UNIVERSITY OF TARTU
Faculty of Science and Technology
Institute of Technology

Luka Bulatović

The use of phospho-degrons for controlled protein expression in the cell cycle

Master's Thesis (30 ECTS)

Curriculum Bioengineering

Supervisor(s):

Assoc. Prof., PhD Ilona Faustova

PhD Mihkel Örd

Prof., PhD Mart Loog

Tartu 2021

The use of phospho-degrons for controlled protein expression in the cell cycle

Abstract:

Cyclin-dependent kinases (Cdks) are the master regulators of the cell cycle, which together with their activating cyclin subunits, phosphorylate hundreds of targets to govern the cell cycle progression. The precise substrate targeting mechanisms of the Cdk complex have led to a hypothesis that the Cdk machinery could be used to control the expression and cellular concentration of any desired protein by fusion with different phospho-regulation modules. The modules consist of barcodes of linear motifs including phosphorylation sites and cyclins docking motifs. These barcodes are scanned by the Cdk complex and depending on the bar-code, the machinery will differentially regulate the protein in the cell cycle. This work shows that one such module, the Far1 degron, can be used to control the cellular concentration of different proteins. We also demonstrate that different versions of the degron allow us to fine-tune the cellular concentration of the targeted protein in different cell cycle phases.

Keywords:

Cell cycle, Phosphorylation, Cyclin-dependent kinase, Cyclin, Kinase specificity

CERCS: P310 Proteins, enzymology

Fosfo-degradatsiooni moodulid valkude tasemete kontrollimiseks rakutsüklis

Lühikokkuvõte:

Tsükliinist sõltuvad kinaasid (Cdk) on rakutsükli kesksed regulaatorid, mis kompleksis tsükliinidega fosforüleerivad sadu valke, et kontrollida rakutsükli kulgu. Cdk kompleksi mitmekesised substraatide äratundmise mehhanismid on viinud hüpoteesini, et Cdk kompleksi saaks kasutada, et kontrollida iga huvipakkuva valgu aktiivsust rakutsüklis. Selleks tuleks lisada valgule moodulid, mis koosnevad erienvatest lineaarsetest motiividest nagu fosforüleerimissaidid ja tsükliini seondumismotiivid, mida saab ennustatavalt kombineerida. Cdk kompleks töötleb neid mooduleid vastavalt nende omadustele ja seekaudu kontrollib huvipakkuva valgu aktiivsust. Käesolevas töös näidatakse, et ühe sellise mooduli abil saab kontrollida erinevate valkude kontsentratsiooni rakutsükli jooksul. Lisaks näidatakse, et degradatsioonimooduli omaduste manipuleerimisega saab piirata valkude aktiivsust erinevatesse rakutsükli faasidesse.

Võtmesõnad:

Rakutsükkel, fosforüleerimine, tsükliinist sõltuvad kinaasid, tsükliin, kinaasi spetsiifika

CERCS: P320 Proteiinid, ensümoloogia

TABLE OF CONTENTS

TERMS, ABBREVIATIONS AND NOTATIONS

APC - Anaphase promoting complex

Cdk – Cyclin-dependent kinases

CKI - Cyclin-dependent kinase inhibitor

CSM - Complete supplement mixture

EDTA - Ethylenediaminetetraacetic acid

eGFP – Enhanced green fluorescent protein

OD – Optical density

pEL222 – VP-EL222 binding promoter

SCF - Skp, Cullin, F-box containing complex

SCRaMbLE - Synthetic Chromosome Rearrangement and Modification by LoxPsym-mediated Evolution

TE buffer - Tris-EDTA buffer

TF – transcriptional factor

YFP – Yellow fluorescence protein

YPD – Yeast extract peptone dextrose

INTRODUCTION

The cell cycle is a timely controlled, multi-step set of events that takes place in four different phases. The progression of the cell cycle is governed by the cell cycle control system. The main component of the cell cycle control system, and the driver of the cell cycle, is a family of proteins known as cyclin-dependent kinases (Cdks). Cdks are proteins that phosphorylate numerous different substrates, thereby changing their activity. However, Cdks are catalytically active only in the presence of an activating protein known as a cyclin. Different cyclins are expressed during different stages of the cell cycle, and together with a Cdk, they form cyclin-Cdk complexes that perform substrate-specific phosphorylation, as different cyclins help guide Cdks to different substrates. This mechanism allows Cdks to phosphorylate numerous substrates during various phases of the cell cycle.

Therefore, as Cdk carries out precise regulation of hundreds of different proteins, the Cdk machinery system could potentially be used to regulate additional proteins of interest. Nonetheless, for the targeted proteins to be recognized by the Cdk machinery, they have to be linearly encoded with a set of different phosphorylation clusters and cyclin docking motifs. Once encoded, they possess a barcode which Cdk later uses to phosphorylate these proteins at the pre-programmed time in the cell cycle.

Many studies have demonstrated current limitations present in the field of protein regulation, especially in the context of regulation and control of different metabolic pathways. If controlled correctly, the cells in which these pathways are expressed can experience notable improvements in the growth and production rates and the final yields of the desired compound. These exciting findings pave the way for the Cdk machinery to be used in synthetic biology and cell factories to control the expression and cellular concentrations of different proteins.

In this work, we focus our attention on the N-terminal domain of a Cdk inhibitor protein, Far1, and its potential use as a modular degron. We use different experimental setups to show that the Far1 degron can control the expression and concentration of a protein of interest during the cell cycle. We also try to show that this degron can be used to restrict the activity of the desired protein to a G1 phase of the cell cycle. Lastly, we work to improve an existing system believed to benefit from a restricted activity of one of its main enzymes.

1 LITERATURE REVIEW

1.1 The cell cycle

A cell is the smallest self-governing structural unit that serves as an essential building block of every living organism. The process of cellular division takes place in all organisms, and it serves as a vital step for the regulation of the development of new cells. The mechanism of cell division in a eukaryotic cell is known as the cell cycle. The cell cycle is a tightly controlled multi-step process that takes place during four different phases. The cell cycle progression is timely regulated, and a multitude of events take place in an ordered manner. The four phases of the cell cycle are G1, S, G2, and M phases **(Figure 1)**.

The most important phases of the cell cycle are the S and M phase. During the synthesis (S) phase, the cell's chromosomal DNA is duplicated into two identical copies referred to as the sister chromatids. These chromatids are later split between two daughter cells during the next critical phase of the cell cycle, the M phase. The M phase is divided into two smaller phases, the nuclear division or mitosis, and the cell division or cytokinesis.

G1 and G2 are commonly known as the gap phases, and their primary purpose is to control the progression of the cell cycle and allow the upcoming phases to occur. They act as separators between the two main phases of the cycle, thereby allowing the cell to check if all previous events were executed correctly before continuing with the later stages. During G1, the cell works on acquiring more biomass and prepares its DNA for the upcoming replication. Moreover, the G1 phase is a phase of 'no return' during which the cell needs to decide to go on with the cycle's progression or start the exit from the cell cycle. The cell might not wish to carry out the progression of the cell cycle due to unfavorable external signals received directly from the neighboring cell or the environment. The cell then stops the progression of the cell cycle and enters into a state of non-division, known as the G0 phase. The second gap phase (G2) is the shortest of all four phases, and during this phase, the cell prepares the newly replicated chromosomes for segregation. (Morgan, 2007).

On the other hand, a cell can be arrested and kept in one of the four different cell cycle stages in the process known as the induced cell cycle arrest. The presence of mating pheromone, α-factor, or other extracellular stress signals cause yeast cells to arrest in the G1 phase. Similarly, problems with DNA replication, for example those caused by hydroxyurea (HU) that

negatively affects the initiation and elongation of chromosomal duplication, cause cell cycle arrest in the S phase. Lastly, problems with chromosome alignment and segregation, like those caused by the presence of nocodazole, which interferes with the assembly/disassembly mechanics of microtubules, result in the cells arresting in the M phase (Morgan, 2007).

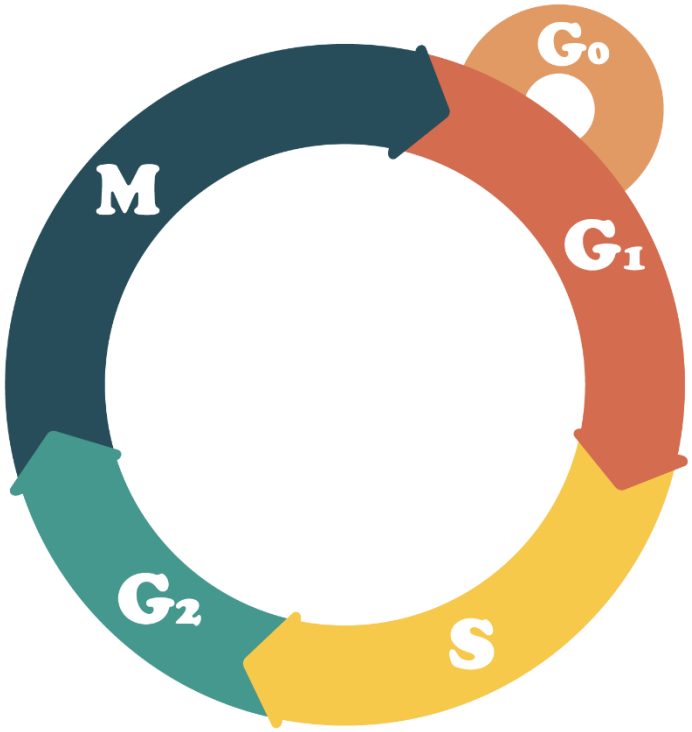


Figure 1. The cell cycle diagram. The schematic representation of different phases of the cell cycle. The cell cycle is a timely triggered set of events that takes place in four phases: G1, S, G2, and M phase. The gap phases (G1 and G2) serve as control phases during which the cells accumulate more biomass and prepare for the events of the upcoming phases. During the S phase, the chromosomes are duplicated and later divided into two daughter cells in the M phase. A cell might choose to enter into a prolonged state of no division (G0 phase) due to the presence of unfavorable environmental signals (Morgan, 2007).

1.2 The cell cycle control system

The cell cycle control system oversees the timely progression of the cell cycle and ensures that different events occur in an ordered manner. The control system is in charge of triggering and controlling a series of different biochemical events that aid the cell's progression through the three major checkpoints in the cell cycle. The main driver of the cell cycle control system is a family of proteins called cyclin-dependent kinases (Cdks). They control all important events from the start of the cell cycle and the DNA replication to the transition into the M phase and various mitotic events. CDK proteins transfer a phosphate group from an ATP molecule to a substrate protein in a process known as phosphorylation. Once phosphorylated, the substrate protein can go through conformational changes, which in turn alter

its activity, or changes in its interactions with other proteins, which can affect the localization and stability of the protein. This fine-tuning of the activity of different cell cycle driving proteins allows the Cdks to have complete control over the events of the cell cycle (Örd & Loog, 2019).

Nonetheless, as their full name states, cyclin-dependent kinases can be catalytically active only if they are in a complex with their activating partners – cyclins (Pines, 1995). Cyclins are proteins that act as activators of Cdks and, together with them, form various cyclin-CDK complexes. Different cyclins are expressed during different stages of the cell cycle, and each cyclin has a unique set of target proteins that they dock to (Bhaduri & Pryciak, 2011; Faustova *et al*., 2021; Loog & Morgan, 2005; Örd *et al*., 2020; Örd & Loog, 2019). This docking interaction presents the target protein to Cdk, which in turn phosphorylates it. Therefore, different cyclins allow Cdks to phosphorylate various substrates during different cell cycle phases, thereby ensuring an ordered progression of the cell cycle. The cell cycle control system is highly robust and adaptable, and it functions through a series of regulatory networks and feedback loops that ensure its progression (Morgan, 2007). The system is also shown to be rather evolutionary conserved as the cell cycle still successfully proceeds when a Cdk1 in *S. cerevisiae* is replaced with its human homolog (Wittenberg & Reed, 1989).

1.3 Cyclins and cyclin-dependent kinases

Cdks are the master regulators of the cell cycle, which phosphorylate different target proteins and, in this manner, control their activity. One of the most studied Cdks is Cdk1 found in the model organism *S. cerevisiae*. Unlike higher eukaryotes which utilize different Cdks during different stages of the cell cycle, *S. cerevisiae* only depends on Cdk1 to regulate the entire progression of the cell cycle. Cdk1 is a proline-directed threonine/serine protein kinase that phosphorylates the full consensus phosphorylation sequence T/S-P-x-K/R, where T/S stands for threonine or serine, P for proline, x for any amino acid, and K/R for positively charged amino acids lysine and arginine, respectively (Songyang *et al*., 1994).

Interestingly, it has also been shown that Cdk1 can phosphorylate the minimal consensus T/S-P sequence as well as some non-proline sites (Suzuki *et al*., 2015). Nonetheless, these phosphorylations both exhibited much lower efficiency than the phosphorylation of the full consensus sequence (Kõivomägi *et al*., 2011).

Cyclins are proteins that bind to Cdk1 and form different cyclin-Cdk1 complexes during different phases of the cell cycle. This binding promotes the catalytic activity of Cdk1,

thereby allowing it to phosphorylate different substrates. Different cyclins are expressed during the cell cycle **(Figure 2)** (Morgan, 2007).

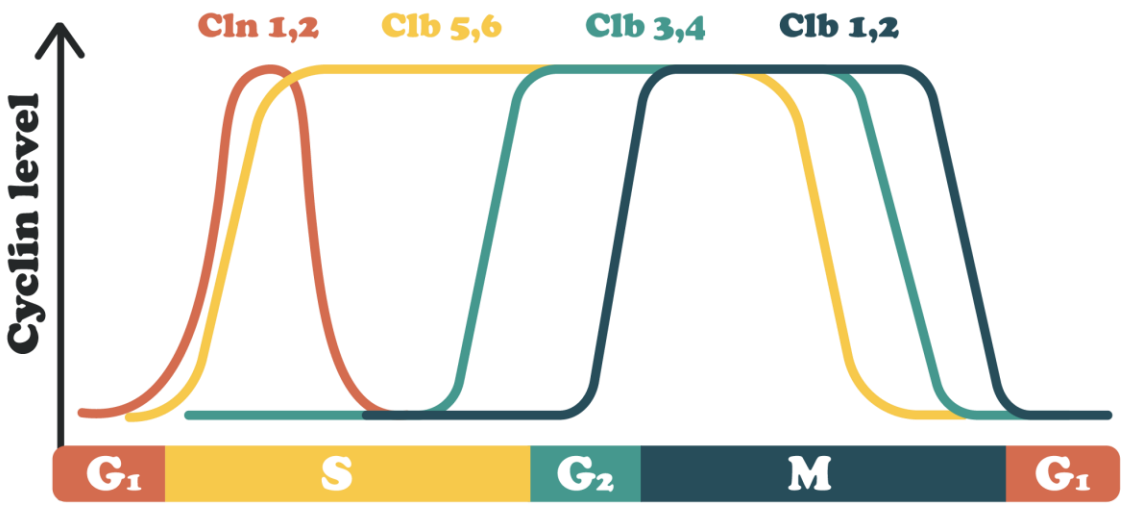


Figure 2. Cyclin expression during different phases of the cell cycle. *S. cerevisiae* expresses nine different Cdk1 cyclins during different phases of the cell cycle. Cln1,2 are the G1/S cyclins responsible for bud morphogenesis and initiation of the cell cycle. S phase cyclins (Clb5,6) promote DNA replication. The late S phase cyclins (Clb3, Clb4) enable the cell's smooth transition into the M phase, during which the M phase cyclins (Clb1,2) initiate different mitotic events. Cln3 (not shown in the figure) is active from the G1 until the early M phase and plays an important role in the transcription of other G1 cyclins, supporting cell cycle entry.

In addition to acting as Cdk1 activators, cyclins allow for the enlisting of different target substrates and increase the efficiency of their binding to the Cdk1. In total, there are nine cyclins expressed in *S. cerevisiae* during the cell cycle: three are known as G1 cyclins and others as B-type cyclins due to their homology to human cyclin B (Enserink & Kolodner, 2010).

The three G1 cyclins (Cln1, Cln2, Cln3) have an important role in cell cycle initiation. The research has demonstrated that only a triple G1 cyclin deletion mutant strain (*cln1Δ cln2Δ cln3Δ)* is inviable, suggesting that any G1 cyclin can trigger the start of the cell cycle (H. E. Richardson *et al*., 1989). However, despite their initial similarity in initiating the early events of the cell cycle, the G1 cyclins, have rather different roles. Cln3 functions as a regulator of transcription, and it is shown to work upstream of other G1 cyclins and promote the transcription of *CLN1* and *CLN2* genes and thus the entry to cell cycle (Marini & Reed, 1992).

On the other hand, the G1/S phase cyclins (Cln1, Cln2) are responsible for initiating the spindle pole body duplication and bud morphogenesis. During the G1 phase, there is a strong

cellular presence of Cdk1 inhibitor (CKIs) proteins, Sic1 and Far1. These CKIs prevent the activation of cyclin-Cdk1 complexes in the early G1 phase. However, due to the rising levels of Cln1 and Cln2 cyclins in the late G1 phase, the activity of the Cln1/Cln2-Cdk1 complexes increases. This increase in activity, accompanied by the expression of S phase cyclins Clb5/Clb6, enables the Cdk1 to phosphorylate CKIs and direct them for ubiquitin-mediated proteolysis (Mendenhall & Hodge, 1998; Schwob *et al*., 1994). From this point on, the activity of Cdk1 keeps on increasing until the start of the anaphase, during which the CKIs are re-expressed, and cyclins are degraded (Amon *et al*., 1994). The G1 cyclins also have a self-regulatory role; they promote their own phosphorylation by the cyclin-Cdk1 complex, after which they are tagged for SCF(Skp1/Cullin/F-box)-mediated destruction (Lanker *et al*., 1996; Tyers *et al*., 1992).

In addition to the G1 cyclins, *S. cerevisiae* expresses six B-type cyclins (Clb1-6). With a notable exception of Clb5 and Clb6, whose expression is already induced during the G1 phase, other B-type cyclins are expressed during the S phase, right after the degradation of the G1 cyclins. Clb5 and Clb6 (S phase cyclins) enable the cell's timely entry into the S phase; they trigger the DNA replication and work on disabling re-replication (Dahmann *et al*., 1995; Schwob *et al*., 1994). Clb3 and Clb4 represent the last S phase cyclins whose main role is to initiate the mitotic entry and spindle assembly. They are expressed in the transitional period between the S and G2 phases and remain active until they get degraded in the anaphase (Mendenhall & Hodge, 1998; H. Richardson *et al*., 1992). Lastly, the M phase cyclins (Clb1, Clb2) have an important role in perpetuating bud morphogenesis and the initiation of different events in mitosis. The M cyclins are expressed at the end of the S phase, and they remain active until the late M phase, during which they are degraded by the anaphase promoting complex (APC) (Lew & Reed, 1993; Mendenhall & Hodge, 1998; Seufert *et al*., 1995).

1.4 Protein degradation in the cell cycle

The anaphase-promoting complex (APC) is a protein ligase that promotes the ubiquitination of securin and S/M cyclins and whose activity starts to increase during the M phase (McLean *et al*., 2011). The destruction of securin by the APC promotes the release of a protease called separase which lyses the cohesion complex. Once the cohesion complex is lysed, the sister chromatids are pulled to two opposite ends of the dividing cell (Morgan, 2007). Another important role that the APC has in controlling the cell cycle is the degradation of S and M cyclins at the end of mitosis. It is worth noting that the S phase cyclins, Clb5 and Clb6, are

degraded during different cell cycle stages. Unlike Clb6, Clb5 does not have a SCF^{Cdc4} degron. This lack of degron enables it to stay stable until anaphase, where the APC degrades it. On the other hand, Clb6 is phosphorylated by Cdk1 and Pho85 complexes at the beginning of the S phase and sent for SCF-mediated destruction (Jackson *et al*., 2006).

The SCF complex, together with the APC, is a central E3 Ub-ligase involved in the control of the cell cycle. Regulated degradation of proteins takes place via the Ub-mediated proteolytic pathway (Hershko & Ciechanover, 1998). The pathway controls the formation of covalently attached Ub-chains, which are added onto a target protein substrate and serve as a degradation signal.

The E3 Ub-ligases are the last step in the ubiquitination pathway, and they facilitate the transfer of ubiquitin (76-amino acid peptide) onto the target substrate (Ang & Wade Harper, 2005). The SCF complex comprises three main subunits: Skp1, Cullin, and Rbx1/Roc1(Koepp *et al*., 2001). The complex acts together with an adopter subunit called F-box protein which serves as a direct recruiter of the substrate. F-box proteins recognize their substrates mainly due to the presence of posttranslational modifications, most notably glycosylation and phosphorylation (Feldman *et al*., 1997; Skowyra *et al*., 1997). One of the examples of the F-box proteins is yeast's Cdc4 protein. Substrate's phosphorylated motifs, known as phospho-degrons, make direct interaction with the F-box proteins, thereby forming a direct link with the ubiquitination machinery (Skowyra *et al*., 1997). The CKI protein Far1 is ubiquitinated in the early S phase by the SCF^{Cdc4} complex. The Clb5-Cdk1 complex phosphorylates Far1's S87 and S91 di-phosphodegrons, thereby enabling Far1 -Cdc4 interaction (Faustova *et al*., 2021). This interaction accounts for a timely and rapid degradation of Far1.

Lastly, the SCF complexes are active throughout the cell cycle, whereas APC is mostly active during the M and G1 phases of the cell cycle **(Figure 3)**. Therefore, it can be noted that the activity of APC and Cdk1 oscillates following a negative feedback loop; when the activity of APC is high (G1 and late M phase), the activity of Cdk1 is low, and vice versa, when the activity of Cdk1 is high, the activity of APC is low (the remaining phases of the cell cycle) (Lin & Scott, 2012)

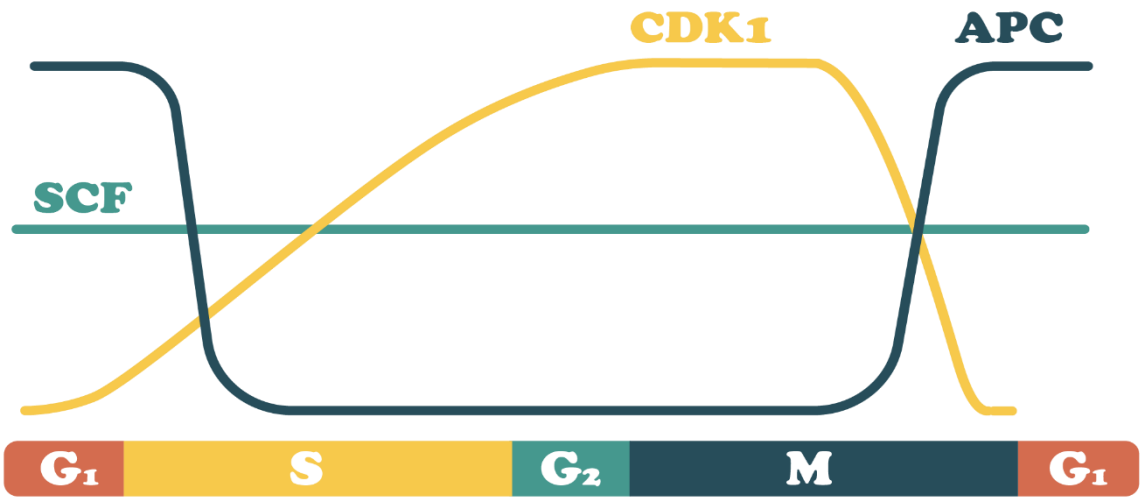


Figure 3. The activity diagram of SCF, Cdk1, and APC during the cell cycle. SCF complex is active during the whole duration of the cell cycle, and with the help of F-box proteins, it ubiquitylates target proteins. Cdk1 and APC function in a negative feedback loop manner, i.e., as the activity of APC rises, the activity of Cdk1 decreases, and vice versa.

1.5 Targeting of Far1 by cyclin-Cdk1 complex

Cyclin-Cdk1 complexes play a crucial role in a timely ordered progression of the cell cycle, and it is believed that Cdk1 alone phosphorylates over 500 different protein targets (Ubersax *et al*., 2003). Temporally controlled phosphorylation of such a large number of different targets is made possible by the presence of cyclins that direct Cdk1 directly to the targeted substrates. S cyclins are known to interact with the highest number of target proteins compared to any other cyclin group (Loog & Morgan, 2005; Schulman *et al*., 1998).

On the other hand, Far1, a multi-mode inhibitor protein, works on inhibiting the activity of Cdk1 as well as the substrate recognition of Cln1/2-Cdk1 complexes (Pope *et al*., 2014). The inhibition of the cyclin-Cdk1 complex is required for the low activity of Cdk1 during the G1 phase. In the presence of a mating pheromone (α-factor), Far1 completely inhibits the activity of Cln-Cdk1 complexes, thus arresting the cell in the G1 phase. Nonetheless, in the absence of the α-factor, the concentration of Far1 is considerably lower, and Far1 is phosphorylated both in the nucleus and the cytoplasm (Blondel *et al*., 2000). It has been shown that an alanine mutation at the Cdk1 phosphorylation site S87 prevents the degradation of Far1 and keeps it stable throughout the course of the cell cycle (Gartner *et al*., 1998). Later, it was found that S87 and a non-proline site S91 form a Cdc4 phospho-degron (Faustova *et al*.,

2021). Therefore, as previously explained, Far1 is phosphorylated by the Cln2-Cdk1 and Clb5-Cdk1 complexes at the phospho-degron sites marking it for degradation by the SCF^{Cdc4} complex (Figure 4). Furthermore, it has been demonstrated that a strain containing a double Clb5 and Clb6 deletion (*Δclb5Δclb6)* arrests a cell in the G1 phase for a much longer period compared to the wildtype strain (Doncic & Skotheim, 2013). This shows that Clb5 plays an important role in the phosphorylation and degradation of the Far1 protein.

Cyclins contain different docking pockets capable of binding short linear motifs (SLiMs) present in the target proteins (Örd & Loog, 2019). Some cyclins contain a hydrophobic patch that binds to the short linear motifs commonly known as the RxL motifs. These motifs have a consensus sequence of K/R-x-L-x-φ or K/R-x-L- φ, where K is lysine, R is arginine, x is any amino acid, L is leucine, and φ is a large hydrophobic amino acid (Lowe *et al*., 2002). For example, the Clb5-Cdk1 complex binds to the RxL motif in CKI protein Sic1 and sends it for the ubiquitin-mediated proteolysis (Schneider *et al*., 1996).

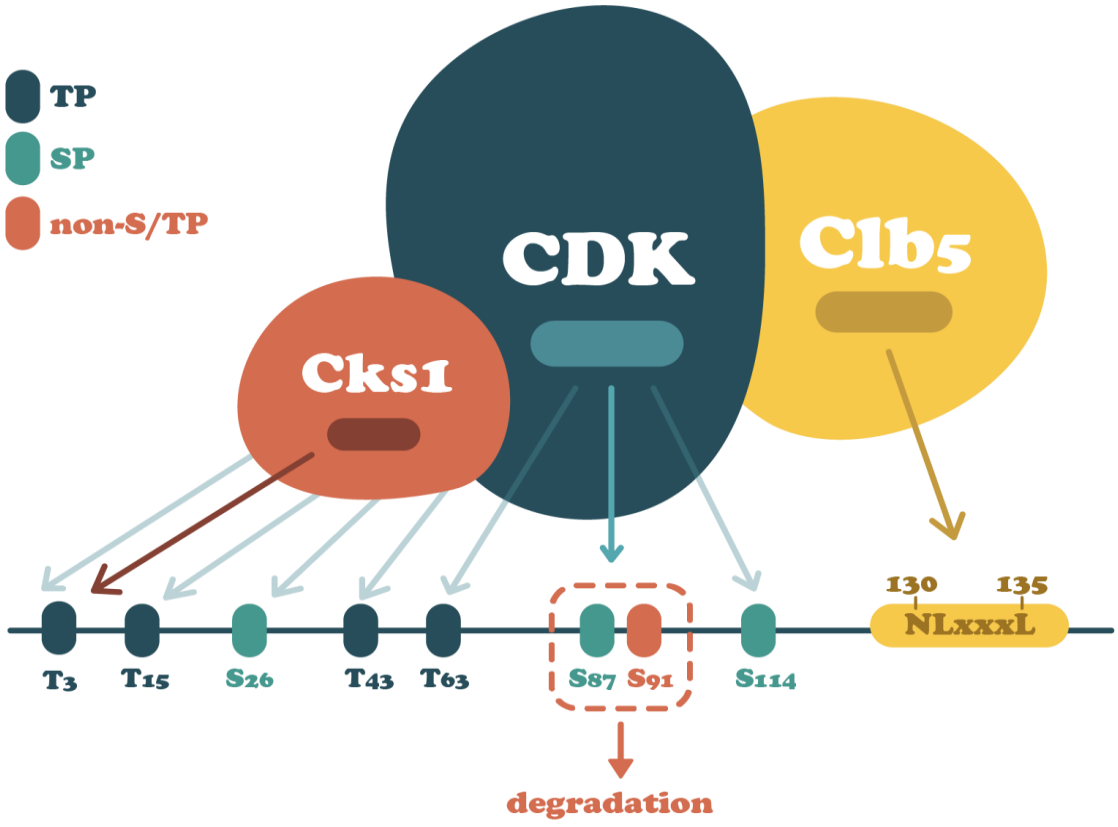


Figure 4. Scheme of the disordered Far1 N terminus (residues 1-150). The NLxxxL motif binds to the docking pocket of Clb5 and enhances phosphorylation of the Far1 N terminus by Cdk1. Cdk1 phosphorylates different S/T-P sites, and upon phosphorylation of the S87-S91 phospho-degron, the protein is sent for SCF-mediated degradation. Cks1, a phospho-adaptor unit, binds to the priming phospho-threonine site, thereby promoting the multisite phosphorylation of the protein.

Similarly, another cyclin binding SLiM is found in the N-terminal region of Far1, and it is referred to as the NLxxxL motif (Faustova *et al*., 2021). The motif promotes substrate docking to Clb5 and Clb6, enhancing the phosphorylation of Far1 by the S-Cdk1 complex **(Figure 4).** In mutants containing an alanine mutation to leucine L135A and L131A, the degradation of Far1 protein is substantially slower than the wildtype, further demonstrating the importance of the NLxxxL motif in the phosphorylation of Far1. The motif is also shown to be modular, meaning that once introduced into a new protein, it gives similar results as in Far1, recruiting S-Cdk1 to phosphorylate the protein (Faustova *et al*., 2021). A key difference between NLxxxL and the previously known S-Cdk1 docking motif RxL is that NLxxxL is exclusively used by S phase cyclins, whereas RxL is also used by later Clb-Cdk1 complexes (Faustova *et al*., 2021). Thus, the substrates with NLxxxL motifs could be phosphorylated only in the time window of S phase cyclin expression.

In addition to Clb5, a phospho-adaptor subunit (Cks1) binds to the Cdk1 and forms a complex. The main role of Cks1 is to promote the multisite phosphorylation of substrates by binding phospho-threonines and using them as docking sites (Kõivomägi *et al*., 2013; McGrath *et al*., 2013). In the example of Far1, the Cks1 binds to the phosphorylated T3 and promotes phosphorylation of other T/S-P sites. A mutant in which all priming threonines (T3, T15, T43, T63) were replaced with serines showed a slower degradation profile compared with the wildtype, suggesting that multisite phosphorylation mediated by Cks1 docking can be used to speed up the degradation profile of a protein (Faustova *et al*., 2021). In addition to this, it has been shown that the multisite phosphorylation code controls the phosphorylation of different Cdk1 targets through the presence of different phosphorylation motif clusters. These clusters behave as timing tags that promote the substrate's site-specific phosphorylation at different Ckd1 thresholds (Örd *et al*., 2019). Furthermore, it has also been demonstrated that the 1-150 amino acid region of Far1 can function as a degron. By introducing different mutations in the NLxxxL motif, the phospho-degron region, or the Cks1 binding sites, one can construct different versions of this degron that could be degraded across different time points throughout the cell cycle (Faustova *et al*., 2021).

In conclusion, the Cdk1 thresholds can be linearly encoded into target substrates through the use of different phosphorylation clusters and cyclin docking motifs. Such linear encoding allows for predictable phosphorylation of the substrates during the duration of the whole cell cycle (Örd *et al*., 2019). This gives rise to a possibility that the cell cycle machinery can be used to limit the activity of any desired protein to a specific time window in the cell cycle.

These linearly encoded proteins, whose cell concentration levels are controlled throughout the cell cycle, promise to bring about revolutionary changes in the fields of synthetic biology and cell factories.

1.6 Synthetic biology and yeast bio-factories

Synthetic biology is an emerging discipline that aims to bridge the gap between biology and engineering. In synthetic biology, microorganisms are given a new set of 'tools' and programmed to perform non-native functions in order to produce the desired product (Delgado & Porcar, 2013). This programming of living organisms holds great potential in enabling the production of valuable chemicals in a more eco-friendly way. One way in which this production can be achieved is through the design of microbial cell factories. *S. cerevisiae* has been used as a model organism for a cell factory due to its versatility, robustness and efficacy. Moreover, *S. cerevisiae* is one of the most well-studied eukaryotic model organisms. It is characterized by a small genome size (12 Mb), short duplication time, and its susceptibility to being genetically modified. In addition to this, and due to its ability to efficiently express proteins with high production rates, *S. cerevisiae* is important in the large-scale industrial production of different enzymes, molecules, and proteins.

An interesting example of how *S. cerevisiae* can be used as a cell factory comes from a study where *S. cerevisiae* was programmed to express the linalool and geraniol biosynthesis pathways. The expression of these metabolic pathways allows the production of flavor molecules derived from hops, thereby eliminating the need for the addition of hop flavor in the beer brewing process (Denby *et al*., 2018). For the right flavor, however, the ratio of linalool and geraniol is important, thus, precisely controlled and not the highest possible expression level of multiple enzymes were required in this cell factory. This new hop-producing yeast strain lowers the costs of the beer brewing process, considering that hops are very expensive additives. It also provides the brewers with an eco-friendly alternative since hops take up many natural resources (water, land, etc.) to be grown.

However, if not controlled efficiently, the expression patterns of many metabolic pathways can dramatically vary in the culture as the conditions change. These changes can, in turn, have a negative effect on yeast's growth and productivity. One study showed that upon the introduction of a controlled delay in the expression of the metabolic modules, the experimental-phase growth rate increased 1.6 times and the post-experimental phase productivity increased twofold (Peng *et al*., 2017). Similarly, a Cln2-based degron was used to improve and optimize the production of muconic acid in *S. cerevisiae*, through the tagging of a bottle

neck enzyme (Pyne *et al*., 2018). Nonetheless, the degron used in the study was poorly described and there was only one version of it, suggesting that the addition of other expression profiles could be more optimal.

Therefore, the precise regulation of expression levels is critical for efficient growth and productivity of a yeast cell factory. Furthermore, a controlled expression can help address the rate limiting reactions in a metabolic pathway, thereby increasing the yields. So far, transcriptional regulation has been widely used to control the expression level and although post-translational regulation of protein activities is ubiquitous in the cell, it has not been utilized much in synthetic biology. Precisely controlled regulation of different proteins can be achieved through the use of Cdk1 machinery that is able to control different linearly encoded phosphorylation modules throughout the cell cycle.

1.7 Blue light-inducible VP-EL222 expression system

An inducible expression system is an important tool in molecular biology used to regulate cellular protein levels. The system allows for a conditional and controllable gene expression, which is important in the synthetic biology and biotechnology industry. There are various inducible expression systems, though the most common ones are based on light-inducible transcription factors, metabolism, or hormones. In an inducible expression system, the expression levels are regulated by altering the strength of different inputs (Benzinger & Khammash, 2018). However, many of these expression systems show high toxicity, limited dynamic range, and rather slow activation/deactivation dynamics (Motta-Mena *et al*., 2014).

VP-EL222 is a light-inducible gene expression system designed to overcome the shortcomings mentioned above and allow for a broad spectrum of use. The system is composed of the transcriptional activator from the VP16 protein fused to the bacterial transcription factor (TF) EL222 (Nash *et al*., 2011). The EL222 TF contains a photosensory light oxygen-voltage (LOV) domain and a Helix-Turn-Helix (HTH) DNA-binding domain, which are both needed for the blue-light dependent transcriptional activation (Huala *et al*., 1997; Motta-Mena *et al*., 2014).

In the dark, the LOV domain binds to the HTH domain and inhibits the dimerization and DNA binding **(Figure 5)** (Nash *et al*., 2011). In the presence of blue light (450-490 nm), a photochemical reaction between the LOV domain and its flavin chromophore takes place, thereby triggering the formation of the VP-EL222 dimer and its binding to the DNA (Zoltowski & Gardner, 2011). The dimer binds to the pEL222 promoter and activates the

transcription of the protein of interest. The cellular concentration of the protein is proportional to the strength of the input signal, i.e., the amount of dimerized VP-EL222 complexes.

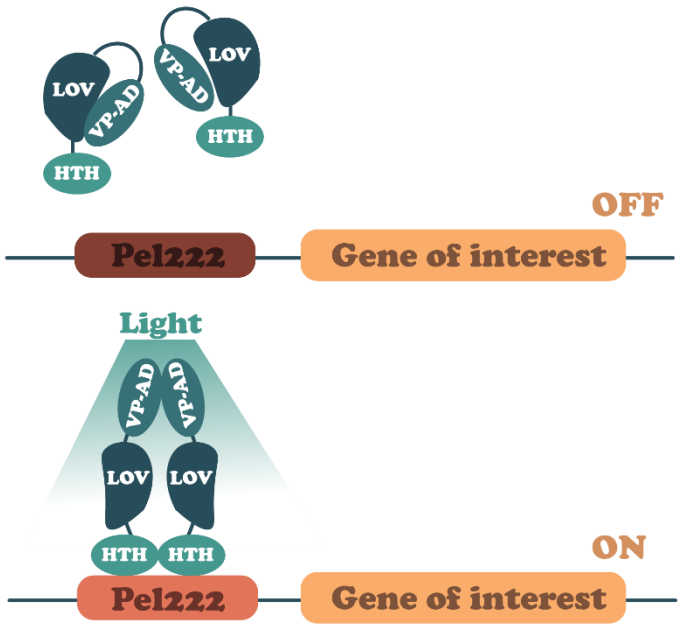


Figure 5. Structure and activation of VP-EL222. VP-EL222, a light-inducible gene expression system, contains a VP16 protein fused to the EL222 TF. In the presence of blue light, the VP-EL222 dimer forms and binds to the pEL222 promoter. This promoter binding activates the transcription of the gene of interest. In the absence of the blue light, the system is non-active, and no transcription takes place.

The more complexes there are, the higher concentration of the protein will be, and vice versa. However, once the blue light signal is depleted, the structural changes are reversed, leading to the inactivation of EL222 TF. The inactivation takes place very rapidly, and it is estimated to be at 11 seconds after the depletion of the blue light (Zoltowski & Gardner, 2011).

Overall, the blue light-inducible VP-EL222 system allows for the light-controlled expression of proteins in a concentration-dependent manner. In addition to this, and thanks to its dynamic expression range, fast activation/deactivation dynamic, and highly linear response to the light induction, the VP-EL222 system can be used to test the effect that different Far1 degrons might have on the cellular concentrations of a protein of interest.

1.8 SCRaMbLE and Cre recombinase

The international Synthetic Yeast Genome Project, Sc2.0, attempts to synthesize fully synthetic yeast chromosomes and study the genome structure, minimal length, and gene content (Annaluru *et al.*, 2014). A central characteristic of the Sc2.0 chromosome is SCRaMbLE

(Synthetic Chromosome Rearrangement and Modification by LoxPsym-mediated Evolution). SCRaMbLE is an inducible evolutionary system that enables recombination, inversion, or deletion of all non-essential yeast genes, hence serving as a genome minimization tool (Dymond & Boeke, 2012). Furthermore, the system allows optimization and development of synthetic pathways used to produce many valuable industrial and pharmaceutical compounds (Keasling, 2010). This optimization is achieved through the elimination or downregulation of interfering metabolic pathways. The SCRaMbLE system is composed of two key components: Cre recombinase protein and LoxP sites.

Cre recombinase is a P1 bacteriophage-derived tyrosine recombinase that uses a topoisomerase I-like mechanism to carry out a series of site-specific recombination events. The recombinase catalysis and carries out the recombination between the two DNA recognition sites (LoxP sites). A LoxP site is 33 base pairs long DNA sequences that contain two 13 base pairs of palindromic sequences separated with eight base pairs long spacer region. Once Cre recombinase recognizes two LoxP sites, it performs a series of different recombination events: inversions, translocations, deletions, and in the presence of new DNA fragments, insertions (Nagy, 2000).

However, the Cre-LoxP system is demonstrated to have some unwanted and potentially toxic effects on the target organism or tissues (Lexow *et al*., 2013). Similarly, the system might result in erroneous recombination, especially during the S and M phases of the cell cycle, where many ongoing events can interfere with its activity. These pieces of evidence go on to show the importance of regulating Cre recombinase's activity to the G1 phase of the cell cycle. This activity restriction to the G1 phase is hypothesized to reduce the toxic effect the system might have and provide a larger variety of accurately executed recombination events.

2 THE AIMS OF THE THESIS

The cell cycle machinery could be used to limit the cellular concentration of any protein during the cell cycle. This controlled expression can be achieved by addition of programmable phosphorylation tags consisting of phosphorylation sites and docking motifs to the desired proteins. The addition of various motifs provides a substrate with a barcode-like tag that enables timely and programmed phosphorylation by the Cdk1 machinery. The N-terminal region (1-150) of Cdk1 inhibitor protein Far1, has been shown to function as a degron. Therefore, in this work we aim to better understand how this motif could be used to control the concentrations of different proteins and its possible advantages in the context of synthetic biology.

The aims of the thesis are:

- to test the Far1 degrons in triggering cell cycle dependent degradation of different proteins using the VP-EL222 transcriptional factor and Cre recombinase as the model substrates;
- to use the synthetic degron tags to fine-tune protein expression levels;
- to minimize the size of the Far1 degron tag to reduce the risk of unspecific effects of tagging;
- to tag the Cre recombinase with Far1 degrons in order to constrain the expression of Cre recombinase to specific cell cycle phases, and test the functionality of the tagged Cre variants to aid strain construction in synthetic biology.

3 EXPERIMENTAL PART

3.1 MATERIALS AND METHODS

3.1.1 Yeast strains

The yeast strains used in this study are based on *w303* and are MATa haploids. The strains are described in Table 1. Strain FC1 contained a plasmid pLM494 that expresses the carotenoid synthesis pathway. pLM494 was a gift from Bernd Müller-Röber (Addgene plasmid # 100539 ; http://n2t.net/addgene:100539 ; RRID:Addgene_100539). The strains and plasmids used in this study were constructed with assistance from other Loog lab members.

Table 1. List of different yeast strains used in the study. The list contains the name of the strain, information on the background strain and a brief description of the strain.

Strain	Background strain	Description
DOM90	w303	*MATa w303, bar1::hisG*
MO172	DOM90	*Padh::Padh-NLSmodule(wt)-mCherry::TRP1*
BY4741	S288C	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0*
FC1	DOM90	*CEN pLM494(pTip1-crtE-tAcs2 + pPgk1-crtI-tAcs2 + pTdh3-crtYB-tCit1 + pZeo1-tHmg1-tAcs2)*
FC2	FC1	*His3: pRS303(Pgal1-Cre-EBD (pSH62-EBD))*
FC3	FC1	*His3:pRS303(Pgal1-Cre-EBD-Far1(85-150-EGFP))*
FC4	FC1	*His3:pRS303(Pgal1-Cre-EBD-Far1(AAxxxA mutant-EGFP))*
FC5	FC1	*His3:pRS303 (Pgal1-Cre-EBD-Far1(85-150 + T3-EGFP)*
FC6	BY4741	*pFA6a-HIS (Pel222-Citrine-tADH)*

		pYM17 (Padh-NLS-NES-(MCM2/3)-mCherry-natNT2)
FC7	FC6	*CEN pRS315 (Pact1-Far1(1-150 lxxxll)-VP-EL222)*
FC8	FC6	*CEN pRS315 (Pact1-Far1(1-150)-VP-EL222)*
FC9	FC6	*CEN pRS315 (Pact1-Far1(1-150 N130A)-VP-EL222)*
FC10	FC6	*CEN pRS315 (Pact1-Far1(1-150 SP RxL)-VP-EL222)*
FC11	FC6	*CEN pRS315 (Pact1-VP-EL222)*
FC12	MO172	*URA3:pRS306 (Psic1-SV40-Far1(1-150 WT)-EGFP)*
FC13	MO172	*URA3:pRS306 (Psic1-SV40-Far1(85-150 WT)-EGFP)*
FC14	MO172	*URA3:pRS306 (Psic1-SV40-Far1(85-150 minimal)-EGFP)*
FC15	MO172	*His3:pRS303(Pgal1-Cre-EBD-Far1(85-150-EGFP))*
FC16	MO172	*His3:pRS303(Pgal1-Cre-EBD-Far1(AAxxxA mutant-EGFP))*
FC17	MO172	*His3:pRS303* (*Pgal1-Cre-EBD-Far1(85-150 + T3)-EGFP)*
FC18	MO172	*CEN* (*Pgal1-Far1(85-150-EGFP)-Cre-EBD)*
FC19	MO172	*CEN* (*Pgal1-Far1(85-150(AAxxA)-EGFP)-Cre-EBD)*

3.1.2 Plasmid constructs

The plasmids used in this study are based on either a centromeric vector pSH62-EBD or an integrative pRS303. pSH62-EBD was a gift from Marc Gartenberg (Addgene plasmid # 49455 ; http://n2t.net/addgene:49455 ; RRID:Addgene_49455).

Table 2. List of plasmids used in the study. The list contains all the plasmids used in this work. It specifies their name, a brief description, a backbone vector, and source.

Name	Description	Vector
pFC1	Pgal1-Cre-EBD (pSH62-EBD)	pSH62-EBD
pFC2	Pgal1-Cre-EBD with N-term NotI/SgsI	CEN HIS
pFC3	Pgal1-Cre-EBD with C-term NotI/SgsI	CEN HIS
pFC4	Pgal1-Far1(85-150-EGFP)-Cre-EBD	CEN HIS
pFC5	Pgal1-Far1(85-150(lxxxll)-EGFP)-Cre-EBD	CEN HIS
pFC6	Pgal1-Cre-EBD-Far1(85-150-EGFP)	CEN HIS
pFC7	Pgal1-Cre-EBD-Far1(85-150(lxxxll)-EGFP)	CEN HIS
pFC8	Pgal1-Cre-EBD-Far1(85-150(Priming SM)-EGFP	CEN HIS
pFC9	Pgal1-Cre-EBD with C-term NotI/SgsI	pRS303
pFC10	Pgal1-Cre-EBD-Far1(85-150-EGFP)	pRS303
pFC11	Pgal1-Cre-EBD-Far1(85-150(lxxxll)-EGFP)	pRS303
pFC12	Pgal1-Cre-EBD-Far1(85-150(Priming SM)-EGFP	pRS303
MO804	Pact1-Far1(1-150)-VP-EL222	pRS315
MO805	Pact1-Far1(1-150 AAxxxA)-VP-EL222	pRS315
MO815	Pact1-Far1(1-150 N130A)-VP-EL222	pRS315
MO816	Pact1-Far1(1-150 SP RxL)-VP-EL222	pRS315
pFC13	Pact1-VP-EL222	pRS315
MO882	Psic1-SV40-Far1(85-150 mini-mal)-EGFP	pRS306
MO833	Psic1-SV40-Far1(85-150)-EGFP	pRS306
MO867	Psic1-SV40-Far1(85-150)-EGFP	pRS306

3.1.2.1 PCR

The inserts used in this study were obtained by PCR of different templates containing the corresponding mutation. The PCR mixture had a final volume of 50 µL. The PCR reagents, as well as their final concentrations, are depicted in **Table 3.** Once the mixture was made, the DNA fragments were amplified in the PCR machine following a three-step program **(Table 4).** The extension time used in the program is determined based on the length of the region to be amplified and by taking into account the synthesis rate of Phusion polymerase (1kb/30s). The annealing temperature is dependent on the composition and length of the primers and was calculated by using Benchling. Once the PCR finished, the mixture was stained with 6x Orange DNA Loading Dye (Thermo Fisher Scientific) and 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)). The gel electrophoresis took place in 1x TAE buffer (40 mM Tris-acetate (pH 8.3), 1 mM EDTA). Afterward, the fragments were observed under UV light. Their length was estimated by using GeneRuler DNA Ladder 1 or DNA Ladder 2 (Thermo Fisher Scientific) that had previously been loaded onto the gel together with the PCR mixture. The bands with the correct size were cut out of the get and placed into separate microcentrifuge tubes. The amplified DNA fragments were purified from the gel using FavorPrep™ GEL/PCR Purification Kit (Favorgen) according to the instructions provided by the manufacturer. Lastly, their concentration was measured using Thermo Fisher's NanoDrop 1000 Spectrophotometer.

Table 3. PCR mixture. PCR components and their final concentrations

Reagent	Final concertation
DNA template	1 ng/µL
5x Phusion HF Buffer	1x
Forward primer	0.5 µM
Reverse primer	0.5 µM
dNTPs	200 µM

High-fidelity Phusion DNA Polymerase	0.02 U/µL
Milli-Q H_2O	Up to 50 µL

Table 4. Three-step PCR program. PCR cycles used to amplify the DNA

Step	Temperature	Time	Number of cycles
Initial Denaturation	98 °C	5 min	1
Denaturation Annealing Extension	98 °C 55 °C 72 °C	20 s 20s 1kb/30s	 32
Final Extension	72 °C	5 min	1

3.1.2.2 Enzyme Digestion

In order to integrate the amplified inserts into the vector, they had to be digested with the corresponding enzymes. During the amplification of DNA inserts, the primers contained the corresponding cutting sites, thereby restricting the inserts with the enzymes of choice. Moreover, to later ligate these inserts into the vector, the vector had to be digested with the same enzymes as the inserts. These digestions were performed using FastDigest™ enzymes. The digestion mixture contained 1-2 ng of vector or 8 µL of PCR amplified insert, 2 µL of 10x FastDigest™, 1 µL of each of FastDigest™ Enzymes, and Milli-Q H_2O that was added to 20 µL. The mixture was incubated at 37 °C for 30-45 minutes.

Once the incubation was over, the mixture containing PCR amplified inserts was further incubated at 80 °C to heat inactivate the enzymes. The inactivation is an important step as otherwise, the enzymes might still be active and re-digest the vector during the ligation. Unlike the insert digestions, the mixture containing the linearized vector was stained with 6x Orange DNA Loading Dye (Thermo Fisher Scientific) and loaded onto 1% Agarose gel. The gel electrophoresis took place in 1x TAE buffer (40 mM Tris-acetate (pH 8.3), 1 mM EDTA). Afterward, the vector was observed under UV light, and its length was estimated by using GeneRuler DNA Ladder 1 or DNA Ladder 2 (Thermo Fisher Scientific). The band

with the correct size was cut out of the gel and placed into separate microcentrifuge tubes. The linearized vector was purified from the gel using FavorPrep™ GEL/PCR Purification Kit (Favorgen) according to the instructions provided by the manufacturer. Later, its concentration was measured using Thermo Fisher's NanoDrop 1000 Spectrophotometer.

3.1.2.3 Ligation

The DNA insert and the vector were ligated together once they were both digested with the same enzymes. The ligation mixture contained around 20 ng of digested vector, 3:1 molar ration over the vector of DNA inserts, 2 µL of 10x T4 DNA Ligase buffer, T4 DNA Ligase (Thermo Fisher Scientific) in a final concentration of 5 U/µL, and Milli-Q H_2O up to 20 µL. The mixture was thoroughly mixed through and incubated at 16 ℃ from 4 to 12 hours.

3.1.2.4 Bacterial transformation and plasmid verification

Bacterial transformation is performed in order to obtain a higher yield of ligated plasmids. A special strain of re-engineered *E. coli* competent cells (Turbo) was used in bacterial transformation. The competent cells were taken out of the -80 ℃ freezer and left to thaw on ice. While the cells were on the ice, 2 µL of ligation mixture was transferred to a new tube. Afterward, 50 µL of bacterial cells were added to the tube and thoroughly re-suspended. The tube containing a re-suspension of ligation mixture and bacterial cells was then placed on ice for 30 minutes. After 30 minutes, the tube was incubated at 42 ℃ for 45 seconds and placed back on the ice for additional 5 minutes. Afterward, 500 mL of LB medium media (10 g/L tryptone (BD Biosciences), 5 g/L yeast extract (Formedium), 10 g/L NaCl) was added to the tube, and the tube was put on a 37 ℃ shaker for one hour. Lastly, the cells were collected by centrifuging the tube for 60 seconds at 6000 rpm and plated on an LB plate containing ampicillin (Sigma), and incubated for not more than 15 hours at 37 ℃.

Once the colonies were visible on the plate, a single colony was taken and grown in a flask containing 4 mL of liquid LB medium in the presence of 100 µg/mL ampicillin. The flasks were then incubated in a 37 ℃ shaker for 8-16 hours. Afterward, the plasmids were extracted using FavorPrep™ Plasmid DNA Extraction Mini Kit (Favorgen) and following the instructions. The concentration of the plasmid was measured with Thermo Fisher's NanoDrop 1000 Spectrophotometer. In order to make sure that the plasmid has the desired insert, the plasmid was digested with the enzymes it was originally restricted with. The digestion mixture contained 2-4 µg of plasmid DNA, 2 µL of 10x FastDigest™ Green Buffer, 2 µL of FastDigest™ enzymes, and Milli-Q H_2O up to 20 µL. The mixture was incubated at 37 ℃ for 30-

45 minutes. Afterward, the mixture was loaded onto 1% Agarose gel. The gel was observed under the UV light, and correctly size bands were estimated using GeneRuler DNA Ladder 1 or DNA Ladder 2 (Thermo Fisher Scientific). The plasmids were sent for Sanger sequencing before being transformed into yeast.

3.1.3 Yeast transformation

Once the correct sequence of the plasmid was confirmed, the plasmid was transformed into a yeast strain. The *S. cerevisiae* strains used in this work were all based on *DOM90* and *MO172* strains **(Table 1).** One day before the transformation, the corresponding strain was taken out of -80℃ stock and streaked onto a fresh YPD plate (10 g/L yeast extract (Formedium), 20 g/L peptones (Formedium), 20 g/L glucose (Oriola)). On the day of transformation, a tip of an inoculation stick was covered with the cells from a fresh YPD plate and transferred into a tube containing 25 mL of YPD media. The inoculum was incubated at 30 ℃ shaker until the optical density reached the range of 0.8-1 (approximately 5-6 hours). The optical density was measured at 600 nm using Ultrospec 10 cell density meter (Amersham Biosciences).

Before the cells were ready for transformation, the non-centromeric plasmids were linearized to be integrated into yeast's genome. The selected enzyme cuts through the amino acid marker gene and allows integrating the plasmid inside the locus through homologous recombination. The restriction contained 1 µg of plasmid DNA, 1 µL of 10x FastDigest™ Buffer, 1 µL of FastDigest™ enzyme, and Milli-Q H_2O up to 10 µL. The restriction was incubated for 30 minutes at 37 ℃. In the case of centromeric (non-integrating) plasmids, they were diluted to 10 ng/mL (50x), and 2 µL was used in the transformation.

Once the yeast cells have grown to the desired OD, they were transferred into a 50 mL tube and centrifuged for 1 minute at 3200 rpm. After the certification, the supernatant was discarded, and the cell pellet was re-suspended in 1 mL of solution I (100 mM lithium acetate in 0.5 x TE (5 mM Tris-HCl (pH 8), 0.5 mM EDTA)). The cells were then centrifuged for 1 minute at 3200 rpm, the supernatant was removed, and the cells were again re-suspended in two times the cell volume of solution I. This mixture was left at room temperature for 10 minutes.

In the meantime, the tube containing plasmid restriction was taken out of the incubator, and 10 µL of boiled Salmon Sperm DNA (SSDNA, pre-boiled at 100 ℃ for 10 minutes) was added. Once the cells had incubated for 10 minutes, 100 mL of the mixture with cells was

added to the tube containing the plasmid and SSDNA. The mixture was re-suspended thoroughly. Afterward, 700 µL of sterile PEG/lithium acetate solution (40% PEG 3350, 100 mM lithium acetate, 10 mM Tris-HCl (pH 8), 1 mM EDTA) was added, followed by 48 µL of DMSO. The mixture was carefully mixed through and incubated at 42 ℃ for 40 minutes. After the incubation, the transformation mixture was heat-shocked on ice for 2 minutes. After the heat shock, the cells were collected by centrifuging the tube for 1 minute at 6000 rpm. The supernatant was removed and cells were washed in 1 mL of 1x TE buffer (10 mM Tris-HCl (pH 8), 1 mM EDTA). The mixture was then centrifuged for the second time. The supernatant was once again removed, and the cells were re-suspended in 200 µL of 1x TE buffer and plated on a plate containing a corresponding selection media. The plates were incubated at 30 ℃ for 2-3 days.

3.1.4 Colony counting and yeast plasmid extraction

β-carotenoid production pathway plasmid (pLM494) contains four genes responsible for β-carotenoid production: *crtE, crtI, crtYB,* and *tHmg1*. The plasmid also has five loxP sites embedded in-between these four genes, allowing Cre recombinase to perform random shuffling or deleting of the genes. To understand the effect of Cre recombinase activity on the β-carotenoid production pathway plasmid, we performed a colony counting experiment and afterward extracted the β-carotenoid producing plasmid from the yeast colonies that did not produce colored colonies. In a strain where there is no Cre activity, the yeast produces β-carotenoids that give a significant orange color to the culture. On the other hand, when there is Cre recombinase activity, the β-carotenoid producing plasmid is recombined on the LoxP sites, resulting in deletion, inversion, or insertion of genes responsible for β-carotenoid production. This, in turn, stops the production of β-carotenoids resulting in white yeast colonies. To prepare for the experiment, *FC1, FC2, FC3* strains **(Table 1)** were taken out of −80 ℃ stock and inoculated in a culture tube containing 2 mL of -HIS -URA (SC-URA/-HIS) raffinose media (7 g/L yeast nitrogen base without amino acids (BD Biosciences), 2 g/L SC-URA SC-HIS powder (MP Biomedicals), 20 g/L raffinose (Oriola)). As the abovementioned strains all have a GAL1 promoter used in the induction of Cre recombinase, they had to be grown on raffinose since glucose has an inhibitory effect on galactose induction. The flasks were left on a 30 ℃ shaker overnight. The next day, the optical density was measured at 600 nm using Ultrospec 10 cell density meter (Amersham Biosciences). The cultures containing different strains were diluted to OD 0.2 and transferred into a new tube. The new culture

tube contained 2 mL of -HIS -URA media and raffinose and galactose sugars in a final concentration of 2%. As Cre recombinase had an estradiol binding domain, the medium also contained β-estradiol in a final concentration of 1 µM (Hochrein *et al*., 2018). The flasks were taken back to a 30 ℃ shaker for 4 hours. After 4 hours, the OD measurement of each flask was taken, and depending on the OD measurement, the number of cells in an individual flask was estimated. Afterward, a series of dilutions were performed to get a solution containing around 100 cells. Later, these diluted solutions were plated on a -HIS -URA glucose plates (7 g/L yeast nitrogen base without amino acids (BD Biosciences), 2 g/L SC-URA SC-HIS powder (MP Biomedicals), 20 g/L glucose (Oriola), 20 g/L agar). The flasks containing the remaining culture were taken back to the 30 ℃ shaker for additional 12 hours. At the 16-hour time point, the flasks were removed from the shaker, and the OD measurements were taken. Afterward, the same set of steps took place as at the 4-hour time point. The plates were incubated at 30 ℃ for two days or until colonies appeared. We counted the number of white vs. the number of orange colonies per plate to estimate the efficiency of Cre recombinase.

Once the cells have grown on the selection plates, we took a couple of white colonies per plate and performed β-carotenoid plasmid extraction. The extraction was performed using Yeast Plasmid Miniprep II (Zymoprep) and following the manufacturer's instructions. Once the plasmid DNA was retrieved, 3-4 µL was transformed into *E. coli* competent cells (Turbo) and later plated onto a plate containing 100 µg/mL ampicillin (Sigma) and incubated overnight at 37 ℃. The next day, a single colony was taken and grown in 4 mL LB medium in the presence of 100 µg/mL ampicillin. The flasks containing the single colonies were incubated at a 37 ℃ shaker for 8-16 hours. Afterward, the plasmids were extracted by using the FavorPrep™ Plasmid DNA Extraction Mini Kit (Favorgen) and following the instructions. The plasmids were restricted with *Pst*I and *Sac*I FastDigest enzymes, and different restriction patterns were observed on 1% Agarose-TAE gel.

3.1.5 Microscopy

In order to measure and observe the levels of GFP/Citrine fluorescence during the cell cycle, we performed time-lapse microscopy. The microscopy experiments allowed real-time changes in the GFP/Citrine levels, indicating either the expression or the degradation of our protein of interest. In this study, we used two different microscopy experimental setups, one for the strains containing different Cre recombinase GFP tagged mutants and the other for the regulation of VP-EL222, which used a pELl222-Citrine-tADH protein as a reporter.

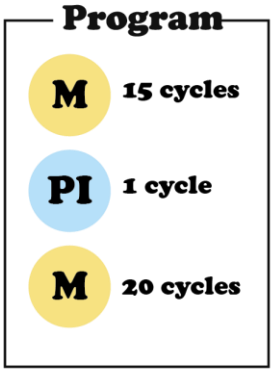
Before starting the microscopy experiment, the colonies containing Cre recombinase that was fused C-terminally to GFP were checked under the microscope to ensure the presence of a GFP signal. Before starting the experiment, the colonies were streaked from the transformation plate onto a new -HIS plate and left overnight at 30 ℃. The next day, a small amount of the cells were transferred to a flask containing 2 mL of SC–HIS 2% raffinose, 2% galactose, 1 µM β-estradiol medium. As the construct was under the Gal1 promoter, galactose was used to induce the expression. The culture was then incubated at 30 ℃ shaker for 5 hours. After the incubation, 0.2 µL of culture were pipetted onto a micro cover glass and covered with a small piece of complete supplement medium (CSM) with 2% raffinose, 2% galactose, 1 µM β-estradiol, and 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)). A smaller cover glass was placed over the gel pieces to cover them, thereby preventing them from moving during the experiment. A plastic cover was placed on the top of everything, thus, effectively keeping the gel pieces from drying out during the experiment. The strains used in the regulation of VP-EL222 experiments followed the same principle as the abovementioned ones. The only difference is that they were grown in SC-LEU 2% glucose medium without the presence of β-estradiol.

The microscopy was conducted using Zeiss Axio Observer.Z1 microscope with 63x/1.4 Oil M27 Plan-Apochromat objective (Zeiss). The experimental setup for Cre recombinase GFP tagged strains included a GFP channel that had an excitation of 15 ms at the wavelength of 470 nm and the mCherry channel with an excitation time of 750 ms at 555 nm. The experiment had 161-time points with 3 minutes between every time point, and it lasted for around 8 hours.

The regulation of VP-EL222 during the cell cycle arrest experiment were performed to see how different mutants behaved in different stages of the cell cycles. The cells were arrested in the G1 phase of the cell cycle with a gel piece containing 10000x α-factor. Other cells were arrested in the S phase of the cycle in presence of a gel piece containing 10x 2mM hydroxyurea (HU). Lastly, nocodazole in 1% DMSO was used to arrest the cells in the M phase of the cell cycle. The setup for the regulation of VP-EL222 cell cycle arrest experiments included two separate blocks. The first block (M) contained a mCherry channel with an excitation time of 450 ms at 555 nm and a Citrine channel with excitation of 10 ms at the wavelength of 510 nm. The second block only had a blue light channel at 470 nm for 1

second (P). The purpose of the second block was to induce the activation of the transcriptional factor, which in turn binds to the pEL222 promoter and promotes the transcription of the citrine gene. The experiment started with 15 cycles of six-minute-long time points (M), during which only the mCherry and Citrine channels were activated. This was followed by one cycle of blue light induction (PI), and another 20 cycles of M **(Table 5)**.

Table 5. VP-EL222 cell cycle arrest setup. The table shows the two different blocks used in the experiment and how they are used from the beginning till the end of the experiment.



Monitoring (M)	PH3	mCherry	Citrine
Excitation time	15 ms	450 ms	10 ms
Excitation wavelength		555 nm	510 nm

Pulse Induction (PI)	Blue light
Excitation time	1000 ms
Excitation wavelength	470 nm

On the other hand, the set up used to measure the cellular concentrations of the VP-EL222 mutants fallowed a similar program as the one for the cell cycle arrest, though there was only one block. The block contained: Citrine channel with excitation of 10 ms at the wavelength of 510 nm, mCherry channel with an excitation time of 450 ms at 555 nm, and a blue light channel at 470 nm for 10 ms. Unlike in the previous experiment, a pulse of blue light was used after every three-minute cycle to induce the expression of Citrine.

The temperature during the experiment was kept at 30 °C with Tempcontrol 37-2 digital (Pecon). On average, 6-10 positions were followed in each experiment and were kept in focus by using the Definite Focus. Upon the finish of the experiment, the program gave three different pictures per one-time point; for the Cre recombinase experiment (phase-contrast, GFP, mCherry) and for VP-EL222 experiments (phase-contrast, Citrinin, mCherry). These pictures would then be converted into grayscale images and analyzed as per (Doncic *et al*., 2013). Later, the graphs depicting different fluorescence levels were plotted using MatLab.

3.2 RESULTS

The N-terminal part (1-150) of Far1 has been shown to bind to the Clb5-Cdk1 complex and promote the phosphorylation of the substrate. The disordered N-terminal part has also been shown to work independently of the rest of the protein and act as a modular degron, thereby promoting the degradation of the encoded substrates. In this study we work with the 1-150 amino acid region of Far1, referred to as the Far1 degron. The wild type of the degron contains a Cbl5 docking motif (NLxxxL), di-phosphodegron (S87 and S91) and multiple different Cks1 binding sites (**Figure 5A**).

3.2.1 Fusion with Far1 degrons control the stability of VP-EL222 transcription factor

To test our hypothesis that different Far1 degron mutants can be used to control the cellular concentrations of different proteins, we set up an experiment using the blue light-inducible VP-EL222 system. Four different Far1 degrons were fused to the transcriptional factor VP-E222, thereby controlling its cellular levels during different cell cycle phases. In the presence of the blue light, the VP-EL222 complex containing different Far1 degrons dimerizes and binds to the pEL222 promoter. Once activated, the pEL222 promoter starts the transcription of the yellow florescent protein (YFP), which remains expressed for as long as there is a blue light pulse. However, once the blue light pulse diminishes, the transcriptional activator becomes inactivated, and the transcription of the YFP stops **(Figure 5B)**. The level of YFP present in the cell corresponds to the amount of Far1-tagged transcriptional activator. For example, suppose the transcriptional activator is already degraded in the S and M phases, and the blue light impulse occurs in the M phase. In that case, there will be no YFP signal detected since there will come no to transcriptional activator present to bind to the promoter and activate the transcription. Therefore, we can use the YFP measurements to get an idea of what amount of transcriptional activator is present in the cell.

Firstly, we wanted to observe the degradation dynamics of different Far1 degrons and test how they function in a different protein context. We fused four Far1 mutants to an eGFP molecule and measured the fluorescence in the time-lapse microscopy. The cell cycle progression was measured using Whi5-mCherry. Whi5 is a transcriptional repressor that is transported out of the nucleus in the late G1 phase, and the start of the cell cycle is defined as a time point when the nuclear levels of Whi5 have decreased by 50% from their peak during the G1 phase (Doncic & Skotheim, 2013).

In microscopy experiments published previously (Faustova *et al*., 2021) it was observed that the Far1 1-150 (wild type) is degraded very rapidly after the start of the cell cycle; it takes around 11 minutes for the nuclear levels to half and around 20 minutes for them to be almost fully depleted **(Figure 5C).** On the other hand, the Far1 **AA**xx**A** mutant takes much longer to degrade: the degradation took almost five times longer than the wild type. This goes on to show the importance the NLxxL motif plays in the phosphorylation of Far1. The N130A mutant that results in partial loss of the NLxxxL docking was a bit slower than the wild type, with a half-life of 20 minutes. The mutant containing an RxL motif instead of the NLxxxL and mutations of the TP sites to SP, leading to loss of Cks1 docking, demonstrated to be more stable than the N130A mutant, with a 50% degradation time of around 35 minutes **(Figure 5C)**. This data shows that it is possible to design different Far1 degron mutants with various degradation profiles throughout the whole cell cycle.

Once we obtained mutants with different degradation profiles, we wanted to check if these mutants can control the expression of YFP inside the cell. To do this, we fused a transcriptional activator VP-EL222 complex with Far1 degrons **(Figure 5A).** We hypothesized that the cellular level of YFP corresponds to the concentration of Far1-degron-VP-EL222 complexes present in the cell. Therefore, a Far1 mutant with a rapid degradation profile (Far1 1-150) will induce a lower expression of YFP compared to a more stable mutant (Far1 1-150 AAxxxA). Time-lapse microscopy experiments proved this hypothesis to be true, as the cellular concentration of YFP induced by the Far1 1-150 mutant was almost five times lower than by the AAxxxA mutant **(Figure 5D).** This data perfectly corresponds to the degradation profiles of the wild type and the AAxxxA mutant; the Far1 1-150 mutant is degraded almost five times faster than the Far1 AAxxxA mutant **(Figure 5C)**.

Similarly, the Far1 RxL SP mutant, whose degradation took almost twice as long as of the Far1 N130A, induced two times higher cellular concentration of YFP than the Far1 N130A **(Figure 5C, 5D).** On the other hand, the VP-EL222 complex with no Far1 degron tag was stable throughout the cell cycle, and it induced a high and constant YFP expression **(Figure 5D).** Therefore, based on the degradation profiles of the Far1-degron tags, one can successfully control the expression of the YFP and its cellular concentration during the cell cycle. Moreover, concentration levels fit perfectly with degradation profiles of corresponding degrons, further supporting our initial hypothesis. Limiting the activity of VP-EL222 in the cell cycle results in different transcriptional factor levels in the cell population and different

YFP levels. As the degron tags can be encoded with different activities, this also enables fine-tuning of the reporter protein expression levels.

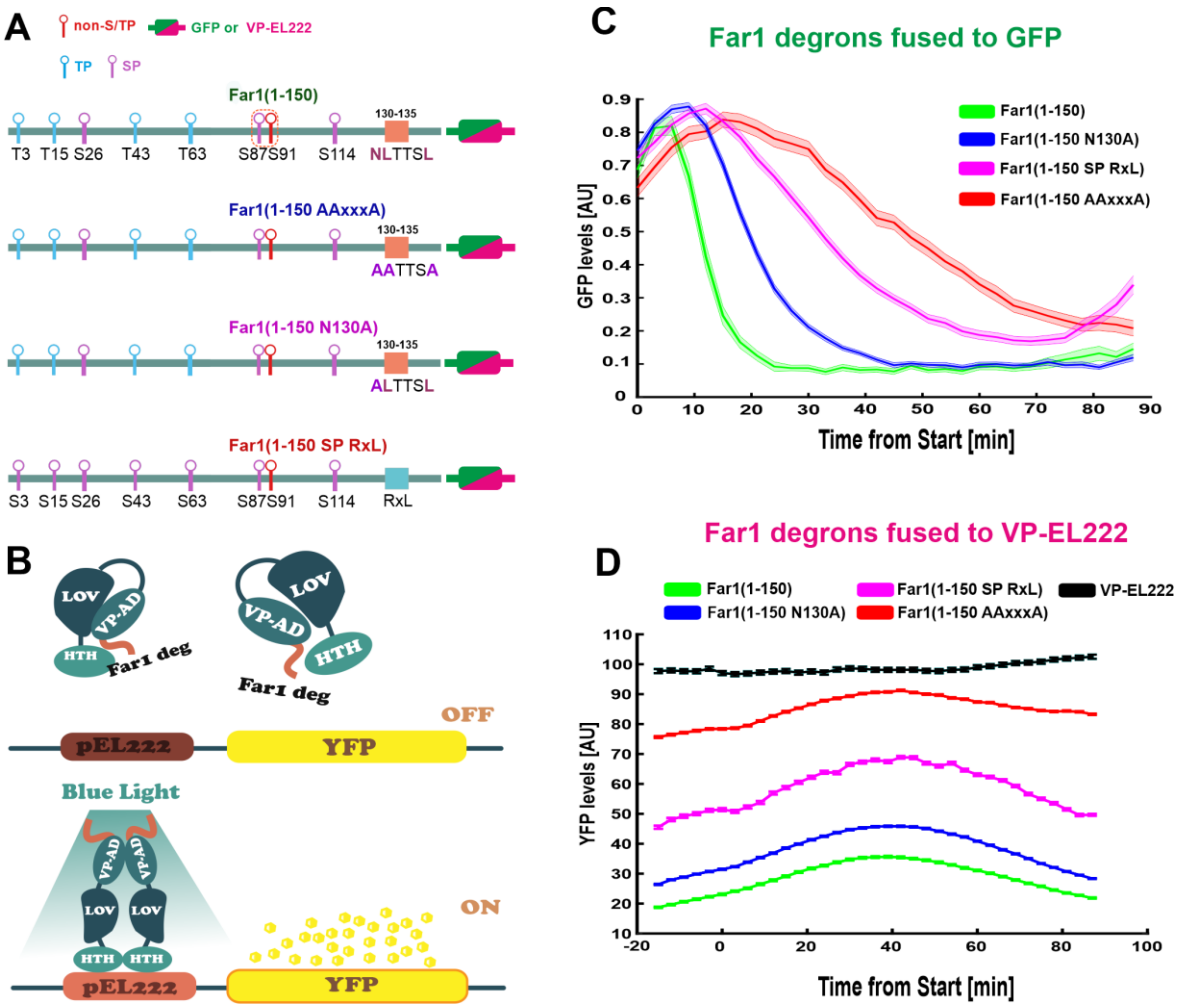


Figure 5. Far1 degrons control the expression of YFP from the VP-EL222-regulated promoter. A) The schematics describe Far1 1-150 wild type, Far1 1-150 AAxxA with alanine mutations in the NLxxxL motif, Far1 1-150 N130A, and Far1 1-150 containing 14 amino acid sequence from Sic1 RXL motif (PKSSVK**RT**L**L**FQFET) integrated into 129 – 142 region and SP (T3S, T15S, T43S, T63S) point mutations. B) The diagram demostrates the mechanism of the blue light-inducible VP-EL2222 system, tagged with the Far1 degrons mentioned in A. C) The graph shows different levels of Far1-GFP mutants during the cell cycle. D) The graph shows different levels of YFP obtained through the blue light-triggered activation with a Far1 tagged VP-E222 transcriptional activator.

However, to understand and test the activity of different Far1 mutants during different cell cycle stages, we performed a series of cell cycle arrest experiments. The main aim of these experiments is to show that the Cdk machinery is responsible for the degradation of the constructs. To do this, we set up an experiment in which cells expressing VP-EL222 tagged

with different Far1 mutants were arrested in three different stages of the cell cycle: G1, S, and M phase. Following the arrest, the VP-EL222 transcription factor was activated by a blue light pulse. All mutants had a high response in the G1 cell cycle arrest and low response in the M phase arrest, suggesting that the increasing Cdk1 activity is responsible for the degradation of different mutants and inactivation of the VP-EL222 in M phase **(Figure 6)**.

Moreover, it can be seen that the concentrations of the YFP induced by the untagged VP-EL222 complex are kept constant during these three phases, whereas the YFP induced by other mutants are decreasing.

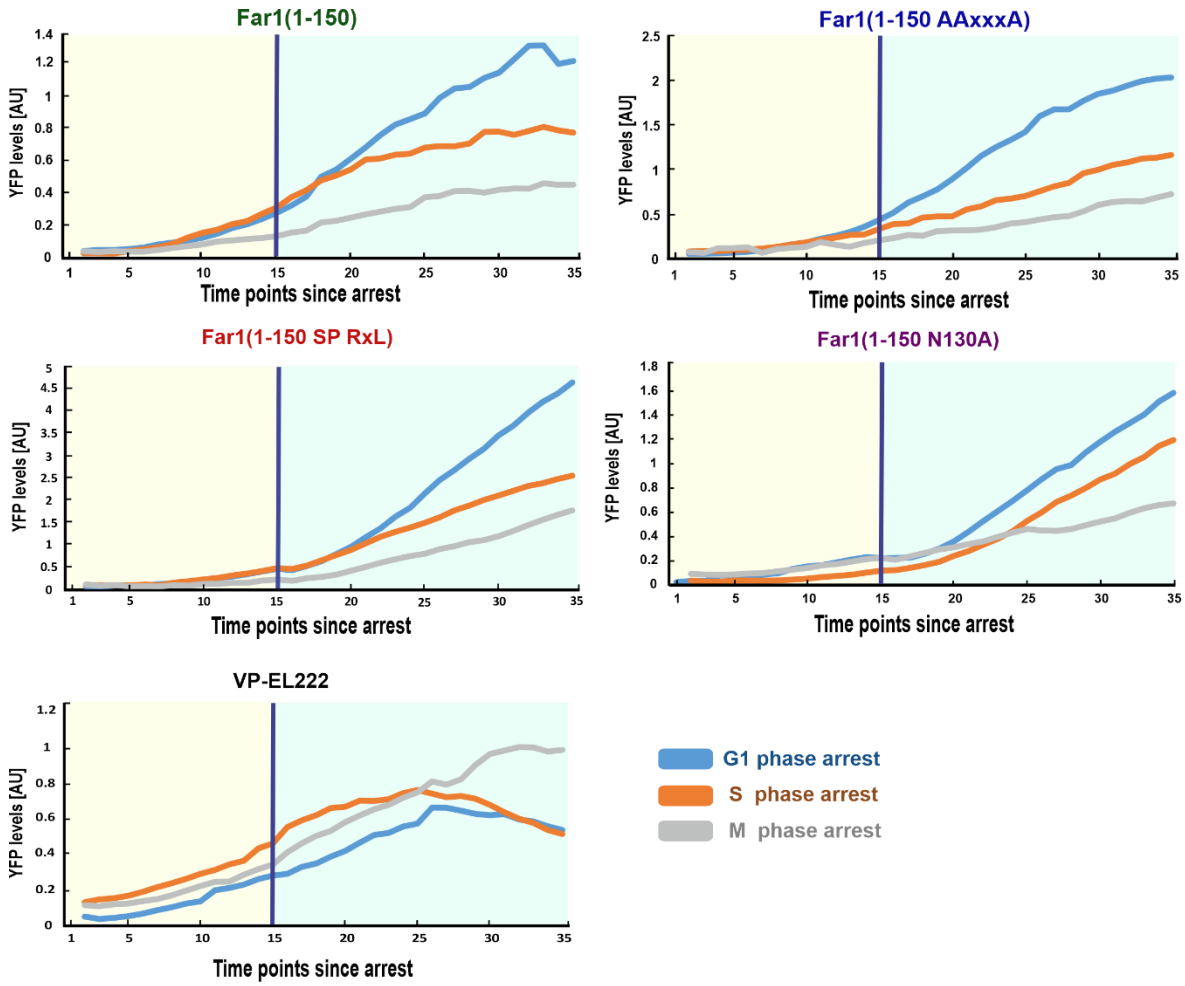


Figure 6. Analysis of the tagged VP-EL222 variants in different cell cycle phases. The figure shows different cell cycle arrest profiles of different mutants. The difference in G1 and M phase YFP response suggests that the increased activity of Cdk1 is responsible for the degradation. The time between the time points is six minutes.

However, due to a noticeable amount of leakage present in the Far1 1-150, AAxxxA, and VP-EL222 constructs, there is a probability that not all cells were arrested in the corresponding cell cycle phase at the time of VP-EL222 activation, thereby increasing the intensity of the YFP signal, especially during the first 15 time points. Therefore, it might be hard to make reliable conclusions about a more specific activity of different mutants arrested in various cell cycle phases. However, the Cdk1 activity is responsible for decreasing the signal suggesting that the Ckd1 machinery is the main controller of the linearly encoded Far1 degron constructs during the cell cycle.

Overall, we have shown that the stability of the VP-EL222 transcriptional factor can be regulated by the cell-cycle machinery by fusion of the protein with the Far1 degron. Importantly, the degradation of different Far1 degron mutants corresponds perfectly to the cellular levels of the reporter protein (YFP). Thus, the Far1 linearly encoded degron can be added to the targeted substrate to promote its timely degradation and control its cellular levels during the cell cycle. These findings help us come one step closer to the design of a protein degron kit that would take advantage of the Cdk machinery to control the concentration of any protein.

3.2.2 Optimization of the Far1 degron tag

As we have previously shown that Fa1 1-150 can function as a modular degron. However, as addition of different tags can affect the functions of the protein in unexpected manner as well, we wanted to see if a smaller 85-150 version of the degron would still be functional. The smaller tag is less likely to have unspecific effects on the protein activity. To decrease the tag size, we constructed an 85-150 Far1 degron containing S87 and S91 degradation sites, as well as the NLxxxL motif. Additionally, to further enhance the degradation, we constructed an 85-150 minimal Far1 degron, which contains a minimal distance between the cyclin docking motif (NLxxxL) phospho-degradation sites. We hypothesized that this version of the degron would be phosphorylated faster, thereby providing a more rapid degradation than the wild type 85-150 **(Figure 7A)**.

As we originally hypothesized, the minimal degron mutant had a half degradation time of around 14 minutes which is faster than the degradation time of the 85-150 mutant, which took around 19 minutes for its concentration to half **(Figure 7B)**. This shows that we can produce a degron of higher efficiency by decreasing the distance between the docking motif and the degradation site. Nonetheless, the 85-150 minimal degron mutant was still a bit

slower compared to the wild type, suggesting that the Cks1 mediated multisite phosphorylation can further shorten the construct's degradation time.

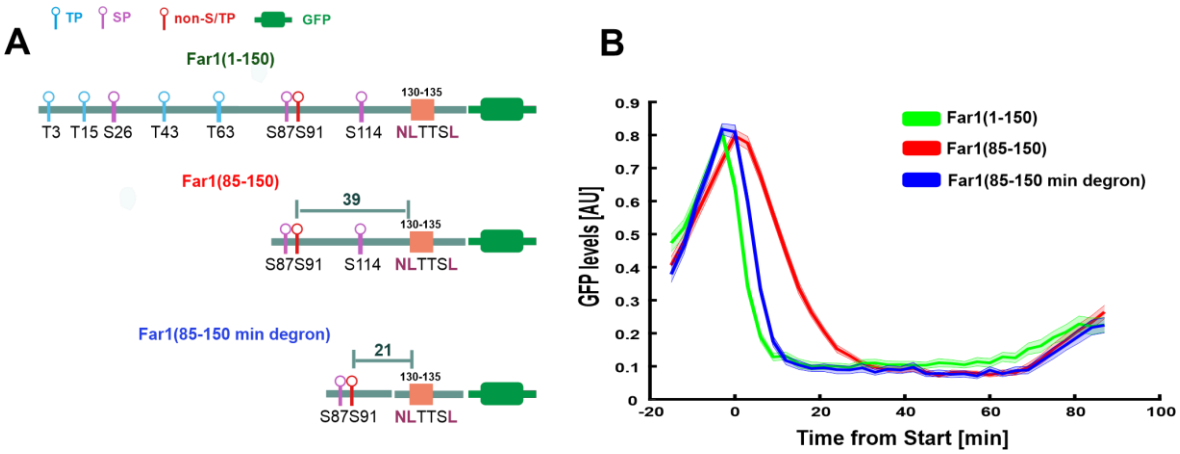


Figure 7. 85-150 minimal mutants degraded faster than 85-150 Far1. A) The schematics show Far1 1-150 wild type, Far1 85-150, and Far1 85-150 with reduced distance between S91 and N130 sites. B) The graph shows different levels of Far1-GFP mutants during the cell cycle.

Thus, we have developed a short degron that functions almost just as efficiently as the Far1 1-150 wild type. This shows that by manipulating the distance between the motif and the phospho-sites and introducing the Cks1 binding domains, we can obtain degrons with considerably different degradation dynamics and fine-tune their concentrations during the cell cycle.

3.2.3 Far1-tagged Cre recombinase activity is limited to the G1 phase

Once we had enough data on the degradation dynamics of different Far1 85-150 degrons, we wanted to use them to improve the function of an existing system. Cre-LoxP system was a suitable candidate as it was reported that the prolonged activity of Cre recombinase outside of the G1 phase can have a toxic effect on the cells and might cause erroneous recombination events during the S phase of the cycle. With this in mind, we decided to construct a Cre enzyme tagged with Far1 degrons hoping that the activity of Cre would be limited to the G1 phase of the cell cycle. The original Cre plasmid contained a GAL1 inducible promoter and an estradiol binding domain (EBD) fused to the C terminus of Cre. GAL1 promoter was used to induce the expression of the enzyme, and the EBD functioned as a nuclear localization signal that transported the recombinase from the cytoplasm to the nucleus.

First, we tagged both the N and the C terminus of the Cre recombinase with the two versions of the Far1 degron: 85-150 wild type and 85-150 AAxxxA mutant. We wanted to see if the enzyme would still be degraded as expected based on the previous data. Interestingly, we observed that only the C-terminally tagged Cre functioned as expected **(Figure 8C)**, whereas the N-terminally tagged enzyme gave a rather noisy signal. Therefore, we decided to use the C-terminally tagged constructs in our further work **(Figure 8A)**.

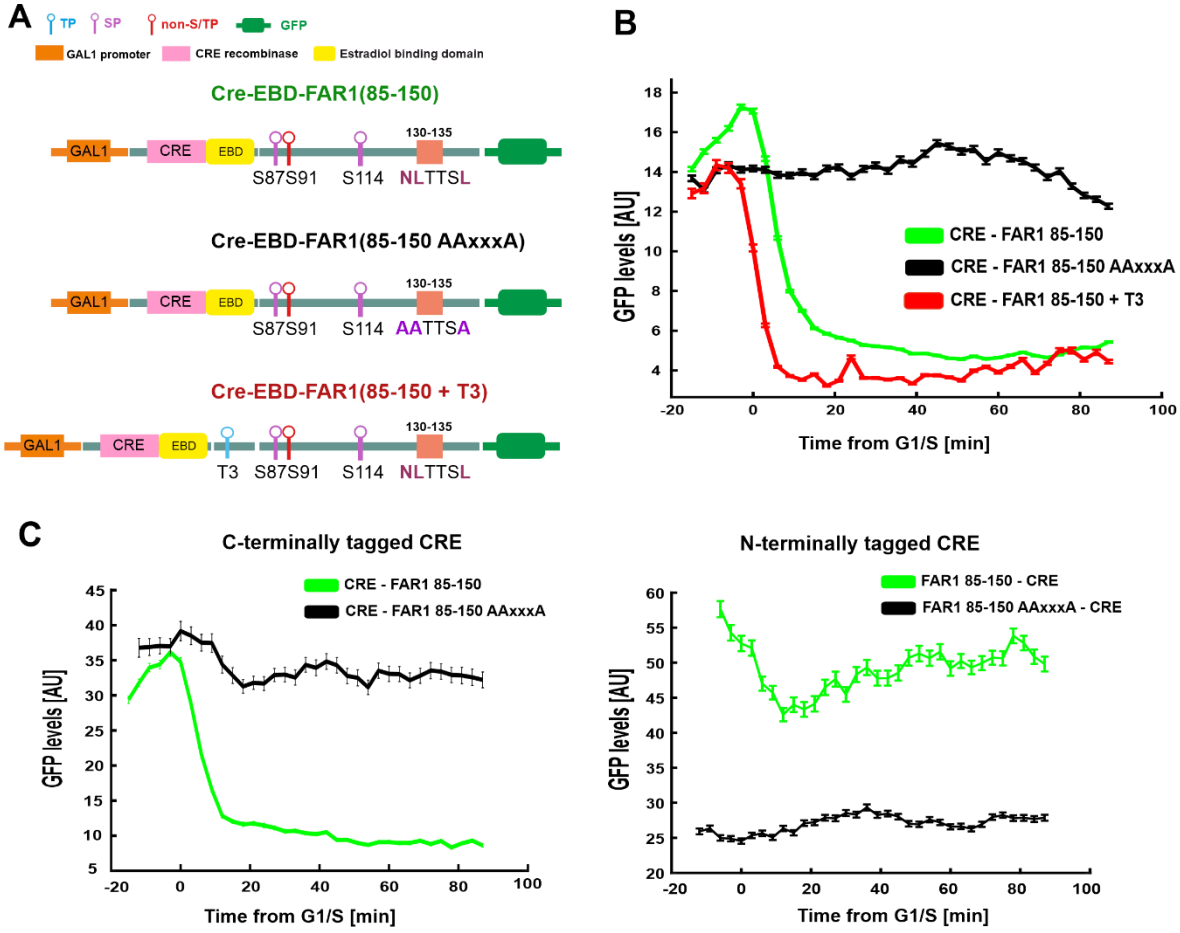


Figure 8. The activity of Cre recombinase can be restricted to the G1 phase with Far1 degrons. A) The scheme of Cre-EBD-Far1(85-150), Cre-EBD-Far1(85-150 AAxxxA), and Cre-EBD-Far1(85-150 + T3) constructs with GAL1 promoter and GFP. B) The graph shows the level of Cre with different C-terminally added Far1-GFP mutants mentioned in panel A during the cell cycle. C) The graphs showing the degradation profiles of C- and N-terminally tagged constructs.

Overall, we observed that the Cre recombinase tagged with the Far1 85-150 behaved as expected and had a degradation half-life of around 15 minutes. Furthermore, in the absence of a functional docking site, the Far1 85-150 AAxxxA mutant was stable during the whole cell cycle. Lastly, due to the addition of the Cks1 binding site T3, in the Cre-Far1(85-150 + T3) mutant, we observed even faster degradation; most of the protein was degraded in 15 minutes after G1/S transition **(Figure 8B)**. Thus, we have successfully show that by tagging the Cre recombinase with two different versions of the Far1 85-150 degron, we can fully control the activity of the enzyme and restrict it to more favorable G1 phase.

3.2.4 Far1 tagging increases the activity of the Cre recombinase

Lastly, we wanted to test our newly developed Cre constructs in a system containing different LoxP sites, thereby checking any potential improvements or defects in Cre recombinase activity. To do this, we took a yeast strain with a β-carotenoid production pathway plasmid and transformed it with our mutants **(Figure 9A).** The plasmid contained four β-carotenoid producing genes (*crtE, crtI, crtYB,* and *tHmg1*) separated by five loxP sites. When there is no Cre recombinase activity present, the pathway is expressed and β-carotenoids are produced, which gives a characteristic orange color to the cells. However, in the presence of Cre recombinase, different recombination events take place, resulting in the deletions or inversions of the genes. These recombination events prevent the production of β-carotenoids, giving the cells their original white color.

First, we set up a colony counting-based assay to understand if there are any changes in the effectiveness of Cre recombinase caused by the addition of different degrons. In addition to the Far1(85-150) and Far1(85-150 AAxxxA) degron-tagged mutants, we used the original Cre construct as a control **(Figure 9A)**. We postulated that the non-tagged Cre construct would have a similar activity level and the Far1(85-150 AAxxxA) mutant, as they should both be active during the whole cell cycle. Moreover, we hypothesized that the Far1(85-150) mutant, which is active only in the G1 phase of the cycle, will demonstrate lower activity and induce fewer recombination events.

However, to our surprise, we observed the opposite happening. One hour after the induction of Cre recombinase with 0.2% galactose and 0.1 μM β-estradiol, the Cre tagged with Far1(85-150) mutant was more active than the non-tagged Cre protein, resulting in around 20.6% of the colonies turning white, while only 8.8% of colonies were white after one hour in the non-tagged version of Cre **(Figure 9B)**.

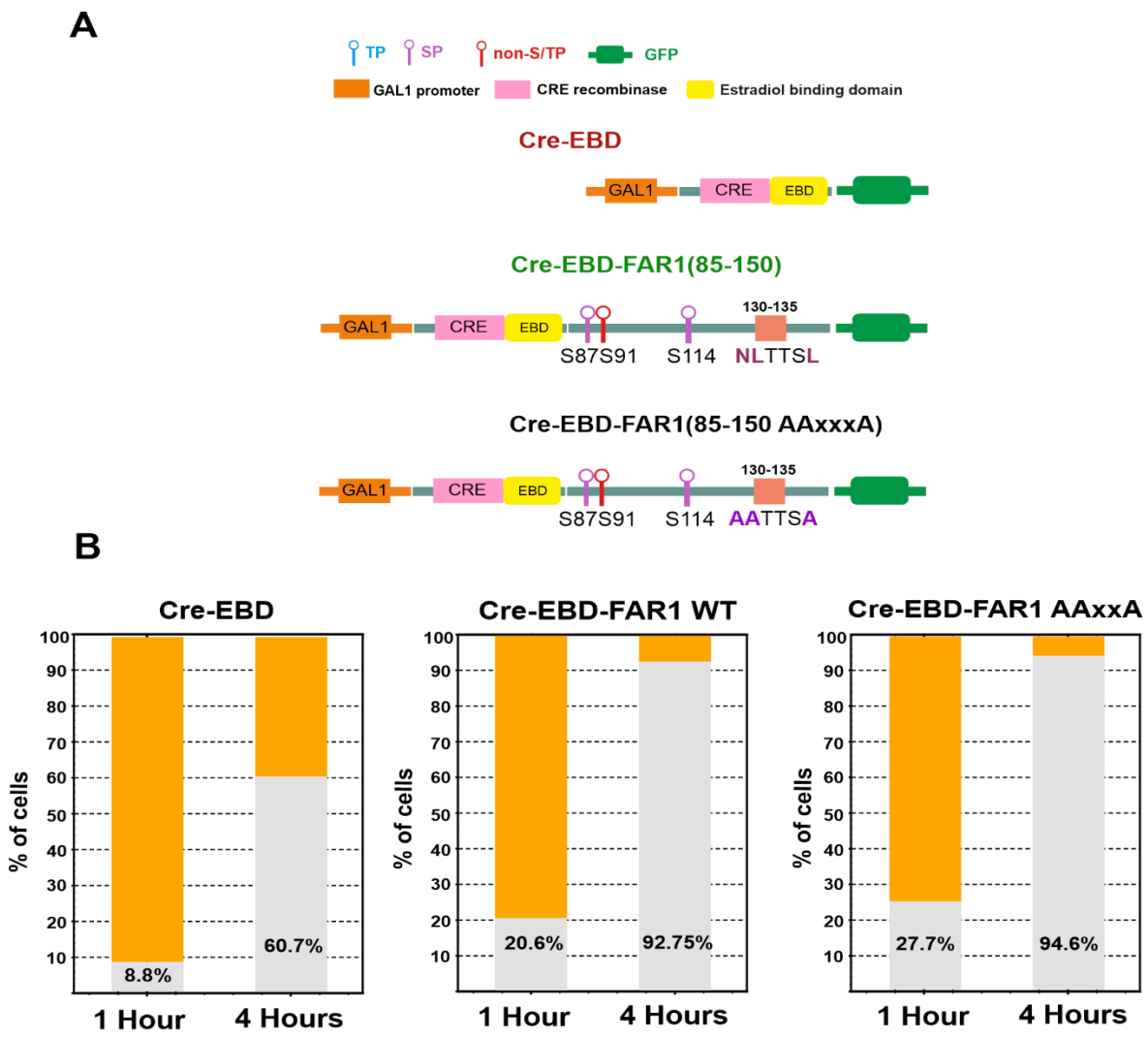


Figure 9. Far1-tagged Cre recombinase initiates more recombination events. A) A panel of different constructs used in the experiment. B) Different recombination experiments with constructs mentioned in A (incubation times of 1 h and 4 h). The cells were grown in -HIS -URA medium overnight and then diluted to the OD 0.2. After induction by addition of 0.2% Galactose and β-estradiol (0.1 μM final concentration), samples were grown for further 1 and 4 h at 30 °C and 230 rpm. The samples were then diluted and plated in -HIS -URA plates. The non-induced controls (not on the graph) yielded more than 98% of the white colonies, showing no leakage in the system.

What was even more surprising was that both the Cre constructs with Far1(85-150) and Far1(85-150 AAxxxA) tags induced recombination events in 92.75% and 94.6% of cells, respectively, four hours after the induction. This suggests that even though Cre recombinase activity was restricted in the G1 phase (Far1 85-150 mutant), the recombinase was still able to trigger more recombination events compared to its non-tagged version. In addition to this,

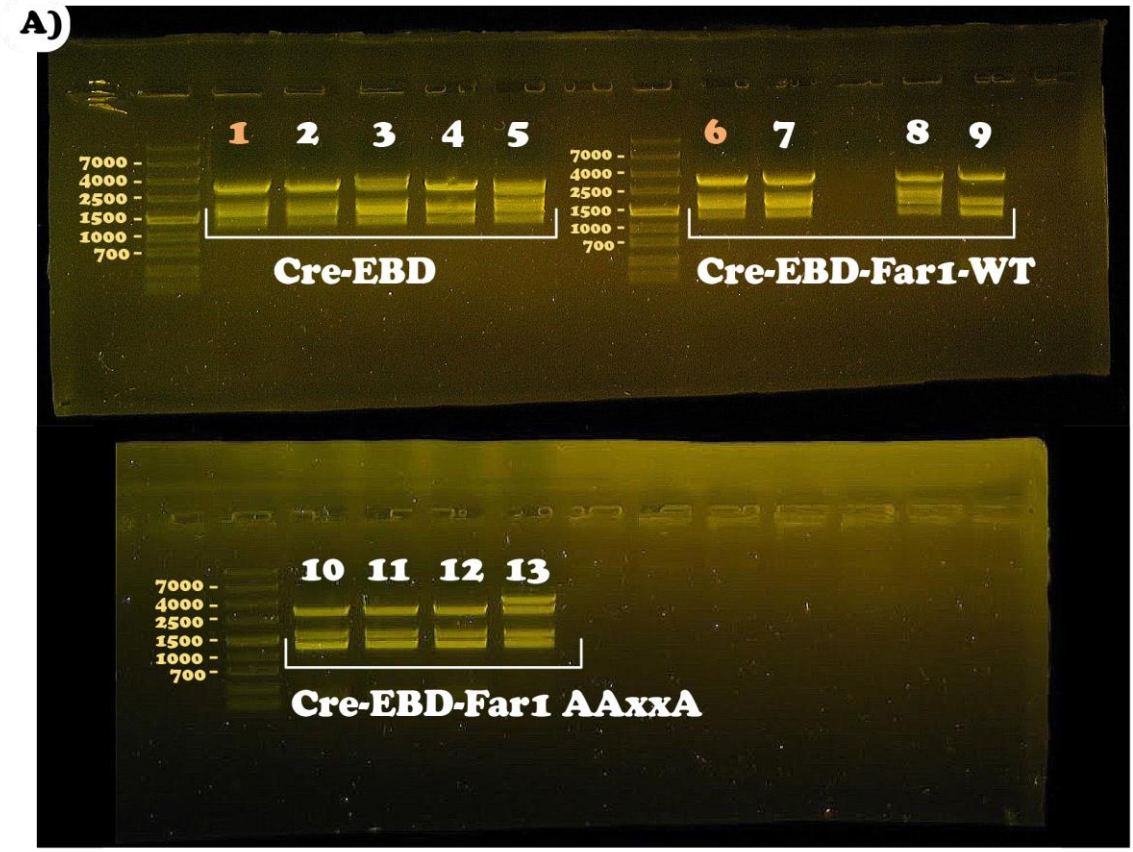
it is worth noting that the number of white colonies in this experiment is likely an underestimate since the β-carotenoid producing plasmid has 4-5 copies within the cell. The recombination needs to happen in all of the plasmids for a colony to be white.

Moreover, the integrative pRS306 plasmid, used in these experiments, is a multi-copy plasmid suggesting that the there might be more than one copy of Cre recombinase in a cell. With this in mind, it could be possible that even though Cre recombinase activity is lowered, there is still a lot of it present in the cell for effect to be seen. A proper setup would have included a plasmid in which the recombinase can only be integrated once.

Nonetheless, the improved activity suggests that the recombinase is working more efficiently to induce different recombination events. However, our main goal is to have different recombination products, thereby achieving variety. To gain a better understanding Cre activity, we must look into the recombination events that took place inside the β-carotenoid producing plasmid to draw reliable conclusions.

The β-carotenoid producing plasmid contains *SacI* and *PstI* restriction sites in between the genes. Depending on the restriction pattern, we can conclude what type of recombination event had taken place through a simple gel analysis. The main goal of a Cre-LoxP system, especially in the context of SCRaMbLE, is to induce different recombinant events, resulting in high variety between different cells. We observed that Cre-EBD gave the greatest variety of plasmid recombination events, suggesting that the low recombinase activity benefits diverse recombination outcomes **(Figure 10A, 10B)**. Similarly, Cre-EBD-Far1 85-150 mutant, which displayed higher activity in the first hour after the induction, gave reasonably diverse outcomes. Lastly, the most active Cre-EBD-Far1 85-150 AAxxxA mutant did not give much plasmid variety, and in three out of four plasmids, the recombination deleted all the carotenoid-producing genes.

Thus, it appears that the lower the activity of the enzyme is, the more diverse the recombination outcomes are. Interestingly, Cre tagged with wild type Far1 85-150 degron resulted in higher variation of recombination events, possibly due to lower amount of this Cre variant compared to the one tagged with the mutated Far1 degron. Importantly, the Cre activity tests show that the Far1-tagged Cre variants are functional, their activity can be limited to specific cell cycle phases and thus they have potential to improve the SCRaMbLE system.



B)

pLM494 – loxP – crtE – loxP – crtI – loxP – crtYB – loxP – tHMG1 – loxP	1,6
loxP	4,9,10,11,12
loxP – crtE – loxP	13
loxP – crtYB – loxP	2,7
loxP – tHMG1 – loxP	5,8
loxP – crtE – loxP – crtI – loxP	3

Figure 10. Cre-EBD and Cre-EBD-Far1 WT induce more variety in recombination products. A) Gel restriction patterns of pLM494 after Cre activity. Plasmids were extracted after 1h time point and digested with *PstI* and *SacI*. The orange colonies are noted in orange numbers. B) A table showing different recombination events. Based on the restriction patterns, we can estimate the event that has taken place. As we have not sequenced plasmids, it is only possible to estimate the events that caused a gene deletion and not inversion.

SUMMARY

This work shows that the Cdk1 machinery can be used to control the expression and concentration of the desired protein in different phases of the cell cycle. A barcode-like tag, consisting of different phosphorylation clusters and cyclin docking motifs, needs to be fused with the desired protein for the Cdk1 machinery to read and trigger degradation of the protein at the pre-programmed time. This mechanism can used in the construction of the industrial yeast strains believed to benefit from a timely controlled expression of their metabolic pathways. In addition to this, the 1-150 amino acid region of the Far1 behaves as a modular degron that can be designed with very different degradation profiles.

To summarize, by fusing the light-inducible VP-EL222 transcriptional factor with different Far1 degron mutants, we were able to control the concentrations of the reporter protein (YFP). The concentration levels were substantially different, depending on the mutant used, suggesting that Far1 degrons can successfully control cellular protein levels and their degradation. Furthermore, we have also shown that depending on the degron, the degradation of the protein can be programmed to a specific phase of the cell cycle, further demonstrating the versatility and accuracy of our model.

In addition to this, we have constructed a minimal 85-150 Far1 degron containing the optimal distance between the docking motif and the phospho-degron sites. The degron was almost as efficient as the Far1 1-150 wild type, suggesting that it is possible to minimize the degron tags.

Lastly, we showed that the Cre recombinase was restricted to the G1 phase of the cell cycle and hypothesized that this activity restriction would benefit the outcome of the recombination events and lower the toxicity in SCRaMbLE experiments. We found that addition of Far1 degron tags efficiently limited the expression of Cre recombinase to the G1 phase. Further, we confirmed that the Far1-tagged Cre variants are functional and seem to have even higher activity than the untagged variant. This setup can be introduced into a synthetic SCRaMbLE strain. This will allow us to understand the effect that the G1-constrained activity has on a strain as a whole.

In conclusion, this work is as a good starting point in designing a toolbox that uses the Cdk and the cell cycle machinery to control the expression and the concentration of any linearly encoded protein. Furthermore, through the use of different degrons, one can obtain very accurate expression and degradation profiles.

REFERENCES

Amon, A., Irniger, S., & Nasmyth, K. (1994). Closing the cell cycle circle in yeast: G2 cyclin proteolysis initiated at mitosis persists until the activation of G1 cyclins in the next cycle. *Cell*, *77*(7), 1037–1050. https://doi.org/10.1016/0092-8674(94)90443-X

Ang, X. L., & Wade Harper, J. (2005). SCF-mediated protein degradation and cell cycle control. *Oncogene*, *24*(17), 2860–2870. https://doi.org/10.1038/sj.onc.1208614

Annaluru, N., Muller, H., Mitchell, L. A., Ramalingam, S., Stracquadanio, G., Richardson, S. M., Dymond, J. S., Kuang, Z., Scheifele, L. Z., Cooper, E. M., Cai, Y., Zeller, K., Agmon, N., Han, J. S., Hadjithomas, M., Tullman, J., Caravelli, K., Cirelli, K., Guo, Z., … Chandrasegaran, S. (2014). Total synthesis of a functional designer eukaryotic chromosome. *Science (New York, N.Y.)*, *344*(6179), 55–58. https://doi.org/10.1126/science.1249252

Benzinger, D., & Khammash, M. (2018). Pulsatile inputs achieve tunable attenuation of gene expression variability and graded multi-gene regulation. *Nature Communications*, *9*(1), 3521. https://doi.org/10.1038/s41467-018-05882-2

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., Lafourcade, C., Longaretti, C., Deshaies, R. J., & Peter, M. (2000). Nuclear-specific degradation of Far1 is controlled by the localization of the F-box protein Cdc4. *The EMBO Journal*, *19*(22), 6085 LP – 6097. https://doi.org/10.1093/emboj/19.22.6085

Dahmann, C., Diffley, J. F. X., & 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*, *5*(11), 1257–1269. https://doi.org/10.1016/S0960-9822(95)00252-1

Delgado, A., & Porcar, M. (2013). Designing de novo: interdisciplinary debates in synthetic biology. *Systems and Synthetic Biology*, *7*(1–2), 41–50. https://doi.org/10.1007/s11693-013-9106-6

Denby, C. M., Li, R. A., Vu, V. T., Costello, Z., Lin, W., Chan, L. J. G., Williams, J., Donaldson, B., Bamforth, C. W., Petzold, C. J., Scheller, H. v, Martin, H. G., & Keasling,

J. D. (2018). Industrial brewing yeast engineered for the production of primary flavor determinants in hopped beer. *Nature Communications*, *9*(1), 965. https://doi.org/10.1038/s41467-018-03293-x

Doncic, A., Eser, U., Atay, O., & Skotheim, J. M. (2013). An Algorithm to Automate Yeast Segmentation and Tracking. *PLoS ONE*, *8*(3), e57970. https://doi.org/10.1371/journal.pone.0057970

Doncic, A., & Skotheim, J. M. (2013). Feedforward Regulation Ensures Stability and Rapid Reversibility of a Cellular State. *Molecular Cell*, *50*(6), 856–868. https://doi.org/10.1016/J.MOLCEL.2013.04.014

Dymond, J., & Boeke, J. (2012). The Saccharomyces cerevisiae SCRaMbLE system and genome minimization. *Bioengineered Bugs*, *3*(3), 168–171. https://doi.org/10.4161/bbug.19543

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

Faustova, I., Bulatovic, L., Matiyevskaya, F., Valk, E., Örd, M., & Loog, M. (2021). A new linear cyclin docking motif that mediates exclusively S-phase CDK-specific signaling. *The EMBO Journal*, *40*(2), e105839. https://doi.org/10.15252/embj.2020105839

Feldman, R. M., Correll, C. C., Kaplan, K. B., & Deshaies, R. J. (1997). A complex of Cdc4p, Skp1p, and Cdc53p/cullin catalyzes ubiquitination of the phosphorylated CDK inhibitor Sic1p. *Cell*, *91*(2), 221–230. https://doi.org/10.1016/s0092-8674(00)80404-3

Gartner, A., Jovanović, A., Jeoung, D.-I., Bourlat, S., Cross, F. R., & Ammerer, G. (1998). Pheromone-Dependent G 1 Cell Cycle Arrest Requires Far1 Phosphorylation, but May Not Involve Inhibition of Cdc28-Cln2 Kinase, In Vivo. *Molecular and Cellular Biology*, *18*(7), 3681–3691. https://doi.org/10.1128/mcb.18.7.3681

Hershko, A., & Ciechanover, A. (1998). THE UBIQUITIN SYSTEM. *Annual Review of Biochemistry*, *67*(1), 425–479. https://doi.org/10.1146/annurev.biochem.67.1.425

Hochrein, L., Mitchell, L. A., Schulz, K., Messerschmidt, K., & Mueller-Roeber, B. (2018). L-SCRaMbLE as a tool for light-controlled Cre-mediated recombination in yeast. *Nature Communications*, *9*(1), 1931. https://doi.org/10.1038/s41467-017-02208-6

Huala, E., Oeller, P. W., Liscum, E., Han, I.-S., Larsen, E., & Briggs, W. R. (1997). Arabidopsis NPH1: A Protein Kinase with a Putative Redox-

Sensing Domain. *Science*, *278*(5346), 2120 LP – 2123. https://doi.org/10.1126/science.278.5346.2120

Jackson, L. P., Reed, S. I., & Haase, S. B. (2006). Distinct Mechanisms Control the Stability of the Related S-Phase Cyclins Clb5 and Clb6. *Molecular and Cellular Biology*, *26*(6), 2456 LP – 2466. https://doi.org/10.1128/MCB.26.6.2456-2466.2006

Keasling, J. D. (2010). Manufacturing Molecules Through Metabolic Engineering. *Science*, *330*(6009), 1355 LP – 1358. https://doi.org/10.1126/science.1193990

Koepp, D. M., Schaefer, L. K., Ye, X., Keyomarsi, K., Chu, C., Harper, J. W., & Elledge, S. J. (2001). Phosphorylation-dependent ubiquitination of cyclin E by the SCFFbw7 ubiquitin ligase. *Science (New York, N.Y.)*, *294*(5540), 173–177. https://doi.org/10.1126/science.1065203

Kõivomägi, M., Örd, M., Iofik, A., Valk, E., Venta, R., & Faustova, I. (2013). Multisite phosphorylation networks as signal processors for Cdk1. *Nat Struct Mol Biol.*, *20*(12), 1–24. https://doi.org/10.1038/nsmb.2706.Multisite

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

Lanker, S., Valdivieso, M. H., & Wittenberg, C. (1996). Rapid Degradation of the G₁ Cyclin Cln2 Induced by CDK-Dependent Phosphorylation. *Science*, *271*(5255), 1597 LP – 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. *The Journal of Cell Biology*, *120*(6), 1305 LP – 1320. https://doi.org/10.1083/jcb.120.6.1305

Lexow, J., Poggioli, T., Sarathchandra, P., Santini, M. P., & Rosenthal, N. (2013). Cardiac fibrosis in mice expressing an inducible myocardial-specific Cre driver. *Disease Models & Mechanisms*, *6*(6), 1470–1476. https://doi.org/10.1242/dmm.010470

Lin, G. G., & Scott, J. G. (2012). The Cdk1-APC/C cell cycle oscillator circuit functions as a timedelayed, ultrasensitive switch. *Nat Cell Biol.*, *100*(2), 130–134. https://doi.org/10.1016/j.pestbp.2011.02.012.Investigations

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

Lowe, E. D., Tews, I., Cheng, K. Y., Brown, N. R., Gul, S., Noble, M. E. M., Gamblin, S. J., & Johnson, L. N. (2002). Specificity Determinants of Recruitment Peptides Bound to Phospho-CDK2/Cyclin A,. *Biochemistry*, *41*(52), 15625–15634. https://doi.org/10.1021/bi0268910

Marini, N. J., & Reed, S. I. (1992). Direct induction of G1-specific transcripts following reactivation of the Cdc28 kinase in the absence of de novo protein synthesis. *Genes and Development*, *6*(4), 557–567. https://doi.org/10.1101/gad.6.4.557

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 &Amp; Molecular Biology*, *20*, 1407. https://doi.org/10.1038/nsmb.2707

McLean, J. R., Chaix, D., Ohi, M. D., & Gould, K. L. (2011). State of the APC/C: organization, function, and structure. *Critical Reviews in Biochemistry and Molecular Biology*, *46*(2), 118–136. https://doi.org/10.3109/10409238.2010.541420

Mendenhall, M. D., & Hodge, A. E. (1998). Regulation of Cdc28 cyclin-dependent protein kinase activity during the cell cycle of the yeast Saccharomyces cerevisiae. *Microbiology and Molecular Biology Reviews : MMBR*, *62*(4), 1191–1243. http://www.ncbi.nlm.nih.gov/pubmed/9841670%0Ahttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC98944

Morgan, D. O. (2007). *The Cell Cycle: Principles of Control* (E. Lawrence, Ed.). New Science Press Ltd.

Motta-Mena, L. B., Reade, A., Mallory, M. J., Glantz, S., Weiner, O. D., Lynch, K. W., & Gardner, K. H. (2014). An optogenetic gene expression system with rapid activation and deactivation kinetics. *Nature Chemical Biology*, *10*(3), 196–202. https://doi.org/10.1038/nchembio.1430

Nagy, A. (2000). Cre recombinase: The universal reagent for genome tailoring. *Genesis*, *26*(2), 99–109. https://doi.org/https://doi.org/10.1002/(SICI)1526-968X(200002)26:2<99::AID-GENE1>3.0.CO;2-B

Nash, A. I., McNulty, R., Shillito, M. E., Swartz, T. E., Bogomolni, R. A., Luecke, H., & Gardner, K. H. (2011). Structural basis of photosensitivity in a bacterial light-oxygen-voltage/helix-turn-helix (LOV-HTH) DNA-binding protein. *Proceedings of the National Academy of Sciences*, *108*(23), 9449–9454. https://doi.org/10.1073/pnas.1100262108

Örd, M., & Loog, M. (2019). How the cell cycle clock ticks. *Molecular Biology of the Cell*, *30*(2), 169–172. 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 & Molecular Biology*, *26*(7), 649–658. https://doi.org/10.1038/s41594-019-0256-4

Örd, M., Puss, K. K., Kivi, R., Möll, K., Ojala, T., Borovko, I., Faustova, I., Venta, R., Valk, E., Kõivomägi, M., & Loog, M. (2020). Proline-Rich Motifs Control G2-CDK Target Phosphorylation and Priming an Anchoring Protein for Polo Kinase Localization. *Cell Reports*, *31*(11), 107757. https://doi.org/https://doi.org/10.1016/j.celrep.2020.107757

Peng, B., Plan, M. R., Carpenter, A., Nielsen, L. K., & Vickers, C. E. (2017). Coupling gene regulatory patterns to bioprocess conditions to optimize synthetic metabolic modules for improved sesquiterpene production in yeast. *Biotechnology for Biofuels*, *10*(1), 43. https://doi.org/10.1186/s13068-017-0728-x

Pines, J. (1995). Cyclins and cyclin-dependent kinases: a biochemical view. *Biochemical Journal*, *308*(3), 697–711. https://doi.org/10.1042/bj3080697

Pope, P. A., Bhaduri, S., & Pryciak, P. M. (2014). Regulation of Cyclin-Substrate Docking by a G1 Arrest Signaling Pathway and the Cdk Inhibitor Far1. *Current Biology*, *24*(12), 1390–1396. https://doi.org/10.1016/J.CUB.2014.05.002

Pyne, M. E., Narcross, L., Melgar, M., Kevvai, K., Mookerjee, S., Leite, G. B., & Martin, V. J. J. (2018). An Engineered Aro1 Protein Degradation Approach for Increased cis,cis-Muconic Acid Biosynthesis in Saccharomyces cerevisiae *Applied and Environmental Microbiology*, *84*(17), e01095-18. https://doi.org/10.1128/AEM.01095-18

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 LP – 562. https://doi.org/10.1126/science.272.5261.560

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*, *95*(18), 10453 LP – 10458. http://www.pnas.org/content/95/18/10453.abstract

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

Seufert, W., Futcher, B., & Jentsch, S. (1995). Role of a ubiquitin-conjugating enzyme in degradation of S- and M-phase cyclins. *Nature*, *373*(6509), 78–81. https://doi.org/10.1038/373078a0

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

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*. https://doi.org/10.1016/S0960-9822(00)00221-9

Suzuki, K., Sako, K., Akiyama, K., Isoda, M., Senoo, C., Nakajo, N., & Sagata, N. (2015). Identification of non-Ser/Thr-Pro consensus motifs for Cdk1 and their roles in mitotic regulation of C2H2 zinc finger proteins and Ect2. *Scientific Reports*, *5*, 1–9. https://doi.org/10.1038/srep07929

Tyers, M., Tokiwa, G., Nash, R., & Futcher, B. (1992). The Cln3-Cdc28 kinase complex of S. cerevisiae is regulated by proteolysis and phosphorylation. *The EMBO Journal*, *11*(5), 1773–1784. http://www.ncbi.nlm.nih.gov/pubmed/1316273%0Ahttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC556635

Ubersax, J. A., Woodbury, E. L., & Quang, P. N. (2003). Targets of the cyclin-dependent kinase Cdk1. *Nature*, *425*(6960), 859–864. https://doi.org/10.1038/nature02057.1.

Wittenberg, C., & Reed, S. I. (1989). Conservation of function and regulation within the Cdc28/cdc2 protein kinase family: characterization of the human Cdc2Hs protein kinase in Saccharomyces cerevisiae. *Molecular and Cellular Biology*, *9*(9), 4064 LP – 4068. https://doi.org/10.1128/MCB.9.9.4064

Zoltowski, B. D., & Gardner, K. H. (2011). Tripping the light fantastic: blue-light photoreceptors as examples of environmentally modulated protein-protein interactions. *Biochemistry*, *50*(1), 4–16. https://doi.org/10.1021/bi101665s

Non-exclusive licence to reproduce thesis and make thesis public

I, Luka Bulatović,

(*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,

The use of phosphor degrons for controlled protein expression in the cell cycle,

(*title of thesis*)

supervised by Assoc. Prof., PhD Ilona Faustova, PhD Mihkel Örd, Prof. Mart Loog,

(*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.

Luka Bulatović

20/05/2021