most closely resembles the one used for S6K1 crystallisation experiments, which contains all required elements for folding and structural stability. Importantly, based

on the structural analysis, the KD and KL constructs I designed based on the sequence alignment with human S6K1 (Section 3.1.2), which start from residue D123 and end at residues P414 and L471, respectively, contain a relatively long flexible segment at both the N-and C-term which could have played a detrimental role in the constructs' solubility. The optimised construct, AtS6K2 KN_136_460_C_C-His, spanning residues 136-460 and containing the two phospho-mimetic mutations S296E, T455E, includes an HRV 3C PreScission protease cleavage site and a C-terminal His-tag. KN_136_460_C_C-His was expressed in LB media, however cell lysis in reducing conditions (5 mM β-mercaptoethanol) and further IMAC purification revealed that, similarly to the AtS6K2 KL_123_471_NC_C-His S296E_T455E mutant, the protein was insoluble (Fig. 3.25C).





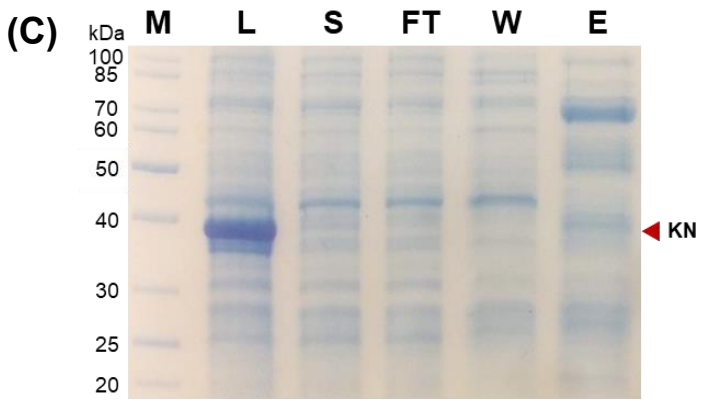


Figure 3.25| AtS6K2 secondary structure prediction and purification of AtS6K2 KN_136_460_C_C-His. **(A)** PSIPRED prediction of full-length AtS6K2. AtS6K2 KL (123-471) is highlighted with a red box and newly computed KN_136_460_C_C-His (136-460) is highlighted with a blue box. **(B)** DISOPRED prediction of full-length AtS6K2. AtS6K2 KD (123-414) and KN are highlighted with red arrows. AtS6K2 KL is also shown. **(C)** Samples of AtS6K2 KN_136_460_C_C-His (39.6 kDa) from different IMAC purification steps. M – Protein ladder; L – Cell lysate; S – Supernatant of centrifuged cell lysate; FT – HisTrap column Flow-through; W – HisTrap column wash; E – Elution. AtS6K2 KN_136_460_C_C-His is highlighted with red arrow on the side of the gel.

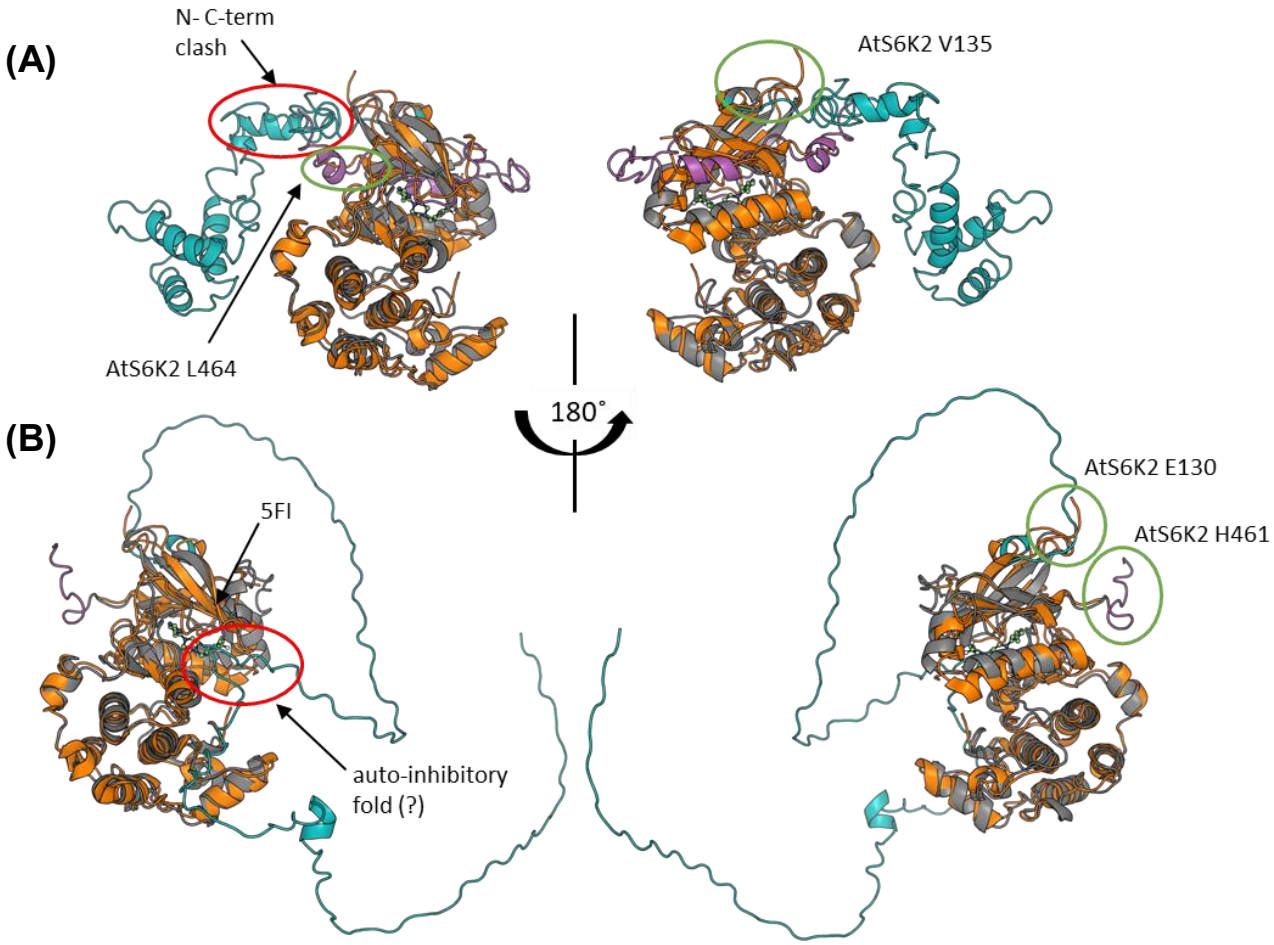


Figure 3.26| AtS6K2 I-TASSER and AlphaFold models. Overposition of AtS6K2 I-TASSER model **(A)** or alpha fold model **(B)** with human S6K1 crystal structure (PDB ID: 4L43), both shown in cartoon representation. The human S6K1 structure is in orange, while AtS6K2 not conserved N-term is in teal, the kinase domain in grey, and the not conserved C-term in violet. Circled in green the N- and C-termini of the S6K1 structure, while in red is circled the AtS6K2 model structural issues. Structures in the two panels are all in the same orientation and same scale.

3.1.7. AlphaFold model and AtS6K2 KDN

Despite the AtS6K2 model I generated for the catalytic domain is consistent with known kinase topology and functional data (Section 3.1.6), I-TASSER failed to recognise flexible and unstructured domains/regions of the protein. From a close inspection, it is possible to note that the N-terminal region has a well-defined 3D structure, not resembling any known structural domain (no match found on PDBeFOLD). Importantly, a noticeable clash between the N- and C-terminal of the protein is visible (Fig. 3.26B), which suggest this may be an artefact resulted by an attempt of I-TASSER to predict folding of all parts of the protein. This lack of reliability on the structure of the N- and C-term of the model may have resulted in a poor re-design of the expression constructs for AtS6K2. With the release of AlphaFold models, showing a much higher overall accuracy in structural prediction (Jumper *et al.*, 2021), I redesigned the AtS6K2 expression construct based on structural alignment with other AGC-kinases (Fig. 3.26B). The catalytic domain shows a classic kinase fold, with high confidence (pLDDT) scores, and it is overall very similar to the I-TASSER prediction, with the all-atom RSMD of 0.95 Å for the kinase domains only. Differently, the N-terminal region in the AlphaFold model is completely unstructured but has low confidence scores. It is interesting to note that residues 73-78 are located inside the active site, mimicking an auto-inhibitory conformation. While it is known that human S6K1 and S6K2 have an autoinhibitory C-terminal domain, which goes through a conformational change upon phosphorylation by mTORC1, the N-term of AtS6K2 and the C-term of both human S6Ks have no sequence similarity. Furthermore, there are no known phosphorylation sites on the N-terminal of AtS6K2 which would suggest presence of an autoinhibitory motif regulated during protein activation, as seen in other human kinases regulated via auto-inhibitory mechanisms. Despite the low confidence of this part of the model, it may still be possible to perform a structural alignment and obtain reliable

information. Based on the alignment with human S6K1 (PDB ID: 4L43 and 4L3I) (Wang *et al.*, 2013), PKCβ2 (PDB ID: 2I0E) (Grodsky *et al.*, 2006), AKT1 (PDB ID: 4EKK) (Lin *et al.*, 2012), AKT2 (PDB ID: 1O6K) (Yang *et al.*, 2002) and mouse PKA (PDB ID 1ATP) (Zheng *et al.*, 1993), a new KDN construct was designed, spanning residues G133-P414. This was cloned into a pETDuet vector which allow the co-expression of AtS6K2 with λ-phosphatase. The KDN sequence with C-terminal His-tag, both cleavable and uncleavable, was cloned into pETDuet-1 vector, generating KDN_133_414_NC_C-His_λ and KDN_133_414_C_C-His_λ constructs.

3.1.7.1. AtS6K2 KDN_133_414_NC_C-His_λ

AtS6K2 KDN_133_414_NC_C-His_λ (Table 3.1) was expressed in LB media and purified with IMAC, IEC and SEC. Due to no improvement in solubility and stability of the protein in using Li_2SO_4, and considering issues with crystallisation, I reverted to using buffer containing NaCl. Strong overexpression bands of ≈33 kDa and 25 kDa indicated successful expression of both KDN_133_414_NC_C-His_λ (33.6 kDa) and λ-phosphatase (25.2 kDa) using a pETDuet-1 vector. λ-phosphatase was removed during the HisTrap column wash steps and KDN_133_414_NC_C-His_λ was present in the HisTrap column elution. However, the fractions containing the largest amount of protein precipitated during overnight storage at 4°C. After centrifuging the precipitate, a reasonable amount of KDN_133_414_NC_C-His_λ was left in the supernatant (Fig. 3.27A), which was further purified via cation exchange chromatography. A peak at 430 mM NaCl appeared (Fig. 3.27B and 3.27C). The collected IEC fractions were further purified via analytical SEC, which resulted in a double peak at 11.2 mL and 12 mL. KDN_133_414_NC_C-His_λ was present in the first peak (Fig. 3.28). Pure protein was produced but the yield was too low (≈70 µg) to attempt crystallisation screening experiments. Further optimization of expression and/or purification conditions is required for structural studies.

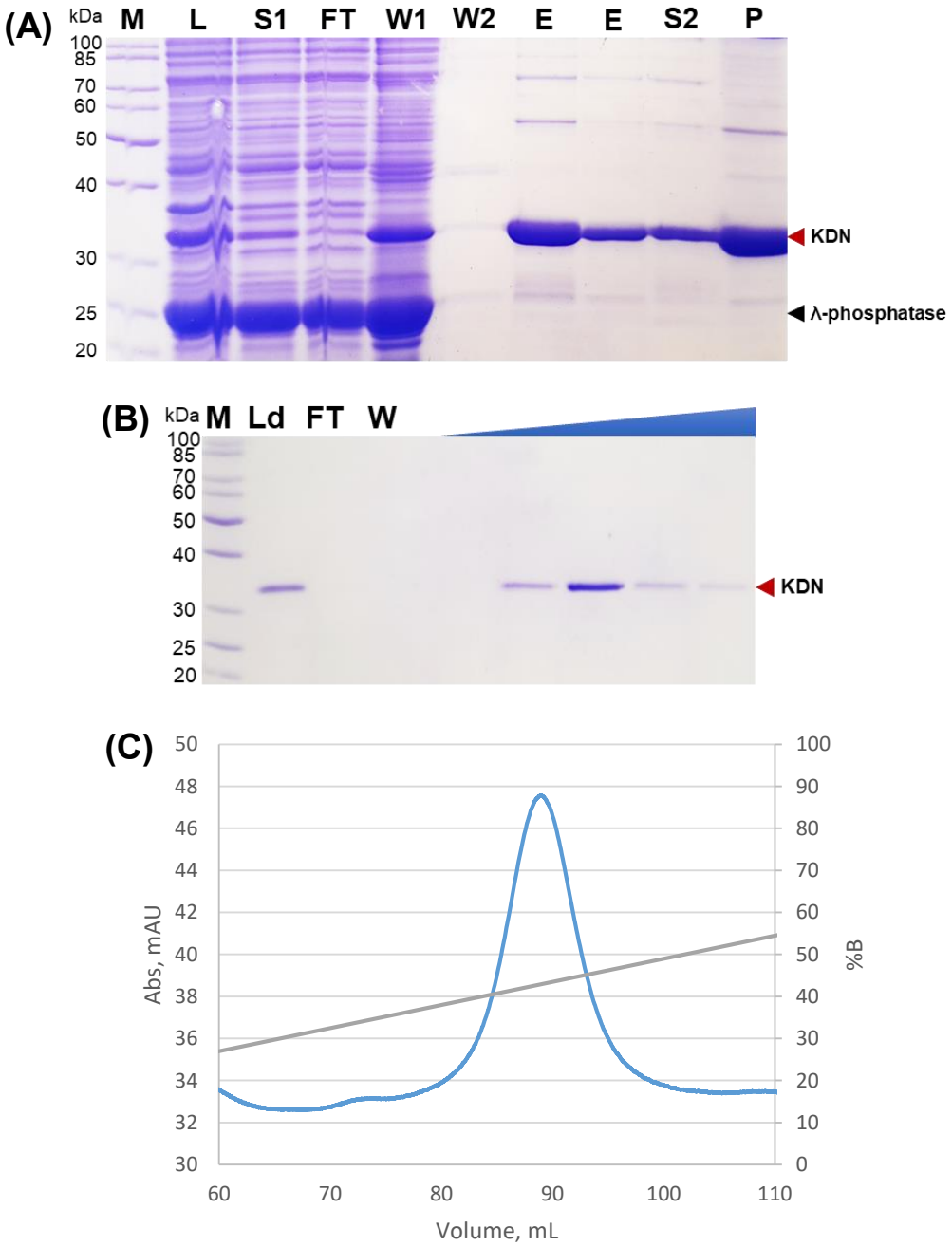


Figure 3.27| Purification of AtS6K2 KDN_133_414_NC_C-His_λ . **(A)** Samples of AtS6K2 KDN_133_414_NC_C-His_λ (33.6 kDa) from different IMAC purification steps. M – Protein ladder; L – Cell lysate; S1 – Supernatant of centrifuged cell lysate; FT – HisTrap column Flow-through; W – HisTrap column wash; E – Elution. Precipitated IMAC elution samples were centrifuged and the supernatant (S2) and precipitated fraction (P) were loaded on the SDS-PAGE. **(B)** Samples from ion exchange chromatography (IEC). Blue triangle represents the NaCl gradient, and samples eluting between 350 mM and 490 mM NaCl were loaded on the SDS-PAGE. Ld – sample loaded on the cation exchange column; FT – cation exchange column flow-through. W – 50 mM NaCl column wash. AtS6K2 KDN_133_414_NC_C-His_λ is highlighted with red arrows on the sides of the gels, λ-phosphatase is highlighted with a black arrow. **(C)** Chromatogram of IEC.

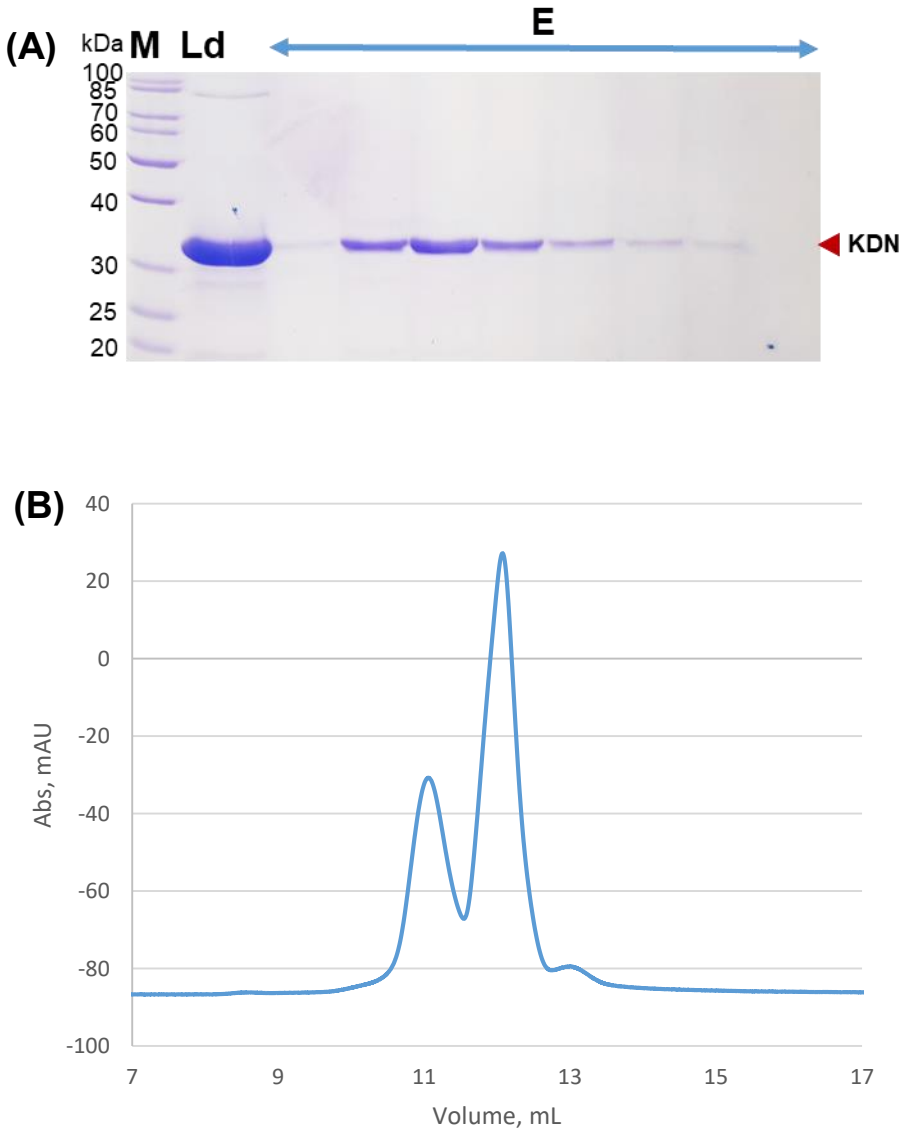


Figure 3.28| SEC purification of AtS6K2 KDN_133_414_NC_C-His_λ. **(A)** Elution Samples of AtS6K2 KDN_133_414_NC_C-His_λ (33.6 kDa) SEC. M – Protein ladder; Ld – sample loaded on the SEC column. E – samples eluting 10-13 mL were loaded on the SDS-PAGE. AtS6K2 KDN_133_414_NC_C-His_λ is highlighted with a red arrow on the side of the gel. **(B)** Chromatogram of SEC with TSKgel G3000SW$_{XL}$ HPLC column.

3.1.7.2. AtS6K2 KDN_133_414_C_C-His_λ

AtS6K2 KDN_133_414_C_C-His_λ (Table 3.1) that contains an HRV 3C PreScission protease site between the protein and C-term His-tag was also tested. Large scale expression in LB media was successful (250 µM IPTG when OD_{600}=0.5, 18°C overnight) for both KDN_133_414_C_C-His_λ (34.7 kDa) and λ-phosphatase. Cell lysis under reducing conditions (5µM β-mercaptoethanol) revealed that only a fraction of KDN_133_414_C_C-His_λ was soluble. IMAC column elution contained the ≈34 kDa band (Fig. 3.29). Unfortunately, similarly to KDN_133_414_NC_C-His_λ, excessive precipitation reduced yields, resulting in an extremely limited amount of soluble pure protein. While especially promising, further optimization of expression and/or purification conditions is required for structural studies.

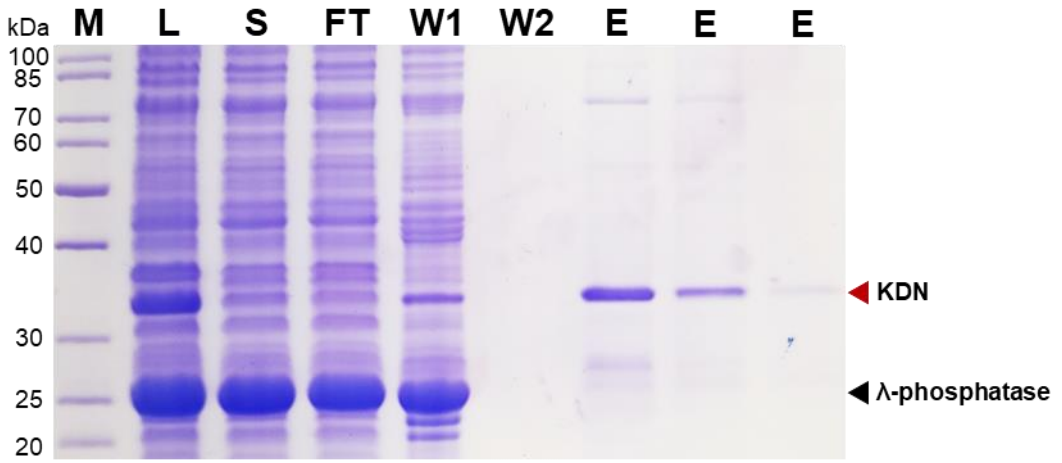


Figure 3.29| IMAC purification of AtS6K2 KDN_133_414_C_C-His_λ. Samples of AtS6K2 KDN_133_414_C_C-His_λ (34.7 kDa) from different IMAC purification steps. M – Protein ladder; L – Cell lysate; S – Supernatant of centrifuged cell lysate; FT – HisTrap column Flow-through; W – HisTrap column wash; E – Elution. AtS6K2 KDN_133_414_C_C-His_λ is highlighted with a red arrow on the side of the gel, while λ-phosphatase with a black arrow.

3.2. Discussion

In order to produce soluble AtS6K2 for crystallisation studies, several factors were taken into account, including the construct length, type of tag and position of tag, vector used for recombinant expression, expression temperature, host strain, expression media and duration of expression (Table 2.1). Production of recombinant proteins with an N-terminal His-tag was not successful. Only a single construct with N-terminal GST tag was purified (KD_123_414_C_N-GST). The C-terminal His-tag constructs (KD_123_414_NC_C-His_N-T7; KD_123_414_C_C-His_N-T7; KL_123_471_NC_ C-His_N-T7; KL_123_471_NC_C-His; KDN_133_414_NC_C-His_λ and KDN_133_414_C_C-His_λ) showed a higher degree of success.

3.2.1. Protein purification tags

Although several soluble truncated variants of AtS6K2 were possesing hexahistidine and GST tags were produced and purifed, crystallisation failed. Fusion tags can improve protein expression, promote folding and increase protein solubility, however, the effect of a specific tag will vary depending on the protein sequence. As such, it may be necessary to empirically test multiple tags (Peti and Page, 2007). The His-tag exerts low metabolic burden on expression hosts and allows a wide range of purification conditions, i.e. mild, denaturing and nondenaturing (Waugh, 2005). The His-tag is the most popular tag for recombinant protein production but a variety of other affinity tags has been developed, such as maltose-binding protein (MBP) tag (Nallamsetty and Waugh, 2007), calmodulin-binding peptide (CBP) tag, intein-chitin binding domain (intein-CBD) tag, FLAG-tag, protein G β1 domain (GB1) tag, protein A IgG ZZ repeat domain (ZZ) tag and others (Kimple, Brill and Pasker, 2013). Additional strategies that exploit tags intrinsic properties to enhance

expression and/or solubility include Thioredoxin (Trx) (Hammarström *et al.*, 2002), N-utilization substance A (NusA) (Davis *et al.*, 1999) and small ubiquitin-like modifier (SUMO) tag (Malakhov *et al.*, 2004), all usually coupled with an His-tag for affinity putification. A number of comparative studies have examined the effects of fusion partner on protein expression. 27 human proteins of small molecular weight (>20 kDa) were cloned into expression vectors with various tags and ranked according to the tags ability to promote expression of soluble protein: Trx ≈ MBP ≈ Gb1 > ZZ > NusA > GST > His_6 (Hammarström *et al.*, 2002). Differently, when comparing protein expression and yield after IMAC purification of 32 human proteins with a molecular weight similar to that of the AtS6K2 constructs I cloned (17-110 kDa), the tags were ranked as GST ≈ MBP > CBP > His_6 (Braun *et al.*, 2002). This is the reason why GST tag was selected as fusion protein tag for initial AtS6K2 purifications, done in parallel to the His-tag experiments. Another study of 40 proteins, including 4 plant proteins, distinguished MBP as the best overall expression and solubility enhancer, while NusA was also largely successful (Shih *et al.*, 2002). Unfortunately, these studies used different combinations of promoters and tags, resulting in variable rates of transcript synthesis and translation rates, making it difficult to compare different proteins expression and solubility experiments. Dyson *et al.*, (2004) used expression vectors with T7 promoters to assess the expression and solubility of 95 mammalian proteins. MBP and Trx tags coupled with the decahistidine tag were found to be most effective in promoting protein solubility. A range of new tags that could potentially improve the yield of soluble AtS6K2 and be used to produce the protein for structural studies, but that requires developing additional cloning, expression and purification strategies, which, considering the vast range of constructs and condition described in this thesis, is unlikely to improve yealds of protein amenable for structural studies.

In addition to the affinity tags, several AtS6K2 constructs contained the T7 epitope tag on their N-termini (Table 3.1). The T7 epitope tag (MASMTGGQQMGRGS) encodes the first 12 amino acids of phage T7 major capsid protein (gene 10), often used for immune purification and blotting (Enomoto, Chen and Berman, 1998; Jarvik and Telmer, 1998). Despite the fact that examples of crystal structures of fusion proteins with N-terminal T7 tag are available (Homma and Moriyama, 2009), its presence in AtS6K2 constructs was negligible (both KL_123_471_NC_C-His_N-T7 and KL_123_471_NC_C-His constructs were expressed and purified).

3.2.2. Tag position

The yield of soluble protein can be heavily influenced by the placement of a tag at either N- or C-terminus (Francis and Page, 2010). In this thesis I described a wide range of studies exploring the benefits of N- or C-terminus His-tag. Position of His-tag within a specific protein has been shown to influence expression, solubility and bioactivity. However, optimal location of His-tag has to be de determined experimentally (Dan, Balachandran and Lin, 2009). Woestenenk *et al.*, (2004) examined expression of 20 human disease-related proteins in E. coli with N- or C-term His-tags. Production yields of proteins with C-term His-tags were found to be the highest after IMAC purification, partly due to fast growth rates of cell cultures. However, cell growth did not correlate with protein expression level, solubility and tertiary structure. Cell density, which is a determinant of protein yield, varies depending on a particular protein expressed (Woestenenk *et al.*, 2004). Proteins act as individuals and are difficult to generalise. There are numerous studies to date that explore expression and solubility of proteins with specific tags, however sequence differences near the start codons of affinity tags and differences in vector backbones make it difficult to compare the studies and it is suggested that multiple

constructs should be tested with each protein sequence (Woestenenk *et al.*, 2004). Contradictory results were presented in a high-throughput expression study showing that C-His tags more frequently have a negative effect on solubility than N-His tags and location of the tag also influences the expression level (Busso *et al*., 2003). However, differently from the experiments described in this thesis, Busso et al. (2003) used a E. coli cell-free system. The mechanism that explains why protein expression is influenced by the position of His-tag is not well understood but it is likely dependent on a specific protein sequence. Although comparative high-throughput studies can elucidate the average effect of a particular tag, proteins act as individuals and generalisations are difficult to make (Dan, Balachandran and Lin, 2009). For the AtS6K2 constructs used, moving His-tag from N- to C-terminus proved to be crucial in order to produce soluble protein. Further data is also available for His-fusion protein-tag combos. Specifically, earlier studies by Dyson *et al.*, (2004) have investigated extensively the different role N- or C-terminal tags in protein expression and solubility. On average, tag placement at the N-terminus was found to be favourable for optimal protein expression. However, this study compared expression of proteins fused to large solubility enhancing tags combined with His-tags (Dyson *et al.*, 2004), while I only used single His- or GST-tags in my experimental work.

3.2.3. Tag cleavage

Fusion protein tags (i.e., MBP – 43 kDa, GST – 26 kDa, Trx – 15 kDa) can be left uncleaved for crystallization trials, but that usually reduces the chance to form well-ordered, diffracting crystals due to conformational heterogeneity allowed by the flexible linker region (Smyth *et al.*, 2003). Although it is possible to solve crystal structures of proteins with uncleaved large size tags, only a few such structures have been solved. In these cases, the fusion junction was rigidified by fusing MBP via a short 3-5 amino acid

spacer that allowed production of high quality crystals required for X-ray diffraction experiments (Smyth *et al.*, 2003). In addition, any tag independent of the size may alter the native structure and activity of the protein or modify its behaviour in some other way (Francis and Page, 2010). For these reasons a protease cleavage site between the tag and the target protein is normally engineered. The tag can then be removed using site-specific proteolysis through an in vitro reaction (Esposito and Chatterjee, 2006)**.** Highly specific proteases such as PreScission (LEVLFQ|GP) have determinants that are located on both termini of the scissile bond (|), which will stay on the target protein after the cleavage of the tag. Consequently, endoproteolytic removal of an N-terminal tag will leave 2 extra residues (GP) on the N-terminus of the target protein and the removal of a C-terminal tag will leave 6 residues on the C-terminus of the target protein (LEVLFQ) (Waugh, 2005). Addition issues may arise during the cleavage step as the reaction may result in a low yield of cleaved protein or induce precipitation requiring optimisation reaction conditions.

Differently, His-tag has less stringent requirments, as it is not necessary to remove the tag for protein structure determination as it is known to have minimal effect on native protein structure and function due to its relatively small size and charge (Dan, Balachandran and Lin, 2009). Statistical analysis of 13.048 available structures with (11.906) and without His-tags (1.142) revealed that there is no difference in *R* factors and resolution between the two groups while *B* factors and solvent contents are slightly higher for the tagged structures (Carson *et al.*, 2007). While His-tag may even be helpful to protein crystallisation (Tajika *et al.*, 2004), some crystallographers believe these tags have to be cleaved (Derewenda, 2004; Waugh, 2005). Indeed, all constructs I generated had high level of flexibility, with the option to cleave or keep tags.

3.2.4. Dephosphorylation

Protein kinases are known to auto-phosphorylate when overexpressed in *E. coli* resulting in heterogeneous protein samples with increased number of phosphorylated states that are not biologically relevant. Very few studies have explored this, but earlied data by Shrestha *et al.*, (2012) supported hypothesis that such phosphorylations may be introduced during protein folding, by an overexpressed kinase. These non-physiological hyper-phosphorylation sites are not accessible as substrate sites in folded proteins and cannot be completely dephosphorylated by *in vitro* phosphatase treatment or reproduced by *in vitro* auto-phosphorylation. Indeed, this was the case for AtS6K2, as Mass Spectrometry data surprisingly revealed presence of 3 phosphorylations, none known to be auto-phosphorylation sites, and 2 of which not known to be phosphorylated in vivo at all (Mahfouz *et al.*, 2006b; Xiong and Sheen, 2012).

The non-physiological phosphorylation sites in AtS6K2 were mapped on a flexible region at the C-term of the construct. Indeed, I was able to produce AtS6K2 in a soluble form without dephosphorylation, despite the presence of non-biologically relevant phosphorylations (Section 3.1.4.4). These samples were, however, highly non homogeneous and potentially unstable, as they all failed to crystallise or stay soluble during purification. To produce homogeneous sample suitable for structural studies, AtS6K2 had to be dephosphorylated. *In vitro* phosphatase treatment (Ali *et al.*, 2011; Prischi *et al.*, 2014) was tested, but results were non optimal or difficult to reproduce. This may suggest that phosphorylations had a detrimental role in protein folding. In these cases, it has been shown that co-expression with phosphatase is often required (Seeliger *et al.*, 2005; Filippakopoulos *et al.*, 2008) as it removes phosphorylations, including those at sites that disrupt protein folding. thus, co-expression with λ-phosphatase were tested with various AtS6K2 constructs. The most promising was the optimised KDN construct that

is co-expressed with λ-phosphatase, and future studies will focus on repeating the experiments done at the end of my PhD and proceed to test different crystallisation conditions.

3.2.5. Expression systems

Production of proteins for structural studies depends on the success of three factors: expression, solubility and purification. *E. coli* recombinant protein expression was chosen as it is faster, cheaper and yields of protein produced are higher. Kinases are known to be particularly challenging to express as recombinant proteins in *E. coli* and many strategies have been proposed over the years to facilitate protein production (Esposito and Chatterjee, 2006)**.**

There are a few reasons why expression of eukaryotic proteins is complicated in *E. coli*: (i) different requirements of chaperones and chaperonins. Unfortunately, there are no commercially available *E. coli* strains for co-expression with plant chaperons. The ArcticExpress (DE3) Competent Cells producing chaperonins Cpn60 and Cpn10 failed to transform. (ii) Effective disulphide bond formation is challenging in the reducing environment of bacterial cytoplasm. However, there are no disulphide bonds in AtS6K2 according to AlphaFold model (all Cys are reduced and far from each other). (iii) A lack of machinery to perform several eukaryotic post-translational modifications. However, there are no known glycosylations or other PTMs that are required for folding of AtS6K2 (Francis and Page, 2010).

As such, the most likely option is to abandon the *E. coli* and move to an eukaryotic expression system: (i) insect cell lines or (ii) tobacco leaves.

Baculovirus expression system in insect cells is well established to produce soluble and active kinases (Pengelley *et al.*, 2006; Prischi *et al.*, 2014). Baculoviruses are dsDNA

viruses, majority of which infect the Lepidoptera order insects, and the most popular baculovirus for protein expression is the Autographa californica multiple nucleopolyhedrovirus (AcMNPV). The most widely used expression host cell lines are Sf9 or the Sf21 from Spodoptera frugiperda and High Five cells (BTI-Tn-5B1-4) from Trichoplusia ni (Vaughn *et al.*, 1977; Granados *et al.*, 1994). Initially, a recombinant baculovirus containing the gene of interest is constructed. Alternatively, bacmid technology can be used, which is faster, but bacmid-derived baculoviruses are less stable. Insect cells are then infected with baculoviruses, and incubated for protein expression (Felberbaum, 2015). Baculovirus expression in Sf21 cells were successfully used for structural studies of RSK1 N-terminal kinase domain (Ikuta *et al.*, 2007) and Sf9 cells were used to produce S6K1 (Wang *et al.*, 2013). However, the baculovirus expression process is expensive and time consuming. Using time-saving bacmid technology takes 6 to 8 weeks from cloning to protein production (up to 11 weeks if recombinant baculoviruses are generated with conventional method) (Mahajan *et al.*, 2014). This is 2 to 3 times longer than *E. coli* expression and, in average, yields are 10 times lower per litre of cells, hence my choice to use bacterial expression system for my project, despite the clear technical challenges extensively described in the results (Burgess-Brown *et al.*, 2014).

The most physiological ways to express recombinant plant proteins relay on (i) development of stable transgenic plants, where the gene of interest is cloned into an expression cassette and introduced into the nuclear or chloroplast genome; and (ii) transient expression, which uses the virus *Agrobacterium tumefaciens*-derived vectors containing the gene of interest. Agrobacterium are then introduced to fully developed plants through agroinfiltration (Leuzinger *et al.*, 2013). Transient expression is usually preferred as it allows production of high amounts of protein in a shorter timescale compared to transgenic plants. Agrobacterium is transformed with either a viral replicon

vector that contains gene of interest or a binary vector system for co-expression with genes known to stimulate cell cycle progression or suppress stress response and gene silencing (Norkunas *et al.*, 2018). Recombinant agrobacteria are introduced into the extracellular leaf spaces by either syringe or vacuum infiltration. The syringe method is simple, does not require expensive equipment and allows infiltration of a single leaf with either one target gene or multiple genes. While syringe infiltration is usually acceptable for research purposes, the vacuum infiltration is more robust and can be scaled-up for commercial manufacturing of recombinant proteins (Leuzinger *et al.*, 2013). Nicotiana benthamiana, a close relative of tobacco from Australia, is a widely used host for transient expression because of its (i) fast growth rate, (ii) ability to express high levels of heterologous proteins and (iii) amenability to transformation (Sheludko *et al.*, 2007). A wide range of options to optimise protein expression in N. benthamiana is available, as cultural conditions significantly impact the yield of recombinant protein (Goulet *et al.*, 2019). Producing AtS6K2 using transient expression in N. benthamiana may be advantageous, but similarly to baculovirus expression system, such system is more expensive and time consuming than expression in E. coli. It is however the most promising one, and it has been chosen for follow-up studies linked to my project.

4. AtS6K2 FL activity

4.1. Results

In vivo studies in A. thaliana showed that AtS6K2 FL WT and T455E, S296E mutant generate different phenotypes (Fig. 4.1 and unpublished data; received from Dr Irabonosi Obomighie). Specifically, AtS6K2 T455E_S296E transgenic plants had increased resistance to cold and salt stresses. This increased resistance was linked to an increased activity and not an increased overexpression of the protein in plants. However, a confirmation that AtS6K2 T455E_S296E has increased activity compared to the WT was needed. As such, attempts to purify recombinant AtS6K2 FL WT and T455E, S296E proteins were carried out for *in vitro* kinase activity measurements.

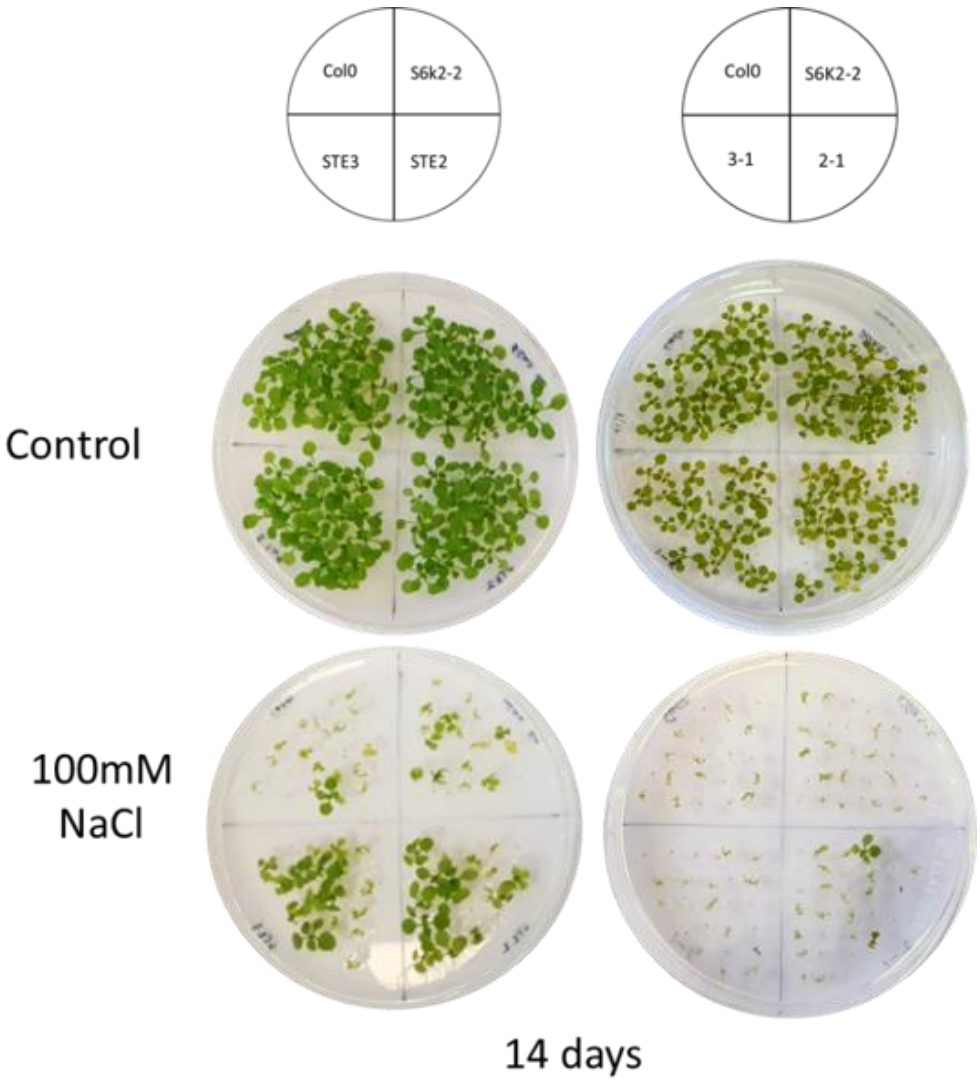


Figure 4.1| Plant growth under salt stress. MS plates containing different genotypes under control and salt stress (100mM NaCl) conditions.

4.1.1. AtS6K2 FL_1_471_NC_C-His_N-T7

Initially, full length AtS6K2 was cloned into the pET-21a+ vector, generating AtS6K2 FL_1_471_NC_C-His_N-T7 construct, that has N-term T7 tag and an uncleavable C-term His-tag. Small scale test expression of FL_1_471_NC_C-His_N-T7 (55.5 kDa; Table 3.1) resulted in a strong overexpression band of expected size (Fig. 4.2A). Large scale expression in AIM yielded similar results and cell lysis indicated that the protein was soluble (Fig. 3.8C). As discussed above (Section 3.1.4.1), site directed mutagenesis was used to produce double phosphomimetic mutant (AtS6K2 FL_1_471_NC_C-His_N-T7 T455E, S296E). AtS6K2 FL_1_471_NC_C-His_N-T7 proteins were purified by IMAC (Table A12, Fig. 4.2B and 4.3A). Due to low yields, samples were not purified further, but simply dialysed in activity assay buffer and used directly for kinase activity measurements. Indeed, the presence of contaminants will not interfere with the activity assay, as the HTRF KinEASE kit can be used directly on cell extracts (Drexler, 2006). 5 µM wild type AtS6K2 and T455E, S296E mutant concentrations were used for the HTRF KinEASE STK discovery kit (Section 2.5). The concentration was increased to 10 and 20 µM but no activity was detected in any of the conditions tested. This unexpected lack of activity suggested a check of folding/aggregation of AtS6K2 FL constructs by analytical SEC. Unfortunately, chromatograms (Fig. 4.2C and 4.3B) showed the presence of a single peak around 5 mL, which corresponds to the void volume of the column, for both WT and double mutant, indicating that proteins were aggregated. This may be due to non-specific hyperphosphorylation, as observed for AtS6K2 KL_123_471_NC_C-His_N-T7 construct (Section 3.1.4.4).

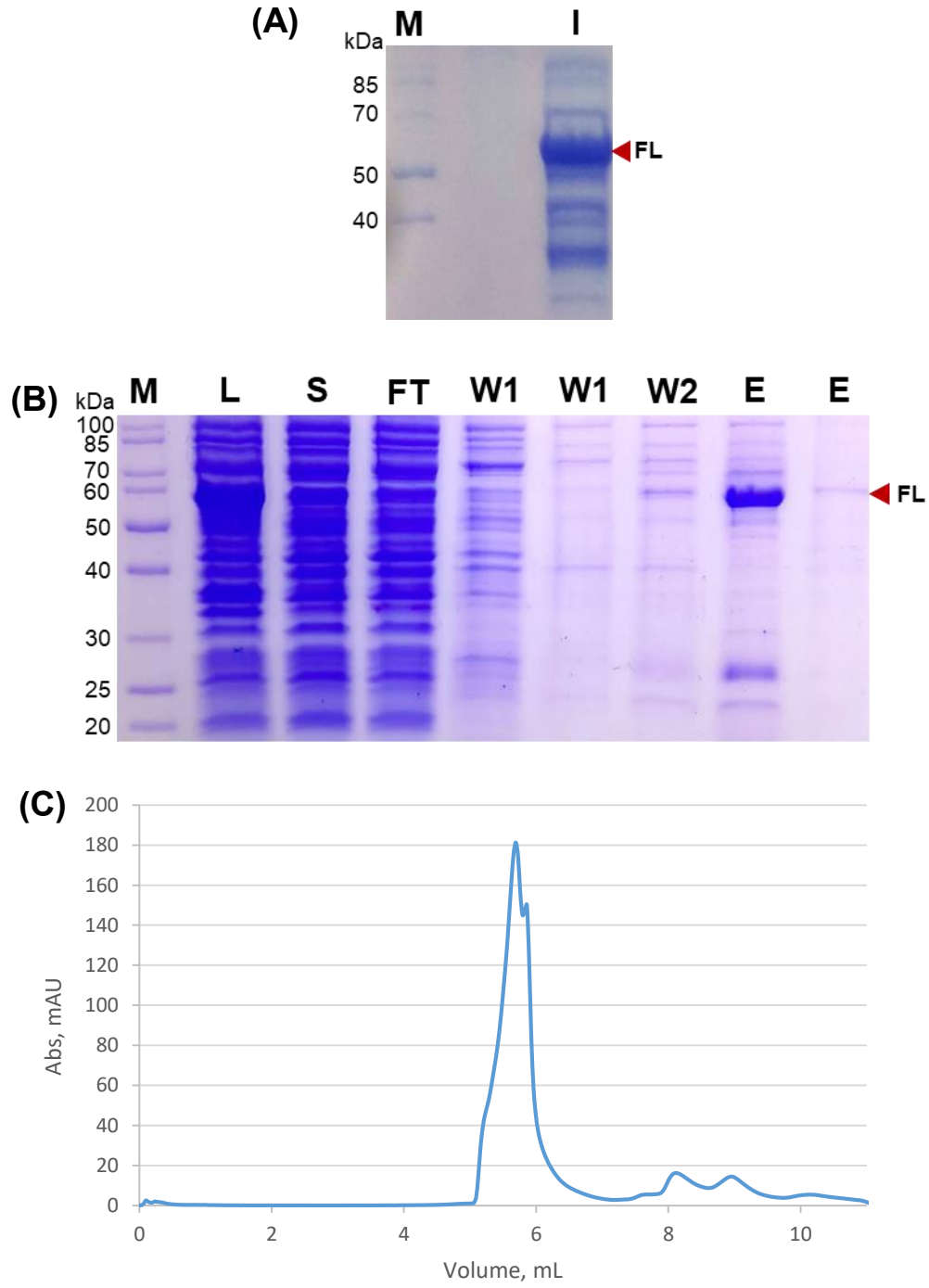


Figure 4.2| Expression and purification of AtS6K2. **(A)** Small scale test expression (using BL21 grown in LB) of wild-type AtS6K2 KL_123_471_NC_C-His_N-T7 (MW=55.5 kDa) showing IPTG induced (I) sample. **(B)** Samples of AtS6K2 KL_123_471_NC_C-His_N-T7 from different IMAC purification steps. L – Cell lysate; S – Supernatant of centrifuged cell lysate; FT – HisTrap column Flow-through; W1 – HisTrap column wash; W2 – HisTrap column wash with 20 mM imidazole; E – Elution; M – Protein ladder. AtS6K2 KL_123_471_NC_C-His_N-T7 is highlighted with red arrows on the sides of the gels. **(C)** Chromatogram of analytical SEC, showing the AtS6K2 KL_123_471_NC_C-His_N-T7 peak at void volume.

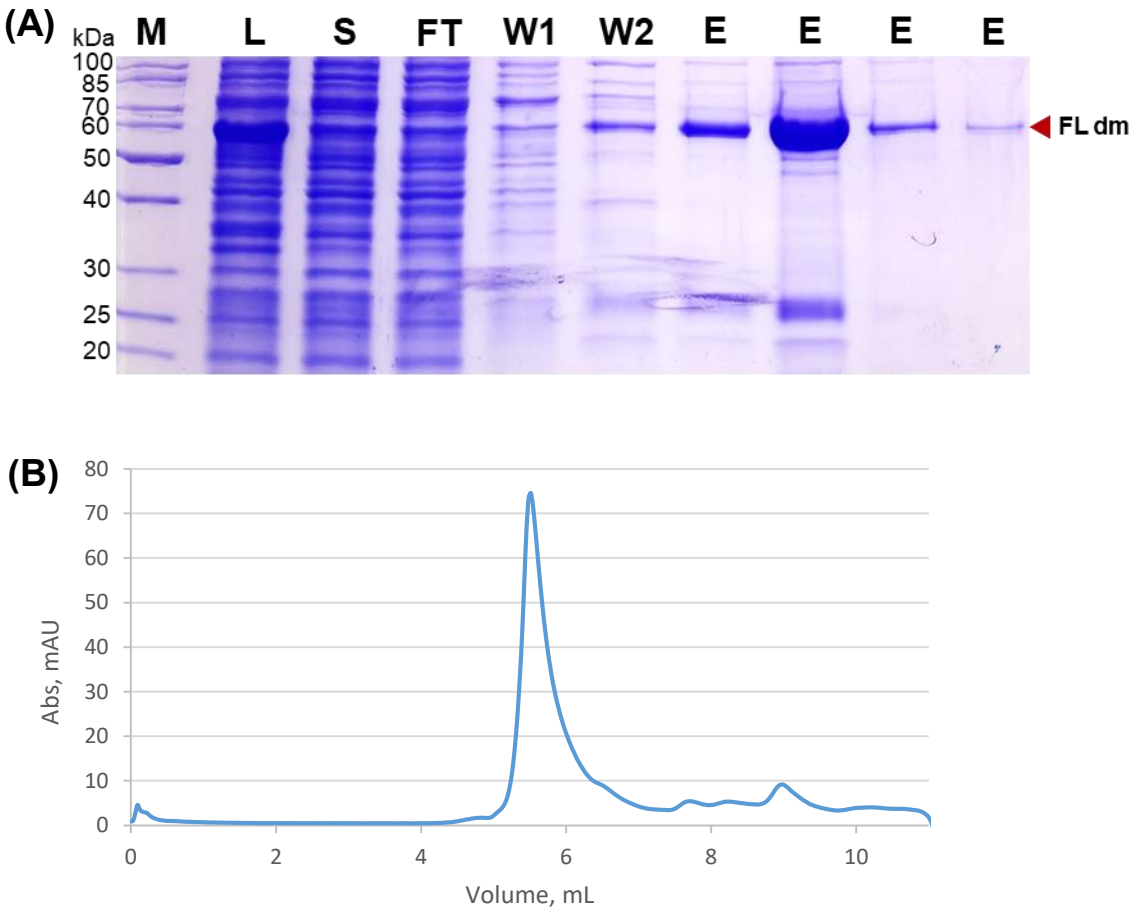


Figure 4.3| Purification of AtS6K2 T455E_S296E mutant. **(A)** Samples of AtS6K2 T455E_S296E (FL_1_471_NC_C-His_N-T7_dm; MW=55.6 kDa) from different IMAC purification steps. M – Protein ladder; L – Cell lysate; S – Supernatant of centrifuged cell lysate; FT – HisTrap column Flow-through; W1 – HisTrap column wash; W2 – HisTrap column wash with 20 mM imidazole; E – Elution. AtS6K2 FL_1_471_NC_C-His_N-T7 T455E_S296E is highlighted with red arrow on the sides of the gel. **(B)** Chromatogram of analytical SEC, showing the AtS6K2 FL_1_471_NC_C-His_N-T7 T455E_S296E peak at void volume.

4.1.2. AtS6K2 FL_1_471_NC_N-His_λ

In order to produce soluble and stable full length AtS6K2, co-expression with λ-phosphatase was tested using a pETDuet-1 vector. The AtS6K2 FL_1_471_NC_N-His_λ construct contains an uncleavable N-term His-tag (Table 3.1). In addition, a double phosphomimetic mutant of AtS6K2 FL_1_471_NC_N-His_λ (T455E, S296E) was generated via site-directed mutagenesis. Expression and purification were carried out in similar manner for both WT protein and mutant. The proteins were co-expressed in LB media and purified by IMAC and SEC. A strong overexpression band of λ-phosphatase at ≈25 kDa was present, however a weak FL_1_471_NC_N-His_λ (54.2 kDa) band between 50 and 60 kDa only appeared in the HisTrap column elution fraction (Fig. 4.4A and 4.6A). To confirm that FL_1_471_NC_N-His_λ is present in the IMAC column elution, western blot with anti-6xHis antibody was carried out for wild type protein. A band above 55 kDa was present in the affinity column elution fraction (Fig. 4.4B). The protein was further purified by analytical SEC, which resulted in multiple overlapped peaks (Fig. 4.5B and 4.6C). A peak at 8.7 mL contained AtS6K2 with reduced amount of contaminants (especially λ-phosphatase that was abundant; Fig. 4.5A and 4.6B), and the corresponding fractions were concentrated, flash frozen and stored at -80°C. Unfortunately, the concentration of the protein was 100 nM, which, considering presence of contaminants, was far too low for the activity measure. This could have been linked to the N-terminal position of the His-tag, which affected protein solubility and stability (Section 3.1.3). Unfortunately, attempts to move the His-tag to the C-terminal failed. Currently, experiments to confirm increased activity for the AtS6K2 T455E, S296E double mutant are being carried out *in vivo* using anti-rpS6 phospho-specific antibody.

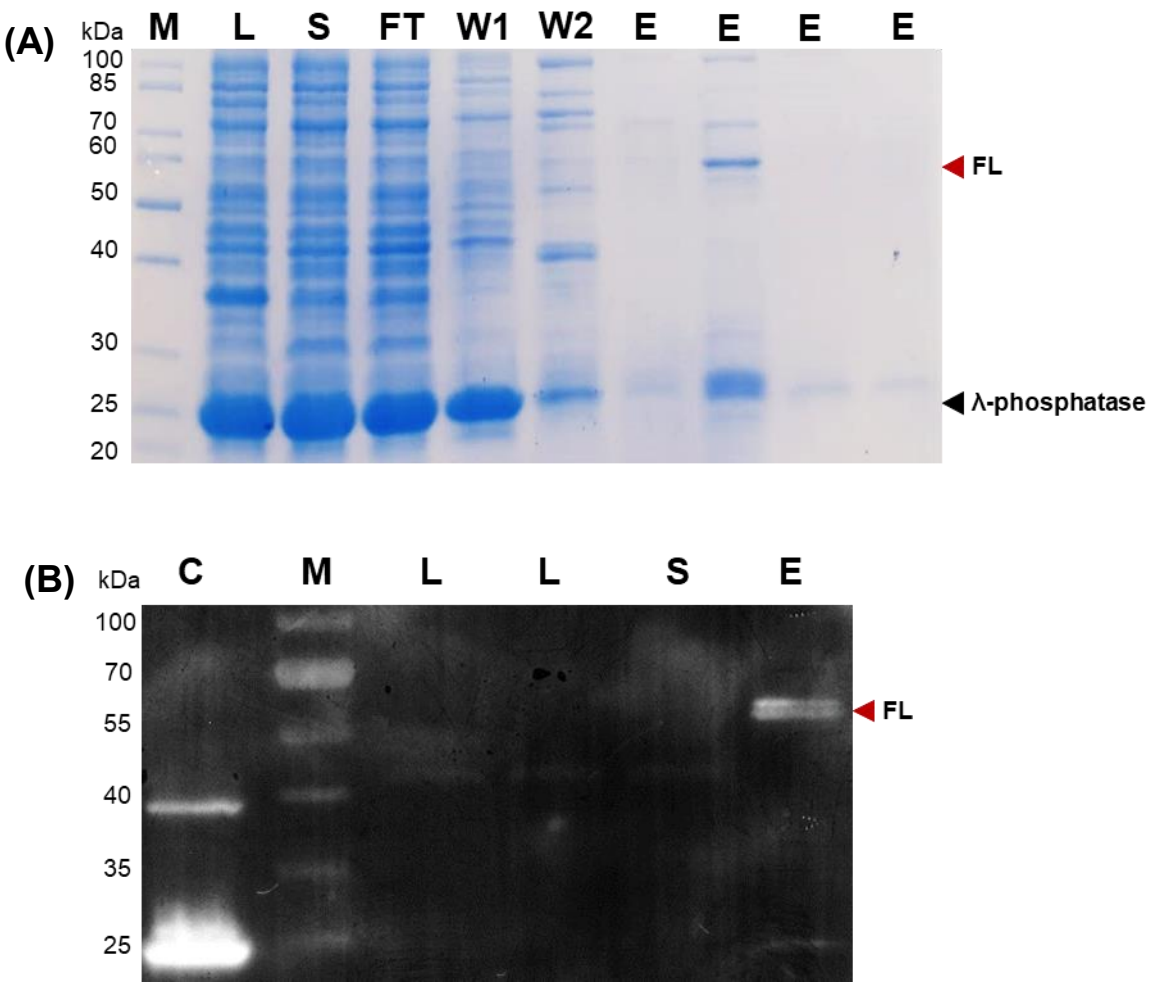


Figure 4.4| Purification of wild-type AtS6K2 FL_1_471_NC_N-His_λ. **(A)** Samples of AtS6K2 FL_1_471_NC_N-His_λ (54.2 kDa) from different IMAC purification steps. L – Cell lysate; S – Supernatant of centrifuged cell lysate; FT – HisTrap column Flow-through; W1 – HisTrap column wash; W2 – HisTrap column wash with 20 mM imidazole; E – Elution. **(B)** Western-blot showing cell lysate (L), supernatant of centrifuged cell lysate (S) and IMAC elution fraction with highest protein content (E). A positive HisTag control (C, MW=37 kDa) was included. AtS6K2 FL_1_471_NC_N-His_λ is highlighted with red arrows on the sides of the figures, while λ-phosphatase with a black arrow. M – Protein ladder.

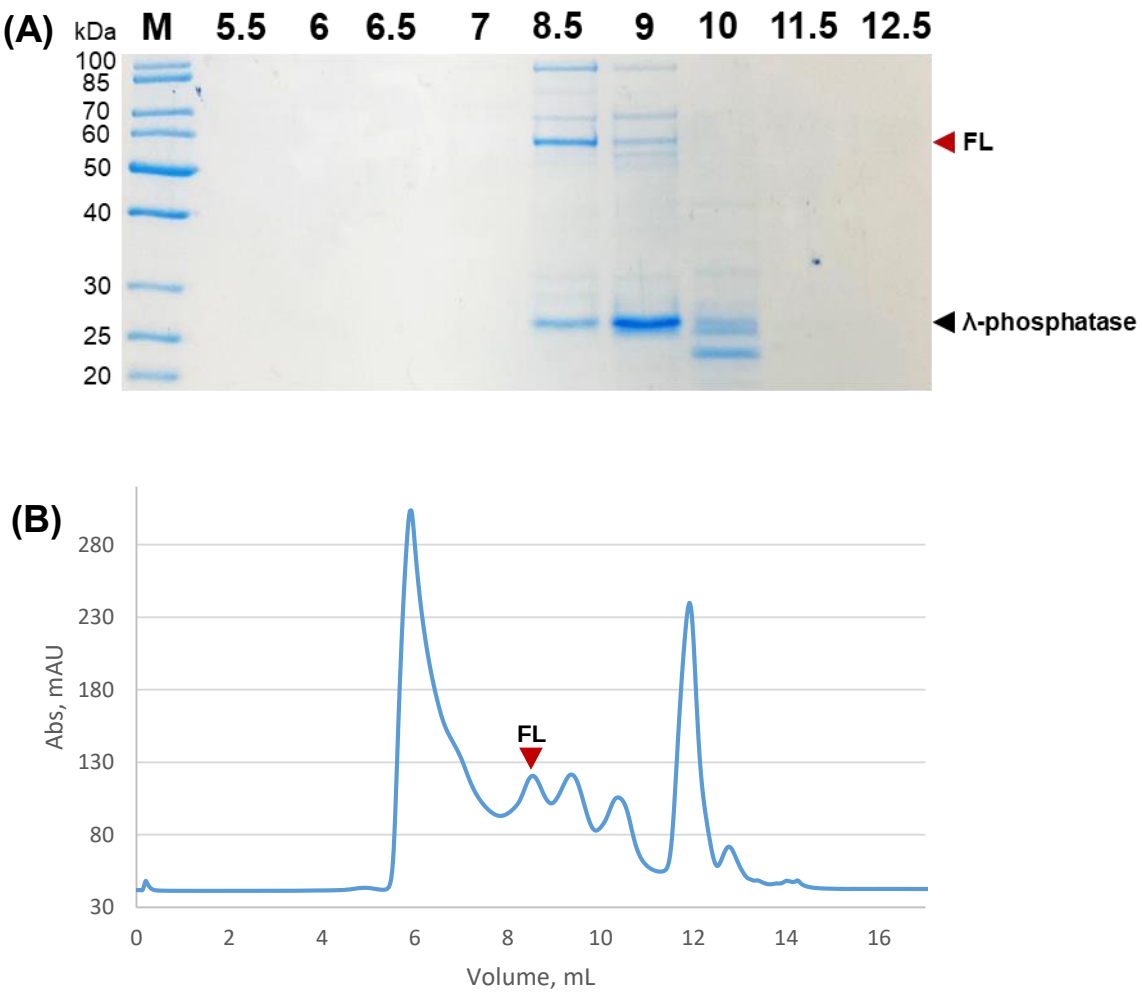


Figure 4.5| Analytical SEC purification of wild-type AtS6K2 FL_1_471_NC_N-His_λ. **(A)** Samples from SEC. M – Protein ladder. Numbers on top of wells represent the elution volume of the fraction in the chromatogram. 0.5 mL fractions were collected. AtS6K2 FL_1_471_NC_N-His_λ (54.2 kDa) is highlighted with red arrows, while λ-phosphatase with a black arrow. **(B)** Chromatogram of SEC, showing the AtS6K2 FL_1_471_NC_N-His_λ elution peak.

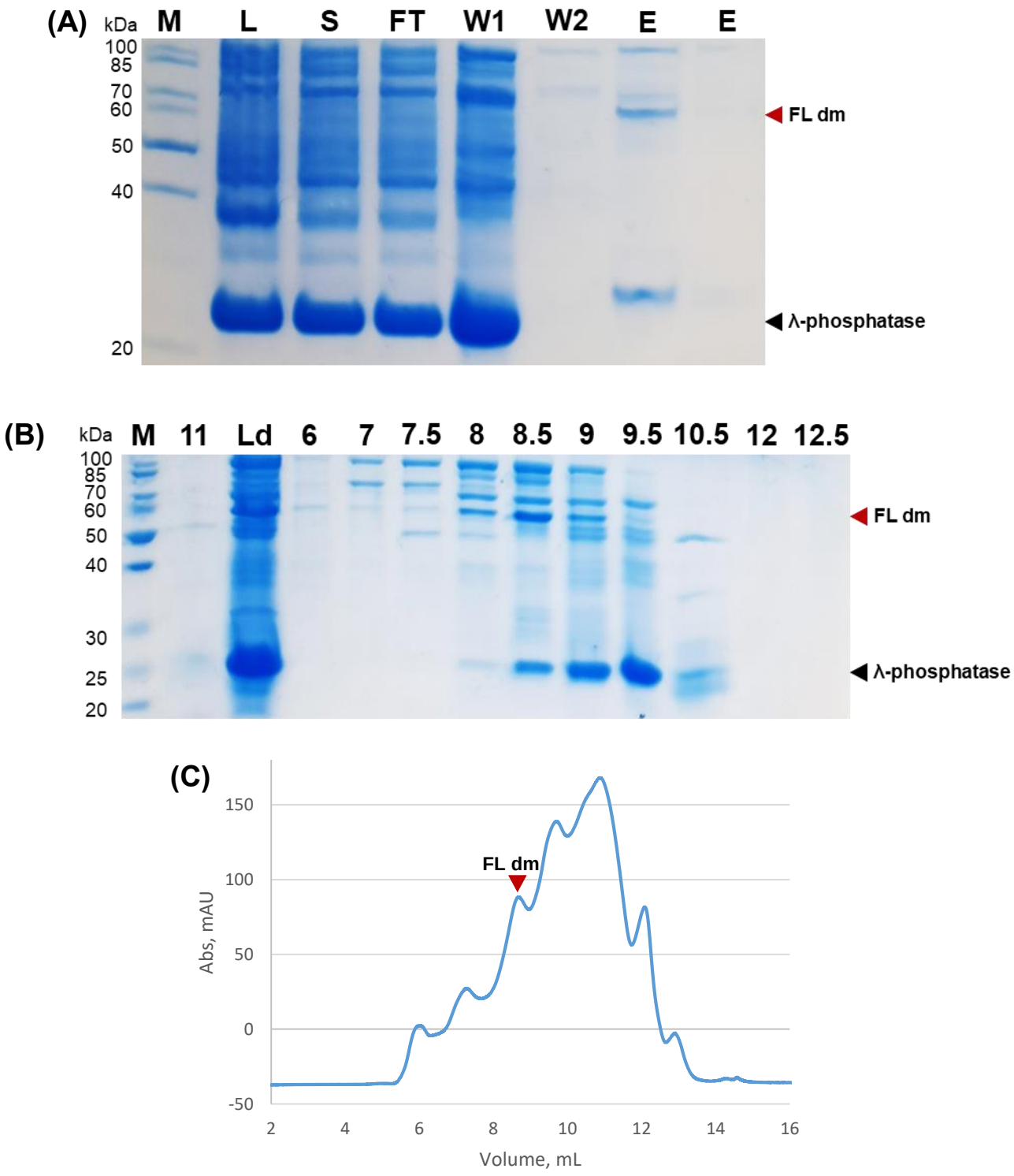


Figure 4.6| Purification of AtS6K2 FL_1_471_NC_N-His_λ T455E_S296E mutant. **(A)** Samples of AtS6K2 FL_1_471_NC_N-His_λ T455E_S296E (54.3 kDa) from different IMAC purification steps. L – Cell lysate; S – Supernatant of centrifuged cell lysate; FT – HisTrap column Flow-through; W1 – HisTrap column wash; W2 – HisTrap column wash with 20 mM imidazole; E – Elution. **(B)** Samples from analytical SEC of AtS6K2 FL_1_471_NC_N-His_λ_dm. Ld- sample loaded on the SEC column. Numbers on top of wells represent the volume in the chromatogram. 0.5 mL fractions were collected. **(A)** Chromatogram of SEC, showing the AtS6K2 FL_1_471_NC_N-His_λ_dm elution peak. AtS6K2 FL_1_471_NC_N-His_λ is highlighted with red arrows on the sides of the figures, while λ-phosphatase with a black arrows. M – Protein ladder.

4.2. Discussion

Producing physiologically active, precisely phosphorylated kinases in sufficient quantity and purity for structure-function determination is a major challenge. Site directed mutations to Asp/Glu are commonly employed to generate constitutively active kinases, however phosphomimetics have been shown to have inconsistent effects (Szewczuk, Tarrant and Cole, 2009). Carboxylates have several differences from phosphates including geometry, size, number of oxygen atoms available for hydrogen bonding and number of negative charges at neutral pH. Efforts to generate an active AtS6K2 kinase using T455E, S296E mutants were unsuccessful. AtS6K2 KL_123_471_NC_C-His and KN_136_460_C_C-His double mutants were insoluble. AtS6K2 FL_1_471_NC_C-His_N-T7 WT and mutant aggregated possibly due to hyperphosphorylation. FL_1_471_NC_N-His_λ was co-expressed with λ-phosphatase but very little protein was produced and optimisation of purification is required.

Recent studies examined activity determinants of plant S6Ks (Yaguchi and Kozaki, 2018, Yaguchi et al., 2020). S6Ks from Arabidopsis (AtS6K1), rice (OsS6K1 and OsS6K2), tomato (SlS6K), soybean (GmS6K) and maize (ZmS6K) were expressed in yeast cells lacking Ypk3, the functional ortholog of mammalian S6K (González *et al.*, 2015). Activity was examined by detecting phosphorylated rpS6 via Western-blotting. OsS6K1, OsS6K2 and GmS6K demonstrated high activity, while SlS6K was less active. ZmS6K and AtS6K1 activity was hardly detectable, while AtS6K2 activity was not detected. OsS6K1 activity was sensitive to rapamycin. Unfortunately, only OsS6K1 expression was confirmed with Anti-OsS6K1 antibody. As such, reduced rpS6 phosphorylation of S6Ks may have been caused by low amount of expressed protein and not a lack of activity (Yaguchi and Kozaki, 2018, Yaguchi et al., 2020). Indeed, heterologous expression in *E. coli* demonstrated that AtS6K2 is difficult to produce in soluble form and may be prone to degradation (Section

3.1.4.5.4) and aggregation (4.1.1). Three important phosphorylation sites are known to be present in plant S6Ks: a hydrophobic motif (HM), activation loop and Turn motif (TM), and HM is phosphorylated by TORC1 (Xiong *et al.*, 2013; Xiong and Sheen, 2014). Phosphorylation of HM (T455 in AtS6K2) in plant S6Ks was examined using anti-human phospho-T389 S6K1 antibody (Yaguchi and Kozaki, 2018, Yaguchi et al., 2020) and, surprisingly, the only phosphorylated HM (pHM) was detected in full length SlS6K (Yaguchi and Kozaki, 2018, Yaguchi et al., 2020). However, it is not clear if the absence of pHM was due to low protein amounts and insensitivity of Western-blots.

In addition, various constructs of truncated S6Ks at the N-terminus were tested, and similarly to FL no activity or HM phosphorylation was detected for truncated variants of AtS6K2. Truncated AtS6K1 and ZmS6K showed increased activity, but no pHM was detected. Truncated SlS6K demonstrated higher phosphorylation of rpS6 and its own HM than the SlS6K FL. In OsS6Ks, phosphorylation in HM was detected when the N-terminal region was removed. Truncated OsS6K1 and OsS6K2 were more resistant to rapamycin treatment than FL proteins, indicating that N-term of OsS6Ks is important to activity regulation by TORC1. This is not surprising, as TORC1 is known to regulate S6Ks through the TOS motif, located on the N-terminus of plant S6Ks (Son *et al.*, 2017b). However, the authors were unable to rationalise why the OsS6K1 HM is found phosphorylated in constructs lacking the N-term with the TOS motif. Furthermore, the importance of the three phosphorylation sites were tested by mutating the HM (S465A), TM (S449A) and activation loop (S307A) (Yaguchi and Kozaki, 2018b). S465A did not disrupt the activity of OsS6K1 while TM and activation loop mutants were inactive, suggesting that HM phosphorylation is not necessary for OsS6K1 activity, while TM and activation loop are important, differently from what shown in human S6Ks (Fenton and Gout, 2011). Unexpectedly, the TM (S449A), activation loop (S307A) or triple mutant (S307A, S449A, S465A) were all active

and had non-phosphorylated HM. The authors speculated that other rapamycin-sensitive phosphorylation sites in plant S6Ks could exist and that these sites could substitute for T-loop and TM sites in truncated OsS6K1 (Yaguchi and Kozaki, 2018b). However, these results could also be an artefact (e.g., yeast have a different PTM control from other eukaryotes, and tend to over phosphorylate and over-glycosylate) (Conde *et al.*, 2004) of heterologous rice S6Ks' expression in yeast, which was not fully appreciated or validated by the authors. Importantly, phosphomimetic mutants were also generated for OsS6K1. Activation loop mutant (S307E) exhibited slightly higher activity than WT, while TM S449E and S449D mutants were less active than WT (S449D demonstrated more activity of the two). Both activation loop and TM phosphomimetic mutants were as sensitive to rapamycin treatment as WT, while the OsS6K1 HM S465E mutant was less sensitive to rapamycin and was similarly active to WT. To further explore the effects of phosphomimetic mutations, double and triple mutants were constructed. Surprisingly the OsS6K1 S307A_S465E (activation loop dead plus HM phosphomimetic) and S449A_S465E (TM dead plus HM phosphomimetic) double mutants were active as the WT, demonstrating that S465E mutation at HM is capable of restoring activity, lost by mutations at TM (S449A) and activation loop (S307A). The triple S307A, S449A, S465E mutant was inactive, indicating that at least two phosphorylated sites were required for active OsS6K1, expressed in yeast (Yaguchi, Ikeya and Kozaki, 2020b).

However, the authors (Yaguchi and Kozaki, 2018, Yaguchi et al., 2020) were unable to draw a clear picture of plant S6Ks' activation due to the use of a heterologous system, the unquantified ectopic expression of plants proteins in ypk3Δ yeast cells, and the lack of protein-specific antibodies. As such, AtS6K2 double phosphomimetic mutant with mutations at HM (T455E) and activation loop (S296E) may be active, but in future different mutants should also be explored and/or investigated in vivo. Inserting aspartates instead

of glutamates, or producing triple or different double mutants should also be tested for AtS6K2. Interestingly, parallel *in vivo* unpublished studies linked to my project, showed Arabidopsis T-DNA insertion mutants of the AtS6K2 gene complemented with $S6K2^{Pro}:AtS6K2^{S296E\_T455E}$ had a significantly higher resistance to cold and salt stresses than the plants complemented with $S6K2^{Pro}:AtS6K2^{WT}$. This would suggest that the mutant has a different activity than the WT protein and that is providing and stress resistance advantage to the plant. Western-blot experiments using phospho-rpS6A/B antibody are being carried out to verify an increased activity of the mutant protein. If confirmed this would be only partially in line with data published by Yaguchi et al. (Yaguchi and Kozaki, 2018, Yaguchi et al., 2020), but it would be significantly more physiologically relevant.

5. S6K kinase domain *in silico* structure and dynamics

5.1. Results

An unusual feature of the crystal structure of the human S6K1 is the presence of a zinc ion near the activation loop (Wang *et al.*, 2013), coordinated by a zinc-finger-like motif two Cys and two His residues (Cys 217, His 222, His 228 and Cys 231). The authors speculated that Zn^{2+} binding might be associated with the inactive state of the protein, as the conformation of the activation loop induced by the Zn^{2+} ion creates a steric hindrance preventing substrate binding. However, the biological role of the Zn^{2+} in kinase activity has not been investigated. Similarly, a Zn^{2+} ion was also observed in a similar location in the crystal structure of the p90 Ribosomal S6 Kinase 4 N-terminal Kinase domain (RSK4-NTKD) (Chrysostomou *et al.*, 2021), an S6K1 homologue. However, in RSK4 the ion is 10Å far from its relative position in S6K1 and it does not prevent the interaction between the phosphorylated activation loop and the HRD motif. While confirming the presence of a “hot spot” (Krantz, 1998; Niccolai *et al.*, 2003; Venditti *et al.*, 2007) for cation binding in ribosomal kinases, the authors suggested that the ion is not biologically relevant, differently from the Zn^{2+} ion in S6K1. Importantly, similar data were also collected for other kinases that are not evolutionary related to rpS6 kinases, for example, Polo-like kinase 1 (Plk1) and dual-specificity tyrosine phosphorylation-regulated kinase 1A (DYRK1A). Zn^{2+} was shown to significantly improve crystallisation of Plk1 and this was linked to the Zn^{2+} binding the protein in a comparable region to that seen in S6K1 (Kothe *et al.*, 2007). Despite structural differences with S6K1 kinases (Plk1 belongs to the other group of Kinases), the ion induces an inactive conformation with a mechanism of inhibition similar to that seen in S6K1 crystal structure. DYRK1A was, instead, crystallised in presence of Li^{+}, which unsurprisingly also is found in the crystal structure in a position resembling that

of Zn^{2+} seen for Plk1 (Kaltheuner *et al.*, 2021). The authors did not investigate or discuss the relevance of this ion. All this evidence supports the hypothesis that cations can bind kinases near the active loop and may play a role in structural stabilisation.

My aim was to investigate how different cations bind to AtS6K2 and S6Ks in general. Despite several published purification protocols (Malakhova *et al.*, 2009; Sunami *et al.*, 2010) and several kinase structures (Tereshko *et al.*, 2001; Kothe *et al.*, 2007; Wang *et al.*, 2013; Chrysostomou *et al.*, 2021), positively charged ions (i.e., Mg^{2+}, Li^{+}, Zn^{2+} and Mn^{2+}) are often bound in sites different from the active site, which is normally occupied by ATP and Mg^{2+}, and their role remains unclear. Extensive and comprehensive attempts to purify recombinant AtS6K2 failed to produce soluble protein for structural and functional studies. In order to overcome this bottleneck and, at the same time, help rationalize results obtained during the optimisation of purification protocols, I performed molecular dynamics (MD) simulations. MD simulations are a powerful tool that allow to study the motions of proteins at atomic-level and explore protein structural features and ligand binding properties (Hollingsworth and Dror, 2018).

5.1.1. AtS6K2 dynamics and metal binding

As mentioned in chapter 3.1.5, human S6K1 was purified in the presence of Li_2SO_4 as this was found to reduce aggregation, although no data was presented (Sunami *et al.*, 2010). However, a zinc, and not Lithium, ion was found to bind near the activation loop of S6K1 crystal structure (PDB ID: 4L46) (Wang *et al.*, 2013). Similarly, Plk1 (Kothe *et al.*, 2007) and RSK4-NTKD (Chrysostomou *et al.*, 2021) were purified with a buffer similar to the one used for AtS6K2 in chapters 3 and 4 (except 3.1.5), and yet the crystal structure contained a Zn^{2+} ions bound to Cys and His residues near the activation loop, with no other clear functional reason to be in that site.

Earlier studies on human S6K1 showed that the residues forming the zinc-finger-like motif are conserved in human S6K2 and S6Ks of other species, but not other AGC kinases (Sunami *et al.*, 2010). To explore the role cations may have on protein stability, and hence be useful additives in purification buffers, I carried out MD simulations of AtS6K2 KD in presence of different ions, i.e. Li^+, Zn^{2+} and, as a control, Mn^{2+}. The structure of AtS6K2 KD was predicted via homology modelling, using human S6K1 as template, and the ATP and Mg^{2+} ions were placed in the active site based on their relative position in PKA X-ray structure (PDB ID: 1Q24). An initial analysis of the 500 ns trajectory shows that the model was stable (Fig 5.1). This was verified via root-mean square deviation (RSMD) and radius of gyration (Rg) analysis, which tend to increase at the beginning of the simulation and then plateau (Assadollahi *et al.*, 2019). The all atom RMSD and the Rg average for AtS6K2 KD were 2.52 ± 0.15 Å and 19.54 ± 0.11 Å (Table 5.1), respectively, suggesting no major conformational changes. I then repeated the simulations in presence of Zn^{2+}, Mn^{2+} and Li^+ placed in a potential “hot-spot” (Krantz, 1998; Niccolai *et al.*, 2003; Venditti *et al.*, 2007) based on the relative position of Zn^{2+} in the S6K1 structure (PDB ID: 1Q24). Interestingly, Li^+, Zn^{2+} and Mn^{2+} form two specific

coordinate bonds with two different residues in the protein and two water molecules. This tetrahedral coordination clearly distinguishes this these ions from an Mg^{2+} ion. The position of the Li^{+}, Zn^{2+} and Mn^{2+} ions during the MD simulations is relatively stable (average distance during the simulation from the centre of the binding pocket (the position of the ion after the equilibration run) 3.17 ± 0.97 Å; 3.19 ± 0.31 Å; 2.94 ± 0.24 Å, respectively (Table 5.2). No changes in protein structure were observed in presence of the ions (Table 5.2, Fig. 5.2 and 5.3). However, the water molecules coordinating the Li^{+} ion were not structural and they stayed in contact with the ion for less than 20% of the simulation. On the other hand, water molecules bound to Zn^{2+} and Mn^{2+} did not detach from the ions for the whole duration of the dynamics (Fig. 5.4).

Taken together, these data may suggest that Zn^{2+} and Mn^{2+} form stable and specific interactions with polar exposed surfaces of AtS6K2, which stabilise the protein locally potentially via formation of structural hydration shells (De Simone *et al.*, 2005).

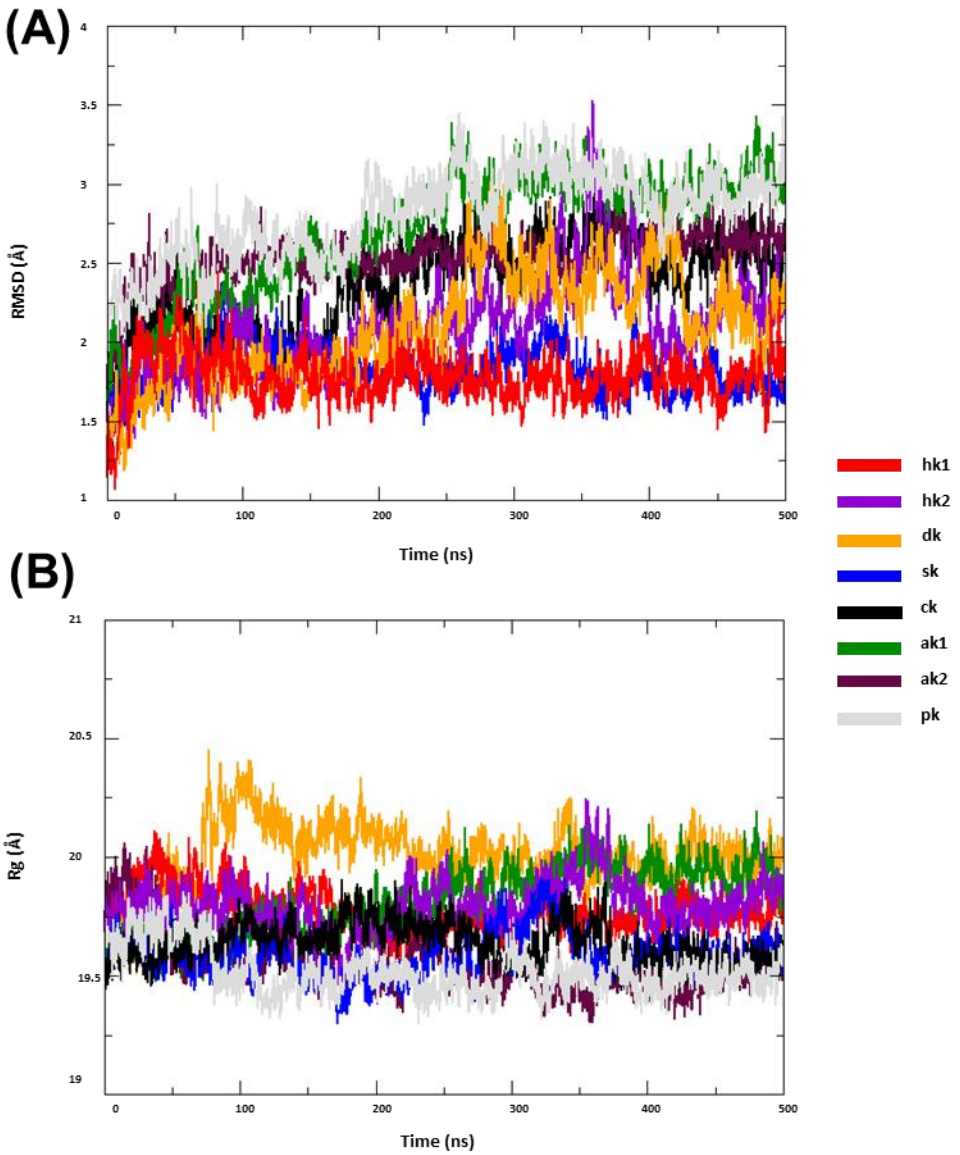


Figure 5.1| Validation of models. Plots show the all-atoms RMSD **(A)** and Rg **(B)** of each model containing ATP and Mg^{2+} (colour coded according to legend) relative to their respective initial frame against simulation time.

Table 5.1| Average RMSD and Rg values for the various protein models in the presence of ATP, Mg2+ (apo) and either Zn^{2+}, Mn^{2+} or Li^{+}.

Orthologues	apo	Zn^{2+}	Mn^{2+}	Li^{+}
RMSD				
hk1	1.77 ± 0.14	1.79 ± 0.24	1.42 ± 0.14	1.53 ± 0.18
hk2	2.08 ± 0.31	1.80 ± 0.20	1.90 ± 0.20	2.24 ± 0.19
dk	2.09 ± 0.35	1.67 ± 0.14	1.83 ± 0.18	2.05 ± 0.32
sk	1.82 ± 0.14	2.21 ± 0.38	2.27 ± 0.28	2.14 ± 0.31
ck	2.36 ± 0.30	2.04 ± 0.19	2.01 ± 0.31	1.71 ± 0.24
ak1	2.69 ± 0.37	3.25 ± 0.39	2.41 ± 0.26	2.79 ± 0.31
ak2	2.52 ± 0.15	2.92 ± 0.58	2.72 ± 0.20	3.19 ± 0.61
pk	2.79 ± 0.27	n.e.*	n.e.*	n.e.*
Rg				
hk1	19.77 ± 0.09	19.59 ± 0.07	19.65 ± 0.07	19.67 ± 0.07
hk2	19.79 ± 0.10	19.84 ± 0.10	19.75 ± 0.07	19.72 ± 0.08
dk	20.00 ± 0.13	19.62 ± 0.07	19.76 ± 0.06	19.72 ± 0.08
sk	19.58 ± 0.09	19.59 ± 0.09	19.68 ± 0.06	19.81 ± 0.08
ck	19.63 ± 0.08	19.68 ± 0.08	19.70 ± 0.06	19.68 ± 0.06
ak1	19.80 ± 0.14	19.96 ± 0.11	19.75 ± 0.07	19.56 ± 0.10
ak2	19.54 ± 0.11	19.91 ± 0.15	19.70 ± 0.15	19.80 ± 0.12
pk	19.51 ± 0.09	n.e.*	n.e.*	n.e.*

Median mean value (Å) ± standard deviation ***not evaluable**

Table 5.2| Ions average distance from centre of binding site

Orthologues	Zn^{2+}	Mn^{2+}	Li^{+}
hk1	1.10 ± 0.21	0.97 ± 0.11	3.66 ± 0.99
hk2	0.82 ± 0.26	1.03 ± 0.56	1.65 ± 0.52
dk	1.50 ± 0.27	1.02 ± 0.18	1.27 ± 0.21
sk	1.76 ± 0.12	2.70 ± 0.44	3.25 ± 2.14
ck	1.33 ± 0.41	0.99 ± 0.19	3.37 ± 1.31
ak1	2.75 ± 0.38	3.30 ± 0.13	10.32 ± 1.77
ak2	3.19 ± 0.31	2.94 ± 0.24	3.17 ± 0.97
pk	n.e.*	n.e.*	n.e.*

Median mean value (Å) ± standard deviation ***not evaluable**

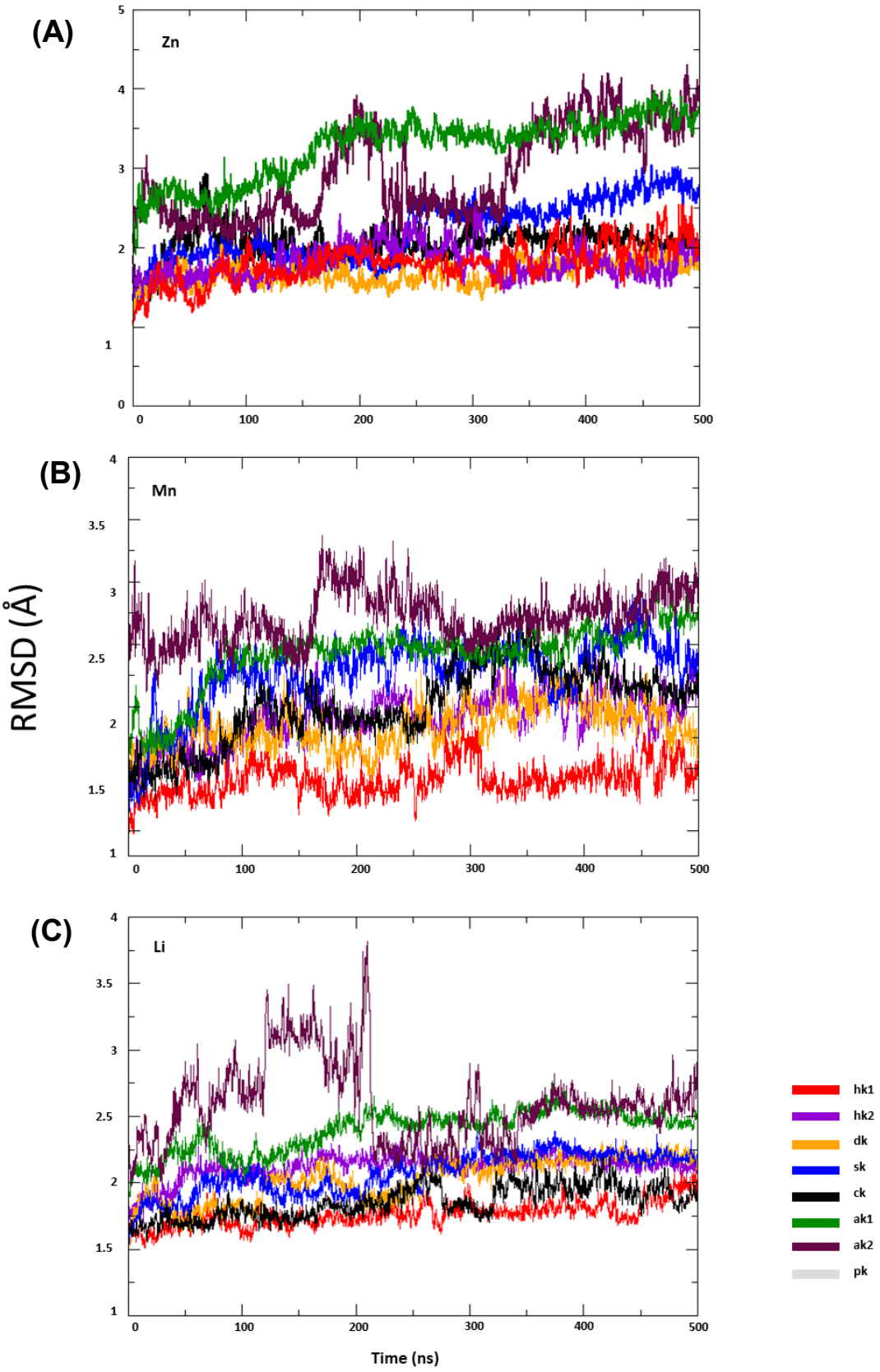


Figure 5.2| Ions' effect on protein flexibility. Plots show the all-atom protein RMSD of each model in presence of Zn^{2+} **(A)**, Mn^{2+} **(B)** and Li^{+} **(C)** (colour coded according to legend) relative to their respective initial frame against simulation time.

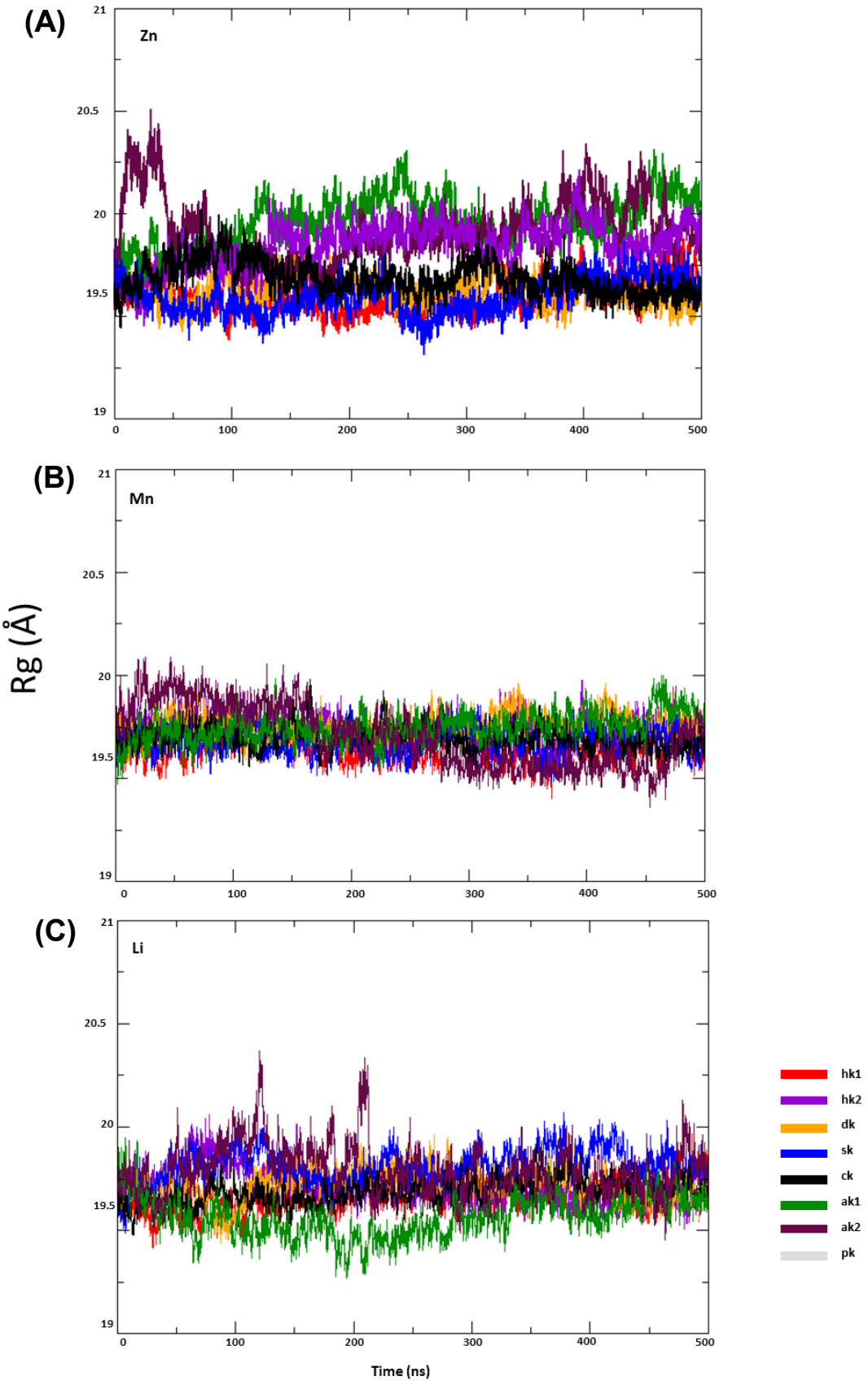


Figure 5.3| Ions' effect on protein structures. Plots show the protein Rg of each model in presence of Zn^{2+} **(A)**, Mn^{2+} **(B)** and Li^{+} **(C)** (colour coded according to legend) relative to their respective initial frame against simulation time.

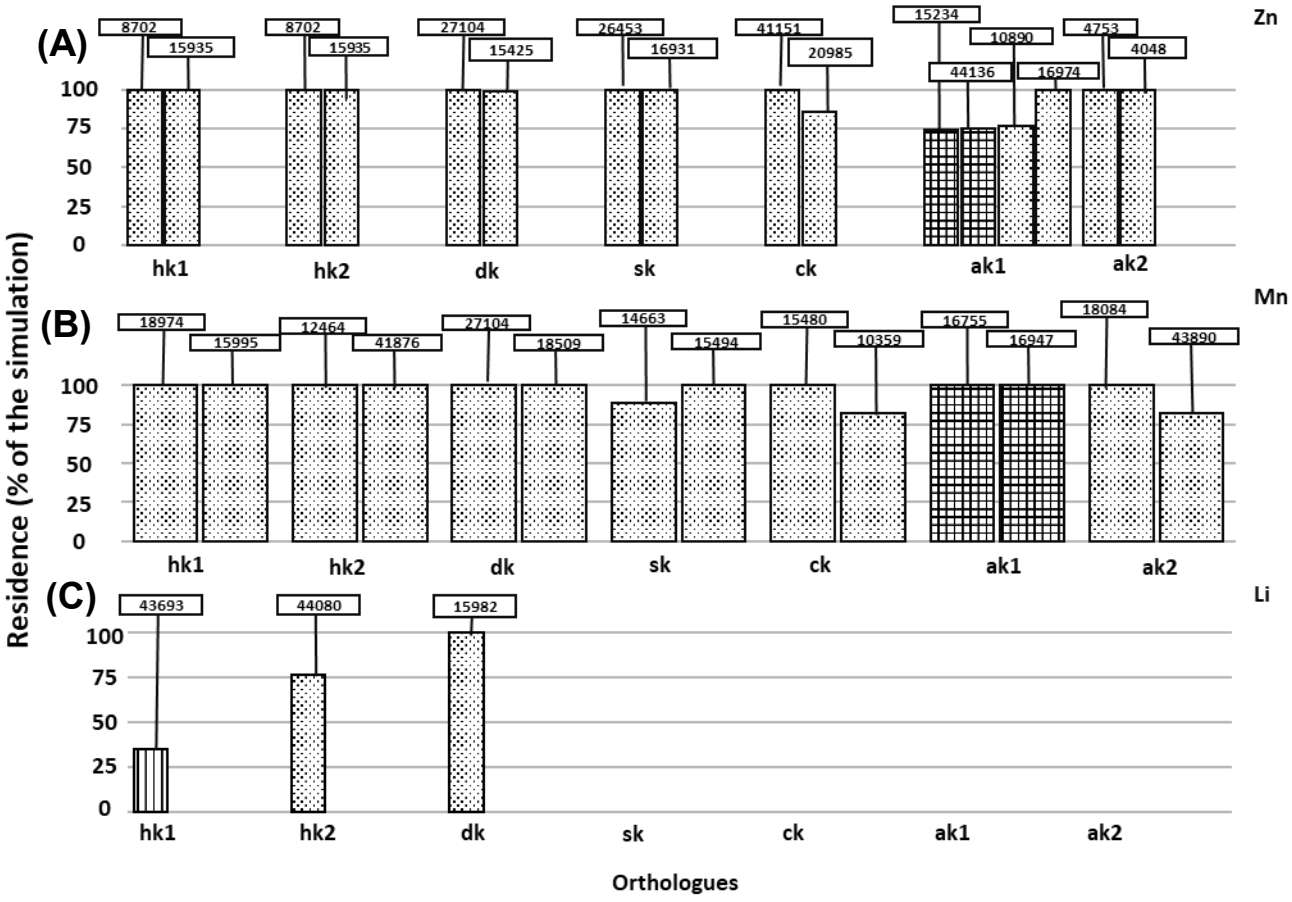


Figure 5.4| Water molecules residence time. Bars show the % of the simulation time (considered as continuous, and not cumulative, time) in which specific water molecules stay coordinated to Zn^{2+} **(A)**, Mn^{2+} **(B)** and Li^{+} **(C)** against the different orthologues. Water molecule identifier code in the simulation is reported above the bars.

5.1.2. Far orthologues

To increase reliability of the predictions and understand if the features identified using AtS6K2 KD model are conserved in the S6K family, I repeated the same simulations described above using different S6K isoforms from different species. To avoid biases due to sequence similarity and highlight unique structural features, I selected evolutionary hyphenated proteins which sequences have significantly diverged during evolution (i.e., overall sequence identity <50%). Specifically, I selected species with two S6Ks isoforms, the *Arabidopsis thaliana* S6K1 (ak1), human S6k1 (hk1) and S6K2 (hk2), and species with only one S6K isoform, *Drosophila Melanogaster* S6K (dk), *Strongylocentrotus purpuratus* S6K (sk), *Caenorhabditis elegans* S6K (ck). While initial studies included the yeast S6K (pk), the generated model didn't bind any of the ions and lacked reliability, thus it was not included in the analysis. Similarly to AtS6K2 KD, all models were stable (Fig. 5.1) and, importantly, showed very similar RMSD and Rg average values (Table 5.1). In presence of Zn^{2+}, Mn^{2+} and Li^{+}, placed in the potential "hot-spot" as identified in the structure of AtS6K2 KD, structural integrity of the models did not change (Table 5.2, Fig. 5.2 and 5.3). Similarly to AtS6K2 KD, Zn^{2+} and Mn^{2+} form stable interactions with all orthologues via a combination of protein residues and water molecule coordination. The binding residues are not conserved, but all proteins have within the binding surface at least two residues which side chains can form specific and stable interaction with Zn^{2+} and Mn^{2+}. Furthermore, for all species studied, coordinated water molecules stay bound to the ions for over 75% of MD simulation (>375 ns) (Fig. 5.4). As observed in AtS6K2, Li^{+} behaves differently, with water molecules residing for short times close to the ion. A closer analysis of the residence time and stability of the Li^{+} ion itself on the protein shows that for all orthologues the ion, in average, moves within the binding surface (for AtS6K1 KD away from it) 3 times more than Zn^{2+} and Mn^{2+} (Table 5.2).

Taken together this data reinforce the observation that Zn^{2+} and Mn^{2+} may have a local stabilisation effect on the proteins, while Li^{+} does not. This may help explain the lack of improvement in protein solubility using Li_2SO_4 instead of NaCL.

5.2. Discussion

The aim of this study was to use a robust *in silico* protocol to overcame the lack of available structures and generate information that could inform protein expression and purification optimisation studies. *In silico* tools for protein structure and dynamics predictions have transformed (and are expected to do so even more in the future) the fields of structural biology and medicinal chemistry (Jumper *et al.*, 2021). Despite an increasing number of tools available for structure prediction (i.e., AlphaFold, I-Tasser, Rosetta) (Yang *et al.*, 2015; Du *et al.*, 2021; Kim *et al.*, 2023), as shown in chapter 3, S6Ks have a conserved kinase fold. For consistency, I used the same homology prediction method and 3D template for all protein studied. However, protein structure prediction doesn't predict hydration patterns and binding of metals or cofactors and their role in protein stability. More reliable information can only be obtained by MD simulations (Gazi *et al*., 2023).

Consistent with Chrysostomou *et al.*, (2021), the surface flanking the activation loop opposite the active site appears to constitute a "hot spot". Indeed, in all structures studied, ions are readily coordinated forming similar bonds with structurally conserved residues in the proteins. Importantly, these residues are not conserved in sequence but are in space, which is evident by comparing far homologues. No comparison study could be carried out with other ribosomal kinase homologues. Indeed, S6Ks and RSKs have different architecture and regulation mechanisms. Specifically, an unusual feature of the RSKs is the presence of a β-sheet that occupies the position of the C helix, conserved element in

classical protein kinase topology (Endicott, Noble and Johnson, 2012; Chrysostomou *et al.*, 2021). Not many studies have focused on this striking feature of the RSK structure, but it has been speculated to mediate the crass-talk among NTKD, linker region and CTKD during activation of proteins composed by two kinase domains (RSKs and MSKs) (Endicott, Noble and Johnson, 2012; Chrysostomou *et al.*, 2021). Furthermore, RSKs are found only in animals (unpublished data), which further reduces the possibility of including RSKs in comparative structure studies.

In this proof of principle study, Zn^{2+} and Mn^{2+}, but not Li^{+}, ions have been shown to form stable interactions with proteins and water molecules, independently of the isoform or species. This would highlight a conserved feature of the S6K family, as suggested by Wang *et al.*, (2013). However, due to the lack of a conventional metal binding motif (in the identified “hot spot” or anywhere else on the protein surface), it is challenging to rationalize the observed differences. Zinc is one of the most abundant metals in biology, being present in about 10% of proteins (Laitaoja *et al.*, 2013). Zinc binding to proteins is usually quite strong (Kd μM-pM) and may be required for (i) correct folding, (ii) catalysis, (iii) protein interaction (Auld, 2009). Coordination usually occurs via Cys, His, Asp and/or Glu, and, mainly in enzymes, water molecules (Maret and Li, 2009). In many zinc metalloproteins, zinc stabilises the folding of zinc finger motifs (Auld, 2009), as also observed by Wang *et al.* (2013) in the human S6K1 structure. While this could suggest that including Zinc in S6Ks expression and purification protocols may be beneficial, data from Mn^{2+} simulations (used as a control) don’t fully support a folding role of zinc. Like zinc, manganese belongs to the d-block of metal cations and, together with iron, cobalt, nickel, copper, are vital for the function of the cell (Udayalaxmi *et al.*, 2020). Manganese is preferentially coordinated by oxygen-containing amino acids (i.e., preferentially Asp and Glu, but also Asn, and Gln) (Brylinski and Skolnick, 2011), backbone carbonyl groups and

water (Dismukes, 1996). Differently from Zn^{2+}, Mn^{2+} predominantly acquires coordination number of six and five (Udayalaxmi *et al.*, 2020) and binds protein with a much lower affinity (Maret, 2010). Its role in kinase binding has been explored extensively, but only in relation to ATP coordination in the active site. Indeed, all kinases require two divalent metal ions (usually Mg^{2+}) as cofactors to phosphorylate substrates. For some kinases, like Tor and leucine-rich repeat kinase 2 (LRRK2), the role on Mn^{2+} was investigated in protein function. Interestingly, in both Tor and LRRK2, Mn^{2+} was shown to lower the *K*m for ATP, activating the kinases better than Mg^{2+} (Lovitt *et al.*, 2010; Nicastro *et al.*, 2022). The presence of Mn^{2+} binding outside the active site of kinases has never been reported. This would suggest that the similarity of results obtained for Mn^{2+} and Zn^{2+} in my simulations are likely linked to some of the chemical similarities of these two metals (with Mn^{2+} mimicking Zn^{2+}) and not to the presence of a genuine metal binding site. Indeed, Mn^{2+} and Zn^{2+} have different coordination preferences, but both metal ions can form complexes with coordination number 4 and may exhibit identical molecular connectivities within polymers (Xu *et al.*, 2020). In contradiction with what was reported for the human S6K1 (Wang *et al.*, 2013), the inability to identify major differences in the simulations for the two cations in any of the S6Ks studied would suggest the binding site is not a conserved zinc-finger-like motif.

Differently from Zn^{2+} and Mn^{2+}, lithium is not a biologically essential metal and, as a consequence, there are no lithium-specific binding sites in proteins. Instead, Lithium exerts its biological functions (e.g., it is used to treat bipolar disorder) by competing for other cation binding sites, usually sodium and magnesium (Jakobsson *et al.*, 2017). In the few instances where molecular data are available (e.g., inositol monophosphatase (IMPase), bisphosphate 3-prime-nucleotidease (BPNT1), glycogen synthase kinase 3beta (GSK3B)), Li^{+} binding induces protein inactivation by displacing Mg^{2+} from its binging site (O'Brien *et*

al., 2004; Spiegelberg *et al.*, 2005; Haimovich, Eliav and Goldbourt, 2012). As such, it is not surprising that Li^+ is not coordinated by any S6Ks in a stable manner in the putative "hot spot" for cation binding. While Li^+ can replace Na^+, sodium chloride in buffers has the role to provide counter-ions to charged residues on protein surfaces, and it is unclear how Li^+ may improve solubility in this context, as reported by Sunami *et al.* (2010). On the contrary, it could be used in substitution of Mg^{2+}, but no study has explored the benefits this could have in protein solubility.

An interesting feature observed in the simulations of all S6KS containing Zn^{2+} and Mn^{2+}, but not Li^+ or apo, is the formation of a hydration shell (i.e., long residence structural waters) (Niccolai *et al.*, 2003) directly mediated by Zn^{2+} and Mn^{2+}. It has been shown that structural waters have a role in protein stability, and, importantly, aggregation. Earlier studies on prion protein (PrPC) by De Simone *et al.* (2005) showed that mutations of exposed residues that remove long residence time water from the protein surface promoted protein unfolding (De Simone *et al.*, 2005). Although, as mentioned above, the by Zn^{2+} and Mn^{2+} may not be bound to a genuine metal binding site that induces stabilisation via folding of a zinc-finger motif, my simulations could suggest a potential stabilisation role via formation of a hydration shell localised around the activation loop of S6Ks. However, the experimental data presented in this thesis don't fully support a strong stabilisation effect for any of the metals tested. More experimental data (i.e., measurement of protein thermostability and activity in presence of different metals) are required to fully clarify the role different cations may have on S6Ks activity.

6. Conclusions

In parallel to my project, Dr I. Obomighie demonstrated that constitutively active AtS6K2 increased resistance to salt treatment compared to the WT and knockout lines (unpublished). This was observed by plants actively growing on plates (figure 4.1) and the rosette area of plants 7 days into the 14-day stress treatment. Similar results were also obtained for cold stress. As such, AtS6K2 is a key player in plant stress response. The aim of my project was to provide a structural and biochemical characterisation of this protein.

For the 23 expression constructs cloned in this work and described in this thesis, 33 different expression experiments were attempted and 26 different purification conditions tested. Although, no viable AtS6K2 protein for X-ray crystallographic studies was produced, some experiments generated promising outcomes:

(i) AtS6K2 KDN_133_414_NC_C-His_λ and KDN_133_414_C_C-His_λ constructs, co-expressed with λ-phosphatase, were expressed as soluble proteins in *E. coli*. Although, in both cases, protein yield was too low to proceed to crystallisation, expression and purification conditions are yet to be optimised. Dialysis after IMAC may prevent precipitation, and further optimisation of expression conditions and purification buffers may improve the protein yield.

(ii) The AtS6K2 FL_1_471_NC_N-His_λ T455E_S296E phosphomimetic mutant was produced, but the yield was too low for activity studies. Optimisation of expression and purification is needed. In addition, optimising the construct, i.e. moving the His-tag from the N- to the C-terminus, which proved to be beneficial for other AtS6K2 constructs, is likely to improve the yield.

The data generated in this work can be used to guide future studies regarding AtS6Ks targeting the elucidation of AtS6K2 involvement in plant stress response.

Bibliography

Albanese, S.K. *et al.* (2018) 'An Open Library of Human Kinase Domain Constructs for Automated Bacterial Expression', *Biochemistry*, 57(31), pp. 4675–4689. Available at: https://doi.org/10.1021/acs.biochem.7b01081.

Albihlal, W.S. *et al.* (2018) 'Arabidopsis HEAT SHOCK TRANSCRIPTION FACTORA1b regulates multiple developmental genes under benign and stress conditions', *Journal of Experimental Botany*, 69(11), pp. 2847–2862. Available at: https://doi.org/10.1093/jxb/ery142.

Alcázar, R. *et al.* (2006) 'Involvement of polyamines in plant response to abiotic stress', *Biotechnology Letters* [Preprint]. Available at: https://doi.org/10.1007/s10529-006-9179-3.

Ali, M.M.U. *et al.* (2011) 'Structure of the Ire1 autophosphorylation complex and implications for the unfolded protein response', *The EMBO journal*, 30(5), pp. 894–905. Available at: https://doi.org/10.1038/emboj.2011.18.

Anderson, G.H., Veit, B. and Hanson, M.R. (2005) 'The Arabidopsis AtRaptor genes are essential for post-embryonic plant growth', *BMC Biology* [Preprint]. Available at: https://doi.org/10.1186/1741-7007-3-12.

Anjum, R. and Blenis, J. (2008) 'The RSK family of kinases: Emerging roles in cellular signalling', *Nature Reviews Molecular Cell Biology* [Preprint]. Available at: https://doi.org/10.1038/nrm2509.

Arakawa, T. and Timasheff, S.N. (1985) 'Theory of protein solubility', *Methods in Enzymology*, 114, pp. 49–77. Available at: https://doi.org/10.1016/0076-6879(85)14005-x.

Assadollahi, V. *et al.* (2019) 'Interaction and molecular dynamics simulation study of Osimertinib (AstraZeneca 9291) anticancer drug with the EGFR kinase domain in native protein and mutated L844V and C797S', *Journal of Cellular Biochemistry*, 120(8), pp. 13046–13055. Available at: https://doi.org/10.1002/jcb.28575.

Auld, D.S. (2009) 'The ins and outs of biological zinc sites', *Biometals: An International Journal on the Role of Metal Ions in Biology, Biochemistry, and Medicine*, 22(1), pp. 141–148. Available at: https://doi.org/10.1007/s10534-008-9184-1.

Baena-González, E. (2010) 'Energy signaling in the regulation of gene expression during stress', *Molecular Plant* [Preprint]. Available at: https://doi.org/10.1093/mp/ssp113.

Baena-González, E. and Sheen, J. (2008) 'Convergent energy and stress signaling', *Trends in Plant Science* [Preprint]. Available at: https://doi.org/10.1016/j.tplants.2008.06.006.

Bechtold, U., Ferguson, J.N. and Mullineaux, P.M. (2018) 'To defend or to grow: Lessons from Arabidopsis C24', *Journal of Experimental Botany*, 69(11), pp. 2809–2821. Available at: https://doi.org/10.1093/jxb/ery106.

Bechtold, U. and Field, B. (2018) 'Molecular mechanisms controlling plant growth during abiotic stress', *Journal of Experimental Botany*, 69(11), pp. 2753–2758. Available at: https://doi.org/10.1093/jxb/ery157.

Beltrán-Peña, E. *et al.* (2002) 'Auxin stimulates S6 ribosomal protein phosphorylation in maize thereby affecting protein synthesis regulation', *Physiologia Plantarum*, 115(2), pp. 291–297. Available at: https://doi.org/10.1034/j.1399-3054.2002.1150216.x.

Biondi, R.M. *et al.* (2000) 'Identification of a pocket in the PDK1 kinase domain that interacts with PIF and the C–terminal residues of PKA', *The EMBO Journal*, 19(5), pp. 979–988. Available at: https://doi.org/10.1093/emboj/19.5.979.

Bita, C.E. and Gerats, T. (2013) 'Plant tolerance to high temperature in a changing environment: scientific fundamentals and production of heat stress-tolerant crops', *Frontiers in Plant Science*, 4(17), pp. 8–1. Available at: https://doi.org/10.3389/fpls.2013.00273.

Boex-Fontvieille, E. *et al.* (2013) 'Photosynthetic Control of Arabidopsis Leaf Cytoplasmic Translation Initiation by Protein Phosphorylation', *PLoS ONE*. Edited by T. Preiss, 8(7), pp. e70692–e70692. Available at: https://doi.org/10.1371/journal.pone.0070692.

Bögre, L. *et al.* (2003) 'Growth signalling pathways in Arabidopsis and the AGC protein kinases', *Trends in Plant Science* [Preprint]. Available at: https://doi.org/10.1016/S1360-1385(03)00188-2.

Braun, P. *et al.* (2002) 'Proteome-scale purification of human proteins from bacteria', *Proceedings of the National Academy of Sciences*, 99(5), pp. 2654–2659. Available at: https://doi.org/10.1073/pnas.042684199.

Brylinski, M. and Skolnick, J. (2011) 'FINDSITE-metal: Integrating evolutionary information and machine learning for structure-based metal binding site prediction at the proteome level', *Proteins*, 79(3), pp. 735–751. Available at: https://doi.org/10.1002/prot.22913.

Buchan, D.W.A. and Jones, D.T. (2019) 'The PSIPRED Protein Analysis Workbench: 20 years on', *Nucleic Acids Research*, 47(W1), pp. W402–W407. Available at: https://doi.org/10.1093/nar/gkz297.

Burgess-Brown, N.A. *et al.* (2014) 'Medium-throughput production of recombinant human proteins: protein production in E. coli', *Methods in Molecular Biology (Clifton, N.J.)*, 1091, pp. 73–94. Available at: https://doi.org/10.1007/978-1-62703-691-7_5.

Busso, D., Kim, R. and Kim, S.-H. (2003) 'Expression of soluble recombinant proteins in a cell-free system using a 96-well format', *Journal of Biochemical and Biophysical Methods*, 55(3), pp. 233–240. Available at: https://doi.org/10.1016/s0165-022x(03)00049-6.

Caldana, C. *et al.* (2013) 'Systemic analysis of inducible target of rapamycin mutants reveal a general metabolic switch controlling growth in Arabidopsis thaliana', *Plant Journal* [Preprint]. Available at: https://doi.org/10.1111/tpj.12080.

Carson, M. *et al.* (2007) 'His-tag impact on structure', *Acta Crystallographica. Section D, Biological Crystallography*, 63(Pt 3), pp. 295–301. Available at: https://doi.org/10.1107/S0907444906052024.

Champion, A. *et al.* (2004) 'Arabidopsis kinome: after the casting', *Functional & Integrative Genomics*, 4(3), pp. 163–187. Available at: https://doi.org/10.1007/s10142-003-0096-4.

Chang, I.-F. *et al.* (2005) 'Proteomic characterization of evolutionarily conserved and variable proteins of Arabidopsis cytosolic ribosomes', *Plant Physiology*, 137(3), pp. 848–862. Available at: https://doi.org/10.1104/pp.104.053637.

Chen, G.-H. *et al.* (2018) 'TOR and RPS6 transmit light signals to enhance protein translation in deetiolating Arabidopsis seedlings', *Proceedings of the National Academy of Sciences of the United States of America*, 115(50), pp. 12823–12828. Available at: https://doi.org/10.1073/pnas.1809526115.

Chrysostomou, S. *et al.* (2021) 'Repurposed floxacins targeting RSK4 prevent chemoresistance and metastasis in lung and bladder cancer', *Science Translational Medicine*, 13(602), p. eaba4627. Available at: https://doi.org/10.1126/scitranslmed.aba4627.

Conde, R. *et al.* (2004) 'A search for hyperglycosylation signals in yeast glycoproteins', *The Journal of Biological Chemistry*, 279(42), pp. 43789–43798. Available at: https://doi.org/10.1074/jbc.M406678200.

Corona-Sánchez, I. *et al.* (2019) 'Cyclodipeptides from Pseudomonas aeruginosa modulate themaize (Zeamays L.) root system and promote S6 ribosomal protein kinase activation', *PeerJ*, 2019(8), pp. e7494–e7494. Available at: https://doi.org/10.7717/peerj.7494.

Craft, J.W. and Legge, G.B. (2005) 'An AMBER/DYANA/MOLMOL Phosphorylated Amino Acid Library Set and Incorporation into NMR Structure Calculations', *Journal of Biomolecular NMR*, 33(1), pp. 15–24. Available at: https://doi.org/10.1007/s10858-005-1199-0.

Creff, A., Sormani, R. and Desnos, T. (2010) 'The two Arabidopsis RPS6 genes, encoding for cytoplasmic ribosomal proteins S6, are functionally equivalent', *Plant Molecular Biology*, 73(4), pp. 533–546. Available at: https://doi.org/10.1007/s11103-010-9639-y.

D.A. Case *et al.* (2022) 'Amber 2022'. University of California, San Francisco.

Dan, H., Balachandran, A. and Lin, M. (2009) 'A Pair of Ligation-Independent Escherichia coli Expression Vectors for Rapid Addition of a Polyhistidine Affinity Tag to the N- or C-Termini of Recombinant Proteins', *Journal of Biomolecular Techniques : JBT*, 20(5), pp. 241–248.

Davis, G.D. *et al.* (1999) 'New fusion protein systems designed to give soluble expression in Escherichia coli', *Biotechnology and Bioengineering*, 65(4), pp. 382–388.

De Simone, A. *et al.* (2005) 'Prion and water: Tight and dynamical hydration sites have a key role in structural stability', *Proceedings of the National Academy of Sciences*, 102(21), pp. 7535–7540. Available at: https://doi.org/10.1073/pnas.0501748102.

Degorce, F. *et al.* (2009) 'HTRF: A Technology Tailored for Drug Discovery –A Review of Theoretical Aspects and Recent Applications', *Current Chemical Genomics*, 3, pp. 22–32. Available at: https://doi.org/10.2174/1875397300903010022.

Deprost, D. *et al.* (2005) 'An Arabidopsis homolog of RAPTOR/KOG1 is essential for early embryo development', *Biochemical and Biophysical Research Communications* [Preprint]. Available at: https://doi.org/10.1016/j.bbrc.2004.11.117.

Deprost, D. *et al.* (2007) 'The Arabidopsis TOR kinase links plant growth, yield, stress resistance and mRNA translation.', *EMBO reports*, 8(9), pp. 864–870. Available at: https://doi.org/10.1038/sj.embor.7401043.

Derewenda, Z.S. (2004) 'The use of recombinant methods and molecular engineering in protein crystallization', *Methods*, 34(3), pp. 354–363. Available at: https://doi.org/10.1016/j.ymeth.2004.03.024.

Dinkova, T.D. *et al.* (2007) 'Dissecting the TOR-S6K signal transduction pathway in maize seedlings: Relevance on cell growth regulation', *Physiologia Plantarum*, 130(1), pp. 1–10. Available at: https://doi.org/10.1111/j.1399-3054.2007.00862.x.

Dismukes, G.C. (1996) 'Manganese Enzymes with Binuclear Active Sites', *Chemical Reviews*, 96(7), pp. 2909–2926. Available at: https://doi.org/10.1021/cr950053c.

Dobrenel, T. *et al.* (2011) 'Regulation of plant growth and metabolism by the TOR kinase', *Biochemical Society Transactions* [Preprint]. Available at: https://doi.org/10.1042/BST0390477.

Dobrenel, T., Mancera-Martínez, E., *et al.* (2016) 'The Arabidopsis TOR Kinase Specifically Regulates the Expression of Nuclear Genes Coding for Plastidic Ribosomal Proteins and the Phosphorylation of the Cytosolic Ribosomal Protein S6', *Frontiers in Plant Science*, 7, pp. 1611–1611. Available at: https://doi.org/10.3389/fpls.2016.01611.

Dobrenel, T., Caldana, C., *et al.* (2016) 'TOR Signaling and Nutrient Sensing', *Annual Review of Plant Biology* [Preprint]. Available at: https://doi.org/10.1146/annurev-arplant-043014-114648.

Drexler, C. (2006) 'HTRF® KinEASETM: a new solution for screening serine-threonine kinases', *Nature Methods*, 3(6), p. I–II. Available at: https://doi.org/10.1038/nmeth883.

Du, Z. *et al.* (2021) 'The trRosetta server for fast and accurate protein structure prediction', *Nature Protocols*, 16(12), pp. 5634–5651. Available at: https://doi.org/10.1038/s41596-021-00628-9.

Dyson, M.R. *et al.* (2004) 'Production of soluble mammalian proteins in Escherichia coli: identification of protein features that correlate with successful expression', *BMC Biotechnology*, 4, p. 32. Available at: https://doi.org/10.1186/1472-6750-4-32.

Endicott, J.A., Noble, M.E.M. and Johnson, L.N. (2012) 'The Structural Basis for Control of Eukaryotic Protein Kinases', *Annual Review of Biochemistry* [Preprint]. Available at: https://doi.org/10.1146/annurev-biochem-052410-090317.

Enganti, R. *et al.* (2018) 'Phosphorylation of Ribosomal Protein RPS6 Integrates Light Signals and Circadian Clock Signals', *Frontiers in Plant Science*, 8, pp. 2210–2210. Available at: https://doi.org/10.3389/fpls.2017.02210.

Enomoto, S., Chen, G. and Berman, J. (1998) 'Vectors for Expressing T7 Epitope- and His6 Affinity-Tagged Fusion Proteins in S. cerevisiae', *BioTechniques*, 24(5), pp. 782–788. Available at: https://doi.org/10.2144/98245st01.

Eschenfeldt, W.H. *et al.* (2010) 'Cleavable C-terminal His-tag vectors for structure determination', *Journal of structural and functional genomics*, 11(1), pp. 31–39. Available at: https://doi.org/10.1007/s10969-010-9082-y.

Esposito, D. and Chatterjee, D.K. (2006) 'Enhancement of soluble protein expression through the use of fusion tags', *Current Opinion in Biotechnology*, 17(4), pp. 353–358. Available at: https://doi.org/10.1016/j.copbio.2006.06.003.

Felberbaum, R.S. (2015) 'The baculovirus expression vector system: A commercial manufacturing platform for viral vaccines and gene therapy vectors', *Biotechnology Journal*, 10(5), pp. 702–714. Available at: https://doi.org/10.1002/biot.201400438.

Fenton, T.R. *et al.* (2010) 'Histone acetyltransferases interact with and acetylate p70 ribosomal S6 kinases in vitro and in vivo', *The International Journal of Biochemistry & Cell Biology*, 42(2), pp. 359–366. Available at: https://doi.org/10.1016/j.biocel.2009.11.022.

Fenton, T.R. and Gout, I.T. (2011) 'Functions and regulation of the 70 kDa ribosomal S6 kinases', *International Journal of Biochemistry and Cell Biology* [Preprint]. Available at: https://doi.org/10.1016/j.biocel.2010.09.018.

Filippakopoulos, P. *et al.* (2008) 'Structural Coupling of SH2-Kinase Domains Links Fes and Abl Substrate Recognition and Kinase Activation', *Cell*, 134(5), pp. 793–803. Available at: https://doi.org/10.1016/j.cell.2008.07.047.

Francis, D. (2007) 'The plant cell cycle - 15 Years on', *New Phytologist* [Preprint]. Available at: https://doi.org/10.1111/j.1469-8137.2007.02038.x.

Francis, D.M. and Page, R. (2010) 'Strategies to Optimize Protein Expression in E. coli', *Current Protocols in Protein Science*, 61(1), pp. 5241–52429. Available at: https://doi.org/10.1002/0471140864.ps0524s61.

Frödin, M. *et al.* (2002) 'A phosphoserine/threonine-binding pocket in AGC kinases and PDK1 mediates activation by hydrophobic motif phosphorylation', *The EMBO journal*, 21(20), pp. 5396–5407. Available at: https://doi.org/10.1093/emboj/cdf551.

Garcia Florez, C. *et al.* (2001) 'A maize insulin-like growth factor signals to a transduction pathway that regulates protein synthesis in maize', *Biochemical Journal*, 358(1), pp. 95–100. Available at: https://doi.org/10.1042/bj3580095.

Garrocho-Villegas, V., Aguilar C, R. and Sánchez de Jiménez, E. (2013) 'Insights into the TOR-S6K Signaling Pathway in Maize (Zea mays L.). Pathway Activation by Effector–Receptor Interaction', *Biochemistry*, 52(51), pp. 9129–9140. Available at: https://doi.org/10.1021/bi401474x.

Gasteiger, E. *et al.* (2005) 'Protein Identification and Analysis Tools on the ExPASy Server', in J.M. Walker (ed.) *The Proteomics Protocols Handbook*. Totowa, NJ: Humana Press (Springer Protocols Handbooks), pp. 571–607. Available at: https://doi.org/10.1385/1-59259-890-0:571.

Gazi, R., Maity, S. and Jana, M. (2023) 'Conformational Features and Hydration Dynamics of Proteins in Cosolvents: A Perspective from Computational Approaches', *ACS omega*, 8(3), pp. 2832–2843. Available at: https://doi.org/10.1021/acsomega.2c08009.

Geerlof, A. *et al.* (2006) 'The impact of protein characterization in structural proteomics', *Acta Crystallographica Section D: Biological Crystallography*, 62(10), pp. 1125–1136. Available at: https://doi.org/10.1107/S0907444906030307.

González, A. *et al.* (2015) 'TORC1 promotes phosphorylation of ribosomal protein S6 via the AGC kinase Ypk3 in Saccharomyces cerevisiae', *PloS One*, 10(3), p. e0120250. Available at: https://doi.org/10.1371/journal.pone.0120250.

González, O. *et al.* (2017) 'Non-ribosomal Peptide Synthases from Pseudomonas aeruginosa Play a Role in Cyclodipeptide Biosynthesis, Quorum-Sensing Regulation, and Root Development in a Plant Host', *Microbial Ecology*, 73(3), pp. 616–629. Available at: https://doi.org/10.1007/s00248-016-0896-4.

Gornall, J. *et al.* (2010) 'Implications of climate change for agricultural productivity in the early twenty-first century', *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365(1554), pp. 2973–2989. Available at: https://doi.org/10.1098/rstb.2010.0158.

Goulet, M.-C. *et al.* (2019) 'Production of Biopharmaceuticals in Nicotiana benthamiana—Axillary Stem Growth as a Key Determinant of Total Protein Yield', *Frontiers in Plant Science*, 10. Available at: https://www.frontiersin.org/articles/10.3389/fpls.2019.00735 (Accessed: 24 March 2023).

Granados, R.R. *et al.* (1994) 'A new insect cell line from Trichoplusia ni (BTI-Tn-5B1-4) susceptible to Trichoplusia ni single enveloped nuclear polyhedrosis virus', *Journal of Invertebrate Pathology*, 64(3), pp. 260–266. Available at: https://doi.org/10.1016/S0022-2011(94)90400-6.

Gressner, A.M. and Wool, I.G. (1974a) 'The phosphorylation of liver ribosomal proteins in vivo. Evidence that only a single small subunit protein (S6) is phosphorylated', *Journal of Biological Chemistry* [Preprint].

Gressner, A.M. and Wool, I.G. (1974b) 'The stimulation of the phosphorylation of ribosomal protein S6 by cycloheximide and puromycin', *Biochemical and Biophysical Research Communications* [Preprint]. Available at: https://doi.org/10.1016/0006-291X(74)90365-9.

Grodsky, N. *et al.* (2006) 'Structure of the catalytic domain of human protein kinase C beta II complexed with a bisindolylmaleimide inhibitor', *Biochemistry*, 45(47), pp. 13970–13981. Available at: https://doi.org/10.1021/bi061128h.

Guijarro-Clarke, C., Holland, P.W.H. and Paps, J. (2020) 'Publisher Correction: Widespread patterns of gene loss in the evolution of the animal kingdom', *Nature Ecology & Evolution*, 4(4), pp. 661–661. Available at: https://doi.org/10.1038/s41559-020-1159-9.

Gwalter, J., Wang, M.-L. and Gout, I. (2009) 'The ubiquitination of ribosomal S6 kinases is independent from the mitogen-induced phosphorylation/activation of the kinase', *The International Journal of Biochemistry & Cell Biology*, 41(4), pp. 828–833. Available at: https://doi.org/10.1016/j.biocel.2008.08.018.

Haimovich, A., Eliav, U. and Goldbourt, A. (2012) 'Determination of the lithium binding site in inositol monophosphatase, the putative target for lithium therapy, by magic-angle-spinning solid-state NMR', *Journal of the American Chemical Society*, 134(12), pp. 5647–5651. Available at: https://doi.org/10.1021/ja211794x.

Hammarström, M. *et al.* (2002) 'Rapid screening for improved solubility of small human proteins produced as fusion proteins in Escherichia coli', *Protein Science : A Publication of the Protein Society*, 11(2), p. 313. Available at: https://doi.org/10.1110/ps.22102.

Heitman, J., Movva, N.R. and Hall, M.N. (1991) 'Targets for cell cycle arrest by the immunosuppressant rapamycin in yeast', *Science* [Preprint]. Available at: https://doi.org/10.1126/science.1715094.

Henriques, R., Magyar, Z., Monardes, A., Khan, S., Zalejski, C., Orellana, J., Szabados, L., de la Torre, C., *et al.* (2010) 'Arabidopsis S6 kinase mutants display chromosome instability and altered RBR1–E2F pathway activity', *The EMBO Journal*, 29(17), pp. 2979–2993. Available at: https://doi.org/10.1038/emboj.2010.164.

van den Heuvel, S. and Dyson, N.J. (2008) 'Conserved functions of the pRB and E2F families', *Nature Reviews. Molecular Cell Biology*, 9(9), pp. 713–724. Available at: https://doi.org/10.1038/nrm2469.

Hollingsworth, S.A. and Dror, R.O. (2018) 'Molecular Dynamics Simulation for All', *Neuron*, 99(6), pp. 1129–1143. Available at: https://doi.org/10.1016/j.neuron.2018.08.011.

Homma, K. and Moriyama, H. (2009) 'Crystallization and crystal-packing studies of Chlorella virus deoxyuridine triphosphatase', *Acta Crystallographica Section F: Structural Biology and Crystallization Communications*, 65(10), pp. 1030–1034. Available at: https://doi.org/10.1107/S1744309109034459.

Hunter, T. (1995) 'Protein kinases and phosphatases: the yin and yang of protein phosphorylation and signaling', *Cell*, 80(2), pp. 225–236. Available at: https://doi.org/10.1016/0092-8674(95)90405-0.

Huse, M. and Kuriyan, J. (2002) 'The conformational plasticity of protein kinases', *Cell*, 109(3), pp. 275–282. Available at: https://doi.org/10.1016/s0092-8674(02)00741-9.

Ikuta, M. *et al.* (2007) 'Crystal structures of the N-terminal kinase domain of human RSK1 bound to three different ligands: Implications for the design of RSK1 specific inhibitors', *Protein Science*, 16(12), pp. 2626–2635. Available at: https://doi.org/10.1110/ps.073123707.

Jacinto, E. and Hall, M.N. (2003) 'TOR signalling in bugs, brain and brawn', *Nature Reviews Molecular Cell Biology* [Preprint]. Available at: https://doi.org/10.1038/nrm1018.

Jacobson, M.P. *et al.* (2004) 'A hierarchical approach to all-atom protein loop prediction', *Proteins: Structure, Function, and Bioinformatics*, 55(2), pp. 351–367. Available at: https://doi.org/10.1002/prot.10613.

Jakobsson, E. *et al.* (2017) 'Towards a Unified Understanding of Lithium Action in Basic Biology and its Significance for Applied Biology', *The Journal of Membrane Biology*, 250(6), pp. 587–604. Available at: https://doi.org/10.1007/s00232-017-9998-2.

Jarvik, J.W. and Telmer, C.A. (1998) 'Epitope tagging', *Annual Review of Genetics*, 32, pp. 601–618. Available at: https://doi.org/10.1146/annurev.genet.32.1.601.

Jenö, P. *et al.* (1988) 'Identification and characterization of a mitogen-activated S6 kinase.', *Proceedings of the National Academy of Sciences of the United States of America* [Preprint]. Available at: https://doi.org/10.1073/pnas.85.2.406.

Jiménez, E.S. de, Beltrán-Peña, E. and Ortíz-López, A. (1999) 'Insulin-stimulated ribosomal protein synthesis in maize embryonic axes during germination', *Physiologia Plantarum*, 105(1), pp. 148–154. Available at: https://doi.org/10.1034/j.1399-3054.1999.105122.x.

Jones, D.T. (1999) 'Protein secondary structure prediction based on position-specific scoring matrices', *Journal of Molecular Biology*, 292(2), pp. 195–202. Available at: https://doi.org/10.1006/jmbi.1999.3091.

Jumper, J. *et al.* (2021) 'Highly accurate protein structure prediction with AlphaFold', *Nature*, 596(7873), pp. 583–589. Available at: https://doi.org/10.1038/s41586-021-03819-2.

Kaltheuner, I.H. *et al.* (2021) 'Abemaciclib is a potent inhibitor of DYRK1A and HIP kinases involved in transcriptional regulation', *Nature Communications*, 12(1), p. 6607. Available at: https://doi.org/10.1038/s41467-021-26935-z.

Kim, D.E. *et al.* (2023) 'De novo design of small beta barrel proteins', *Proceedings of the National Academy of Sciences of the United States of America*, 120(11), p. e2207974120. Available at: https://doi.org/10.1073/pnas.2207974120.

Kim, T.-H. *et al.* (2004) 'Translational regulation via 5' mRNA leader sequences revealed by mutational analysis of the Arabidopsis translation initiation factor subunit eIF3h', *The Plant Cell*, 16(12), pp. 3341–3356. Available at: https://doi.org/10.1105/tpc.104.026880.

Kim, Y.K. *et al.* (2014) 'Ribosomal protein s6, a target of rapamycin, is involved in the regulation of rRNA genes by possible epigenetic changes in arabidopsis', *Journal of Biological Chemistry*, 289(7), pp. 3901–3912. Available at: https://doi.org/10.1074/jbc.M113.515015.

Kimple, M.E., Brill, A.L. and Pasker, R.L. (2013) 'Overview of Affinity Tags for Protein Purification', *Current protocols in protein science / editorial board, John E. Coligan ... [et al.]*, 73, p. Unit-9.9. Available at: https://doi.org/10.1002/0471140864.ps0909s73.

Kimura, Y. *et al.* (2005) 'Brevicompanine C, Cyclo-(d-Ile-l-Trp), and Cyclo-(d-Leu-l-Trp), Plant Growth Regulators from Penicillium brevi-compactum', *Journal of Natural Products*, 68(2), pp. 237–239. Available at: https://doi.org/10.1021/np040178p.

Kliebenstein, D.J. (2016) 'False idolatry of the mythical growth versus immunity tradeoff in molecular systems plant pathology', *Physiological and Molecular Plant Pathology*, 95, pp. 55–59. Available at: https://doi.org/10.1016/j.pmpp.2016.02.004.

Kothe, M. *et al.* (2007) 'Structure of the Catalytic Domain of Human Polo-like Kinase 1', *Biochemistry*, 46(20), pp. 5960–5971. Available at: https://doi.org/10.1021/bi602474j.

Kozma, S.C. and Thomas, G. (2002) 'Regulation of cell size in growth, development and human disease: PI3K, PKB and S6K', *BioEssays* [Preprint]. Available at: https://doi.org/10.1002/bies.10031.

Krantz, A. (1998) 'Probing protein surfaces for `hot spots': a new frontier', *Trends in Biotechnology*, 16(5), pp. 198–199. Available at: https://doi.org/10.1016/S0167-7799(98)01177-9.

Laitaoja, M., Valjakka, J. and Jänis, J. (2013) 'Zinc coordination spheres in protein structures', *Inorganic Chemistry*, 52(19), pp. 10983–10991. Available at: https://doi.org/10.1021/ic401072d.

Lee, D.H. *et al.* (2017) 'MRF family genes are involved in translation control, especially under energy-deficient conditions, and their expression and functions are modulated by the TOR signaling pathway', *Plant Cell* [Preprint]. Available at: https://doi.org/10.1105/tpc.17.00563.

Lehti-Shiu, M.D. and Shiu, S.-H. (2012) 'Diversity, classification and function of the plant protein kinase superfamily', *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367(1602), pp. 2619–2639. Available at: https://doi.org/10.1098/rstb.2012.0003.

Leppek, K., Das, R. and Barna, M. (2018) 'Functional 5' UTR mRNA structures in eukaryotic translation regulation and how to find them', *Nature Reviews Molecular Cell Biology* [Preprint]. Available at: https://doi.org/10.1038/nrm.2017.103.

Leuzinger, K. *et al.* (2013) 'Efficient Agroinfiltration of Plants for High-level Transient Expression of Recombinant Proteins', *JoVE (Journal of Visualized Experiments)*, (77), p. e50521. Available at: https://doi.org/10.3791/50521.

Levy, S. *et al.* (1991) 'Oligopyrimidine tract at the 5' end of mammalian ribosomal protein mRNAs is required for their translational control', *Proceedings of the National Academy of Sciences of the United States of America*, 88(8), pp. 3319–3323. Available at: https://doi.org/10.1073/pnas.88.8.3319.

Li, T. *et al.* (2020) 'A gain-of-function mutation in Brassinosteroid-insensitive 2 alters Arabidopsis floral organ development by altering auxin levels', *Plant Cell Reports*, 39(2), pp. 259–271. Available at: https://doi.org/10.1007/s00299-019-02489-9.

Li, X. *et al.* (2017) 'Differential TOR activation and cell proliferation in Arabidopsis root and shoot apexes', *Proceedings of the National Academy of Sciences of the United States of America* [Preprint]. Available at: https://doi.org/10.1073/pnas.1618782114.

Liao, X.H. *et al.* (2008) 'Growth control via TOR kinase signaling, an intracellular sensor of amino acid and energy availability, with crosstalk potential to proline metabolism', *Amino Acids* [Preprint]. Available at: https://doi.org/10.1007/s00726-008-0100-3.

Lin, K. *et al.* (2012) 'An ATP-site on-off switch that restricts phosphatase accessibility of Akt', *Science Signaling*, 5(223), p. ra37. Available at: https://doi.org/10.1126/scisignal.2002618.

Lovitt, B. *et al.* (2010) 'Differential effects of divalent manganese and magnesium on the kinase activity of leucine-rich repeat kinase 2 (LRRK2)', *Biochemistry*, 49(14), pp. 3092–3100. Available at: https://doi.org/10.1021/bi901726c.

Madhavi Sastry, G. *et al.* (2013) 'Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments', *Journal of Computer-Aided Molecular Design*, 27(3), pp. 221–234. Available at: https://doi.org/10.1007/s10822-013-9644-8.

Maegawa, K. *et al.* (2015) 'Evolutionary conservation of TORC1 components, TOR, Raptor, and LST8, between rice and yeast', *Molecular genetics and genomics: MGG*, 290(5), pp. 2019–2030. Available at: https://doi.org/10.1007/s00438-015-1056-0.

Mahajan, P. *et al.* (2014) 'Medium-Throughput Production of Recombinant Human Proteins: Protein Production in Insect Cells', in Y.W. Chen (ed.) *Structural Genomics: General Applications*. Totowa, NJ: Humana Press (Methods in Molecular Biology), pp. 95–121. Available at: https://doi.org/10.1007/978-1-62703-691-7_6.

Mahfouz, M.M. *et al.* (2006) 'Arabidopsis TARGET of RAPAMYCIN interacts with RAPTOR, which regulates the activity of S6 kinase in response to osmotic stress signals', *Plant Cell* [Preprint]. Available at: https://doi.org/10.1105/tpc.105.035931.

Malakhov, M.P. *et al.* (2004) 'SUMO fusions and SUMO-specific protease for efficient expression and purification of proteins', *Journal of Structural and Functional Genomics*, 5(1–2), pp. 75–86. Available at: https://doi.org/10.1023/B:JSFG.0000029237.70316.52.

Malakhova, M. *et al.* (2009) 'Structural Diversity of the Active N-Terminal Kinase Domain of p90 Ribosomal S6 Kinase 2', *PLOS ONE*, 4(11), p. e8044. Available at: https://doi.org/10.1371/journal.pone.0008044.

Manning, B.D. and Cantley, L.C. (2007) 'AKT/PKB signaling: navigating downstream', *Cell*, 129(7), pp. 1261–1274. Available at: https://doi.org/10.1016/j.cell.2007.06.009.

Manning, G. *et al.* (2002) 'The protein kinase complement of the human genome', *Science* [Preprint]. Available at: https://doi.org/10.1126/science.1075762.

Maret, W. (2010) 'Metalloproteomics, metalloproteomes, and the annotation of metalloproteins', *Metallomics*, 2(2), pp. 117–125. Available at: https://doi.org/10.1039/B915804A.

Maret, W. and Li, Y. (2009) 'Coordination dynamics of zinc in proteins', *Chemical Reviews*, 109(10), pp. 4682–4707. Available at: https://doi.org/10.1021/cr800556u.

Margalha, L., Confraria, A. and Baena-González, E. (2019) 'SnRK1 and TOR: Modulating growth–defense trade-offs in plant stress responses', *Journal of Experimental Botany* [Preprint]. Available at: https://doi.org/10.1093/jxb/erz066.

McPherson, A. (2017) 'Protein Crystallization', in A. Wlodawer, Z. Dauter, and M. Jaskolski (eds) *Protein Crystallography: Methods and Protocols*. New York, NY: Springer (Methods in Molecular Biology), pp. 17–50. Available at: https://doi.org/10.1007/978-1-4939-7000-1_2.

Meagher, K.L., Redman, L.T. and Carlson, H.A. (2003) 'Development of polyphosphate parameters for use with the AMBER force field', *Journal of Computational Chemistry*, 24(9), pp. 1016–1025. Available at: https://doi.org/10.1002/jcc.10262.

Meharena, H.S. *et al.* (2013) 'Deciphering the Structural Basis of Eukaryotic Protein Kinase Regulation', *PLOS Biology*, 11(10), p. e1001680. Available at: https://doi.org/10.1371/journal.pbio.1001680.

Menand, B. *et al.* (2002) 'Expression and disruption of the Arabidopsis TOR (target of rapamycin) gene', *Proceedings of the National Academy of Sciences of the United States of America* [Preprint]. Available at: https://doi.org/10.1073/pnas.092141899.

Mendoza, M.C., Er, E.E. and Blenis, J. (2011) 'The Ras-ERK and PI3K-mTOR pathways: cross-talk and compensation', *Trends in Biochemical Sciences*, 36(6), pp. 320–328. Available at: https://doi.org/10.1016/j.tibs.2011.03.006.

Meyuhas, O. (2015) 'Ribosomal Protein S6 Phosphorylation: Four Decades of Research.', *International review of cell and molecular biology*, 320, pp. 41–73. Available at: https://doi.org/10.1016/bs.ircmb.2015.07.006.

Miller, C.J. and Turk, B.E. (2018) 'Homing in: Mechanisms of Substrate Targeting by Protein Kinases', *Trends in Biochemical Sciences*, 43(5), pp. 380–394. Available at: https://doi.org/10.1016/j.tibs.2018.02.009.

Mittler, R., Finka, A. and Goloubinoff, P. (2012) 'How do plants feel the heat?', *Trends in Biochemical Sciences*, 37(3), pp. 118–125. Available at: https://doi.org/10.1016/j.tibs.2011.11.007.

Mizoguchi, T. *et al.* (1995) 'Two genes that encode ribosomal-protein S6 kinase homologs are induced by cold or salinity stress in Arabidopsis thaliana', *FEBS Letters* [Preprint]. Available at: https://doi.org/10.1016/0014-5793(94)01423-X.

Mockler, T.C. *et al.* (2007) 'The diurnal project: Diurnal and circadian expression profiling, model-based pattern matching, and promoter analysis', in *Cold Spring Harbor Symposia on Quantitative Biology*. Available at: https://doi.org/10.1101/sqb.2007.72.006.

Moreau, M. *et al.* (2012) 'Mutations in the Arabidopsis homolog of LST8/GβL, a partner of the target of Rapamycin kinase, impair plant growth, flowering, and metabolic adaptation to long days', *The Plant cell*. 2012/02/03 edn, 24(2), pp. 463–481. Available at: https://doi.org/10.1105/tpc.111.091306.

Nallamsetty, S. and Waugh, D.S. (2007) 'Mutations that alter the equilibrium between open and closed conformations of Escherichia coli maltose-binding protein impede its ability to enhance the solubility of passenger proteins', *Biochemical and biophysical research communications*, 364(3), pp. 639–644. Available at: https://doi.org/10.1016/j.bbrc.2007.10.060.

Nemenoff, R.A. *et al.* (1988) 'An S6 kinase activated during liver regeneration is related to the insulin-stimulated S6 kinase in H4 hepatoma cells', *Journal of Biological Chemistry* [Preprint]. Available at: https://doi.org/10.1016/s0021-9258(19)77655-8.

Nicastro, R. *et al.* (2022) 'Manganese is a physiologically relevant TORC1 activator in yeast and mammals', *eLife*, 11, p. e80497. Available at: https://doi.org/10.7554/eLife.80497.

Niccolai, N. *et al.* (2003) 'NMR studies of protein hydration and TEMPOL accessibility', *Journal of Molecular Biology*, 332(2), pp. 437–447. Available at: https://doi.org/10.1016/s0022-2836(03)00852-0.

Nishi, H., Shaytan, A. and Panchenko, A.R. (2014) 'Physicochemical mechanisms of protein regulation by phosphorylation', *Frontiers in Genetics*, 5. Available at: https://www.frontiersin.org/articles/10.3389/fgene.2014.00270 (Accessed: 1 February 2023).

Nolen, B., Taylor, S. and Ghosh, G. (2004) 'Regulation of protein kinases; controlling activity through activation segment conformation', *Molecular Cell*, 15(5), pp. 661–675. Available at: https://doi.org/10.1016/j.molcel.2004.08.024.

Norkunas, K. *et al.* (2018) 'Improving agroinfiltration-based transient gene expression in Nicotiana benthamiana', *Plant Methods*, 14(1), p. 71. Available at: https://doi.org/10.1186/s13007-018-0343-2.

Nukarinen, E. *et al.* (2016) 'Quantitative phosphoproteomics reveals the role of the AMPK plant ortholog SnRK1 as a metabolic master regulator under energy deprivation', *Scientific Reports* [Preprint]. Available at: https://doi.org/10.1038/srep31697.

Obomighie, I. *et al.* (2021) 'The Role of Ribosomal Protein S6 Kinases in Plant Homeostasis', *Frontiers in Molecular Biosciences*, 8. Available at: https://doi.org/10.3389/fmolb.2021.636560.

O'Brien, W.T. *et al.* (2004) 'Glycogen synthase kinase-3beta haploinsufficiency mimics the behavioral and molecular effects of lithium', *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 24(30), pp. 6791–6798. Available at: https://doi.org/10.1523/JNEUROSCI.4753-03.2004.

Ortiz-Castro, R. *et al.* (2011) 'Transkingdom signaling based on bacterial cyclodipeptides with auxin activity in plants', *Proceedings of the National Academy of Sciences*, 108(17), pp. 7253–7258. Available at: https://doi.org/10.1073/pnas.1006740108.

Ortiz-Castro, R., Campos-García, J. and López-Bucio, J. (2019) 'Pseudomonas putida and Pseudomonas fluorescens Influence Arabidopsis Root System Architecture Through an Auxin Response Mediated by Bioactive Cyclodipeptides', *Journal of Plant Growth Regulation*, 39(1), pp. 254–265. Available at: https://doi.org/10.1007/s00344-019-09979-w.

Pardo, O.E. and Seckl, M.J. (2013) 'S6K2: The Neglected S6 Kinase Family Member', *Frontiers in Oncology* [Preprint]. Available at: https://doi.org/10.3389/fonc.2013.00191.

Pearce, L.R., Komander, D. and Alessi, D.R. (2010) 'The nuts and bolts of AGC protein kinases', *Nature Reviews Molecular Cell Biology* [Preprint]. Available at: https://doi.org/10.1038/nrm2822.

Peña-Uribe, C.A. *et al.* (2012) 'Oligogalacturonides inhibit growth and induce changes in S6K phosphorylation in maize (Zea mays L. var. Chalqueño)', *Plant Growth Regulation*, 67(2), pp. 151–159. Available at: https://doi.org/10.1007/s10725-012-9672-8.

Peng, P. *et al.* (2008) 'Regulation of the Arabidopsis GSK3-like kinase BRASSINOSTEROID-INSENSITIVE 2 through proteasome-mediated protein degradation', *Molecular Plant*, 1(2), pp. 338–346. Available at: https://doi.org/10.1093/mp/ssn001.

Pengelley, S.C. *et al.* (2006) 'A suite of parallel vectors for baculovirus expression', *Protein Expression and Purification*, 48(2), pp. 173–181. Available at: https://doi.org/10.1016/j.pep.2006.04.016.

Peti, W. and Page, R. (2007) 'Strategies to maximize heterologous protein expression in Escherichia coli with minimal cost', *Protein Expression and Purification*, 51(1), pp. 1–10. Available at: https://doi.org/10.1016/j.pep.2006.06.024.

Pinna, L.A. and Ruzzene, M. (1996) 'How do protein kinases recognize their substrates?', *Biochimica Et Biophysica Acta*, 1314(3), pp. 191–225. Available at: https://doi.org/10.1016/s0167-4889(96)00083-3.

Prischi, F. *et al.* (2014) 'Phosphoregulation of Ire1 RNase splicing activity', *Nature Communications*, 5, p. 3554. Available at: https://doi.org/10.1038/ncomms4554.

Qin, J. *et al.* (2015) 'Development of Organometallic S6K1 Inhibitors', *Journal of Medicinal Chemistry*, 58(1), pp. 305–314. Available at: https://doi.org/10.1021/jm5011868.

Rademacher, E.H. and Offringa, R. (2012) 'Evolutionary adaptations of plant AGC kinases: From light signaling to cell polarity regulation', *Frontiers in Plant Science* [Preprint]. Available at: https://doi.org/10.3389/fpls.2012.00250.

Rajamäki, M.L. *et al.* (2017) 'Differential requirement of the ribosomal protein s6 and ribosomal protein s6 kinase for plant-virus accumulation and interaction of s6 kinase with potyviral Vpg', *Molecular Plant-Microbe Interactions* [Preprint]. Available at: https://doi.org/10.1094/MPMI-06-16-0122-R.

Ren, G., Ke, N. and Berkmen, M. (2016) 'Use of the SHuffle Strains in Production of Proteins', *Current Protocols in Protein Science*, 85, p. 5.26.1-5.26.21. Available at: https://doi.org/10.1002/cpps.11.

Ren, M. *et al.* (2013) 'Target of rapamycin signaling regulates metabolism, growth, and life Span in arabidopsis', *Plant Cell*, 24(12), pp. 4850–4874. Available at: https://doi.org/10.1105/tpc.112.107144.

Rexin, D. *et al.* (2015) 'TOR signalling in plants', *Biochemical Journal* [Preprint]. Available at: https://doi.org/10.1042/BJ20150505.

Reyes de la Cruz, H., Aguilar, R. and Sánchez de Jiménez, E. (2004) 'Functional Characterization of a Maize Ribosomal S6 Protein Kinase (ZmS6K), a Plant Ortholog of Metazoan p70S6K', *Biochemistry*, 43(2), pp. 533–539. Available at: https://doi.org/10.1021/bi035222z.

Rhoads, R.E., Dinkova, T.D. and Jagus, R. (2007) 'Approaches for Analyzing the Differential Activities and Functions of eIF4E Family Members', in *Methods in Enzymology*. Available at: https://doi.org/10.1016/S0076-6879(07)29013-5.

Robaglia, C. *et al.* (2004) 'Plant growth: The translational connection', in *Biochemical Society Transactions*. Available at: https://doi.org/10.1042/BST0320581.

Robaglia, C., Thomas, M. and Meyer, C. (2012) 'Sensing nutrient and energy status by SnRK1 and TOR kinases', *Current Opinion in Plant Biology* [Preprint]. Available at: https://doi.org/10.1016/j.pbi.2012.01.012.

Roux, P.P. *et al.* (2007) 'RAS/ERK signaling promotes site-specific ribosomal protein S6 phosphorylation via RSK and stimulates cap-dependent translation', *Journal of Biological Chemistry* [Preprint]. Available at: https://doi.org/10.1074/jbc.M700906200.

Ryabova, L.A., Robaglia, C. and Meyer, C. (2019) 'Target of Rapamycin kinase: Central regulatory hub for plant growth and metabolism', *Journal of Experimental Botany*, 70(8), pp. 2211–2216. Available at: https://doi.org/10.1093/jxb/erz108.

S., U. *et al.* (2020) 'Investigation of manganese metal coordination in proteins: a comprehensive PDB analysis and quantum mechanical study', *Structural Chemistry*, 31(3), pp. 1057–1064. Available at: https://doi.org/10.1007/s11224-020-01488-x.

Scharf, K.D. and Nover, L. (1982) 'Heat-shock-induced alterations of ribosomal protein phosphorylation in plant cell cultures', *Cell*, 30(2), pp. 427–437. Available at: https://doi.org/10.1016/0092-8674(82)90240-9.

Schepetilnikov, M. *et al.* (2011) 'Viral factor TAV recruits TOR/S6K1 signalling to activate reinitiation after long ORF translation', *EMBO Journal* [Preprint]. Available at: https://doi.org/10.1038/emboj.2011.39.

Schepetilnikov, M. *et al.* (2013) 'TOR and S6K1 promote translation reinitiation of uORF-containing mRNAs via phosphorylation of eIF3h', *EMBO Journal* [Preprint]. Available at: https://doi.org/10.1038/emboj.2013.61.

Schrödinger, LLC (2015) 'The PyMOL Molecular Graphics System, Version 1.8'.

Seeliger, M.A. *et al.* (2005) 'High yield bacterial expression of active c-Abl and c-Src tyrosine kinases', *Protein Science : A Publication of the Protein Society*, 14(12), pp. 3135–3139. Available at: https://doi.org/10.1110/ps.051750905.

Sheludko, Y.V. *et al.* (2007) 'Comparison of several Nicotiana species as hosts for high-scale Agrobacterium-mediated transient expression', *Biotechnology and Bioengineering*, 96(3), pp. 608–614. Available at: https://doi.org/10.1002/bit.21075.

Shi, L., Wu, Y. and Sheen, J. (2018) 'TOR signaling in plants: conservation and innovation', *Development (Cambridge, England)* [Preprint]. Available at: https://doi.org/10.1242/dev.160887.

Shih, Y.-P. *et al.* (2002) 'High-throughput screening of soluble recombinant proteins', *Protein Science*, 11(7), pp. 1714–1719. Available at: https://doi.org/10.1110/ps.0205202.

Shrestha, A. *et al.* (2012) 'Analysis of conditions affecting auto-phosphorylation of human kinases during expression in bacteria', *Protein Expression and Purification*, 81(1), pp. 136–143. Available at: https://doi.org/10.1016/j.pep.2011.09.012.

Sievers, F. *et al.* (2011) 'Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega', *Molecular Systems Biology*, 7(1), p. 539. Available at: https://doi.org/10.1038/msb.2011.75.

Smith, D.B. and Johnson, K.S. (1988) 'Single-step purification of polypeptides expressed in Escherichia coli as fusions with glutathione S-transferase', *Gene*, 67(1), pp. 31–40. Available at: https://doi.org/10.1016/0378-1119(88)90005-4.

Smyth, D.R. *et al.* (2003) 'Crystal structures of fusion proteins with large-affinity tags', *Protein Science*, 12(7), pp. 1313–1322. Available at: https://doi.org/10.1110/ps.0243403.

Son, O. *et al.* (2017) 'Molecular details of the Raptor-binding motif on Arabidopsis S6 kinase', *Biochemical and Biophysical Research Communications* [Preprint]. Available at: https://doi.org/10.1016/j.bbrc.2017.03.013.

Søndergaard, C.R. *et al.* (2011) 'Improved Treatment of Ligands and Coupling Effects in Empirical Calculation and Rationalization of pKa Values', *Journal of Chemical Theory and Computation*, 7(7), pp. 2284–2295. Available at: https://doi.org/10.1021/ct200133y.

Spiegelberg, B.D. *et al.* (2005) 'Alteration of lithium pharmacology through manipulation of phosphoadenosine phosphate metabolism', *The Journal of Biological Chemistry*, 280(7), pp. 5400–5405. Available at: https://doi.org/10.1074/jbc.M407890200.

Studier, F.W. (2005) 'Protein production by auto-induction in high density shaking cultures.', *Protein expression and purification* [Preprint]. Available at: https://doi.org/10.1016/j.pep.2005.01.016.

Studier, F.W. (2014) 'Stable Expression Clones and Auto-Induction for Protein Production in E. coli', in Y.W. Chen (ed.) *Structural Genomics: General Applications*. Totowa, NJ: Humana Press (Methods in Molecular Biology), pp. 17–32. Available at: https://doi.org/10.1007/978-1-62703-691-7_2.

Studier, F.W. and Moffatt, B.A. (1986) 'Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes', *Journal of Molecular Biology*, 189(1), pp. 113–130. Available at: https://doi.org/10.1016/0022-2836(86)90385-2.

Sun, L. *et al.* (2016) 'Ribosomal protein S6 kinase1 coordinates with TOR-Raptor2 to regulate thylakoid membrane biosynthesis in rice', *Biochimica et Biophysica Acta - Molecular and Cell Biology of Lipids*, 1861(7), pp. 639–649. Available at: https://doi.org/10.1016/j.bbalip.2016.04.009.

Sunami, T. *et al.* (2010) 'Structural Basis of Human p70 Ribosomal S6 Kinase-1 Regulation by Activation Loop Phosphorylation', *Journal of Biological Chemistry*, 285(7), pp. 4587–4594. Available at: https://doi.org/10.1074/jbc.M109.040667.

Szewczuk, L.M., Tarrant, M.K. and Cole, P.A. (2009) 'Chapter 1 Protein Phosphorylation by Semisynthesis: From Paper to Practice', in T.W. Muir and J.N. Abelson (eds) *Methods in Enzymology*. Academic Press (Non-Natural Amino Acids), pp. 1–24. Available at: https://doi.org/10.1016/S0076-6879(09)62001-2.

Tajika, Y. *et al.* (2004) 'Crystal structure of hypothetical protein PH0828 from Pyrococcus horikoshii', *Proteins: Structure, Function, and Bioinformatics*, 57(4), pp. 862–865. Available at: https://doi.org/10.1002/prot.20255.

Tavares, M.R. *et al.* (2015) 'The S6K protein family in health and disease', *Life Sciences* [Preprint]. Available at: https://doi.org/10.1016/j.lfs.2015.03.001.

Tegel, H. *et al.* (2010) 'Increased levels of recombinant human proteins with the Escherichia coli strain Rosetta(DE3)', *Protein Expression and Purification*, 69(2), pp. 159–167. Available at: https://doi.org/10.1016/j.pep.2009.08.017.

Tereshko, V. *et al.* (2001) 'Crystal structures of the catalytic domain of human protein kinase associated with apoptosis and tumor suppression', *Nature Structural Biology*, 8(10), pp. 899–907. Available at: https://doi.org/10.1038/nsb1001-899.

Thiébeauld, O. *et al.* (2009) 'A new plant protein interacts with eIF3 and 60S to enhance virus-activated translation re-initiation', *The EMBO Journal*, 28(20), pp. 3171–3184. Available at: https://doi.org/10.1038/emboj.2009.256.

Turck, F. *et al.* (1998) ' A Heat-Sensitive Arabidopsis thaliana Kinase Substitutes for Human p70 s6k Function In Vivo ', *Molecular and Cellular Biology* [Preprint]. Available at: https://doi.org/10.1128/mcb.18.4.2038.

Turck, F. *et al.* (2004) 'Phytohormones participate in an S6 kinase signal transduction pathway in arabidopsis', *Plant Physiology* [Preprint]. Available at: https://doi.org/10.1104/pp.103.035873.

Turkina, M.V., Klang Årstrand, H. and Vener, A.V. (2011) 'Differential phosphorylation of ribosomal proteins in Arabidopsis thaliana plants during day and night', *PloS One*, 6(12), p. e29307. Available at: https://doi.org/10.1371/journal.pone.0029307.

Tzeng, T.-Y. *et al.* (2009) 'Overexpression of the Lily p70s6k Gene in Arabidopsis Affects Elongation of Flower Organs and Indicates TOR-Dependent Regulation of AP3, PI and SUP Translation', *Plant and Cell Physiology*, 50(9), pp. 1695–1709. Available at: https://doi.org/10.1093/pcp/pcp114.

Van Leene, J. *et al.* (2019) 'Capturing the phosphorylation and protein interaction landscape of the plant TOR kinase', *Nature Plants* [Preprint]. Available at: https://doi.org/10.1038/s41477-019-0378-z.

Vaughn, J.L. *et al.* (1977) 'The establishment of two cell lines from the insect Spodoptera frugiperda (Lepidoptera; Noctuidae)', *In Vitro*, 13(4), pp. 213–217. Available at: https://doi.org/10.1007/BF02615077.

Venditti, V. *et al.* (2007) 'MD and NMR studies of alpha-bungarotoxin surface accessibility', *Biochemical and Biophysical Research Communications*, 356(1), pp. 114–117. Available at: https://doi.org/10.1016/j.bbrc.2007.02.094.

Verma, S., Nizam, S. and Verma, P.K. (2013) 'Biotic and Abiotic Stress Signaling in Plants', *Stress Signaling in Plants: Genomics and Proteomics Perspective*, 1. Available at: https://doi.org/10.1007/978-1-4614-6372-6.

Vile, D. *et al.* (2012) 'Arabidopsis growth under prolonged high temperature and water deficit: Independent or interactive effects?', *Plant, Cell and Environment*, 35(4), pp. 702–718. Available at: https://doi.org/10.1111/j.1365-3040.2011.02445.x.

Wang, J. *et al.* (2013) 'Crystal structures of S6K1 provide insights into the regulation mechanism of S6K1 by the hydrophobic motif', *Biochemical Journal* [Preprint]. Available at: https://doi.org/10.1042/BJ20121863.

Wang, M.-L. *et al.* (2008) 'Regulation of ribosomal protein S6 kinases by ubiquitination', *Biochemical and Biophysical Research Communications*, 369(2), pp. 382–387. Available at: https://doi.org/10.1016/j.bbrc.2008.02.032.

Waugh, D.S. (2005) 'Making the most of affinity tags', *Trends in Biotechnology*, 23(6), pp. 316–320. Available at: https://doi.org/10.1016/j.tibtech.2005.03.012.

Williams, A.J. *et al.* (2003) 'Regulated Phosphorylation of 40S Ribosomal Protein S6 in Root Tips of Maize', *Plant Physiology*, 132(4), p. 2086 LP – 2097. Available at: https://doi.org/10.1104/pp.103.022749.

Winter, D. *et al.* (2007) 'An "Electronic Fluorescent Pictograph" Browser for Exploring and Analyzing Large-Scale Biological Data Sets', *PLoS ONE*. Edited by I. Baxter, 2(8), p. e718. Available at: https://doi.org/10.1371/journal.pone.0000718.

Woestenenk, E.A., Hammarström, M., van den Berg, S., *et al.* (2004) 'His tag effect on solubility of human proteins produced in Escherichia coli: a comparison between four expression vectors', *Journal of Structural and Functional Genomics*, 5(3), pp. 217–229. Available at: https://doi.org/10.1023/b:jsfg.0000031965.37625.0e.

Xiong, F. *et al.* (2017) 'Brassinosteriod Insensitive 2 (BIN2) acts as a downstream effector of the Target of Rapamycin (TOR) signaling pathway to regulate photoautotrophic growth in Arabidopsis', *New Phytologist*, 213(1), pp. 233–249. Available at: https://doi.org/10.1111/nph.14118.

Xiong, Y. *et al.* (2013) 'Glucose-TOR signalling reprograms the transcriptome and activates meristems', *Nature* [Preprint]. Available at: https://doi.org/10.1038/nature12030.

Xiong, Y. and Sheen, J. (2012) 'Rapamycin and glucose-target of rapamycin (TOR) protein signaling in plants', *Journal of Biological Chemistry* [Preprint]. Available at: https://doi.org/10.1074/jbc.M111.300749.

Xiong, Y. and Sheen, J. (2014) 'The role of target of rapamycin signaling networks in plant growth and metabolism', *Plant Physiology*, 164(2), pp. 499–512. Available at: https://doi.org/10.1104/pp.113.229948.

Xu, B.-W. *et al.* (2020) 'Similarities and differences between Mn(II) and Zn(II) coordination polymers supported by porphyrin-based ligands: synthesis, structures and nonlinear optical properties', *Dalton Transactions*, 49(36), pp. 12622–12631. Available at: https://doi.org/10.1039/D0DT02450C.

Yaguchi, M., Ikeya, S. and Kozaki, A. (2020) 'The activation mechanism of plant S6 kinase (S6K), a substrate of TOR kinase, is different from that of mammalian S6K', *FEBS Letters*, 594(4), pp. 776–787. Available at: https://doi.org/10.1002/1873-3468.13661.

Yaguchi, M. and Kozaki, A. (2018) 'Plant S6 kinases do not require hydrophobic motif phosphorylation for activity in yeast lacking Ypk3', *FEBS Letters*, 592(4), pp. 610–620. Available at: https://doi.org/10.1002/1873-3468.12980.

Yang, J. *et al.* (2002) 'Crystal structure of an activated Akt/Protein Kinase B ternary complex with GSK3-peptide and AMP-PNP', *Nature Structural Biology* [Preprint]. Available at: https://doi.org/10.1038/nsb870.

Yang, J. *et al.* (2015) 'The I-TASSER Suite: protein structure and function prediction', *Nature Methods*, 12(1), pp. 7–8. Available at: https://doi.org/10.1038/nmeth.3213.

Zhang, S.H., Lawton, M.A., *et al.* (1994) 'atpk1, a novel ribosomal protein kinase gene from Arabidopsis. I. Isolation, characterization, and expression', *The Journal of Biological Chemistry*, 269(26), pp. 17586–17592.

Zhang, S.H., Broome, M.A., *et al.* (1994) 'atpk1, a novel ribosomal protein kinase gene from Arabidopsis. II. Functional and biochemical analysis of the encoded protein', *Journal of Biological Chemistry* [Preprint].

Zhang, Y. (2008) 'I-TASSER server for protein 3D structure prediction', *BMC Bioinformatics*, 9(1), p. 40. Available at: https://doi.org/10.1186/1471-2105-9-40.

Zhang, Z. *et al.* (2016) 'TOR Signaling Promotes Accumulation of BZR1 to Balance Growth with Carbon Availability in Arabidopsis', *Current Biology* [Preprint]. Available at: https://doi.org/10.1016/j.cub.2016.05.005.

Zheng, J. *et al.* (1993) '2.2 Å refined crystal structure of the catalytic subunit of cAMP-dependent protein kinase complexed with MnATP and a peptide inhibitor', *Acta Crystallographica Section D: Biological Crystallography*, 49(3), pp. 362–365. Available at: https://doi.org/10.1107/S0907444993000423.

Zulawski, M. *et al.* (2014) 'The Arabidopsis Kinome: phylogeny and evolutionary insights into functional diversification', *BMC Genomics*, 15(1), p. 548. Available at: https://doi.org/10.1186/1471-2164-15-548.

Zulawski, M. and Schulze, W.X. (2015) 'The plant kinome', *Methods in Molecular Biology (Clifton, N.J.)*, 1306, pp. 1–23. Available at: https://doi.org/10.1007/978-1-4939-2648-0_1.

Appendix

Table A1| PCR Primers

Name	Sequence, 5' to 3'
Construct insert PCR	
FL N-His Fw (NcoI)	GGATC**CCATGG**TAAGCTCCCAATGCAG
FL/KL N-His Rev (NotI)	GAATC**GCGGCCGC**CTACAAATTCGAGGT
KL/KD N-His Fw (NcoI)	GGATC**CCATGG**ACTCAGAGAAGAGCCCCG
KD N-His Rev (NotI)	GAATC**GCGGCCGC**CTAGGGCTTGAAGCTA
FL C-His Fw (BamHI)	GCT**GGATCC**ATGGTAAGCTCCC
FL/KL C-His Rev (XhoI)	GCT**CTCGAG**TTACAAATTCGAGGTGGT
KL/KD C-His Fw (BamHI)	GCT**GGATCC**GACTCAGAGAAGAG
KD C-His Rev (XhoI)	**CTCGAG**GGGCTTGAAGCTAGGTTGTA
KL C-His Fw (NheI)	GCT**GCTAGC**GACTCAGAGAAGAGCC
KDN Fw (NcoI)	**CCATGG**GTGTGGTAGGGATTGAGGA
KDN Rev (NotI)	**GCGGCCGC**TCAGTGGTGGTGGTGGT
Mutagenesis of phosphomimetic mutations	
T455E Sense	GTGGCGGGCGTACGTA**TTC**GAAGTTGGTGAAAGGGT
T455E Antisense	ACCCTTTCACCAACTTC**GAA**TACGTACGCCCGCCAC
S296E Sense	CGGTCGTCCCGCACAT**TTC**GTTTGAACGAGTGTTCTCCTCAA
S296E Antisense	TTGAGGAGAACACTCGTTCAAAC**GAA**ATGTGCGGGACGACCG
pET-21(3C) production	
HRV 3C site Fw	TCGAGCGGGCCCTGAAACAGCACTTCCAGC
HRV 3C site Rev	GCTGGAAGTGCTGTTTCAGGGCCCGCTCGA
XhoI site mutagenesis Sense	AGGGCCCG**CTC**C**AG**CACCACCAC
XhoI site mutagenesis Antisense	GTGGTGGTG**CT**G**GAG**CGGGCCCT

Fw – forward primer; Rev – reverse primer.
Restriction endonuclease sites are highlighted in red.
Nucleotide mutations are highlighted in blue.

Table A2| AtS6K2 KL_123_471_C_N-GST IMAC buffers

Stage	Buffer name	Buffer composition
IMAC	Lysis buffer	50 mM Tris-HCl pH 7.0, 300 mM NaCl, 5 % (v/v) glycerol, 0.01% Triton X-100, 2 mM $MgCl_2$, 0.5 mM ATP, 0.5 mM EDTA, 4 µg/ml Dnase I
	Wash buffer	50 mM Tris-HCl pH 7.0, 300 mM NaCl, 5 % (v/v) glycerol, 0.01% Triton X-100, 2 mM $MgCl_2$, 0.5 mM ATP, 0.5 mM EDTA
	Elution buffer	50 mM Tris-HCl pH 7.0, 300 mM NaCl, 5 % (v/v) glycerol, 0.01% Triton X-100, 2 mM $MgCl_2$, 0.5 mM ATP, 0.5 mM EDTA, 10 mM L-glutathione reduced

Table A3| AtS6K2 KL_123_471_C_N-His IMAC buffers

Stage	Buffer name	Buffer composition
IMAC	Lysis buffer	50 mM Tris-HCl pH 7.5, 300 mM NaCl, 5 mM β-mercaptoethanol, 5 % (v/v) glycerol, 0.01% Triton X-100, 1 mM $MgCl_2$, 0.5 mM ATP, 8 µg/ml DNase I, EDTA-free protease inhibitors
	Wash buffer	50 mM Tris-HCl pH 7.5, 300 mM NaCl, 5 mM β-mercaptoethanol, 5 % (v/v) glycerol, 0.01% Triton X-100, 1 mM $MgCl_2$, 0.5 mM ATP
	Elution buffer	50 mM Tris-HCl pH 7.5, 300 mM NaCl, 5 mM β-mercaptoethanol, 5 % (v/v) glycerol, 0.01% Triton X-100, 1 mM $MgCl_2$, 0.5 mM ATP, 300 mM imidazole

Table A4| AtS6K2 FL_1_471_C_N-His IMAC buffers

Stage	Buffer name	Buffer composition
IMAC	Lysis buffer	50 mM Tris-HCl pH 8.0, 300 mM NaCl, 5 % (v/v) glycerol, 0.01% Triton X-100, 1 mM $MgCl_2$, 0.5 mM ATP, 8 µg/ml Dnase I, EDTA-free protease inhibitors
	Wash buffer	50 mM Tris-HCl pH 8.0, 300 mM NaCl, 5 % (v/v) glycerol, 0.01% Triton X-100, 1 mM $MgCl_2$, 0.5 mM ATP
	Elution buffer	50 mM Tris-HCl pH 8.0, 300 mM NaCl, 5 % (v/v) glycerol, 0.01% Triton X-100, 1 mM $MgCl_2$, 0.5 mM ATP, 300 mM imidazole

Table A5| KL_123_471_NC_C-His_N-T7 Purification buffers

Stage	Buffer name	Buffer composition
IMAC	Lysis buffer	50 mM Tris-HCl pH 7.5, 250/300 mM NaCl, 5 % (v/v) glycerol, 0.01% Triton X-100, 0.4/1 mM $MgCl_2$, 0.2/0.5 mM ATP, 1-6 µg/ml DNase I, EDTA-free protease inhibitos
	Wash buffer	50 mM Tris-HCl pH 7.5, 250/300 mM NaCl, 5 % (v/v) glycerol, 0.01% Triton X-100, 0.4/1 mM $MgCl_2$, 0.2/0.5 mM ATP
	Elution buffer	50 mM Tris-HCl pH 7.5, 150/300 mM NaCl, 5 % (v/v) glycerol, 0.01% Triton X-100, 0.4/1 mM $MgCl_2$, 0.2/0.5 mM ATP, 300 mM imidazole
IEC	Buffer A	50 mM Tris-HCl pH 7.5, 5 % (v/v) glycerol, 1 mM $MgCl_2$, 0.5 mM ATP, 5 mM β-mercaptoethanol
	Buffer B	50 mM Tris-HCl pH 7.5, 5 % (v/v) glycerol, 1 mM $MgCl_2$, 0.5 mM ATP, 5 mM β-mercaptoethanol, 1 M NaCl
SEC	SEC buffer	25 mM Tris-HCl pH 7.5, 5 % (v/v) glycerol, 1 mM $MgCl_2$, 5 mM β-mercaptoethanol, 300 mM NaCl

Table A6| KL_123_471_NC_C-His Purification buffers

Stage	Buffer name	Buffer composition
IMAC	Lysis buffer	50 mM Tris-HCl pH 7.5/8.0 or 50 mM MES Na salt pH 6.5, 300 mM NaCl, 5 % (v/v) glycerol, 0.01% Triton X-100, 1 mM $MgCl_2$, 0.5 mM ATP, 0/5 mM β-mercaptoethanol, 1-6 µg/ml DNase I, EDTA-free protease inhibitors
	Wash buffer	50 mM Tris-HCl pH 7.5/8.0 or 50 mM MES Na salt pH 6.5, 300 mM NaCl, 5 % (v/v) glycerol, 1 mM $MgCl_2$, 0.5 mM ATP, 0/5 mM β-mercaptoethanol
	Elution buffer	50 mM Tris-HCl pH 7.5/8.0 or 50 mM MES Na salt pH 6.5, 300 mM NaCl, 5 % (v/v) glycerol, 1 mM $MgCl_2$, 0.5 mM ATP, 0/5 mM β-MEt, 300 mM imidazole
	Dialysis buffer	50 mM Tris-HCl pH 7.5, 50 mM NaCl, 5 % (v/v) glycerol, 5 mM β-mercaptoethanol
IEC	Buffer A	50 mM Tris-HCl pH 8.0, 5 % (v/v) glycerol, 1 mM $MgCl_2$, 0.5 mM ATP, 5 mM β-mercaptoethanol
	Buffer B	50 mM Tris-HCl pH 8.0, 5 % (v/v) glycerol, 1 mM $MgCl_2$, 0.5 mM ATP, 5 mM β-mercaptoethanol, 1 M NaCl.
SEC	SEC buffer	25 mM Tris-HCl pH 8.0, 5 % (v/v) glycerol, 1 mM $MgCl_2$, 5 mM β-mercaptoethanol, 250 mM NaCl

Table A7| KD_123_414_NC_C-His_N-T7 and KD_123_414_C_C-His_N-T7 Purification buffers

Stage	Buffer name	Buffer composition
IMAC	Lysis buffer	50 mM Tris-HCl pH 8.0, 250 mM Li_2SO_4, 0,01% Triton X-100, 5 % (v/v) glycerol, 2 mM $MgCl_2$, 0,5 mM ATP, 5 mM β-mercaptoethanol, 2 µg/mL DNase I, EDTA-free protease inhibitors
	Wash 1 buffer	50 mM Tris-HCl pH 8.0, 250 mM Li_2SO_4, 5 % (v/v) glycerol, 2 mM $MgCl_2$, 0,5 mM ATP, 5 mM β-mercaptoethanol
	Wash 2 buffer	50 mM Tris-HCl pH 8.0, 250 mM Li_2SO_4, 5 % (v/v) glycerol, 2 mM $MgCl_2$, 0,5 mM ATP, 5 mM β-mercaptoethanol, 20 mM imidazole
	Elution buffer	50 mM Tris-HCl pH 8.0, 250 mM Li_2SO_4, 5 % (v/v) glycerol, 2 mM $MgCl_2$, 0,5 mM ATP, 5 mM β-mercaptoethanol, 300 mM imidazole
IEC	Buffer A	50 mM Tris-HCl pH 8.0, 5 % (v/v) glycerol, 2 mM $MgCl_2$, 0.5 mM ATP, 5 mM β-mercaptoethanol
	Buffer B	50 mM Tris-HCl pH 8.0, 5 % (v/v) glycerol, 2 mM $MgCl_2$, 0.5 mM ATP, 5 mM β-mercaptoethanol, 1M Li_2SO_4
SEC	SEC buffer	25 mM Tris-HCl pH 8.0, 5 % (v/v) glycerol, 2 mM $MgCl_2$, 5 mM β-mercaptoethanol, 250 mM Li_2SO_4

Table A8| KD_123_414_C_N-GST Purification buffers

Stage	Buffer name	Buffer composition
IMAC	Lysis buffer	50 mM Tris-HCl pH 8.0, 250 mM Li_2SO_4, 0,01% Triton X-100, 5 % (v/v) glycerol, 2 mM $MgCl_2$, 0,5 mM ATP, 5 mM β-mercaptoethanol, 2 µg/mL DNase I, 0.5 mM EDTA
	Wash buffer	50 mM Tris-HCl pH 8.0, 250 mM Li_2SO_4, 5 % (v/v) glycerol, 2 mM $MgCl_2$, 0,5 mM ATP, 5 mM β-mercaptoethanol, 0.5 mM EDTA
	Elution buffer	50 mM Tris-HCl pH 8.0, 250 mM Li_2SO_4, 5 % (v/v) glycerol, 2 mM $MgCl_2$, 0,5 mM ATP, 5 mM β-mercaptoethanol, 0.5 mM EDTA, 10 mM L-glutathione reduced
IEC	Buffer A	50 mM Tris-HCl pH 8.0, 5 % (v/v) glycerol, 2 mM $MgCl_2$, 0.5 mM ATP, 5 mM β-mercaptoethanol
	Buffer B	50 mM Tris-HCl pH 8.0, 5 % (v/v) glycerol, 2 mM $MgCl_2$, 0.5 mM ATP, 5 mM β-mercaptoethanol, 1M Li_2SO_4
SEC	SEC buffer	25 mM Tris-HCl pH 8.0, 5 % (v/v) glycerol, 2 mM $MgCl_2$, 5 mM β-mercaptoethanol, 250 mM Li_2SO_4

Table A9| AtS6K2 KN_136_460_C_C-His IMAC buffers

Stage	Buffer name	Buffer composition
IMAC	Lysis buffer	50 mM Tris-HCl pH 7.5, 300 mM NaCl, 0/5 mM β-mercaptoethanol, 5 % (v/v) glycerol, 0.01/0.05% Triton X-100, 1 mM $MgCl_2$, 0.5 mM ATP, 8 µg/ml DNase I, EDTA-free protease inhibitors
	Wash buffer	50 mM Tris-HCl pH 7.5, 300 mM NaCl, 5 % (v/v) glycerol, 0.01% Triton X-100, 1 mM $MgCl_2$, 0.5 mM ATP
	Elution buffer	50 mM Tris-HCl pH 7.5, 150 mM NaCl, 5 % (v/v) glycerol, 0.01% Triton X-100, 1 mM $MgCl_2$, 0.5 mM ATP, 300 mM imidazole

Table A10| KDN_133_414_NC_C-His_λ and KDN_133_414_C_C-His_λ Purification buffers

Stage	Buffer name	Buffer composition
IMAC	Lysis buffer	50 mM Tris-HCl pH 8.0, 300 mM NaCl, 0,01% Triton X-100, 5 % (v/v) glycerol, 1 mM $MgCl_2$, 0,5 mM ATP, 5 mM β-mercaptoethanol, 2 µg/mL DNase I, EDTA-free protease inhibitors
	Wash 1 buffer	50 mM Tris-HCl pH 8.0, 300 mM NaCl, 5 % (v/v) glycerol, 1 mM $MgCl_2$, 0,5 mM ATP, 5 mM β-mercaptoethanol
	Wash 2 buffer	50 mM Tris-HCl pH 8.0, 300 mM NaCl, 5 % (v/v) glycerol, 1 mM $MgCl_2$, 0,5 mM ATP, 5 mM β-mercaptoethanol, 20 mM imidazole
	Elution buffer	50 mM Tris-HCl pH 8.0, 300 mM NaCl, 5 % (v/v) glycerol, 1 mM $MgCl_2$, 0,5 mM ATP, 5 mM β-mercaptoethanol, 300 mM imidazole
IEC	Buffer A	50 mM Tris-HCl pH 8.0, 5 % (v/v) glycerol, 1 mM $MgCl_2$, 0.5 mM ATP, 5 mM β-mercaptoethanol
	Buffer B	50 mM Tris-HCl pH 8.0, 5 % (v/v) glycerol, 1 mM $MgCl_2$, 0.5 mM ATP, 5 mM β-mercaptoethanol, 1M NaCl
SEC	SEC buffer	25 mM Tris-HCl pH 8.0, 5 % (v/v) glycerol, 1 mM $MgCl_2$, 5 mM β-mercaptoethanol, 250 mM NaCl

Table A11| FL_1_471_NC_C-His_N-T7 WT and T455E+S296E IMAC buffers

Stage	Buffer name	Buffer composition
IMAC	Lysis buffer	50 mM Tris-HCl pH 7.5/8.0*, 300 mM NaCl, 5 % (v/v) glycerol, 0.01% Triton X-100, 1 mM $MgCl_2$, 0.5 mM ATP, 2 µg/ml DNase I, EDTA-free protease inhibitors
	Wash buffer 1	50 mM Tris-HCl pH 7.5/8.0*, 300 mM NaCl, 5 % (v/v) glycerol, 1 mM $MgCl_2$, 0.5 mM ATP
	Wash buffer 2	50 mM Tris-HCl pH 7.5/8.0*, 300 mM NaCl, 5 % (v/v) glycerol, 1 mM $MgCl_2$, 0.5 mM ATP, 20 mM imidazole
	Elution buffer	50 mM Tris-HCl pH 7.5/8.0*, 300 mM NaCl, 5 % (v/v) glycerol, 1 mM $MgCl_2$, 0.5 mM ATP, 300 mM imidazole
Dialysis	Storage buffer	25 mM Tris-HCl pH 7.5/8.0*, 150 mM NaCl, 5 % (v/v) glycerol, 1 mM MgCl2, 0,5 mM ATP, 5 mM β-mercaptoethanol

* - Tris-HCl pH 7.5 was used for AtS6K2 FL_1_471_NC_C-His_N-T7 T455E+S296E mutant and Tris-HCl pH 8.0 was used for wild-type AtS6K2 FL_1_471_NC_C-His_N-T7.

Table A12| FL_1_471_NC_N-His_λ WT and T455E+S296E purification buffers

Stage	Buffer name	Buffer composition
IMAC	Lysis buffer	50 mM Tris-HCl pH 7.0/7.5*, 300 mM NaCl, 5 % (v/v) glycerol, 0.01% Triton X-100, 1 mM $MgCl_2$, 0.5 mM ATP, 2 µg/ml DNase I, EDTA-free protease inhibitors
	Wash buffer 1	50 mM Tris-HCl pH 7.0/7.5*, 300 mM NaCl, 5 % (v/v) glycerol, 1 mM $MgCl_2$, 0.5 mM ATP
	Wash buffer 2	50 mM Tris-HCl pH 7.0/7.5*, 300 mM NaCl, 5 % (v/v) glycerol, 1 mM $MgCl_2$, 0.5 mM ATP, 20 mM imidazole
	Elution buffer	50 mM Tris-HCl pH 7.0/7.5*, 300 mM NaCl, 5 % (v/v) glycerol, 1 mM $MgCl_2$, 0.5 mM ATP, 300 mM imidazole
SEC	SEC buffer	25 mM Tris-HCl pH 7.0/7.5*, 250 mM NaCl, 5 % (v/v) glycerol, 1 mM MgCl2, 0,5 mM ATP, 5 mM β-mercaptoethanol

* - Tris-HCl pH 7.0 was used for AtS6K2 FL_1_471_NC_N-His_λ T455E+S296E mutant and Tris-HCl pH 7.5 was used for wild-type AtS6K2 FL_1_471_NC_N-His_λ .

Table A13| Protease inhibitors

Inhibitor	Target protease
AEBSF	Serine proteases
Aprotinin	Serine proteases
Bestatin	Aminopeptidases
E-64	Cysteine proteases
Leupeptin	Serine and cysteine proteases
Pepstatin A	Aspartic acid proteases

Table A14| Western-blot antibodies

Antibody	Primary or secondary	Supplier, identification code	Dilution	AtS6K2 Construct
Anti-6xHis (Ms)	Primary	Abcam, ab15149	1:5000	KL_123_471_C_N-His
Anti-mouse (Gt)	Secondary	Abcam, ab205719	1:2000	KL_123_471_C_N-His
Anti-6xHis (Rb)	Primary	Abcam, ab137839	1:5000	FL_1_471_NC_N-His_λ WT
Anti-rabbit (Gt)	Secondary	Agisera, AS09 602	1:10000	FL_1_471_NC_N-His_λ WT

Table A15| Crystallisation screening kits

Kit name	Supplier	Commercial code
Structure Screen 1	Molecular Dimensions	MD1-01
Structure Screen 2	Molecular Dimensions	MD1-02
PACT premier™	Molecular Dimensions	MD1-29
JCSG-plus™	Molecular Dimensions	MD1-37
ProPlex™	Molecular Dimensions	MD1-38
Morpheus®	Molecular Dimensions	MD1-46
PGA Screen™	Molecular Dimensions	MD1-50
PEG/Ion Screen™	Hampton Research	HR2-126
PEG/Ion 2 Screen™	Hampton Research	HR2-098
Grid Screen™ Ammonium Sulfate	Hampton Research	HR2-211
Crystal Screen Cryo™	Hampton Research	HR2-122
Crystal Screen 2 Cryo™	Hampton Research	HR2-121

Table A16. Crystallisation screening tests of AtS6K2 KL_123_471_NC_C-His_N-T7

Protein conc., mg/mL	Protein vol., µL	Buffer vol., µL	Reservoir vol., µL	Temp, °C	Screening kits
10	0.1	0.1	60	16	PACT, PEG/Ion, ProPlex, PGA, Structure, JCSG
10	0.2	0.2	70	16	PACT, PEG/Ion, PGA, ProPlex, Strucutre, JCSG
10	0.1	0.2	70	16	PGA, ProPlex, PEG/Ion
20	0.1	0.1	60	18	PACT, PEG/Ion, ProPlex, PGA, Structure, JCSG
20	0.1	0.2	70	18	ProPlex, PGA

Structure: Structure Screen 1 and Structure Screen 2. PEG/Ion: PEG/Ion Screen and PEG/Ion 2 Screen

Table A17. Crystallisation screening tests of AtS6K2 KL_123_471_NC_C-His

Protein conc., mg/mL	Protein vol., µL	Buffer vol., µL	Reservoir vol., µL	Temp, °C	Screening kits
10	0.1	0.1	60	18	PACT
8	0.1	0.1	60	18	PACT, PGA, JCSG
7	0.1	0.1	60	4	PACT, PGA, JCSG, ProPlex, PEG/Ion, Crystal, Ammonium Sulphate
7	0.2	0.2	70	4	PACT

PEG/Ion: PEG/Ion Screen and PEG/Ion 2 Screen. Crystal: Crystal Screen Cryo and Crystal Screen

Table A18. Crystallisation screening tests of AtS6K2 KD constructs

Protein conc., mg/mL	Protein vol., µL	Buffer vol., µL	Reservoir vol., µL	Temp, °C	Screening kits	Construct
10	0.1	0.1	60	16	PACT, JCSG	KD_123_414_C_N-GST
10	0.1	0.1	60	16	PACT, PGA, JCSG, PEG/Ion	KD_123_414_NC_C-His_N-T7
10	0.1	0.1	60	4	PACT, JCSG, PEG/Ion	KD_123_414_C_C-His_N-T7
10	1	1	1000	4	Ammonium Sulphate	KD_123_414_NC_C-His_N-T7

Table A19| KL_123_471_C_C-His_N-T7 and KL_123_471_C_C-His IMAC buffers

Stage	Buffer name	Buffer composition
IMAC	Lysis buffer	50 mM Tris-HCl pH 7.5, 300 mM NaCl, 5 % (v/v) glycerol, 0.01/0.05% Triton X-100, 0.4/1 mM $MgCl_2$, 0.2/0,5 mM ATP, 0/5 mM β-mercaptoethanol, 6-8 µg/ml Dnase I, EDTA-free protease inhibitors
	Wash buffer	50 mM Tris-HCl pH 7.5, 300 mM NaCl, 5 % (v/v) glycerol, 0.01% Triton X-100, 0.4/1 mM $MgCl_2$, 0.2/0.5 mM ATP, 0/5 mM β-mercaptoethanol
	Elution buffer	50 mM Tris-HCl pH 7.5, 150/300 mM NaCl, 5 % (v/v) glycerol, 0.01% Triton X-100, 0.4/1 mM $MgCl_2$, 0.2/0.5 mM ATP, 0/5 mM β-mercaptoethanol, 300 mM imidazole

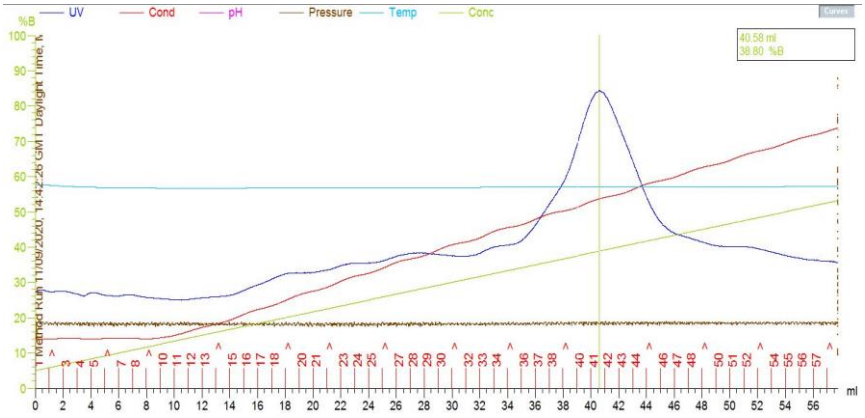


Figure A1| Chromatogram of cation exchange chromatography of AtS6K2 KL_123_471_NC_C-His_N-T7