UNIVERSITY OF TARTU
Faculty of Science and Technology
Institute of Technology

Irina Borovko

Cell cycle dependent nuclear localization

Bachelor's Thesis (12 ECTS)

Curriculum Science and Technology

Supervisor:
MSc, Mihkel Örd

Tartu 2020

Cell cycle dependent nuclear localization

Abstract:

Cell cycle is a sequence of precisely timed events, controlled mainly by the cyclin-dependent kinase (CDK) activity. CDK is activated by binding to another protein - cyclin and regulates the cellular processes via phosphorylation. One of the important tools in mediating enzyme activity is localization. Some proteins change their subcellular localization according to the cell cycle phase, for example, shuttle between the cytoplasm and nucleus, which is also mediated by CDK activity. Nucleocytoplasmic shuttling depends on the presence of the nuclear localization signal (NLS) and nuclear export signal (NES) in the target protein, and phosphoresidues can tune the activity of these signals. In this work, the role of CDK sites overlapping with bipartite NLS in Psy4 and Dna2 proteins was studied, and truncated localization modules that are sufficient for replicating the wild type proteins shuttling were identified.

Keywords: Cell cycle, Phosphorylation, Cyclin-dependent kinase, Cyclin, Nuclear localization signal

CERCS: P310 proteins, enzymology

Tuuma lokalisatsiooni signaalide regulatsioon tsükliinist sõltuva kinaasi Cdk1 poolt

Lühikokkuvõte:

Rakutsükkel on täpselt järjestatud sündmuste jada, mis on koordineeritud tsükliinist sõltuvate kinaaside (CDK) poolt. CDK on aktiivne vaid koos tsükliiniga ning vastav kompleks reguleeritud rakutsüklit valkude fosforüleerimise kaudu. Rakusisene paiknemine on oluline tegur valkude funktsioonide piiramisel. Mitmete valkude paiknemine muutub rakutsükli jooksul, näiteks transporditakse valke tuumast sisse ja välja CDK-vahendatud fosforüleerimise kaudu. Selleks on vajalikud tuuma impordi ja ekspordi signaalid, mille aktiivsust võib mõjutada nende fosforüleerimine. Käesolevas töös uuritakse, kuidas CDK reguleerib tuuma lokalisatsiooni mooduleid valkudes Psy4 ja Dna2.

Võtmesõnad: rakutsükkel, fosforüleerimine, tsükliinist sõltuv kinaas, tsükliin, tuuma lokalisatsiooni signaal

CERCS: P310 proteiinid, ensümoloogia

Table of Contents

TERMS, ABBREVIATIONS AND NOTATIONS

APC - Anaphase promoting complex

CDK – Cyclin-dependent kinase

CKI - Cyclin-dependent kinase inhibitor

CSM - Complete supplement mixture

SCF - Skp, Cullin, F-box containing complex

EDTA - Ethylenediaminetetraacetic acid

eGFP – eukaryotic green fluorescent protein

LLPP – Short linear Cln2 docking motif

NES - nuclear export signal

NLS – nuclear localization signal

NLxxxL - Short linear Clb5-specific docking motif

OD – Optical density

RxL - Short linear Clb5- docking motif with consensus sequence of K/R-x-L- φ or K/R-x-L-x-φ

pADH1 – Adh1 promoter

TAE buffer - Tris-acetate-EDTA buffer

TE buffer - Tris-EDTA buffer

TBS-T - Tris-buffered saline – Tween buffer

YPD – Yeast extract peptone dextrose

INTRODUCTION

The cell cycle is a complex sequence of interconnected cellular processes resulting in cell division. Resolving the mechanisms of its regulation and being able to manipulate them can open a lot of new possibilities for both medicine and synthetic biology.

The eukaryotic cell cycle consists of four phases: G1, S, G2 and M. Correct timing of different phases is crucial for cell proliferation. Most cell cycle events are controlled by phosphorylation of proteins by cyclin-dependent kinases (CDKs). CDKs catalyze the phosphorylation of serine and threonine residues in target proteins. Phosphorylation is a common posttranslational modification that regulates cellular processes, such as protein degradation, protein to protein interactions, and localization of proteins.

CDKs are activated by forming complexes with cyclin proteins. Fluctuations in cyclin concentrations during the cell cycle result in the recurrence in CDK activation. Cyclin abundance is regulated by cyclin gene expression and degradation rates, leading to different cyclins being present in different cell cycle stages. *S. cerevisiae* has one essential CDK – Cdk1,that is activated by nine cyclins. It is still not fully clear what governs the correct timing of phosphorylation events, allowing just one CDK to precisely control the flow of the cell cycle.

One of the mechanisms that confer to the correct targeting is subcellular localization. Different cyclin-Cdk1 complexes are positioned in different parts of the cell and target proteins can be colocalized with one complex and isolated from others.

The control over subcellular localization of proteins is important for the correct cell functioning since the rate of reactions largely depends on the availability of reactants. A change in localization can work as a switch, as moving an enzyme away from its target hinders the reaction, therefore the ability to manipulate it can become a powerful tool in biology. Regulation of localization is a much faster switch than transcriptional control that is mainly used in synthetic biology as the protein itself remains intact.

All proteins are synthesized in the cytosol, but some have to function in the nucleus. Nucleocytoplasmic transport of macromolecules is controlled by the multiprotein nuclear pore complex (NPC). Large macromolecules have to be actively ferried via transport proteins, called karyopherins – importins, catalyzing the transport inside the nucleus and exportins, regulating the movement back to the cytoplasm. To be recognized a target protein has to carry NLS (nuclear localization signal) and NES (nuclear export signal) amino acid sequences.

Previous studies have shown that cell cycle-dependent localization is mainly regulated by CDK. Several *S. cerevisiae* proteins (Whi5, Swi5, Mcm3, Acm1, Dna2, Psy4, etc.) are showing this property. The activity of NLS and NES sequences in such proteins depends on the

phosphorylation of the overlapping CDK sites. In this work, the role of CDK sites in the regulation of NLS activity on the example of two proteins, Dna2 and Psy4, is studied. Also, the work aims to achieve more control over CDK-dependent nuclear localization with the aim to understand the mechanisms controlling ordered cell cycle phosphorylation and also to design controllable NLS modules for use in synthetic biology and research.

1 LITERATURE REVIEW

1.1 Cell cycle

Cell is the smallest independent unit of any living organism. The proliferation of an organism largely depends on the ability of cells to grow and divide. Cell cycle is a tightly regulated sequence of processes that results in cell division.

Eukaryotic cell cycle is divided into four phases: G1, S, G2 and M. G1, S and G2 are commonly grouped to form the interphase. Typically, cell spends most of its lifetime, around 95%, in this stage of the cell cycle. During interphase cell prepares for the division, grows and duplicates its cellular components and genetic material (Cooper, 2000). M phase is the division itself, and can be subdivided into mitosis (nuclear division) and cytokinesis (cell division). S phase (S for synthesis) is a crucial stage during which DNA replication takes place. G1 and G2 phases, also called gap phases, separate S and M. During gap phases cells are provided with necessary time to accumulate biomass, produce proteins and cytoplasmic organelles.

The duration of the stages largely depends on the type of cell. For example, in *S. cerevisiae* doubling time is around 90 minutes, with G1 phase taking from 20 to 40 minutes, S phase and M phase – 30 minutes each (Brewer et al., 1984).

Cell cycle also includes a number of checkpoints as it is crucial for each stage to be correctly completed before entering the next one. Certain signaling molecules help to indicate whether the cell cycle can proceed to the next stage. Most important checkpoints are in G1, G2 and M phases. G1 is an important decision point for a cell, as G1 cells can get arrested and not go into division, but rather go into the nondividing state, also called G0 phase. For example, in yeast the decision to enter S phase happens in the late G1 at the START point and is mostly determined by the cell size, availability of nutrients and presence of mating factors (Pringle,1981; Cooper, 2000). G2 checkpoint prevents cell from entering the M phase before all of the genetic material is correctly duplicated. In M phase the correct alignment of the chromosomes is checked. The cell is restricted from initiating nuclear division until all the chromosomes are properly aligned. (Morgan,2007)

1.2 Cyclin-dependent kinases

Cell cycle requires strict control for all involved processes to start and finish at a well-defined timepoint. Therefore, there are a number of regulatory molecules controlling the correct sequence of cell cycle events. One of the key components of this control system are cyclin-dependent kinases (CDKs) that regulate the localization, interactions, stability and activity of hundreds of cell cycle proteins through phosphorylation (Enserink & Kolodner, 2010).

Two substrates, ATP and the target protein, bind to CDK catalytic core in a highly specific manner, so that γ-phosphate from ATP and hydroxylated side chain of the protein are facing each other, enabling transfer of the γ-phosphate from ATP to the substrate protein.

Cdks contain two lobes: a smaller N-terminal and a larger C-terminal lobe. ATP is positioned in the hydrophobic pocket in the cleft in between of two lobes. In case Cdk is in the active state, the protein substrate can bind near the entrance of the cleft next to the γ -phosphate. However, in deactivated state, ATP phosphates are not positioned correctly, and large disordered T-loop prevents protein from binding the cleft(Jeffrey et al., 1995). Cdks are activated by forming complexes with cyclin proteins (Morgan, 1997). The binding of cyclin changes the Cdk active site conformation in the way that T-loop does not occupy the cleft anymore and that ATP phosphates are oriented properly for phosphorylation(Jeffrey et al., 1995). In addition, it exposes a conserved threonine on the T-loop allowing it to be phosphorylated, which is required for the complete activation of the complex. Presumably, the activating phosphorylation further improves substrate protein binding to the CDK active site. (Morgan, 1997)

As the core cell cycle control machinery is conserved in eukaryotes, *S. cerevisiae* is widely used as a model organism for the cell cycle studies. It has only one essential Cdk – Cdk1 (Cdc28) (Kitazono & Kron, 2002).

1.3 Regulation of Cdk1 activity

CDK activity fluctuates over the cell cycle, giving rise to a difference in phosphorylation patterns. Oscillations in phosphorylation of the cell cycle machinery components result in correct initiation of cell cycle processes(Swaffer et al., 2016). Recurrence in CDK activation is primarily due to changes in the concentrations of cyclins during the cell cycle, which is mainly controlled by cyclin gene expression and degradation rates

(Morgan,2007). Cyclins are divided into four classes – G1/S cyclins, S cyclins, M cyclins and G1 cyclins – based on their abundance at certain points of the cell cycle (Fig. 1).

Additionally, CDK activity is regulated via inhibitory phosphorylation and inhibitor proteins called CDK inhibitory subunits (CKIs). They bind cyclin-CDK complexes preventing them from interacting with the substrates (Peter & Herskowitz, 1994).

In *S. cerevisiae* a CKI Far1 regulates the activity of cyclin-Cdk1 complexes in G1. Its activity is responsible for the G1 arrest in presence of mating factor (Schneider et al., 1996; Schwob et al., 1994). Sic1 inhibits later cyclin-Cdk1 complexes, suppressing S phase Cdk1 activity in G1 (Morgan,2007).

1.3.1 The sequential expression waves of cyclins

There are nine cyclins activating Cdk1 in yeast which are generally grouped into two classes – G1 cyclins (Cln1, Cln2, Cln3) and B-type cyclins (Clb1, Clb2, Clb3, Clb4, Clb5, Clb6). During the cell cycle they emerge in a highly controlled order in a wavelike fashion (Morgan, 2007).

Cln3 is a first step in a complex sequence of events controlling Cdk1 activity throughout the cell cycle.Cln3 present in low levels throughout the cell cycle, but the highest concentration is reached in late M–early G1(Bállega et al., 2019). Cln3-Cdk1 activity regulates the entry to the cell cycle (Polymenis & Schmidt, 1997). Cln3 is followed by G1 cyclins Cln1, Cln2, and S phase cyclins Clb5 and Clb6 whose expression peaks around the time of G1/S (Morgan, 2007). Their transcription is regulated by two related heterodimeric transcription factors SBF (composed of Swi4 and Swi6) and MBF (a complex of Mbp1 and Swi6) (Koch et al., 1993). Cln3 is predominantly localized in the nucleus (Bloom & Cross, 2007; Edgington & Futcher, 2001; Miller & Cross, 2001), where it phosphorylates Whi5, an SBF repressor, forcing its nuclear export and thus inducing expression of Cln1 and Cln2 (Costanzo et al., 2004; de Bruin et al., 2004).

Cln1/2-Cdk1 complexes drive the cell through the G1 phase (Wittenberg et al., 1990). They are mainly responsible for the bud emergence and spindle pole body duplication (Schneider et al., 1996; Schwob et al., 1994). Cln1, Cln2 and Cln3 can take over the functions of each other, as only a cln1Δ cln2Δ cln3Δ triple deletion strain is inviable (H. E. Richardson et al., 1989). Cln1 and Cln2 are also necessary for the cell to proceed to the S phase.

In S phase DNA is replicated for the cell to have a double set of chromosomes before the division. Clb5,6-Cdk1 complexes are necessary to start the replication (Donaldson et al., 1998). They also prohibit reinitiation on DNA replication origins (Dahmann et al., 1995). In

some cases, Clb5 and Clb6 may also play a role in spindle formation in later stages of the cell cycle (Schwob et al., 1994).

Clb5 and Clb6 are coexpressed with Cln1 and Cln2, however, they are inactive until degradation of the Clb-Cdk1 inhibitor Sic1 (Schneider et al., 1996; Schwob et al., 1994). Sic1 is targeted to ubiquitin-mediated degradation by phosphorylation Cdk1-dependent phosphorylation of Sic1, which requires both activity of Cln1,2-Cdk1 and Clb5,6-Cdk1. Initial phosphorylation by Cln1,2-Cdk1 triggers a phosphorylation cascade, which creates a positive feedback, where the released Clb5,6-Cdk1 activity further phosphorylates Sic1 (Kõivomägi et al., 2011).

Moreover, Clb5 and Clb6 control the efficient completion of G1. Clb6-Cdk1 phosphorylates Swi6, forcing it into the cytoplasm, thus deactivating SBF and MBF and inhibiting the expression of G1 cyclins (Geymonat et al., 2004).

Next two waves of cyclins are G2 cyclins Clb3,4 and and M cyclins Clb1,2. Clb1-4 are responsible for entering the M phase and for the progression through mitosis. They also inhibit mitotic exit and cytokinesis (Bloom & Cross, 2007). These cyclins share a lot of functions and can also partially take over Clb5 and Clb6 functions in the absence of S phase cyclins (Amon et al., 1993; H. Richardson et al., 1992).

Expression of mitotic cyclins is regulated by Fkh2 transcription factor and its coactivator Ndd1 (Kumar et al., 2000; Lim & Pic-Taylor, 200). Clb5-Cdk1 phosphorylates Fkh2 and Ndd1, directing it to bind Clb2 promoter (Pic-Taylor et al., 2004). They can also be phosphorylated by other Clb-Cdk complexes, creating positive feedback loops that promote the switch-like behavior of cyclin expression waves (Linke et al., 2017).

Clb3 and Clb4 appear during the S phase and their levels are kept high until anaphase. They are mainly responsible for spindle morphogenesis (Amon et al., 1993; H. Richardson et al., 1992). Clb1 and Clb2 concentrations peak around 10 minutes before anaphase (Fitch et al., 1992; Ghiara et al., 1991; H. Richardson et al., 1992; Surana et al., 1991). Clb2 is much more abundant compared to Clb1 and contributes the most to entering the M phase. Clb2 also promotes spindle elongation and negatively regulates bud emergence and SBF activity (Amon et al., 1993; Lew & Reed, 1993; Lim & Pic-Taylor, 2000). To finish the division, Cdk1 must be inactivated, which is reached by cyclin degradation and expression of CKIs.

1.4 Cyclin degradation

For the progression of the cell cycle, precise amounts of the involved proteins at certain timepoints are required. These quantities are not only controlled by the transcription machinery, but accurately timed degradation is also important. Phosphorylation is one of the tools to control protein destruction.

Ubiquitin-mediated proteolysis is one of the mechanisms that regulates the concentration of cyclins, CKIs and other regulators of the cell cycle. Multiple copies of the small protein ubiquitin are added on the protein targeted for the degradation(Schrader et al., 2009). Protease complexes called proteasomes recognise the ubiquitinated proteins and degrade them. Two main proteins catalysing the ubiquitination of the proteins involved in cell cycle regulation are related multi-subunit ubiquitin-protein ligases APC (anaphase-promoting complex) and SCF (Skp, Cullin, F-box containing complex) (Morgan,2007).

Early cyclins are targeted by Cdk1 phosphorylation for SCF (Lanker et al., 1996; Skowyra et al., 1999). SCF is constitutively active throughout the cell cycle and its target proteins carry a short linear motif called phosphodegron (Berset et al., 2002). Phosphorylation of these sequences allows ubiquitin ligase to recognize the target and trigger ubiquitin-mediated proteolysis (Holt, 2012). SCF targets include Cln1, Cln2, Far1 and Sic1(Berset et al., 2002; Deshaies, 1999). Their degradation depends on the Cln1,2-Cdk1 activity in the G1 phase and it is one of the mechanisms of G1/S switch (Yang et al., 2013).

There are several forms of SCF, depending on which F-box protein is in the complex. In *S. cerevisiae* there are two complexes involved in cell cycle control, SCF^{Cdc4} and SCF^{Grr1} (Reed, 2003). Their targets are different: Sic1 and Far1 are recognized by SCF^{Cdc4} (Feldman et al., 1997; Henchoz et al., 1997), whereas Cln1,2 are targeted by SCF^{Grr1} (Barral et al., 1995). There is evidence that this specificity is partially dictated by subcellular localization (Landry et al., 2012). For instance, Cln2 is constantly cytoplasmic (Miller & Cross, 2001b), and therefore recognised by SCF^{Grr1}, which is both cytoplasmic and nuclear, but not by SCF^{Cdc4}, which is solely nuclear (Blondel et al., 2000).

Mammalian G1 cyclins are regulated in a similar manner, as cyclin E carries a phosphodegron that is recognised by SCF^{Cdc4} in both mammalian cells and *S. cerevisiae* (van Drogen et al., 2006).

In contrast, degrons in APC substrates are active constantly, while APC activity oscillates during the cell cycle (Berset et al., 2002; Vodermaier, 2004). APC catalyzes the ubiquitin attachment to S and M cyclins that contain APC destruction box. Their degradation results in the loss of Cdk1 activity that is necessary for mitotic exit (Morgan,2007).

Correct timing of the APC activity is controlled by Cdk1 (Morgan,2007). APC has two activator – molecules Cdh1 and Cdc20 (Vodermaier, 2004). In G1, activity of this ubiquitin ligase is necessary to prevent early cell cycle entry. Upon the rise of Cdk1 activity in late G1 phase, Cdh1 is phosphorylated and deactivated, allowing the accumulation of S and M cyclins. In metaphase, M phase cyclin-Cdk1 complexes phosphorylate APC and promote Cdc20 binding, restoring its activity (Morgan,2007). This results in a drop of Cdk1 activity: cell division can be finished, and cell cycle can restart.

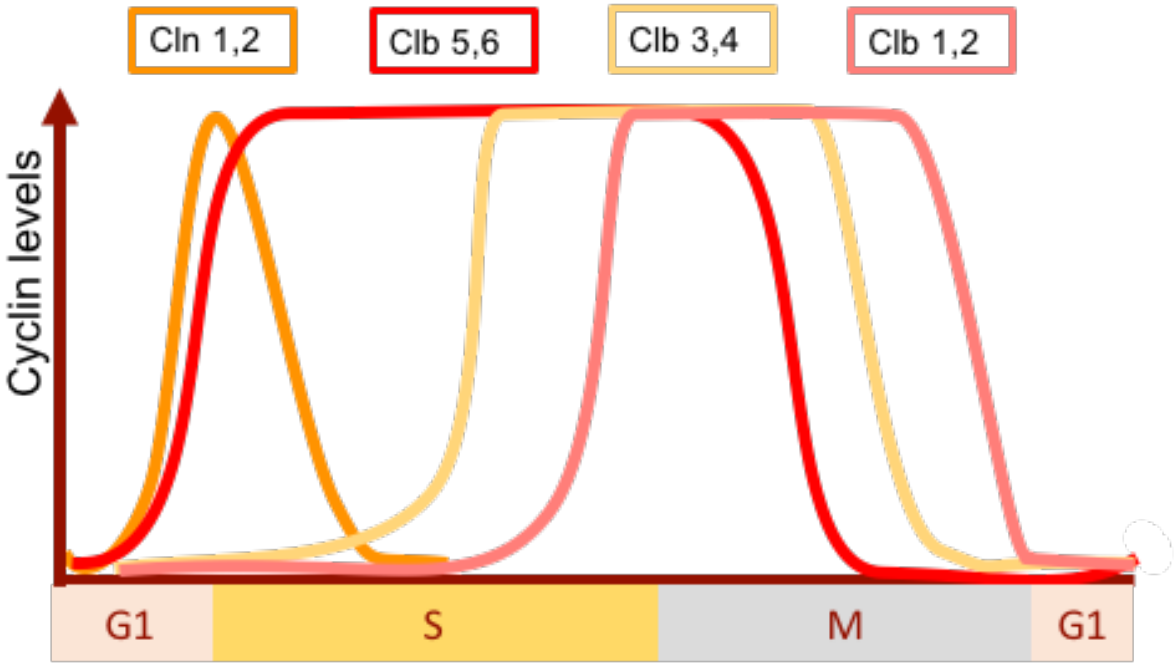

Figure 1. Cdk1 is activated by forming complexes different cyclins. Cdk1 activity fluctuates throughout the cell cycle, giving rise to a difference in phosphorylation patterns. This happens due to the oscillations in cyclin concentrations during the cell cycle. The cyclin levels are mainly controlled by cyclin gene expression and degradation rates. Different types of cyclins are predominant in different cell cycle phases. (Morgan, 2007)

1.5 Cdk1 substrate specificity

CDK is a serine/threonine protein kinase. CDK active site binds a specific site called the consensus phosphorylation motif in the target protein. Typically, it is a serine or threonine residue that is followed by proline and a basic amino acid, lysine or arginine, forming a consensus phosphorylation motif [S/T]Px[K/R], where x is any amino acid (Songyang et al., 1994; Mok et al., 2010) (Fig. 2). Sites that match the full consensus are called optimal. However, phosphorylation, although less effectively, can still happen in the absence of the basic

residue, and in some cases, also in the absence of the proline in +1 position from the S/T. Phosphorylation sites lacking the +3 basic residue are called suboptimal or minimal consensus sites. Notably, serine-based phosphorylation sites are predominant in human, fly, worm and yeast (Ubersax & Ferrell, 2007).

Another protein that is involved in the formation of the CDK complex, CKS (Cyclin-dependent kinase regulatory subunit), makes substrate targeting more specific. In budding yeast this role is taken by Cks1. Similar to CDK, the abundance of CKS is not cell cycle dependent (Hadwiger, Wittenberg, Mendenhall, et al., 1989). Conserved in a lot of organisms, CKS binds CDK target proteins at already phosphorylated CDK sites and potentiates the phosphorylation of secondary sites (Kõivomägi et al., 2011, 2013; McGrath et al., 2013). Notably, Cks1 is only capable of binding to phospho-threonines and not -serines and the Cks1 binding sites are called priming sites. The distance between primer and acceptor phosphosites is crucial, with the optimum being 12-16 amino acids (Kõivomägi et al., 2011) (Fig. 2).

However, as it was mentioned before, in *S. cerevisiae* cell cycle control can be fulfilled by just one CDK, Cdk1. During different cell cycle phases different targets are involved and different processes are triggered. That leads to the question of what determines the phosphorylation timing of a Cdk1 substrate in the cell cycle. One of the explanations is that instead of cyclin specificity, the Cdk1-mediated switches are triggered at specific Cdk1 activity thresholds. During the cell cycle, a gradual increase in Cdk1 activity occurs, so that later targets need higher kinase activity that is mediated by accumulation of cyclins (Örd & Loog, 2019; Stern & Nurse, 1996). In addition, Cdk1 activity towards a target peptide is also not constant and depends on the cyclin that forms the complex. It has been shown that the cyclin-Cdk intrinsic activity increases in the order of the cyclins appearing in the cell cycle. For example, S phase Clb5–Cdk1 has lower activity towards the substrates without docking motifs than mitotic Clb2–Cdk1 (Kõivomägi et al., 2011; Loog & Morgan, 2005). These thresholds make sure the cell cycle phases do not overlap with each other and that earlier events are triggered before later ones.

In addition, it was shown that not all target proteins are equally recognized by all cyclin-Cdk1 complexes (Loog & Morgan, 2005). Using docking pockets, the cyclins can enhance Cdk substrate specificity by binding linear docking motifs in target proteins (Loog & Morgan, 2005; Schulman et al., 1998) (Fig. 2). In budding yeast, docking motifs specific for four major cyclins have been described (Bhaduri & Pryciak, 2011; Kõivomägi et al., 2011; Loog & Morgan, 2005; Örd et al., 2019). G1 cyclin Cln2 targets leucine- and proline-rich motif LP, S-CDK targets carry Clb5-specific motif RxL, Clb3 acting in G2/M recognizes PxF, and M phase substrates carry LxF motif recognized by Clb2 (Bhaduri & Pryciak, 2011; Kõivomägi et al., 2011; Loog

& Morgan, 2005). The specific positioning of the docking sites relative to the phosphorylation sites along the disordered substrate proteins also affects the phosphorylation rate, as docking sites are more efficient 20–40 amino acids downstream from the CDK site (Kõivomägi et al., 2013).

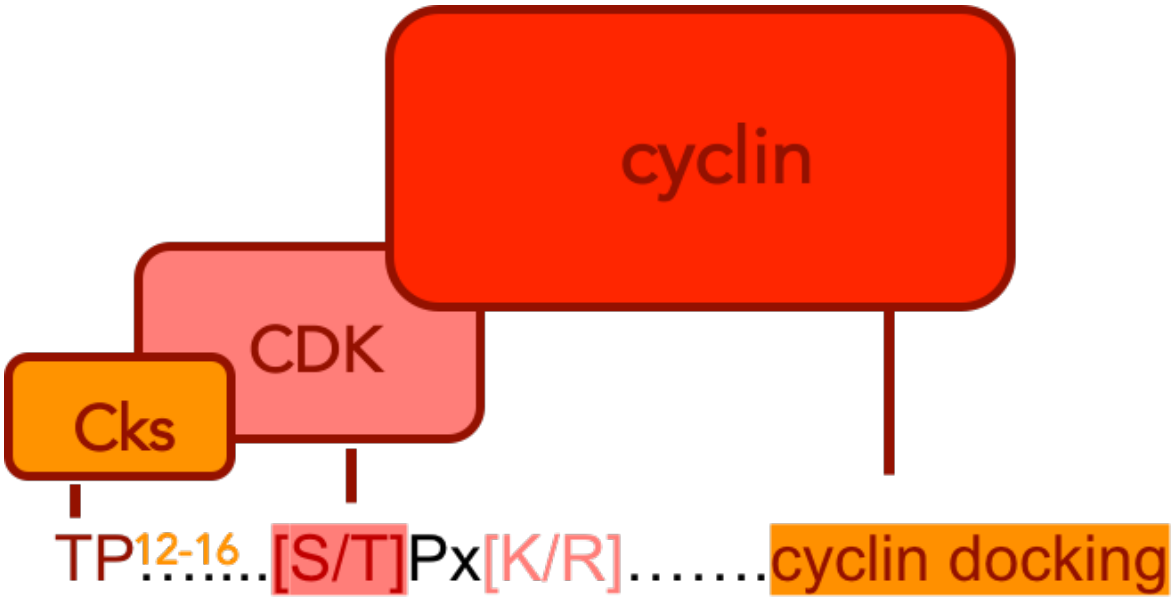

Figure 2. Substrate interactions mechanisms of the CDK complex. Cdk active site binds a specific site in the target protein: a serine or threonine residue followed by the proline and a basic amino acid. Sites of this form are called optimal. Phosphorylation of suboptimal (minimal) sites lacking the basic residue, is also possible. Cks binds Cdk target proteins at specific priming phosphosites sites and enhances Cdk activity towards the pre-phosphorylated substrate proteins. Notably, Cks1 is only capable of binding to phospho-threonines. The distance between primer and acceptor phosphosites is crucial. Some target proteins also carry cyclin-specific linear docking motifs that can promote phosphorylation by specific cyclin-Cdk1 complexes.

Another interesting mechanism of cyclin specificity is subcellular localization, as cyclins govern the localization of the CDK complex. Cyclins in different cellular compartments have access to a certain targets (Bloom & Cross, 2007). In fact, as it has been shown on the examples of Whi5, Grr1, Cln3, Swi6, subcellular localization is one of the most important tools in cell cycle regulation.

Although there are examples of cyclin-specific substrate targeting and also of a minimal cell cycle in fission yeast that can function with just one cyclin-Cdk1 complex, it is still not clear how the phosphorylation timings of specific substrates are governed(Bloom & Cross, 2007; Coudreuse & Nurse, 2010; Swaffer et al., 2016). Recent work aiming to design Cdk1 targets with specific phosphorylation timings revealed that by manipulating the Cks1 and cyclin docking specificities, Cdk1-dependent switches covering the whole cell cycle could be linearly

encoded into substrate proteins. This work, however, was based on a phosphodegron output, in which case the phosphorylated substrate is rapidly degraded. Thus, it is still unclear how the switches with a reversible output, such as change in localization or a conformational switch, are regulated.

1.6 Subcellular localization

Local concentration of macromolecules, including proteins, plays a huge role in cellular processes, as concentration of the reactants determines the rate of reaction. Thus, correct subcellular localization is crucial for the proper protein function. The subcellular localization of proteins is controlled by several mechanisms, like trafficking proteins or anchors. All proteins are synthesized in cytosol, however, enzymes that act on chromosomes have to accumulate in the nucleus. Consequently, nuclear localization of certain proteins, such as transcription factors and DNA polymerase, is essential for cell viability. Transport of macromolecules in and out of the nucleus is mediated by the multiprotein nuclear pore complex (NPC). Large macromolecules, above 60 kDa, cannot passively diffuse in and out of the nucleus and need to be actively ferried in a highly regulated manner (Lim et al., 2015).

Karyopherins is a class of transport proteins that regulate nuclear import and export. In *S. cerevisiae* this family has 14 members (Mosammaparast & Pemberton, 2004). Karyopherins are correspondingly subdivided into importins and exportins, which recognize specific amino acid sequences, NLS (nuclear localization signal) and NES (nuclear export signal), respectively. Importins are further subdivided into α and β importins. Importin α functions as an adapter and binds both the cargo and importin β1, which in turn binds NPC. Importin β can also bind the cargo directly, however NLS recognized in that fashion have proved to be harder to define (Fried & Kutay, 2003). To be able to shuttle in and out of the nucleus, proteins need to carry both NLS and NES sequences (Fig. 3).

Nuclear transport is controlled by the Ran-GTP gradient, which is more concentrated in nucleus. In the cytoplasm in the absence of Ran-GTP importin can bind its cargo and move it to the nucleus, where binding of the Ran-GTP releases the cargo. Inversely, Ran-GTP bound to the exportin allows it to bind to NES of the target protein. The complex dissociates in the cytoplasm (Fried & Kutay, 2003).

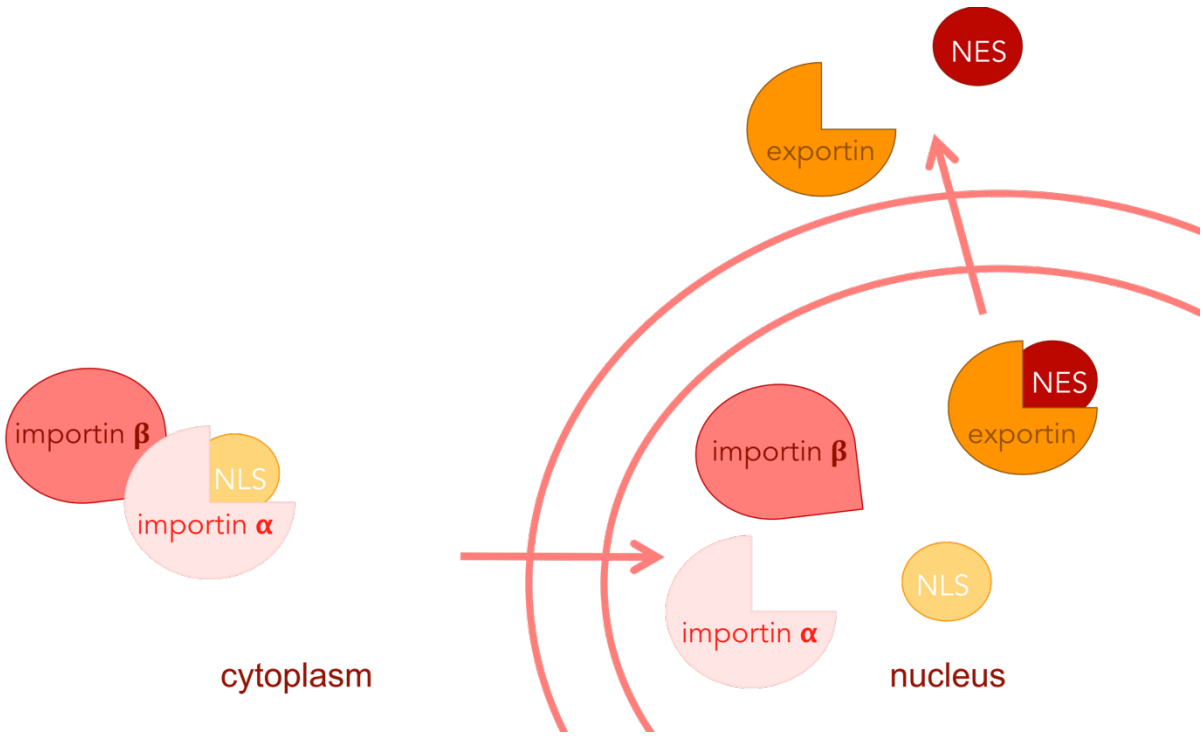

Figure 3. Nucleocytoplasmic shuttling is mediated by transport proteins exportins and importins. In cytoplasm, importins bind a protein carrying NLS amino acid sequence and transfer it into nucleus. Proteins with NES amino acid sequence are brought back to cytoplasm by exportins.

1.7 NLS and NES motifs

NLS recognition by importin α depends on the interactions based on charge and hydrophobicity (Pemberton & Paschal, 2005). Typically, NLS motifs are short lysine enriched basic stretches. This kind on NLS sequences are called classical (cNLS) and can be monopartite or bipartite (two basic stretches 10-12 amino acids apart from each other) (Dingwall & Laskey, 1991; Hahn et al., 2008). Importin α has two NLS binding sites: major and minor. Both of them are required for the binding of bipartite NLS, while monopartite binds only one, predominantly the major site (E Conti et al., 1998; Elena Conti, 2002; Fontes et al., 2000, 2003; Kosugi, Hasebe, Matsumura, et al., 2009). Monopartite NLS can consist out of long stretches or short ones, only two to three amino acids. Monopartite NLS also requires specific residues surrounding the basic core to improve binding specificity (Kosugi, Hasebe, Matsumura, et al., 2009; Makkerh et al., 1996).

Bipartite NLS consists of short N- and C- terminal basic stretches with optimal combinations being KRx(10-12)K[K/R][K/R] and KRx(10-12)K[K/R]x[K/R]. Linker region has a conserved length due to the distance between minor and major binding site of importin α.

Residues in the linker region also affect the strength of the signal. For instance, for optimal activity, the central linker region has to be rich in acidic residues, but lack them in terminal part. In contrast, hydrophobic residues in the central region reduce the strength of the signal and proline in the terminal region positively affects NLS. (Kosugi, Hasebe, Matsumura, et al., 2009)

In *S. cerevisiae* there are two known nuclear exporters, Crm1 and Msn5 (Kosugi et al., 2009). For Msn5,no consensus NES sequence has been identified and the mechanism of its interaction with the targets is not yet fully understood (Bakhrat et al., 2008; Quilis et al., 2019). Crm1-recognized NES typically consists of conserved hydrophobic residues with the traditional consensus sequence being Lx(2-3)[L/I/V/F/M]x(2-3)Lx[L/I] (Bogerd et al., 1996; la Cour et al., 2003). Some variations include Φx(1-3)Φx(2-3)Φx(1-3)Φ, where x is any amino acid and Φ is L, I, V, M, F, W, C, T or A. Some restrictions include that L or I should fill at least two positions out of four and C, T, A and W should not appear not more than once (Kosugi, Hasebe, Tomita, et al., 2008).

1.8 Phosphorylation of localization signal

Phosphorylation of NLS can affect the efficiency of nuclear transport via several mechanisms. Nuclear import can be stimulated by phosphorylation both by the phosphorylation of NLS itself or by the phosphorylation of the adjacent phosphosites. For example, phosphorylation of the NLS itself can increase the binding affinity for the importin and phosphorylation of the upstream sites can promote the recognition. In addition, phosphorylation can induce conformational changes that make NLS more accessible (Nardozzi et al., 2010). Downregulation by phosphorylation is also possible via NLS inactivation. Serine or threonine sites can overlap the NLS and mimic the action of acidic residues when phosphorylated (Kosugi, Hasebe, Tomita, et al., 2009).

1.9 Cell cycle dependent nuclear localization

Some proteins need to be localized according to the cell cycle phase due to their functions. The cell cycle dependent nuclear transport is mainly regulated by CDK. In *S. cerevisiae* there are a number of proteins that exhibit cell cycle dependent nucleocytoplasmic

shuttling. Some examples are Whi5 (SBF inhibitor), Mcm3 (a protein involved in DNA replication) and Acm1 (APC^{Cdh1} inhibitor).

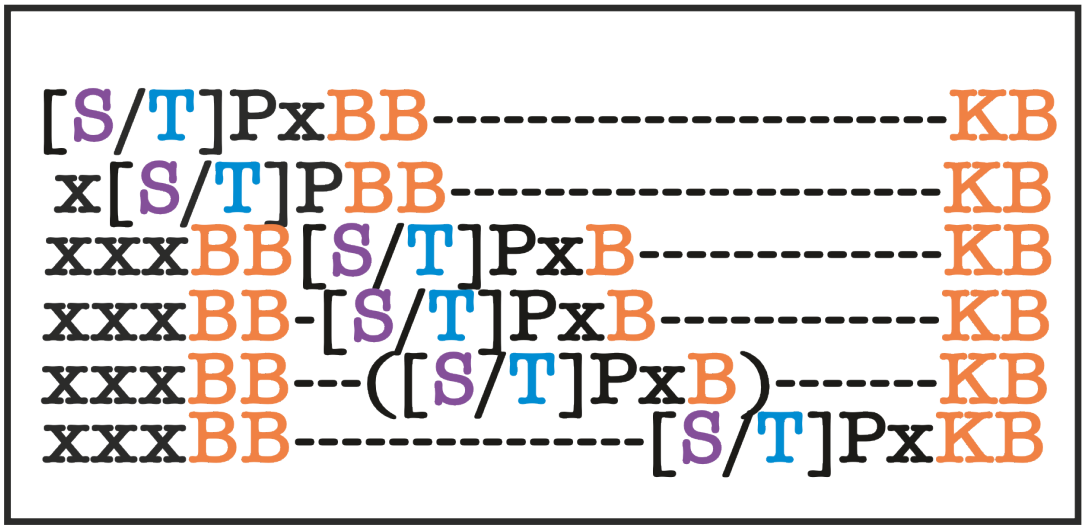

Figure 4. NLS and CDK sites. The scheme shows possible positions of CDK sites overlapping the bipartite NLS (Kosugi et al., 2009).

Kosugi et al., 2009 have tried to predict the CDK1-regulated NLS motifs in *S. cerevisiae* by looking for the overlaps between NLS and consensus CDK site. Using this method, the team has identified five proteins (Yen1, Psy4, Pds1, Msa1, and Dna2) that are showing cell cycle dependent shuttling. Mutations of serine and threonine residues to alanine resulted in loss of localization oscillations, proving that it is mediated by phosphorylation. Three proteins from the list carry a monopartite NLS, the rest – bipartite. It has been shown, that due to the presence of the linker region, bipartite NLSs have more flexibility in the position of the CDK sites (Fig.4). Different positions can also result in the different effect. For example, the presence S/TP sites in the upstream flanking region and the terminal region of the linker sequence result makes NLS phospho-inhibited. On the other hand, CDK sites in the central linker region have an activating effect (Kosugi, Hasebe, Entani, et al., 2008; Kosugi, Hasebe, Matsumura, et al., 2009). As bipartite NLS have more possible CDK sites and NLS overlaps with different localization patterns, it was decided to concentrate on Dna2 and Psy4 that have a bipartite NLS that is overlapping with Cdk concensus sites (Kosugi, Hasebe, Tomita, et al., 2009).

Psy4 is a regulatory subunit of protein phosphatase PP4, involved in the repair of crosslinked DNA (Vázquez-Martin et al., 2008). Dna2 is an ATP-dependent nuclease involved in the repair of DNA double-strand breaks (Stewart et al., 2008; Zhu et al., 2008). NLS sequence in Psy4 is deactivated by phosphorylation and the protein is nuclear in the beginning of the cell cycle, G1/S. In contrast, Dna2 NLS is activated by CDK activity and it is imported to the nucleus during the cell cycle. This behaviour, in principle, corresponds with their functions, as

Dna2 has to be active in S phase, because it is also involved in Okazaki fragment maturation. As for Psy4, there is evidence that the crosslinked repair is restricted to G1 (Kosugi, Hasebe, Tomita, et al., 2009).

2 AIMS

The ability of controlling protein localization during the cell cycle would be beneficial both for the research and industry. It can be used for activation and deactivation of the proteins in the cell cycle dependent manner. Studying the Dna2 and Psy4 NLS and CDK motifs could allow us to develop the localization modules that can be fused with the proteins of interest. The combinations with other known motifs and modules would allow us to create wide variety of the localization patterns. Dna2 and Psy4 CDK sites can still be studied further to better understand of how their localization is mediated and to achieve more control over it. In addition, there is a possibility of identifying CDK-dependent NES.

The aims of this work are:

- Design CDK-dependent nuclear localization modules for use in synthetic biology and cell cycle research.
- Study the phospho-regulation of nuclear localization and export signals in Dna2 and Psy4 in the cell cycle context.
- Analyze the threshold encoding of localization-dependent cell cycle switches.

3 EXPERIMENTAL PART

3.1 MATERIALS AND METHODS

3.1.1 Plasmid construction

All the constructs were made on the backbone of the pRS304 yeast integrative vector. The plasmids carry an ampicillin resistance gene for bacterial selection and *TRP1* gene for yeast selection. Target genes were expressed under ADH1 promoter and fused with C-terminal eGFP. The original vector was provided by Mihkel Örd. Coding sequences for the disordered regions of Dna2 and Psy4 were initially amplified from yeast genomic DNA. For the study of Dna2 we took amino acids-1-100 and 314-441 for Psy4. Both contain bipartite NLS sequence with neighboring SP/TP sites as well as a putative NES sequence.

3.1.2 PCR

Different mutants of the proteins were obtained via PCR from the original plasmids (p18,p21). The total reaction mixture volume was 25 µl. For the reaction, 1 µl of the template was taken (approximately 15 ng). For most of the reactions 5x Phusion HF Buffer (Thermo Fisher Scientific) was used, with some exceptions when 5x Phusion GC Buffer (Thermo Fisher Scientific) was preferred, with final concentration for both of them being 1x. The mixture also contains 200 µM dNTPs (stock concentration 10 mM), 0.75 µM of forward and reverse primers, and 0.04 U/µl. high-fidelity Phusion DNA Polymerase (Thermo Fisher Scientific). Milli-Q H2O was added to adjust the volume to 25 µl.

The following three-step PCR program was used for the amplification (Table 1). The annealing temperature is specific for each set of primers and was estimated using an online tool (*NEB Tm Calculator*). The extension time was calculated based on the synthesis rate of Phusion polymerase approximation (1kb/30s).

After the completion of the program, 6x Orange DNA Loading Dye (Thermo Fisher Scientific) was added to the mixture (final concentration 1x) and loaded on to 1% Agarose TAE gel (40 mM Tris-acetate (pH 8.3), 1 mM EDTA, 1% agarose, 0.05 µl/ µl Atlas ClearSight DNA Stain (BioAtlas)) along with the ZipRuler Express DNA Ladder 2 (Thermo Fisher Scientific). The gel electrophoresis was carried in 1x TAE buffer (40 mM Tris-acetate (pH 8.3), 1 mM EDTA). The gel picture was taken under the UV light. Bands were cut and purified with the FavorPrep™ GEL/PCR Purification Kit (Favorgen) according to the manufacturers' protocol.

Upon purification, PCR product concentrations were measured with Thermo Fisher's NanoDrop 1000 Spectrophotometer.

Cycle step	Temperature	Time	Number of cycles
Initial Denaturation	98 °C	5 min	1
Denaturation Annealing Extension	98 °C 56-60 °C 72 °C	20 s 20 s 9 s (Dna2)/ 12 s (Psy4)	 33
Final extension	72°C	5 min	1

Table 1. PCR program

In most of the clonings the forward primer for the insert carried SalI cutting site and the reverse - SmaI cutting site. However, to avoid long primers for some mutations other cutting enzymes positioned closer to the mutation site were utilized. Another method that was used to mutate the central region of the proteins was stitching PCR. This method includes two rounds of PCR. The sequence is divided into two parts. During the first one 10-15 nucleotides complementary to the end region of another part are introduced to each of them. The protocol and mix were as described above. During the second round of PCR two products from the first step are annealed together via overlap sequences and amplified with the outermost primers. Taking into account their length and concentration, templates are added to the mix in approximately a 1 to 1 ratio. The rest of the mix composition is left unchanged. PCR program includes 3 additional steps (Table 2). For better performance, primers are added only before the second denaturation, so they are not interfering with the annealing of the complementary regions.

3.1.3 Restriction

ThermoFisher's FastDigest™ enzymes were used for the restriction of PCR products. The restriction mixture with a total volume of 20 µl contained 100 ng of the obtained PCR product, 2 µl of 10x FastDigest™ Buffer (final concentration 1x), 1 µl of FastDigest™ enzymes corresponding to primers used and Milli-Q H2O. The mixture was incubated at 37 °C for 45 minutes. Subsequently, enzymes were inactivated to prevent them from interfering with the

ligation. Thermal inactivation conditions were taken from the table provided by the manufacturer.

Cycle step	Temperature	Time	Number of cycles
Initial annealing	72°	2 min	1
Denaturation Annealing Extension	98 °C 60 °C 72 °C	10 s 20 s 9 s (Dna2)/ 12 s (Psy4)	3
Denaturation Annealing Extension	98 °C 56-60 °C 72 °C	10 s 20 s 9 s (Dna2)/ 12 s (Psy4)	33
Final extension	72°	10 min	1

Table 2. PCR program for overlap extension

Vector was prepared from the pRS304-ADH-GFP plasmid, provided by Mihkel Örd). In some cases, other plasmids from this study were used to obtain double mutants. The restriction mixture consisted of approximately 1000 ng of the vector, 3 µl of 10x FastDigest™ Buffer (final concentration 1x), 1 µl of FastDigest™ enzymes matching those of the insert and 1 µl of FastAP (Thermo Fisher Scientific). Milli-Q H2O was added to the final volume of 30 µl. Fast AP is an alkaline phosphatase, that catalyzes the removal of phosphates from the DNA ends, which prevents self-ligation. The reaction was carried out at 37 °C for 1 hour and afterward mixed with Orange DNA Loading Dye (Thermo Fisher Scientific) and loaded on to 1% Agarose TAE gel (40 mM Tris-acetate (pH 8.3), 1 mM EDTA, 1% agarose, 0.05 µl/µl Atlas ClearSight DNA Stain (BioAtlas)) along with the ZipRuler Express DNA Ladder 2 (Thermo Fisher Scientific). The gel electrophoresis took place in 1x TAE buffer (40 mM Tris-acetate (pH 8.3), 1 mM EDTA). Gel electrophoresis also deactivates the enzymes participating in the reaction. The band corresponding to the restricted vector was cut and purified using the same protocol as previously described. The concentration of the purified fragment was measured using Thermo Fisher's NanoDrop 1000 Spectrophotometer.

3.1.4 Ligation

After the preparation of the vector and the insert, they were ligated together. The mixture of the total volume of 10 µl contained the vector and insert in 1 to 3 volume ratio, approximately 400 ng, 1 µl of 10x T4 DNA Ligase buffer (final dilution 1x), T4 DNA Ligase (Thermo Fisher Scientific) in a final concentration of 5 U/µl, and Milli-Q H2O. For the reactions involving inserts restricted with SmaI (blunt end cutters), 1 µl of 50% PEG 4000 solution was also added. PEG is a hydrophobic molecule that helps concentrating the reaction components by taking up extra space. Therefore, it increases the probability of the correct ligation. The reaction was carried at 16 °C for an hour to 3 hours.

3.1.5 Whole-plasmid PCR to introduce mutations

Another method that was used to obtain some of the point mutations is a whole plasmid PCR (*Site Directed Mutagenesis | NEB*). This method allows going straight to the ligation step without the need to restrict the vector and inserts. PCR mixture and program are identical to the ones described above, except for extension time being elongated to 3 minutes (the whole plasmid is approximately 6 kb long). For the ligation, 1 µl of the PCR product is mixed with 2 µl of 10x T4 DNA Ligase buffer (final dilution 1x), 1 µl of FastDigest™ DpnI, 1 µl of PNK (Thermo Fisher Scientific), and 12 µl of Milli-Q H2O. PNK catalyzes the addition of 5′-phosphates that allows subsequent ligation. DpnI cleaves methylated DNA making sure that there is no template DNA left. The mixture is incubated at 37 °C for 30 minutes. Then 2 µl of 50% PEG 4000 solution and 1 µl of T4 DNA Ligase (Thermo Fisher Scientific) (final concentration being 5 U/µl) are added, to make up to the final volume of 20 µl. The mixture is incubated for 10 minutes at room temperature.

3.1.6 Bacterial transformation

Bacterial transformation was performed with E. coli DH5α and NEB Turbo competent cells. Before the transformation competent cells were taken out from the -80 °C fridge and thawed on ice. In a new tube, 1 µl of the ligation mixture was resuspended with 50 µl of cells (England Biolabs). Subsequently, the tube was kept on ice for 30 minutes, followed by the heat shock at 42 °C. The duration of the heat shock is different for two strains: 75 seconds for DH5α and 30 seconds for NEB Turbo. Afterwards, the tubes were again kept on ice for 5 minutes.

Cells were incubated in 400 µl of LB media (10 g/L tryptone (BD Biosciences), 5 g/L yeast extract (Formedium), 10 g/L NaCl) at 37 °C for 45 minutes and then centrifuged at 6000 rpm for 1 minute. Concentrated cells were plated on 100 µg/mL ampicillin (Sigma) LB plates and incubated overnight at 37 °C.

Two colonies from the transformation plate were resuspended in 4 ml of LB media with 100 µg/mL ampicillin. After 12 to 16 hours of growth for DH5α (alternatively 6 to 9 hours for NEB Turbo) cells were collected by centrifugation. Plasmids were extracted using FavorPrep™ Plasmid DNA Extraction Mini Kit (Favorgen) according to the manufacturer's protocol. The concertation of the purified DNA (ng/µl) was measured with Thermo Fisher's NanoDrop 1000 Spectrophotometer. To check for the length of the insert, the plasmid was restricted again using SalI and SmaI enzymes. The reaction mixture contained 1500 ng of the plasmid, 1 µl of 10x FastDigest™ Green Buffer (final concentration 1x), 0,5 µl of SalI and SmaI FastDigest™ enzymes and Milli-Q H2O up to 10 µl. The reaction was carried out at 37 °C for 30 minutes and afterwards loaded on to 1% Agarose gel (40 mM Tris-acetate (pH 8.3), 1 mM EDTA, 1% agarose, 0.05 µl/µl Atlas ClearSight DNA Stain (BioAtlas)) along with the ZipRuler Express DNA Ladder 2 (Thermo Fisher Scientific). The gel electrophoresis was performed in 1x TAE buffer (40 mM Tris-acetate (pH 8.3), 1 mM EDTA). The DNA was visualized with UV and plasmids containing the correct length insert were sent for Sanger sequencing to the Estonian Biocentre.

3.1.7 Yeast transformation

RV298 strain obtained from Rainis Venta was used as a basis for the strains used in the study. RV298 is a haploid derivative of w303 with bar1 deletion and Whi5 tagged with mCherry. The genotype of RV298 is leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15 [phi+] bar1::hisG WHI5-mCherry-SpHIS5). Constructs were transformed to RV298 or clb5::CLB2 strains. All the strains used in this study are listed in the Supplementary Table 1.

On the day of transformation, cells from the fresh plate were transferred by the streaking stick into a flask with 50 ml YPD (10 g/L yeast extract (Formedium), 20 g/L peptone (Formedium), 20 g/L glucose (Oriola)). The culture was grown in the 30 °C shaker till the optical density reaching 0.8-1, approximately 6 hours. The OD was measured at 600 nm wavelength with Ultrospec 10 cell density meter (Amersham Biosciences).

In order to be integrated into the yeast's genome prior to the transformation plasmids have to be linearized via restriction with FastDigest™ enzyme PacI (Thermo Fisher Scientific).

The reaction mixture contained 4ul of plasmid DNA, 1 µl of 10x FastDigest™ Buffer, 1 µl of PacI, and Milli-Q H2O up to 10 µl. The reaction was carried out at 37 °C for 30 minutes.

The grown yeast culture was transferred to 50 mL centrifuge tubes and centrifuged for 2 minutes at 2000 rpm. The supernatant was discarded and the cells were resuspended in 1 mL of sterile buffer I (100 mM lithium acetate in 0.5 x TE (5 mM Tris-HCl (pH 8), 0.5 mM EDTA)) buffer and transferred to 1,5 ml tubes. Afterwards, cells were centrifuged down again at 3600 rpm for 60 seconds. The supernatant was discarded, cells were resuspended in two times the cell volume of buffer I and left for 10 minutes at room temperature to incubate.

Meanwhile, Salmon Sperm DNA (SSDNA) was kept at 100 °C to boiling. 1 µl of the linearized plasmid was mixed with 10 ul of SSDNA. After the end of the incubation 100 µl of the cells were added along with 700 µl of sterile PEG/lithium acetate solution (40% PEG 3350, 100mM lithium acetate, 10 mM Tris-HCl (pH 8), 1 mM EDTA) and resuspended thoroughly. Subsequently, 48 µl of DMSO was added and the mixture was resuspended again. It was then kept at 42 °C for 40 minutes. Afterwards, the mixture was left on ice for 2 minutes to chill and was then centrifuged for 60 seconds at 6000 rpm. The supernatant was discarded and cells were re-suspended in 1 mL 1x TE buffer (10 mM Tris-HCl (pH 8), 1 mM EDTA). The mixture was centrifuged one more time for 1 minute at 3600 rpm at room temperature. Supernatant was removed again. Cells were resuspended in 200 µl of 1x TE buffer and plated on synthetic complete –TRP glucose (SC-TRP) plates (7 g/L yeast nitrogen base without amino acids (BD Biosciences), 2 g/L SC-TRP powder (MP Biomedicals), 20 g/L glucose (Oriola), 20 g/L agar). Plates were incubated at 30 °C for 2-3 days.

3.1.8 Time lapse microscopy

Time lapse microscopy was used to measure the fluorescence intensity and track the localization of the the GFP signal in unsynchronized cells. The experiments allowed to see the cell growth and the changes in the GFP fluorescence intensity in real time. As previously mentioned, in the constructs used in this study mutant NLS sequences from Dna2 and Psy4 proteins were fused to the eGFP signal to track the changes in the protein location. All the proteins were expressed under the ADH1 promoter. During the course of the cell cycle, GFP signal was changing its intensity by either concentrating in the nucleus or dispersing in the cytoplasm. By studying these oscillations, it is possible to draw conclusions about the strength of the NLS and the cyclins responsible for the phosphorylation of the sites of interest.

Two colonies from transformation plates were streaked out on the new –TRP glucose plates and checked under the microscope for the presence of the GFP signal. Positive colonies

were then transferred with a streaking stick to the 3 ml of SC–TRP 2% glucose medium. 200 µl of the culture was then diluted in 3 ml of the same media and incubated on a 30 °C shaker overnight. After incubation cells were thoroughly vortexed.

From each sample, 0.4 µl was pipetted on a 24 x 50 x 0.08 mm micro cover glass and covered with a piece of complete supplement mixture (CSM) with 2% glucose 1.5 % agarose gel (7 g/L yeast nitrogen base (BD Biosciences, 0.79 g/L CSM powder (Formedium), 20 g/L glucose (Oriola), 1.5% NuSieve GTG agarose (Lonza)). Additional longer slices CSM glucose agar were put around the pieces with cells to prevent them from drying during the experiment. A 20 x 20 mm micro cover glass was put on top of the gel pieces to restrict the movement of the gel pieces and also avert drying. Finally, a plastic cover was placed on top of the 24 x 50 glass and fixed.

The experiment was held for 8 hours, pictures were taken for 161 timepoints with 3 minutes gaps in between. Each time 3 pictures were obtained: phase-contrast, GFP and mCherry. Zeiss Axio Observer.Z1 microscope with 63x/1.4 Oil M27 Plan-Apochromat objective (Zeiss) was used for all the experiments. Colibri LED modules at 25% power were used for the fluorescence imaging. GFP images were obtained at the wavelength of 470 nm with 15 ms exposure time, mCherry - at 555 nm and 750 ms exposure. During the experiment, the temperature has to remain at 30 °C for the optimal growth of the yeast cells. This is reached via Tempcontrol 37-2 digital (Pecon). Each experiment contains on average 10 positions, 1 or 2 per agar piece. Positions were tracked through the experiment using the automated stage and ZEN software. Focus was stabilized using Definite Focus.

All the obtained pictures were converted to grayscale and quantified as described in Doncic et al., 2013. The graphs of nuclear GFP were obtained and plotted in Matlab. The values were normalized to the total signal to avoid the mistake from the plasmid copynumber difference between the strains

3.1.9 Western Blot

For the detection of proteins of interest, western blot technique was used. Cultures for the experiment were grown in 50 ml YPD at 30 °C to OD 0.8. 5 ml of the culture was taken to 15 ml tubes, collected by centrifugation at 3600 rpms for 1 min and flash-frozen with liquid nitrogen. Cell pellet was resuspended in 200 µl of urea lysis buffer. 200 µl of the glass bead were added to the tube and cells were disrupted in a bead beater at 4 m/s for 40 s.

Obtained lysates were loaded on the 10% acrylamide gel (separating gel: 0.375 M Tris-HCl (pH 8,8), 10% acrylamide [29:1 acrylamide:bis-acrylamide], 0,1% SDS; stacking gel:

0.125 M Tris-HCl (pH 6,8), 5% acrylamide [29:1 acrylamide:bis-acrylamide], 0,1% SDS)) along with the 2 µl of PageRuler™ Prestained Protein Ladder. Proteins were separated using SDS-PAGE for 60 min at 15 mA. The gel was soaked in the Semi-Dry buffer(25 mM Tris, 192 mM glycine, 0,1% SDS), proteins were transferred to nitrocellulose membranes with Pierce G2 Fast Blotter (Thermo Scientific) for 60 min using the standard semi-dry transfer program. After the transfer the membrane was left at room temperature on a tilting shaker in a blocking solution (5% Milk powder in TBS-T buffer) for 1 hour. Then the solution was replaced with the 3% Milk powder in TBS-T buffer containing 1 to 500 dilution of the anti-HA.11 epitope tag antibody (clone 16B12, BioLegend Cat. No. 901501) and left on the shaker for 30 to 60 mins. Afterwards, the membrane was thoroughly washed with TBS-T and put into the secondary antibody solution for 30 mins. After that the membrane was washed again. The results were visualized on the film using SuperSignal™ West Pico PLUS Chemiluminescent Substrate (Thermo Fisher Scientific).

3.1.10 Phos tag

Based on the obtained graphs, we have hypothesized that the phosphorylation in both proteins, especially Psy4, that has more sites, can be happening in the stepwise fashion. To test this, it was decided to run Phos-tag experiments.

Phos-tag™ SDS-PAGE method allows the separation of the proteins based on the phosphorylation level and the number of phosphorylated sites. Phos-tag is a is an acrylamide pendant that is able to co-polymerize in acrylamide gels. Phos-tag forms complexes with Zn^{2+} or Mn^{2+} ions and selectively binds phosphorylated serines, threonines, and tyrosines. The more phosphorylated sites a protein has the lower its electrophoretic mobility will be.

To perform Phos-tag™ SDS-PAGE two wild type proteins were tagged with the 6HA tag using pYM plasmids as described in the article (Janke et al., 2004). After the transformation, 4 colonies from each plate were checked with the Western blot as described in the section above to prove the presence of the tag.

Positive colonies were grown overnight in the 5 ml YPD. In the morning the starters were diluted in 50 ml till the OD 0.1 and grown in the 30 °C shaker till the OD 0.3 µl. Afterwards, the cells were treated with 1 µg/ml α-factor for 3 hours. The cells were then washed for three times with 50 ml YPD and returned to the shaker. 5 ml aliquots were taken at 0, 10, 20, 30, 40, 50, 60 and 70 min timepoints. Cultures were immediately centrifuged down at 3600 rpms for 1 min and flash-frozen.

Cell pellets were lysed as previously described. Lysates were separated with Phos-tag SDS-PAGE for 1.5 hours at 15 mA. The 8% acrylamide 50 uM Phos-tag Mn^{2+} gels were prepared according to ”Detection of Multisite Phosphorylation of Intrinsically Disordered Proteins Using Phos-tag SDS-PAGE”.

Phos-tag decreases the transferring efficiency. Therefore, after the run the gel was soaked in the TBS-T buffer containing 10 mM EDTA to wash off Mn^{2+}. Proteins were transferred to nitrocellulose membranes using iBlot™ Gel Transfer Device (Thermo Fisher Scientific). The membrane was soaked in the Ponceau S dye to check the transfer efficiency, the dye was then removed using TBS-T buffer. Subsequently, the membrane was washed with the primary and secondary antibodies as described above.

3.2 RESULTS

Several Cdk1-regulated nuclear localization modules have been identified, however, the mechanisms controlling their phosphorylation dynamics in the cell cycle have not been studied (Harreman et al., 2004). Further, it is often not thoroughly studied how one or another phosphorylation site contributes to the regulation of the protein localization. In this work, the phosphorylation sites proximal to the NLS sequences in Dna2 and Psy4 were separately studied to estimate how they affect the localization dynamics. As the main technique, time-lapse fluorescence microscopy was used during which the intensity of nuclear and cytoplasmic GFP fluorescence was measured. As the cells in the experiment were growing asynchronously, the nucleocytoplasmic shuttling of Whi5-mCherry was also monitored and used to synchronize the individual cell cycles. Whi5 is a nucleocytoplasmic shuttling protein with cell cycle-dependent localization, whose export marks the START point in late G1 (Doncic et al., 2011). On average, *S.cerevisiae* cell cycle lasts for 90 minutes, and the transition to the S phase happens around 12 minutes after the START (Örd et al., 2019).

3.2.1 Dna2 – phospho-activated nuclear import

Dna2 nuclear import is activated by phosphorylation, and Dna2 was found to be cytoplasmic in G1, but mainly nuclear in S and M phase cells (Kosugi et al., 2009).

In this study, the intrinsically disordered region containing 100 N-terminal amino acids of Dna2 protein was taken as a model for Cdk1-activated nuclear import module (Fig. 5A**).** The performed microscopy experiments show that this segment is enough for regulatedlocalization, as the localization of the GFP signal was similar to the wild type Dna2 published previously (Kosugi et al., 2009). The Dna2(1-100)-GFP protein is cytoplasmic in G1 phase when Whi5-mCherry is nuclear, but shortly after the START point it starts being imported to the nucleus, reaching maximal nuclear concentration in S phase at 20 minutes after the START, and at around 60 minutes from START it starts to accumulate to the cytoplasm again (Fig. 5B).

The Dna2 1-100 region contains two overlapping bipartile NLS motifs and two Cdk1 phosphorylation sites: T at the position 4 and S at the position 17 (Fig. 5A). Importantly, both sites have a K in the +3 position, which makes them optimal Cdk1 sites. Previous work has shown that S17A mutant, leading to loss of phosphorylation at this position, is constantly cytoplasmic (Kosugi et al., 2009).

To estimate whether the phosphorylation of T4 is important for the NLS activation, the effect of T4A mutation on the localization dynamics of Dna2(1-100)-GFP was tested. As seen from the graph, the nuclear GFP levels of T4A mutant do not reach the wild type levels, indicating that the absence of phosphorylation at the position 4 does not allow a full activation of NLS. Interestingly, T4 is located 13 amino acids upstream from the SP site, which makes it a putative Cks1 priming site for S17 (Fig. 5A, B).

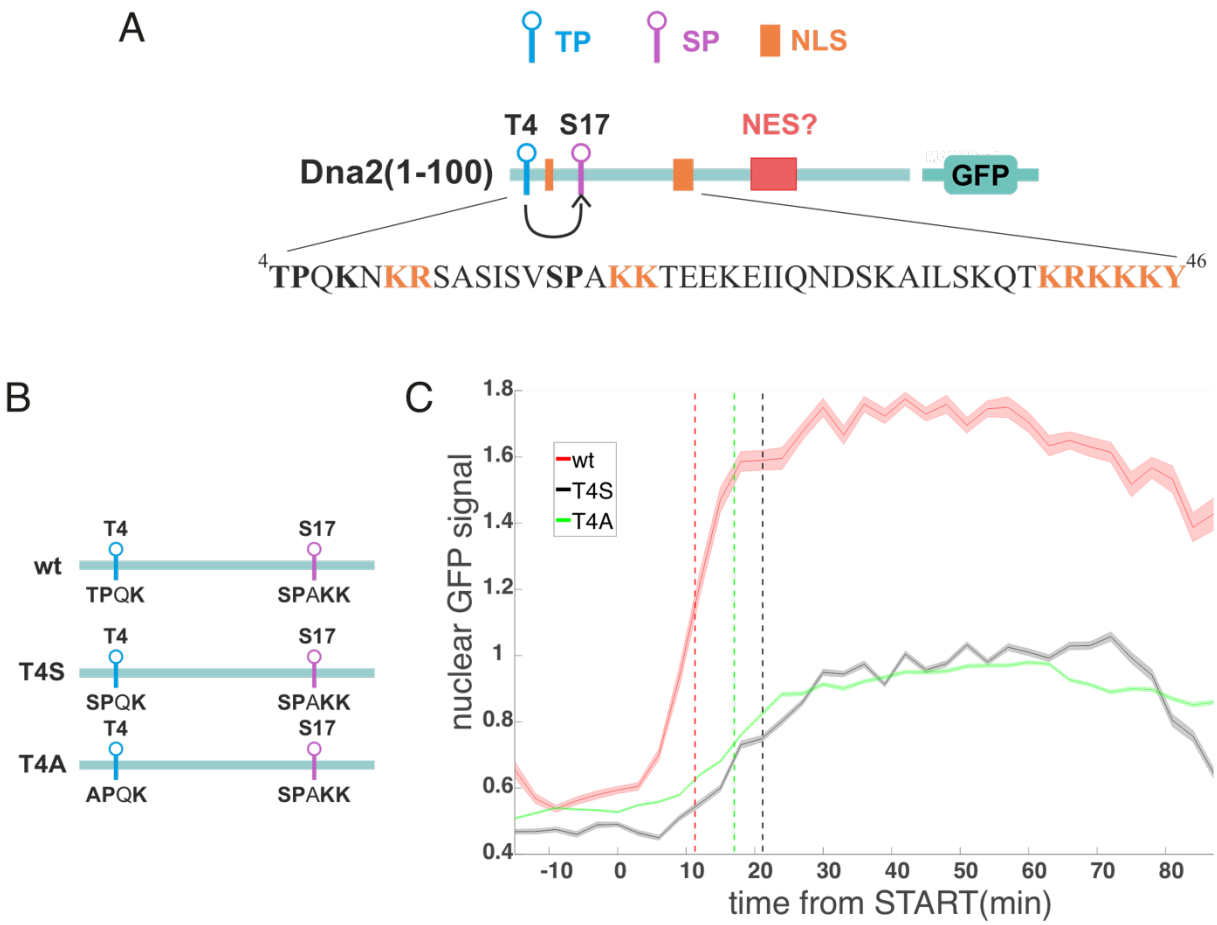

Figure 5. Dna2 T4 primes phosphorylation of S17 for efficient nuclear import of Dna2(1-100)-GFP. A) The scheme shows the positions of the bipartite NLS, putative NES and Cdk1 sites in the Dna2 localization module. Docking connection between the SP site and the putative Cks1 priming site is shown with the black arrow. B) The diagrams show the mutations of the Cks1 priming site. C) The mean relative intensity of the nuclear GFP signal in different strains calculated from cells synchronized by the START point. Here and further, SEM error bars are used. Dashed line shows the time of 50% of signal being imported.

To test the hypothesis, T4 was mutated to serine, as Cks1 can only recognize phosphothreonines (Kõivomägi et al., 2011).Both T4A and T4S mutants show similar trends, which leads to the conclusion that TP site is a priming site for S17 rather than having a direct contribution to NLS activation, as the T4S mutation is not expected to affect the phosphorylation at this site. In comparison to the wild type, nuclear accumulation of Dna2(1-

100)-GFP in T4A and T4S mutants is much weaker, as the maximal intensity of the nuclear signal is almost two times lower and maximal nuclear concentration in mutants is achieved about 15-20 minutes later than in the wild type. This indicates that without the Cks1 docking, phosphorylation of S17 is less effective (Fig. 5C).

To further analyze how CDK substrate targeting mechanisms affect the dynamics of Dna2(1-100)-GFP nuclear import, S17 (SPAKK), which is an optimal site in wild type module, was mutated to suboptimal consensus site (K20Q K21Q) (Fig. 6A). It was previously discovered that K21A mutation did not affect the nuclear localization (Kosugi et al., 2009). In K20Q K21Q mutant, with both disrupted middle part of NLS and the suboptimal SP site, the localization of the protein was similar to the wild type, however, the intensity of the nuclear GFP signal was lower and accumulated slower. Suboptimal sites have lower affinity towards cyclin-CDK complexes and require more kinase activity for phosphorylation, which can explain the detected delay (Örd et al., 2019).

To test the efficiency of suboptimal SP site phosphorylation without the support of Cks1 docking, the priming site T4 was mutated to serine in the construct with K20Q K21Q mutation. This led to no nuclear accumulation of the module during the cell cycle. This indicates that for the sufficient phosphorylation of S17, the site has to be either optimal or primed by a Cks1 docking site.

For Cks1 to enhance phosphorylation of secondary sites, the Cks1 binding site should be positioned 12-16 amino acids upstream from the phosphorylation site. Therefore, the next step was to test the effect of the distances between the two sites to potentially tune the NLS activity during the cell cycle (Fig. 6C). In the wild type background, the insertion of four amino acids downstream of T4 delayed the import for a few minutes. Combining the increased distance with the suboptimal SP site led to very inefficient nuclear import (Fig. 6D). In comparison to T4S K20Q K21Q, where GFP signal was uniform throughout the cell cycle, in T4+4 K20Q K21Q there is a slight increase in nuclear GFP signal around 15 minutes from the START (Fig. 6D). This is in agreement with previously published results showing that increasing the distances between the two sites decreases the efficiency of Cks1-mediated phosphorylation, whereas mutation of the priming site to serine abolishes Cks1 binding. Therefore, these experiments combined with data from Kosugi et al. 2009 indicate that the nuclear import is activated by phosphorylation of S17, and that this is primed by phosphorylated T4 via Cks1 docking. Interestingly, the efficiency and timing of nuclear import can be tuned to some extent by manipulating the Cks1 docking efficiency and the specificity of S17.

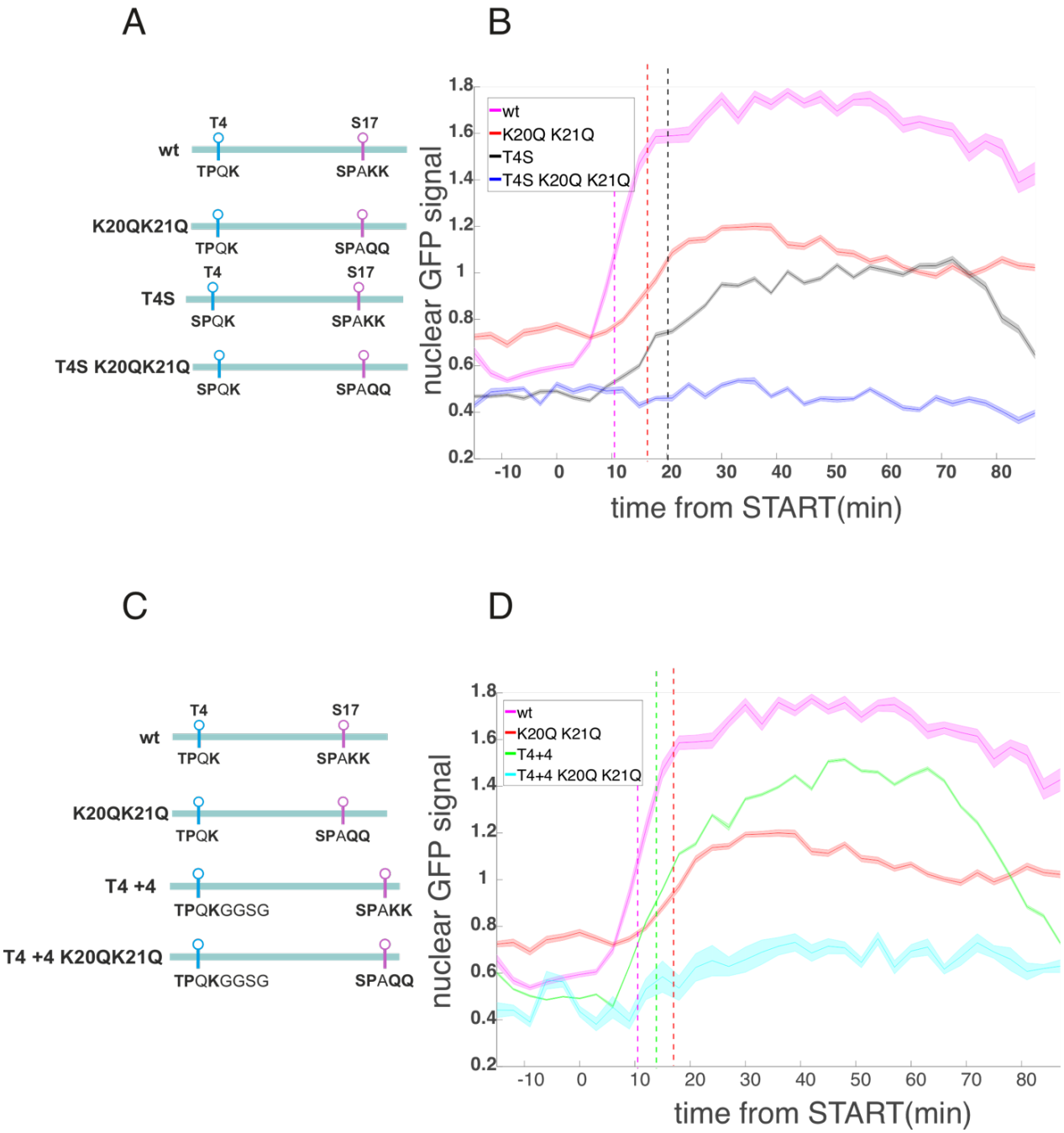

Figure 6. Cdk1 specificity determinants affect the timing and extent of Dna2(1-100)-GFP nuclear import. A)The diagrams show the mutations of two K residues in the linker region of bipartite NLS, that make SP site optimal. C) The diagrams show the changes in the distance between the priming site and S17. B,D) The figures show the relative intensity of the nuclear GFP signal in different strains. Dashed line shows the time of 50% of signal being imported.

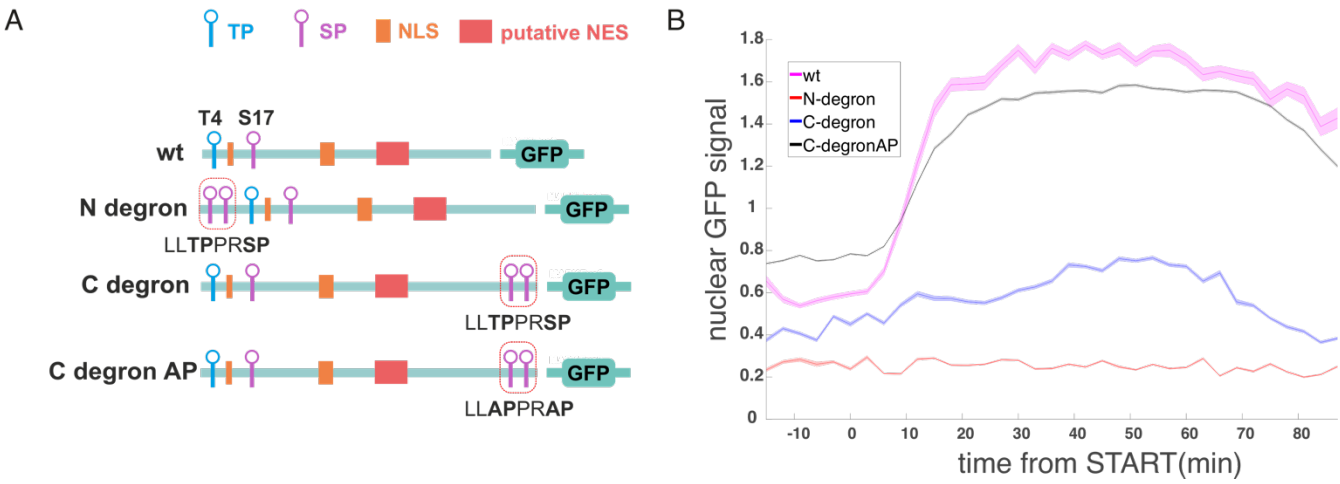

Figure 7. Phospho-regulated nuclear import and degradation of cytoplasmic proteins. A) The diagrams show the position of the cyclin E degron (LLTPPRSP) fused to the Dna2 localization module. B) The figure shows the relative intensity of the nuclear GFP signal in different strains.

3.2.1.1 Degradation of cytoplasmic proteins via combination of nuclear degron and NLS

One of the possible ways to change and study the existing metabolic fluxes is to degrade the involved enzymes. To achieve controllable degradation, phosphodegrons can be fused to proteins to target them for degradation. Proteins are degraded in the nucleus via SCF^{Cdc4}, which has a well-defined effective phosphodegron sequence (Berset et al., 2002). Controllable degradation of cytoplasmic proteins is more difficult to achieve, as SCF^{Grr1} phosphodegrons are less defined and are partially active even without phosphorylation and therefore not suitable for switch-like response (Berset et al., 2002). One way to overcome this complication is to induce degradation by controlled nuclear import. This way, the cytoplasmic protein is first shuttled to the nucleus where it will be then degraded by SCF^{Cdc4}, which targets the well-defined modular degrons. Phosphoactivated Dna2 NLS module could be used for this purpose. To test this, mammalian cyclin E degron was fused to the Dna2 NLS either N- or C-terminally (Fig. 7A). For negative control, to test the specific effect of the degron, an inactive version of the phosphodegron was used with the phosphorylation sites mutated to A (Fig. 7A).

As can be seen from the graph, the construct with the mutant degron (C degron AP) mimics wild type. The N-terminal degron is more efficient, as there is no nuclear accumulation of the GFP, pointing to quick degradation of the protein upon nuclear import. The GFP signals are lower also before and at the START point, likely because GFP has a slow maturation time and if it is degraded during the cell cycle, it will not give fluorescent signal during the short G1 period (Fig. 7B).

The effect of the C-terminal degron is weaker and some of the GFP still accumulates in the nucleus (Fig. 7B). This shows that the fusions of nuclear import and degradation modules are functional and that the fused modules can potentially be used to regulate the levels of metabolic proteins.

3.2.2 Psy4 – phospho-inhibited nuclear localization

In contrast to Dna2, nuclear import of the second protein, Psy4, is phospho-inhibited and the protein is nuclear only during the G1 phase (Kosugi et al., 2009).

As the aim was to find a minimal regulated protein localization module, initially the amino acids in positions 314 to 375 from the Psy4 were used. This region contains a bipartile NLS with four surrounding phosphorylation sites, S337 and S364, and T320 and T347 (Fig. 8A). Both TP sites are optimal, whereas the SP sites are suboptimal. The GFP signal for the Psy4(314-375)-GFP was constantly nuclear, which could be due to lack of nuclear export activity of this region. Therefore, although Psy4 does not have an identified NES, it was decided to take a longer stretch, amino acids 314 to 441 from Psy4, containing the whole C-terminal disordered region. This was enough to recover the cell-cycle-dependent oscillations in localization, which leads to the assumption, that C terminus contains a putative NES. Besides, two Cdk1 phosphorylation sites, S409 and S434, are present in the C-terminal region (Fig. 8A). As shown on the graph, the GFP signal of Psy4(314-441) is nuclear in G1, and shortly after the START, it is quickly exported into the cytoplasm, reaching maximal cytoplasmic concentrations at 30 minutes after the START (Fig. 8B).

It is expected that two TP sites in the upstream flanking region and the terminal region of the linker sequence (T320 and T347) are the most important in NLS inhibition. In fact, double T320A T347A mutants were previously shown to have a constant nuclear signal during the cell cycle (Kosugi et al., 2009).

In this work, the contribution of each phosphorylation site was studied to gain a better understanding of how the localization of this protein is regulated. Time-lapse microscopy analysis revealed that in all single mutants, the nuclear GFP signal is still decreasing during the cell cycle, however, in T320A, T347A and S364A the export is slowed down. On the other hand, S337A mutant behaves similarly to the wild type, which suggests that this SP site does not contribute to the NLS regulation.

The experiments indicate that T347 is the most important phosphorylation site for the NLS deactivation. In its absence, the lowest nuclear concentration is reached 50 minutes later than in the wild type and it never reaches the wild type baseline. In T320A after a fast start,

around 20 minutes from START nuclear export is slowing down, but it is more effective than T347A and reaches the wild-type levels. S364A mutant shows a similar trend, however, the phosphorylation of this residue seems to be important in later phases of the cell cycle, as the GFP signal is imported back to nucleus earlier and faster than in the wild type. Besides, in the early G1 it is also more nuclear (Fig. 8B).

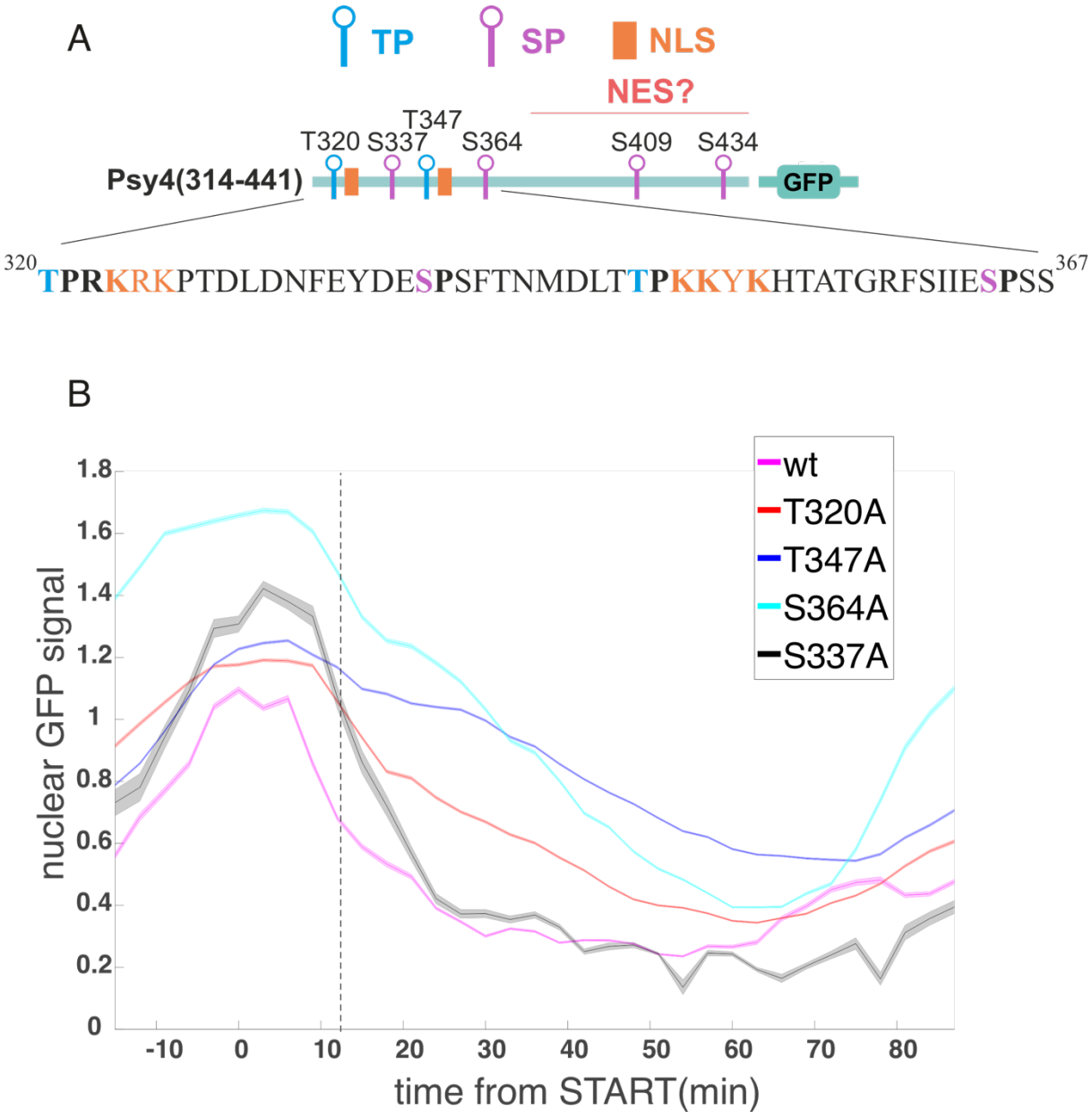

Figure 8. Contribution of different phosphorylation sites to the nuclear export. A) The diagrams show the positions of Cdk1 sites, bipartite NLS and putative NES position in Psy4 localization module. B) The figure shows the relative intensity of the nuclear GFP signal in single mutants. The dashed line shows the timing of 50% nuclear GFP signal in the wild-type strain.

Interestingly, in the phosphorylation site mutants, the phosphorylation seems to happen in a stepwise manner, as there is a period when GFP signal is stable, approximately between 15 and 30 minutes after START, which corresponds to the time of the cyclin waves emergence (Fig. 8B).

3.2.2.1 The impact of different cyclin-Cdk1 complexes on Psy4(314-441) nuclear export

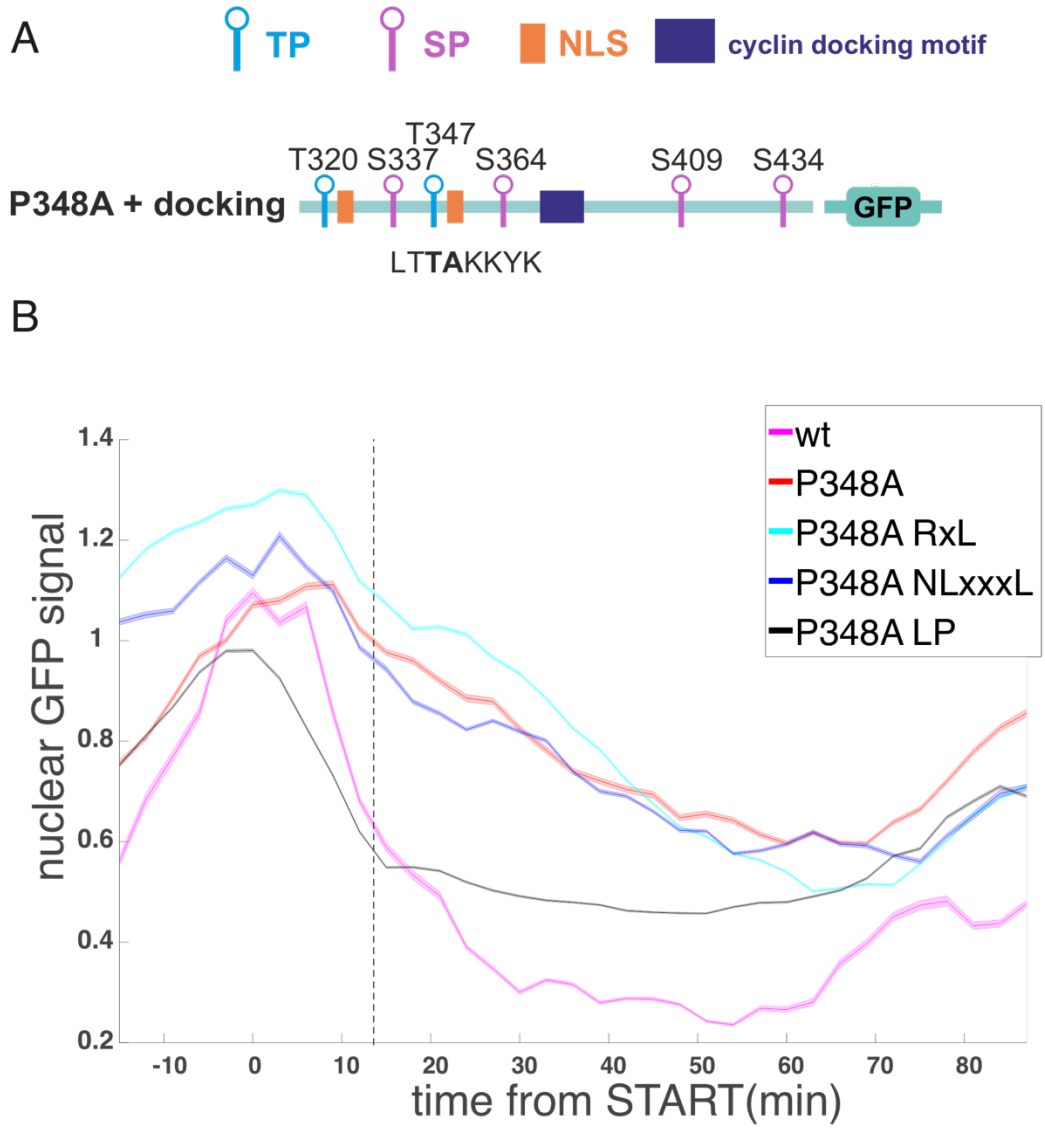

Figure 9. The effect of cyclin docking motifs in Psy4 P348A nuclear export. A) The scheme shows the positions of cyclin docking sites and the mutation of the P348. B) The figure shows the relative intensity of the nuclear GFP signal in different strains. The dashed line shows the time of 50% nuclear GFP signal in the wild-type strain.

With the aim to understand how the timing of Cdk1-dependent nuclear localization switches is controlled, it was next decided to try to modulate the speed of nuclear export by

adding cyclin docking sites next to the S364 site (Fig. 9A). All docking sites were added to the P348A mutant that is similar to the T347A as the phosphorylation of this site is inhibited. Initially, P348A was chosen as a background, as it was thought that a proximal cyclin docking site can restore the phosphorylation of T even in the absence of P, as seen previously with phospho-degrons (Örd et al., 2019).

As expected, the addition of the Cln2 docking site LP motif results in a fast export of the GFP signal after START. However, further phosphorylation remains inhibited and wild type NLS activity is not restored. Presumably, without the presence of the phosphate residue in the linker region nuclear import cannot be fully inhibited (Fig. 9B).

S-Cdk1 docking motifs NLxxxL and RxL affect the phosphorylation in later phases. Interestingly, mutant carrying NLxxxL show increased nuclear export around 10 mins from START, as soon as Clb5 has emerged,and the effect is more clear than that of RxL. This may point to the higher affinity of NLxxxL towards Clb5 in comparison to RxL (Ilona Faustova, unpublished data). In both mutants though, the nuclear concentration of the protein stays higher than in the wild type, which again shows that maximal inhibition can only be reached when T347 is phosphorylated (Fig. 9B).

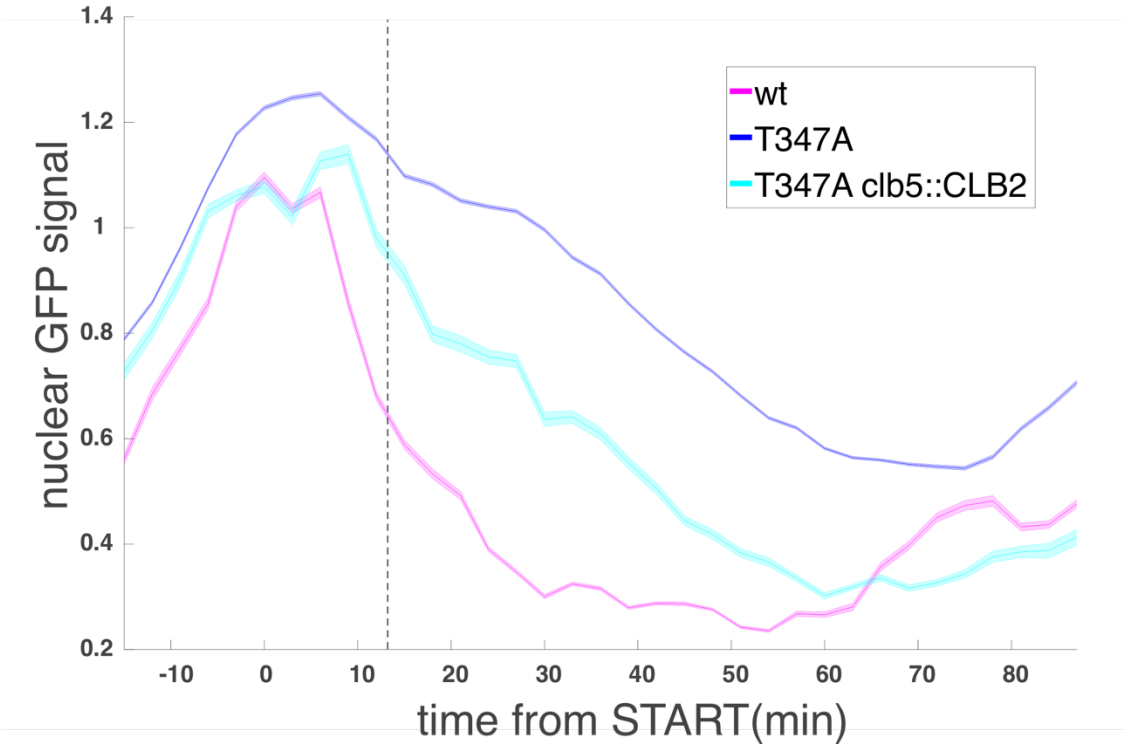

Figure 10. Changing the order of cyclin expression changes the dynamics of Psy4(314-441 T347A)-GFP nuclear export. The figure shows the relative intensity of the nuclear GFP signal in different strains. The dashed line shows the level of 50% nuclear GFP signal in the wild-type strain.

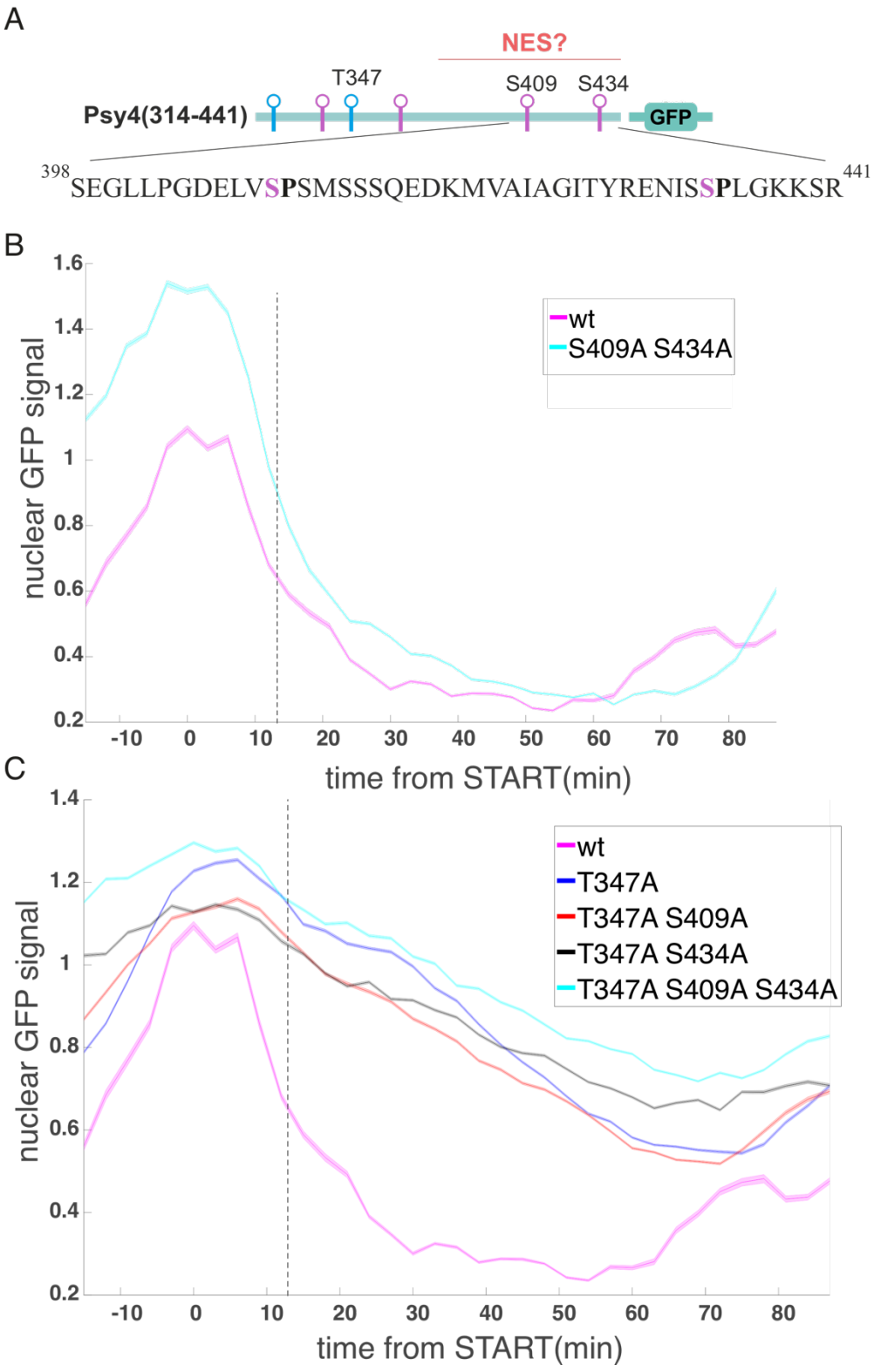

Figure 11. Phosphoregulated export in Psy4. A) The scheme shows a Psy4 segment putatively containing NES and the Cdk1 sites presumably influencing the export. B,C) The figures show the relative intensity of the nuclear GFP signal in different strains.

To better understand the contribution of different cyclins in phosphorylation of these modules, T347A plasmid was also transformed to the strain, where Clb2 is expressed from the Clb5 promoter. As the Clb2-Cdk1 complex has higher activity, it was expected to that GFP is

exported to greater extent. Around 10 minutes after the START, the GFP signal starts to be exported quickly, slightly slowing down about 10 minutes later. In contrast to the mutants with docking motifs, around 60 mins from the START nuclear GFP signal drops to the wild type level. The clb5::CLB2 strain contains two copies of mitotic cyclin CLB2, and Clb2-Cdk1 has higher intrinsic activity compared to Clb5-Cdk1. These results show that by increasing CDK activity, the kinase activity threshold for complete nuclear export of the T347A mutant is reached.

The results obtained from these experiments again point to the stepwise phosphorylation of Psy4 and the effect of different cyclin-Cdk1 complexes having different affinities.

3.2.2.2 Phospho-regulated export

As stated before, the truncated version of the Psy4 module, amino acids 314 to 375 is constantly nuclear, indicating that the C-terminal part of the protein seems to be responsible for its export. The presence of two SP sites at the position 409 and 434 can lead to the assumption that export of Psy4 is also phosphoregulated (Fig. 11A).

To test the hypothesis, S434A S409A mutant strains was made. However, these mutations did not affect the localization, only in the early G1 the signal was slightly more nuclear than in the wild type (Fig. 11B). However, when the combinations of the same mutations with the more active NLS in T347A, made phosphodependent export noticeable. The phosphorylation of S434 residue seems to be the more important, as the GFP signal of T347A S434A mutant is almost constantly nuclear. Besides, T347A S434A S409A behaves similarly to the T347A S434A, which also points to the dominant role of this residue. In T347A S409A strain GFP signal is exported in the similar fashion with T347A. (Fig. 11C)

High number of phosphorylation sites in Psy4 localization module can putatively allow a more precise control over localization. Additional regulation of export can give a rise to more potential combinations and localization patterns.

3.2.3 Phos tag

To directly track the phosphorylation of the proteins of interest, Phos-tag SDS-PAGE Western blot experiments were conducted with the wild type localization modules tagged with 3HA. The cultures were arrested in G1 by α-factor, released and samples were collected at 0, 10, 20, 30, 40, 50, 60, and 70 minutes after release to the cell cycle (Fig. 12).

For Dna2, there is already one phosphorylated site, presumably priming site T4 in the G1-arrested cells. At 20 minutes after the release from α-factor arrest, there is a sharp increase in the amount of double phosphorylated form and by 30 mins there is almost no unphosphorylated protein. This supports the data from the microscopy experiments where a sharp increase in the nuclear signal was seen at around 15 minutes after the START. The fraction of protein being in double phosphorylated form has only started decreasing at 70 minutes upon start, which also aligns with the obtained microscopy data, where a slow export of the GFP signal starts around 70 mins (Fig. 12).

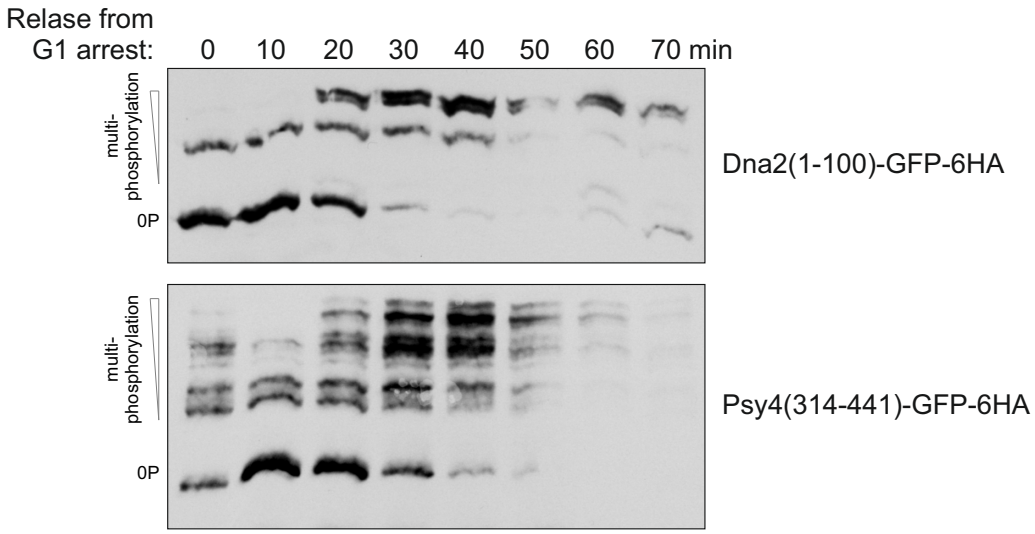

Figure 12. Multisite phosphorylation of wild-type localization modules. The pictures above show the results of the Phos-tag Western blot experiments with the wild type localization modules. The lowest bands correspond to unphosphorylated proteins, the more sites are phosphorylated, the higher the band is located. Band intensity depends on the protein concentration.

The main difference between Psy4 and Dna2 localization modules is the steepness of response. In Dna2 phosphorylation and activated import happens at once, in a switch-like manner. For Psy4, as expected, phosphorylation is happening step by step, with only an insignificant amount of protein being fully phosphorylated at 20 mins after START, and its amount gradually growing for the next 20 minutes. At 40 minutes after START the most prominent band is a fully phosphorylated protein, with almost no protein being partially phosphorylated. At 0 timepoint there are also some phosphoforms present, which can be phosphorylated non-CDK sites or residual CDK phosphorylation. Further Phos tag experiments with mutant strains can help to identify the order of sites being phosphorylated (Fig. 12).

4 DISCUSSION

Dna2 and Psy4 localization modules described in this study can be potentially used as a synthetic biology tool. However, there is still a lot of room for research, as it is not fully understood to what extent the import and export can be tuned by mutating the modules. Aspects that are important for future applications are the timing of export and import, rate of response, and protein levels in the nucleus and cytoplasm. Furthermore, in this study, the mechanisms of the timing of nucleocytoplasmic shuttling were studied. Previously, several ways of encoding the different thresholds were described using phosphodegrons (Örd et al., 2019). However, unlike the degrons, which are rapidly ubiquitinated upon phosphorylation, the phosphorylation sites in localization signals constantly dephosphorylated, thus the kinase activity threshold of these sites is higher. This hypothesis is supported by the finding that the threshold of non-consensus sites is not reached for the reversible localization output, as seen with the P348A mutation (Fig. 9B), whereas non-consensus sites are phosphorylated in degrons of very early Cdk1 substrates (Örd et al., 2019).

In this work, a Cks1 priming site in the Dna2 sequence was identified. The importance of the distance between priming and CDK sites for the activation of NLS was studied, and it was shown that it is possible to mediate the levels of nuclear protein by the addition of extra amino acids (Fig. 6D). It can be further investigated whether the gradual change in distance will result in a gradual change of levels. Previously, on the example of degrons, it was shown that the distance between the priming site and the phosphorylation site confers to the CDK thresholding (Örd et al., 2019). This study shows that Cks1 connections are also critical in regulating phosphorylation of localization switches, however, loss of specificity has a stronger effect on the level of nuclear accumulation rather than the timing of the switch, in contrast to the degron-based output published previously (Örd et al., 2019).

Also in this work, the importance of the optimal CDK sites for the localization was studied. Suboptimal sites could encode for later targets, as they require higher cyclin-CDK activity, which explains delay in the import in the Dna2 K20Q K21Q mutant (Fig. 6B). Furthermore, in the case of NLS regulation, basic residues in the center of the linker region most probably have a dual function, both enhancing CDK binding and working as a part of NLS itself. Psy4 localization module also has several optimal CDK sites, their mutations might have interesting results, as the absence of K makes NLS weaker and at the same time the deactivating phosphorylation is also lower.

The second localization module, Psy4, has more phosphorylation sites and therefore potentially more room for regulation (Fig. 8A). In the wild type, export occurs in a switch-like manner, and mutations of different sites decrease the rate of export. T347A mutant is fully

exported in the clb5::CLB2 strain with higher CDK activity, indicating that full phosphorylation of the remaining sites is sufficient for export. This suggests that the remaining sites have quite high kinase acitivity thresholds and indicates that the wild-type module is exported rapidly at G1/S due to redundancy of phosphorylation sites in NLS inactivation. This is leads to a hypothesis that the number of redundant phosphorylation sites can be an important determinant in timing of the phosphorylation switch in the cell cycle.

An interesting feature with the Psy4 modules is a plateau appearing at around 15 minutes after the START, at the time when cyclin Clb5 levels are still increasing (Fig. 13) The plateau might be explained by the product inhibition, which could inhibit Clb5-Cdk1 activity after accumulation of phosphorylated TP sites that bind to the Cks1 pocket. Clb5-Cdk1 could be more sensitive for such product inhibition because in addition to Cks1, it utilizes high affinity binding with RxL motifs in many S phase targets. Due to the lack of high affinity RxL interaction, the mitotic CDK could be less sensitive towards product inhibition. Importantly, this mechanism could explain why S phase CDK cannot reach mitotic thresholds, even when overexpressed.

In wild type, the plateau is less defined and seems shorter. Interestingly, Clb5-mediated phosphorylation results in a fast export, although only happening for a short time, while Clb2-Cdk1 complex phosphorylates the target for a longer period, but its effect on the localization is less defined (Fig. 13). However, Clb2-mediated phosphorylation in T347A strain has a bigger impact (Fig. 13). Possibly, phosphorylation sites that are the most important for the NLS deactivation are fired already in G1/S, and therefore later phosphorylation does not have a huge effect.

Currently, only the single mutants were made, however, their combinations can also show more time patterns of export and import. There is also a possibility to tune the distances between the CDK sites and NLS, which can tune the inhibiting effect of phosphorylation.

Besides, further manipulations on the localization modules can give more information on the role of separate cyclin-CDK complexes in the cell-cycle dependent nucleocytoplasmic shuttling of proteins. This study once again underlines the importance of cyclin localization in Cdk1 specificity. For instance, Dna2 is presumably phosphorylated in the cytoplasm by Cln2-Cdk1 in G1, and after being imported it is most probably further phosphorylated by nuclear cyclin-CDK complexes to be kept in the nucleus. It is interesting, however, that in Cln deletion strain Dna2 import is not fully inhibited, which leads to the question of which other cyclins can mediate the cytoplasmic phosphorylation. Future work with deletion strains and cyclin docking motifs can shed more light on this question. The use of the docking motifs can also be useful

for the creation of controllable localization modules. As was shown on the example of Psy4, cyclin targeting can tune the steepness of response and protein levels.

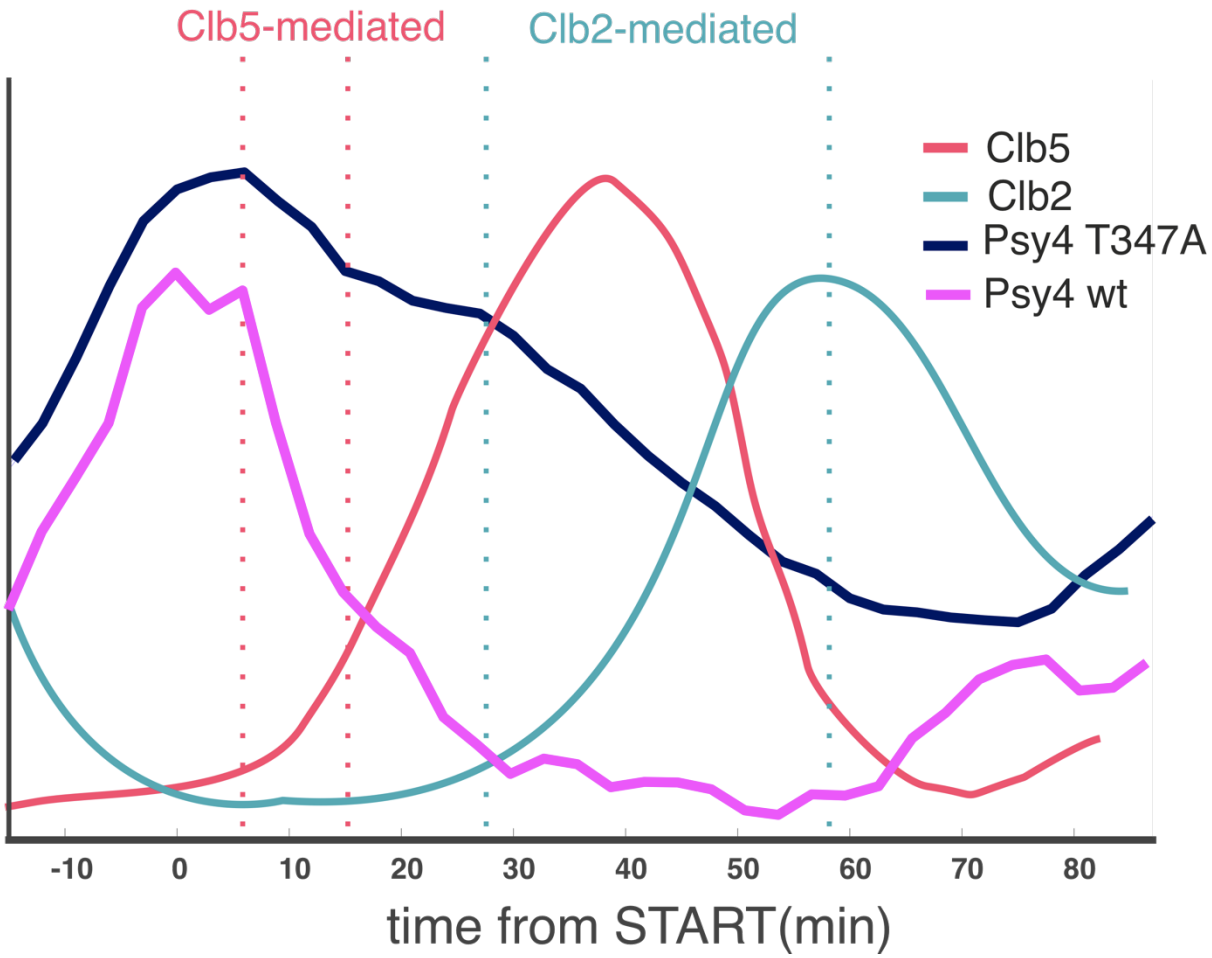

Figure 13. Psy4 phosphorylation by different cyclin-CDK complexes. On the scheme trends for Psy4 wt and T347A are overlayed with the graphs of expression of cyclins (Örd et al., 2019).

Another part of the localization module, NES, was not thoroughly studied in this work. Although Psy4 was shown to have a phospho-dependent nuclear export, the NES has not been mapped in Psy4 and Dna2. In the C-terminal of the Dna2 module, there is a sequence of hydrophobic residues, that are defining for NES, and the mutations in this region can be done to prove the hypothesis (Fig. 5A). Besides, identification of NES can allow making more truncations to shorten the module to make it easier to use.

Precise mapping of the localization signals is important for the creation of accurately controlled modules. Besides, there is another possibility to mediate localization modules, that were not described in this study, is to mutate NLS and NES sequences. Combining CDK sites mutations with NLS and NES sequences of different strengths can result in a wide range of localization modules.

Future research should not be restricted to Dna2 and Psy4 only. Bipartite NLS allows for a wide variety of CDK sites. Further study of similarly regulated proteins can help to create

a full range of localization modules with different time of import and export and different levels of nuclear and cytoplasmic protein. It is also yet to be defined whether two differently regulated NLSs can be fused to create short impulses of nuclear import or export.

The next step is to fuse localization modules to metabolic proteins to test their efficiency with different proteins to estimate to which extent it is possible to predict their localization in the cell at a certain point in the cell cycle. As these modules were yet only fused with the eGFP, it is unknown how does the size of the protein, its conformation, and maturation time can affect the modules.

This work can be the first step in the creation of a set of localization modules that can be used to regulate metabolic enzymes. Potentially, this approach can also be used in biomedical studies, as the NLS sequences in eukaryotes are similar.

SUMMARY

Subcellular localization plays an important role in protein regulation. Proteins containing NLS and NES signals can be actively shuttled in and out of the nucleus, the presence of CDK phosphorylation sites next to them might point to the cell-cycle-dependent nucleocytoplasmic shuttling. In this work specifically, the overlaps between CDK sites and bipartite NLS were studied in two proteins, Dna2 and Psy4, with differently regulated localization. NLS in Dna2 is activated by phosphorylation, and in Psy4 – deactivated. For both proteins were identified short localization modules, sufficient for the reproduction of the wild-type-like subcellular localization patterns. These modules can be fused to other proteins to alter their activity.

In this study, Cks1 priming site, important for the activation of import, was identified in Dna2(1-100). The meaning of the distance between the discovered site and the phosphorylation site was also researched. Movement of the Cks1 priming site can result in the changes in nuclear levels of protein and the time of import. Additionally, it was proved that Dna2 localization module can be fused to the phosphodegron for the targeted degradation of cytoplasmic proteins in the nucleus, which proves the possible versatile applications of the discovered modules.

On the other hand, in Psy4(314-441) localization module the significance of four phosphorylation sites proximal to NLS was studied by eliminating phosphorylation sites one by one. The presence of several phosphorylation sites cumulatively creates a switch-like response in the wild type profile. The addition of cyclin docking sites enhancing the phosphorylation can partially restore wild-type dynamics, and can be potentially used to regulate the time of export, and can also give insights to the localization of cyclin-Cdk1 complexes. Furthermore, Psy4(314-441) localization module has a putative NES in the C-terminus, which is also regulated by CDK activity, which makes its localization module even more tunable and diverse.

Overall, this work can be the first step in the creation of a synthetic biology tool that can find a lot of implications in design and research. Additionally, it can give more understanding on the importance of the subcellular localization of proteins in the cell cycle control.

REFERENCES

Amon, A., Tyers, M., Futcher, B., & Nasmyth, K. (1993). Mechanisms that help the yeast cell cycle clock tick: G2 cyclins transcriptionally activate G2 cyclins and repress G1 cyclins. *Cell*, *74*(6), 993–1007. https://doi.org/10.1016/0092-8674(93)90722-3

Bakhrat, A., Baranes-Bachar, K., Reshef, D., Voloshin, O., Krichevsky, O., & Raveh, D. (2008). Nuclear export of Ho endonuclease of yeast via Msn5. *Current Genetics*, *54*(5), 271–281. https://doi.org/10.1007/s00294-008-0216-8

Basco, R. D., Segal, M. D., & Reed, S. I. (1995). Negative regulation of G1 and G2 by S-phase cyclins of Saccharomyces cerevisiae. *Molecular and Cellular Biology*, *15*(9), 5030–5042. https://doi.org/10.1128/mcb.15.9.5030

Bállega, E., Carballar, R., Samper, B., Ricco, N., Ribeiro, M. P., Bru, S., Jiménez, J., & Clotet, J. (2019). Comprehensive and quantitative analysis of G1 cyclins. A tool for studying the cell cycle. *PLOS ONE*, *14*(6), e0218531. https://doi.org/10.1371/journal.pone.0218531

Berset, C., Griac, P., Tempel, R., la Rue, J., Wittenberg, C., & Lanker, S. (2002). Transferable Domain in the G1 Cyclin Cln2 Sufficient To Switch Degradation of Sic1 from the E3 Ubiquitin Ligase SCFCdc4 to SCFGrr1. *Molecular and Cellular Biology*, *22*(13), 4463–4476. https://doi.org/10.1128/mcb.22.13.4463-4476.2002

Bhaduri, S., & Pryciak, P. M. (2011). Cyclin-specific docking motifs promote phosphorylation of yeast signaling proteins by G1/S Cdk complexes. *Current Biology*, *21*(19), 1615–1623. https://doi.org/10.1016/j.cub.2011.08.033

Blondel, M., Galan, J., Chi, Y., … C. L.-T. E., & 2000, undefined. (n.d.). Nuclear-specific degradation of Far1 is controlled by the localization of the F-box protein Cdc4. *Embopress.Org*. Retrieved April 30, 2020, from https://www.embopress.org/doi/full/10.1093/emboj/19.22.6085

Bogerd, H. P., Fridell, R. A., Benson, R. E., Hua, J., & Cullen, B. R. (1996). Protein sequence requirements for function of the human T-cell leukemia virus type 1 Rex nuclear export signal delineated by a novel in vivo randomization-selection assay. *Molecular and Cellular Biology*, *16*(8), 4207–4214. https://doi.org/10.1128/mcb.16.8.4207

Brewer, B. J., Chlebowicz-Sledziewska, E., & Fangman, W. L. (1984). Cell cycle phases in the unequal mother/daughter cell cycles of Saccharomyces cerevisiae. *Molecular and Cellular Biology*, *4*(11), 2529–2531. https://doi.org/10.1128/mcb.4.11.2529

Conti, E, Uy, M., Leighton, L., Blobel, G., & Kuriyan, J. (1998). Crystallographic analysis of the recognition of a nuclear localization signal by the nuclear import factor karyopherin alpha. *Cell*, *94*(2), 193–204. https://doi.org/10.1016/s0092-8674(00)81419-1

Conti, Elena. (2002). Structures of importins. In *Results and problems in cell differentiation* (Vol. 35, pp. 93–113). Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-44603-3_5

Cooper, G. M. (2000). *The Eukaryotic Cell Cycle*.

Coudreuse, D., & Nurse, P. (2010). Driving the cell cycle with a minimal CDK control network. *Nature*, *468*(7327), 1074–1080. https://doi.org/10.1038/nature09543

Cross, F. R. (1990). Cell cycle arrest caused by CLN gene deficiency in Saccharomyces cerevisiae resembles START-I arrest and is independent of the mating-pheromone signalling pathway. *Molecular and Cellular Biology*, *10*(12), 6482–6490. https://doi.org/10.1128/mcb.10.12.6482

Dahmann, C., Diffley, J. F., & Nasmyth, K. A. (1995). S-phase-promoting cyclin-dependent kinases prevent re-replication by inhibiting the transition of replication origins to a pre-replicative state. *Current Biology : CB*, *5*(11), 1257–1269. https://doi.org/10.1016/s0960-9822(95)00252-1

Dingwall, C., & Laskey, R. A. (1991). Nuclear targeting sequences--a consensus? *Trends in Biochemical Sciences*, *16*(12), 478–481. https://doi.org/10.1016/0968-0004(91)90184-w

England Biolabs, N. (n.d.). *Datasheet for NEB Turbo Competent E. coli (High Efficiency) (C2984; Lot 98)*.

Enserink, J. M., & Kolodner, R. D. (2010). An overview of Cdk1-controlled targets and processes. In *Cell Division* (Vol. 5, p. 11). BioMed Central. https://doi.org/10.1186/1747-1028-5-11

Fitch, I., Dahmann, C., Surana, U., Amon, A., Nasmyth, K., Goetsch, L., Byers, B., & Futcher, B. (1992). Characterization of four B-type cyclin genes of the budding yeast Saccharomyces cerevisiae. *Molecular Biology of the Cell*, *3*(7), 805–818. https://doi.org/10.1091/mbc.3.7.805

Fontes, M. R. M., Teh, T., Jan, D., Brinkworth, R. I., & Kobe, B. (2003). Structural basis for the specificity of bipartite nuclear localization sequence binding by importin-α. *Journal of Biological Chemistry*, *278*(30), 27981–27987. https://doi.org/10.1074/jbc.M303275200

Fontes, M. R. M., Teh, T., & Kobe, B. (2000). Structural basis of recognition of monopartite and bipartite nuclear localization sequences by mammalian importin-α. *Journal of Molecular Biology*, *297*(5), 1183–1194. https://doi.org/10.1006/jmbi.2000.3642

Fried, H., & Kutay, U. (2003). Nucleocytoplasmic transport: Taking an inventory. In *Cellular and Molecular Life Sciences* (Vol. 60, Issue 8, pp. 1659–1688). https://doi.org/10.1007/s00018-003-3070-3

Ghiara, J. B., Richardson, H. E., Sugimoto, K., Henze, M., Lew, D. J., Wittenberg, C., & Reed, S. I. (1991). A cyclin B homolog in S. cerevisiae: Chronic activation of the Cdc28 protein kinase by cyclin prevents exit from mitosis. *Cell*, *65*(1), 163–174. https://doi.org/10.1016/0092-8674(91)90417-W

Hadwiger, J. A., Wittenberg, C., Mendenhall, M. D., & Reed, S. I. (1989). The Saccharomyces cerevisiae CKS1 gene, a homolog of the Schizosaccharomyces pombe suc1+ gene, encodes a subunit of the Cdc28 protein kinase complex. *Molecular and Cellular Biology*, *9*(5), 2034–2041. https://doi.org/10.1128/mcb.9.5.2034

Hadwiger, J. A., Wittenberg, C., Richardson, H. E., de Barros Lopes, M., & Reed, S. I. (1989). A family of cyclin homologs that control the G1 phase in yeast). *Proceedings of the National Academy of Sciences of the United States of America*, *86*(16), 6255–6259. https://doi.org/10.1073/pnas.86.16.6255

Hahn, S., Maurer, P., Caesar, S., & Schlenstedt⁎, G. (2008). *Classical NLS Proteins from Saccharomyces cerevisiae*. https://doi.org/10.1016/j.jmb.2008.04.038

Harreman, M. T., Kline, T. M., Milford, H. G., Harben, M. B., Hodel, A. E., & Corbett, A. H. (2004). Regulation of nuclear import by phosphorylation adjacent to nuclear localization signals. *Journal of Biological Chemistry*

Holt, L. J. (2012). Regulatory modules: Coupling protein stability to phopshoregulation during cell division. In *FEBS Letters* (Vol. 586, Issue 17, pp. 2773–2777). No longer published by Elsevier. https://doi.org/10.1016/j.febslet.2012.05.045

Jeffrey, P. D., Russo, A. A., Polyak, K., Gibbs, E., Hurwitz, J., Massagué, J., & Pavletich, N. P. (1995). Mechanism of CDK activation revealed by the structure of a cyclinA-CDK2 complex. *Nature*, *376*(6538), 313–320. https://doi.org/10.1038/376313a0

Kitazono, A. A., & Kron, S. J. (2002). An essential function of yeast cyclin-dependent kinase Cdc28 maintains chromosome stability. *Journal of Biological Chemistry*, *277*(50), 48627–48634. https://doi.org/10.1074/jbc.M207247200

Koch, C., & Nasmyth, K. (1994). Cell cycle regulted transcription in yeast. *Current Opinion in Cell Biology*, *6*(3), 451–459. https://doi.org/10.1016/0955-0674(94)90039-6

Koch, C., Schleiffer, A., Ammerer, G., & Nasmyth, K. (1996). Switching transcription on and off during the yeast cell cycle: Cln/Cdc28 kinases activate bound transcription factor SBF (Swi4/Swi6) at Start, whereas Clb/Cdc28 kinases displace it from the promoter in G2. *Genes and Development*, *10*(2), 129–141. https://doi.org/10.1101/gad.10.2.129

Kõivomägi, M., Örd, M., Iofik, A., Valk, E., Venta, R., Faustova, I., Kivi, R., Balog, E. R. M., Rubin, S. M., & Loog, M. (2013). Multisite phosphorylation networks as signal processors for Cdk1. *Nature Structural and Molecular Biology*, *20*(12), 1415–1424. https://doi.org/10.1038/nsmb.2706

Kõivomägi, M., Valk, E., Venta, R., Iofik, A., Lepiku, M., Morgan, D. O., & Loog, M. (2011). Dynamics of Cdk1 Substrate Specificity during the Cell Cycle. *Molecular Cell*, *42*(5), 610–623. https://doi.org/10.1016/j.molcel.2011.05.016

Kosugi, S., Hasebe, M., Entani, T., Takayama, S., Tomita, M., & Yanagawa, H. (2008). Design of Peptide Inhibitors for the Importin α/β Nuclear Import Pathway by Activity-Based Profiling. *Chemistry and Biology*, *15*(9), 940–949. https://doi.org/10.1016/j.chembiol.2008.07.019

Kosugi, S., Hasebe, M., Matsumura, N., Takashima, H., Miyamoto-Sato, E., Tomita, M., & Yanagawa, H. (2009). Six classes of nuclear localization signals specific to different binding grooves of importinα. *Journal of Biological Chemistry*, *284*(1), 478–485. https://doi.org/10.1074/jbc.M807017200

Kosugi, S., Hasebe, M., Tomita, M., & Yanagawa, H. (2008). Nuclear export signal consensus sequences defined using a localization-based yeast selection system. *Traffic*, *9*(12), 2053–2062. https://doi.org/10.1111/j.1600-0854.2008.00825.x

Kosugi, S., Hasebe, M., Tomita, M., & Yanagawa, H. (2009). Systematic identification of cell cycle-dependent yeast nucleocytoplasmic shuttling proteins by prediction of composite motifs. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(25), 10171–10176. https://doi.org/10.1073/pnas.0900604106

la Cour, T., Gupta, R., Rapacki, K., Skriver, K., Poulsen, F. M., & Brunak, S. (2003). NESbase version 1.0: a database of nuclear export signals. *Nucleic Acids Research*, *31*(1), 393–396. https://doi.org/10.1093/nar/gkg101

Lanker, S., Valdivieso, M. H., & Wittenberg, C. (1996). Rapid degradation of the G1 cyclin Cln2 induced by CDK-Dependent phosphorylation. *Science*, *271*(5255), 1597–1601. https://doi.org/10.1126/science.271.5255.1597

Lew, D. J., & Reed, S. I. (1993). Morphogenesis in the yeast cell cycle: Regulation by Cdc28 and cyclins. *Journal of Cell Biology*, *120*(6), 1305–1320. https://doi.org/10.1083/jcb.120.6.1305

Lim, F.-L., & Pic-Taylor, A. (2000). The forkhead protein Fkh2 is a component of the yeast cell cycle transcription factor SFF. *The EMBO Journal*.

Lim, W., Mayer, B., & Pawson, T. (2015). *Cell signaling : principles and mechanisms*.

Loog, M., & Morgan, D. O. (2005). Cyclin specificity in the phosphorylation of cyclin-dependent kinase substrates. *Nature*, *434*(7029), 104–108. https://doi.org/10.1038/nature03329

Makkerh, J. P., Dingwall, C., & Laskey, R. A. (1996). Comparative mutagenesis of nuclear localization signals reveals the importance of neutral and acidic amino acids. *Current Biology : CB*, *6*(8), 1025–1027. https://doi.org/10.1016/s0960-9822(02)00648-6

McGrath, D. A., Balog, E. R. M., Kõivomägi, M., Lucena, R., Mai, M. v., Hirschi, A., Kellogg, D. R., Loog, M., & Rubin, S. M. (2013). Cks confers specificity to phosphorylation-dependent CDK signaling pathways. *Nature Structural and Molecular Biology*, *20*(12), 1407–1414. https://doi.org/10.1038/nsmb.2707

Mok, J., Kim, P. M., Lam, H. Y. K., Piccirillo, S., Zhou, X., Jeschke, G. R., Sheridan, D. L., Parker, S. A., Desai, V., Jwa, M., Cameroni, E., Niu, H., Good, M., Remenyi, A., Ma, J. L. N., Sheu, Y. J., Sassi, H. E., Sopko, R., Chan, C. S. M., … Turk, B. E. (2010). Deciphering protein kinase specificity through large-scale analysis of yeast phosphorylation site motifs. *Science Signaling*, *3*(109), ra12. https://doi.org/10.1126/scisignal.2000482

Morgan, D. O. (1997). CYCLIN-DEPENDENT KINASES: Engines, Clocks, and Microprocessors. *Annual Review of Cell and Developmental Biology*, *13*(1), 261–291. https://doi.org/10.1146/annurev.cellbio.13.1.261

Morgan, D.O. (2007). The Cell Cycle: Principles of Control (London: New Science Press Ltd).
Mosammaparast, N., & Pemberton, L. F. (2004). Karyopherins: From nuclear-transport mediators to nuclear-function regulators. In *Trends in Cell Biology* (Vol. 14, Issue 10, pp. 547–556). https://doi.org/10.1016/j.tcb.2004.09.004
Nardozzi, J. D., Lott, K., & Cingolani, G. (2010). Phosphorylation meets nuclear import: A review. In *Cell Communication and Signaling* (Vol. 8, Issue 1, pp. 1–17). BioMed Central. https://doi.org/10.1186/1478-811X-8-32
Nasmyth, K. (1996). At the heart of the budding yeast cell cycle. In *Trends in Genetics* (Vol. 12, Issue 10, pp. 405–412). Elsevier Ltd. https://doi.org/10.1016/0168-9525(96)10041-X
NEB Tm Calculator. (n.d.). Retrieved April 22, 2020, from https://tmcalculator.neb.com/#!/main
Örd, M., Venta, R., Möll, K., Valk, E., & Loog, M. (2019). Cyclin-Specific Docking Mechanisms Reveal the Complexity of M-CDK Function in the Cell Cycle. *Molecular Cell*, *75*(1), 76-89.e3. https://doi.org/10.1016/j.molcel.2019.04.026
Örd, M., & Loog, M. (2019). How the cell cycle clock ticks. In *Molecular Biology of the Cell* (Vol. 30, Issue 2, pp. 169–172). American Society for Cell Biology. https://doi.org/10.1091/mbc.E18-05-0272
Örd, M., Möll, K., Agerova, A., Kivi, R., Faustova, I., Venta, R., Valk, E., & Loog, M. (2019). Multisite phosphorylation code of CDK. *Nature Structural and Molecular Biology*, *26*(7), 649–658. https://doi.org/10.1038/s41594-019-0256-4
Pemberton, L. F., & Paschal, B. M. (2005). Mechanisms of receptor-mediated nuclear import and nuclear export. In *Traffic* (Vol. 6, Issue 3, pp. 187–198). https://doi.org/10.1111/j.1600-0854.2005.00270.x
Peter, M., & Herskowitz, I. (1994). Direct inhibition of the yeast cyclin-dependent kinase Cdc28-Cln by Far1. *Science*, *265*(5176), 1228–1231. https://doi.org/10.1126/science.8066461
Petsko, G. A., Ringe, D., Defranco, A., Locksley, R., Robertson, M., Craig, N. L., Cohen-Fix, O., Green, R., Greider, C. W., Storz, G., Wolberger, C., Ingham, P., Whitfield, T., Lim, W., Mayer, B., & Pawson, A. (n.d.). *Primers in Biology: Published titles Protein Structure and Function Forthcoming titles: Immunity Molecular Biology Cell Signaling*.
Quilis, I., Taberner, F. J., Martínez-Garay, C. A., Alepuz, P., & Igual, J. C. (2019). Karyopherin Msn5 is involved in a novel mechanism controlling the cellular level of cell cycle regulators Cln2 and Swi5. *Cell Cycle*, *18*(5), 580–595. https://doi.org/10.1080/15384101.2019.1578148
Richardson, H. E., Wittenberg, C., Cross, F., & Reed, S. I. (1989). An essential G1 function for cyclin-like proteins in yeast. *Cell*, *59*(6), 1127–1133. https://doi.org/10.1016/0092-8674(89)90768-x
Richardson, H., Lew, D. J., Henze, M., Sugimoto, K., & Reed, S. I. (1992). Cyclin-B homologs in Saccharomyces cerevisiae function in S phase and in G2. *Genes and Development*, *6*(11), 2021–2034. https://doi.org/10.1101/gad.6.11.2021
Schneider, B. L., Yang, Q. H., & Futcher, A. B. (1996). Linkage of replication to start by the Cdk inhibitor Sic1. *Science*, *272*(5261), 560–562. https://doi.org/10.1126/science.272.5261.560
Schrader, E. K., Harstad, K. G., & Matouschek, A. (2009). Targeting proteins for degradation. In *Nature Chemical Biology* (Vol. 5, Issue 11, pp. 815–822). Nature Publishing Group. https://doi.org/10.1038/nchembio.250
Schulman, B. A., Lindstrom, D. L., & Harlow, E. (1998). Substrate recruitment to cyclin-dependent kinase 2 by a multipurpose docking site on cyclin A. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(18), 10453–10458. https://doi.org/10.1073/pnas.95.18.10453
Schwob, E., Böhm, T., Mendenhall, M. D., & Nasmyth, K. (1994). The B-type cyclin kinase inhibitor p40SIC1 controls the G1 to S transition in S. cerevisiae. *Cell*, *79*(2), 233–244. https://doi.org/10.1016/0092-8674(94)90193-7

Site Directed Mutagenesis | NEB. (n.d.). Retrieved April 12, 2020, from https://www.nebiolabs.com.au/applications/cloning-and-synthetic-biology/site-directed-mutagenesis

Skotheim, J. M., di Talia, S., Siggia, E. D., & Cross, F. R. (2008). Positive feedback of G1 cyclins ensures coherent cell cycle entry. *Nature*, *454*(7202), 291–296. https://doi.org/10.1038/nature07118

Skowyra, D., Koepp, D. M., Kamura, T., Conrad, M. N., Conaway, R. C., Conaway, J. W., Elledge, S. J., & Harper, J. W. (1999). Reconstitution of G1 cyclin ubiquitination with complexes containing SCF(Grr1) and Rbx1. *Science*, *284*(5414), 662–665. https://doi.org/10.1126/science.284.5414.662

Songyang, Z., Blechner, S., Hoagland, N., Hoekstra, M. F., Piwnica-Worms, H., & Cantley, L. C. (1994). Use of an oriented peptide library to determine the optimal substrates of protein kinases. *Current Biology*, *4*(11), 973–982. https://doi.org/10.1016/S0960-9822(00)00221-9

Stern, B., & Nurse, P. (1996). A quantitative model for the cdc2 control of S phase and mitosis in fission yeast. In *Trends in Genetics* (Vol. 12, Issue 9, pp. 345–350). Elsevier Ltd. https://doi.org/10.1016/S0168-9525(96)80016-3

Stewart, J. A., Miller, A. S., Campbell, J. L., & Bambara, R. A. (2008). Dynamic removal of replication protein A by Dna2 facilitates primer cleavage during okazaki fragment processing in Saccharomyces cerevisiae. *Journal of Biological Chemistry*, *283*(46), 31356–31365. https://doi.org/10.1074/jbc.M805965200

Surana, U., Robitsch, H., Price, C., Schuster, T., Fitch, I., Futcher, A. B., & Nasmyth, K. (1991). The role of CDC28 and cyclins during mitosis in the budding yeast S. cerevisiae. *Cell*, *65*(1), 145–161. https://doi.org/10.1016/0092-8674(91)90416-V

Swaffer, M. P., Jones, A. W., Flynn, H. R., Snijders, A. P., & Nurse, P. (2016). CDK Substrate Phosphorylation and Ordering the Cell Cycle. *Cell*, *167*(7), 1750-1761.e16. https://doi.org/10.1016/j.cell.2016.11.034

Ubersax, J. A., & Ferrell, J. E. (2007). Mechanisms of specificity in protein phosphorylation. In *Nature Reviews Molecular Cell Biology* (Vol. 8, Issue 7, pp. 530–541). Nature Publishing Group. https://doi.org/10.1038/nrm2203

Vázquez-Martin, C., Rouse, J., & Cohen, P. T. W. (2008). Characterization of the role of a trimeric protein phosphatase complex in recovery from cisplatin-induced versus noncrosslinking DNA damage. *FEBS Journal*, *275*(16), 4211–4221. https://doi.org/10.1111/j.1742-4658.2008.06568.x

Vodermaier, H. C. (2004). APC/C and SCF: Controlling each other and the cell cycle. In *Current Biology* (Vol. 14, Issue 18). https://doi.org/10.1016/j.cub.2004.09.020

Zhu, Z., Chung, W. H., Shim, E. Y., Lee, S. E., & Ira, G. (2008). Sgs1 Helicase and Two Nucleases Dna2 and Exo1 Resect DNA Double-Strand Break Ends. *Cell*, *134*(6), 981–994. https://doi.org/10.1016/j.cell.2008.08.037

SUPPLEMENTARY MATERIALS

Supplementary table 1

Strain	Background strain	Description	Source
RV298	w303	*MATa, bar1::hisG WHI5-mCherry-SpHIS5*	Rainis Venta
MO367	RV298	*clb5::CLB2 swe1::URA3 Whi5-mCh-SpHis5*	Mihkel Örd
IB1	RV298	*trp1::pRS306(pADH1-Dna2(1-100)-eGFP)*	This study
IB2	RV298	*trp1::pRS306(pADH1-Dna2[1-100 (*T4A*)]-eGFP)*	This study
IB3	RV298	*trp1::pRS306(pADH1-Dna2[1-100 (*T4S*)]-eGFP)*	This study
IB4	RV298	*trp1::pRS306(pADH1-Dna2[1-100 (K20Q K21Q)]-eGFP)*	This study
IB5	RV298	*trp1::pRS306(pADH1-Dna2[1-100 (T4S K20Q K21Q)]-eGFP)*	This study
IB6	RV298	*trp1::pRS306(pADH1-Dna2[1-100 (T4+4)]-eGFP)*	This study
IB7	RV298	*trp1::pRS306(pADH1-Dna2[1-100 (T4+4 K20Q K21Q)]-eGFP)*	This study
IB8	RV298	*trp1::pRS306(pADH1-LLTPPRSP-Dna2(1-100)-eGFP)*	This study
IB9	RV298	*trp1::pRS306(pADH1-Dna2(1-100)-LLTPPRSP-eGFP)*	This study
IB10	RV298	*trp1::pRS306(pADH1-Dna2(1-100)-LLAPPRAP-eGFP)*	This study
IB11	RV298	*trp1::pRS306(pADH1-Psy4(314-375)-eGFP)*	This study
IB12	RV298	*trp1::pRS306(pADH1-Psy4(314-441)-eGFP)*	This study

IB13	RV298	*trp1::pRS306(pADH1-Psy4[314-441 (T320A)]-eGFP)*	This study
IB14	RV298	*trp1::pRS306(pADH1-Psy4[314-441 (T347A)]-eGFP)*	This study
IB15	RV298	*trp1::pRS306(pADH1-Psy4[314-441 (S337A)]-eGFP)*	This study
IB16	RV298	*trp1::pRS306(pADH1-Psy4[314-441 (S364A)]-eGFP)*	This study
IB17	RV298	*trp1::pRS306(pADH1-Psy4[314-441 (P348A)]-eGFP)*	This study
IB18	RV298	*trp1::pRS306(pADH1-Psy4[314-441 (P348A)]-eGFP)*	This study
IB19	RV298	*trp1::pRS306(pADH1-Psy4[314-441 (P348A) NLxxxL]-eGFP)*	This study
IB20	RV298	*trp1::pRS306(pADH1-Psy4[314-441 (P348A) RxL]-eGFP)*	This study
IB21	RV298	*trp1::pRS306(pADH1-Psy4[314-441 (S409A)]--eGFP)*	This study
IB22	RV298	*trp1::pRS306(pADH1-Psy4[314-441 (S409A S434A)]-eGFP)*	This study
IB23	RV298	*trp1::pRS306(pADH1-Psy4[314-441 (T347A S409A)]-eGFP)*	This study
IB24	RV298	*trp1::pRS306(pADH1-Psy4[314-441 (T347A S434A)]-eGFP)*	This study
IB25	RV298	*trp1::pRS306(pADH1-Psy4[314-441 (T347A S409A S434A)]-eGFP)*	This study

NON-EXCLUSIVE LICENCE TO REPRODUCE THESIS AND MAKE THESIS PUBLIC

I, Irina Borovko,

(*author's name*)

1. herewith grant the University of Tartu a free permit (non-exclusive licence) to:

reproduce, for the purpose of preservation, including for adding to the DSpace digital archives until the expiry of the term of copyright,

Cell cycle dependent nuclear localization,

(*title of thesis*)

supervised by MSc, Mihkel Örd.

(*supervisor's name*)

2. I grant the University of Tartu a permit to make the work specified in p. 1 available to the public via the web environment of the University of Tartu, including via the DSpace digital archives, under the Creative Commons licence CC BY NC ND 3.0, which allows, by giving appropriate credit to the author, to reproduce, distribute the work and communicate it to the public, and prohibits the creation of derivative works and any commercial use of the work from ***20/05/2023*** until the expiry of the term of copyright.

3. I am aware of the fact that the author retains the rights specified in p. 1 and 2.

4. I certify that granting the non-exclusive licence does not infringe other persons' intellectual property rights or rights arising from the personal data protection legislation.

Irina Borovko

20/05/2020