Université de Montréal

Regulation of the RNA/DNA helicase Sen1 by proteasome-mediated degradation

Par

Marjorie Andrea Aleman Alvarado

Programme de biologie moléculaire

Faculté de médecine

Mémoire présenté en vue de l'obtention du grade de Maitrise

en biologie moléculaire

Avril, 2021

Université de Montréal

Unité académique : Faculté de Médecine

Ce mémoire intitulé

Regulation of the RNA/DNA helicase Sen1 by proteasome-mediated degradation

Présenté par

Marjorie Andrea Aleman Alvarado

A été évalué(e) par un jury composé des personnes suivantes

Marlene Oeffinger, PhD
Président-rapporteur

Francois Robert, PhD
Directeur de recherche

Jason Tanny, PhD
Membre du jury

Résumé

Sen1 est une hélicase ARN/ADN impliquée dans la protection du génome de la levure en résolvant les hybrides ARN/ADN et dans la terminaison de la transcription de courts ARN non codants. Malgré la demande cellulaire généralisée pour l'action Sen1, son abondance cellulaire est très faible, ce qui suggère que des mécanismes régulent les niveaux de protéine Sen1 dans la cellule. Nous avons confirmé que Sen1 est dégradé via une voie dépendante du protéasome. Ce mécanisme dépend de l'activité catalytique de Glc7, une protéine phosphatase dont il a été précédemment démontré qu'elle déphosphoryle Sen1 *in vitro* et qu'elle interagit avec Sen1 via un motif RVxF selon des expériences à deux hybrides.

Notre hypothèse de travail est que Glc7 contrôle les niveaux de protéine Sen1 via la déphosphorylation d'un phospho-dégron. Fait intéressant, un site potentiel dans la région N-terminale de Sen1 (sérine 863) qui peut fonctionner comme un phospho-dégron a été identifié dans une analyse à l'échelle du protéome de la co-occurrence de phosphorylation et d'ubiquitylation. Afin d'identifier les sites de phosphorylation Sen1 qui sont enrichis en l'absence de Glc7, nous avons réalisé une immunoprécipitation de Sen1 suivie d'une spectrométrie de masse. Cette analyse a identifié un site de phosphorylation dans Sen1 à la sérine 1505 qui pourrait agir comme un site de dégron potentiel. A noter que ce site, a également été signalé dans les études antérieures phosphoprotéomiques sur la levure. De plus, l'interaction entre Sen1 et Glc7 et l'importance du motif RVxF (par la mutation du résidu F2003) pour cette interaction ont été confirmée par co-immunoprécipitation. De manière surprenante, la prévention de cette interaction n'affecte pas la stabilité de Sen1 ou la croissance cellulaire.

Dans l'ensemble, nous avons identifié un petit groupe de sites de phosphorylation Sen1 avec une pertinence biologique potentielle. Nos résultats confirment également que la mutation du motif RVxF de Sen1 altère l'interaction avec Glc7 *in vivo*. Ces données approfondissent notre compréhension de la régulation de la protéine Sen1 par Glc7 dans

les cellules de levure qui peuvent fournir des indices sur le rôle de la sénataxine, orthologue humaine, dans les troubles neurodégénératifs.

Mots clés: Sen1, Glc7, système ubiquitine-protéasome, phosphodegron

Abstract

Sen1 is an RNA/DNA helicase involved in protecting the yeast genome by resolving RNA/DNA hybrids and in the transcription termination of short non-coding RNAs. Despite the widespread cellular demand for Sen1 action, its cellular abundance is very low, suggesting that mechanisms regulate Sen1 protein levels in the cell. We have confirmed that Sen1 is degraded via a proteasome-dependent pathway. This mechanism depends on the catalytic activity of Glc7, a protein phosphatase that was previously shown to dephosphorylate Sen1 *in vitro*, and to interact with Sen1 through an 'RVxF' motif.

Our working hypothesis is that Glc7 controls Sen1 protein levels via dephosphorylation of a phospho-degron. Interestingly, a potential site in the N-terminal region of Sen1 (serine 863) that may work as a phospho-degron has been identified in a proteome-wide analysis of phosphorylation and ubiquitylation cross-talk. In order to identify Sen1 phosphorylation sites enriched in the absence of Glc7, we conducted immunoprecipitation of Sen1 followed by mass spectrometry. This analysis identified one phosphorylation site within Sen1 at serine 1505 that could act as a potential degron site. Note that this site has also been reported in previous phosphoproteomic studies on yeast. Furthermore, the interaction between Sen1 and Glc7 and the importance of the RVxF motif (by mutation of residue F2003) for this interaction was confirmed by co-immunoprecipitation. Surprisingly, prevention of this interaction does not affect the stability of Sen1 or cell growth.

Overall, we have identified a small group of Sen1 phosphorylation sites with potential biological relevance. Our findings also confirm that mutating the RVxF motif of Sen1 impairs the interaction with Glc7 *in vivo*. These data further our understanding of Sen1 protein regulation by Glc7 in yeast cells that may provide clues to the role of senataxin, human orthologue, in neurodegenerative disorders.

Keywords: Sen1, Glc7, ubiquitin-proteasome system, phosphodegron

Table of contents

List of tables

List of figures

List of abbreviations

ALS4: Amyotrophic lateral sclerosis type 4
AOA2: Ataxia with oculomotor apraxia type 2
APC: Anaphase promoting complex
APT: Associated with Pta1
BRCA1: Breast cancer type 1
CF: Cleavage factor I A and B
CP: Core particle
CPF: Cleavage and polyadenylation factor
CTD: Carboxy-terminal domain
CUT: Cryptic unstable transcript
D-boxes: Destruction boxes
DDR: DNA damage response
DSB: Double strand break
E1: Ubiquitin-activating enzyme
E2: Ubiquitin-transferring enzyme
E3: Ubiquitin ligase
HR: Homologous recombination
IGHMBP2: Immunoglobulin Mu-Binding Protein 2
Lys: Lysines
ncRNA: Non-coding RNA
NLS: Nuclear localization signal
NMD: Nonsense-mediated mRNA
NNS: Nrd1-Nab3-Sen1
PIPs: PP1 interacting proteins
PP1: Phosphoprotein phosphatase type 1
PP1c: PP1 phosphatases catalytic subunit
PP2A: Phosphoprotein phosphatase type 2A
PP2B: Phosphoprotein phosphatase type 2B
PPases: Protein phosphatases
PPM: Protein phosphatase metal-dependent

PPP: Phosphoprotein phosphatase

PRU: Plecktrin-like receptor of ubiquitin

PTMs: Post-translational modifications

PTP: Phospho-tyrosine phosphatase

PTT: Premature transcription termination

RENT1: Regulator of Nonsense Transcripts-1

RF: Replication fork

RNAPII: RNA polymerase II

RP: Regulatory particle

RRM: RNA recognition motif

SF1B: Superfamily 1B of helicases

SMARD: Severe spinal muscular atrophy with respiratory distress

SMN: Survival of motor neuron protein

snoRNA: Small nucleolar RNA

snRNA: Small nuclear RNA

SUMOs: Small ubiquitin-like modifiers

TCR: Transcription-coupled nucleotide excision repair

TRAMP: Trf4-Air2-Mtr4

tRNA: Transfer RNA

Ub: Ubiquitin

UIM: Ubiquitin-interacting motif

UPS: Ubiquitin-proteasome system

UTR: Untranslated region

Acknowledgements

First of all, I would like to thank my research director François Robert for hosting me in his laboratory during my master's degree. Without his help, this thesis would never have taken shape. Besides my advisor, I would like to thank the members of my Follow-up Committee: Prof Jason Tanny and Prof Woong-Kyung Suh for their encouragement, insightful comments, and their advice.

I would like to give special thanks to two members of the laboratory: Élie Lambert and Noé Seija, thank you for always being there to listen to me and advise me. It is said that friends are the family one chooses and I am glad to have enjoyed your company, pleasant moments, and laughs with you. Thank you for making my stay here so much happier.

Célia Jeronimo, thank you for your patience and for showing me different techniques and tools that facilitated my work in the laboratory. I would also like to thank previous members and friends from other laboratories: Micaela Miyauchi, for being a great partner and making my stay more enjoyable. Pierre Collin, thank you for starting this project, none of this would be possible without your initial experiments. Ulises Rodriguez, thanks for the scientific discussions, your time, and especially for your friendship, these years would not have been the same without you just a few steps away.

I want to infinitely thank my family and friends, without whom I would not have had the strength and energy to continue on this path. To my parents and my sister, for being the main engines of my dreams, thanks for always trusting me, believing in me and my expectations. To my lifelong friends, Lesly and Izabó, who have always given me great moral and human support, necessary in difficult times, despite being 6 405 km away.

Finally, I would like to thank my life partner Andrés for his motivation during all these years. Thank you for listening to me, supporting me, and always being there despite everything. You have been my greatest source of motivation and strength in difficult situations and thanks to you I am now a better version of myself.

RATIONALE

Sen1 is a low abundance protein with various functions in the cell. Thus, there might be mechanisms that regulate its low levels. It has been previously shown that Sen1 protein levels are regulated in a proteasome-dependent manner during the G1 phase of the cell cycle (Mischo et al., 2018). However, this study was not conducted under physiological conditions as Sen1 was overexpressed to overcome the technical problems related to its low protein level.

Regarding the regulation of Sen1, preliminary data from our lab has confirmed that Sen1 is degraded via a proteasome-dependent pathway under physiological conditions (Collin, 2019). This mechanism is associated with Sen1 stability in the nucleus and relies on the catalytic activity of Glc7.

Upon a time course of Glc7 nuclear depletion with rapamycin (0, 30, 60, 90 min), Sen1 was destabilized. To determine whether the phosphatase activity of Glc7 was responsible for this effect on Sen1, two previously described catalytic mutants were used. E101Q has partial catalytic activity while D94N is considered catalytically dead (Logan et al., 2008). These mutants were used as part of the Glc7 complementation after nuclear depletion (CAND) system. That is, mutants were integrated at the *LEU2* locus into a strain where the endogenous *GLC7* gene is tagged with FRB, so that it can be depleted from the nucleus upon the addition of rapamycin. After nuclear depletion of endogenous Glc7 (90 minutes), Sen1 levels decreased in both mutants compared to WT. The most severe effect was shown by the catalytically dead mutant, where the level of Sen1 was dramatically reduced. These results demonstrated that the phosphatase activity of Glc7 was correlated with the stability of Sen1 in the nucleus (Collin, 2019).

These preliminary results suggested that the dephosphorylation of Sen1 by Glc7 prevented its proteasome-mediated degradation. Interestingly, serine 863 in Sen1 (S863) has been reported as a potential phosphodegron in a proteome-wide analysis of phosphorylation and ubiquitylation cross-talk (Swaney et al., 2013). This led us to evaluate the importance of S863 in Sen1 stability.

INTRODUCTION

1. The budding yeast nuclear helicase Sen1

Senataxin is a human RNA/DNA helicase involved in regulating different cellular processes. The importance of this protein came out in 2004 when it was reported that mutations in its *SETX* gene could be the cause of neurological disorders (Moreira et al., 2004). Shortly after this discovery, researchers identified a high degree of conservation between the *SETX* helicase domain and Sen1, its ortholog in *Saccharomyces cerevisiae*. Like senataxin, Sen1 is a helicase that fulfills several functions in the cell, the main and most prominent being the role it plays in terminating transcription of non-coding RNAs. Considerable research on Sen1 has provided the starting point to study Senataxin in mammalian and human cells. For this reason, I will first describe Sen1 in detail and the different works that have led to a better understanding of the wide spectrum of roles that Senataxin plays in humans. Then, I will describe Senataxin and its functions, focusing on the similarities and differences with Sen1.

1.1 Structure and domains of Sen1

Sen1 is a 252-kDa protein that has the ability to translocate in a 5′ to 3′ direction and exhibits helicase activity on both DNA and RNA duplexes as well as on RNA/DNA hybrids *in vitro* (Martin-Tumasz and Brow, 2015). It contains an N- and C-terminal domains involved in protein-protein interactions, along with an essential helicase region (DeMarini et al., 1992) (Figure 1).

The helicase domain (residues 1095 – 1896) comprises the minimal essential region of Sen1. It can bind to both single-stranded RNA and DNA, having a more stable

binding to the former (Martin-Tumasz and Brow, 2015; Mischo et al., 2011). This affinity for nucleic acids is characteristic of the superfamily 1B (SF1B) Upf1-like family of helicases, to which Sen1 belongs. This family contains two RecA domains (RecA1 and RecA2) in the helicase region, that are associated with nucleic acid binding and ATP hydrolysis, as well as other two SF1B-specific subdomains involved in RNA binding (Cheng et al., 2007; Chakrabarti et al., 2011). In addition to these common domains for SF1B, Sen1 has an accessory N-terminal “brace” subdomain which is necessary to adopt a correct conformation when binding and unwinding the RNA (Leonaitė et al., 2017) (Figure 1). Mutations in the helicase domain cause RNA polymerase II (RNAPII) readthrough (DeMarini et al. 1992; Han et al., 2017; Hazelbaker et al. 2013; Mischo et al. 2011; Steinmetz et al. 2001), suggesting that this region could be involved in regulating transcription termination. In the same manner, this region has been determined to be sufficient to induce dissociation of the elongation complex *in vitro* (Han et al., 2017).

The C-terminal domain has been much more studied and is known to contain important sequence motifs for nuclear localization (NLS). This bipartite domain is responsible for targeting and directing Sen1 nuclear localization (Chen et al., 2014; Nedea et al., 2008; Ursic et al., 1995). Importantly, proteins like the phosphatase Glc7, Nab3, and Nrd1, which are all involved in transcription termination, have been identified as Sen1 interactors. Nab3 and Nrd1 are known to interact with each other and form a complex with Sen1 to promote transcription termination in some genes. For instance, Nab3 strongly interacts with the C-terminal domain of Sen1 (Nedea et al., 2008) and it has recently been suggested that other sequence motifs in this region also allow Sen1 direct interaction with Nrd1 (Han et al 2020; Zhang et al., 2019). Further, it has been proposed that Glc7 can dephosphorylate Sen1 *in vitro* and that this interaction is essential for the termination process in noncoding transcripts (Nedea et al., 2008).

The N-terminal domain is important for Sen1 function *in vivo* (Ursic et al., 2004), however, its presence is not essential for exercising its helicase activity *in vitro* (Leonaitė et al., 2017). This terminal region has been mostly described as an interaction platform between Sen1 and several Sen1-interacting proteins that allow either the recruitment or the regulation of it. For instance, it has been proposed that the N-terminal domain can

mediate the interaction with Rpb1, the largest subunit of RNAPII (Conrad et al., 2000; Steinmetz et al., 2006). This interplay is correlated with the fact that the deletion or mutation of this region can generate specific transcription termination defects *in vivo* (Finkel et al., 2010; Steinmetz *et al.,* 2006). Specifically, it was first proposed that Sen1 can interact with phosphorylated serine 2 of the carboxy-terminal domain (CTD) of Rpb1 and that this allows its recruitment to the elongating RNAPII (Chinchilla et al., 2012). This was reported using two-hybrid assays, however, the interaction with the CTD is not fully defined since other studies state that the ability of Sen1 to interact with this phosphorylated residue is not essential for its recruitment or transcription termination (Collin et al., 2019; Han et al., 2020). Despite Sen1 function is more associated with transcription termination, other interactor proteins have been discovered that further expand the spectrum of functions that Sen1 fulfills in various cellular processes. Proteins like Rnt1, a ribonuclease involved in processing transcripts; Rad2, a DNA repair enzyme, and SmD3, required for mRNA splicing, have also been identified as N-terminal interacting proteins (Fromont-Racine et al., 1997; Lamontagne et al., 2000). Recently, it has been identified that both Ctf4 and Mrc1 (necessary for the formation of the replication fork complex) bind to the N-terminal domain of Sen1, and this promotes its association with the replisome (Appanah et al., 2020).

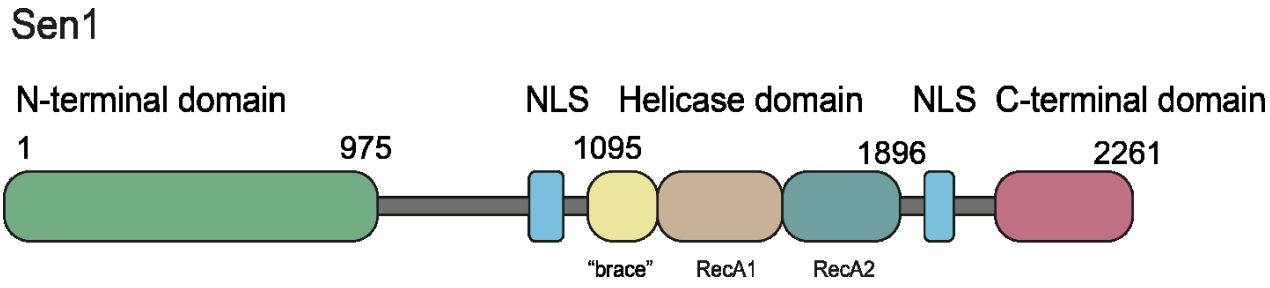


Figure 1. – Functional domains of the helicase Sen1. Schematic representation of Sen1 and its different domains. The N and C terminal domains are involved in protein-protein interactions. The helicase domain contains two regions, RecA1 and RecA2, which allow the binding to nucleic acids and ATP hydrolysis; and a "brace" region necessary to adopt a good conformation. Motifs for nuclear localization (NLS) flank the helicase region and target Sen1 to the nucleus.

1.2 Cellular processes involving Sen1

Sen1 was firstly reported as a transfer RNA (tRNA) splicing endonuclease in a direct screening for mutations that affect this activity. In this study, two new genes were identified (*SEN1* and *SEN2*), whose products were associated with this catalytic reaction (Winey and Culbertson, 1987). Years later, it was shown that deficiencies in the splicing of tRNA precursors were caused by defects in *SEN2* gene termination that provoke a decrease in its expression and were not directly related to *SEN1* (DeMarini et al., 1992). The first clue of Sen1 function came from its homologous *sen1*$^{+}$ gene in *Schizosaccharomyces pombe*. The *sen1*$^{+}$ gene codes for an enzyme able to unwind RNA and DNA, along with its ability to hydrolyze ATP from nucleotides and translocate in the 5’ to 3’ direction.

Similar to its fission yeast and human homologs, Sen1 protein levels are present in very low amounts in the cell; hence, its overexpression using different systems can generate atypical abundance, eventually altering various cellular aspects. Maintaining the physiological levels of Sen1 is thus important for its interaction with several proteins that act in distinct biological processes such as RNAs processing and maturation (Ursic et al., 1997), transcription termination (reviewed in Mischo and Proudfoot, 2013; Porrua and Libri, 2015) and DNA damage response (Li t al., 2016; Ursic et al., 2004).

1.2.1 Role in transcription termination

Transcription termination is a key step that allows the dissociation of RNAPII from DNA at the end of transcription, this involves a series of events that regulate the boundaries of the transcript enabling its release for further processing.

In budding yeast, there are two mechanisms by which transcription termination occurs. Protein production from coding genes generates stable transcripts that are mostly directed to the cytoplasm. Processing at its 3’ terminal end starts with the addition of adenosines known as the polyA tail, this comprises the recognition of a polyA site, also known as polyA signal. The polyA tail is added in the 3' untranslated region (UTR) of the

nascent RNA by the CPF (cleavage and polyadenylation factor) and CF (cleavage factor I A and B) complexes. On the other hand, termination of short non-coding transcripts and some protein-coding genes is independent of the CPF and CF and relies on the Nrd1-Nab3-Sen1 (NNS) complex. This includes functional RNAs such as small nuclear RNAs (snRNAs) and small nucleolar RNAs (snoRNAs), which are processed by the nuclear exosome (Allmang et al., 1999; Gudipati et al., 2012). The exosome is a multiprotein ribonuclease in charge of degrading several types of RNA, mainly molecules that have been inefficiently processed, but also it is in charge of processing and converting snRNAs and snoRNAs into mature transcripts.

Moreover, this pathway also allows the termination of cryptic unstable transcripts (CUTs), which generally results in the total degradation of the transcript by the nuclear exosome to prevent the accumulation of non-functional RNAs (Jensen et al., 2013; Wyers et al., 2005). Strikingly, the NNS complex has also been involved in the termination of some protein-coding genes, specifically inducing premature RNAPII termination (attenuation) (Colin et al., 2011; Contreras et al., 2013) or termination of mRNAs (Steinmetz et al., 2006).

1.2.1.1 Transcription termination in coding genes

In budding yeast, most of the protein-coding genes are terminated via the CPF pathway which relies on the cleavage and polyadenylation of the nascent RNA. This mechanism is independent of Sen1 action and rather includes other components that allow the recognition and cleavage of the polyA site, as well as other mechanisms to dissociate the RNAPII from the template.

Briefly, the first step involves recruiting the CPF complex to the 3' UTR of the nascent RNA (polyA site), which generally has the AAUAAA sequence. The recognition of the polyA signal provides the specificity for this termination path and relies on several components of the CPF complex (Rna15, Cft1, Cft2, Yth1, Mpe1, and Hrp1). The CPF complex can also be recruited through its Pcf11 subunit, which recognizes the phosphorylated form of Serine 2 of the CTD (Kuehner et al., 2011; Porrua and Libri, 2015; Tong and Manley, 2014). Both processes, the recognition of the polyA site and the interaction with RNAPII are of vital importance for the recruitment of CPF to the 3' end of

protein-coding genes. Subsequently, the pre-mRNA is cleaved by the Ysh1 endonuclease at the polyA site while the Pap1 polymerase is responsible for the synthesis of adenosine nucleotides at the 3' end (Figure 2).

Released RNAs are rapidly exported into the cytoplasm and the RNAPII is released from the DNA. In this context, the CPF-pathway can induce transcription termination by two non-mutually exclusive mechanisms. One involves the release of elongation factors, thus reducing the processivity of RNAPII due to conformational changes that lead to termination (allosteric model) (reviewed in Richard and Manley, 2009). In the second proposed model, the newly generated 5' end of the cleaved nascent RNA, which is still attached to RNAPII, is targeted by the 5'-3' exonuclease Rat1 which degrades the remaining RNA and induce termination by interacting with RNAPII (torpedo model) (Kim et al., 2004; West and Proudfoot, 2004).

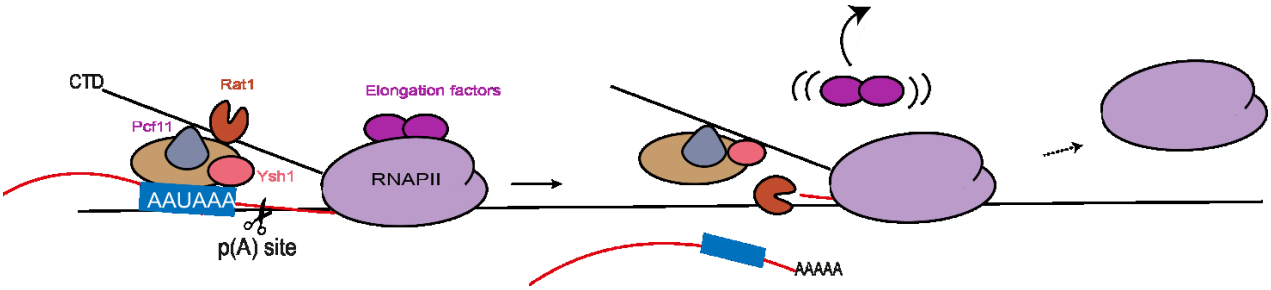


Figure 2. – Schematization of transcriptional termination mechanisms for coding genes in *S. cerevisiae.* Both models propose the recruitment of the CPF-CF complex by an interaction between Pcf11 and the CTD. This recruitment then allows cleavage of the pre-mRNA by the endonuclease Ysh1 and polyadenylation by Pap1, which then leads to two non-mutually exclusive events that allow termination. Either by the dissociation of elongation factors (allosteric model), the degradation of nascent RNA by Rat1 exonuclease (torpedo model) or both (allosteric/torpedo model).

There is evidence supporting each model independently, however, it has also been proposed that both mechanisms are acting through a unifying allosteric-torpedo model. In this sense, this model suggests that the Rat1 exonuclease degrades the remaining

nascent RNA, consistent with the torpedo model. Nevertheless, this degradation is not sufficient to induce the release of RNAPII from the template, for which the action of the CPF components is necessary (Deng and Cramer, 2009; Luo et al., 2006). This joint model has recently gained more strength since two studies have shown that the elongation complex undergoes a conformational change (allosteric model), which is mainly mediated by dephosphorylation of the elongation factor Spt5, that causes the deceleration of RNAPII favoring termination by Xrn2 (Rat1 in budding yeast) (Cortazar et al., 2019; Eaton et al., 2020) (Figure 2).

Even though Sen1 does not play a fundamental role in the termination of mRNAs, there is increasing evidence suggesting that it acts in the termination of some protein-coding genes and also in other termination mechanisms at specific genes. Premature transcription termination (PTT) is an example of an alternative mechanism that can also control the expression of certain protein-coding genes. Also known as attenuation, this process prevents the formation of full-length transcripts and therefore generates shorter products that are either rapidly degraded, or stabilized by polyadenylation (reviewed in Kamieniarz-Gdula and Proudfoot, 2019). PTT in budding yeast is mainly mediated by components of the NNS complex, specifically by Sen1, however, the molecular mechanisms and how this process is triggered are still unclear. Some examples of protein-coding genes regulated by attenuation include *GPH1* (glycogen phosphorylase), *IMD2* (nucleotide biosynthesis), and more recently *DEF1* (DNA damage response) (Chen et al., 2017; Jenks et al., 2008; Whalen et al., 2018).

Interestingly, genome-wide studies identified that Sen1 cross-linking sites with RNA were located at the 3' end of protein-coding transcripts (Creamer et al., 2011; Jamonnak et al., 2011), suggesting that Sen1 could have a role in polyA-dependent transcription termination. In line with this evidence, it has been indicated that the substitution of an amino acid in the helicase domain of Sen1 (*sen1*-E1597K) generates readthrough transcription in a small group of protein-coding genes (Steinmetz *et al.,* 2006). This same termination defect in some coding genes was also observed in CTD Tyrosine1 mutants, which was rescued by slowing down transcription (Collin et al., 2019). This indirectly suggests a role for Sen1 in regulating termination in these genes.

Additionally, when the polyA-dependent termination mechanism fails, Sen1 can promote efficient termination in some protein-coding genes, thereby generating a fail-safe mechanism (Rondón et al., 2009). This involves RNA cleavage by the RNAse III Rnt1, which allows Rat1 to degrade the nascent transcript and Sen1 to dissociate RNAPII from the template (Kawauchi et al., 2008). Yet, the exact function of Sen1 in mRNA termination is far from being fully understood and requires further investigation.

1.2.1.2 Transcription termination in non-coding genes

RNAPII also generates transcripts other than those encoded by coding regions, among them are several classes of non-coding RNA (ncRNA) such as snRNA, snoRNA, and CUTs. The first two exert important functions in the cell, while the last one is mainly associated with unexpected products of transcription, thus its termination is highly regulated to avoid interference with the coding transcriptome (Jensen et al., 2013).

Termination of ncRNAs relies on the NNS pathway (Nrd1-Nab3-Sen1) (Figure 3). This pathway begins with the recruitment of Nrd1 to the elongation complex via the interaction with phosphorylated Serine 5 (Ser5) of the CTD (Han et al., 2020; Mayer et al., 2011; Tietjen et al., 2010). This process is controversial since no termination defects were detected in *nrd1* mutants (Vasiljeva et al., 2008) nor in CTD Ser5 mutants (Collin et al, 2019). After being recruited, Nrd1 binds specific motifs in the nascent RNA that are targeted as they emerge from RNAPII. In this sense, Nrd1 and Nab3 possess an RNA recognition motif (RRM), which recognizes these short sequences in the RNA (GUAA/G and UCUUG, respectively) (Carroll et al., 2007; Porrua et al., 2012). This provides the specificity for the NNS termination pathway.

These transcripts are not cleaved or polyadenlylated at this stage, and rather what is proposed is that the helicase Sen1 is in charge of releasing RNAPII by mechanisms that are still poorly understood. Given that Sen1 is the only component of the NNS complex that has catalytic activity, along with its ability to dissociate the elongation complex *in vitro* (Porrua and Libri, 2013), the most accepted model proposes that Sen1 takes advantage of RNAPII pausing to track along the nascent RNA and catch up the transcribing RNAPII to induce termination (Collin et al., 2019; Han et al., 2017; Hazelbaker *et al*., 2013; Mischo and Proudfoot, 2013).

Most of the RNAs produced by this pathway are degraded by the nuclear exosome after being released. The released transcripts are first processed by the TRAMP complex (Trf4-Air2-Mtr4), thanks to its interaction with Nrd1. This allows the polyadenylation of the target RNAs (Wyers et al., 2005; Tudek et al., 2014). TRAMP also might allow the recruitment of the exosome to favor the subsequent degradation of the transcript. The exosome trims the 3' ends using its two nucleases, Rrp6, and Dis3, to either transform them into mature transcripts (snRNAs and snoRNAs) or degrade them (CUTs) (Allmang et al., 1999).

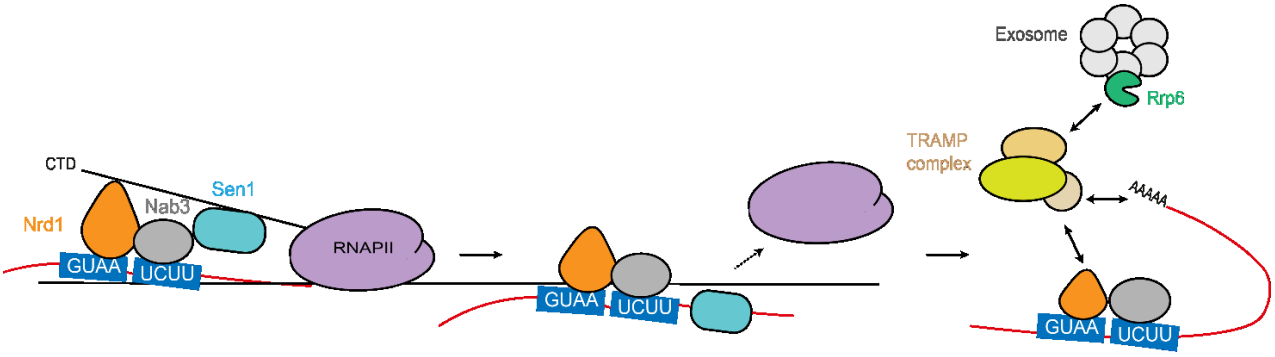


Figure 3. – Schematization of transcriptional termination mechanism for non-coding genes in *S. cerevisiae*. Mechanism of termination in snoRNA, snRNA, CUTs, as well as certain coding genes (attenuation). This model involves the NNS complex (Nrd1-Nab3-Sen1). Nrd1 is recruited by the CTD and by an RNA interaction, which allows the recruitment of Nab3 and Sen1. Then, Sen1, via its helicase activity, catches up with RNAPII to induce termination. Once released, the transcripts are polyadenylated by TRAMP and trimmed or degraded by the exosome.

1.2.2 Role in genome stability

Sen1 is best known for its function in transcription termination, nonetheless, it also has other functions associated with maintaining the stability and integrity of the genome. These functions imply the total and accurate replication of DNA in each cell cycle. In this sense, multiple obstacles can block the progression of replication forks (RF) leading to DNA damage and genome instability. For instance, transcription represents one of the

main barriers to replication, not only because RNAPII can obstruct RF progression, but also because it can generate various structures, such as DNA:RNA hybrids (R-loops), which interfere with its advancement (Gomez- Gonzales and Aguilera, 2019).

R-loops are formed by the re-hybridization of nascent RNA to the DNA template, which generates triple-stranded structures (Aguilera and Garcia-Muse, 2012). Many of them are transient and naturally formed during the transcription process, however, their abnormal accumulation can cause deleterious effects on genome stability, especially during replication, since it has been shown that they promote DNA replication machinery stalling and increased DNA rearrangements (Hegazy et al., 2020; Gan et al., 2011; Garcia-Pichardo et al., 2017). The enzymes in charge of removing these hybrids are the RNAse H, which digest the RNA directly. Other helicases can also unwind R-loops *in vitro*, among these we can find Sen1 (Mischo et al., 2011). Supporting Sen1 role in promoting genome stability, it was early demonstrated that it is associated with RFs (Alzu et al. 2012) and that its function here is mainly linked with preventing the accumulation of R-loops during the S phase of the cell cycle, where a collision between the hybrids and the replisome is more prone. Recently, a study has shown that the N-terminal domain of Sen1 is necessary for its association with the replisome and that it mediates its interaction through the replisome Ctf4 and Mrc1 subunits (Appanah et al., 2020). Despite its clear role in dissolving R-loops and replisome association, it remains to be defined whether replication defects in the absence of Sen1 are due exclusively to their role in RFs or occur as a consequence of abnormal transcription termination.

Sen1 is also part of the DNA damage response (DDR). Supporting this role, the *sen1-1* (G1747D) mutation in the helicase domain not only generates termination defects (Steinmetz et al, 2006) but also an increase in the rate of recombination and accumulation of R-loops (Mischo et al., 2011). By identifying the DNA repair mechanisms involved in maintaining the viability of this mutant, it was determined that the Rad50 and Mre11 proteins - involved in homologous recombination - caused synthetic lethality with the *sen1-1* mutant (Mischo et al., 2011). Besides, other mutations such as *sen1-2* (Δ1–975), which lacks the N-terminal domain, caused enhanced sensitivity to several DNA-damaging

agents as well as defects in the activation of the DNA damage checkpoint (Golla et al., 2013).

Because *SEN1* mutations can make the cell more sensitive to DNA-damaging agents, it may be possible for Sen1 to coordinate one or more DNA repair pathways. For instance, Sen1 is known to interact with Rad2 and Rad1, two proteins involved in transcription-coupled nucleotide excision repair (TCR) (Li et al., 2016; Ursic et al., 2004). Interestingly, R-loop formation together with Mre11, has been recently implicated in double-strand break (DSB) repair by recruiting Sen1 to the damaged site (Rawal et al., 2020).

1.3 Cell cycle regulation of Sen1

Protein levels of Sen1 can vary -depending on the context- from 64 to 500 molecules per cell. This number is much lower compared to the abundance of some other proteins involved in transcription termination, such as Nrd1 (550-20,000), Nab3 (2,000-6,000), and RNAPII (600-1,000) (Chong et al., 2015; Ghaemmaghami et al., 2003; Newman et al., 2006). This may be because Sen1 is responsible for disengaging the RNAPII from the template, making its levels much more fine-tuned to effectively complete termination (Mischo et al., 2018). Additionally, Sen1 has other functions beyond those related to NNS, as previously mentioned, and it may also be involved in promoting genome stability and response to DNA damage.

Being Sen1 of vital importance among different cellular processes, it was unexpected that it was found at such low levels. This suggested that Sen1 abundance could be regulated based on the requirements of the cell. In fact, it was early demonstrated that the ubiquitin-proteasome system was responsible for keeping its levels low (DeMarini et al., 1995). Recent work has confirmed that Sen1 levels fluctuate throughout the cell cycle, with the ubiquitin-proteasome system being responsible for reducing Sen1 levels in the G1 phase, compared to the S/G2 phase where the levels are higher (Mischo et al., 2018) (Figure 4).

The Anaphase Promoting Complex (APC) has been reported as the ubiquitin ligase that specifically targets Sen1 for its programmed degradation during G1 (Mischo et al., 2018). Proteolysis by APC begins in early mitosis, and persists throughout the entire G1 phase until it is terminated when cells enter S phase (Zachariae, et al., 1998). The APC recognizes linear motifs in the protein, which contribute to its destabilization and subsequent degradation by the proteasome; however, to date, the existence of a death signal or other similar motifs has not been identified in Sen1, and the mechanism that triggers its degradation remains unclear.

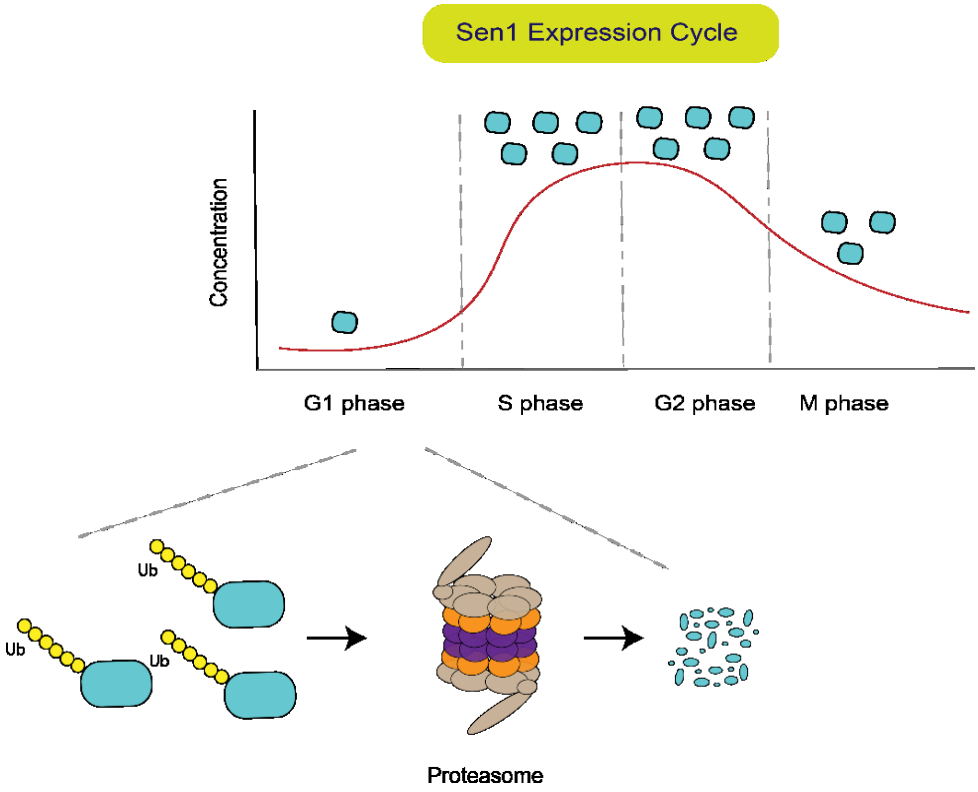


Figure 4. – Sen1 is ubiquitinated and degraded in a proteasome-dependent manner. Sen1 levels increase during the S and G2 phases of the cell cycle. This level is controlled by the ubiquitin-proteasome, which degrades Sen1 preferentially during G1. Sen1 degradation relies, in part, on the ubiquitin ligase anaphase-promoting complex (APC) that targets it to the proteasome.

The overexpression of Sen1 has also led to a reduction of RNAPII occupancy in CUTs, SUTs and to a lesser degree in non-coding transcripts such as snRNAs and snoRNAs (Mischo et al., 2018). Furthermore, it is known that high levels of Sen1 cause cell lethality possibly due to the accumulation of coding transcripts (Mischo et al., 2018), however, it has not been possible to identify whether abnormal levels of Sen1 would affect NNS-independent functions.

2. The human RNA/DNA helicase senataxin

The *SETX* gene encodes senataxin, a 302-kDa RNA/DNA helicase whose helicase domain shows homology to the budding yeast protein Sen1. Although Sen1 is found only in the nucleus, Senataxin can be found in both the nucleus and cytoplasm. Despite this, the abundance of Senataxin in the cell remains as low as that of Sen1, and its level has to be highly regulated to prevent accumulation leading to abnormal processes.

Most of the SETX functions derive from studies done on Sen1 in yeast. However, unlike Sen1, which is mainly involved in NNS termination, Senataxin is not required for snRNA termination (Suraweera et al., 2009). This is due, in part, to the fact that the NNS pathway is not conserved in mammals, mainly because the RNA-binding protein Nab3 does not have a known human homologous protein (Kuehner et al., 2011). In this sense, the main functional processes attributed to Senataxin are associated with its role in regulating gene expression, this includes the processing of coding transcripts through its interaction with RNAPII and the nuclear exosome. On the other hand, there is also evidence of the role of Senataxin in promoting genome integrity through the resolution of RNA/DNA hybrids (R-loops) and the formation of nuclear foci during the S/G2 phase generated by replicative stress.

Interestingly, *SETX* mutations have been reported to cause various neuropathologies, including ataxia with oculomotor apraxia type 2 (AOA2) and amyotrophic lateral sclerosis type 4 (ALS4) (Chen et al., 2006; Moreira et al., 2004). The interest in knowing *SETX* function in these neurological disorders led to the generation of

the *SETX* interactome (Figure 5), which includes association with DNA repair factors, SUMOylation and ubiquitination-associated factors, splicing factors, and RNA-binding proteins, which will be addressed throughout this section.

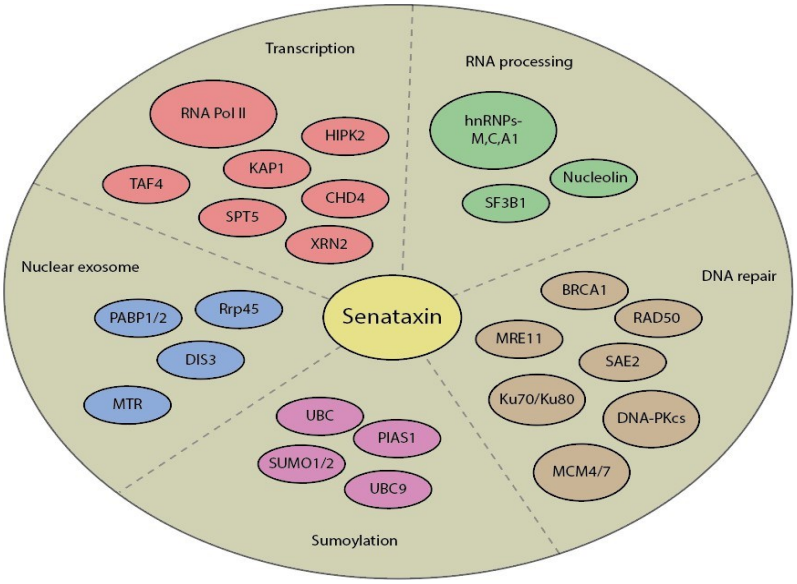


Figure 5. – SETX protein-protein interactome network. Different proteomics analysis have identified several SETX interactors. These factors are mostly associated with: DNA repair (brown), SUMOylation and ubiquitination (purple), RNA exosome (blue), RNAPII and transcription (red), splicing factor and RNA processing (green), adapted from Dutta et al,. 2020.

2.1 Characterization and biological processes regulated by senataxin

Senataxin has three known domains: the N-terminal protein interaction domain, the C-terminal helicase domain, and only one nuclear localization signal (NLS) domain (Bennett et al., 2018) (Figure 6).

The C-terminal domain is the best characterized as it shares homology with the helicase region of two other human proteins, Immunoglobulin Mu-Binding Protein 2 (IGHMBP2) and Regulator of Nonsense Transcripts-1 (RENT1) (Chen et al., 2006;

Moreira et al., 2004). Given this, all three, Senataxin, IGHMBP2, and RENT form a helicase subfamily. RENT1 (the yeast Upf1 ortholog) is required for nonsense-mediated mRNA decay (NMD) and its activity is dependent on phosphorylation (Weng et al., 1996), whereas mutations in IGHMBP2 helicase have been reported to cause severe spinal muscular atrophy with respiratory distress (SMARD) (Grohmann et al., 2001). Interestingly, mutations in this helicase subfamily have been involved in either embryonic lethality (RENT1) or neurodegenerative disorders caused by neuronal susceptibility (IGHMBP2 and Senataxin) (Bennet et al., 2018), suggesting a possible role for this subfamily in the malfunction of the central nervous system.

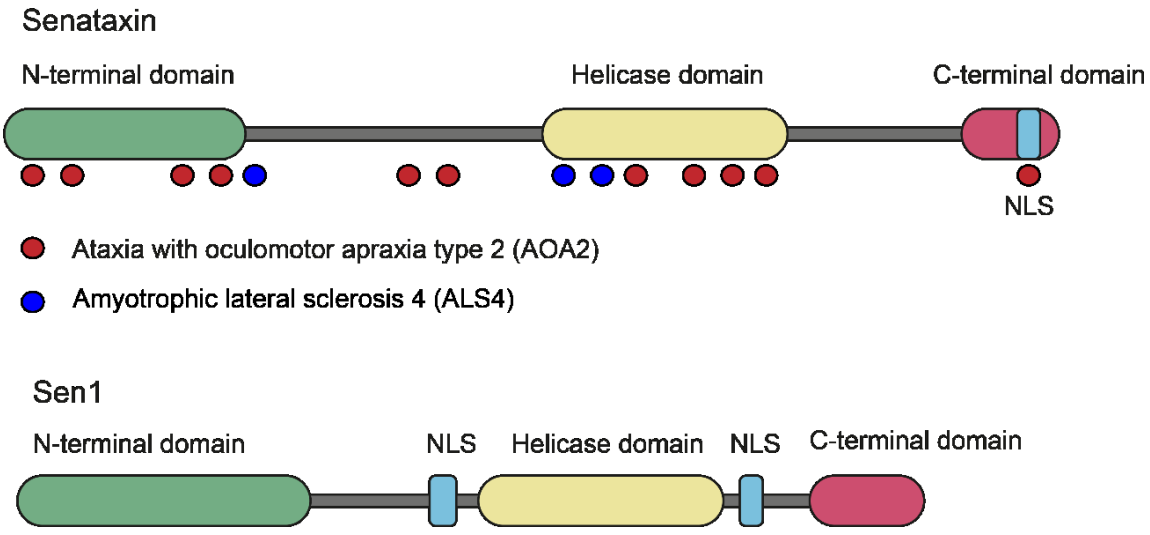


Figure 6. – Schematic diagram of SETX and Sen1 indicating their different domains. Both, SETX and Sen1 have an N-terminal domain that acts as a platform for protein-protein interactions. More than 100 mutations throughout SETX have been associated with ALS4 and AOA2.

The N-terminal domain is not only essential for the interaction of Senataxin with different proteins but also for the formation and localization of different foci during the S/G2 phase (Yuce and West, 2013). Among the proteins that interact with this region are SPT5, RRP45, and Nucleolin (Hein et al., 2015; Richard et al., 2013; Suraweera et al., 2009), which are involved in transcription and RNA processing. Likewise, interaction with

proteins that play a role in DNA damage repair has been reported, these include RBBP8/SAE2, UBC, and UBC9 (Bennett et al., 2013). This region can also be modified by the addition of small ubiquitin-like modifiers (SUMOs) required for the interaction of Senataxin with the RRP45 subunit of the exosome (Richard et al., 2013).

2.1.1 Role in transcription and RNA processing

The high degree of conservation within the helicase domain between budding yeast Sen1 and human *SETX* suggested that the latter could also be involved in regulating gene expression (Groh et al., 2017; Suraweera et al., 2009). In this sense, the role of Senataxin in transcription termination has been extensively described in several studies. During termination, Senataxin acts by unwinding R-loops at the termination site, which allows for pausing, RNA to be released, and degradation by exonuclease XRN2 which leads to a consequent release of RNAPII from the DNA template (Skourti-Stathaki et al., 2011). It has been shown that Senataxin interact with various components of the transcription machinery, mainly with RNAPII (Suraweera et al., 2009), but also with other processing factors such as SPT5 and TAF4 (Hein et al., 2015; Miller et al., 2015), which strengthens its role in this process.

Senataxin can interact with several proteins to trigger efficient transcription termination. For instance, it has been demonstrated that it can also be recruited to termination sites through its interaction with the breast cancer type 1 susceptibility protein (BRCA1) and survival of motor neuron protein (SMN), this recruitment depends on the accumulation of R-loops as well as on the RNAPII CTD, respectively (Hatchi et al., 2015; Zhao et al., 2016). Interestingly, SMN is essential for the assembly of the pre-mRNA splicing machinery, and similar to Senataxin, it can present neuropathy-associated mutations (Zhao et al., 2016).

While the interaction of Senataxin with XRN2, BRCA1, and SMN can promote transcription termination in different cellular contexts, interaction with other proteins might

induce transcriptional readthrough. This is the case of the association with the PERIOD complex, involved in the regulation of circadian clock genes. This association prevents RNAPII from being released and therefore causes reduced mRNA production (Padmanabhan et al., 2012).

Similar to Sen1 in budding yeast, Senataxin has also been involved in premature transcription termination (PTT) and can also act as a transcription inhibitor (Groh et al., 2017). One of these cases occurs upon viral infection, where Senataxin negatively regulates the transcription of certain genes associated with the antiviral response (Miller et al., 2015). In this sense, Senataxin recruitment promotes the early release of RNAPII in genes that encode antiviral molecules, thus limiting its expression to ensure an appropriate response. Likewise, Senataxin in complex with XRN2, RRP6, and DROSHA/DGCR8 (microprocessor) can be recruited to the promoter-proximal region of endogenous retroviruses causing early transcription termination and thus repressing the expression of certain viruses (Wagschal et al., 2012).

2.1.2 Role in maintenance of genome stability

Studies in the budding yeast homolog Sen1 allowed to propose a role for Senataxin in solving R-loops. Given the large size of genes in humans, transcription might last longer and take place even during the S phase of the cell cycle (Helmrich et al., 2013). Due to this, an abnormal accumulation of R-loops in this phase can cause deleterious effects on the stability of the genome. Those mechanisms mainly involve the action of RNAse H1/2, however, there is growing evidence suggesting that Senataxin is also in charge of tightly regulating these levels.

R-loops can be originated by either stalling of RNAPII, disruption of the transcription-coupled process, or head-on collisions between the DNA replication and transcription machinery during the S phase (Hatchi et al., 2015; Yuce and West, 2013). Moreover, many of these R-loops can be associated with the formation of DNA double-strand breaks (DSBs). In this sense, it has recently been shown that Senataxin is

specifically recruited to these DSBs in actively transcribed genes to promote DNA repair via homologous recombination (HR) (Cohen et al., 2018).

The displacement of a single-stranded DNA in the structure of the R-loop makes this strand highly susceptible to deleterious molecules. Endorsing this, some studies suggest Senataxin would play a more direct role in inducing DNA damage repair since both Sen1 and Senataxin associate and even co-localize with repair proteins during replication (Richard et al., 2013; Yuce and West, 2013). As an example of this, Senataxin can be recruited to R-loops formed at the transcription termination sites by factors involved in DNA damage response, such as BRCA1 (Hatchi et al., 2015). Here, BRCA1 is responsible for repairing the single-strand breaks generated in the exposed strand, while Senataxin could work in parallel with the RAD51 recombinase to repair DSBs via the resolution of R-loops and HR at these sites, respectively (Cohen et al. al., 2018).

Senataxin forms subnuclear foci in the S/G2 phase and the number of these increases in the presence of transcription-associated replicative stress and R-loops. Besides, it co-localizes with DNA damage response proteins, such as 53BP1, suggesting that these foci would correspond to sites of replication-transcription collision (Yuce and West, 2013).

Finally, Senataxin also contributes to nuclear RNA surveillance and genome stability through its interaction with the nuclear exosome (Richard et al., 2013). Rrp45, a component of the exosome, is able to interact with senataxin when this is SUMOylated (Richard et al., 2013). Small ubiquitin-like modifiers (SUMO) facilitates the formation of protein complexes when there is DNA damage (Jackson and Durocher, 2013), these complexes tend to form nuclear foci in which the co-location of Senataxin-Rrp45 has been demonstrated. This showed that the interaction between SETX and the exosome targets the exosome to sites where there is transcription-related DNA damage, for instance, R-loop formation, which allows its effective resolution by degrading the RNA moiety (Richard et al., 2013).

2.2 Localization and function in cyclin cells

Senataxin is located mainly in the nucleus in unsynchronized cells, however, some studies have reported that it can be located specifically in the nucleolus, which is consistent with the location of Sen1 and its role in processing rRNA and snoRNAs (Ursin et al., 1997). Subsequent studies confirmed this by observing that Senataxin co-localized with fibrillarin, a component of many ribonucleoproteins situated in the nucleolus (Chen et al., 2006).

In an analysis of nuclear foci formation in synchronized cells, the presence of GFP-tagged *SETX* was visualized in the nucleolus during the S phase. However, as cells progressed into the G2 phase, the location of Senataxin-GFP changed to the nucleoplasm (Yuce and West, 2013). The presence of Senataxin-associated nuclear foci under replicative stress conditions has also been evidenced during the S/G2 phase, and other studies state that by blocking RNAPII-mediated transcription, the number of these foci decreased. This would support a critical role of Senataxin in solving, together with other proteins, the collisions between the replication and transcription machinery that cause genome instability.

Some studies indicate that Senataxin may be initially targeted to the nucleus and it would then be exported to the cytoplasm (Chen et al., 2006). This cytoplasmatic function of Senataxin has not yet been fully described, but it would be evolutionarily distant from those of Sen1, which locates exclusively in the nucleus. Another difference with its yeast counterpart is that the concentration of Senataxin is not dependent on the cell cycle, however, a different localization of Senataxin has been seen throughout the cell cycle (Yuce and West, 2013).

3. Protein phosphatases in budding yeast

Phosphatases and kinases dictate the phosphorylation status of proteins, which can affect their stability, localization, or activity. Protein phosphatases (PPases) are enzymes in charge of removing the phosphate of phosphorylated residues. In *S. cerevisiae*, there are 43 PPases in total (Offley and Schmidt, 2019), which have been classified into five distinct evolutionary families based on their DNA sequence, structural and catalytic data. Among these, three have been widely studied: Phosphoprotein phosphatase (PPP), Protein phosphatase metal-dependent (PPM), and Phospho-tyrosine phosphatase (PTP). The PPP family is mostly represented by serine/threonine phosphatases, which can be classified as type 1 (PP1), type 2A (PP2A), or type 2B (PP2B) (Cohen, 1989). So far, 13 phosphatases have been identified as members of the PPP family in budding yeast.

The PP1 group has only one representative in yeast, encoded by a single gene, *GLC7*. Glc7 is an essential protein, which has different functions. Many of these depend on different regulatory subunits, which modulate the specificity with the target proteins and the localization in the cell (Cannon et al., 2010). In this section, I will focus on the main functions of Glc7 in the cell and its possible connection to Sen1.

3.1 The PP1 family

PP1 family member's protein sequence is highly conserved among all eukaryotes. Specifically, PP1 phosphatases contain a catalytic subunit (PP1c) that is in charge of hydrolyzing the phosphoester bond by using two metal ions: Mn^{2+} and Fe^{2+} (Egloff et al., 1995; Goldberg et al., 1995) (Figure 7).

The PP1 crystal structure was first solved in mammals and it showed a central β-sandwich located between two α-helix domains, which are the N- and C-terminal regions (Goldberg et al., 1995; Ye et al., 2013). The intersection of the α-helix with the β- sandwich

forms a Y-shaped cleft, in which the catalytic site is located. There, the two metal ions are coordinated by other residues (His, Asp, and Asn) to carry out the catalytic reaction (Egloff et al., 1995; Shi et al., 2009). Likewise, there are three grooves in the Y-cleft known as hydrophobic, acidic, and C-terminal.

A striking feature of PPP phosphatases, and specifically of the PP1 family, is that they form multiple complexes, known as holoenzymes, which are made up of a single PP1c and one or more regulatory subunits (Cannon et al., 2010). Thus, the quantity and identity of the holoenzymes depend on their regulatory subunits. These proteins bind to PP1c and are necessary to modulate its activity, subcellular location, substrate specificity or to be substrates themselves (Cannon, 2010; Ariño et al., 2019).

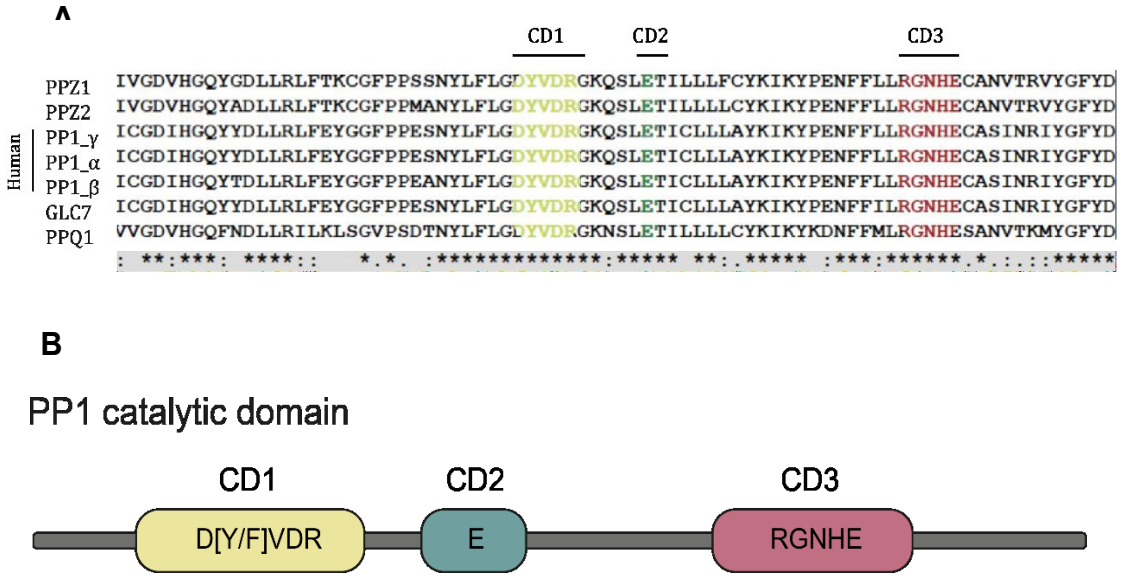


Figure 7. – Protein sequence alignment and domains of PP1 homologs. (A) The amino acid sequence of the catalytic domain of *S. cerevisiae* Glc7, PPQ1, PPZ1, PPZ2, and three isoforms (α, β and γ) of human PP1 were aligned. **(B)** Conserved catalytic domains are shown in yellow (CD1), green (CD2) and red (CD3). Modified from Logan et al., 2008

3.2 The phosphatase Glc7

In mammals, PP1c is encoded by three genes, while in budding yeast is limited to a single gene known as *GLC7*. Glc7 is composed of 312 amino acids and presents 85% identity to the human PP1c (Cannon et al., 2010). Contrarily, the N and C terminal regions of the majority of PP1 members vary in size and show little similarity (Offley and Schmidt, 2019).

In most cases, Glc7 does not act alone in the cell, but fulfills its different roles by interacting with a large number of PP1 interacting proteins (PIPs). In budding yeast, there are at least 30 of these proteins identified for Glc7 (Cannon, 2010). PIPs can serve for the formation of multimeric complexes, as modulators of catalytic activity or as substrates themselves. However, their identification represents a great challenge because of the little similarity between their protein sequences (Ariño et al., 2019; Offley and Schmidt, 2019). PIPs tend to have long intrinsically disorder regions, this characteristic facilitates their wrapping around the PP1c and the formation of multiple contacts on the surface (Heroes et al., 2012)

Glc7 is a relatively stable protein (Nigavekar et al., 2002; Christiano et al., 2014), however little is known about the mechanisms that regulate its abundance in the cell. In mammals, PP1 has phosphorylatable residues in its C-terminal region (Ceulemans and Bollen, 2004), which inhibit the protein under certain conditions. Some phosphorylation and ubiquitination sites have recently been identified in Glc7 (Alburquerque et al., 2008; Chen et al., 2010; Swaney et al., 2013), but it remains unknown if these involve any mechanism to control its abundance or location in the cell.

3.2.1 RVxF-regulatory subunits

Binding to Glc7 can be mediated by short sequence motifs (4-8 residues) present on PIPs (Egloff et al., 1997; Terrak et al., 2004). This interaction takes place in the hydrophobic region of Glc7, which showed reduced affinity with some regulatory subunits,

when point mutations and deletions were done in this region (Wu and Tatchell, 2001). When looking for the specific PIPs motifs in budding yeast, no consensus sequence has been reported, but rather a series of sequences that varied in affinity for Glc7 (Ariño et al., 2019). In mammals, however, there are at least 10 PP1-docking motifs identified.

So far, only one motif has been characterized in yeast, the RVxF motif. An analysis of this motif in yeast PIPs revealed a relative consensus amino acid sequence [KR]xVRF, which is present in at least 10 of the Glc7-interacting proteins and is necessary for the stability of the multimeric complexes (Moorhead et al., 2007; Zhao and Lee, 1997; Hendrickx et al., 2009). Importantly, the conserved valine and phenylalanine residues are responsible for binding to Glc7, since mutations in these amino acids eliminated or reduced this binding (Chang et al., 2002; Kozubowski et al., 2003; Wu and Tatchell , 2001). It should be noted that the hydrophobic region of Glc7, is located on the opposite side of the catalytic site of the enzyme (Egloff et al., 1997; Terrak et al., 2004) , suggesting that binding to Glc7 is independent of its catalytic action (Figure 8).

Some PIPs contain more than one RVxF motif, which can contribute to binding at varying degrees. However, crystallographic structures predict that these motifs cannot bind Glc7 simultaneously, suggesting that PIPs with more than one RVxF motif could form different complexes with distinct specificity or localization (Akiyoshi et al., 2009). This correlates with the protein abundance of PIPs compared to Glc7, which suggests that these could compete with each other for a limited amount of the phosphatase (Ghaemmaghami et al., 2003).

Among the best characterized PIPs are Ref2, Gip2, Sla1, Bud14, Fin1 and Reg1. Additionally, the KHVCF sequence in Sen1 has been identified as a potential RVxF motif. Its interaction with Glc7 was tested in *SEN1* mutants via yeast two-hybrid assays, which showed that the mutation in phenylalanine (F2003A) resulted in a loss of Glc7 binding (Nedea et al., 2008). Thus confirming the importance of this residue in the RVxF motif for its interaction with Glc7 under non-physiological conditions and suggesting a possible function of Sen1 as PIP.

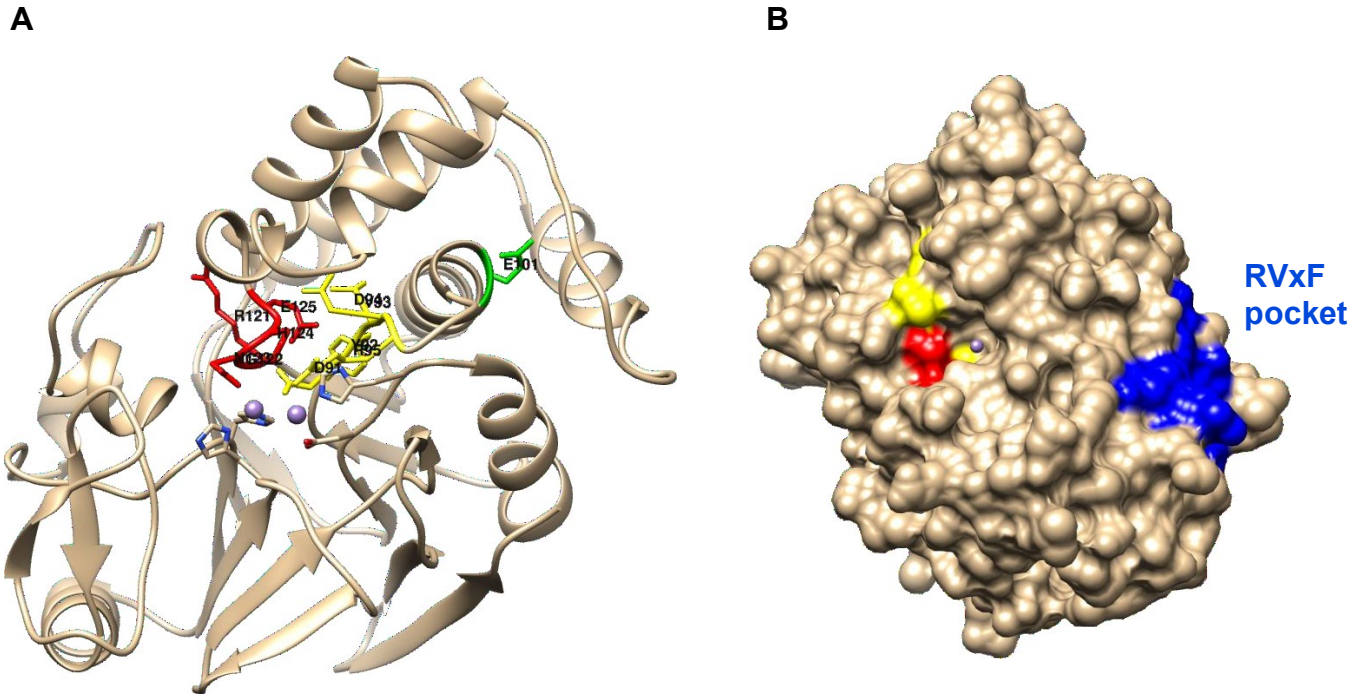


Figure 8. – Structure of Glc7 catalytic and hydrophobic domains. The Glc7 catalytic domains were modeled using the crystal structure of the catalytic subunit of human PP1 (PP1C-γ) (PDB Model: 1IT6). **(A)** The conserved catalytic domains are represented as in Figure 7. The two metal ions, Mn^{2+} and Fe^{2+} (purple spheres), are coordinated by the His, Asp, and Asn residues. **(B)** The catalytic domains are shown in red and yellow, while the Glc7 hydrophobic pocket in blue is on the opposite side.

3.2.2 Non-RVxF regulatory subunits

Although the RVxF motif has been the best characterized in terms of its ability to bind to the hydrophobic groove, not all PIPs possess it and therefore there are additional elements that are required for the interaction to take place. Due to the wide range of yeast PIPs and the difficulty in identifying their interaction motifs, there is no specific classification for regulatory subunits that have motifs different from RVxF. In fact, many of

these domains have been discovered by point mutations in regions of the Glc7 surface other than the hydrophobic groove, which affected its association with PIPs (Baker et al., 1997; Connor et al., 2000). For instance, Sds22 uses a motif based on leucine-rich repeats (Ceulemans et al., 2002; Ghosh et al., 2013) and other proteins such as Pti1, a component of CPF, despite interacting with Glc7 its interaction region has not been revealed, but clearly, it is not based on an RVxF motif (He and Moore, 2005).

Additionally, some of these PIPs can associate with Glc7 and form larger complexes. This is the case of the trimeric complex formed by Glc7-Sds22-Ypi1, which is responsible for translocating Glc7 to the nucleus (Pedelini et al., 2007). Another well-known complex is CPF, in which Glc7 interacts with Ref2 and Pta1 through RVxF motifs and with the non-RVxF motif protein, Pti1, for the 3' ends processing of pre-mRNAs in eukaryotes (He and Moore, 2005; Nedea et al., 2008). Recently, another complex known as APT (associated with Pta1) has been described, this involves Glc7, and only differs from the CPF phosphatase module in one subunit, Syc1. The APT complex would play a more important role in non-coding RNA production (Lidschreiber et al., 2018) (Figure 9).

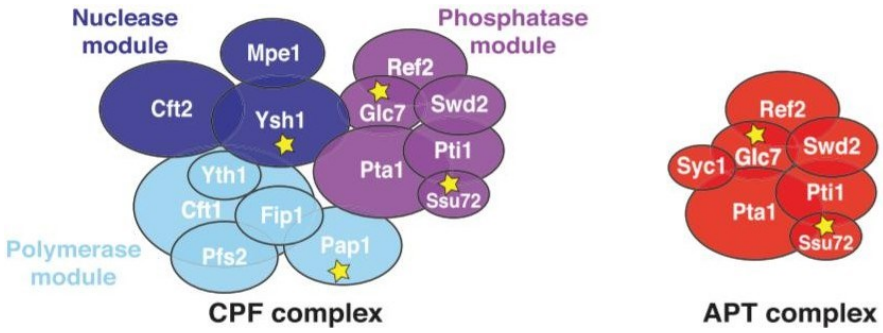


Figure 9. – Schematic representations of CPF and APT complexes. The APT contains the Syc1 subunit, which is not found in the CPF complex. Both are involved in transcription termination, either of protein-coding genes (CPF) or in non-coding RNAs (APT). Yellow stars denote enzymes. Taken from Lidschreiber et al., 2018.

3.2.3 Glc7 functions

Glc7 fulfills roles both inside and outside the nucleus. Among its cytosolic functions are those related to glucose repression, bud-site selection, endocytosis, and actin organization. However, most of its functions are fulfilled in the nucleus, these being related to transcriptional regulation (independent and dependent of the CPF complex), microtubule attachment to kinetochores, and cell cycle checkpoints.

3.2.3.1 Cytosolic functions

Glucose is par excellence the primary source of carbon in yeast. Because of this, under high glucose conditions, the transcription of many genes associated with different carbon source utilization is repressed. This repression is carried out mainly by the transcription factor Mig1 (Ahuatzi et al., 2007; Carlson, 1999). Briefly, when yeast cells grow in low glucose concentrations, the Snf1 kinase phosphorylates Mig1, which deactivates its capacity for transcriptional repression of its target genes. Under high glucose conditions, the Glc7-Reg1 holoenzyme deactivates the Snf1 kinase by dephosphorylating its Thr-210 residue, which generates a derepression of these genes (Sanz et al., 2000).

Similarly, to fulfill other roles Glc7 can bind to different regulatory subunits, thus forming different holoenzymes that intervene in the process of regulation of septins, bud-site selection, and actin organization. It is known that the association with Afr1 allows modulating the structure of septins in the cell cortex, to extend a projection directed toward the extracellular medium (Bharucha et al., 2008). By binding to the regulatory subunit Bud14, it can control bud-site selection and the shape of these by stabilizing microtubule attachments to the cell cortex thanks to its affinity for proteins located in this region (Ho et al., 2002; Knaus et al., 2005). Finally, Scd5-Glc7 is able to regulate endocytosis and actin organization through the binding and dephosphorylation of some endocytic proteins located in the cell cortex (Zeng et al., 2007)

3.2.3.3 Nuclear functions

The production of cleaved and polyadenylated mRNAs depends on the CPF complex, which includes Glc7 in its phosphatase module (Nedea et al., 2008) (Figure 8). Dissociation of Glc7 from the CPF complex results not only in termination defects but also destabilization of the Pta1 and Fip1 subunits and loss of polyadenylation activity (He and Moore, 2005). Pta1 has been identified as a substrate for Glc7, which is dephosphorylated to promote transcript polyadenylation. The importance of Glc7 lies in its ability to maintain the affinity of subunits with polyadenylation activity to the CPF. Additionally, it is known that the components Swd2 and Ref2 are responsible for maintaining Glc7 associated with the complex (Nedea et al., 2008).

Once polyadenylated, the mRNAs must be exported into the cytoplasm. This function depends, in part, on Glc7. mRNAs are recognized by an RNA binding protein, Npl3, which shuttles between the nucleus and the cytoplasm. Glc7 is responsible for dephosphorylating Npl3 in the nucleus, which increases its affinity for the transcript and facilitates transport to the cytoplasm (Gilbert and Guthrie, 2004). However, it is unknown whether Glc7 performs this action by itself, as part of the CPF complex or in association with a different holoenzyme.

Glc7, via the CPF multimeric complex, also dephosphorylates CTD Tyr1 (Tyr1) at the end of the transcription cycle to promote transcription termination. It is thought that this action on Tyr1 could promote the recruitment of termination factors such as Pcf11 or Rtt103 (Schreieck et al., 2014). Additionaly, we and others have previously shown that Glc7 anchor-away causes a strong termination defect on genes of the NNS pathway (Collin et al., 2019; Schreieck et al., 2014). Tyr1 also appears to play a critical role in NNS termination by mediating RNAPII pausing at the 5' end of non-coding genes. It has been proposed that RNAPII pausing is necessary for Sen1, to catch up RNAPII and thus carry out transcription termination in these genes (Collin et al., 2019; Han et al., 2017). Interestingly, these defects in non-coding gene termination have been observed by mutating Tyr1 to phosphomimetic or non-phosphorylatable residues, suggesting a possible phosphorylation/dephosphorylation role of Tyr1 in the process (Collin et al., 2019). However, this NNS termination deficiency may, in the same way, be associated

with Glc7 action on other substrates. Specifically, it has been proposed that Glc7 can be recruited to the NNS complex and thus dephosphorylate the Sen1 helicase, which would allow its activation to carry out transcription termination (Nedea et al., 2008). Yet, many of these studies were performed using *in vitro* models or overexpressing Sen1 due to its low abundance in the cell, thus no *in vivo* evidence has demonstrated this.

Glc7, as part of the APT complex, is involved in snoRNA transcription. The APT complex has a unique subunit that differs from the phosphatase module of CPF, Syc1, which is recruited and preferentially interacts with sn/snoRNAs (Lidschreiber et al., 2018). Although the APT complex, through its Syc1 subunit, engages the RNA during transcription, there are no transcription termination defects in *SYC1Δ* cells (Lidschreiber et al., 2018). However, ncRNAs termination defects occur when depleting Glc7 from the APT, which can be corrected when the C-terminal domain of Sen1 is overexpressed (Nedea et al., 2008). Thus suggesting a possible role between the APT complex, through its Glc7 subunit, and Sen1 in regulating ncRNA transcription termination.

Most of the PP1 proteins are essential for cell cycle progression (Stark, 1996). The presence of checkpoints allows the cell cycle to take place only when certain specific conditions have been met, thus ensuring precision at each step. The regulation of each checkpoint is due, in part, to the phosphorylation of proteins, which promote cell cycle arrest (Cannon, 2010). There are at least three checkpoints in which Glc7's participation has been confirmed, being its main role to revert the spindle and DNA damage checkpoint and one meiotic checkpoint. Specifically, Glc7 antagonizes the action of several kinases by dephosphorylating proteins involved in these checkpoints (Akiyoshi et al., 2009). For instance, several protein kinases phosphorylate proteins such as Rad53 and histone H2A (also encoded by Hta1) once DNA damage is detected (Longhese et al., 2006). Histone H2A phosphorylation allows to recruit repair factors and thus repair the damage. Once repaired, histone H2A is dephosphorylated by Glc7, which allows the cell to continue dividing (Bazzi et al., 2010).

In fission yeast, one of the PP1 subunits is Dis2. Dis2 governs dephosphorylation of the conserved elongation factor Spt5. Spt5 dephosphorylation allows the transition from elongation to termination by RNAPII, although it is not very clear how (Parua et al., 2018).

One possibility is that Dis2 dephosphorylates CTD Thr4 at the end of transcription, which in turn promotes the recruitment of Seb1, a termination factor (Keckman et al., 2018).

4. Phospho-regulation of proteins by the Ubiquitin-Proteasome System

Protein phosphorylation is one of the most common PTMs. In budding yeast, there are about 2 300 phosphoproteins and approximately 40 000 identified phosphorylation sites involved in different processes (Bodenmiller et al., 2010; Lanz et al., 2021). Phosphorylation can affect protein regulation through changes in protein abundance or availability, it can determine their stability, interactions, cell location, and even their enzymatic activity against other substrates (Johnson, 2009).

Despite being the most common PTM in yeast, many of these still do not have a defined biological function. To solve this, several approaches have been developed. A classic way to biochemically corroborate the biological relevance of a phosphorylation site is by inducing mutations that mimic or prevent phosphorylation. These phosphomutants proteins can inform us about the importance of different phosphorylation sites throughout a protein. More recent studies take advantage of different chemistries techniques to enrich phosphorylated peptides, which are then detected by mass spectrometry (MS). Among these, the generation of an in-depth phosphoproteome dataset for budding yeast stands out (Lanz et al., 2021; Swaney et al., 2013).

4.1 Ubiquitination

Ubiquitin is a 76 amino acid protein that can be conjugated to protein lysine residues (Peng et al. 2003). The protein modification, known as ubiquitination, can alter the activity and stability of the substrate through different processes. Once bound to the

protein, ubiquitin can be itself coupled to additional ubiquitin monomers at one of its 7 lysines (Lys), thus forming polyubiquitin chains (Kwon and Ciechanover, 2016).

The ubiquitination mark generally directs proteins to the proteasome for degradation, however many of these modifications can act through non-proteolytic mechanisms. This is due to the nature of the ubiquitin chain assembly. A different outcome will appear if a single ubiquitin is attached to a protein (monoubiquitination), if a protein is polyubiquitinated or modified by phosphorylation or acetylation (Kwon and Ciechanover, 2016).

These various ways of conjugating ubiquitin make possible the regulation and signaling of multiple proteins in eukaryotes. In general, monoubiquitination is associated with non-proteolytic processes that acts as a mark that regulates the activity of the protein, however, in some cases it can lead to protein degradation by the proteasome (Braten et al., 2016). On the other hand, poly ubiquitination can have different outcomes, ubiquitins bound through Lys48 target the protein to degradation by the proteasome, while Lys63 linkages can also modulate non-proteolytic processes such as DNA repair and membrane proteins trafficking (Kwon and Ciechanover, 2016).

4.1.1 Ubiquitin-Protein conjugation

Ubiquitin conjugation involves the action of three enzymes: the ubiquitin-activating enzyme (E1), the ubiquitin-transferring enzyme (E2), and the ubiquitin ligase (E3), known as the E1 – E2 – E3 cascade (Ciechanover, 2015). The process begins with the formation and transferring of covalent bonds between the ubiquitin (Ub) and the ubiquitin-activating enzyme E1, which requires ATP hydrolysis. The activated ubiquitin is then transferred to E2 enzymes, and together with E3 enzymes, they catalyze the formation of the isopeptide bond between Ub and the Lys residue of the substrate (Hershko and Ciechanover, 1998).

4.1.2 E1-E2-E3 enzymes

In budding yeast, the activating enzyme E1 is encoded by a single *UBA1* gene (McGrath et al., 1991), while there are at least 11 genes that encode conjugating enzymes (E2), known as *UBC* genes. The E3 ligases constitute the largest group of proteins in the ubiquitination process since they mediate the specificity and selectivity of the substrate ubiquitination. About 100 potential E3 enzymes in yeast can be classified into two groups: RING or HECT, which differ in their mechanism for catalyzing the reaction. Substrate recruitment by the RING E3 ligases is carried out either by binding the substrate directly to RING domains or by forming multimeric receptors that contain RING domains (Deshaies and Joazeiro, 2009).

4.1.2.1 The APC ubiquitin ligase

The anaphase-promoting complex (APC), which belongs to the RING E3 group, is composed of 13 subunits and its main function is to trigger the degradation of multiple proteins starting in mitosis, continuing throughout the G1 phase until the cells reach the S phase, where it is degraded (Pesin and Orr-Weaver, 2008; Zachariae, et al., 1998). APC recognizes its substrates in association with one of its two co-activators, either Cdh1 or Cdc20 during the cell cycle (Visitin et al., 1997). Binding of APC to Cdh1 is controlled by phosphorylation, once bound, the complex targets the protein for proteasome degradation. Among the most prominent targets are the mitotic cyclins, cyclin A and cyclin B, which drive mitotic entry (Peters, 2006). These substrates usually have degradation motifs that allow their recognition by APC, for instance, the destruction box (D-box, RxxLxxxxN) and KEN box (RxxxxxKEN) represent the best characterized degradation motifs (Glotzer et al. 1991; Pfleger and Kirschner, 2000; Sivakumar and Gorbsky, 2015). Interestingly, it has been shown that APC-Cdh1 is responsible for Sen1 degradation during G1 (Mischo et al., 2018), however, none of these recognition motifs has been identified in Sen1, thus the motif that targets Sen1 for degradation remains unknown.

4.2 The Ubiquitin-proteasome system

Protein homeostasis is dictated by two different systems in eukaryotic cells: the Ubiquitin-proteasome system (UPS) and autophagy. The UPS allows proteolysis by specifically targeting ubiquitylated proteins to the proteasome (Hershko and Ciechanover, 1998). On the other hand, intracellular proteins can also be degraded by autophagy, which involves the use of vacuoles or lysosomes (Nakatogawa et al. 2009). Both mechanisms are responsible for the regulation of intracellular protein levels, the proteasome being responsible for 80-90% of proteolysis, while autophagy for 10-20% (Hershko and Ciechanover, 1998).

Target processes of the UPS include proteins involved in cell cycle regulation, antigen processing, and removal of misfolded proteins. Many of them are marked with one or more ubiquitins, which are then recognized by the proteasome. This stage of target recognition activates the ubiquitin-cascade system (described above), which is followed by degradation by the proteasome. It is important to note that the ubiquitination of a protein does not guarantee its degradation, and proteolysis by the UPS is only one of the possible outcomes (Geng et al., 2012)

4.2.1 The proteasome

The proteasome is a multimeric complex consisting of at least 33 different subunits (Finley, 2009). Proteasomes are organized into two different compartments, the 20S core particle (CP) and one or two 19S regulatory particles (RP) flanking the CP (Figure 10). The union of one or two RP forms the 26S and 30S proteasomes, respectively (Finley et al., 2016). The RP is further subdivided into lid and base components and recognizes the substrates to be degraded; while the CP harbors the catalytic active site. This proteolytic site is found inside the CP, which allows the process to be under control and prevents unspecific degradation (Finley et al., 2012).

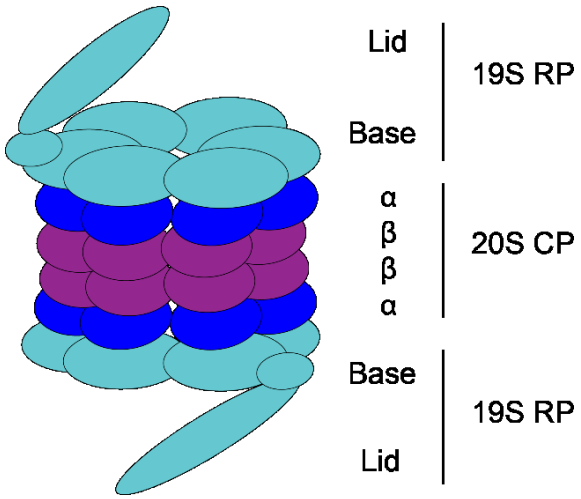


Figure 10. – Schematic representation of the proteasome structure. The proteasome has a cylindrical structure composed of α and β rings (20S CP), this core particle contains the catalytic site where proteins are degraded. The 19S regulatory particles (RP) are composed of Base and Lid, which recognize the polyubiquitin chain of the protein and initiate the degradation process.

The RP is composed of 19 subunits, nine of them belong to the lid and 10 to the base subcompartments (Glickman et al., 1998; Finley 2009). Potential substrates are recognized by at least 13 components of the RP (Rpn1-Rpn13), which bind and/or process the ubiquitin chain of the proteins. For instance, the Rpn10 and Rpn13 subunits recognize their substrates via its α-helical ubiquitin-interacting motif (UIM) and pleckstrin-like receptor of ubiquitin (PRU) domain, respectively (Mayor et al., 2007; Husnjak et al., 2008). To be degraded, substrates must be translocated to the CP catalytic site via a narrow channel and globular proteins are unfolded before being translocated. Both processes, unfolding and translocation are carried out by the six ATPases located in the base (Rpt1-Rpt6 subunits) through ATP hydrolysis (Sauer and Baker, 2011) and the isopeptidase activity of the Rpn11 subunit located in the lid.

The CP or 20S proteasome has 14 subunits, which form a barrel-like structure of four staggered rings (Groll et al., 1997). These rings can be either α or β type, which are found outside and inside the structure, respectively. Also, inside this structure is the proteolytic site of the proteasome. This catalytic activity resides in the β rings, which are

classified as threonine proteases, typically generating peptides of between 4 to 25 amino acids in length (Groll et al, 2005). The α rings regulate the access of the substrate to the interior of the 20S and serve as a docking site for the RP proteins which also allow the passage of the substrate (Finley, 2009).

4.2.2 Phospho-degrons

Phosphodegrons are the best example of cross-talk between phosphorylation and ubiquitynation, where one or more phosphorylation sites can induce the consecutive ubiquitination of a protein, eventually leading to degradation (Swaney et al., 2013) (Figure 11).

The abundance of proteins in the cell is regulated by the rates of synthesis and degradation. This degradation, in many cases, is carried out by the UPS, which determines the half-life time of the proteins through the interaction between substrates and E3 ligases. In this sense, phosphodegrons are key to coordinate protein degradation via the UPS.

Phosphodegrons can be classified into two groups: inherent and acquired (Ella et al., 2019). Inherent degrons are always present and are generally exposed when the proteins are improperly folded. This exposure can trigger its degradation by the UPS (Shiber et al., 2014). The destruction boxes (D-boxes) present in cyclins are an example of inherent degron sequences (Barford et al., 2011). On the other hand, acquired degrons are the result of the cross-talk between phosphorylation and ubiquitination, and they occur transiently in proteins, that is, under certain biological conditions. These degrons have been identified in proteins that undergo a rapid variation in their levels, such as those that promote cell-cycle progression (Skowyra et al., 1997).

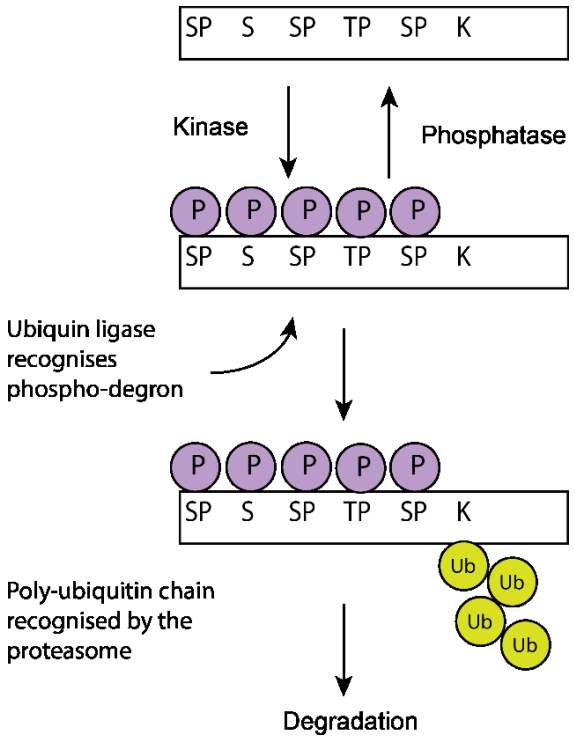


Figure 11. – Functioning of the phosphodegron motif. Phosphodegrons are short sequences of amino acids where the phosphorylation of certain residues induces the subsequent ubiquitination of the protein, which leads to its degradation. Taken from Holt et al., 2012

4.2.2.1 Sen1 has potential phosphodegron motifs

An important characteristic of degrons is that they are recognized by Ub-ligases to trigger their degradation (Hershko and Ciechnaover, 1998). Therefore, the discovery of them is also linked to the identification of their specific E3 ligase. In the case of Sen1, the APC complex has been identified as the Ub-ligase responsible for targeting it to the proteasome (Mischo et al., 2018). Given this preliminary evidence, Sen1 could be a carrier of one or more phosphodegron motifs.

Recent studies have identified thousands of phosphorylation sites in yeast proteins through proteomic approaches (Lanz et al., 2021). Within these, many sites were previously reported as potential phosphodegron motifs (Swaney et al., 2013). This last study used a proteomic-based approach with two enrichment strategies that allowed

identifying co-modified proteins (phosphorylated and ubiquitinated). Among these, a potential phosphodegron motif was reported in Sen1, where serine at position 863 (S863) was phosphorylated (Swaney et al., 2013). Thus, the Sen1-S863A mutant could be evaluated to see whether it makes the protein stable.

Sen1 levels are known to be kept low by the proteasome (DeMarini et al., 1995). In fact, Sen1 levels fluctuate according to the requirements of the cell, specifically during the cell cycle. That is, during the S phase, when the transcription and replication machinery is in place, the abundance of Sen1 increases compared to that found in the G1 phase, where Sen1 is degraded by UPS (Mischo et al., 2018). Additionally, altering the degradation of Sen1 leads to cell death due to a reduction in the termination efficiency and ncRNA abundance (Mischo et al., 2018). Therefore, the regulation of Sen1 levels during the cell cycle is crucial for the proper functioning of the cell.

Interestingly, a previous study showed that deleting the 40 residues within amino acids 459 to 498 of Sen1 (*sen1*-459-498Δ) leads to the stabilization of the protein in the G1 phase, where it is usually degraded (Mischo et al., 2018). This region is located in the N-terminal domain and a KEN box was identified within this region (residues 480-493), however, alanine substitutions in this box failed to stabilize the protein. This suggests that there might be additional residues within this motif that could stabilize the Sen1 levels.

5. Hypothesis and aims

Recently, it has been demonstrated that Sen1 protein levels are regulated in a proteasome-dependent manner during the G1 phase of the cell cycle (Mischo et al., 2018). Nonetheless, this study was carried out using a Sen1 overexpression system to overcome the problems related to its low protein level. Regarding the regulation of Sen1, a previous member of the lab showed that Sen1 is degraded under physiological conditions via a proteasome-dependent pathway (Collin, 2019). This mechanism is associated with Sen1 stability in the nucleus and relies on the action of Glc7, a protein phosphatase that was previously shown to dephosphorylate Sen1 *in vitro*, and to interact with Sen1 in two-hybrid experiments (Nedea et al., 2008).

Our data so far suggest that the dephosphorylation of Sen1 by Glc7 prevents its proteasome-mediated degradation. Interestingly, a potential site in the N-terminal region of Sen1 (S863) that may work as a phospho-degron has been identified in a proteome-wide analysis of phosphorylation and ubiquitylation cross-talk (Swaney et al., 2013). Additionally, others have shown that an F2003A mutation in Sen1 (phenylalanine to alanine at position 2003) prevents its association with Glc7 (Nedea et al., 2008). If the effect of Glc7 on Sen1 is direct (by dephosphorylating Sen1 on a phospho-degron), we would predict that preventing Glc7 from interacting with Sen1 should lead to destabilization of it.

For these reasons, our working hypothesis is that Glc7 (and a yet-to-be-identified kinase) controls Sen1 stability via dephosphorylation of a phospho-degron. In order to answer this research hypothesis, two objectives have been established:

1. To identify the Sen1 phospho-degron

2. To test whether the interaction of Sen1 with Glc7 affects its stability

MATERIAL AND METHODS

1. Yeast strains

All the experiments described here were carried out in the model organism *Saccharomyces cerevisiae*. All epitope tags were introduced at the C-terminal ends of endogenous genes. Generally, one colony was inoculated in 5 ml YPD medium (Yeast extract peptone) supplemented with 2% glucose for pre-culture grown overnight at 30°C. Cell density was measured using a spectrophotometer at OD_{600} and cells were diluted to OD_{600} of 0.05 in 50 or 100 ml YPD medium (depending on the experiment) and grown until OD_{600} reached 0.9.

1.1 Generation of epitope-tagged yeast strains

Yeast strains used and generated in this study have a W303 background and are listed in Table 2 (Annex A). Typically, epitope tagging was achieved through the introduction of exogenous DNA by homologous recombination (Gietz and Woods, 2002). Plasmids pMPY-3xMYC (Schneider et al., 1995), p3FLAG-KanMX (Gelbart et al., 2001), and pFA6a-FRB-KanMX6 (Haruki et al., 2008) were used to add the MYC, FLAG, and FRB protein tags, respectively. Briefly, for Myc-tagged genes, a PCR product corresponding to the tag along with flanking regions homologous to the upstream and downstream region of *SEN1* stop codon was generated using pMPY-3xMYC as a template. The amplified product was then transformed into yeast cells following Gietz and Woods's protocol. The same procedure was followed to add the FLAG and FRB tags.

All strains were genotyped by PCR, as well as validated by immunoblotting and sequencing.

2. Molecular Cloning

In order to test different *SEN1* mutations, plasmids containing the wild-type (WT) version and mutants were cloned into the donor vector pRS313 (CEN, HIS3) (Sikorski and Hieter, 1989) following two steps. First, the coding sequence was cloned together with the promoter and terminator regions, then, the epitope tag was added at the C-terminal end by cloning, thus allowing the expression of the protein in fusion with the hemagglutinin (HA) tag. Mutations were introduced using inverse PCR. *SEN1* plasmids used and generated are listed in Table 1 (Annex A).

For experiments involving the expression of Sen1 plasmids, strains were grown in YNB medium (Yeast Nitrogen Base) lacking histidine and supplemented with 2% glucose at 30°C. Cell culture conditions were similar to those described in Section 1, pre-cultures were grown for 24 h instead of overnight.

2.1 Cloning *SEN1*-WT-3xHA into pRS313

2.1.1 Cloning the whole *SEN1* sequence into pRS313

The whole sequence of the *SEN1* gene was amplified by PCR from the yFR116 strain (W303 background). This strain contains the endogenous wild-type version of *SEN1* (Annex B). The total amplified sequence has a length of 7 889 base pairs (bp), where 6 696 bp corresponds to the coding region of *SEN1*, together with 589 bp of the promoter region and 604 bp of the terminator.

The oligonucleotides used for the amplification (Table 3, Annex A) contained the restriction sites sequence for SacII and XhoI, upstream and downstream the region to be amplified. The PCR amplification reaction was carried out using the KOD Hot Start DNA Polymerase enzyme (MilliporeSigma) according to the manufacturer's recommendations. The PCR product was then purified using the QIAquick PCR Purification Kit (Qiagen). Both the purified product and the plasmid pRS313 were digested with the restriction

enzymes SacII and XhoI (New England Biolabs) in CutSmart digestion buffer (New England Biolabs) overnight at 37 °C. Both fragments were purified on agarose gel 1% using the QIAquick Gel Extraction Kit (Qiagen) before the ligation reaction.

The ligation of the PCR product and pRS313 followed a 5: 1 ratio (insert: vector) using T4 DNA ligase (Invitrogen) for 3 hours at 16 °C. The pRS313-*SEN1*-WT (pFR711) plasmid was screened by enzymatic digestion and sequenced in the Molecular Biology and functional genomics Core Facility at IRCM.

2.1.2 Adding HA tag at the C-terminal end of pFR711

The pFR711 plasmid was cloned in order to express the 3xHA protein tag. For this purpose, we took advantage of the yeast strain *SEN1*-3xHA (yFR3195), which has the HA tag integrated in frame at the 3' end of *SEN1*. A pair of oligonucleotides were used to amplify a region of the C-terminal domain (including the tag and part of the terminator sequence), which was 1 722 bp in length. These primers contained the restriction sites for NcoI and XhoI in the forward and reverse, respectively. The reaction was carried out by the KOD Hot Start DNA Polymerase enzyme (MilliporeSigma) according to the manufacturer's recommendations. Both the PCR product and pRS313-*SEN1*-WT were digested with NcoI and XhoI (New England Biolabs) using the CutSmart Buffer (New England Biolabs) overnight at 37 °C. The digested plasmid was purified on agarose gel using the QIAquick Gel Extraction Kit (Qiagen) and the digested PCR product was cleaned using the QIAquick PCR Purification Kit (Qiagen). A ligation reaction as previously described (section 2.1.1) was carried out to insert the C-terminal end and the HA tag downstream of the *SEN1* coding sequence in pFR711. The pRS313-*SEN1*-WT-3xHA (pFR737, Annex C) was screened by enzymatic digestion and sequenced in the Molecular Biology and functional genomics Core Facility at IRCM.

2.2 Generating *sen1-F2003A*-3xHA mutant

2.2.1 Subcloning *SEN1* fragment into pUC18

In order to generate the *sen1*-F2003A mutant, we used *SEN1*-WT-3xHA (pFR737) as a template. The idea was to subclone a fragment of interest in the pUC18 plasmid to generate the mutation at residue 2 003 (position 6 009) and then clone the fragment back to the original plasmid.

Due to the large size of this plasmid (13 kb), we decided to work with a smaller fragment that comprised the site to be mutated (F2003). This fragment is located in the C-terminal domain, so we used the restriction enzyme SexAI to clone a 2.5 kb region into a smaller plasmid such as pUC18.

Oligonucleotides were used to add the SexAI restriction site sequence into pUC18 by PCR. The product was purified using the QIAquick PCR Purification Kit (Qiagen). Both the newly generated pUC18 and pFR737 were digested with SexAI in CutSmart Buffer digestion buffer (New England Biolabs) overnight at 37 °C. The pUC18 plasmid was dephosphorylated by the Calf Intestinal Alkaline Phosphatase (CIAP, Invitrogen) for 30 minutes at 37 °C and 30 minutes at 50 °C. 10mM EDTA was added and a 15 minutes incubation was carried out at 65 °C to inactivate the enzymes. Both pUC18 and pFR737 were purified on agarose gel with the QIAquick Gel Extraction Kit (Qiagen) and the ligation reaction was similar to that already described (section 2.1.1). The generated plasmid was screened by enzymatic digestion.

2.2.2 Mutating *sen1*-F2003A

We introduced the mutation by changing the 2 003 residue (phenylalanine for alanine). For this, amplification was carried out by inversed PCR, with the forward primer containing the mutation. This same oligonucleotide was phosphorylated at its 5' end in

order to perform the ligation reaction. The PCR product was purified on agarose gel and then re-ligated on itself using the T4 DNA ligase (Invitrogen) for 3 hours at 16 °C. The generated plasmid (pFR830) was screened by enzymatic digestion and sequenced in the Molecular Biology and functional genomics Core Facility at IRCM.

2.2.3 Subcloning the *SEN1* fragment back to pFR737

The fragment containing the *SEN1* mutation was cloned back into *SEN1*-WT-3xHA (pFR737) to generate the s*en1*-F2003A-3xHA mutant (pFR850). The procedure followed was the same as previously described (section 2.2.1), using the SexAI restriction enzyme (New England Biolabs). The pFR850 plasmid was screened by enzymatic digestion and sequenced in the Molecular Biology and functional genomics Core Facility at IRCM.

3. Protein extraction

Cells expressing the chromosomal *SEN1*-WT-3xMyc allele and any of the other *SEN1*-WT-3xHA or *sen1*-F2003A-3xHA alleles were grown in either YPD or YNB -HIS following the procedure described in section 1. When cells reached OD_{600} of 0.9, they were washed with ice-cold water, resuspended in 700 µL of Lysis buffer (50 mM HEPES-KOH 1 M pH 7.5, 140 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% Na-deoxycholate) supplemented with protease inhibitors (1 mM PMSF, 1 mM benzamidine, 10 µg/mL aprotinin, 1 µg/mL leupeptin and 1 µg/mL pepstatin). Cells were then lysed by bead beating and lysates were centrifuged 5 min at 4 °C. To enrich Sen1 protein levels, 650 µL of whole-cell extract (WCE) was incubated with 50 µL of Dynabeads coated with Pan Mouse IgG antibodies (Invitrogen), pre-coupled with 2.5 µL of anti-Myc 9E10 (Collin et al., 2019) or 3 µg of anti-HA F7 (Santa Cruz Biotechnology) antibody overnight at 4°C with agitation. Beads were then washed six times with Lysis buffer.

4. Protein detection

4.1 Polyacrylamide gel electrophoresis (SDS-page)

The beads were incubated at 95 °C for 5 minutes in 50 µL of 1X Laemmli buffer (250 mM Tris-HCl pH 6.8, 10% SDS, 50% Glycerol, 0.05% bromophenol blue, 5% β-mercaptoethanol). 10 ul of 5X Laemmli buffer were added to 40 µL of supernatant, which preceded the incubation at 95 °C for 5 minutes. The samples were briefly centrifuged and separated by SDS-PAGE gel. A resolving gel (6% acrylamide: bis, 0.37 M Tris pH 8.8, 0.1% SDS) and a stacking gel (5% acrylamide: bis, 0.37 M Tris pH 6.8, 0.1% SDS) were poured in that order. Migration was done with a 1X electrophoresis buffer (30 mM Tris, 144 mM glycine, 0.1% SDS) at 130 v for 100 minutes.

4.2 Protein transfer and detection

The proteins were transferred on a nitrocellulose membrane (Bio-rad). The transfer was carried out at 80 v for 120 minutes using the Mini Trans-Blot system (Bio-rad) in a 1X transfer buffer (25 mM Tris, 200 mM glycine, and 20% ethanol)

Membranes were incubated in 5% (w/v) milk or BSA in TBS 1X (20mM Tris pH 7.5, 65 mM NaCl) supplemented with 0.1% Tween-20 (TBST) containing primary antibody (anti-Myc 9E10 (1: 2 000), anti-HA F7 (1: 1 000), anti-FLAG (1: 2 000) or anti-Spt16 (1: 1 000)) (Table 1) overnight at 4°C. Membranes were washed three times with TBST and incubated for 30 minutes at room temperature with secondary antibody coupled to anti-mouse IRDye 680RD and anti-rabbit IRDye 800CW (LI-COR Biosciences), diluted in TBST plus 5% (w/v) milk. Membranes were washed with TBST and scanned on the Odyssey infrared imaging system (LI-COR Biosciences).

Table 1. – List of antibodies used for the detection of proteins by immunoblotting

Protein	Primary Antibody	Secondary Antibody
3xMyc	Mouse monoclonal anti-Myc clone 9E10 (Collin et al., 2019)	IRDye 680RD Donkey anti-mouse (Li-Cor Biosciences)
3xHA	Mouse monoclonal anti-HA clone F7 (Santa Cruz Biotechnology)	IRDye 680RD Donkey anti-mouse (Li-Cor Biosciences)
3xFLAG	Mouse monoclonal anti-FLAG clone M2 (Sigma)	IRDye 680RD Donkey anti-mouse (Li-Cor Biosciences)
Spt16	Rabbit polyclonal antibody (Tim Formosa Lab)	IRDye 800CW Donkey anti-rabbit (Li-Cor Biosciences)

4.3 Silver-stained SDS polyacrylamide gels

After the protein samples were run by electrophoresis on a 6% polyacrylamide gel, the gel was first fixed in a fixation solution (50% methanol, 10% acetic acid) stirring overnight. The next day the gel was rinsed with 20% ethanol for 20 minutes, followed by a second wash of water for 20 minutes. After the washes, the gel was reduced for 2 minutes at room temperature in sodium thiosulfate (0.2 g/L). Two water washes of the gel were then followed by a 30 minutes incubation in silver nitrate (2 g/L). After this incubation, the gel was rinsed in water and the stain was developed using a 5 % formaldehyde solution. The development was controlled to prevent overstaining and it was stopped by exchanging the developing solution with 1% acetic acid for at least 30 minutes. The stained gel was immediately photographed.

5. Assays for Sen1 expression and stability

To assess s*en1*-F2003A-3xHA (pFR850) stability, we used the Complementation After Nuclear Depletion (CAND) system, an *in vivo* approach of ectopic expression (Jeronimo et al., 2014), which is described in the next section. The strains were generated by transforming yFR116 with the different Sen1 plasmids. Additionally, as a control, the CEN/HIS3 plasmid pRS313 was transformed. It should be noted that all Sen1 plasmids contain the endogenous promoter and terminator of the *SEN1* gene.

5.1 Anchor-away

We have used endogenous Sen1-3xMyc and mutants in two *in vivo* systems, which took advantage of the anchor-away approach (Haruki et al., 2008). In the anchor-away system, the target protein is tagged with the FRB epitope at the C-terminus. This allows the tagged protein to exit the nucleus using a rapamycin treatment. The technique is based on the dimerization of the FRB domain of the mTOR protein with the FKBP12 protein in the presence of rapamycin. FKBP12 is bound to the ribosomal protein RPL13A, which is characterized by fluctuating between the nucleus and the cytoplasm during its maturation process. By adding rapamycin, the FRB tag and RPL13A-FKB12 are fused and this complex is translocated to the cytoplasm after approximately 90 minutes. In order to suppress the potential toxic effect of rapamycin, it is important to mutate the homologous gene of mTOR (Tor1) and the homolog of FKBP12 (Frp1) to avoid competition with the system. Using the anchor-away approach mainly allows evaluating possible phenotypes.

The second system, CAND, is based on the anchor-away technique. In this case, the HA-tagged Sen1 plasmids were ectopically expressed in a strain where the endogenous *SEN1* was tagged with FRB. The presence of the endogenous gene makes

it possible to guarantee the survival of the cell, even when the ectopic alleles could be potentially lethal. This allows the analysis to be independent of the endogenous protein.

5.2 Serial-growth dilution assay

Yeast cells expressing the endogenous Sen1-FRB tag were transformed with ectopic plasmids carrying Sen1-WT-3xHA or Sen1-F3003A-3xHA and grown in 5 ml YNB HIS- for 48h. The cultures were brought to the same density (OD_{600} = 1), which was followed by 10 fold serial dilutions. Successively, these dilutions were spotted on YNB HIS- or YNB HIS- + rapamycin (90 µg/mL) plates, incubated at 30 °C and checked daily.

6. Interaction validation by Co-immunoprecipitation

Strains carrying Glc7-3xFLAG were transformed with the Sen1-WT-3xHA plasmid (pFR737). The specificity of the interaction tested by co-immunoprecipitation (co-IP) was validated by transforming pFR737 into strains non-tagged with FLAG and by reverse co-IP.

Each of the strains grew in a final volume of 50 ml YNB HIS- at 30 °C before proceeding to the co-IP experiment. The cells were collected at an optical density at 600 nm of 0.8 to 0.9 and washed once with cold water. For the Glc7-3xFLAG co-IP, cells were lysed in 700 µl of Co-IP buffer (50 mM HEPES pH 7.5, 100 mM NaCl, 20% glycerol, 0.5 mM EDTA and 0.05% NP-40) supplemented with protease inhibitors (1 mM PMSF, 1 mM benzamidine, 10 µg/mL aprotinin, 1 µg/mL leupeptin and 1 µg/mL pepstatin) and 1mM DTT just before use. Cells were then lysed by bead beating for 5 minutes and lysates were put into ice-water bath for 5 minutes. The samples were centrifuged for 15 minutes at 4 °C at 18 000 x *g* to eliminate cell debris. We used 50 µl of Dynabeads coated with Pan Mouse IgG antibodies (Invitrogen), pre-coupled with 5 µg of anti-FLAG M2 (Sigma)

antibody overnight at 4 °C with agitation, which was washed with the co-IP buffer before incubation with the lysates. An approximate amount of 3-3.5 mg of protein from the total extract (WCE = 600 µl) was used for the co-IP. Likewise, an aliquot of 10% (60 µl) of this quantity (Input) was conserved for the analysis by immunoblotting. The beads were incubated with the WCE by rotation for 1 hour at 4 °C. Subsequently, the beads were washed three times with 1 ml co-IP buffer before releasing the affinity-purified proteins.

For the HA co-IP, we used the co-IP buffer II (10mM Tris-HCl pH 7.9, 100mM NaCl, 20% glycerol, 0.2 mM EDTA and 0.1% Triton x-100) supplemented with protease inhibitors (1 mM PMSF , 1 mM benzamidine, 10 µg/mL aprotinin, 1 µg/mL leupeptin and 1 µg/mL pepstatin) and 0.5 mM DTT just before use (Krogan et al., 2002). Cells were grown and lysed following the same procedure indicated above. For the HA co-IP we used 50 µl of Dynabeads coated with Pan Mouse IgG antibodies (Invitrogen), pre-coupled with 3 µg of anti-HA F7 (Santa Cruz Biotechnology) antibody overnight at 4 °C with agitation, which was washed with the co-IP buffer II. The incubation of 600 µl of WCE with beads was carried out with agitation for 1 hour at 4 °C. The washes that followed the co-IP were performed in the same way as for the co-IP FLAG.

The co-PIs were analyzed by immunoblotting, using the FLAG and HA antibodies described in section 4.2. First, the proteins were eluted from the beads by adding 1X Laemmli buffer and heating them at 95 °C for 5 minutes. Similarly, 15 µl of 5X Laemmli buffer was added to the Input samples. An amount of the total extract corresponding to 5% of the amount used for the co-IP was deposited on a 6% polyacrylamide SDS gel, as well as 50% of the eluate from the beads. The following steps were carried out as described in section 4 of this chapter.

7. Mass spectrometry

7.1 Yeast culture and treatments:

Yeast strain yFR3097 (Table 3, Annex A) growth conditions for proteomics were similar to those described in section 1. This strain contains the Sen1-3xMYC and Glc7-FRB tagged proteins to be able to perform the anchor-away experiments. When cells reached OD_{600} = 0.5, they were treated (or not) with 90 µg/mL rapamycin for 90 minutes. Additionally, the entire *PDR5* gene was replaced with the LEU2 selectable marker by homologous recombination. This was done to improve proteasome inhibitor effectiveness. The cells were treated with a cocktail of proteasome inhibitors (PS341 150 µM and MG115 150 µM) for 30 minutes before rapamycin treatment. Cells were washed twice with ice-cold water and each pellet was resuspended in 700 µL of Lysis Buffer containing protease inhibitors. Cells were lysed by bead beating for 5 minutes and lysates were centrifuged for 5 minutes at 4 °C. The supernatant was recovered and affinity purification was carried out.

7.2 Sen1 affinity purification

50 µL of Dynabeads coated with Pan Mouse IgG antibodies (Invitrogen), pre-coupled with 2.5 µL of anti-Myc were used for the affinity purification of Sen1. The beads were washed twice with 0.5% BSA (w/v) in cold PBS and centrifuged at 3 000 rpm for 3 minutes at 4 °C before being incubated with anti-Myc. After overnight incubation at 4 °C, the coupled-beads were washed twice with Lysis Buffer and added to 650 µL of soluble cell lysate. The tubes were placed on a rotating wheel and incubated overnight at 4 °C. After binding, the beads were washed six times in Lysis buffer by inversion. Successively, they were washed five times in 50 mM ammonium bicarbonate and resuspended in 50 µL of cold ammonium bicarbonate before proceeding with enzymatic on-bead digestion.

7.3 On-bead digestion and LC-MS/MS

The on-bead proteins were first diluted in 2M Urea/50mM ammonium bicarbonate and on-bead trypsin digestion was performed overnight at 37ºC. The samples were then reduced with 13 mM dithiothreitol at 37ºC and, after cooling for 10 minutes, alkylated with 23 mM iodoacetamide at room temperature for 20 minutes in the dark. The supernatants were acidified with trifluoroacetic acid and cleaned from residual detergents and reagents with MCX cartridges (Waters Oasis MCX 96-well Elution Plate) following the manufacturer's instructions. After elution in 10% ammonium hydroxide/90% methanol (v/v), samples were dried with a Speed-vac, re-solubilized under agitation for 15 min in 11 µL of 2%ACN/1% formic acid and loaded onto a 75 µm i.d. × 150 mm Self-Pack C18 column installed in the Easy II-nLC system (Proxeon Biosystems). The buffers used for chromatography were 0.2% formic acid (buffer A) and 90% acetonitrile/0.2% formic acid (buffer B). Peptides were eluted with a two slope gradient at a flowrate of 250 nL/min. Solvent B first increased from 1 to 35% in 90 min and then from 35 to 86% in 10 min. The HPLC system was coupled to Orbitrap Fusion mass spectrometer (Thermo Scientific) through a Nanospray Flex Ion Source. Nanospray and S-lens voltages were set to 1.3-1.7 kV and 50 V, respectively. Capillary temperature was set to 225 °C. Full scan MS survey spectra (m/z 360-1560) in profile mode were acquired in the Orbitrap with a resolution of 120,000 with a target value at 1e6. The 22 most intense peptide ions were fragmented in the HCD collision cell and analyzed in the linear ion trap with a target value at 1e4 and a normalized collision energy at 30 V. A MS3 scanning was performed upon detection of a neutral loss of phosphoric acid (48.99, 32.66 or 24.5 Th) in MS2 scans. Target ions selected for fragmentation were dynamically excluded for 40 sec after three MS2 events.

7.4 Data processing

Protein database searching was performed with Mascot 2.5 (Matrix Science) against the UniProt *S. Cerevisia*e S288c txid 559292. The peak list files were generated

with Proteome Discoverer (version 2.3) using the following parameters: minimum mass set to 500 Da, maximum mass set to 6000 Da, no grouping of MS/MS spectra, precursor charge set to auto, and minimum number of fragment ions set to 5. The mass tolerances for precursor and fragment ions were set to 10 ppm and 0.6 Da, respectively. Trypsin was used as the enzyme allowing for up to 1 missed cleavage. Cysteine carbamidomethylation was specified as a fixed modification, and methionine oxidation, serine, threonine and tyrosine phosphorylation modifications as variable modifications. Data analysis was performed using Scaffold (version 4.7.2). Spectra of phosphopeptides were manually inspected to confirm phosphorylation sites.

RESULTS

1. The Ubiquitin-Proteasome System degrades Sen1

To determine whether Sen1 levels are regulated by the Ubiquitin-Proteasome system (UPS), yeast cells were treated with the proteasome inhibitors PS341 (150 µM) and MG115 (150 µM), which resulted in Sen1 accumulation (Figure 12, lane 2). This confirmed that the degradation of Sen1 is regulated by the proteasome (DeMarini et al., 1995; Mischo et al., 2018). Furthermore, the destabilization of Sen1 in the absence of Glc7 was rescued when cells were co-treated with proteasome inhibitors (Figure 12, lane 4). Altogether, these results showed that Sen1 protein levels are controlled by Glc7 in a proteasome-dependent manner.

Since most UPS targets are recognized for degradation by polyubiquitination and phosphorylation marks within the protein (Holt et al., 2012), we sought to assess Sen1 post-translational modifications in yeast cells.

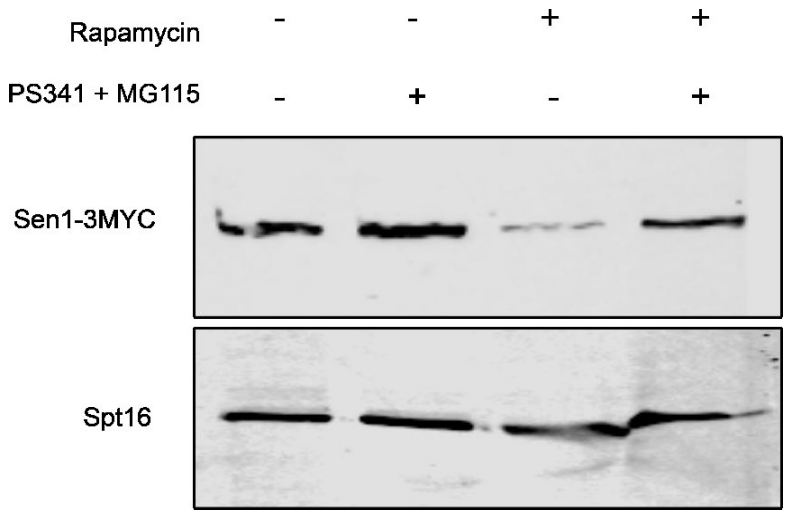


Figure 12. – Proteasomal inhibition using PS341 and MG115 treatment upon Glc7 depletion rescues Sen1 protein levels. Comparison of Sen1 IP after treatment with proteasome inhibitors (120 minutes), depletion of the endogenous Glc7 protein by anchor-away (90 min) or co-treatment with both

drugs. Sen1 protein levels are increased when cells were treated with proteasome inhibitors (lane 2), and are rescued when cells are treated with rapamycin in combination with proteasome inhibitors (lane 4). The figure shows representative blots of two independent experiments. The *PDR5* gene was deleted to improve proteasome inhibition efficiency. A fraction of the flowthrough was used as the loading control (Spt16). These conditions were also used as a small-scale control for the LC-MS/MS experiment described below.

2. Sen1 phosphorylation site identified by mass spectrometry

To map potential post-translational modifications in Sen1, we treated yFR3097 strain (*SEN1*-3MYC; *GLC7*-FRB) with proteasome inhibitors to enrich phosphorylation sites. We had a specific interest in identifying potential phosphorylation sites regulated by Glc7. For this, we also used a second sample of the yFR3097 strain that was co-treated with proteasome inhibitors and rapamycin to deplete Glc7 from the nucleus. We seek to distinguish among all phosphorylation sites those potentially regulated by Glc7. Sen1 was isolated by immunoprecipitation with anti-MYC antibody, and digestion was carried out on-beads using trypsin (Figure 13).

Mass spectrometry analysis identified 177 unique Sen1 peptides with 62% amino acid coverage in the first Sen1 IP (proteasome inhibitors), and 150 unique Sen1 peptides with 55% amino acid coverage for the second sample (proteasome inhibitors and rapamycin). In both Sen1 IP samples a single (at position S1505) phosphopeptide was identified (Table 5, Annex D). Consistent with phosphorylation of S1505 being regulated by Glc7, this phosphopeptide was more abundant when Glc7 was depleted from the nucleus.

Phosphorylated S1505 has been previously reported in phosphoproteomics studies in budding yeast (Swaney et al., 2013; Lanz et al., 2021). Interestingly, many of the phosphorylation sites already reported occur in the same cluster (Figure 14), slightly

overlapping the N-terminal part of the helicase domain, near the center of the protein. The S1505 phosphorylation site is located outside this cluster, positioning itself in the center of the helicase domain. Scaffold analysis (Scaffold_4.7.2 Proteome Software, Inc.) did not identify other previously reported phosphorylation sites in our experiment, in part, because tryptic digestion did not produce the fragments necessary to identify peptides in the N-terminal region. Using two successive digestions with LysC and trypsin also did not improve coverage (data not shown).

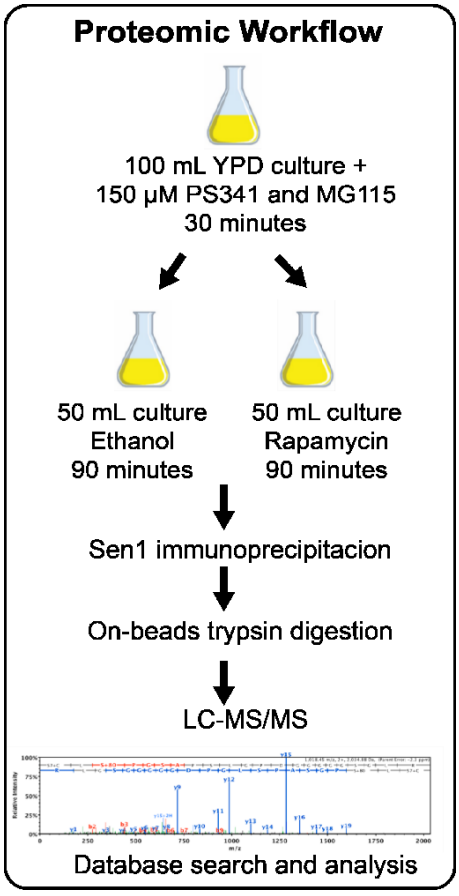


Figure 13. – IP coupled to Mass Spectrometry analysis identifies Sen1 phosphorylation site. Schematic workflow of the mass spectrometry (LC-MS/MS) experiment. Yeast cells were treated with the PS341 and MG115 proteasomal inhibitors before adding rapamycin. Sen1 immunoprecipitation was performed with anti-MYC antibody before on-beads tryptic digestion. Protein database searching was performed with Mascot 2.5 (Matrix Science) and Data analysis was performed using Scaffold (version 4.7.2).

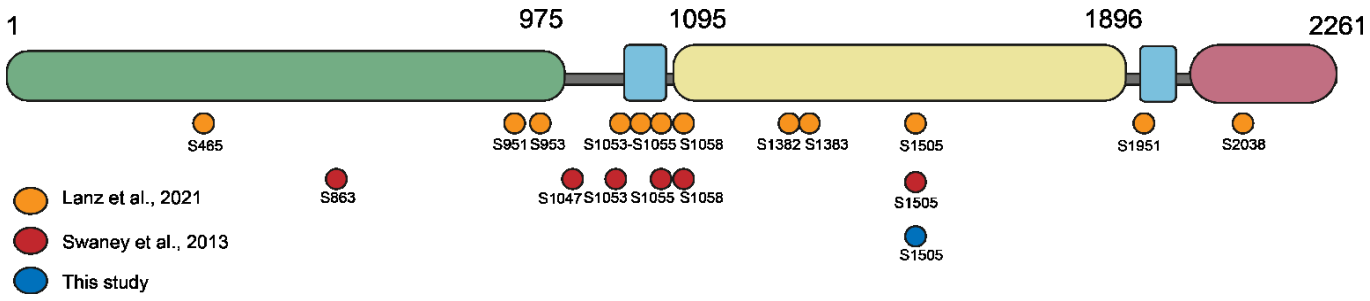


Figure 14. – Schematic representation of Sen1 phosphorylation sites identified in proteasome inhibited yeast cells. The graph shows the Sen1 phosphorylation sites reported so far, in their different domains and conserved regions. The green region indicates the N-terminal domain, the two blue regions the NLS, the yellow region shows the helicase domain and in pink the C-terminal. The circles denote the phosphorylation sites, orange, the phosphorylation sites recently reported by Lanz et al., 2021; in red the sites reported in Swaney et al., 2013 and in blue the sites that were confirmed in this study. The phosphorylated residue S1505 was observed on several peptide fragments. Many of the phosphorylation sites are concentrated in a cluster, slightly overlapping the N-terminal part of the helicase domain.

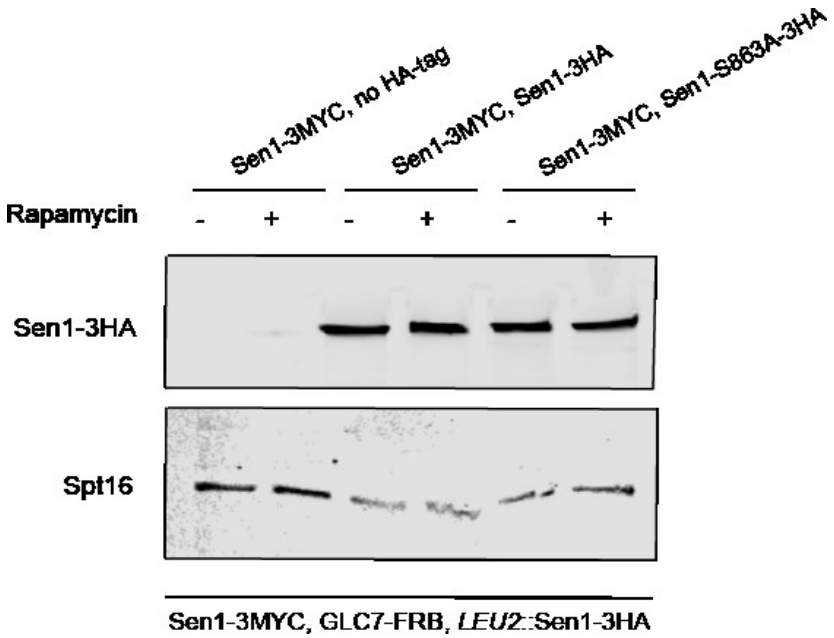
3. Functional analysis of Sen1 phosphorylation sites

Despite reported in several phosphoproteomics studies, these phosphorylation sites in Sen1 have not been functionally evaluated. In order to assess the stability of the Sen1 phospho-mutants, Sen1 IP followed by Western Blot analysis was performed using WT and phospho-mutant Sen1 constructs that were integrated into the *LEU2* locus of the yFR2348 strain. Sen1 phospho-mutants containing the HA epitope tag at the C-terminal were generated with the following mutations: S863A, S1505A, and double mutant S863A/S1505A. Phosphorylation at the S1505 site was identified in our mass spectrometry analysis, while the S863 site was reported as the first potential

phosphodegron in a study of ubiquitination and phosphorylation cross-talk (Swaney et al., 2013).

We first tested whether the S863A mutant affected the stability of Sen1 by Glc7 by treating with rapamycin to remove Glc7 from the nucleus. Two *SEN1* alleles were present in this yeast strain: 1) the phospho-mutant *sen1*-S863A (or WT as a control) tagged with HA was expressed from its own promoter and integrated at the *LEU2* locus and; 2) the endogenous WT *SEN1* tagged with MYC. We were expecting that if the *sen1*-S863A mutant turned out to be a real phospho-degron, it should have stabilized Sen1 protein levels in the absence of Glc7. However, depletion of Glc7 with rapamycin did not lead to the degradation of the *SEN1*-WT-HA allele (Figure 15A), compared to the *SEN1*-WT-MYC allele (see Figure 12, lane 3). To rule out any defects in the rapamycin treatment, we visualized the endogenous *SEN1*-WT-MYC allele from the same extracts, as an internal control. The *SEN1*-WT-MYC allele showed the expected effect (Figure 15B). Therefore, because our mutants were engineered in an HA-tagged *SEN1*, we were not able to test the stability of Sen1 in the presence of phospho-mutations.

A



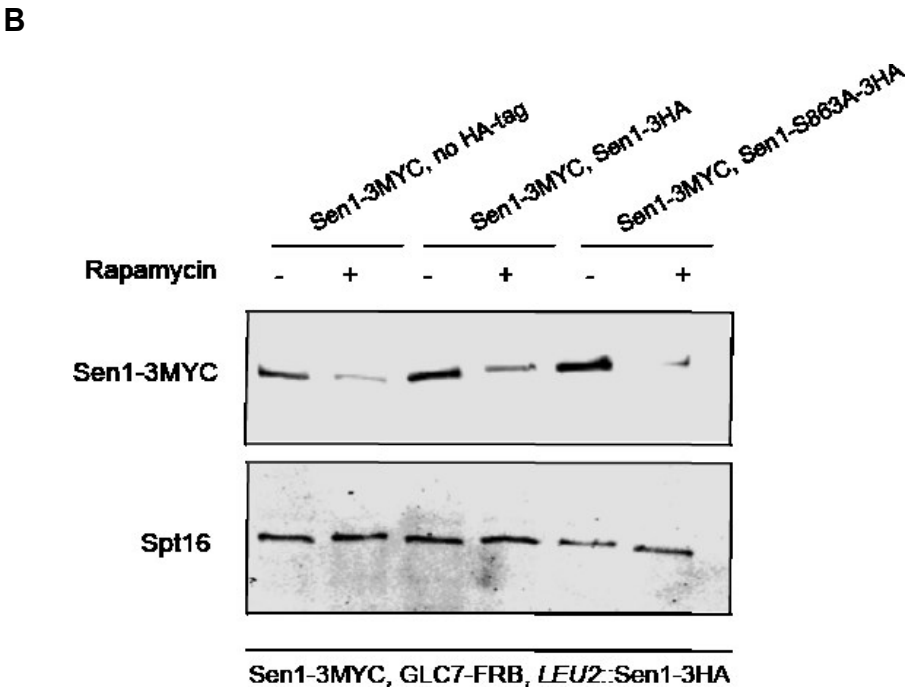


Figure 15. – Sen1-WT-3HA does not show changes in protein stability upon Glc7 nuclear depletion. The yFR2348 strain was treated with rapamycin for 90 minutes to deplete Glc7 from the nucleus. **A)** The *SEN1*-WT-HA allele was not destabilized after Glc7 nuclear depletion by rapamycin treatment. **B)** Using the same extracts from Figure 15A, we saw the destabilization of the *SEN1*-WT-MYC allele after removing Glc7 from the nucleus. The figure shows representative blots of two independent experiments. A fraction of the flowthrough was used as the loading control in both cases (Spt16).

Since the regulation of Sen1 was compromised by being fused with the HA tag, we decided to change this for a FLAG tag in all mutants. These mutants (as well as the WT allele) were also integrated into the *LEU2* locus under the control of its own promoter. Similar to what was done previously, we first tested the response of the *SEN1*-WT-FLAG allele to nuclear depletion of Glc7 by rapamycin. Removing Glc7 from the nucleus did not lead to degradation of the *SEN1*-WT-FLAG allele (Figure 16A). In the same way, we used the internal control from the same extracts to confirm the effect of rapamycin on the endogenous *SEN1*-WT-MYC allele, which resulted as expected (Figure 16B).

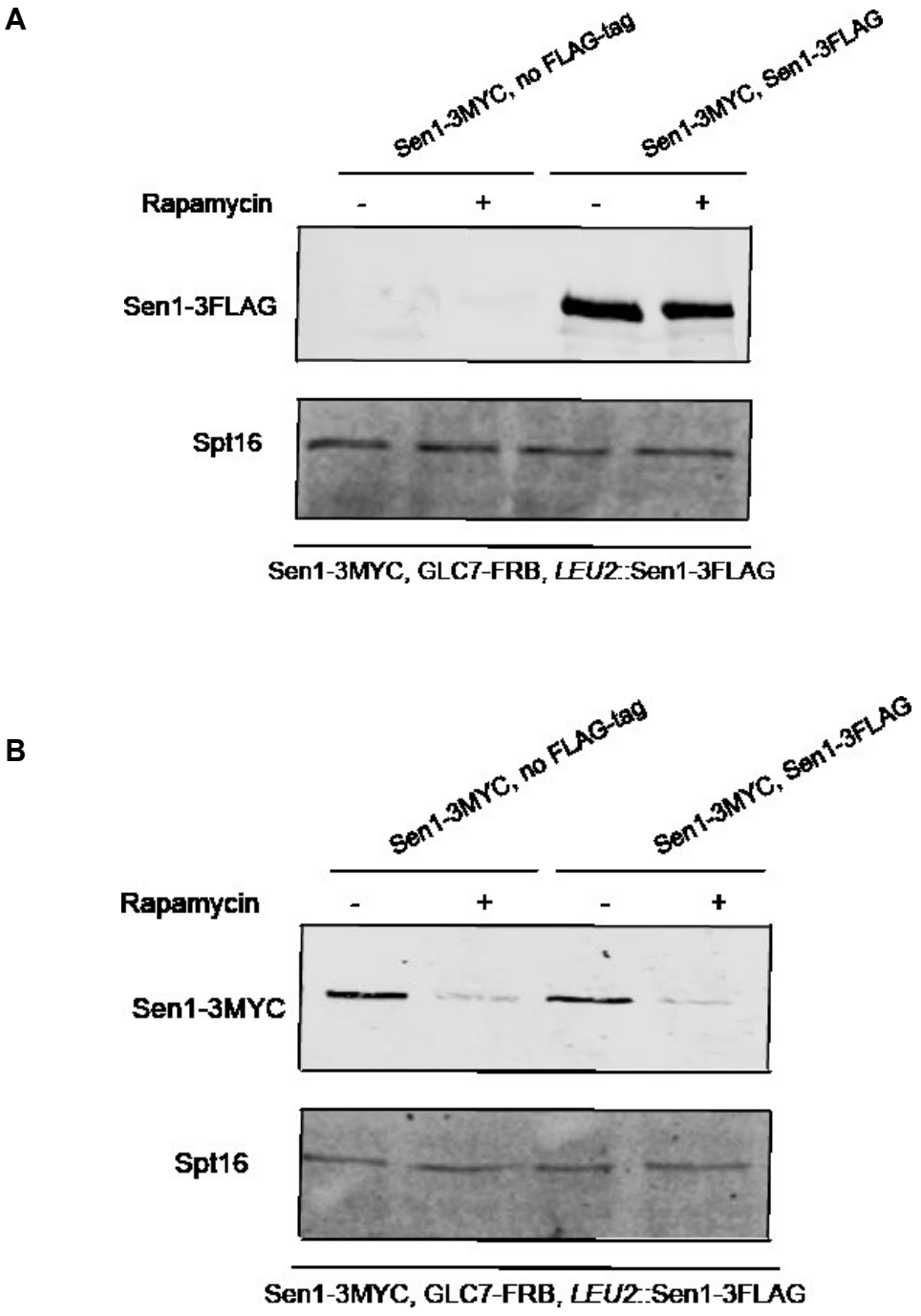


Figure 16. – Sen1-WT-3FLAG does not show changes in protein stability upon Glc7 nuclear depletion. The yFR2348 strain was treated with rapamycin for 90 minutes to deplete Glc7 from the nucleus. **A)** The *SEN1*-WT-FLAG allele was not destabilized after Glc7 nuclear depletion by rapamycin treatment. **B)** Using the same extracts from Figure 16A, we saw the destabilization of the *SEN1*-WT-MYC allele after removing Glc7 from the nucleus. The figure shows

representative blots of two independent experiments. A fraction of the flowthrough was used as the loading control in both cases (Spt16).

Why the addition of an HA or FLAG tag prevents the degradation of Sen1 upon nuclear depletion of Glc7 remains unknown and will require further investigations. However, we believe that the destabilization of the endogenous *SEN1*-WT-MYC allele is due to the removal of Glc7 and not to an artifact of the MYC tag (see Section 4, Discussion).

Given this tag problem, we thought that the generation of these phospho-site mutations could require manipulation at the endogenous *SEN1*-MYC. For this, we used CRISPR-Cas9 to integrate them directly into the endogenous *SEN1* locus (MYC tagged). However, none of these mutants generated colonies, suggesting that the mutations were lethal.

4. F2003A mutation in *SEN1* impairs interaction with Glc7 *in vivo*

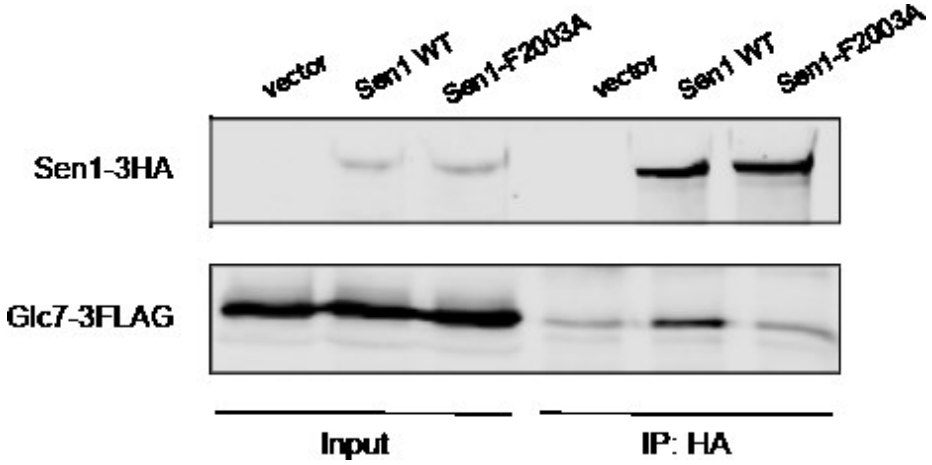
Potential physical interactions between Sen1 and Glc7 have been previously reported using yeast two-hybrid and Sen1-IP coupled to mass spectrometry analyses (Ursic et al., 2004; Ho et al., 2002; Nedea et al., 2008). As previously mentioned, most proteins interact with Glc7 via a RVXF motif of the target proteins. Importantly, the conserved phenylalanine residue has been identified as required for this interaction, since mutations in this amino acid greatly reduced binding to Glc7 (Chang et al., 2002, Wu and Tatchell, 2001).

A potential RVXF motif was identified in Sen1 in the RKHVCF sequence (residues 1999-2003) (Nedea et al., 2008), which has been tested in a two-hybrid assay. Specifically, the two-hybrid assays analyzed the interaction between sen1-WT and *sen1*-F2003A with Glc7 using only the C-terminal region of Sen1 (residues 1890 to 2092) tagged with GFP as prey. This showed that the WT version strongly interacted with Glc7, while the F2003A mutant did not. Although this approach has allowed the identification of

potential interactions, this was not carried out under physiological conditions since only a fragment of Sen1 was used.

To confirm this interaction, we carried out a co-IP of Sen1 to identify its protein complexes under physiological conditions. To examine *sen1*-F2003A, plasmids carrying the whole *SEN1*-WT gene or the *sen1*-F2003A mutant (fused with triple HA epitope) were transformed into an epitope-tagged Glc7-3FLAG strain. Whole-cell extracts were bound to anti-HA antibody, eluted from beads, and analyzed by immunoblotting against anti-HA and anti-FLAG antibodies. The results show that sen1-WT was co-purified along with the Glc7-3FLAG protein. However, the F2003A mutation prevented this ability of the proteins to interact (Figure 17A). The control condition, which did not express the HA construction, confirmed the specificity of the interaction between Sen1 and Glc7. Additionally, to confirm the reversibility of the reaction, we carried out a Glc7-FLAG IP in cells transformed with either *SEN1*-WT or *sen1*-F2003A plasmids. The resulting immunoblotting confirmed that the F2003A mutation abolished the interaction with Glc7 (Figure 17B).

A



B

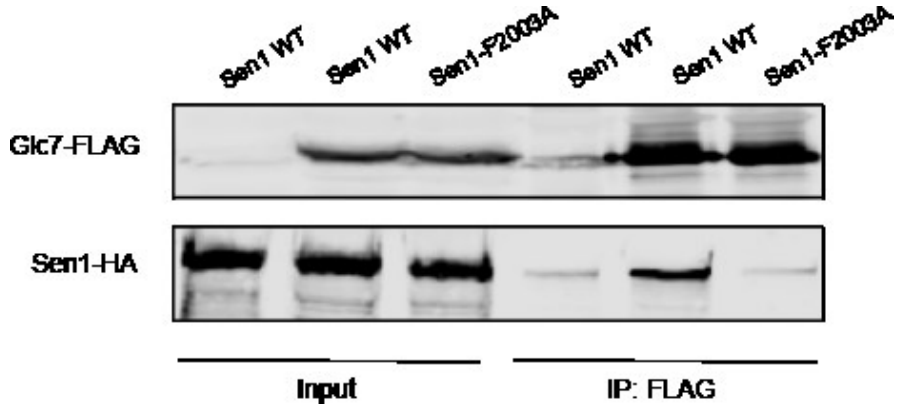
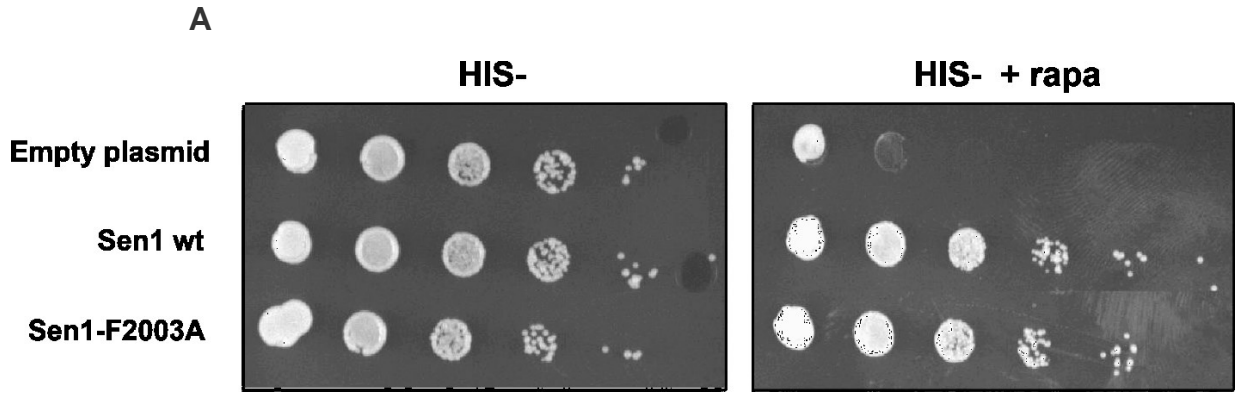


Figure 17. – F2003A substitution in Sen1 abolishes interaction with Glc7. A) Yeast strain carrying Glc7-3FLAG was transformed with either pFR2 (empty vector), *SEN1*-WT-HA, or *sen1*-F2003A-HA. The protein complexes were immunoprecipitated using anti-HA from whole-cell extracts. Western blotting with anti-FLAG showed that *sen1*-F2003A impairs the interaction between Sen1 and Glc7. **B)** Reverse co-IP experiment using anti-FLAG antibody to immunoprecipitate Glc7. Immunoblotting with anti-HA confirmed the impairment of the interaction. The figure shows representative blots of two independent experiments.

Given this result, we thought that abolishing the interaction between Sen1 and Glc7 could be involved in the regulation of processes that make it essential for growth. To compare the cell growth rate of yeast in wild-type and mutant cells, we also took advantage of the anchor-away approach. Since *SEN1* gene is essential for yeast viability, applying the anchor-away technique allows the protein to be transiently translocated into the cytoplasm. To this end, the endogenous Sen1 protein was C-terminally tagged with FRB. This strain also carried an FKBP12 tag to the ribosomal protein RPL13A that served to anchor Sen1-FRB to the cytoplasm in response to rapamycin. In this strain, we transformed plasmids carrying the *SEN1*-WT-HA and *sen1*-F2003A-HA and tested the effect of these plasmids by anchoring Sen1 and evaluating its growth using spot assay on control HIS- or HIS- plates containing rapamycin. After 3 days of growth, rapamycin treatment did not show any defect on cell growth in the strain carrying the *sen1*-F2003A mutant (Figure 18A). Additionally, we analyzed the Sen1-WT and Sen1-F2003A protein

levels by immunoprecipitation followed by western blotting, which also did not reveal significant differences at the expression level (Figure 18B). We are aware that these experiments were performed on strains in which Sen1 is HA tagged. If the stability of Sen1 is compromised by this tag, the lack of phenotype and the expression level of Sen1-F2003A observed here might not reflect what occurs under physiological conditions.

To resolve this ambiguity and evaluate the possible phenotype caused by the lack of interaction between Sen1 and Glc7, we consider that it would be necessary to replace the genomic *SEN1* with *sen1*-F2003A allele in a plasmid shuffle assay. Because *SEN1* is necessary to maintain viability, *sen1Δ* cells die when an empty vector is added but survive when a *SEN1*-WT (URA3-marked) plasmid is added. This can be selected against with 5-FOA to allow its replacement with the F2003A mutant allele. Since the background strain lacks *SEN1* or has a non-functional one, we could see the effect caused by the F2003A mutation on cell growth.



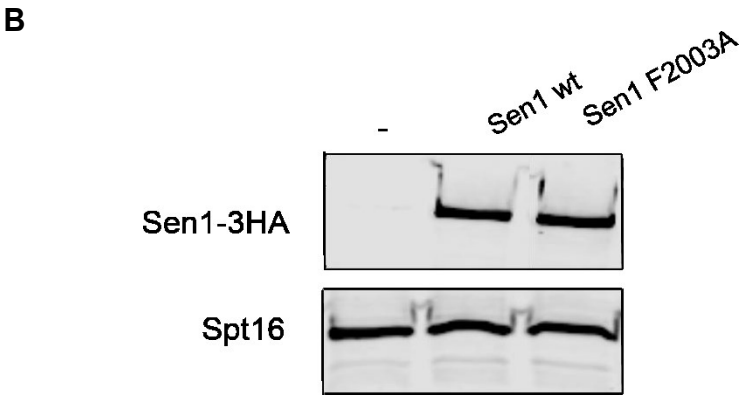


Figure 18. – *sen1*-F2003A mutation does not affect cell growth or protein stability. **A)** Growth assay of Sen1 anchor-away strain. *SEN1*-FRB parental strain was transformed with an empty plasmid, *SEN1*-WT or *sen1*-F2003A and grown at 30 °C overnight. 200 µl was taken at OD_{600} = 1 and 10-fold serially dilutions cultures were spotted on control HIS- or HIS- plates containing rapamycin (1µg/ml), which activates the anchor-away system. Photographs were taken after 72 hours of incubation at 30 °C. **B)** Wild-type and *sen1*-F2003A strains were grown to log-phase in HIS-. Protein levels in wild-type and mutant were analyzed by western blot using anti-HA antibody. A fraction of the flowthrough was used as the loading control (Spt16). The figure show a blot of one independent experiment.

Using a bioinformatic predictor we saw that Sen1 consists of two ordered regions flanked by intrinsic disorder regions. The IUPred analysis (Dosztányi et al., 2005) predicted two disordered regions in the C-terminal region and between the N-terminal and helicase domain, while the N-terminal regions and the helicase domain are more ordered (Figure 19). Interestingly, the C-terminal region encompasses not only the RVxF motif but also an interaction motif with Nrd1 (Han et al., 2020). The cluster of phosphorylation sites is placed in a region that is also predicted to be intrinsically disordered (980-1050). Both regions are therefore accessible for interaction with Glc7 or other proteins.

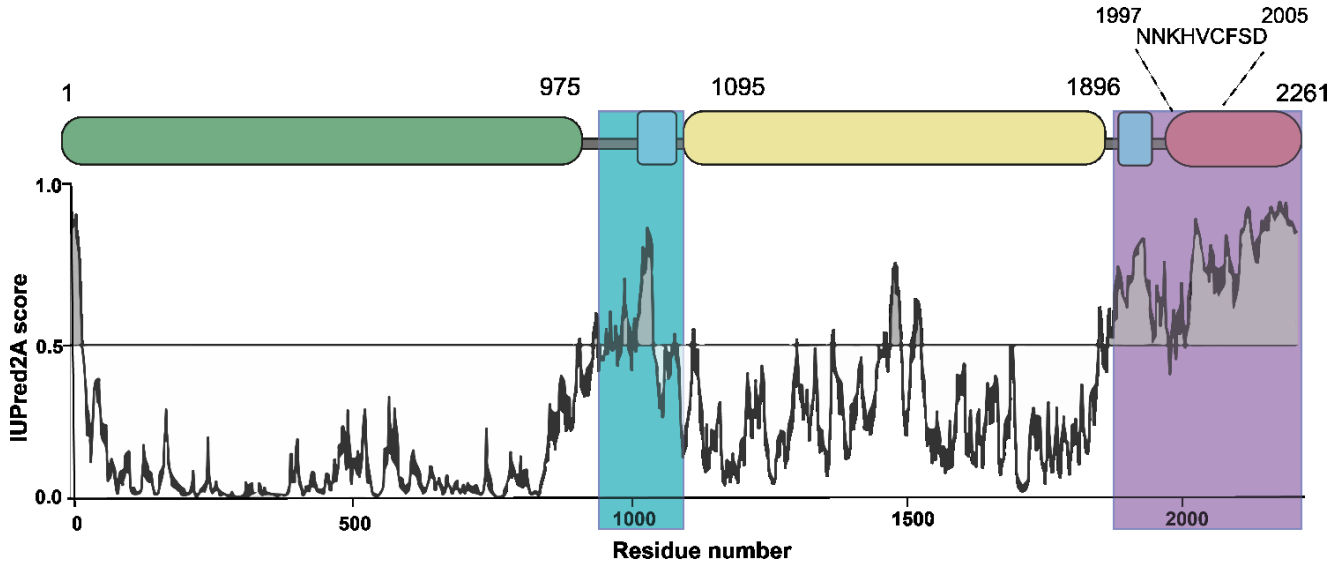


Figure 19. – Scheme of Sen1 predicted intrinsically disordered regions. Intrinsically disordered regions are highlighted in blue and pink. The disorder prediction was obtained using IUPred (Dosztányi et al, 2005). The sequence of the Sen1 RVxF motif is shown on the top. The green region indicates the N-terminal domain, the two blue regions the NLS, the yellow region shows the helicase domain and in pink the C-terminal.

DISCUSSION

1. The role of Glc7 in the stabilization of Sen1

In the framework of my project, we wanted to further explore the mechanism involving the effect of Glc7 on Sen1. For this, we performed all the experiments under physiological conditions, evaluating the stability of Sen1 when Glc7 was depleted from the nucleus. Since Sen1 is a low abundance protein (500 molecules/cell), we carried out an immunoprecipitation to enrich Sen1 levels before its detection by western blotting. The first results allowed us to corroborate the stability of Sen1 upon a cycloheximide time-course (Collin, 2019), however, this stability was highly compromised when removing Glc7 from the nucleus (Figure 12, lane 3). These results, together with the fact that Glc7 can dephosphorylate Sen1 *in vitro* (Nedea et al., 2008), allowed us to propose that Glc7 would be responsible for the stabilization of Sen1 in the nucleus. Another additional result that further supports this model is the fact that the stability of Sen1 is directly associated with the catalytic activity of Glc7 (Collin, 2019). Moreover, it has been proposed that Glc7 can be recruited to the NNS complex and interact with other components of it (Nedea et al., 2008). However, it remains unclear whether the effect of Glc7 on Sen1 is necessary for its termination function or is independent of the NNS pathway.

2. Sen1 protein levels are regulated via the UPS

The degradation of Sen1 via the ubiquitin-proteasome was first suggested by DeMarini and colleagues (1995), who proposed the existence of one or more specific degradation signals within Sen1. Subsequently, it was shown that Sen1 protein levels fluctuate throughout the cell cycle, with Sen1 being more unstable during the G1 phase. This Sen1 destabilization was prevented by treatment with proteasome inhibitors (Mischo et al., 2018). Although this study was conducted under conditions in which Sen1 was overexpressed using inducible plasmids, our results have confirmed that this also occurs

under physiological conditions. As shown in Figure 12, the *in vivo* expression of Sen1 is regulated by the ubiquitin-proteasome system. Thus, suggesting a mechanism where the accumulation of Sen1 levels is jointly regulated by Glc7 and the proteasome.

3. Sen1 contains potential phosphodegron sites

Sen1 participates in different processes in the cell, however, its protein levels are less abundant compared to other proteins of the NNS pathway. Additionally, it has been reported that its overexpression results in a reduction in cell viability, due to an excess in termination (Misho et al., 2018). This limitation in Sen1 levels is carried out by the proteasome. Likewise, we know that post-translational modifications such as phosphorylation and ubiquitylation govern degradation via the ubiquitin-proteasome system in eukaryotic cells (Finley et al., 2012). Polyubiquitin accumulation of endogenous Sen1 has been demonstrated (Mischo et al., 2018), however, its dephosphorylation by Glc7 has only been suggested by *in vitro* experiments (Nedea et al., 2008). Therefore, we wanted to explore the post-transcriptional regulation of Sen1, specifically its phosphorylation, to understand the mechanism behind its degradation via the proteasome.

We wanted to explore the mechanism by which Sen1 levels are maintained in the cell. We hypothesized that Sen1 is stabilized by dephosphorylation by Glc7 (or another phosphatase) and once phosphorylation increases, it is degraded through the proteasome. Both pathways have been discussed above and our results have confirmed what has been shown in the literature.

Sen1 phosphorylation has also been explored in some studies, mainly in experiments based on phosphoproteomics in budding yeast, which is certainly a challenge given that this helicase is low abundant and is the second-largest protein in this organism. Some years ago the first potential phosphodegron was identified in serine 863 (S863) of Sen1, from a study that identified phosphorylation and ubiquitination sites upon proteasome inhibition (Swaney et al., 2013). Subsequently, other phosphorylation sites

have been reported at the N-terminal part of the helicase domain as well as within it (Figure 14), thus supporting the idea that Sen1 contains one or more potential phosphodegrons.

Despite participating in multiple processes, its low abundance seems to be essential for the cell, since its overexpression causes cell death (Mischo et al 2018). Given that many phosphodegrons are found in short-lived cell-cycle proteins such as Sen1 and based on our results and that observed by others, we believe that it is very likely that the stability of Sen1 is controlled via dephosphorylation/phosphorylation of a phosphodegron by Glc7 (and a yet-to-be-identified kinase).

In order to test our hypothesis, we mapped the abundance of phosphorylation sites in Sen1 when Glc7 was absent. To cover the largest number of sites, we analyzed protein complexes purified from a Sen1 immunoprecipitation, coupled to mass spectrometry. The yeast cell cultures were treated with proteasome inhibitors and with rapamycin (or not) to compare the effect between the presence and absence of Glc7 in Sen1. To reduce dephosphorylation during purification, cells were lysed with buffers containing phosphatase inhibitors.

We have identified one Sen1 phosphorylation site by mass spectrophotometry in budding yeast. Previously, at least three phosphoproteomic screens had shown evidence of Sen1 post-translational modifications, however, the biological or functional relevance of these sites had not been evaluated (Swaney et al., 2013; MacGilvray et al., 2020; Lanz et al. al., 2021). Most of these sites are in the N-terminal region (residues 1 to 1000 approximately). In our experiment it was not possible to identify any of these, in part because tryptic digestion did not produce the fragments necessary to identify peptides in this region. Using two successive digestions with LysC and trypsin did not improve the protein coverage. In addition to this, the homemade anti-MYC antibody did not allow adequate recovery of Sen1 (compare Figures 16A and 16B for difference between anti-MYC and anti-FLAG recovery).

Phosphorylation sites S951 to S1058 are located very close to each other, forming a small cluster. The S1505 site resides outside of this cluster, in the center of the helicase domain. Other sites such as S465 and S863 are in the N-terminal region (Figure 14). S863

has been the only site reported as a possible phosphodegron. Interestingly, site S465 is located within the region proposed as the cause of Sen1 stabilization during G1 (*sen1*-459-498Δ) (Mischo et al., 2018). Mutations within amino acids 480-493 failed to stabilize the protein, and they could not identify the specific residues involved in the degradation of Sen1. Although it was not identified in our study, we believe that the S465 site could act as a possible phosphodegron during the G1 phase.

The small cluster of phosphorylated residues overlaps N-terminal part of the helicase domain, in one of the nuclear localization signals (NLS) of Sen1, which functions as the signal responsible for the nuclear localization of the protein. Specifically, the N-terminal NLS has been proposed as the signal that allows Sen1 to be located in the nucleolus, thus making it possible for Sen1 to interact with snoRNP proteins (Chen et al., 2014). In addition, it is known that phosphorylation of NLS in budding yeast proteins can affect their recruitment site (Harreman et al., 2004). Therefore, we think that phosphorylation in this region might occur transiently when Sen1's specific action is needed in the nucleus. This effect has also been described for the human homolog of Sen1, Senataxin, which changes its cellular location according to its site of action (Yuce and West, 2013).

Sites S1382, S1383, and S1505 are located in the helicase domain. S1505 is the only one that has been reported in three studies, including this one. In the case of budding yeast, it is known that modifications in helicase domains are usually associated with a change in the catalytic activity of the enzyme or they can also marked the protein for degradation via the ubiquitin-proteasome system (Sommers et al., 2015). Additionally, it has been reported that these modifications can affect the subcellular localization of the enzyme, especially in helicases that are recruited to nuclear foci generated by replicative stress, as is the case of Senataxin.

Bioinformatic predictors indicate that Sen1 consists of defined protein interaction modules (N-terminal and helicase domains) linked by one region of intrinsic disorder, as well as a large disordered region at the C-terminal end (Figure 19). Analysis with IUPRED (Dosztányi et al, 2005) predicted the first disordered region between residues 950 to 1060 and the second within the C-terminal domain of Sen1. This is consistent with the fact that

most phosphorylation sites are found within unstructured or intrinsically disordered regions (Iakoucheva et al., 2004).

A Clustal analysis was carried out in order to identify if the residues, in which the phosphorylation mark is found, are conserved in the human homologous Senataxin. This approach was restricted to the helicase domain since it is the only region where it is possible to identify homologous residues due to their high amino acid similarity. The analysis revealed that none of the three residues of the helicase domain are conserved in Senataxin. This could be because these sites reside within or very close to a disordered region, which coincides with the fact that phosphorylated proteins in these regions may not show conservation of these sites. However, this changes when phosphorylation occurs in structured domains (Niu et al., 2012), which might apply to residues located at the N-terminus.

In the same way, we wanted to identify the location of these amino acids in the crystal structure of Sen1 to detect if they were near or on the interaction surface of Sen1. However, only part of the structure of the helicase domain of Sen1 is available, thus it was not possible to identify the residues. Unfortunately, there is still no crystal structure for Senataxin or its helicase domain.

4. Functional analysis of potential phospho-degron sites

We selected the phosphorylated residues S863 and S1505 for further evaluation of their stability and possible action as phosphodegron. S1505 was the only phosphorylation site identified in our mass spectrometry study and S863 is the only site that has been specifically identified as a potential phosphodegron. However, because of the increasing number of phosphorylation sites reported in more recent studies, we know that other combinations of potential phosphodegron sites should be tested.

To further validate these phosphorylation sites, we generated *SEN1* phospho-mutants plasmids carrying Serine to Alanine mutations that were integrated into the

genome at the *LEU2* locus. With this, we wanted to assess the impact of the mutations on protein degradation by preventing the occurrence of phosphorylation.

SEN1 plasmids containing the S863A, S1505A as well as double site mutations tagged with the HA epitope were expressed endogenously to evaluate their effect on the degradation of Sen1 compared to the WT version. Unexpectedly, the stability of Sen1-WT-HA and Sen1-WT-FLAG was not affected by the removal of Glc7 from the nucleus (Figure 15 and Figure 16). Suggesting that both tags affected the stability of the protein and that they might represent inappropriate substrates of the proteasome, unlike the endogenous Sen1-MYC allele.

We checked whether the MYC tag contained any region that made it susceptible to degradation by the proteasome. The MYC protein does indeed have phosphorylation sites associated with proteasome degradation, located in the N-terminal region (LPTPPLSP, residues 56 to 64) (Chakraborty et al., 2015). However, the MYC epitope that we use as a tag corresponds to the C-terminal region of the protein (EQKLISEEDL). Therefore, we rule out that the MYC tag has produced a degradation artifact. Additionally, this is supported by the fact that Glc7 catalytic mutants affected Sen1 stability at different levels (Collin, 2019). However, it remains unknown how the different tags (either HA or FLAG) could affect protein stability.

Although our study and that of others (Mischo et al., 2018; Swaney et al., 2013; Lanz et al., 2021) have demonstrated the main role played by the ubiquitin-proteasome system in the regulation of Sen1 levels, whether or not this involves the dephosphorylation by Glc7 remains elusive. Moreover, the biological context of the action of Glc7 on Sen1 must be established, either at the level of the cell cycle, NNS pathway, or other NNS-independent processes.

5. *sen1*-F2003A mutation impairs the interaction with Glc7 *in vivo*

Our results so far suggest that the stabilization of Sen1 by Glc7 prevents its proteasome-mediated degradation. Preliminary evidence predicted an RVxF motif in Sen1 that could mediate its interaction with Glc7 (Nedea et al., 2008). We hypothesized that preventing Glc7 from interacting with Sen1 should lead to the destabilization of Sen1. In order to test this, we generated a plasmid carrying an F2003A mutation in Sen1 (phenylalanine to alanine at position 2003), which has been proposed by others to prevent its association with Glc7 (Nedea et al., 2008). The interaction of wild-type Sen1 and Glc7 was confirmed by co-immunoprecipitation and the effect on the stability of the mutant was then carried out by western blot. The results showed that Sen1 and Glc7 interact with each other and that this interaction is abolished by mutating the 'RVxF' motif of Sen1 (Figure 17). Surprisingly, preventing the interaction of Sen1 with Glc7 does not affect the stability of Sen1 or cell growth (Figure 18). However, we do not rule out the possibility that this is could be due to a tag problem, which would generate greater stability in the protein, as discussed in the previous section. Experiments involving the use of mutants via a plasmid-shuffle assay would serve to resolve this ambiguity.

In the case that we confirm that Sen1-F2003A is a stable protein and does not affect cell growth, this would suggest that the interaction between Sen1 and Glc7 is independent of its dephosphorylating activity. This coincides with the fact that the valine and phenylalanine residues of the RVxF motif fit into the hydrophobic channel located on the surface of Glc7, while the catalytic site of the enzyme is on the opposite side (see Figure 8) (Offley and Schmidt, 2019; Cannon, 2010). In this scenario, Glc7 could still play a role in the dephosphorylation of Sen1. However, since Sen1 is a stable protein in non-synchronized cells, we consider that it would be important to confirm the destabilization of Sen1-F2003A under these same conditions by using a cycloheximide assay.

Finally, Sen1 could also act as a regulatory protein in some biological conditions, that is, its interaction with Glc7 would serve to dephosphorylate other substrates. This is

supported by the fact that the substrate activity, location, and specificity of many PP1 phosphatases are regulated by the binding of one or more different subunits (Abd-Rabbo and Michnick, 2017). In the case of the effect of Glc7 on the stabilization of Sen1, we know that it is not associated with multiprotein complexes such as APT or CPF (Collin, 2019). However, among the factors associated with Sen1 are two Glc7 regulatory proteins: Glc8 and Sds22 (Nedea et al., 2008). Interestingly, these regulators do not contain the RVxF sequence and therefore they would not compete with Sen1 to bind Glc7. This could indicate a possible link between the phosphatase activity of Glc7 and other regulatory proteins on NNS components.

CONCLUSION

Sen1 has been involved in protecting the yeast genome by resolving RNA/DNA hybrids naturally formed during transcription and, as part of the Nrd1–Nab3–Sen1 complex, to mediate the transcription termination of short non-coding RNAs. Regarding its abundance regulation, we have confirmed that Sen1 is degraded under physiological conditions via a proteasome-dependent pathway. Moreover, this mechanism relies on the catalytic activity of Glc7, a protein phosphatase that has been shown to dephosphorylate Sen1 *in vitro*, and to interact with Sen1 through an 'RVxF' motif in two-hybrid experiments. Therefore, we have provided an explanation of the *in vivo* role of this phosphatase on Sen1.

This research project provides the first specific analysis of Sen1 phosphorylation peptides in yeast cells. The results obtained have allowed us to highlight the residues that could act in the regulation of Sen1 protein levels and could strengthen our understanding of the modulation of Sen1 in its different cellular processes.

It would also be important for future studies to be able to determine the biological relevance of these phosphorylation sites. Similarly, a larger question that remains unanswered is the identification of the kinase that phosphorylates Sen1. The mass spectrometry analysis of this study has generated necessary information that could allow the identification of kinases potentially associated with Sen1, as well as other proteins that could open the possibility of new regulatory mechanisms. Validation of the specificity of these interactions with Sen1 should be performed as well.

Our findings also confirm that mutating the RVxF motif of Sen1 impairs the interaction with Glc7 *in vivo*. However, additional experiments are necessary to determine the existence of a phenotype associated with this lack of interaction.

Finally, to highlight the significance of these results, we believe that deciphering the regulatory mechanism of Sen1 could help us to further dissect its impact on the different phases of the cell cycle, as well as on other processes. Therefore, further studies will be necessary to fill these gaps.

FUTURE DIRECTIONS

Although it was possible to identify a potential phosphodegron in Sen1 (S1505), the biological relevance of this or the other previously reported in position S863 was not be confirmed. This was due to the effect of the tags on the stability of the protein and it remains unclear why the addition of an HA or FLAG tag prevents Sen1 degradation. Generating the mutation directly into the genome using CRISPR-Cas9 also failed, suggesting their lethality. Due to this and given the importance of testing these potential phosphodegrons, it is necessary to first determine the viability of these mutants. We propose to use a *sen1Δ* background strain transformed with the query constructs (vector, sen1-WT, sen1-S863A, sen1-S1505A, or sen1-S863A/S1505A) in a plasmid shuffle assay. Since *SEN1* is necessary to maintain viability, *sen1Δ* cells can be kept alive if they carry a SEN1-WT on a plasmid. If this plasmid has the URA3 marker, it can be selected against with 5FOA in the presence of other *sen1* alleles to test for their function. Once the viability of these mutations is determined, the stability of the MYC-tagged mutants in the absence of Glc7 could be tested by generating the mutation directly into the genome. Additionally, this same approach can be used to evaluate the possible phenotype caused by the lack of interaction between Sen1 and Glc7, seeing the effect caused by the F2003A mutation on cell growth.

The existence of a phosphodegron in Sen1 would allow testing whether Glc7 nuclear depletion causes termination defects due to its role in the dephosphorylation of Sen1. As mentioned, defects in the termination of ncRNAs were observed upon anchor-away of Glc7 (Collin et al., 2019). One way to test this hypothesis would be by using a Sen1 phosphodegron mutant. If the dephosphorylation of Sen1 by Glc7 is necessary for transcription termination of ncRNAs, the phosphodegron mutant should rescue the defect in the termination of these genes upon Glc7 anchor-away. That would allow proposing a role of Glc7 with at least three of its different complexes. As part of the CPF, in the dephosphorylation of CTD Tyr1 (Schreieck et al., 2014); as part of the APT, in transcription of snoRNAs (Lidschreiber et al., 2018) and as part of the NNS pathway (Nedea et al., 2008) it would regulate the stability of Sen1.

REFERENCES

Ahuatzi, D., Riera, A., Pelaez, R., Herrero, P., and Moreno, F. (2007). Hxk2 regulates the phosphorylation state of Mig1 and therefore its nucleocytoplasmic distribution. *J. Biol. Chem. 282,* 4485–4493.

Akiyoshi, B., Nelson, C. R., Ranish, J. A., and Biggins, S. (2009). Quantitative proteomic analysis of purified yeast kinetochores identifies a PP1 regulatory subunit. *Genes Dev. 23,* 2887–2899.

Albuquerque, C. P., Smolka, M. B., Payne, S. H., Bafna, V., Eng, J., and Zhou, H. (2008). A multidimensional chromatography technology for in-depth phosphoproteome analysis. *Molecular and cellular proteomics: MCP, 7(7),* 1389–1396.

Allmang, C., Kufel, J., Chanfreau, G., Mitchell, P., Petfalski, E., and Tollervey, D. (1999). Functions of the exosome in rRNA, snoRNA and snRNA synthesis. *The EMBO journal*, *18*(19), 5399–5410.

Ariño, J., Velázquez, D., and Casamayor, A. (2019). Ser/Thr protein phosphatases in fungi: structure, regulation and function. *Microbial Cell. 6.* 217-256.

Bailis, J.M., and Roeder, G. S. (2000). Pachytene exit controlled by reversal of Mek1-dependent phosphorylation. *Cell* 101, 211–221.

Baker, S.H., Frederick, D.L., Bloecher, A., and Tatchell, K. (1997). Alanine-scanning mutagenesis of protein phosphatase type 1 in the yeast Saccharomyces cerevisiae. *Genetics* 145, 615–626.

Barford, D. (2011). Structural insights into anaphase-promoting complex function and mechanism. *Philosophical Transactions: Biological Sciences, 366*(1584), 3605-3624.

Bazzi, M., Mantiero, D., Trovesi, C., Lucchini, G., and Longhese, M. P. (2010). Dephosphorylation of gamma H2A by Glc7/protein phosphatase 1 promotes recovery from inhibition of DNA replication. *Mol. Cell. Biol.* 30, 131–145.

Bennett, C. L., Chen, Y., Vignali, M., Lo, R. S., Mason, A. G., Unal, A., Huq Saifee, N. P., Fields, S., and La Spada, A. R. (2013). Protein interaction analysis of senataxin and the ALS4 L389S mutant yields insights into senataxin post-translational modification and uncovers mutant-specific binding with a brain cytoplasmic RNA-encoded peptide. *PloS one*, 8(11), e78837.

Bharucha, J. P., Larson, J. R., Konopka, J. B., and Tatchell, K. (2008). *Saccharomyces cerevisiae* Afr1 protein is a protein phosphatase 1/Glc7-targeting subunit that regulates the septin cytoskeleton during mating*. Eukaryot. Cell* 7, 1246–1255

Bodenmiller B, Campbell D, Gerrits B et al. (2008) PhosphoPep–a database of protein phosphorylation sites in model organisms. *Nat Biotechnol* 26: 1339–1340.

Braten, O., Livneh, I., Ziv, T., Admon, A., Kehat, I., Caspi, L. H., Gonen, H., Bercovich, B., Godzik, A., Jahandideh, S., Jaroszewski, L., Sommer, T., Kwon, Y. T., Guharoy, M., Tompa, P., and Ciechanover, A. (2016). Numerous proteins with unique characteristics are degraded by the 26S proteasome following monoubiquitination. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(32), E4639–E4647.

Bridges, A. A., and Gladfelter, A. S. (2015). Septin Form and Function at the Cell Cortex. *The Journal of biological chemistry*, *290*(28), 17173–17180.

Cannon, J.F. (2010). Function of Protein Phosphatase-1, Glc7, in Saccharomyces cerevisiae. *Adv. Appl. Microbiol., 1st ed. Elsevier B.V.*; pp 27–59.

Carlson, M. (1999). Glucose repression in yeast. *Curr. Opin. Microbiol*. 2, 202–207.

Carroll, K. L., Ghirlando, R., Ames, J. M., and Corden, J. L. (2007). Interaction of yeast RNA-binding proteins Nrd1 and Nab3 with RNA polymerase II terminator elements. *RNA (New York, N.Y.)*, *13*(3), 361–373.

Ceulemans, H., Vulsteke, V., De Maeyer, M., Tatchell, K., Stalmans, W., and Bollen, M. (2002). Binding of the concave surface of the Sds22 superhelix to the alpha 4/alpha 5/alpha 6-triangle of protein phosphatase-1. *The Journal of biological chemistry, 277(49),* 47331–47337.

Chakrabarti, S., Jayachandran, U., Bonneau, F., Fiorini, F., Basquin, C., Domcke, S., Le Hir, H., and Conti, E. (2011). Molecular mechanisms for the RNA-dependent ATPase activity of Upf1 and its regulation by Upf2. *Molecular cell,* 41(6), 693–703.

Chakraborty, A. A., Scuoppo, C., Dey, S., Thomas, L. R., Lorey, S. L., Lowe, S. W., and Tansey, W. P. (2015). A common functional consequence of tumor-derived mutations within c-MYC. *Oncogene*, *34*(18), 2406–2409.

Chang, J. S., Henry, K., Wolf, B. L., Geli, M., and Lemmon, S. K. (2002). Protein phosphatase-1 binding to Scd5p is important for regulation of actin organization and endocytosis in yeast. *J. Biol. Chem*. 277, 48002–48008.

Chen, S. H., Albuquerque, C. P., Liang, J., Suhandynata, R. T., and Zhou, H. (2010). A proteome-wide analysis of kinase-substrate network in the DNA damage response. *The Journal of biological chemistry*, *285*(17), 12803–12812.

Chen, Y. Z., Hashemi, S. H., Anderson, S. K., Huang, Y., Moreira, M. C., Lynch, D. R., Glass, I. A., Chance, P. F., and Bennett, C. L. (2006). Senataxin, the yeast Sen1p orthologue: characterization of a unique protein in which recessive mutations cause ataxia and dominant mutations cause motor neuron disease. *Neurobiology of disease*, *23*(1), 97–108.

Chen, X., Poorey, K., Carver, M.N., Müller, U., Bekiranov, S. (2017). Transcriptomes of six mutants in the Sen1 pathway reveal combinatorial control of transcription termination across the *Saccharomyces cerevisiae* genome. *PLOS Genetics 13(6):* e1006863.

Cheng, Z., Muhlrad, D., Lim, M. K., Parker, R., and Song, H. (2007). Structural and functional insights into the human Upf1 helicase core. *The EMBO journal*, *26*(1), 253–264.

Chong, Y.T., Koh, J.L., Friesen, H., Duffy, S.K., Cox, M.J., Moses, A., Moffat, J., Boone, C., and Andrews, B.J. (2015). Yeast Proteome Dynamics from Single Cell Imaging and Automated Analysis. *Cell 161,* 1413–1424.

Christiano, R., Nagaraj, N., Fröhlich, F., and Walther, T. C. (2014). Global proteome turnover analyses of the Yeasts *S. cerevisiae and S. pombe*. *Cell reports, 9(5),* 1959–1965.

Ciechanover, A. (2015). The unravelling of the ubiquitin system. *Nat. Rev. Mol. Cell Biol*. *16*, 322–324.

Cohen, P. (1989). The Structure and Regulation of Protein Phosphatases. *Annu Rev Biochem 58(1):* 453–508.

Colin, J., Libri, D., and Porrua, O. (2011). Cryptic transcription and early termination in the control of gene expression. *Genetics research international*, *2011*, 653494.

Collin, P., Jeronimo, C., Poitras, C., and Robert, F. (2019). RNA Polymerase II CTD Tyrosine 1 Is Required for Efficient Termination by the Nrd1-Nab3-Sen1 Pathway. *Molecular cell*, 73(4), 655–669.e7.

Collin, P. (2019). Caractérisation du domaine C-terminal de l'ARN polymérase II et de la phosphatase Glc7 dans la terminaison transcriptionnelle chez Saccharomyces cerevisiae [Doctoral dissertation, Université de Montréal].

Connor, J. H., Frederick, D., Huang, H. b., Yang, J., Helps, N. R., Cohen, P. T., Nairn, A. C., DePaoli-Roach, A., Tatchell, K., and Shenolikar, S. (2000). Cellular mechanisms regulating protein phosphatase-1. A key functional interaction between inhibitor-2 and the type 1 protein phosphatase catalytic subunit. *The Journal of biological chemistry, 275(25),* 18670–18675.

Conrad, N. K., Wilson, S. M., Steinmetz, E. J., Patturajan, M., Brow, D. A., Swanson, M. S., and Corden, J. L. (2000). A yeast heterogeneous nuclear ribonucleoprotein complex associated with RNA polymerase II. *Genetics*, *154(2*), 557–571.

Contreras, X., Benkirane, M., Kiernan, R. (2013). Premature termination of transcription by RNAP II: the beginning of the end. *Transcription 4:* 72–76.

Creamer, T. J., Darby, M. M., Jamonnak, N., Schaughency, P., Hao, H., Wheelan, S. J., and Corden, J. L. (2011). Transcriptome-wide binding sites for components of the Saccharomyces cerevisiae non-poly(A) termination pathway: Nrd1, Nab3, and Sen1. *PLoS genetics*, *7*(10), e1002329.

de Almeida, S. F., García-Sacristán, A., Custódio, N., and Carmo-Fonseca, M. (2010). A link between nuclear RNA surveillance, the human exosome and RNA polymerase II transcriptional termination. *Nucleic acids research*, *38*(22), 8015–8026.

DeMarini, D. J., Winey, M., Ursic, D., Webb, F., and Culbertson, M. R. (1992). SEN1, a positive effector of tRNA-splicing endonuclease in *Saccharomyces cerevisiae*. *Molecular and cellular biology*, *12*(5), 2154–2164.

DeMarini, D.J., Papa, F.R., Swaminathan, S., Ursic, D., Rasmussen, T.P.,Culbertson, M.R., and Hochstrasser, M. (1995). The yeast SEN3 gene encodes a regulatory subunit of the 26S proteasome complex required for ubiquitindependent protein degradation in vivo. *Molecular and cellular biology, 15,* 6311–6321.

Dengl, S., and Cramer, P. (2009). Torpedo nuclease Rat1 is insufficient to terminate RNA polymerase II in vitro. *The Journal of biological chemistry*, *284*(32), 21270–21279.

Deshaies, R. J., and Joazeiro, C. A. (2009). RING domain E3 ubiquitin ligases. *Annual review of biochemistry*, *78*, 399–434.

Dosztányi, Z., Csizmok, V., Tompa, P., and Simon, I. (2005). IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content. *Bioinformatics (Oxford, England)*, *21*(16), 3433–3434.

Egloff, M. P., Cohen, P. T., Reinemer, P., and Barford, D. (1995). Crystal structure of the catalytic subunit of human protein phosphatase 1 and its complex with tungstate. *Journal of molecular biology, 254(5),* 942–959.

Egloff, M. P., Johnson, D. F., Moorhead, G., Cohen, P. T., Cohen, P., and Barford, D. (1997). Structural basis for the recognition of regulatory subunits by the catalytic subunit of protein phosphatase 1. *EMBO J. 16,* 1876–1887.

Finkel,J.S., Chinchilla,K., Ursic,D. and Culbertson,M.R. (2010) Sen1p performs two genetically separable functions in transcription and processing of U5 small nuclear RNA in *Saccharomyces cerevisiae. Genetics, 184*, 107–118

Finley, D. (2009). Recognition and processing of ubiquitin-protein conjugates by the proteasome. *Annual review of biochemistry*, *78*, 477–513.

Finley, D., Chen, X., and Walters, K. J. (2016). Gates, Channels, and Switches: Elements of the Proteasome Machine. *Trends in biochemical sciences*, *41*(1), 77–93.

Finley, D., Ulrich, H. D., Sommer, T., and Kaiser, P. (2012). The ubiquitin-proteasome system of Saccharomyces cerevisiae. Genetics, 192(2), 319–360.

Fromont-Racine, M., Rain, J. C., and Legrain, P. (1997). Toward a functional analysis of the yeast genome through exhaustive two-hybrid screens. Nature genetics, 16(3), 277–282.

Gan, W., Guan, Z., Liu, J., Gui, T., Shen, K., Manley, J. L., and Li, X. (2011). R-loop-mediated genomic instability is caused by impairment of replication fork progression. *Genes & development*, 25(19), 2041–2056.

Gelbart, M. E., Rechsteiner, T., Richmond, T. J., and Tsukiyama, T. (2001). Interactions of Isw2 chromatin remodeling complex with nucleosomal arrays: analyses using recombinant yeast histones and immobilized templates. *Molecular and cellular biology*, *21*(6), 2098–2106.

Geng, F., Wenzel, S., and Tansey, W. P. (2012). Ubiquitin and proteasomes in transcription. *Annual review of biochemistry*, *81*, 177–201.

Ghaemmaghami, S., Huh, W.K., Bower, K., Howson, R.W., Belle, A., Dephoure, N., O'Shea, E.K., and Weissman, J.S. (2003). Global analysis of protein expression in yeast. Nature 425, 737–741.

Ghosh, A., and Cannon, J. F. (2013). Analysis of protein phosphatase-1 and aurora protein kinase suppressors reveals new aspects of regulatory protein function in *Saccharomyces cerevisiae. PloS one, 8(7),* e69133.

Gietz, R. D., and Woods, R. A. (2002). Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method. *Methods in enzymology*, *350*, 87–96.

Gilbert, W., and Guthrie, C. (2004). The Glc7p nuclear phosphatase promotes mRNA export by facilitating association of Mex67p with mRNA. *Mol. Cell 13*, 201–212.

Glickman, M. H., Rubin, D. M., Coux, O., Wefes, I., Pfeifer, G., Cjeka, Z., Baumeister, W., Fried, V. A., and Finley, D. (1998). A subcomplex of the proteasome regulatory particle required for ubiquitin-conjugate degradation and related to the COP9-signalosome and eIF3. *Cell*, *94*(5), 615–623.

Glotzer, M., Murray, A. W., and Kirschner, M. W. (1991). Cyclin is degraded by the ubiquitin pathway. *Nature*, *349*(6305), 132–138.

Goldberg J, Huang HB, Kwon YG, Greengard P, Nairn AC, and Kuriyan J. (1995). Three-dimensional structure of the catalytic subunit of protein serine/threonine phosphatase-1. *Nature 376,* 745–753.

Golla, U., Singh, V., Azad, G. K., Singh, P., Verma, N., Mandal, P., Chauhan, S., and Tomar, R. S. (2013). Sen1p contributes to genomic integrity by regulating expression of ribonucleotide reductase 1 (RNR1) in *Saccharomyces cerevisiae. PloS one*, *8*(5), e64798.

Grohmann, K., Schuelke, M., Diers, A., Hoffmann, K., Lucke, B., Adams, C., Bertini, E., Leonhardt-Horti, H., Muntoni, F., Ouvrier, R., Pfeufer, A., Rossi, R., Van Maldergem, L., Wilmshurst, J. M., Wienker, T. F., Sendtner, M., Rudnik-Schöneborn, S., Zerres, K., and Hübner, C. (2001). Mutations in the gene encoding immunoglobulin mu-binding protein 2 cause spinal muscular atrophy with respiratory distress type 1. *Nature genetics*, *29*(1), 75–77.

Groll, M., Bochtler, M., Brandstetter, H., Clausen, T., and Huber, R. (2005). Molecular machines for protein degradation. *Chembiochem: a European journal of chemical biology*, *6*(2), 222–256.

Groll, M., Ditzel, L., Löwe, J., Stock, D., Bochtler, M., Bartunik, H. D., and Huber, R. (1997). Structure of 20S proteasome from yeast at 2.4 A resolution. *Nature*, *386*(6624), 463–471.

Gudipati, R. K., Xu, Z., Lebreton, A., Séraphin, B., Steinmetz, L. M., Jacquier, A., and Libri, D. (2012). Extensive degradation of RNA precursors by the exosome in wild-type cells. *Molecular cell*, *48*(3), 409–421.

Hamperl, S., and Cimprich, K. A. (2016). Conflict Resolution in the Genome: How Transcription and Replication Make It Work. *Cell*, *167*(6), 1455–1467.

Han, Z., Libri, D., and Porrua, O. (2017). Biochemical characterization of the helicase Sen1 provides new insights into the mechanisms of non-coding transcription termination. *Nucleic Acids Res 45:* 1355–1370

Han, Z., Jasnovidova, O., Haidara, N., Tudek, A., Kubicek, K., Libri, D., Stefl, R., and Porrua, O. (2020). Termination of non-coding transcription in yeast relies on both an RNA Pol II CTD interaction domain and a CTD-mimicking region in Sen1. *The EMBO journal*, *39*(7), e101548.

Harreman, M. T., Kline, T. M., Milford, H. G., Harben, M. B., Hodel, A. E., and Corbett, A. H. (2004). Regulation of nuclear import by phosphorylation adjacent to nuclear localization signals. *The Journal of biological chemistry*, *279*(20), 20613–20621.

Haruki, H., Nishikawa, J., and Laemmli, U. K. (2008). The anchor-away technique: rapid, conditional establishment of yeast mutant phenotypes. *Molecular cell*, *31*(6), 925–932.

Hatchi, E., Skourti-Stathaki, K., Ventz, S., Pinello, L., Yen, A., Kamieniarz-Gdula, K., Dimitrov, S., Pathania, S., McKinney, K. M., Eaton, M. L., Kellis, M., Hill, S. J., Parmigiani, G., Proudfoot, N. J., and Livingston, D. M. (2015). BRCA1 recruitment to transcriptional pause sites is required for R-loop-driven DNA damage repair. *Molecular cell*, *57*(4), 636–647.

He, X., and Moore, C. (2005). Regulation of Yeast mRNA 3′ End Processing by Phosphorylation. *Mol Cell 19(5):* 619–629.

Hegazy, Y. A., Fernando, C. M., and Tran, E. J. (2020). The balancing act of R-loop biology: The good, the bad, and the ugly. *The Journal of biological chemistry*, 295(4), 905–913.

Hein, M. Y., Hubner, N. C., Poser, I., Cox, J., Nagaraj, N., Toyoda, Y., Gak, I. A., Weisswange, I., Mansfeld, J., Buchholz, F., Hyman, A. A., and Mann, M. (2015). A

human interactome in three quantitative dimensions organized by stoichiometries and abundances. *Cell*, *163*(3), 712–723.

Helmrich, A., Ballarino, M., Nudler, E., and Tora, L. (2013). Transcription-replication encounters, consequences and genomic instability. *Nature structural and molecular biology*, *20*(4), 412–418.

Hendrickx, A., Beullens, M., Ceulemans, H., Den Abt, T., Van Eynde, A., Nicolaescu, E., Lesage, B., and Bollen, M. (2009). Docking motif-guided mapping of the interactome of protein phosphatase-1. *Chemistry and biology*, *16*(4), 365–371.

Hendriks, I. A., D'Souza, R. C., Yang, B., Verlaan-de Vries, M., Mann, M., and Vertegaal, A. C. (2014). Uncovering global SUMOylation signaling networks in a site-specific manner. *Nature structural and molecular biology, 21(10*), 927–936.

Heroes, E., Lesage, B., Görnemann, J., Beullens, M., Van Meervelt, L., and Bollen, M. (2013). The PP1 binding code: a molecular-lego strategy that governs specificity. *The FEBS journal*, *280*(2), 584–595.

Hershko, A. and Ciechanover, A. (1998). The ubiquitin system. *Annu. Rev. Biochem. 67,* 425–479

Ho, Y., Gruhler, A., Heilbut, A., Bader, G. D., Moore, L., Adams, S. L., Millar, A., Taylor, P., Bennett, K., Boutilier, K., Yang, L., Wolting, C., Donaldson, I., Schandorff, S., Shewnarane, J., Vo, M., Taggart, J., Goudreault, M., Muskat, B., Alfarano, C., … Tyers, M. (2002). Systematic identification of protein complexes in Saccharomyces cerevisiae by mass spectrometry. *Nature, 415(6868),* 180–183.

Husnjak, K., Elsasser, S., Zhang, N., Chen, X., Randles, L., Shi, Y., Hofmann, K., Walters, K. J., Finley, D., and Dikic, I. (2008). Proteasome subunit Rpn13 is a novel ubiquitin receptor. *Nature*, *453*(7194), 481–488.

Iakoucheva, L. M., Radivojac, P., Brown, C. J., O'Connor, T. R., Sikes, J. G., Obradovic, Z., and Dunker, A. K. (2004). The importance of intrinsic disorder for protein phosphorylation. *Nucleic acids research*, *32*(3), 1037–1049

Jackson, S.P., Durocher, D. (2013). Regulation of DNA damage responses by ubiquitin and SUMO. *Mol Cell 49:* 795–807.

Jamonnak, N., Creamer, T. J., Darby, M. M., Schaughency, P., Wheelan, S. J., and Corden, J. L. (2011). Yeast Nrd1, Nab3, and Sen1 transcriptome-wide binding maps suggest multiple roles in post-transcriptional RNA processing. *RNA (New York, N.Y.)*, *17*(11), 2011–2025.

Jenks, M. H., T. W. O'Rourke, and D. Reines. (2008). Properties of an intergenic terminator and start site switch that regulate IMD2 transcription in yeast. *Mol. Cell. Biol. 28*: 3883–3893.

Jensen, T. H., Jacquier, A., and Libri, D. (2013). Dealing with pervasive transcription. *Molecular cell, 52(4),* 473–484.

Jeronimo, C., and Robert, F. (2014). Kin28 regulates the transient association of Mediator with core promoters. *Nature structural and molecular biology*, *21*(5), 449–455.

Johnson L. N. (2009). The regulation of protein phosphorylation. *Biochemical Society transactions*, *37*(Pt 4), 627–641.

Kawauchi, J., Mischo, H., Braglia, P., Rondon, A., and Proudfoot, N. J. (2008). Budding yeast RNA polymerases I and II employ parallel mechanisms of transcriptional termination. *Genes and development*, *22*(8), 1082–1092.

Kecman, T., Kuś, K., Heo, D. H., Duckett, K., Birot, A., Liberatori, S., Mohammed, S., Geis-Asteggiante, L., Robinson, C. V., and Vasiljeva, L. (2018). Elongation/Termination Factor Exchange Mediated by PP1 Phosphatase Orchestrates Transcription Termination. *Cell reports,* 25(1), 259–269.e5.

Kim, M., Krogan, N. J., Vasiljeva, L., Rando, O. J., Nedea, E., Greenblatt, J. F., and Buratowski, S. (2004). The yeast Rat1 exonuclease promotes transcription termination by RNA polymerase II. *Nature*, *432*(7016), 517–522.

Knaus, M., Cameroni, E., Pedruzzi, I., Tatchell, K., De Virgilio, C., and Peter, M. (2005). The Bud14p-Glc7p complex functions as a cortical regulator of dynein in budding yeast. *EMBO J. 24*, 3000–3011.

Kozubowski, L., Panek, H., Rosenthal, A., Bloecher, A., DeMarini, D. J., and Tatchell, K. A. (2003). Bni4-Glc7 phosphatase complex that recruits chitin synthase to the site of bud emergence. *Mol. Biol. Cell* 14, 26–39.

Kuehner, J. N., Pearson, E. L., and Moore, C. (2011). Unravelling the means to an end: RNA polymerase II transcription termination. *Nature reviews. Molecular cell biology*, *12*(5), 283–294.

Lamontagne, B., Tremblay, A., Abou Elela, S. (2000) The N-terminal domain that distinguishes yeast from bacterial RNase III contains a dimerization signal required for efficient double-stranded RNA cleavage. *Mol. Cell. Biol.* 20: 1104 –1115.

Lanz, M. C., Yugandhar, K., Gupta, S., Sanford, E. J., Faça, V. M., Vega, S., Joiner, A., Fromme, J. C., Yu, H., and Smolka, M. B. (2021). In-depth and 3-dimensional exploration of the budding yeast phosphoproteome. *EMBO reports*, *22*(2), e51121.

Li, W., Selvam, K., Rahman, S. A., and Li, S. (2016). Sen1, the yeast homolog of human senataxin, plays a more direct role than Rad26 in transcription coupled DNA repair. *Nucleic acids research*, *44*(14), 6794–6802.

Lidschreiber, M., Easter, A. D., Battaglia, S., Rodríguez-Molina, J. B., Casañal, A., Carminati, M., Baejen, C., Grzechnik, P., Maier, K. C., Cramer, P., and Passmore, L. A. (2018). The APT complex is involved in non-coding RNA transcription and is distinct from CPF. *Nucleic acids research*, 46(21), 11528–11538.

Livneh, I., Kravtsova-Ivantsiv, Y., Braten, O., Kwon, Y. T., and Ciechanover, A. (2017). Monoubiquitination joins polyubiquitination as an esteemed proteasomal targeting signal. *BioEssays: news and reviews in molecular, cellular and developmental biology*, *39*(6), 10.1002/bies.201700027.

Longhese, M. P., Mantiero, D., and Clerici, M. (2006). The cellular response to chromosome breakage. *Mol. Microbiol.* 60, 1099–1108.

Luo, W., Johnson, A. W., and Bentley, D. L. (2006). The role of Rat1 in coupling mRNA 3'-end processing to transcription termination: implications for a unified allosteric-torpedo model. *Genes and development*, *20*(8), 954–965.

MacGilvray, M. E., Shishkova, E., Place, M., Wagner, E. R., Coon, J. J., and Gasch, A. P. (2020). Phosphoproteome Response to Dithiothreitol Reveals Unique *Versus* Shared Features of *Saccharomyces cerevisiae* Stress Responses. *Journal of proteome research*, *19*(8), 3405–3417.

Martin-Tumasz, S., and Brow, D. A. (2015). Saccharomyces cerevisiae Sen1 Helicase Domain Exhibits 5'- to 3'-Helicase Activity with a Preference for Translocation on DNA Rather than RNA. *The Journal of biological chemistry*, *290*(38), 22880–22889.

Mayer, A., Lidschreiber, M., Siebert, M., Leike, K., Söding, J., and Cramer, P. (2011). *Uniform transitions of the general RNA polymerase II transcription complex. Nature Structural and Molecular Biology, 17(10), 1272–1278.*

Mayor, T., Graumann, J., Bryan, J., MacCoss, M. J., and Deshaies, R. J. (2007). Quantitative profiling of ubiquitylated proteins reveals proteasome substrates and the substrate repertoire influenced by the Rpn10 receptor pathway. *Molecular and cellular proteomics: MCP*, *6*(11), 1885–1895.

McGrath, J. P., Jentsch, S., and Varshavsky, A. (1991). UBA 1: an essential yeast gene encoding ubiquitin-activating enzyme. *The EMBO journal*, *10*(1), 227–236.

Miller, M. S., Rialdi, A., Ho, J. S., Tilove, M., Martinez-Gil, L., Moshkina, N. P., Peralta, Z., Noel, J., Melegari, C., Maestre, A. M., Mitsopoulos, P., Madrenas, J., Heinz, S., Benner, C., Young, J. A., Feagins, A. R., Basler, C. F., Fernandez-Sesma, A., Becherel, O. J., Lavin, M. F., … Marazzi, I. (2015). Senataxin suppresses the antiviral transcriptional response and controls viral biogenesis. *Nature immunology*, *16*(5), 485–494.

Mischo, H. E., Chun, Y., Harlen, K. M., Smalec, B. M., Dhir, S., Churchman, L. S., and Buratowski, S. (2018). Cell-Cycle Modulation of Transcription Termination Factor Sen1. *Molecular cell*, *70*(2), 312–326.e7.

Moorhead, G. B., Trinkle-Mulcahy, L., and Ulke-Lemee, A. (2007). Emerging roles of nuclear protein phosphatases. *Nat. Rev. Mol. Cell Biol.* 8, 234–244.

Moreira, M. C., Klur, S., Watanabe, M., Németh, A. H., Le Ber, I., Moniz, J. C., Tranchant, C., Aubourg, P., Tazir, M., Schöls, L., Pandolfo, M., Schulz, J. B., Pouget, J., Calvas, P., Shizuka-Ikeda, M., Shoji, M., Tanaka, M., Izatt, L., Shaw, C. E., M'Zahem, A., … Koenig, M. (2004). Senataxin, the ortholog of a yeast RNA helicase, is mutant in ataxia-ocular apraxia 2. *Nature genetics*, *36*(3), 225–227.

Nakatogawa, H., Suzuki, K., Kamada, Y., and Ohsumi, Y. (2009). Dynamics and diversity in autophagy mechanisms: lessons from yeast. *Nature reviews. Molecular cell biology*, *10*(7), 458–467.

Newman, J.R., Ghaemmaghami, S., Ihmels, J., Breslow, D.K., Noble, M., DeRisi, J.L., and Weissman, J.S. (2006). Single-cell proteomic analysis of *S. cerevisiae* reveals the architecture of biological noise. *Nature 441*,840–846

Nigavekar, S. S., Tan, Y. S. H., and Cannon, J. F. (2002). Glc8 is a glucose-repressible activator of Glc7 protein phosphatase-1. *Arch. Biochem. Biophys*. 404, 71–79.

Niu, S., Wang, Z., Ge, D., Zhang, G., and Li, Y. (2012). Prediction of functional phosphorylation sites by incorporating evolutionary information. *Protein & cell*, *3*(9), 675–690.

Offley, S. R., and Schmidt, M. C. (2019). Protein phosphatases of *Saccharomyces cerevisiae. Current genetics, 65(1*), 41–55.

Parua, P. K., Booth, G. T., Sansó, M., Benjamin, B., Tanny, J. C., Lis, J. T., and Fisher, R. P. (2018). A Cdk9-PP1 switch regulates the elongation-termination transition of RNA polymerase II. *Nature,* 558(7710), 460–464

Pearson, E., and Moore, C. (2014). The evolutionarily conserved Pol II flap loop contributes to proper transcription termination on short yeast genes. *Cell reports*, *9*(3), 821–828.

Pedelini L, Marquina Rodríguez M, Ariño J, Casamayor A, Sanz L, Bollen M, Sanz P, and García-Gimeno A (2007). YPI1 and SDS22 Proteins Regulate the Nuclear Localization and Function of Yeast Type 1 Phosphatase Glc7. *J Biol Chem 282(5)*:3282-92.

Pesin, J. A., and Orr-Weaver, T. L. (2008). Regulation of APC/C activators in mitosis and meiosis. Annual review of cell and developmental biology, 24, 475–499.

Peters J. M. (2006). The anaphase promoting complex/cyclosome: a machine designed to destroy. *Nature reviews. Molecular cell biology,* 7(9), 644–656.

Pfleger, C. M., and Kirschner, M. W. (2000). The KEN box: an APC recognition signal distinct from the D box targeted by Cdh1. *Genes and development*, *14*(6), 655–665.

Porrua, O., and Libri, D. (2013). A bacterial-like mechanism for transcription termination by the Sen1p helicase in budding yeast. *Nat Struct Mol Biol 20*, 884–891

Porrua, O., Hobor, F., Boulay, J., Kubicek, K., D'Aubenton-Carafa, Y., Gudipati, R. K., Stefl, R., and Libri, D. (2012). *In vivo* SELEX reveals novel sequence and structural determinants of Nrd1-Nab3-Sen1-dependent transcription termination. *The EMBO journal*, *31*(19), 3935–3948.

Rawal, C. C., Zardoni, L., Di Terlizzi, M., Galati, E., Brambati, A., Lazzaro, F., Liberi, G., and Pellicioli, A. (2020). Senataxin Ortholog Sen1 Limits DNA:RNA Hybrid Accumulation at DNA Double-Strand Breaks to Control End Resection and Repair Fidelity. *Cell reports*, *31*(5), 107603.

Richard, P. and Manley, J. L. (2009). Transcription termination by nuclear RNA polymerases. *Genes Dev. 23,* 1247–1269

Richard, P., Feng, S., and Manley, J. L. (2013). A SUMO-dependent interaction between Senataxin and the exosome, disrupted in the neurodegenerative disease AOA2, targets the exosome to sites of transcription-induced DNA damage. *Genes and development*, *27*(20), 2227–2232.

Rondón, A. G., Mischo, H. E., Kawauchi, J., and Proudfoot, N. J. (2009). Fail-safe transcriptional termination for protein-coding genes in S. cerevisiae. *Molecular cell*, *36*(1), 88–98.

Sanz, P., Alms, G. R., Haystead, T. A., and Carlson, M. (2000). Regulatory interactions between the Reg1-Glc7 protein phosphatase and the Snf1 protein kinase. *Mol. Cell. Biol. 20,* 1321–1328.

Sauer, R. T., and Baker, T. A. (2011). AAA+ proteases: ATP-fueled machines of protein destruction. *Annual review of biochemistry*, *80*, 587–612.

Schaughency, P., Merran, J., and Corden, J. L. (2014). Genome-wide mapping of yeast RNA polymerase II termination. *PLoS genetics*, *10*(10), e1004632.

Schneider, R., Massow, M., Lisowsky, T., and Weiss, H. (1995). Different respiratory-defective phenotypes of Neurospora crassa and Saccharomyces cerevisiae after inactivation of the gene encoding the mitochondrial acyl carrier protein. *Current genetics*, *29*(1), 10–17.

Schreieck, A., Easter, A. D., Etzold, S., Wiederhold, K., Lidschreiber, M., Cramer, P., & Passmore, L. A. (2014). RNA polymerase II termination involves C-terminal-domain tyrosine dephosphorylation by CPF subunit Glc7. *Nature structural & molecular biology*, 21(2), 175–179.

Shi, Y. (2009). Assembly and structure of protein phosphatase 2A. *Science in China. Series C, Life sciences*, *52*(2), 135–146.

Shiber, A., and Ravid, T. (2014). Chaperoning proteins for destruction: diverse roles of Hsp70 chaperones and their co-chaperones in targeting misfolded proteins to the proteasome. *Biomolecules*, *4*(3), 704–724.

Sikorski, R. S., and Hieter, P. (1989). A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in *Saccharomyces cerevisiae. Genetics,* 122(1), 19–27.

Sivakumar, S., and Gorbsky, G.J. (2015). Spatiotemporal regulation of the anaphase-promoting complex in mitosis. *Nat. Rev. Mol. Cell Biol.* 16, 82–94.

Skourti-Stathaki, K., Proudfoot, N.J., and Gromak, N. (2011). Human senataxin resolves RNA/DNA hybrids formed at transcriptional pause sites to promote Xrn2-dependent transcription. *Mol. Cell 42*: 794–805.

Skowyra, D., Craig, K. L., Tyers, M., Elledge, S. J., and Harper, J. W. (1997). F-box proteins are receptors that recruit phosphorylated substrates to the SCF ubiquitin-ligase complex. *Cell*, *91*(2), 209–219.

Sommers, J. A., Suhasini, A. N., and Brosh, R. M., Jr (2015). Protein degradation pathways regulate the functions of helicases in the DNA damage response and maintenance of genomic stability. *Biomolecules*, *5*(2), 590–616.

Stark, M. J. (1996). Yeast protein serine/threonine phosphatases: Multiple roles and diverse regulation. *Yeast 12,* 1647–1675.

Steinmetz, E. J., Conrad, N. K., Brow, D. A., and Corden, J. L. (2001). RNA-binding protein Nrd1 directs poly(A)-independent 3'-end formation of RNA polymerase II transcripts. *Nature*, *413*(6853), 327–331.

Steinmetz,E.J., Warren,C.L., Kuehner,J.N., Panbehi,B., Ansari,A.Z. and Brow,D.A. (2006). Genome-wide distribution of yeast RNA polymerase II and its control by Sen1 helicase*. Mol. Cell, 24*, 735–746.

Swaney, D. L., Beltrao, P., Starita, L., Guo, A., Rush, J., Fields, S., Krogan, N. J., and Villén, J. (2013). Global analysis of phosphorylation and ubiquitylation cross-talk in protein degradation. *Nature methods*, *10*(7), 676–682.

Terrak, M., Kerff, F., Langsetmo, K., Tao, T., and Dominguez, R. (2004). Structural basis of protein phosphatase 1 regulation. *Nature 429,* 780–784.

Tietjen, J. R., Zhang, D. W., Rodríguez-Molina, J. B., White, B. E., Akhtar, M. S., Heidemann, M., Li, X., Chapman, R. D., Shokat, K., Keles, S., Eick, D., and Ansari, A. Z. (2010). Chemical-genomic dissection of the CTD code. *Nature structural and molecular biology*, *17*(9), 1154–1161.

Tu, J., Song, W., and Carlson, M. (1996). Protein phosphatase type 1 interacts with proteins required for meiosis and other cellular processes in Saccharomyces cerevisiae. *Mol. Cell. Biol. 16,* 4199–4206.

Tudek, A., Porrua, O., Kabzinski, T., Lidschreiber, M., Kubicek, K., Fortova, A., Lacroute, F., Vanacova, S., Cramer, P., Stefl, R., and Libri, D. (2014). Molecular basis for coordinating transcription termination with noncoding RNA degradation. *Molecular cell*, *55*(3), 467–481.

Ursic, D., Chinchilla, K., Finkel, J. S., and Culbertson, M. R. (2004). Multiple protein/protein and protein/RNA interactions suggest roles for yeast DNA/RNA helicase Sen1p in transcription, transcription-coupled DNA repair and RNA processing. *Nucleic acids research*, *32*(8), 2441–2452.

Ursic, D., Himmel, K. L., Gurley, K. A., Webb, F., and Culbertson, M. R. (1997). The yeast SEN1 gene is required for the processing of diverse RNA classes. *Nucleic acids research*, *25*(23), 4778–4785.

Vance, C., Rogelj, B., Hortobágyi, T., De Vos, K. J., Nishimura, A. L., Sreedharan, J., Hu, X., Smith, B., Ruddy, D., Wright, P., Ganesalingam, J., Williams, K. L., Tripathi, V., Al-Saraj, S., Al-Chalabi, A., Leigh, P. N., Blair, I. P., Nicholson, G., de Belleroche, J., Gallo, J. M., … Shaw, C. E. (2009). Mutations in FUS, an RNA processing protein, cause familial amyotrophic lateral sclerosis type 6. *Science (New York, N.Y.)*, *323*(5918), 1208–1211.

Visintin, R., Prinz, S., and Amon, A. (1997). CDC20 and CDH1: a family of substrate-specific activators of APC-dependent proteolysis. *Science (New York, N.Y.)*, *278*(5337), 460–463.

Wagschal, A., Rousset, E., Basavarajaiah, P., Contreras, X., Harwig, A., Laurent-Chabalier, S., Nakamura, M., Chen, X., Zhang, K., Meziane, O., Boyer, F., Parrinello, H., Berkhout, B., Terzian, C., Benkirane, M., and Kiernan, R. (2012). Microprocessor, Setx, Xrn2, and Rrp6 co-operate to induce premature termination of transcription by RNAPII. *Cell*, *150*(6), 1147–1157.

Wendler, P., and Enenkel, C. (2019). Nuclear Transport of Yeast Proteasomes. *Frontiers in Molecular Biosciences, 6.*

Weng, Y., Czaplinski, K., and Peltz, S. W. (1996). Genetic and biochemical characterization of mutations in the ATPase and helicase regions of the Upf1 protein. *Molecular and cellular biology*, *16*(10), 5477–5490.

West, S., Gromak, N., and Proudfoot, N. J. (2004). Human 5' --> 3' exonuclease Xrn2 promotes transcription termination at co-transcriptional cleavage sites. *Nature*, *432*(7016), 522–525.

Wu, X., and Tatchell, K. (2001). Mutations in Yeast Protein Phosphatase Type 1 that Affect Targeting Subunit Binding. *Biochemistry 40(25):* 7410–7420.

Wyers, F., Rougemaille, M., Badis, G., Rousselle, J. C., Dufour, M. E., Boulay, J., Régnault, B., Devaux, F., Namane, A., Séraphin, B., Libri, D., and Jacquier, A. (2005). Cryptic pol II transcripts are degraded by a nuclear quality control pathway involving a new poly(A) polymerase. *Cell*, *121*(5), 725–737.

Xiang, K., Tong, L., and Manley, J. L. (2014). Delineating the structural blueprint of the pre-mRNA 3'-end processing machinery. *Molecular and cellular biology*, *34*(11), 1894–1910.

Zachariae, W., Schwab, M., Nasmyth, K., and Seufert, W. (1998). Control of cyclin ubiquitination by CDK-regulated binding of Hct1 to the anaphase promoting complex. *Science (New York, N.Y.)*, *282*(5394), 1721–1724.

Zeng, G., Huang, B., Neo, S. P., Wang, J., and Cai, M. (2007). Scd5p mediates phosphoregulation of actin and endocytosis by the type 1 phosphatase Glc7p in yeast. *Mol. Biol. Cell 18,* 4885–4898.

Zhao, D. Y., Gish, G., Braunschweig, U., Li, Y., Ni, Z., Schmitges, F. W., Zhong, G., Liu, K., Li, W., Moffat, J., Vedadi, M., Min, J., Pawson, T. J., Blencowe, B. J., and Greenblatt, J. F. (2016). SMN and symmetric arginine dimethylation of RNA polymerase II C-terminal domain control termination. *Nature*, *529*(7584), 48–53.

Zhao, S., and Lee, E. Y. (1997). A protein phosphatase-1-binding motif identified by the panning of a random peptide display library. *J. Biol. Chem. 272*, 28368–28372.

Annex A - Strains, plasmids and oligos

Table 2. – Plasmids used in this study

Plasmid	Description	Source
pMPY-3xMyc	3xMyc C-terminal tagging	Schneider et al., 1995
p3Flag-KanMX	3xFlag C-terminal tagging	Gelbart et al., 2001
pFA6a-FRBkanMX6	FRB C-terminal tagging	Euroscarf
pRS313	CEN, HIS3	Sikorski and Hieter, 1989
pFR711	CEN, HIS3, Sen1-WT	Collin et al., 2018
pFR737	CEN, HIS3, Sen1-WT-3xHA	Collin et al., 2018
pUC18	Cloning and manipulation of DNA fragments	
pFR830	pUC18, Sen1-F2003A fragment	This study
pFR850	CEN, HIS3, Sen1-F2003A-3xHA	This study

Table 3. – Yeast strains used in this study

Strain	Genotype	Source
yFR116	MATa, ade2-1, trp1-1, can1-100, leu2-3,112, his3-11,15, ura3, GAL+, psi+	Young Laboratory
yFR1894	MATα, ade2-1, trp1-1, can1-100, leu2-3,112, his3-11,15, ura3, GAL, psi+, tor1-1, fpr1Δ::NAT, RPL13A-2xFKBP12::TRP1, GLC7-FRB::KanMX6	Schreieck et al., 2014
yFR2277	MATα ade2-1 trp1-1 can1-100 leu2-3,112 his3-11,15 ura3 GAL psi+ tor1-1 fpr1::NAT RPL13A-2xFKBP12::TRP1 SEN1-FRB-3xFLAG::KanMX6	Collin et al., 2018
yFR3097	MATα ade2-1 trp1-1 can1-100 leu2-3,112 his3-11,15 ura3 GAL psi+ tor1-1 fpr1::NAT RPL13A-2xFKBP12::TRP1 GLC7-FRB::kanMX6 Sen1-3MYC::URA3 pdr5D::LEU2	Collin et al., 2018
yFR3195	MATa, ade2-1, trp1-1, can1-100, leu2-3,112, his3-11,15, ura3, GAL+, psi+, SEN1-3HA, pFR754	Collin et al., 2018
yFR3313	MATa, ade2-1, trp1-1, can1-100, leu2-3,112, his3-11,15, ura3, GAL+, psi+, GLC7-3xFLAG::KanMX6	This study

Table 4. – List of oligonucleotides used in this study

Use	Direction	Sequence (5' ⟶ 3')
Tagging Sen1-3xMYC	F	GGAATGCTTCATCTAGCCCATTTATCCCAAAAAAAAGAAAGCCTAGATCAAGGGAACAAAAGCTGG
	R	TATACACCAATATATATGCAGGTATAATTCCTAACACTTTTACTTCAAGACTATAGGGCGAATTGG
Screening Sen1-3xMYC	F	TTCGAGACGGAATGCTTCAT
	R	CAAAGATTGGCATTTCTGAAA
Amplify Sen1 ORF	F	ATCGCCGCGGTACATAAGACGGCGGCGATA
	R	ATCGCTCGAGTAACCAAACTTCGGAGCAGC
Add 3xHA to Sen1 plasmid	F	TAAGCTACCCATGGAATATATTACGCAAGG
	R	TGAATCATGTTCACCTGTCCA
Add SexAI restriction site to pUC18	F	CGATCGACCTGGTGGTCGACTCTAGAGGATC
	R	TATCGTACCAGGTTGCAGGCATGCAAGCTTG
Inversed PCR to make Sen1-F2003A	F	GCATCGGATGATGTTAGTTTCA
	R	ACAAACGTGTTTGTTATTTTT
To sequence F2003A mutation	F	AAAGTCCAAGGCAGATGATAAGA
	R	TGGTTATGAGAAGCCACCCT
Tagging Glc7-3xFLAG	F	AGCCAGCCCAAAAAAGTCTACCAAGGCAAGCTGGGGGTAGAAAGAAAAAAAGGGAACAAAAGCTGG
	R	TGTTAATAAGTATTTTCCTTTTTAAACTTTGATTTAGGACGTGAATCTATCTATAGGGCGAATTGG
Screening Glc7-3xFLAG	F	GTGGTGAATTTGATAACGCTGG
	R	TCTTCGCTTAAAGAATTGACT

Annex B – Sen1-WT Sequence

Sen1 sequence (including promoter and terminator)

GTGAATCATGTTCACCTGTCCACTAACAAAAAAGGAAACAAAATGTTGCTTTCCCACAGGAAAAAGAATGTTTTTGAC
AAAACACTAAGCGTTATGTGACCAATGTATAATTGTCCTCATCATATAGTTTTATGCAAGCATAAATAAGTCTTTCTT
TTAGTTAAAACGCGTTACATAGTAAAGGTAACCCTGAATGAATTTTTATTTATATTTTTCTTCGCATATTTTAGGATC
TTGAAAAATTAACCACAAGTTAGAAGTTGATGAGATCAGGAATTCTAACAAAAGGTCATTGGTTAGGTTTATTTTGT
AGGTGCACGTTTGCAACTTTATTCCTTTTGTTTCAAAGTTATAGAGCTTTCATTTTCTCATTCTTCTCTAGTGTAGGT
CATTTCTCCTGCATATTAACCTACTAGAAAAGTTTACGAAGCTTCATAGGAACGAATAATCGATCACTTTAAAGGCTT
CCCGCACTTTAATCTTATCATCTACACCTACAAAAAAGCTCTACTTTGTCATTTATTTTCTACTTCATCTCTGTATGGC
ATCTATCTCTATATATATAAAAAAGCGCATCTGTTTATTATAAATGAATTCCAACAATCCTGATAATAATAATAGCAA
CAACATTAATAATAATAATAAGGATAAGGATATTGCCCCCAATAGCGATGTTCAGCTAGCTACCGTCTATACCAAGGC
AAAGTCTTACATTCCTCAGATTGAACAAGTTTATCAAGGCACAAATCCAAATATTCAAGAGGCAAAATTATTAGGGG
AGTTATTACAAGTTCTAGCAGAAGTTCCCAAGGGTACGCATTTATTTTGCGACCCCATACTTGAACCAATATCAATTT
TTTCCCTAACCATTTTCTCATTCAATGAAGAAGCTACGGCCACATGGCTGAAGAATCATTTTAATCCAATTCTCTCTG
TTTGTGATAAATGCATTCTAAATTTTGCGAGAGGGAAGTGCAAAATGTTACAACATTTTGCCATTCAAAGACATGTGC
CGCATGAACATGTAGCCAAGTTTAATGACATCGTATGCCAATGGAGGGTGGAGGCTGTTTTTCCTATTCTTAGAAATA
TTTCCGTTAACGATAATACGGGAATTAACATTACAAATGAAATTGAAACTGCAATGTACGAGTGTCTCTGTAATCCGC
ATATGTTGAGGCTTAACAAGCAACTTAAGGCTACTTTCGAAGCAATCTTCAAATTTTTCTACGACACAAAGCATCGTT
TGTTGGACGTTACGAATCCATTAAGCATAAAAACCTTTATATCAGGTGTCATATTTTGCTGGTGTGAAGGTTCAAAAG
AGGAAAATGAGTGGTCACGAGCTTTTCTAAAGGATTTATACTCAAGAAATTTCCATATAAATCTTTCTAACCTGACAC
CGGATATTATTGAGGAAGTTTATATCCATATTTTATTCTTACAGAATCCGGCAAACTGGACAGAGATTGTGGTATCTC
AATTCTGGTCGAGATTACTACCGGTTTTCAATTTATTTGATAAAGACGTCTTCATAGAGTACTTTCAGGTCCCAAAAA
ATGTGGAATCTTTAAAAAAAACATTCAAATTTCCGTTAGAGCCAATTTTTAAAATGTGGTACAATCACTTAAGCAAA
TCATATCATGATAAGCCACTAGACTTTTTGTTAAGAGGATTAACAATGTTCTTGAATAAATTTGGCTCCGAGTTTTGG
TCCAAAATTGAACCCTTCACATTCCATAGCATTTTGGATATAATTTTCAATAGAGATTCATTCCCAATCAAACTAATC
AAAATTCAAGACAATCCAATTGTTGAGCATCAAACAGAAGTTTACTTTCAACTTACCGGTTCTGTAACCGATTTACTT
TCTTGGACCTTACCATTTTATCATGCACTGTCACCATCCAAAAGAATTCAAATGGTTAGAAAGGTCTCAATGGCGTTC
CTGCGTATAATAGCCAATTATCCATCCCTAAAATCCATTCCCAAAGCTTGTTTAATGAATTCAGCAACTGCCTTGTTG
AGGGCTGTGTTGACTATTAAGGAAAATGAAAGGGCAATGCTTTATAAGAATGATGAATTTGAAACAGTACTGTTGAC
TAAAACAGATTCTAGAGCTTTGCTCAACAATCCTTTGATCCAAGACATAATAATTAGATCAGCTTCGAATCCTAACGA
CTTTTATCCTGGGCTAGGTGCTGCTTCTGCATCGGTGGCCACATCCACGATGATGGTATTGGCAGAATGTATCGATTT
TGATATACTGTTATTATGTCATCGAACGTTTAAGCTTTATTCAGGTAAACCAATTAGCGAAATTCCAATTTCAACCAA
CGTCTTGGAAAATGTTACAAATAAAATTGACCTGAGATCTTTTCATGACGGTCCATTGCTGGCCAAACAGCTGCTAGT
ATCTTTAAAAAATATTAATGGTCTTTTGATAGTACCGTCAAATACTGCCGTAGCTGAGGCACACAACGCATTGAATCA
GAAGTTTTTGCTGCTATCAACGAGATTAATGGAAAAATTTGCGGATATTTTGCCCGGTCAACTGTCTAAAATCCTTGC
GGATGAAGATGCATCGCAGGGTTTTTGGTCCTGCATATTCTCCTCTGATAAACACTTGTATCAAGCAGCTACTAATAT
TTTATACAATACATTTGACGTTGAGGGTAGACTGGAAGGTATCTTAGCGATTCTGAATAGTAATTTGACCGTTAACTT
GAAGAATATAAACGTGATGCTACAAAGATTAATCAATTGTGAATTCTATGAGCCTTGCCCCCGTGCTGTTCGTGTTTT
GATGGACGTAGTGAGTGCATTTGTAGACCCGATTTCTGGTGTGTTTGCTAATTTCCAAACTTTGAAGAGCCAAAATAC
CGAAAAGGAATTTTTAAAATTCTGGGAATCATGTTGGTTATTCTTGGATACGATTTATAAGTTTACTTTGAAATGGG
CTTCTAAATATGACTATTCTGAACTGGAAAATTTCACCAAAGATACTTTAGATCTAAGTCGCTCTTTGGTGGACTCCT
TTAGAGAATTTTCAGATATTCTTCATGACCAAACGAAAAATTTACTTCTCAACGTTTTAGAAACCTTTAAGAATATGT
TGTATTGGCTGAGATTGAGCGACGAAGTACTTTTGGAATCTTGTGTTAGATTAATTATAAGCACGTCTGATTTAGCAC
ATGAAAAGCATGTGAAAGTGGATGATTCCTTGGTAGAAATGATGGCCAAGTACGCCTCGAAGGCAAAGAGATTTTCA
AATAAGTTGACGGAGCAACAGGCGAGTGAGATTCTACAAAAGGCTAAAATATTCAATAAAGCACTCACGGAAGAGGT
CGCTACAGAAGCTGAAAACTACAGAAAAGAAAAGGAATTATCAAGACTTGGGAAGGTGATCGACTTAACGGATTCTG
TGCCGGCGTCACCATCCCTCTCGCCCTCTCTGTCTTCAACCATTGCTAGTTCTTCTGCTGAATCCAGGGCAGACTATTT
ACAAAGAAAAGCTTTATCTTCTTCAATAACTGGGAGACCAAGAGTAGCTCAGCCTAAAATAACATCTTTTGGTACGTT
CCAATCATCAGCTAATGCTAAACTGCATCGTACAAAACCTGTTAAGCCTTTATCCAAAATGGAATTAGCTAGGATGCA
ATTATTAAACAACAGAGTTGTTCATCCACCTAGTGCGCCAGCTTTCCATACAAAGAGTAGGGGTTTGTCTAACAAAAA
CGACGATAGCAGTAGTGAGGAGAGTGACAACGATATAGAAAGCGCTCGGGAGCTTTTTGCTATTGCTAAGGCAAAAG

GCAAAGGAATACAAACTGTCGATATTAATGGTAAGGTTGTCAAGAGACAAACTGCGGCTGAATTGGCTAAACAGGAG
TTAGAACATATGAGAAAGAGATTAAATGTTGATATGAACCCTCTATATGAAATCATATTGCAATGGGATTACACCAG
AAATAGTGAATATCCTGATGATGAGCCTATTGGTAATTATTCTGACGTAAAGGATTTCTTCAACTCTCCTGCTGACTA
CCAAAAAGTCATGAAGCCTTTACTACTTCTAGAATCTTGGCAAGGTCTATGTTCTTCACGTGATAGAGAGGACTATAA
ACCTTTCAGTATTATTGTAGGAAATCGAACTGCTGTGTCAGATTTCTATGATGTGTATGCCTCTGTCGCTAAACAAGT
CATACAGGATTGTGGTATTTCTGAATCTGACTTAATTGTAATGGCTTATTTACCCGATTTTCGACCAGACAAAAGGTT
GTCCAGTGATGATTTCAAGAAAGCACAACATACATGTTTAGCAAAGGTGAGAACATTGAAAAATACCAAAGGTGGTA
ATGTTGATGTCACATTACGAATTCATAGAAATCATTCTTTCAGTAAATTTTTGACACTAAGATCAGAGATATATTGTG
TGAAGGTTATGCAAATGACGACTATTGAAAGAGAGTATTCAACACTGGAAGGGCTGGAATACTATGACTTAGTTGGC
CAAATTCTACAGGCTAAACCATCGCCTCCCGTAAATGTGGATGCGGCAGAAATTGAAACAGTTAAGAAAAGCTATAAG
CTCAATACGTCGCAGGCTGAGGCCATCGTTAATTCGGTATCTAAAGAAGGTTTTTCTTTAATTCAGGGTCCACCTGGT
ACGGGTAAAACAAAGACTATACTGGGTATTATTGGTTACTTTTTGTCTACAAAGAATGCCTCATCTTCAAATGTCATC
AAAGTGCCACTAGAAAAGAACTCATCGAATACTGAACAACTGTTAAAAAAGCAAAAAATTCTAATCTGTGCCCCCAGT
AATGCCGCTGTGGACGAAATCTGCCTCAGACTGAAGAGTGGAGTGTATGACAAGCAGGGGCATCAATTTAAACCTCAA
TTGGTCCGTGTAGGTAGGTCAGACGTTGTAAACGTTGCAATTAAGGACCTTACTTTAGAAGAACTTGTGGATAAAAG
GATAGGTGAAAGGAATTACGAAATCAGAACAGATCCGGAACTAGAACGTAAATTTAATAATGCTGTGACTAAAAGAA
GAGAATTGAGAGGAAAATTAGACTCTGAGAGTGGTAATCCAGAAAGTCCAATGTCAACCGAAGATATCAGTAAATTA
CAACTTAAGATTAGGGAGCTAAGTAAAATAATCAATGAACTTGGCCGTGATAGGGACGAAATGCGTGAGAAGAATTC
CGTAAATTATAGGAATAGGGATTTGGATAGAAGGAATGCACAGGCCCATATCTTGGCGGTTAGTGACATCATATGTTC
GACTTTATCTGGTTCAGCCCATGATGTACTCGCAACAATGGGTATTAAATTTGATACGGTCATTATAGATGAAGCATG
TCAATGTACCGAACTTTCTTCAATTATTCCTCTCAGGTATGGAGGGAAGCGTTGTATTATGGTTGGTGATCCTAACCA
ACTACCACCAACTGTTCTTTCAGGTGCAGCAAGCAATTTTAAATACAATCAGTCACTGTTCGTTAGAATGGAGAAAAA
TAGCTCCCCATACTTGCTGGATGTTCAATACCGTATGCATCCTTCCATCAGTAAATTCCCCTCTTCTGAATTTTATCAG
GGACGACTAAAAGACGGCCCTGGTATGGATATCTTGAACAAAAGACCCTGGCACCAGCTTGAGCCATTAGCTCCCTAC
AAATTTTTTGACATCATTAGTGGTAGGCAAGAACAAAATGCGAAGACTATGTCTTATACAAATATGGAGGAAATACG
TGTTGCCATTGAGTTAGTTGACTACCTTTTCAGAAAATTCGATAATAAAATTGATTTCACGGGAAAAATCGGTATAAT
TTCACCTTATAGAGAGCAAATGCAAAAAATGCGTAAAGAGTTTGCCCGCTATTTTGGAGGCATGATCAACAAATCAAT
CGATTTCAATACAATCGACGGTTTCCAAGGTCAAGAAAAAGAAATCATCTTGATATCGTGTGTCCGTGCCGATGATAC
GAAATCTAGTGTCGGTTTTTTAAAGGACTTTCGTCGTATGAATGTGGCTCTGACTAGAGCCAAGACAAGTATTTGGGT
TTTGGGACATCAAAGGTCTCTGGCTAAAAGCAAGTTATGGAGAGATTTGATTGAAGATGCAAAAGATAGAAGTTGTC
TTGCATATGCTTGCTCGGGTTTCCTGGATCCCAGAAATAATAGAGCTCAAAGTATTTTAAGAAAGTTCAATGTACCCG
TACCATCAGAGCAAGAAGATGATTATAAGCTACCCATGGAATATATTACGCAAGGCCCCGATGAGGTAAAATCCAATA
AGGACACAAAGAAGAGAAGAGTTGTCGATGAAGGTGAAGAGGCAGATAAAGCTGTAAAGAAAAAGAAGAAAGAGAA
GAAGAAGGAAAAGAAAAAGTCCAAGGCAGATGATAAGAAAAAGAATAACAAGAAGGCAGAATCGCCCAGCACCTCAT
CTGGTACAAAGAAAAAGTCATCCATTTTTGGTGGAATGAGTGTACCGAGTGCCGTAGTTCCTAAAACTTTTCCTGATG
TGGACAGCAACAAAAAGGCAGCAGCAGTTGTTGGTAAAAAGAAAAATAACAAACACGTTTGTTTTTCGGATGATGTT
AGTTTCATACCAAGAAATGACGAACCTGAAATTAAAGTGACCAGAAGTTTATCCTCTGTATTGAAGGAAAAGCAGTT
AGGTTTGAAGGAGACGAGAACAATTTCTCCGCCTGAAATCAGTAATAATGAGGATGATGATGACGAAGATGATTATA
CGCCGTCAATCTCCGATTCTTCATTAATGAAAAGTGAAGCAAATGGTCGTAACAACAGGGTGGCTTCTCATAACCAAA
ATTTCAGCGCTAGTATATACGACGATCCACAAGTTTCGCAAGCAAAGCAGACGCAGGTACCTGCTGCTATTACAAAAC
ACAGAAGTTCAAATAGTGTTTTATCCGGTGGTTCTTCGAGGATCCTGACTGCTTCCGATTACGGTGAACCCAATCAAA
ATGGACAGAATGGAGCAAATAGGACCCTCTCTCAACATGTGGGTAATGCGAATCAATACTCAACGGCTCCTGTGGGTA
CAGGCGAACTGCATGAAACTCTGCCCGCCCATCCGCAAGACAGCTATCCTGCTGAGGCAGAAGATCCATATGATTTAA
ACCCCCATCCGCAACCACAATCTTCAGCATTCAAAGGACCTGGTAGTGGGCCCACAGGCACACGCAACAGTTCGAGACG
GAATGCTTCATCTAGCCCATTTATCCCAAAAAAAAGAAAGCCTAGATCATGATCTTGAAGTAAAAGTGTTAGGAATTA
TACCTGCATATATATTGGTGTATACATTTATATTGAATTTAATAGCAACATTATTTTTCATTATTTTTCATTATTTCA
GCAAAAAGAACATTTTTTTTCAATTTGACGTGAAAGAGTTTGACTATTACGTGGGCATACAATGTTTAAATTTACATT
ATCCTCATGGCTACTCTAGCTATTTGCATTTCATTCAACTTTTTTTGATATGGCATCATTTTTATTAGCGTTTACTAC
ACTTTCAGAAATGCCAATCTTTGGTGGCGATTTGCTAGCAACCAAGAATGAGGGAGTCGAGGGCTGTTTTAAATTTGA
ACTCGCGTTCGACCCTTTATGGCTTTTGTGTATATCAATTGGTTTAGAAGGTAGTGCGCTTGACCCATCTTGTGCAGT
GGTACTTTCGCTATGGAGTTCTTCACAAGCTTTCGAAAGTACGTCCCTTTCGTTTGATATTAAAGTTGTCGTGAGCTT
TTCATCAGAAGAGTCAAGAAGGTTATTTAGGTACTGTATTGTAGAATTGATCATTTCTGACATTTCATTTGGAGGTAT
TTTGGAGTTGCTGCTGCTCCGAAGTTTGGTTA

Annex C – Sen1-WT-3xHA Plasmid Sequence

pFR737 sequence (Sen1-WT-3xHA)

GTGAATCATGTTCACCTGTCCACTAACAAAAAAGGAAACAAAATGTTGCTTTCCCACAGGAAAAAGAATGTTTTTGAC
AAAACACTAAGCGTTATGTGACCAATGTATAATTGTCCTCATCATATAGTTTTATGCAAGCATAAATAAGTCTTTCTT
TTAGTTAAAACGCGTTACATAGTAAAGGTAACCCTGAATGAATTTTTATTTATATTTTTCTTCGCATATTTTAGGATC
TTGAAAAATTAACCACAAGTTAGAAGTTGATGAGATCAGGAATTCTAACAAAAGGTCATTGGTTAGGTTTATTTTGT
AGGTGCACGTTTGCAACTTTATTCCTTTTGTTTCAAAGTTATAGAGCTTTCATTTTCTCATTCTTCTCTAGTGTAGGT
CATTTCTCCTGCATATTAACCTACTAGAAAAGTTTACGAAGCTTCATAGGAACGAATAATCGATCACTTTAAAGGCTT
CCCGCACTTTAATCTTATCATCTACACCTACAAAAAAGCTCTACTTTGTCATTTATTTTCTACTTCATCTCTGTATGGC
ATCTATCTCTATATATATAAAAAAGCGCATCTGTTTATTATAAATGAATTCCAACAATCCTGATAATAATAATAGCAA
CAACATTAATAATAATAATAAGGATAAGGATATTGCCCCCAATAGCGATGTTCAGCTAGCTACCGTCTATACCAAGGC
AAAGTCTTACATTCCTCAGATTGAACAAGTTTATCAAGGCACAAATCCAAATATTCAAGAGGCAAAATTATTAGGGG
AGTTATTACAAGTTCTAGCAGAAGTTCCCAAGGGTACGCATTTATTTTGCGACCCCATACTTGAACCAATATCAATTT
TTTCCCTAACCATTTTCTCATTCAATGAAGAAGCTACGGCCACATGGCTGAAGAATCATTTTAATCCAATTCTCTCTG
TTTGTGATAAATGCATTCTAAATTTTGCGAGAGGGAAGTGCAAAATGTTACAACATTTTGCCATTCAAAGACATGTGC
CGCATGAACATGTAGCCAAGTTTAATGACATCGTATGCCAATGGAGGGTGGAGGCTGTTTTTCCTATTCTTAGAAATA
TTTCCGTTAACGATAATACGGGAATTAACATTACAAATGAAATTGAAACTGCAATGTACGAGTGTCTCTGTAATCCGC
ATATGTTGAGGCTTAACAAGCAACTTAAGGCTACTTTCGAAGCAATCTTCAAATTTTTCTACGACACAAAGCATCGTT
TGTTGGACGTTACGAATCCATTAAGCATAAAAACCTTTATATCAGGTGTCATATTTTGCTGGTGTGAAGGTTCAAAAG
AGGAAAATGAGTGGTCACGAGCTTTTCTAAAGGATTTATACTCAAGAAATTTCCATATAAATCTTTCTAACCTGACAC
CGGATATTATTGAGGAAGTTTATATCCATATTTTATTCTTACAGAATCCGGCAAACTGGACAGAGATTGTGGTATCTC
AATTCTGGTCGAGATTACTACCGGTTTTCAATTTATTTGATAAAGACGTCTTCATAGAGTACTTTCAGGTCCCAAAAA
ATGTGGAATCTTTAAAAAAAACATTCAAATTTCCGTTAGAGCCAATTTTTAAAATGTGGTACAATCACTTAAGCAAA
TCATATCATGATAAGCCACTAGACTTTTTGTTAAGAGGATTAACAATGTTCTTGAATAAATTTGGCTCCGAGTTTTGG
TCCAAAATTGAACCCTTCACATTCCATAGCATTTTGGATATAATTTTCAATAGAGATTCATTCCCAATCAAACTAATC
AAAATTCAAGACAATCCAATTGTTGAGCATCAAACAGAAGTTTACTTTCAACTTACCGGTTCTGTAACCGATTTACTT
TCTTGGACCTTACCATTTTATCATGCACTGTCACCATCCAAAAGAATTCAAATGGTTAGAAAGGTCTCAATGGCGTTC
CTGCGTATAATAGCCAATTATCCATCCCTAAAATCCATTCCCAAAGCTTGTTTAATGAATTCAGCAACTGCCTTGTTG
AGGGCTGTGTTGACTATTAAGGAAAATGAAAGGGCAATGCTTTATAAGAATGATGAATTTGAAACAGTACTGTTGAC
TAAAACAGATTCTAGAGCTTTGCTCAACAATCCTTTGATCCAAGACATAATAATTAGATCAGCTTCGAATCCTAACGA
CTTTTATCCTGGGCTAGGTGCTGCTTCTGCATCGGTGGCCACATCCACGATGATGGTATTGGCAGAATGTATCGATTT
TGATATACTGTTATTATGTCATCGAACGTTTAAGCTTTATTCAGGTAAACCAATTAGCGAAATTCCAATTTCAACCAA
CGTCTTGGAAAATGTTACAAATAAAATTGACCTGAGATCTTTTCATGACGGTCCATTGCTGGCCAAACAGCTGCTAGT
ATCTTTAAAAAATATTAATGGTCTTTTGATAGTACCGTCAAATACTGCCGTAGCTGAGGCACACAACGCATTGAATCA
GAAGTTTTTGCTGCTATCAACGAGATTAATGGAAAAATTTGCGGATATTTTGCCCGGTCAACTGTCTAAAATCCTTGC
GGATGAAGATGCATCGCAGGGTTTTTGGTCCTGCATATTCTCCTCTGATAAACACTTGTATCAAGCAGCTACTAATAT
TTTATACAATACATTTGACGTTGAGGGTAGACTGGAAGGTATCTTAGCGATTCTGAATAGTAATTTGACCGTTAACTT
GAAGAATATAAACGTGATGCTACAAAGATTAATCAATTGTGAATTCTATGAGCCTTGCCCCCGTGCTGTTCGTGTTTT
GATGGACGTAGTGAGTGCATTTGTAGACCCGATTTCTGGTGTGTTTGCTAATTTCCAAACTTTGAAGAGCCAAAATAC
CGAAAAGGAATTTTTAAAATTCTGGGAATCATGTTGGTTATTCTTGGATACGATTTATAAGTTTACTTTGAAATGGG
CTTCTAAATATGACTATTCTGAACTGGAAAATTTCACCAAAGATACTTTAGATCTAAGTCGCTCTTTGGTGGACTCCT
TTAGAGAATTTTCAGATATTCTTCATGACCAAACGAAAAATTTACTTCTCAACGTTTTAGAAACCTTTAAGAATATGT
TGTATTGGCTGAGATTGAGCGACGAAGTACTTTTGGAATCTTGTGTTAGATTAATTATAAGCACGTCTGATTTAGCAC
ATGAAAAGCATGTGAAAGTGGATGATTCCTTGGTAGAAATGATGGCCAAGTACGCCTCGAAGGCAAAGAGATTTTCA
AATAAGTTGACGGAGCAACAGGCGAGTGAGATTCTACAAAAGGCTAAAATATTCAATAAAGCACTCACGGAAGAGGT
CGCTACAGAAGCTGAAAACTACAGAAAAGAAAAGGAATTATCAAGACTTGGGAAGGTGATCGACTTAACGGATTCTG
TGCCGGCGTCACCATCCCTCTCGCCCTCTCTGTCTTCAACCATTGCTAGTTCTTCTGCTGAATCCAGGGCAGACTATTT
ACAAAGAAAAGCTTTATCTTCTTCAATAACTGGGAGACCAAGAGTAGCTCAGCCTAAAATAACATCTTTTGGTACGTT
CCAATCATCAGCTAATGCTAAACTGCATCGTACAAAACCTGTTAAGCCTTTATCCAAAATGGAATTAGCTAGGATGCA
ATTATTAAACAACAGAGTTGTTCATCCACCTAGTGCGCCAGCTTTCCATACAAAGAGTAGGGGTTTGTCTAACAAAAA
CGACGATAGCAGTAGTGAGGAGAGTGACAACGATATAGAAAGCGCTCGGGAGCTTTTTGCTATTGCTAAGGCAAAAG

GCAAAGGAATACAAACTGTCGATATTAATGGTAAGGTTGTCAAGAGACAAACTGCGGCTGAATTGGCTAAACAGGAG
TTAGAACATATGAGAAAGAGATTAAATGTTGATATGAACCCTCTATATGAAATCATATTGCAATGGGATTACACCAG
AAATAGTGAATATCCTGATGATGAGCCTATTGGTAATTATTCTGACGTAAAGGATTTCTTCAACTCTCCTGCTGACTA
CCAAAAAGTCATGAAGCCTTTACTACTTCTAGAATCTTGGCAAGGTCTATGTTCTTCACGTGATAGAGAGGACTATAA
ACCTTTCAGTATTATTGTAGGAAATCGAACTGCTGTGTCAGATTTCTATGATGTGTATGCCTCTGTCGCTAAACAAGT
CATACAGGATTGTGGTATTTCTGAATCTGACTTAATTGTAATGGCTTATTTACCCGATTTTCGACCAGACAAAAGGTT
GTCCAGTGATGATTTCAAGAAAGCACAACATACATGTTTAGCAAAGGTGAGAACATTGAAAAATACCAAAGGTGGTA
ATGTTGATGTCACATTACGAATTCATAGAAATCATTCTTTCAGTAAATTTTTGACACTAAGATCAGAGATATATTGTG
TGAAGGTTATGCAAATGACGACTATTGAAAGAGAGTATTCAACACTGGAAGGGCTGGAATACTATGACTTAGTTGGC
CAAATTCTACAGGCTAAACCATCGCCTCCCGTAAATGTGGATGCGGCAGAAATTGAAACAGTTAAGAAAAGCTATAAG
CTCAATACGTCGCAGGCTGAGGCCATCGTTAATTCGGTATCTAAAGAAGGTTTTTCTTTAATTCAGGGTCCACCTGGT
ACGGGTAAAACAAAGACTATACTGGGTATTATTGGTTACTTTTTGTCTACAAAGAATGCCTCATCTTCAAATGTCATC
AAAGTGCCACTAGAAAAGAACTCATCGAATACTGAACAACTGTTAAAAAAGCAAAAAATTCTAATCTGTGCCCCCAGT
AATGCCGCTGTGGACGAAATCTGCCTCAGACTGAAGAGTGGAGTGTATGACAAGCAGGGGCATCAATTTAAACCTCAA
TTGGTCCGTGTAGGTAGGTCAGACGTTGTAAACGTTGCAATTAAGGACCTTACTTTAGAAGAACTTGTGGATAAAAG
GATAGGTGAAAGGAATTACGAAATCAGAACAGATCCGGAACTAGAACGTAAATTTAATAATGCTGTGACTAAAAGAA
GAGAATTGAGAGGAAAATTAGACTCTGAGAGTGGTAATCCAGAAAGTCCAATGTCAACCGAAGATATCAGTAAATTA
CAACTTAAGATTAGGGAGCTAAGTAAAATAATCAATGAACTTGGCCGTGATAGGGACGAAATGCGTGAGAAGAATTC
CGTAAATTATAGGAATAGGGATTTGGATAGAAGGAATGCACAGGCCCATATCTTGGCGGTTAGTGACATCATATGTTC
GACTTTATCTGGTTCAGCCCATGATGTACTCGCAACAATGGGTATTAAATTTGATACGGTCATTATAGATGAAGCATG
TCAATGTACCGAACTTTCTTCAATTATTCCTCTCAGGTATGGAGGGAAGCGTTGTATTATGGTTGGTGATCCTAACCA
ACTACCACCAACTGTTCTTTCAGGTGCAGCAAGCAATTTTAAATACAATCAGTCACTGTTCGTTAGAATGGAGAAAAA
TAGCTCCCCATACTTGCTGGATGTTCAATACCGTATGCATCCTTCCATCAGTAAATTCCCCTCTTCTGAATTTTATCAG
GGACGACTAAAAGACGGCCCTGGTATGGATATCTTGAACAAAAGACCCTGGCACCAGCTTGAGCCATTAGCTCCCTAC
AAATTTTTTGACATCATTAGTGGTAGGCAAGAACAAAATGCGAAGACTATGTCTTATACAAATATGGAGGAAATACG
TGTTGCCATTGAGTTAGTTGACTACCTTTTCAGAAAATTCGATAATAAAATTGATTTCACGGGAAAAATCGGTATAAT
TTCACCTTATAGAGAGCAAATGCAAAAAATGCGTAAAGAGTTTGCCCGCTATTTTGGAGGCATGATCAACAAATCAAT
CGATTTCAATACAATCGACGGTTTCCAAGGTCAAGAAAAAGAAATCATCTTGATATCGTGTGTCCGTGCCGATGATAC
GAAATCTAGTGTCGGTTTTTTAAAGGACTTTCGTCGTATGAATGTGGCTCTGACTAGAGCCAAGACAAGTATTTGGGT
TTTGGGACATCAAAGGTCTCTGGCTAAAAGCAAGTTATGGAGAGATTTGATTGAAGATGCAAAAGATAGAAGTTGTC
TTGCATATGCTTGCTCGGGTTTCCTGGATCCCAGAAATAATAGAGCTCAAAGTATTTTAAGAAAGTTCAATGTACCCG
TACCATCAGAGCAAGAAGATGATTATAAGCTACCCATGGAATATATTACGCAAGGCCCCGATGAGGTAAAATCCAATA
AGGACACAAAGAAGAGAAGAGTTGTCGATGAAGGTGAAGAGGCAGATAAAGCTGTAAAGAAAAAGAAGAAAGAGAA
GAAGAAGGAAAAGAAAAAGTCCAAGGCAGATGATAAGAAAAAGAATAACAAGAAGGCAGAATCGCCCAGCACCTCAT
CTGGTACAAAGAAAAAGTCATCCATTTTTGGTGGAATGAGTGTACCGAGTGCCGTAGTTCCTAAAACTTTTCCTGATG
TGGACAGCAACAAAAAGGCAGCAGCAGTTGTTGGTAAAAAGAAAAATAACAAACACGTTTGTTTTTCGGATGATGTT
AGTTTCATACCAAGAAATGACGAACCTGAAATTAAAGTGACCAGAAGTTTATCCTCTGTATTGAAGGAAAAGCAGTT
AGGTTTGAAGGAGACGAGAACAATTTCTCCGCCTGAAATCAGTAATAATGAGGATGATGATGACGAAGATGATTATA
CGCCGTCAATCTCCGATTCTTCATTAATGAAAAGTGAAGCAAATGGTCGTAACAACAGGGTGGCTTCTCATAACCAAA
ATTTCAGCGCTAGTATATACGACGATCCACAAGTTTCGCAAGCAAAGCAGACGCAGGTACCTGCTGCTATTACAAAAC
ACAGAAGTTCAAATAGTGTTTTATCCGGTGGTTCTTCGAGGATCCTGACTGCTTCCGATTACGGTGAACCCAATCAAA
ATGGACAGAATGGAGCAAATAGGACCCTCTCTCAACATGTGGGTAATGCGAATCAATACTCAACGGCTCCTGTGGGTA
CAGGCGAACTGCATGAAACTCTGCCCGCCCATCCGCAAGACAGCTATCCTGCTGAGGCAGAAGATCCATATGATTTAA
ACCCCCATCCGCAACCACAATCTTCAGCATTCAAAGGACCTGGTAGTGGGCCCACAGGCACACGCAACAGTTCGAGACG
GAATGCTTCATCTAGCCCATTTATCCCAAAAAAAAGAAAGCCTAGATCATACCCATACGATGTTCCTGACTATGCGGG
TATCCCTATGACGTCCCGGACTATGCAGGATCCTATCCATATGACGTTCCAGATTACGCTTAATCTTGAAGTAAAAGT
GTTAGGAATTATACCTGCATATATATTGGTGTATACATTTATATTGAATTTAATAGCAACATTATTTTTCATTATTTT
TCATTATTTCAGCAAAAAGAACATTTTTTTTCAATTTGACGTGAAAGAGTTTGACTATTACGTGGGCATACAATGTTT
AAATTTACATTATCCTCATGGCTACTCTAGCTATTTGCATTTCATTCAACTTTTTTTGATATGGCATCATTTTTATTA
GCGTTTACTACACTTTCAGAAATGCCAATCTTTGGTGGCGATTTGCTAGCAACCAAGAATGAGGGAGTCGAGGGCTGT
TTTAAATTTGAACTCGCGTTCGACCCTTTATGGCTTTTGTGTATATCAATTGGTTTAGAAGGTAGTGCGCTTGACCCA
TCTTGTGCAGTGGTACTTTCGCTATGGAGTTCTTCACAAGCTTTCGAAAGTACGTCCCTTTCGTTTGATATTAAAGTT
GTCGTGAGCTTTTCATCAGAAGAGTCAAGAAGGTTATTTAGGTACTGTATTGTAGAATTGATCATTTCTGACATTTCA
TTTGGAGGTATTTTGGAGTTGCTGCTGCTCCGAAGTTTGGTTA

Annex D – Sen1 phoshorylated peptides identified in Scaffold Analysis

Table 5. – Sen1 phosphorylated peptides identified in mass spectrometry analysis

Biological sample	Treatment	Protein accession number	Protein identification probability	# of unique peptides	# of unique spectra	Peptide sequence	Modifications identified by spectrum	Peptide start index	Peptide stop index
S1	Proteasome Inh.	SEN1_YEAST	100,0%	177	266	GKLDSESGNPESPMSTEDISKLQLK	Phospho (+80)	1494	1518
	Proteasome Inh.	SEN1_YEAST	100,0%	177	266	GKLDSESGNPESPMSTEDISKLQLK	Phospho (+80), Oxidation (+16)	1494	1518
	Proteasome Inh.	SEN1_YEAST	100,0%	177	266	GKLDSESGNPESPMSTEDISKLQLK	Phospho (+80), Oxidation (+16)	1494	1518
	Proteasome Inh.	SEN1_YEAST	100,0%	177	266	GKLDSESGNPESPMSTEDISKLQLK	Phospho (+80)	1494	1518
	Proteasome Inh.	SEN1_YEAST	100,0%	177	266	GKLDSESGNPESPMSTEDISKLQLK	Phospho (+80)	1494	1518
S2	Proteasome Inh + rapa	SEN1_YEAST	100,0%	150	222	GKLDSESGNPESPMSTEDISK	Phospho (+80)	1494	1514
	Proteasome Inh + rapa	SEN1_YEAST	100,0%	150	222	GKLDSESGNPESPMSTEDISK	Phospho (+80)	1494	1514
	Proteasome Inh + rapa	SEN1_YEAST	100,0%	150	222	GKLDSESGNPESPMSTEDISKLQLK	Phospho (+80), Oxidation (+16)	1494	1518
	Proteasome Inh + rapa	SEN1_YEAST	100,0%	150	222	GKLDSESGNPESPMSTEDISKLQLK	Phospho (+80), Oxidation (+16)	1494	1518
	Proteasome Inh + rapa	SEN1_YEAST	100,0%	150	222	GKLDSESGNPESPMSTEDISKLQLK	Phospho (+80)	1494	1518
	Proteasome Inh + rapa	SEN1_YEAST	100,0%	150	222	GKLDSESGNPESPMSTEDISKLQLK	Phospho (+80), Oxidation (+16)	1494	1518
	Proteasome Inh + rapa	SEN1_YEAST	100,0%	150	222	GKLDSESGNPESPMSTEDISKLQLK	Phospho (+80)	1494	1518
	Proteasome Inh + rapa	SEN1_YEAST	100,0%	150	222	GKLDSESGNPESPMSTEDISKLQLK	Phospho (+80)	1494	1518