UNIVERSITY OF TARTU
Faculty of Science and Technology
Institute of Technology

Aleksandra Panfilova

Cln2-based phosphodegron tags for the regulation of protein stability

Bachelor's Thesis (12 ECTS)

Curriculum Science and Technology

Supervisor(s):
Research Fellow, Ph.D. Nastassia Shtaida
Research Fellow, Ph.D. Rait Kivi

Tartu 2019

Cln2-based phosphodegron tags for the regulation of protein stability

Abstract:

There are several levels of the regulation of protein expression. Protein phosphorylation is a common way of signal transduction in cell signaling pathways and it provides a fast response. Some proteins possess so-called phosphodegron sequences, which being phosphorylated send protein for destruction. In this study, phosphodegron from *Saccharomyces cerevisiae* cyclin Cln2 was used to generate a series of protein tags to affect EGFP fluorescence level. As it was shown by FACS, a fusion of EGFP with degron tags containing either 4 or 5 phosphorylation sites, resulted in 30% decrease in EGFP fluorescence, without significant difference between 4 and 5 phosphosites. Surprisingly, the addition of Cln2 docking motif from Sic1 protein to the degron led to increasing of EGFP fluorescence, in comparison to alanine mutant or non-tagged control. These results require further investigation.

In comparison to the fast response from protein phosphorylation, regulation of protein expression at the promoter level is slow. Combination of an inducible promoter, which activity can be tightly regulated, with phosphodegron tag can provide additional control lever for fine-tuning of protein expression. To do that, yeast strains carrying EGFP under control of inducible pLexAt promoter was created using CRISPR/Cas9-mediated transformation. These strains will be further used in the laboratory for tagging with phosphodegrons.

Keywords:

Cell cycle, Cln2, cyclin degradation, regulation of degradation

CERCS:

P310 Proteins, enzymology

Lühikokkuvõte:

Rakus kontrollitakse valgutasemeid mitmel erineval moel. Signaaliülekandes on levinud meetod valkude fosforüülimine, mis võimaldab väga kiiret vastust keskkonnas toimunud muutustele. Mõnedes valkudes on olemas niiöelda fosfodegronid – valgujärjestused, mille fosforüülimise tagajärjel suunatakse valk lagundamisele.

Antud töös kasutati S. cerevisiae Cln2 valgust pärineva fosfodegroni erinevaid mutante, mis seoti GFP külge ja jälgiti, kuidas GFP fluorestsents selle tagajärjel muutub. FACSiga mõõtmised demonstreerisid, et 4 või 5 fosfosaidiga degronid vähendasid GFP fluorestsentsi

30%. Üllatava tulemuseni viis Sic1 valgust pärineva Cln2 seondumissaidi lisamine degroni järjestusele, mille tagajärjel võis näha valgu degradatsiooniefektiivsuse vähenemist. Saadud tulemused väärivad edasist uurimistööd.

Võrreldes valkude fosforüülimisega on promootori tasemel valguekspressiooni kontroll suhteliselt aeglane. Kombineerides aga valguekspressiooni indutseeritava promootori alt fosfodegronitega, saame tekitada lisakontrolli valguekpressiooni väga täpsel reguleerimisel. Sellise süsteemi loomiseks viidi CrispR/Cas9 meetodi abil pärmi genoomi eGFP koos indutseeritava pLexA promootoriga. Saadud tüve kasutatakse edasiste katsete käigus, kus plaanitakse GFP-d fosfodegronitega modifitseerida.

Võtmesõnad:

Rakutsükkel, Cln2, tsükliini degradatsioon, degradatsiooni regulatsioon

CERCS:

P310 Proteiinid, ensümoloogia

TERMS, ABBREVIATIONS AND NOTATIONS

CDK is an abbreviation from cyclin-dependent kinases, a family of serine/threonine protein kinases playing a key role in cell cycle regulation.

Cdc28 is a single central CDK of *S. cerevisiae.*

NES is an abbreviation from the nuclear export signal. It is a short peptide sequence targeting protein for nuclear transportation out of the nucleus through the nuclear pores.

GSA is a flexible linker created by Waldo, Standish, Berendzen, & Terwilliger (1999) to allow correct folding for GFP fused to any heterologous sequence.

Cyclins are a family of proteins with a common feature of binding and activating CDKs. They do not possess any enzymatic activities by themselves, but they do contain binding regions for different CDK's substrates and localization sequences, which define substrate specificity and subcellular locations of CDK-cyclin complex.

Cln2 is a *S. cerevisiae* G1/S cyclin, contributing to Sic1 degradation and localized mostly in the cytoplasm.

Cks1 is a small polypeptide that binds to a larger lobe of the Cdc28 and promotes multisite phosphorylation by binding downstream phosphosites after phosphorylation.

Grr1 is an F-box protein, one of the possible subunits of SCF ubiquitin-protein ligase, responsible for the binding of target protein.

LP is a region of Sic1 (Figure 3B) that was shown to increase the rate of Cln2-dependent phosphorylation of this inhibitor in comparison to proteins with deleted or mutated motif. It is suggested to be a docking motif for Cln2 cyclin, which promotes Sic1 phosphorylation by binding it. Length of LP motif may vary in different researches, but VLLPP sequence probably has the biggest influence on the cyclin binding.

INTRODUCTION

Precise regulation of protein concentrations in the cell could allow for tight control over its metabolism, or make possible to program cells to perform the specific task through signal transduction. Transcriptional regulation, widely used in synthetic biology, does not allow for rapid change in the protein concentrations. One of the possible solutions can be found in the cell cycle regulatory network. Cyclin-dependent kinases are the key players of this network. Oscillations in their activity, localization and substrate specificity are precisely controlled by cyclin subunit currently bound to it. Combination of transcriptional and post-translational regulations results in rapid turnover of these proteins. Cyclins contain special regions, which being phosphorylated become targets for proteolytic machinery and are sent for destruction. These sequences are called phosphodegrons.

In this study, a variety of phosphodegrons were designed on a base of C-terminal part of Saccharomyces cerevisiae cyclin Cln2. Highly-stable EGFP was used as a reporter in order to visualize the effect of phosphodegrons on protein stability. Cln2-based degrons of two different lengths were used — one containing four and another one five phosphosites. They both showed about 30% decrease in EGFP fluorescence. Constructs with Ser and Thr phosphosites changed to Ala were used as references. In order to increase degradation efficiency, LP, Cln2 docking motif from Sic1 inhibitor protein, was fused to C-terminus of degrons, but the addition of this motif, on the contrary, made EGFP more stable. This was unexpected and requires further investigation. These results provide us a set of phosphodegrons which after further research could be developed for tagging proteins of interest.

Degrons were tested on strains with constitutive expression of EGFP, whereas precisely controlled system would require gene expression regulation as well. Two strains with integrated LexA-ER-B112 system were created for further testing of developed degrons and effect of LP docking site positioning.

1 LITERATURE REVIEW

1.1 Overview of a mitotic cell cycle

One of the tenets of the cell theory, first formulated in the 19th century, states that all cells originate from pre-existing cells. As a cell is a basic structural unit of life, the cell cycle resulting in a division is a fundamental process, required for reproduction of unicellular organisms and genesis, development and repair of multicellular ones.

Differences in cell structures and especially in the complexity of the organization of genetic material predetermine the differences in cell cycles between prokaryotes and eukaryotes. Prokaryotic division, called binary fission, is relatively quick, and basically consists of doubling of a circular chromosome and reconstruction of membrane and cell wall. In contrast to simple division of prokaryotes, eukaryotic mitotic cell cycle consists of four phases: G1, S, G2, and M (Figure 1). G1 and G2 are so-called gap phases, separating two major ones — S, in which DNA replication occurs, and M, where the division of the nucleus and the formation of the new cell takes place. Gap phases are required for cell growth. Altogether, G1, S, and G2 are called interphase. In G1, cell makes a decision about cell cycle progression. If environmental or intracellular conditions are unsatisfying, it can freeze in G1 or even enter so-called G0 (quiescent) phase until conditions will not become favorable. When cell cycle proceeds to S phase, chromosomes are duplicated forming sister chromatids associated in centromeric region. In G2, before entering M phase, cells continue to grow and make sure that DNA is replicated completely and correctly using feedback regulation mechanisms (Morgan, 2007). M phase consists of nuclear (karyokinesis) and cellular (cytokinesis) divisions. The karyokinesis is a highly regulated multistage process. Sister chromatids are separated and brought to the opposite poles of a dividing cell, where each set is packed into a new nucleus. During or after these events, the cytoplasm of a mother cell is divided between two daughters and a membrane is formed between them (Balasubramanian & Bi, 2004; Boettcher & Barral, 2013)

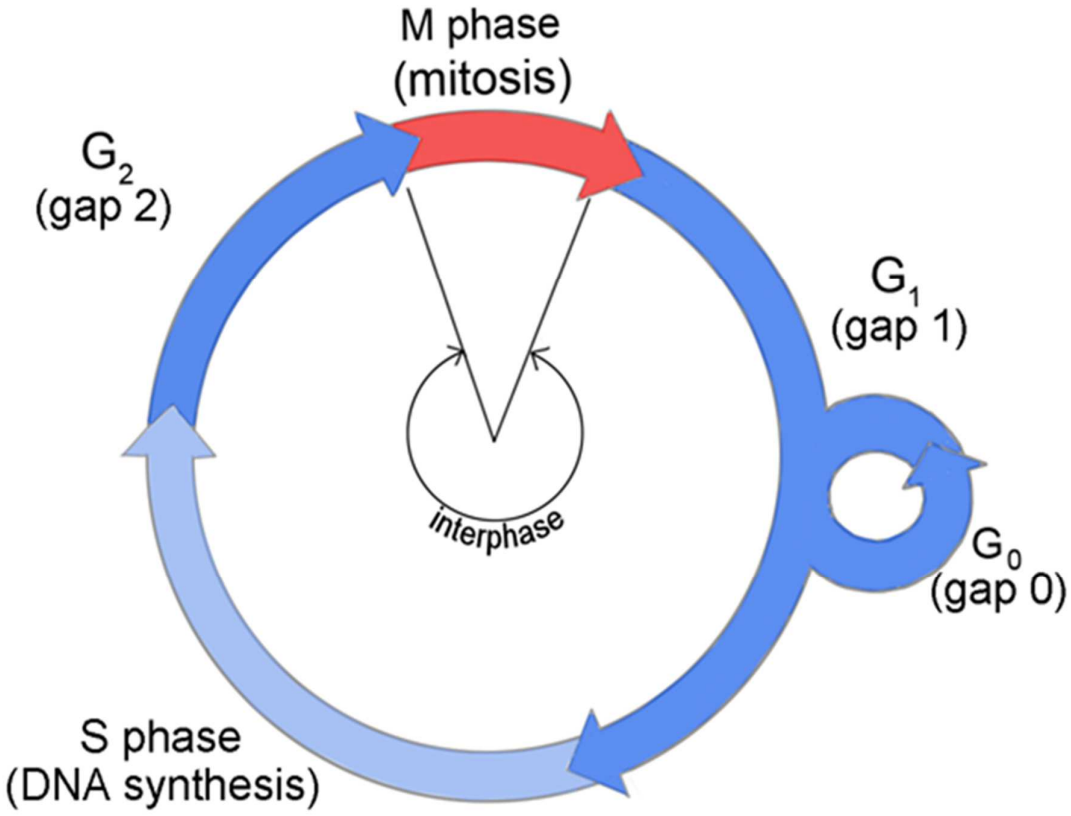

Figure 1. Schematic representation of the mitotic cell cycle flow. Durations of cell cycle phases may vary not only among different species, but also among different cell types of multicellular organisms. (http://cyberbridge.mcb.harvard.edu/mitosis_3.html)

1.2 Introduction to the cell cycle control system

Correct and full transmission of the genetic information and organelles from mother to daughter cell is crucial for the vitality of the offspring, and a cell can precisely control the order and timing of these events. The cell-cycle controlling system is a regulatory signaling network that analyzes both the state of a cell and its environment in every given moment and performs or does not perform certain actions as an output. The main regulators of the system are protein kinases (Jordan, Landau, & Iyengar, 2000; Morgan, 2007).

Protein kinases are enzymes that catalyze a transfer one or multiple phosphoryl groups (PO_3^{2-}) from an ATP molecule to a specific amino acid in a substrate. Main kinases responsible for the cell cycle regulation are serine/threonine kinases, which phosphorylate OH group of either Ser or Thr residues of a protein molecule by exchanging H^+ from the OH group to the PO_3^{2-}. Protein kinases have a bi-lobate tertiary structure, with a smaller lobe

from the N-terminus and a larger one from the C-terminus. The ATP molecule used as a phosphate group donor is located in the cleft between two lobes, which is an active site of protein kinases (Morgan, 2007). Since ATP molecules in living cells are usually bound to the Mg^{2+} ions, there is a complex of ATP with Mg^{2+} inside an active site of protein kinases. By hydrogen and Mg^{2+} ion bonding with several amino-acid residues, an ATP molecule is held in a position oriented toward the exit of a cleft, near which a substrate binds larger lobe. One of the amino acids in the activation loop facilitates the transfer of phosphoryl group from ATP to the target phosphorylation site (phosphosite). Residue of this amino acid causes partial deprotonation of the OH group and subsequent nucleophilic attack by oxygen in the phosphoryl group (De Bondt et al., 1993; Smith, Ke, Guo, & Hengge, 2011).

Phosphorylation is a reversible modification. There is another group of enzymes, protein phosphatases, which are able to remove phosphate groups from the site by exchanging PO_3^{2-} to H^+ from water, bringing the amino acid to its initial state.

Protein phosphorylation is a common way of signal transduction in cell signaling pathways that may result in different outcomes. For example, phosphorylation of one sites of a protein can send it for degradation (Skowyra, Craig, Tyers, Elledge, & Harper, 1997), while other sites of the same protein, being phosphorylated, can activate or stabilize the protein (Bononi et al., 2011) and it results in different cell response.

1.3 Structure and function of Cdc28, *S. cerevisiae* central CDK

As phosphorylation usually have a significant impact on the protein's fate and especially important for the cell cycle progression, the specificity and the activity of protein kinases are tightly controlled. The most important components involved in the cell cycle progression, are cyclin-dependent kinases (CDKs), a family of serine/threonine protein kinases. By definition, all CDKs require association with a regulatory cyclin subunit for their enzymatic activation. Cyclins are a family of proteins with a common feature of binding and activating CDKs. They do not possess any enzymatic activities by themselves, but they do contain binding regions for different CDK's substrates and localization sequences, which define substrate specificity and subcellular locations of CDK-cyclin complex (Morgan, 2007).

In different organisms, number, size, and structure of CDKs may vary. In *S. cerevisiae*, common model organism to study the cell cycle, the single central CDK is Cdc28, encoded by *CDC28* gene. The Cdc28 protein (homolog of human Cdk1 and it is often called Cdk1 in yeast too) is 34 kDa and contains 298 amino acids (Saccharomyces Genome Database,

https://www.yeastgenome.org). Similar to other CDKs, Cdc28 need to be associated with a cyclin subunit for the activation (Morgan, 2007).

For its activation, Cdc28 requires an association with a cyclin subunit because cyclin binding will induce several essential structural modifications of Cdc28. Prior to association with a cyclin, a flexible loop in a larger lobe, known as the activation loop or T-loop, at inactivated state blocks access of a substrate to ATP and prevents substrate binding. In addition to that, amino acids in the Cdc28 active site are positioned in a way that ATP is wrongly oriented. After the cyclin binding, T-loop and ATP are brought into positions required for binding and phosphorylation of a substrate (Morgan, 2007).

There are nine different cyclins in budding yeast cells. Sequences of different cyclins are very diverse, except the region of about 100 amino acids in length called a cyclin box. It is required for common function of all cyclins: CDK binding and activation. Despite the sequence diversity, the structure of different cyclins possesses some common features. Part of their tertiary structure is relatively well-conserved: a core of two domains, known as cyclin fold, each consisting of five alpha-helixes. The first domain contains a cyclin box, the second one, although it repeats the arrangement of the first, has different sequence from cyclin to cyclin. Outside of the cyclin fold structure vary as well, as it contains regulatory and binding domains, which define substrate specificity of cyclin-Cdc28 complex. (Morgan, 2007).

Complete activation of Cdc28 and its proper functioning requires also phosphorylation of a specific Thr residue in T-loop. This phosphorylation is catalyzed by Cdk-activating kinase CAK1 and, along with cyclin binding, changes the T-loop position (De Bondt et al., 1993; Morgan, 2007; Russo, Jeffrey, & Pavletich, 1996).

In addition to cyclins and CAK1, Cdc28 acts in a complex also with Cks1 proteins. Cks1 protein binds a larger lobe of a CDK and contains a phosphate-binding pocket (Figure 2.). By binding both CDK and phosphorylated site of a substrate, Cks1 increases CDK affinity towards substrates with one of the sites phosphorylated, therefore promoting downstream multisite phosphorylation. (McGrath et al., 2013)

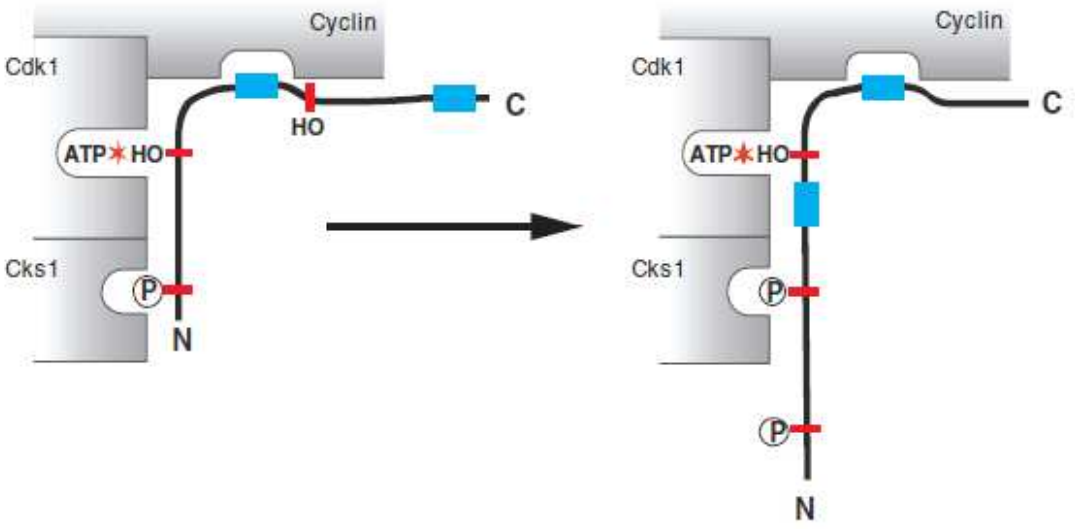

Figure 2. Structure of Cdk1-cyclin-Cks1 complex. By binding to phosphorylated site of the substrate, Cks1 increases CDK1 affinity towards this substrate, therefore, promoting downstream multisite phosphorylation. Blue boxes indicate cyclin docking motifs of the substrate sequence. (Figure was taken from the article *Multistep phosphorylation systems: tunable components of biological signaling circuits,* 2014 by Valk et al.)

1.4 The regulation of Cdc28 activity

Cdc28 is a single kinase regulator facilitating the whole cell cycle, and its activity has to oscillate, to allow Cdc28 to perform different tasks in every phase. For example, in late G1/early S phase Cdc28-cyclin complex activates transcription of genes, which code for parts of replication machinery, while later in S phase, Cdc28-cyclin complex initiates DNA replication (Epstein & Cross, 1992; Morgan, 2007).

There are several levels of the regulation of protein phosphorylation by Cdc28. The first one is a primary amino acid sequence of a substrate. Due to the presence of a hydrophobic pocket near its active site, Cdc28 has more affinity to phosphorylation sites with Pro at +1 position. Thus, so called full consensus site for Cdc28 phosphorylation is [S/T*]PX[K/R], where S/T* is phosphorylated amino acid, P is proline, X is any amino acid and K/R is basic amino acid lysine or arginine. At the same time, it can phosphorylate [S/T*]P site when it reaches higher activity threshold (Malumbres, 2014; Morgan, 2007).

Apart from the sequence motif, which is recognized by the Cdc28 itself, substrates have docking motifs for cyclin binding that allows tighter contact of substrate and cyclin-Cdc28 complex. (Kõivomägi et al., 2011; Morgan, 2007).

Enzymatic activity of Cdc28 depends on a cyclin which is currently bound to the kinase, as cyclin determines location and substrate specificity of a complex. However, as was shown in several experiments with deletions and mislocalization of cyclins, there is a partial overlap between cyclin functions (Edgington & Futcher, 2001; Fisher & Nurse, 1996). These observations resulted in emergence of two different models explaining differences in performances of Cdc28-cyclin complexes throughout the cell cycle. Quantitative model explains cell cycle flow by different thresholds of Cdc28 activity at the different cell cycle phases (it is low in S phase and high in M phase). The differences between cyclins and their role comes from different levels of their expression and timing (Stern & Nurse, 1996). According to qualitative model, cyclins change CDK1 affinity towards different substrate and affect its localization (Loog & Morgan, 2005; Miller & Cross, 2001).

Cyclins are differentially expressed at the different phases of the cell cycle: there are G1, G1/S, S and M cyclins. In *S. cerevisiae*, there are nine cyclins which form complexes with Cdc28: Cln3 is a G1, Cln1 and 2 are G1/S, Clb5 and 6 are S and Clb1, 2, 3 and 4 are M phase cyclins. Concentration of a cyclin usually peaks at a particular stage of the cell cycle, and during this stage probability of complex formation between this cyclin and Cdc28 is increased. (Morgan, 2007).

Inhibition of Cdc28 activity also contributes to the activity oscillations. CDK inhibitor proteins (CKIs), in *S. cerevisiae* represented by Sic1 and Far1 proteins, bind and inactivate cyclin-Cdc28 complexes. Sic1 inhibits S- or M-Cdc28 activity, and Far1 inhibits G1/S-Cdc28 in the presence of mating pheromones, causing the cell cycle freeze. Sic1 is marked for degradation by G1/S-Cdc28, which allows cell cycle to move forward(Morgan, 2007).

Cdc28-cyclin complexes trigger important transitions in the cell cycle. These transitions from one set of related cellular processes to another are called checkpoints, as they cannot be passed until all required intra- and extracellular signals are gained. Start checkpoint in late G1 triggers the transition to S phase, enter to M phase is controlled by G2/M checkpoint, and metaphase-to-anaphase transition point is called spindle assembly checkpoint. (Morgan, 2007).

1.5 Protein degradation

Abundance of the cyclin depends on the rates of several processes: gene expression, translation and protein degradation. Gene expression is controlled through transcriptional factors, molecules which influence RNA polymerase affinity to the promoter region. Transcription

factors themselves are activated as a consequence of cell cycle events happened in the previous stage. This contributes to the ordering of the cycle and ensures its progression.

The degradation rate of cyclins and other cell-cycle regulators is controlled by ubiquitin-dependent proteolysis. In this proteolytic pathway proteins have to be tagged by multiple molecules of small protein ubiquitin in order to be processed by huge protease complexes – proteasomes. Ubiquitin tagging is performed in three steps by three types of enzymes — E1, E2 and E3 corresponding to the number of catalyzed step. E1, ubiquitin activator, binds ubiquitin, and then interacts with ubiquitin-conjugating enzyme E2, which catalyze transfer of ubiquitin to E2 active site. Finally, E3, ubiquitin-protein ligase, mediate transfer of ubiquitin from E2 to target protein.

There are several ubiquitin-protein ligases, but two of them are especially important for cell cycle regulation. The first one is anaphase-promoting complex (APC), which facilitates metaphase-to-anaphase transition. The second one is a multisubunit enzyme complex called SCF, which is responsible for marking of Sic1, Far1 and G1/S cyclins in early S phase for degradation (Morgan, 2007).

SCF consists of following subunits: Cul1 (cullin), Rbx1 (domain where conjugate E2-ubiquitin binds), Skp1 with an associated F-box protein. F-box proteins bind target protein and therefore determine SCF specificity (Skowyra et al., 1997). F-box proteins involved in the degradation of G1 inhibitors and G1/S cyclins are Cdc4 and Grr1, respectively. However, there is some overlap in substrate specificity as Cdc4 can target some of G1/S cyclins. This difference is partially defined by localization of F-box proteins — SCF^{Cdc4} can be found mostly in nucleus, while SCF^{Grr1} in cytoplasm, where G1/S cyclin-Cdc28 complexes are located (Landry, Doyle, Toczyski, & Benanti, 2012).

Grr1 substrates, namely Cln1-3, contain destruction PEST signal (sequence rich with proline (P), glutamic acid (E), serine (S), and threonine (T)) with three phosphosites (Lanker, Valdivieso, & Wittenberg, 1996; Rogers, Wells, & Rechsteiner, 1986). PEST sequence in G1/S cyclins is part of the degron (region, which is known to shorten protein half-life). Degrons which are performing their function only after phosphorylation of certain sites inside their amino acid sequence are called phosphodegrons (Holt, 2012).

G1 and G1/S cyclins in *S. cerevisiae* possess phosphodegrons sequences with PEST regions that makes them highly unstable and ensures oscillations in their expression levels during cell cycle (Quilis & Igual, 2017).

2 THE AIMS OF THE THESIS

- To analyze the efficiency of Cln2-based multisite phosphorylation degron tags for the regulation of EGFP fluorescence levels.
- Create *S. cerevisiae* strains with integrated EGFP under control of inducible pLexAt promoter for further tagging with multisite phosphorylation degrons.

3 EXPERIMENTAL PART

3.1 MATERIALS AND METHODS

3.1.1 Media

Following media were used for **bacterial** growth and selection:

LB (lysogeny broth) media: 10 g/L tryptone (Formedium), 5 g/L microgranulated yeast extract (Formedium), 10 g/L NaCl (Chempur). To prepare agar plates with ampicillin, 15 g/L of bacto agar (Formedium) and 100 mg/L of ampicillin were added to the mixture.

Following media were used for **yeast** growth and selection:

YPD (Yeast Extract–Peptone–Dextrose) media: 20 g/L peptone (Formedium), 10 g/L micro granulated yeast extract (Formedium), 20 g/L D-Glucose anhydrous. To prepare agar plates, 15 g/L bacto agar (Formedium) were added to the mixture.

CSM (Complete Supplement Mixture): 20 g/L peptone (Formedium), 10 g/L CSM (Formedium), 20 g/L D-Glucose anhydrous. To prepare agar plates with G-418, 15 g/L bacto agar (Formedium) and 200 mg/L G-418 were added to the mixture. To prepare TRP dropout CSM plates, CSM, -TRP (Formedium) was used instead of the complete CSM.

3.1.2 Bacterial strains

Competent DH5ά *E. coli* cells were prepared in our laboratory.

3.1.3 Yeast strains

S. cerevisiae strains used in the course of current project are listed in the Table 1.

Table 1. Yeast strains implemented in this project.

Strain	**Genotype***	**Short description/Trasformation**	**Source/Reference**
IMX672	CEN.PK: *MATa ura3-52 trp1-289 leu2-3,112 his3Δ can1Δ::cas9-natNT2*	Initial strain with constitutive Cas9 expression. All other strains were generated by trasformation of IMX672	Euroscarf (AN Y40595)
NS229	*leu2*::*pADH1-EGFP-NES*	Constitutive expression of EGFP fused with NES/CRISPR-Cas9 mediated integration of a PCR fragment	Our lab

NS301	*trp1::pACT1-LexA-ER-B112-Trp1*	LexA-ER-B112 transcription factor coding sequence under ACT1 promoter in TRP locus/Plasmid integration into the genome	This study
NS325	*leu2::pLexAt-EGFP-NES*	On a base of NS229/integration of a PCR fragment through homologous recombination (Janke et al., 2004)	
NS322	*leu2::pLexAt-EGFP-GSA-LP-NES*	On a base of NS229/integration of a PCR fragment through homologous recombination (Janke et al., 2004)	This study
NS230 NS231	*leu2::pADH1-EGFP-NES-Cln2_5WT-LP-Leu2*		
NS232 NS233	*leu2::pADH1-EGFP-NES-Cln2_5A-LP-Leu2*		
NS234 NS235 NS376 NS377 NS378	*leu2::pADH1-EGFP-NES-Cln2_5WT-Leu2*		
NS236 NS279 NS280	*leu2::pADH1-EGFP-NES-Cln2_4WT-Leu2*		
NS237 NS238 NS239	*leu2::pADH1-EGFP-NES-Cln2_4WT-LP-Leu2*		
NS262 NS288 NS289 NS290 NS292	*leu2::pADH1-EGFP-NES-Cln2_4A-Leu2*		
NS263 NS293 NS294	*leu2::pADH1-EGFP-NES-Cln2_4A-LP-Leu2*		

* All the strains created in this work are based on IMX672 and common genetic background is not described.

pADH1 – sequence located 716 bp upstream of *ADH1* gene start codon and referred as ADH1 promoter. **GSA** is a flexible protein linker. Four variants of Cln2 degron: **Cln2_5WT** – 373-445 of Cln2 amino acid sequence, **Cln2_4WT** – 393-445 of Cln2 amino acid sequence, **Cln2_5A** and **Cln2_4A** are similar sequences with 5 or 4 phosphorylation sites mutated to alanine, respectively. **LP** is Cln2 docking site from *S. cerevisiae* Sic1 protein (131-145 amino acid sequence). Here **NES** stands for nuclear export signal, DNA regions coding for short peptide sequence targeting protein for transportation from nucleus to cytosol.

Strains, constructed in this study, were generated using lithium acetate mediated transformations (see the protocol below).

3.1.4 Plasmids

All the plasmids used in this study ate listed in the Table 2.

Table 2. Plasmids, used in this research.

Plasmid	**Yeast marker gene/mode of action**	**Short description**	**Primers for cloning**	**Source/ Reference**
pNS106	*Trp*/Complementation of mutated *trp* gene	Contains LexA-ER-B112 synthetic transcription factor under ACT1 promoter	-	Our lab
pNS53	*Ura*/Complementation of mutated *ura* gene	LexA-target promoter, containing 4 LexA protein binding sites (referred below as pLexAt)	-	
pNS77	*KanMX*/resistance to G-418 antibiotic	for CRISPR/Cas9 gRNA1 for LEU2	-	
pRK47		for CRISPR/Cas9 gRNA2 for LEU2	-	
pNS86		Cln2_5WT-LP	4719/4721	This study
pNS87		Cln2_5A-LP	4719/4721	
pNS97		Cln2_5WT	4719/4722	
pNS98		Cln2_4WT	4724/4722	
pNS99		Cln2_4WT-LP	4724/4721	
pNS112		Cln2_5A	4719/4722	

AmpR was present as bacterial selection marker in all plasmids.

To construct plasmids pNS86, pNS87, pNS97, pNS98 and pNS99, two synthetic DNA were ordered: *BamHI*-Cln2_5WT-LP-*SgsI* and *BamHI*-Cln2_5A-LP-*SgsI* (351 bp each). Both synthetic DNAs contained *BamHI* restriction site at 5'-end and *SgsI* restriction site at 3'-end for cloning. pFA6a vector with *KanMX* selection marker gene (confers yeast resistance to G-418 antibiotic) was used as a backbone. After verification by sequencing, constructed pNS86, pNS87, pNS97, pNS98 and pNS99 plasmids were used as a template for PCR to amplify target sequence alone with *KanMX* selection marker gene. Primers were designed according to Janke et al., 2004. All *in silico* cloning, primer design and sequencing results analysis was performed with Benchling platform (www.benchling.com).

pNS86 and pNS87 were constructed by direct cloning of synthetic fragments into pFA6a-KanMX plasmid. Prior ligation, both synthetic DNA and backbone pFA6A-KanMX vector were restricted with *BamHI/SgsI* restriction enzymes. To construct pNS97, pNS98, pNS99 and pNS112, fragments for cloning were first PCR amplified, ran on a gel and purified, restricted and then ligated into pFA6A-KanMX vector. Primers and template used for PCR are listed in the Table 2.

For the restriction, 2 µL of 300 ng/µL of vector DNA and 10 µL of 20 ng/µL of amplified fragments or synthetic DNA were incubated with *BamHI/SgsI* restriction enzymes. Restriction reactions were performed in 20 µL mixture with 0.5 µL of *BamHI*, 0.5 µL of *SgsI*, 2 µL of 10x Fast Digest buffer, and up to 20 µL Milli-Q. In case of plasmid DNA restriction reaction, 1 µL of FastAP Thermosensitive Alkaline Phosphatase (1 U/µL) was used. All the reagents used for restriction were from Thermo Fisher Scientific. Reactions were incubated at 37° C for 30 minutes and then ran on a gel (look below for gel electrophoresis protocol), and bands of expected size were cut out and purified.

After purification from the gel using GEL/PCR purification kit (Favorgen) according to the protocol provided by the manufacturer, DNA concentration was measured using Nanodrop (Thermo Fisher). Amount of vector and insert, used for the reaction were calculated according to the following formula using 3:1 insert/vector molar ratio:

$$Volume_{insert} = \frac{\frac{mass_{vector} \times length_{insert}}{length_{vector}} \times insert\ to\ vector\ ratio}{mass\ concentration_{insert}}$$

After that vector and insert were ligated in 20 µl mixture that contained 1 µl of T4 DNA ligase (Thermo Fisher Scientific, 5 U/µl), 2 µl of 10x T4 ligation buffer, vector and insert according to calculations and Milli-Q to 20 µL. Ligations were incubated overnight at 18° C.

After overnight incubation, ligation mixture was transformed into competent DH5ά cells using heat shock. Bacterial cells were taken from -80° C, left for 10 minutes on ice to melt, then 50 µl of cells were mixed with 2 µl of ligation mixture and left on ice for additional 30 minutes. After incubation, cells were subjected to 42° C heat shock for 1.5 minutes, followed by 2 minutes of chilling cells on ice. After that, cells were diluted with 500 µl of LB media and incubated for an hour in 220 rpm shaker at 37° C, then centrifuged for 1 minute at 6000 rpm, resuspended at 200 µl of LB, plated on LB agar plates with ampicillin and incubated at 37° C for 12-15 hours.

When colonies appeared, liquid LB cultures of 3 mL were inoculated with single colonies and incubated in shaker at 37° C 220 rpm for 12-15 hours. Afterwards plasmids were extracted from cells using Favorgen plasmid extraction mini kit according to the protocol provided by manufacturer. Correct plasmids were found by sequencing with primer 2736.

3.1.5 Construction of PCR-amplified cassettes for chromosomal integration

PCR primers used for amplification of degron sequences together with *KanMX* contained overhangs at their 5'-ends: 40 bp overhang for the NES region and 40 bp overhang for LEU2 locus in forward and reverse primer, respectively. Amplified fragments were subsequently used for homologous recombination-mediated transformation of the strain NS229. PCR protocols can be found in the PCR section (3.1.8).

Table 3. PCR amplified cassettes with template plasmids and primers, used for the amplification.

Fragments amplified	**Template plasmid**	**Primers used**
NES-Cln2_5WT-LP-*KanMX*	pNS86	4731/4826
NES-Cln2_5A-LP-*KanMX*	pNS87	4723/4826
NES-Cln2_5WT-*KanMX*	pNS97	4731/4826
NES-Cln2_4WT-*KanMX*	pNS98	4732/4826
NES-Cln2_4WT-LP-*KanMX*	pNS99	4732/4826
NES-Cln2_4A-LP-*KanMX*	pNS87	4733/4826
NES-Cln2_5A-*KanMX*	pNS112	4723/4826
NES-Cln2_4A-*KanMX*	pNS112	4733/4826

3.1.6 Genomic DNA extraction

In order to amplify EGFP-NES fragment, genomic DNA of pADH1-EGFP-NES containing strain NS229 was used. For its extraction, protocol by Lõoke, Kristjuhan, & Kristjuhan (2011) was used.

NS229 was plated on a YPD agar plate from the stock and grown overnight at 37° C. After that, single yeast colony was picked from the YPD plate. Cells were suspended in a lysis solution (100 µL 200 mM lithium acetate (LiAc) with 1% SDS) and briefly vortexed. Cell suspension was incubated for 10 min at room temperature. After that 300 µL of 96% ethanol was added, and tube was vortexed briefly. DNA was collected by centrifugation at 13400 rpm for 3 min. Supernatant was removed followed by addition to the pellet of 500 µL of

70% ethanol. Tube was briefly turned up-side-down for DNA washing. After that tube was centrifuged at 13400 rpm for 1 min. Supernatant was removed by pipetting, and tube was left opened on a bench for 2 minutes for drying. Pellet was suspended in 100 µL water, and after that debris was removed by centrifugation at 13400 rpm for 1 min. Afterwards cleared supernatant was transferred to a clean tube (contains soluble part of DNA and RNA).

3.1.7 Generation of LEU2-pLexAt-EGFP-NES-LEU2 and LEU2-pLexAt-EGFP-GSA-LP-NES-LEU2 by overlap-extension PCR

Fragments further used for overlap extension PCR were amplified from pNS53 (pLexAt) and NS229 genomic DNA (EGFP). Fragments and primers used for overlap-extension PCR are listed in the Table 4.

Table 4. Construction of fragments for transformations by series of PCR reactions.

N	PCR template	Primers	Product	Brief description
1.	pNS53	5028/5030	*pLexAt*	pLexAt promoter with *leu2* (60 bp homology) and EGFP (16 bp) regions from 5'- and 3'-end, respectively
2.	NS229 genomic DNA	5031/5035	*EGFP-NES*	EGFP-NES fragment with pLexAt (23 bp) and *leu2* locus (29 bp) regions from 5'- and 3'-end, respectively
3.	NS229 genomic DNA	5031/5032	*EGFP-GSA*	EGFP-GSA partial fragment with pLexAt (23 bp) sequences from 5'- end and with GSA linker from its 3'-end
4.	*pLexAt* + *EGFP-NES*	No primers at first three cycles, then 4699/5036	*leu2-pLexAt-EGFP-NES-leu2*	The whole fragment for yeast transformation with 60 bp homology regions for *leu2* locus at 5'- and 3'-ends
5.	*pLexAt* + *EGFP-GSA*	No primers at first three cycles, then 4699/5033	*leu2-pLexAt-EGFP-GSA-LP*	pLexAt-EGFP-GSA fragment with *leu2* (60 bp) and LP (40 bp) regions from 5'- and 3'-end, respectively
6.	*leu2-pLexAt-*	4699/5034,	*leu2-pLexAt-EGFP-GSA-LP-NES-leu2*	The whole fragment for transformation into *leu2* locus with 60 bp homology regions at

	EGFP-GSA-LP	4699/5035, 4699/5036		5' and 3' ends. This fragments were obtained by three sequential PCR reactions. Primers used to PCR amplify this fragment had overhangs at their 5-'ends for step-by-step addition of LP, NES, and leu2 sequences

3.1.8 PCR

To amplify fragments from NS229 genomic DNA or plasmids using Phusion polymerase, following protocol was used for 50 µL reactions: 10 µL of 5xHF buffer (Thermo Fisher Scientific), 1 µL of 10 µM forward primer, 1 µL of 10 µM reverse primer, 1 µL of 5 ng/µl template DNA, 0.4 µL of 25 mM dNTPs, 0.5 µL of Phusion polymerase made in our lab (2 U/µL), up to 50 MilliQ water.

For yeast colony PCR, first step was to lyse yeast cells and get genomic DNA. To do that, colonies from the agar plate were suspended in 20 µL of 20 mM NaOH and boiled for 10 minutes at 100° C. PCR reactions themselves were mixed in 20 µL using 1 µL of extracted DNA, 2 µL of 10x Direct Load PCR buffer (Solis BioDyne), 1 µL of 50 mM MgCl, 0.25 µL of 25 mM dNTPs, 1 µL of 10 µM forward primer, 1 µL of 10 µM reverse primer, 0.3 µL of Taq polymerase made in our lab (5 U/µL) and 13.45 µL of Milli-Q.

Table 4.

PCR step	**Temparature**	**Time**	**Number of cycles**
Taq polymerase, yeast colony PCR			
Initial denaturation	95° C	5 minutes	
Denaturation	95° C	30 seconds	44
Annealing	X*° C	15 seconds	
Elongation	72° C	1 minute per kilobase	
Final elongation	72° C	5 minutes	
Final hold	15° C	until program is stopped manually	
Phusion polymerase, amplification from plasmids and yeast genomic DNA			
Initial denaturation	98° C	5 minutes	
Denaturation	98° C	30 seconds	30
Annealing	X*° C	15 seconds	
Elongation	72° C	1 minute per kilobase	
Final elongation	72° C	5 minutes	
Final hold	15° C	until program is stopped manually	
Phusion polymerase, overlap extension PCR			

Initialization	95° C	5 minutes	
Denaturation	95° C	30 seconds	3
Annealing	X*° C	15 seconds	
Elongation	72° C	1 minute per kilobase	
Denaturation	95° C	30 seconds	30
Annealing	X*° C	15 seconds	
Elongation	72° C	1 minute per kilobase	
Final elongation	72° C	5 minutes	
Final hold	15° C	until program is stopped manually	

*To calculate annealing temperature for Phusion reactions, 3° C were added to the lower melting temperature from the pair of primers. To calculate annealing temperature for Taq reactions, the higher temperature of two is used.

** Initial melting and first 3 cycles run without primers, after that primers are added.

3.1.9 Primers

Table 5.

Number in lab database	Sequence	Tm
5028	ATT ATG GAG AAA AAC TGT GGA GGA AAG GCC GCA ATA ATA TAT AAA C	75.4
5030	CTT CTC CTT TAC TCA TAT AGA AGT ATA GTA ATT TAT GCT G	79.7
5031	AGC ATA AAT TAC TAT ACT TCT ATA TGA GTA AAG GAG AAG AAC	71.3
5035	GGA GAT GAT ATC ACC AAA CAT GTT GCT GGT TAT TTG TTA ATG TCC AGA CCA GCC AAT TTC	81.6
5029	CCT TGG ATA AAG CTA ATG TTT TGG CCT CTT CAA GAT TAT GGA GAA AAA CTG TGG	79.7
5036	CCA AGG AAC CTG GGA TAA CGG AGG CTT CAT CGG AGA TGA TAT CAC CAA AC	83
4699	CCT TGG ATA AAG CTA ATG TTT TGG C	62.5
4700	CCA AGG AAC CTG GGA TAA CGG	63.2
5032	TTC GAA TTC GCC AGA ACC AGC AGC GGA GCC AGC GGA TCC TTT GTA TAG TTC ATC CAT GCC	87
5033	ATG TAG GTC TGC TGG GGG GAA GAA GAA CTT CTC CTT CTT CTT CGA ATT CGC CAG AAC CAG	85.7
5034	GTC CAG ACC AGC CAA TTT CAA AGC CAA TTC ATT CAT AGA TGT AGG TCT GCT GGG GGG AAG	84.9

4719	CAG AGG ATC CAT TTC GAG AAA GCT TAC CAT ATC AAC CCC	76,7
4724	CAG AGG ATC CTC CAT TCC TTC GCC CGC TTC	76,2
4721	TGT ATG GCG CGC CTC AAG ATG TAG GTC TGC TGG GGG GAA GAA GAA CTT CTC CTT CTT C	86,5
4722	TGT ATG GCG CGC CTC AAC ATA TAC TGT TTG ACT GCT GCT GAC CAA ATT G	81,4
4731	TGAATTGGCTTTGAAATTGGCTGGTCTG-GACATTAACAAAATTTCGAGAAAGCTTACCATATC	81,1
4732	TGA ATT GGC TTT GAA ATT GGC TGG TCT GGA CAT TAA CAA ATC CAT TCC TTC GCC CGC TTC	83,7
4733	TGA ATT GGC TTT GAA ATT GGC TGG TCT GGA CAT TAA CAA ATC CAT TCC TGC ACC CGC TTC	83,7
4723	CAG AGG ATC CAT TTC GAG AAA GCT TAC CAT ATC AGC CCC	78
4826	GAG GCT TCA TCG GAG ATG ATA TCA CCA AAC ATG TTG CTG GGC AGA TCC GCG GCC GCA TAG	87,8
2750	GCA GGA AGA GGA AGA AGG GT	60,5
2751	TCA AAT CCA CAA AGC CTG GC	58,4
2736	GCG CAC GTC AAG ACT GTC AAG G	65.8
4734	CCC TTT CGA AAG ATC CCA ACG	61.2
4762	CAA TTA ATA CAT AAC CTT ATG TAT CAT AC	60.5
4699	CCT TGG ATA AAG CTA ATG TTT TGG C	62.5
4700	CCA AGG AAC CTG GGA TAA CGG	63.2

3.1.10 Agarose gel electrophoresis

To check whether PCR or restriction succeeded and to separate DNA fragments of different lengths, agarose gel electrophoresis was performed. DNA samples were mixed with 10x loading dye (Thermo Fisher) and loaded to 1% or 2% (depending on the length of target fragment) agarose gel in TAE buffer (40 mM Tris-acetate with pH 8.3, 1 mM EDTA, 1%-2% agarose, 5 µL/L Atlas ClearSight DNA Stain, BioAtlas). Gels were placed into electrophoresis chamber filled with 1x TAE buffer (40 mM Tris-acetate with pH 8.3, 1 mM EDTA). Gels were running for 25 minutes at 180V and checked under UV light (280 nm). If necessary, DNA bands were cut from the gel and purified using GEL/PCR purification kit (Favorgen) according to the protocol provided by manufacturer.

3.1.11 Lithium acetate-mediated yeast transformation

For transformations, NS229 strain was taken from – 80 ºC stock (see below the protocol) using sterile wooden stick and was spread on YPD agar plate followed with overnight incubation at 30 ºC. Next day, NS229 was inoculated from the plate to 50 mL of YPD and grown in a 150 ml flask at 160 rpm shaker at 30° C. When optical density reached 0.6-0.8 at 600 nm (approx. after 12-16 hours of incubation), culture was transferred into 50 mL falcon tube and centrifuged for 2 min at 3600 rpm. Supernatant was removed and the cell pellet was resuspended in 1 mL of sterile 100mM solution of lithium acetate (LiAc, source) in TE buffer (10 mM Tris, 1mM EDTA), transferred into 1.5 mL eppendorf tube and centrifuged again at 3600 rpm for 60 seconds. The supernatant was removed and cell were resuspended in 100mM LiAc in TE buffer that exceeded cell volume approximately two times. Yeast competent cells were incubated at room-temperature for 10 minutes.

In parallel, salmon sperm DNA was boiled for 10 minutes at 100° C and was put on ice to chill immediately after to keep DNA single-stranded (SS-DNA). SS-DNA was kept one ice through the whole transformation procedure.

In a separate 1.5 mL tube, DNA to be transformed and 10 µL of SS-DNA were mixed. Then 100 µl of yeast competent cells were added to the tube and gently mixed by pipetting. At the next step, 700 ml of sterile PEG/LiAc solution (40% PEG 3350 + 100 mM LiAc in 1x TE) and 48 µL of DMSO were added to the cells. Solution was mixed by gentle pipetting and incubated at 42° C for 40 min. After the incubation, cells were chilled on ice for 2 minutes, and centrifuged for 1 minute at 6000 rpm. The supernatant was removed, and pellet was resuspended in 1 mL of sterile 1xTE buffer for washing and centrifuged for 2 minutes at 2000 rpm. The supernatant was removed and cells were resuspended in 200 µL of sterile 1xTE buffer and plated on the agar plates.

In order to integrate LexA-ER-B112 transcription factor into Trp locus in *S. cerevisiae* genome, pNS106 vector was used. Plasmid was restricted with *Eco81I*: 7 µL (1.4 µg) of the plasmid were mixed with 10.5 µl of Milli-Q water, 2 µL of FastDigest 10x buffer and 0.5 µL of *Eco81I* FastDigest restriction enzyme (ThermoFisher Scientific) and incubated for 30 minutes at 37° C. After that restriction enzyme was inactivated by incubation at 80° C for 10 minutes. 7 µL of restriction mixture were used to transform into NS228 yeast strain. After the transformation, cells were plated on –TRP CSM agar plates for selection of cells with insert and incubated upside down at 30°C until colonies appeared in 2 days. In order to confirm insertions, yeast colony PCR with primers 2750 and 2751 were performed. Bands of

expected size (683 bp) were cut out and sent for sequencing. Glycerol stocks were prepared for correct strains. The resulted strain was named as NS301.

The second set of yeast transformations to generate strains expressing GFP under inducible pLexAt promoter was performed using CRISRP-Cas9 method as it is described in (Mans et al., 2015). In this study, strain with constant Cas9 and LexA-ER-B112 transcription factor expression, NS301, was used for transformation. In addition to target DNA, two pMEL13 plasmids with gRNAs were co-transformed into NS301 strain. The plasmids were required to guide Cas9 nuclease to *leu2* locus inside yeast genome. pLexAt-EGFP-NES and pLexAt-EGFP-GSA-LP-NES DNA fragments used for transformation were generated by overlap-extension PCR (see the description above). The amplified fragments contained at the 5'- and 3'-ends 60 bp overhangs for integration of the fragments into *leu2* by homologous recombination. Overhangs corresponded to 566-625 and 730-789 bp of leu2 locus, for 5'- and 3'-overhangs, respectively. 200 ng of PCR product were co-transformed with 200 ng of each p-MEL13-gRNA plasmid. After transformation, cells were spread on non-selective CSM plates and left at room temperature overnight. The next day, replica plating on G-418-containing CSM plates was performed. G-418 antibiotic was used for selection of cells containing pMEL13 plasmids, while selection of target DNA fragments were not required. After appearance of colonies (2 days after transformation), they were transferred to the new G-418 plates. Separate colonies were screened by FACS for GFP fluorescence.

The third set of transformations to integrate degrons was performed using NS229 as initial strain. Before transformations, fragments were tagged with NES region from 5' end and *KanMX* region from 3' end by PCR. 200 ng of PCR products were transformed. After the transformations cells were plated to YPD plates and kept at room temperature overnight. The next day replica plating to G-418 containing CSM plates was performed. After colonies appeared, they were transferred to the new G-418 plates. Selection of colonies with insert was performed by yeast colony PCR using primers 4734/2736. Band of size (508 bp). Glycerol stocks were prepared for correct strains.

3.1.12 Preparing yeast glycerol stocks

To prepare stocks of yeast strains, freshly grown cultures from YPD plate were inoculated to 1 mL of 15% glycerol, vortexed and put in -80° C.

3.1.13 OD measurement

For OD measurements, spectrophotometer Ultrospec 10 from Amersham Biosciences was used. Measurements were taken at 600 nm wavelength (a suspension containing $1*10^6$ cells ml^{-1} will give an OD600 of 0.1). For blank measurements, 1 mL of growth medium was used.

3.1.14 DNA concentration measurements

NanoDrop 1000 Spectrophotometer was used to measure DNA concentrations. NanoDrop 1000 3.8.1 software was used.

For each measurement, 1.5 µL of DNA solution were used. For blank measurements, 1.5 µL of elution buffer from the kit used to purify DNA was used.

3.1.15 Flow cytometry

To perform flow cytometry, cells were inoculated in liquid CSM media, containing 2% of glucose, and incubated for 2 hours in a shaker at 160 rpm, 30° C until cultures reached OD600 = 0.4. After that, 200 µL of cultures were transferred to 96-round-wells plate and analyzed using Attune NxT flow cytometer. To measure EGFP fluorescence, 470 nm exitation laser (100mW) was used with BL1 emission filter (530/30) at 380V. For data recording, Attune NxT software was used. For each sample, we recorded 100 000 events with the sample pessure of 12.5 µL/min.

To analyze the data, statistical functions of Microsoft Excel were used.

3.1.16 Time-lapse quantitative microscopy

Single cell microscopy was performed using Zeiss Axio Observer Z1 microscope and Cellasic microfluidic device. Y04C plates were used. Following settings were used for the experiments: 25 ms exposure time phase contrast for phase contrast channel, 15 ms exposure time for EGFP channel. Colibri LED modules was used as a source of 470 nm wavelength at 25% intensity. 12 positions were collected. Definite focus was used in order to keep the same focusing for all positions. To analyze data, images were processed by custom MATLAB software published by Doncic et al. (2013) and modified in our lab.

3.2 RESULTS

3.2.1 LP motif fused to the degron increased the stability of the EGFP

In the current work, the effect of Cln2-based degron tag sequences on protein degradation was analyzed. Several degrons were constructed and inserted to the strain with constant expression of EGFP, tagging the fluorescent protein from the C-terminus (Figure 3C). The effect of addition of Cln2 docking motif from yeast Sic1 protein was analyzed. Some of the tags included LP docking motif, a region of Sic1 (Figure 3B) that was shown to increase the rate of Cln2-dependent phosphorylation of Sic1 in comparison to proteins with deleted or mutated motif. It was proposed as a docking motif for Cln2 cyclin, which promotes Sic1 phosphorylation by binding it. Length of LP motif may vary in different researches, but VLLPP sequence probably has the biggest influence on the cyclin binding (Kõivomägi et al., 2011). In this study, sequence of 15 aa (5 aa before and 5 after VLLPP motif from Sic1) was used, as surrounding of docking motif may influence the efficiency of binding as well.

Cln2 degron was previously used as a protein-destabilizing tag in multiple variations. PEST sequence plays major role in the effect of Cln2 N-terminus on the rate of cyclin's degradation, but phosphosites T430 and S518 (Figure 3A) enhance it (Berset et al., 2002; Lanker et al., 1996; Mateus & Avery, 2000; Salama, Hendricks, & Thorner, 1994). In this study, regions from 373 to 445 (5WT) and from 393 to 445 (4WT) were used as degrons, truncating S518 and, in case of the shorter degron, T381. As one of the aims of this study was to develop phosphodegron tags which can be used for metabolic enzymes control, degron should be as short as possible to avoid interference with enzyme functioning. EGFP was used as a reporter protein to visualize influence of a degron on the concentration levels of protein it is fused to.

Eight different tags were constructed, and seven of them were successfully integrated in leu2 locus of *S. cerevisiae*. First four degrons are either Cln2_5WT or Cln2_4WT with or without LP docking motif. Another four constructs repeat them, but all of their phosphorylation sites are mutated to alanine (Figure 3C). Mutations of serine or threonine residues in phosphosites to alanine prevent phosphorylation. Alanine mutant strain Cln2_4A was used as a reference strain. GFP_NES strain was used as an additional control to test an effect of tagging Transformation of 5A failed multiple times, and fluorescence signal from 4A strain was used as a reference.

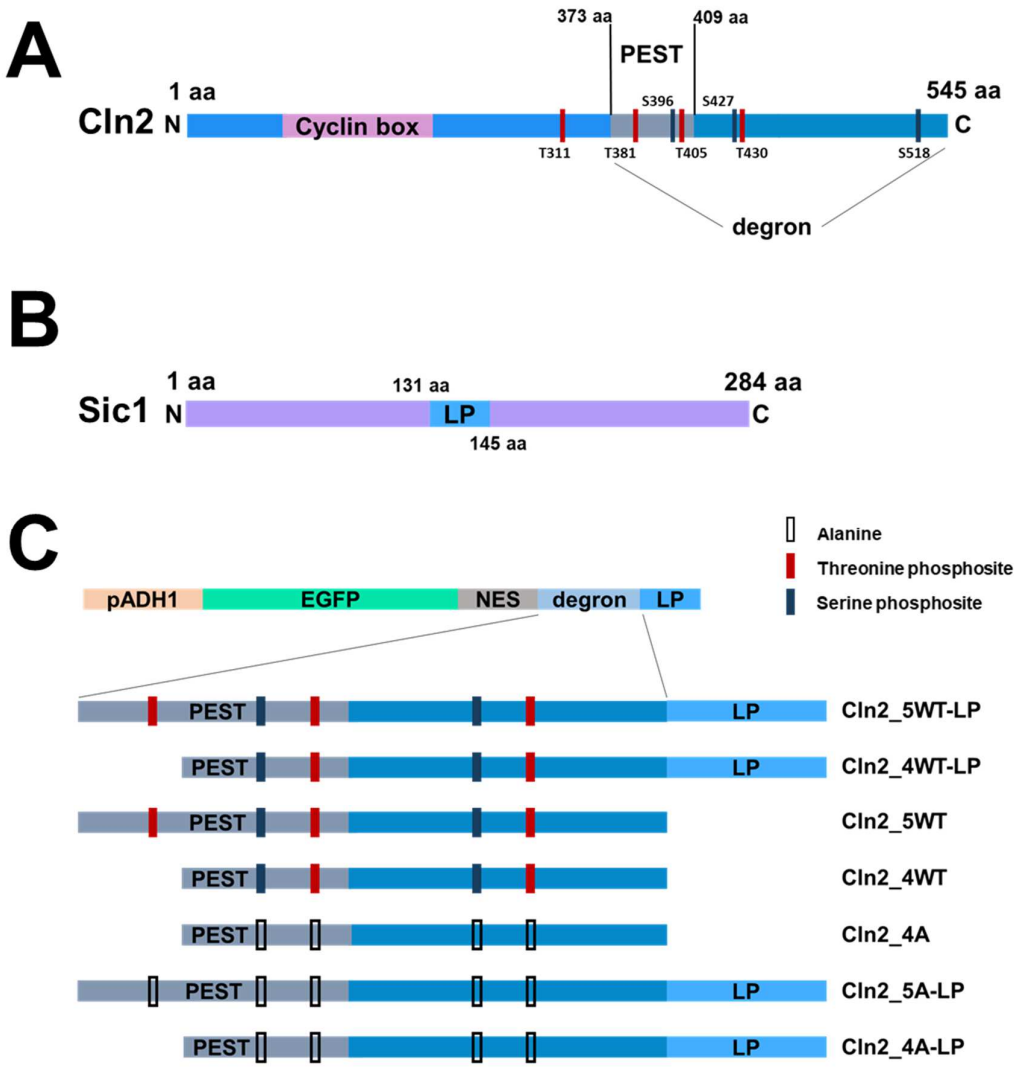

Figure 3. Schematic structures of cyclin Cln2 (A), inhibitor protein Sic1 (B) and constructs, designed in this research (C). (**A**) Primary structure of **Cln2** cyclin. Elements, marked on the figure: **cyclin box**, Cdc28 binding domain; **T311**, phosphosite involved in Cln2-Cdc28 binding; **PEST sequence**, region reach with proline (P), glutamic acid (E), serine (S), and threonine (T). PEST was observed to shorten half-life of proteins containing it; **degron,** region showed to be responsible for protein tagging for proteolysis; **phosphosites,** threonines or serines showed to be targets for Cdc28 phosphorylation. (**B**) Primary structure of **Sic1** inhibitor. **LP** docking motif, a site at which Cln2 binds to Sic1, increasing probability of Sic1 being phosphorylated. (**C**) EGFP-NES sequence under the ADH1 promoter and variety of degradation promoting fragments, constructed in this research.

Fluorescence signal of positive control strain, where EGFP under the same promoter was expressed without any degron tags or docking motifs, showed 210% of 4A fluorescence; 5WT-LP showed 110%, showing the most similar result to 4A; 4WT showed 80% and 5WT

70%, being the lowest results of all; 4A-LP, 4WT-LP, 5A-LP showed 140%, 150% and 130% of 4A fluorescence, respectively.

Results of fluorescence measurements with flow cytometry are presented in Figure 4.

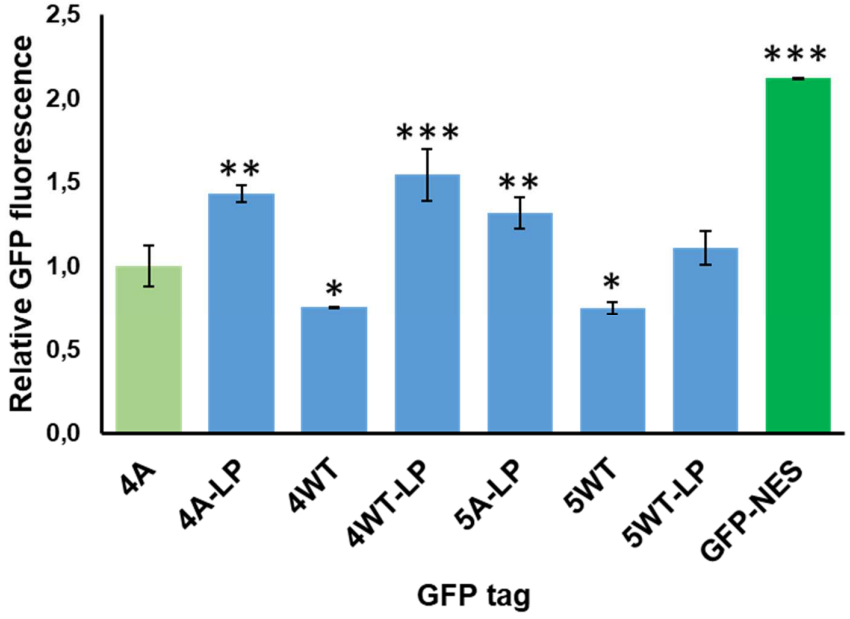

Figure 4. Relative fluorescence of strains with different tags in relation to 4A containing strain. Ratios were calculated by dividing mean of the fluorescence signal showed by strain of interest by mean of 4A. From 2 to 3 strains with the same degron tag inserted were used for the measurements, and 2 technical replicas were made from every strain.

* — $p < 0.05$

** — $p < 0.01$

*** — $p < 0.001$

3.2.2 Strains with EGFP-NES and EGFP-GSA-LP-NES constructs under LexAt inducible promoter

For further developing of systems with combined post-translational and transcriptional regulations, two strains with integrated LexA-ER-B112 system (Ottoz, Rudolf, & Stelling, 2014) were created. LexA-ER-B112 is a transcriptional factor, where LexA is DNA binding domain of bacterial inhibitor protein LexA, ER is β-estradiol receptor and B112 is a transcription activator. This transcriptional factor binds to promoter with embedded LexA binding sites and activates transcription only after the conformational change caused by β-estradiol binding to ER domain.

GSA is a flexible linker created by Waldo, Standish, Berendzen, & Terwilliger (1999) to allow correct folding for EGFP fused to any heterologous sequence. It was added to the construct to allow for proper EGFP folding.

In order to check colonies, appeared after the transformation, for EGFP expression after the induction, flow cytometry was used. Three positive colonies were detected for both of the constructs. For colonies with integrated GSA-LP containing constructs, fluorescence level after induction was around 430% from the background fluorescence, and for colonies with EGFP-NES construct fluorescence was increased from around 230% to 335% depending on a colony.

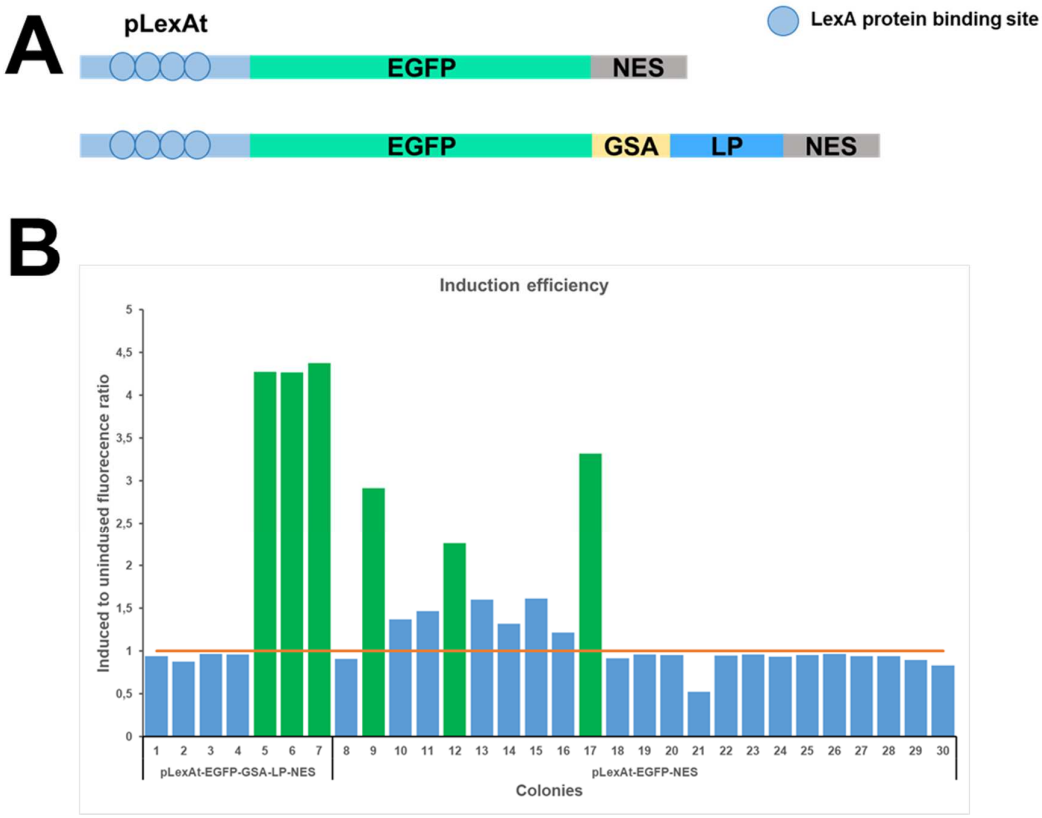

Figure 5. EGFP-NES and EGFP-GSA-LP-NES constructs under the inducible LexAt promoter. (**A**) Schematic DNA sequence of constructs. (**B**) Ratios of signals from induced cultures over non-induced ones. Positive colonies are colored in green. From the colonies transformed with pLexAt-EGFP-NES constructs, 9, 12 and 17 showed EGFP expression after the induction with β-estradiol, and from the pLexAt-EGFP-GSA-LP-NES 5, 6 and 7.

3.2.3 EGFP tagged with GSA-LP showed enhanced fluorescence signal

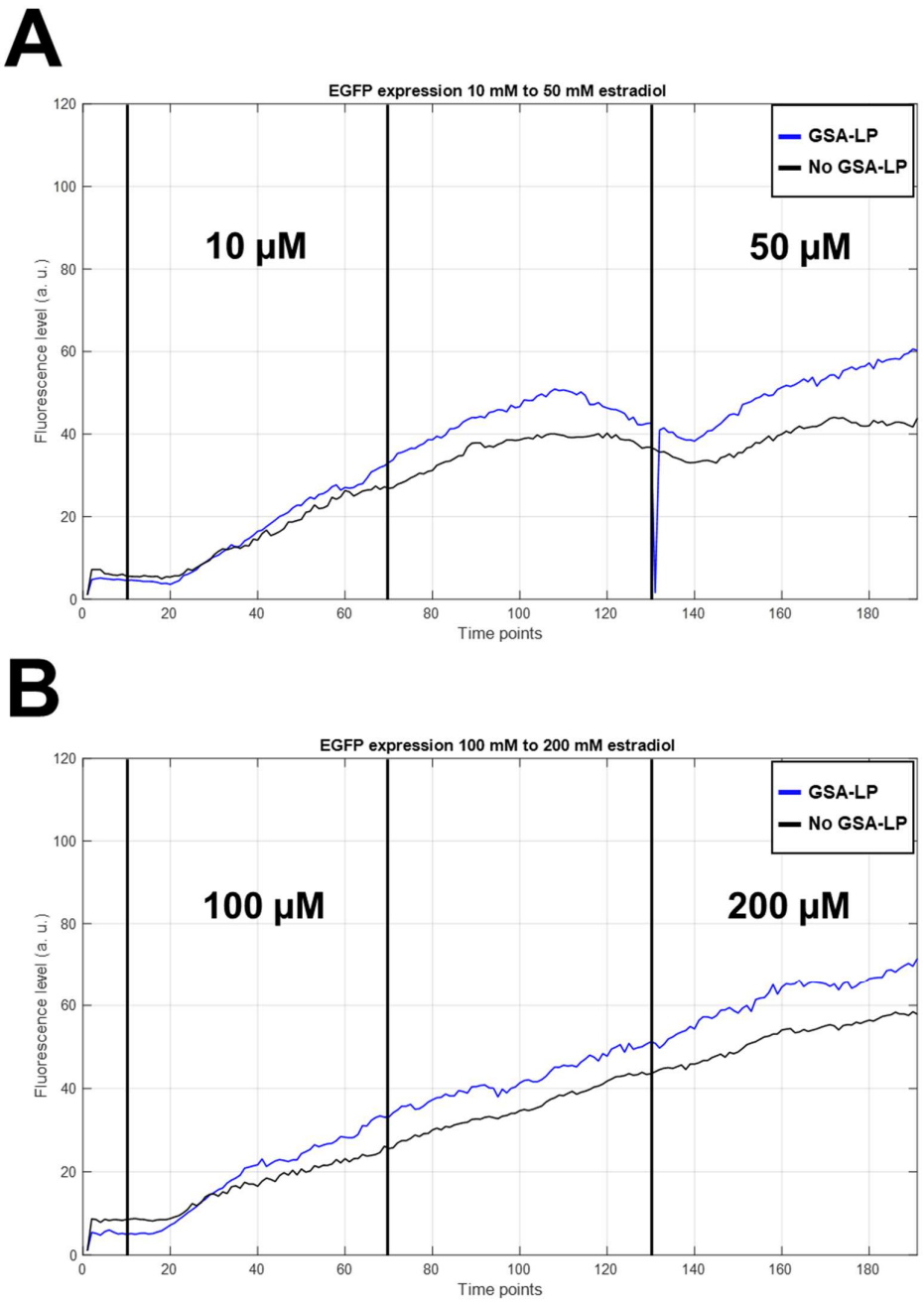

Figure 6. EGFP-NES and EGFP-GSA-LP-NES under LexAt promoter indirectly inducible by β-estradiol. Both strains (pLexAt-EGFP-NES and pLexAt-EGFP-GSA-LP-NES) were grown for 30 minutes (10 timepoints, 3 minutes per timepoint) in media without β-estradiol, then at lower concentration (10 mM for **A** and 100 for **B**) for 180 minutes, followed by 180 minutes of media without inducer and another 180 minutes of media with higher concentration.

To investigate the fluctuations of EGFP level in response to different inducer concentrations and compare strains with and without GSA-LP fragment between EGFP and NES, time lapse microscopy experiment was performed. Fluorescence intensity-estradiol concentration dependency was not linear, the most significant difference between two sets of concentrations was observed after timepoint 110. About two hours after inducer supply stopped, signal of both strains induced with lower concentrations started to go down and decreased until about 30 minutes from the start of supply (Figure 6A).

This experiment, as well as the Result 3.2.3, showed a difference between strains with and without GSA-LP. EGFP fluorescence in the strain with GSA-LP was higher than in the strain without Cln2 docking motif. It began to appear shortly after accumulation of EGFP started and once established, fluctuated in the range of 20 to 30% for both concentration sets.

3.3 DISCUSSION

3.3.1 Phosphodegron efficiency

In the current work, several Cln2-based phosphodegron tags were compared for their efficiency at promoting tagged protein phosphorylation.

All strains containing tagged EGFP showed significantly lower results than positive control (construct without any tags), including tags with phosphosites mutated to alanine.

Combination of phosphodegron with LP motif was expected to increase the rate of degradation of the tagged protein since it can enhance binding of Cln2-Cdc28 complex with the substrate. Specificity of Cln2-Cdc28 complex towards protein containing LP motif have to be increased in comparison to protein without it.

Surprisingly, 4WT and 5WT were the most efficient in promoting degradation, whereas strains with LP docking motif fused to degron seem to produce more stable EGFP. This effect was phosphorylation-independent since it was observed in alanine mutant too (Cln2_4A-LP).

That could be possibly due to change of LP positioning from original central region of Sic1 (Figure 3B) to the C-terminus of the tagged protein (Figure 3C). Further examination of the influence of LP positioning is required to understand the reason behind these results.

3.3.1 Strains with inducible EGFP expression

In the course of this study, two strains with integrated LexA-ER-B112 system (Ottoz, Rudolf, & Stelling, 2014) were created for further developing of systems with combined post-translational and transcriptional regulation and subsequent research on LP and degron positioning role in protein stability.

The advantages of this inducible system is that it is composed of parts orthogonal to yeast and does not interfere with its metabolism. Moreover, expression level depends on concentration of inducer, which provides additional possibilities for design of complex systems.

Time-lapse microscopy with this inducible promoter containing strains, aimed to investigate expression efficiency for different concentrations, showed the delayed decrease of the signal after the end of induction and absence of the decrease in strains induced with higher concentrations (Figure 6B). That could be explained by high EGFP stability and accumulation of

inducer bound transcription factors and should be taken in consideration while designing systems with pLexAt.

SUMMARY

Multisite phosphorylation degron tags analyzed in this work can ensure different levels of EGFP fluorescence. LP docking site did not show positive effect on degradation rate, and reasons behind it require further investigation.

Two yeast strains with EGFP under inducible promoter were generated for further tagging with phosphodegrons. Strain containing pLexAt-EGFP-GSA-LP-NES construct can be used to gather more data on contribution of LP positioning into protein stability.

Further research on the topic of this study could result in discovering more about cell cycle regulatory mechanisms. The practical implementation would be development of tools for precise regulation of proteins' expression and half-life. That would allow for better design of cell factories (control of metabolic enzymes), biomedical usage (cell programming to perform very specific tasks), cellular computing and more.

REFERENCES

Balasubramanian, M. K., & Bi, E. (2004). *Comparative Analysis of Cytokinesis in Budding Yeast, Fission Yeast and Animal Cells*. *14*, 806–818. https://doi.org/10.1016/j.cub.2004.09.022

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

Boettcher, B., & Barral, Y. (2013). The cell biology of open and closed mitosis. *Nucleus (United States)*. https://doi.org/10.4161/nucl.24676

Bononi, A., Agnoletto, C., De Marchi, E., Marchi, S., Patergnani, S., Bonora, M., … Pinton, P. (2011). Protein kinases and phosphatases in the control of cell fate. *Enzyme Research*, *2011*(1). https://doi.org/10.4061/2011/329098

De Bondt, H. L., Rosenblatt, J., Jancarik, J., Jones, H. D., Morgant, D. O., & Kim, S. H. (1993). Crystal structure of cyclin-dependent kinase 2. *Nature*. https://doi.org/10.1038/363595a0

Doncic, A., Eser, U., Atay, O., & Skotheim, J. M. (2013). An Algorithm to Automate Yeast Segmentation and Tracking. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0057970

Edgington, N. P., & Futcher, B. (2001). Relationship between the function and the location of G1 cyclins in S. cerevisiae. *J Cell Sci*, *114*(24), 4599–4611. Retrieved from http://jcs.biologists.org/cgi/content/abstract/114/24/4599

Epstein, C. B., & Cross, F. R. (1992). CLB5: A novel B cyclin from budding yeast with a role in S phase. *Genes and Development*, *6*(9), 1695–1706. https://doi.org/10.1101/gad.6.9.1695

Fisher, D. L., & Nurse, P. (1996). A single fission yeast mitotic cyclin B p34cdc2 kinase promotes both S-phase and mitosis in the absence of G1 cyclins. *The EMBO Journal*, *15*(4), 850–860. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8631306%0Ahttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC450283

Holt, L. J. (2012). Regulatory modules: Coupling protein stability to phopshoregulation during cell division. *FEBS Letters*. https://doi.org/10.1016/j.febslet.2012.05.045

Janke, C., Magiera, M. M., Rathfelder, N., Taxis, C., Reber, S., Maekawa, H., … Knop, M. (2004). A versatile toolbox for PCR-based tagging of yeast genes: New fluorescent proteins, more markers and promoter substitution cassettes. *Yeast*, *21*(11), 947–962. https://doi.org/10.1002/yea.1142

Jordan, J. D., Landau, E. M., & Iyengar, R. (2000). Signaling networks: The origins of cellular multitasking. *Cell*, *103*(2), 193–200. https://doi.org/10.1016/S0092-8674(00)00112-4

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

Landry, B. D., Doyle, J. P., Toczyski, D. P., & Benanti, J. A. (2012). F-box protein specificity for G1 cyclins is dictated by subcellular localization. *PLoS Genetics*, *8*(7). https://doi.org/10.1371/journal.pgen.1002851

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

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

Lõoke, M., Kristjuhan, K., & Kristjuhan, A. (2011). Extraction of genomic DNA from yeasts for PCR-based applications. *BioTechniques*. https://doi.org/10.2144/000114497

Mans, R., van Rossum, H. M., Wijsman, M., Backx, A., Kuijpers, N. G. A., van den Broek, M., … Daran, J. M. G. (2015). CRISPR/Cas9: A molecular Swiss army knife for simultaneous introduction of multiple genetic modifications in Saccharomyces cerevisiae. *FEMS Yeast Research*, *15*(2), 1–15. https://doi.org/10.1093/femsyr/fov004

Mateus, C., & Avery, S. V. (2000). Destabilized green fluorescent protein for monitoring dynamic changes in yeast gene expression with flow cytometry. *Yeast*. https://doi.org/10.1002/1097-0061(200010)16:14<1313::AID-YEA626>3.0.CO;2-O

Miller, M. E., & Cross, F. R. (2001). Cyclin specificity: how many wheels do you need on a unicycle? *Journal of Cell Science*.

Morgan, D. O. (2007). *The Cell Cycle: Principles of Control* (E. Lawrence, Ed.). London:

New Science Press Ltd.

Ottoz, D. S. M., Rudolf, F., & Stelling, J. (2014). Inducible, tightly regulated and growth condition-independent transcription factor in Saccharomyces cerevisiae. *Nucleic Acids Research*, *42*(17). https://doi.org/10.1093/nar/gku616

Quilis, I., & Igual, J. C. (2017). A comparative study of the degradation of yeast cyclins Cln1 and Cln2. *FEBS Open Bio*. https://doi.org/10.1002/2211-5463.12157

Rogers, S., Wells, R., & Rechsteiner, M. (1986). Amino acid sequences common to rapidly degraded proteins: The PEST hypothesis. *Science*. https://doi.org/10.1126/science.2876518

Russo, A. A., Jeffrey, P. D., & Pavletich, N. P. (1996). Structural basis of cyclin-dependent kinase activation by phosphorylation. *Nature Structural Biology*. https://doi.org/10.1038/nsb0896-696

Salama, S. R., Hendricks, K. B., & Thorner, J. (1994). G1 cyclin degradation: the PEST motif of yeast Cln2 is necessary, but not sufficient, for rapid protein turnover. *Molecular and Cellular Biology*. https://doi.org/10.1128/mcb.14.12.7953

Skowyra, D., Craig, K. L., Tyers, M., Elledge, S. J., & Harper, J. W. (1997). F-box proteins are receptors that recruit phosphorylated substrates to the SCF ubiquitin-ligase complex. *Cell*, *91*(2), 209–219. https://doi.org/10.1016/S0092-8674(00)80403-1

Smith, G. K., Ke, Z., Guo, H., & Hengge, A. C. (2011). Insights into the Phosphoryl Transfer Mechanism of Cyclindependent Protein Kinases from Ab Initio QM/MM Free-energy Studies. *J Phys Chem B.*, *115*(46), 13713–13722. https://doi.org/10.1038/mp.2011.182.doi

Stern, B., & Nurse, P. (1996). A quantitative model for the cdc2 control of S phase and mitosis in fission yeast. *Trends in Genetics*. https://doi.org/10.1016/S0168-9525(96)80016-3

Waldo, G. S., Standish, B. M., Berendzen, J., & Terwilliger, T. C. (1999). Rapid protein-folding assay using green fluorescent protein. *Nature Biotechnology*, *17*(7), 691–695. https://doi.org/10.1038/10904

Non-exclusive licence to reproduce thesis and make thesis public

I, Aleksandra Panfilova,

(*author's name*)

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

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

1.2. make 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 ***20/05/2024*** until the expiry of the term of copyright,

Cln2-based phosphodegron tags for the regu-lation of protein stability,

(*title of thesis*)

supervised by Nastassia Shtaida, PhD, and Rait Kivi, PhD,

(*supervisor's name*)

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

3. 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.

Aleksandra Panfilova

20/05/2019