Functional characterization of the N-terminal KH domains of the vigilin-protein Scp160

Dissertation

der Mathematisch-Naturwissenschaftlichen Fakultät

der Eberhard Karls Universität Tübingen

zur Erlangung des Grades eines

Doktors der Naturwissenschaften

(Dr. rer. nat.)

vorgelegt von

Jonathan Feicht

aus Saulgau jetzt Bad Saulgau

Tübingen

2025

Gedruckt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Eberhard Karls Universität Tübingen.

Tag der mündlichen Qualifikation:	03.06.2025
Dekan:	Prof. Dr. Thilo Stehle
1. Berichterstatter/-in:	Prof. Dr. Ralf-Peter Jansen
2. Berichterstatter/-in:	Prof. Dr. Thorsten Stafforst

Declaration

This thesis describes my work conducted in the laboratory of Prof. Ralf-Peter Jansen at Interfakultäres Institut für Biochemie (IFIB), University of Tübingen, Tübingen, Germany from September 2019 to February 2024. The work was co-supervised by Prof. Dr. Thorsten Stafforst also at IFIB, University of Tübingen, Tübingen, Germany.

I declare that this thesis is the product of my own work. The parts that have been published or where other sources have been used were cited accordingly.

Contents

Abstract

The vigilin proteins represent a family of RNA-binding proteins that are characterized by their high degree of conservation. These proteins possess a unique feature, namely the presence of up to 15 KH domains, which is a distinctive attribute among RNA-binding proteins. The vigilin proteins have been associated with numerous cellular processes in various organisms, including translation regulation, ploidy maintenance, mRNA localization and stabilization, ribosome association, and numerous others.

The present study was motivated by the observation that Scp160 exerts a substantial influence on the polyglutamine aggregation and re-use of t-RNA. The central objective of this study was to investigate the impact of KH domains on the scope of their influence on RNA, on interacting proteins and finally on each other. To address these questions, KH domains were deleted using a CRISPR/Cas9 approach and the impact of KH domain absence was subsequently examined. The strains lacking specific KH domains were analyzed for the localization of Scp160 in the cell. Scp160, which normally localizes at the ER, showed delocalization for deletion of specific KH domains. This result suggests an impairment in the protein's function. For further investigation, analysis of the ploidy profiles was performed, displaying a change in ploidy for certain KH deletions. Furthermore, an investigation was conducted to ascertain the emulsifying properties of Scp160 in the absence of specific KH domains. The strains that developed an emulsifying effect and also showed changes in their ploidy and localization, of Scp160 were further analyzed. Mass spectrometry revealed a relatively minor impact of KH domain deletion on the pattern of proteins that interact with Scp160. To further investigate the impact of deletion of KH domains on RNA interaction, RNA sequencing and CRAC were performed, which indicated that RNA binding is severely impacted by the absence of certain KH domains but not for other deletions with an equivalent numbers of KH domains. This result suggests that there is a difference in the importance of KH domains within Scp160. The loss of the middle KH domains (KH domains four and five) resulted in the forementioned phenotypes, including a severe loss of RNA binding. In contrast, the deletion of the first two KH domains did not alter the function of Scp160 and therefore didn´t exhibit any phenotype.

The findings of this study have demonstrated that both the position and the number of KH domains are of significant importance for the functionality and localization of Scp160. Consequently, future research on this protein should concentrate on analyzing the function of individual KH domains and their interaction with RNA, proteins and each other.

Zusammenfassung

Die Vigilin-Proteine stellen eine Familie RNA-bindender Proteine dar, die sich durch einen hohen Konservierungsgrad auszeichnen. Diese Proteine besitzen ein einzigartiges Merkmal: sie bestehen aus bis zu 15 KH-Domänen, welche ein charakteristisches Merkmal von RNA-bindenden Proteinen darstellen. Die Vigilin-Proteine wurden mit zahlreichen zellulären Prozessen in verschiedenen Organismen in Verbindung gebracht, darunter Stressgranulatbildung, Translationsregulation, Aufrechterhaltung der Ploidie, Pheromonerkennung, Telomer-„Silencing“, mRNA-Lokalisierung und -Stabilisierung, P-Körper-Lokalisierung, Ribosomenassoziation und zahlreiche andere.

Die vorliegende Studie entstand durch die Beobachtung, dass Scp160 einen erheblichen Einfluss auf die Wiederverwendung von t-RNA und auf die Aggregation von Polyglutamin ausübt. Das zentrale Ziel dieser Studie bestand darin, den Einfluss von KH-Domänen auf RNA und interagierende Proteine sowie ihren Einfluss untereinander zu untersuchen. Um diese Fragen zu beantworten, wurden KH-Domänen mit Hilfe von CRISPR/Cas9 deletiert und anschließend die Auswirkung der fehlenden KH-Domänen untersucht. Die Stämme, denen spezifische KH-Domänen fehlen, wurden auf die Lokalisierung von Scp160 in der Zelle untersucht. Scp160, welches normalerweise am ER lokalisiert ist, zeigte eine Delokalisierung im Falle der Deletion spezifischer KH-Domänen. Dieses Ergebnis deutet auf eine Beeinträchtigung der Funktion des Proteins hin. Zur weiteren Untersuchung wurde eine Analyse des Ploidiegrades durchgeführt, welche eine Änderung der Ploidie für bestimmte KH-Deletionen zeigte. Darüber hinaus wurde ein Experiment durchgeführt, um die emulgierenden Eigenschaften von Scp160 in Abwesenheit spezifischer KH-Domänen zu ermitteln. Die Stämme, welche eine emulgierende Wirkung entwickelten und auch Veränderungen in ihrer Ploidie und Lokalisierung zum ER zeigten, wurden weitergehend untersucht. Massenspektrometrie ergab einen geringen Einfluss der KH-Domänen-Deletion auf die Protein-Protein-Interaktion von Scp160. Um den Einfluss der Deletion von KH-Domänen auf die RNA-Interaktion weiter zu untersuchen, wurden RNA-Sequenzierung und CRAC durchgeführt. Die Ergebnisse deuteten darauf hin, dass die RNA-Bindung durch das Fehlen bestimmter KH-Domänen stark beeinträchtigt wird, bei anderen Deletionen mit einer äquivalenten Anzahl von KH-Domänen jedoch nicht. Dies legt nahe, dass es einen Unterschied in der Bedeutung von KH-Domänen innerhalb von Scp160 gibt.

Die Ergebnisse dieser Studie haben gezeigt, dass sowohl die Position als auch die Anzahl der KH-Domänen von entscheidender Bedeutung für die Funktionalität und Lokalisierung von Scp160 sind. Daher sollte sich die zukünftige Forschung zu diesem Protein auf die Analyse der Funktion einzelner KH-Domänen und ihrer Interaktion mit RNA, Proteinen und untereinander konzentrieren.

Introduction

Translation provides the basis for protein synthesis

As two of the most fundamental processes in any living organism, transcription and translation are universal steps in the production of all required proteins in a cell. The importance of translation can be measured by its high energy consumption. With a 50:1 ratio of RNA to DNA, divided into 80% rRNA (representing ~60% of the total transcription in the cell), 15% tRNA and 5% mRNA and an estimated 200,000 ribosomes per yeast cell, a high level of energy must be used to maintain the process of translation. To sustain the high level of ribosomes alone, the yeast cell must produce 2,000 new ribosomes per minute (Warner, 1999). A highly regulated translation machinery is therefore essential. Translation regulation is critical for the ability of cells to adapt to stress situations, cell development and differentiation. These regulatory elements have been found in almost every step of the eukaryotic translation machinery but are most concentrated in the initiation step (Jackson et al., 2010). Inactivation of eukaryotic initiation factors (eIFs) reduces translation for most mRNAs under stress conditions such as starvation (Sonenberg and Hinnebusch, 2009). Localized translation allows specific changes in cell morphology through spatially restricted protein synthesis, for example during early stages of embryogenesis (Cho et al., 2005). In general, translation regulation can be divided into two control domains: global control of protein synthesis and mRNA-specific translation regulation. Global control includes RNA-independent levels of regulation during normal growth as well as under stress conditions. In order to adjust to changes in their environment, yeast cells have a wide range of possible gene expression adaptations. This may be in response to self-generated problems such as nutrient deprivation due to growth (Paz and Choder, 2001; Svitkin et al., 2005) or in response to external factors, which include a wide range of different stresses (Smirnova et al., 2005).

The process of translation consists of several steps

Translation can be divided into three distinct phases: initiation, elongation and termination (Figure 1).

Initiation: First, the PIC (pre-initiation complex) is formed. The 40S small subunit of the ribosome binds to the eukaryotic initiation factors eIF1, -1A, -3, -5 and the TC (ternary complex), which consists of eIF2-GTP and initiator tRNA. The pre-initiation complex then recognizes the polyadenylated mRNA, which has been preactivated by eIF4F. Propelled by eIF5, the PIC begins to search for the 5' UTR of the

mRNA for the anticodon of the initiator tRNA (Met). Upon detection of the start codon, the 60S ribosomal subunit is recruited by eIF5B, resulting in the functional 80S ribosome (Hinnebusch, 2014).

Elongation: GTP hydrolysis and eukaryotic elongation factors (eEF1A, eEF2, etc.) allow the ribosome to move along the mRNA. The movement is codon-based, where a codon consists of a specific nucleotide triplet. Of the 64 possible codon combinations, 61 code for canonical amino acids. The codons of the mRNAs bind complementary anticodons of the charged tRNAs. These transfer RNAs are charged with an amino acid corresponding to the mRNA codon, resulting in the growing peptide chain. Codon anticodon binding occurs at the A site of the large subunit of the ribosome. The amino acid is attached to the growing peptide chain after the tRNA moves to the P site. The uncharged tRNA then moves to the E site. Here it is released from the ribosome.

Termination: This process continues until the ribosome encounters a stop codon. At this point, the eukaryotic release factor rEF1 is bound to the A site along with other release factors. The nascent protein chain is released for further processing. The ribosome is then dissociated and recycled into 40S and 60S units by Rli1 in *Saccharomyces cerevisiae* (Schuller and Green 2018).

A notable problem of common translation elongation is ribosome stalling. This translation interruption can have several causes: slow kinetics in peptidyl transfer, poor A-site occupancy due to insufficiently expressed tRNAs or poor tRNA aminoacylation, a queue of successive tRNA codon pairs that is suboptimal relative to synonymous subsequent pairs, and mRNA secondary structures such as stem loops or pseudoknot (Schuller and Green, 2018).

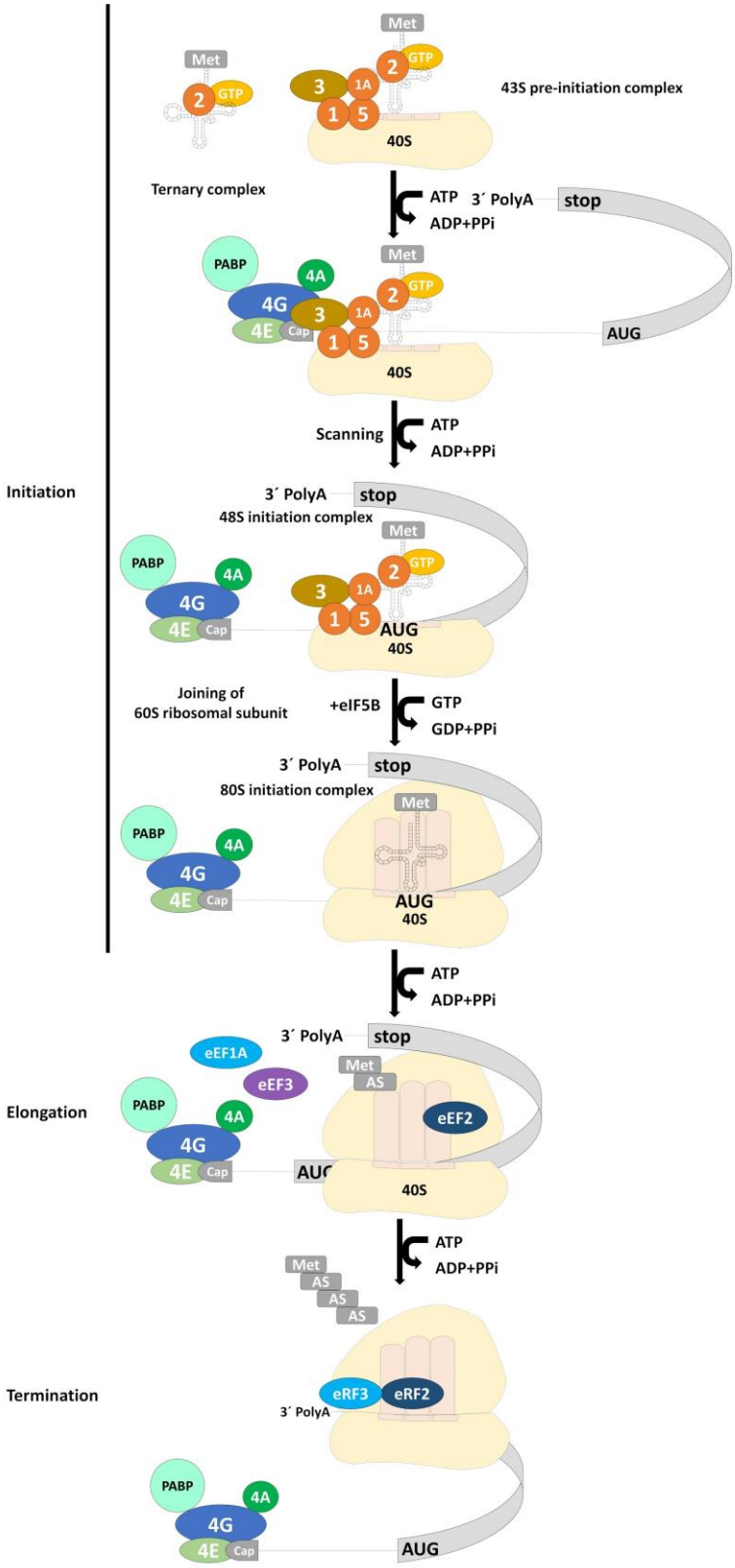


Figure 1: The three steps of translation: initiation, elongation and termination. Initiation: the ternary complex binds to the 40S subunit and forms the 43S pre-initiation complex; this complex associates with the mRNA. This 48S pre- initiation complex scans the RNA for the initiator "AUG". The 60S ribosomal subunit joins afterwards; Elongation: see text; Termination: After recognition of the stop codon, translation is terminated. Release of the polypeptide chain; Eukaryotic initiation factors (eIFs) are portrayed as colored, numbered circles; PABP: poly(A) binding protein; PPi: pyrophosphate; Illustration modified from "Molecular mechanisms of translational control" (Gebauer and Hentze, 2004) and "Translational control by the multi-KH domain protein Scp160" (Schreck 2010).

ER localized translation is dependent on different factors

Apart from differences in early translation at the mechanistic level (see below), the differences between cytosolic and ER translation are minor. There is evidence that ER translation is more efficient than its cytosolic counterpart. The steady-state mRNA load of ribosomes at the ER is higher than that of cytosolic ribosomes (Reid and Nicchitta, 2012). Another difference is the translation of specific mRNAs in relation to cytosolic and ER-bound ribosomes: mRNAs for cytosolic proteins could be found to be translated at the ER, but not vice versa (Reid and Nicchitta, 2012). Further investigation into the intricacies of translation revealed discrepancies in the regulation of translation in the cytosol and at the ER. A reporter mRNA was shown to be translated by a higher number of ribosomes at the ER compared to the cytosol, indicating either slower elongation or faster translation initiation (Voigt et al., 2017).

Another ongoing investigation suggests that the ribosome composition in terms of different sets of components may lead to specialized ribosomes. These specialized ribosomes can focus on functional aspects, like translation initiation, elongation speed, fidelity and selectivity of mRNA or they can be separated by their localization ribosomes (Genuth and Barna, 2018). A difference in cytosolic and ER-localized ribosomes was found in mouse embryonic stem cells (mESCs) were the ribosome associated proteins (RAP) PKM2, a glycolytic enzyme, was found for the ER localized ribosomes but not for cytosolic ribosomes (Genuth and Barna, 2018). Another example for heterogeneity of ribosomes was in the case of the ribosomal RNA extension segment ES27L. Cryo-electron tomography shows a conformational difference for ER-bound and cytosolic, suggesting a structural adaptation for ribosomes depending on their location (Pfeffer et al., 2012).

While the cytosol contains a largely soluble fraction of proteins, mRNAs translated at the endoplasmic reticulum encode among other proteins that are integrated into membranes. The ER-Golgi pathway is particularly important for membrane-targeted proteins such as transmembrane and peripheral membrane proteins, including receptor proteins, channel proteins and pumps, but also for soluble proteins destined for secretion (Viotti et al., 2016). Protein secretion and protein integration into membranes are presumably catalyzed by the same ER machinery (Figure 2) (Caro and Palade, 1964; Rapoport, 2007; Reid and Nicchitta, 2015; Walter and Blobel, 1981). The synthesis of a targeting signal initiates the secretory pathway. The signal recognition particle (SRP) can then bind to the nascent peptide chain, resulting in the RNC-SRP complex (Walter and Blobel, 1981). Under certain circumstances, elongation can be modified by factors such as SRP. In 1980, the composition of the SRP or signal recognition particle was identified as consisting of SRP RNA and six proteins, which were

named SRP9, SRP14, SRP19, SRP54, SRP68, SRP72 according to their size of 72,000, 68,000, 54,000, 19,000, 14,000 and 9,000 daltons (Mutka and Walter, 2001; Walter and Blobel, 1981).

SRP binds ribosomes independently of the presence of a signal sequence. However, the presence of this signal, which is an N-terminal hydrophobic sequence of a specific newly synthesized protein, positively influences the binding affinity of SRP. The region for the SRP 5' Alu domain overlaps with the binding site of the elongation factor between the two ribosomal subunits. This allows SRPs to modulate translational elongation after binding the nascent chains of membrane/secretory proteins. Translation is halted to allow transport of the newly formed SRP-RNC to the ER (Egea et al., 2005; Halic and Beckmann, 2005). The complex can then be translocated to the ER membrane (Zhang and Shan, 2014). It should be noted that targeting of RNCs to the ER in general can occasionally occur without the SRP (Neuhof et al., 1998). In yeast, SRP-independent pathways are known. Besides the homologous yeast guided entry of tail-anchored proteins (GET) another pathway constitutes of the SND proteins. Deletion of these proteins depleted the localization of proteins with more downstream transmembrane domain (TMDs). This pathway is known as SRP-independent targeting, or SND for short. The shifting of the TMD seems to alter the dependency of the different pathway which directs substrates to the to the endoplasmic reticulum (Aviram et al., 2016). Targeting of mRNAs to the ER is still not fully understood. There have been numerous publications investigating mRNA localization, indicating multiple ER localization mechanisms, but the full extent of mRNA localization pathways is still unknown (Costa et al., 2018; Lakkaraju et al., 2007; Mutka and Walter, 2001).

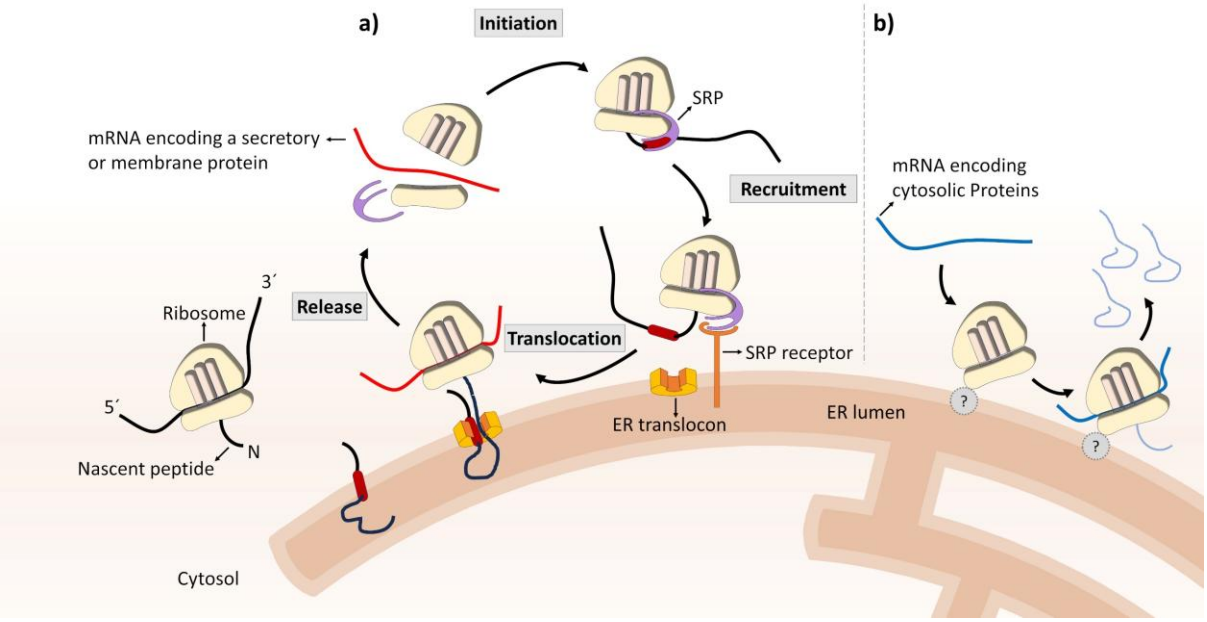


Figure 2: Translational compartmentalization; a) Model of translational compartmentalization; starting of translation initiation (cytosol). Ribosomes remain in the cytosol when translating mRNAs encoding cytosolic proteins (lack of a topogenic ER-targeting signal); ribosome are targeted to the ER (co-translationally by the SRP) when translating mRNAs that encode transmembrane domains (integral membrane and secretory proteins) or proteins containing a signal peptide (both indicated as red barrel); topogenic signals serve as a targeting signal to the ER; membrane and secretory proteins are translocated through/into the ER membrane after docking onto the ER translocon; subunits of the ribosome are recycled into the cytosol upon completion of protein synthesis; b) Model of a pan-transcriptomic role for the ER in mRNA translation; translation initiation can occur right on the ER; mRNAs encoding cytosolic proteins also can be translated at the ER (How this association comes to pass is unknown), hence a major part of the proteome can be translated at the ER; ER: endoplasmic reticulum; SRP: signal recognition particle; illustration modified from "Diversity and selectivity in mRNA translation on the endoplasmic reticulum; figure 1" (Reid and Nicchitta, 2015).

The mRNAs encoding the secretory proteins are targeted to the ER via an N-terminal signal peptide that mediates SRP binding. However, there are numerous other regulatory elements that influence the ER targeting of mRNAs. A characteristic feature of ER-targeted mRNAs is their increased uracil content in the coding region (Prilusky and Bibi 2009). In addition, the adenine content in the signal sequence coding regions (SSCRs) is reduced. Taking together, these findings suggest that mRNAs of secreted/membrane proteins (mSMPs) can be distinguished from other mRNAs by their nucleobase composition (Palazzo et al., 2007). More recently, a *cis*-acting sequence motif in RNAs (SECReTE; secretion-enhancing cis regulatory targeting element) has been found to enhance mSMP localization to the ER (Cohen-Zontag et al., 2019). In addition to localization, it appears to enhance RNA stability and secretome protein synthesis/secretion. SECReTE consists of approximately ten nucleotide triplet repeats with an enrichment of pyrimidines (C or U) at every third base. Interestingly, in an unbiased attempt to find the motif in cell wall proteins using the motif search tool MEME, a general U/C enrichment was found not only for the third base (Cohen-Zontag et al., 2019). Similarly, codons with U in the second position correspond to hydrophobic amino acids (Koonin and Novozhilov, 2009). Taken

together, the C/U rich motifs of RNAs often result in hydrophobic proteins, implying their association with the cell wall and membrane.

Protein levels in subcellular compartments are regulated by localized translation

For a significant number of proteins, the site of translation affects the localization of the functional protein. In *Drosophila* oocytes, almost 10% of RNAs are localized. These RNAs include important examples of localized RNAs such as *bicoid*, *oskar* and *gurken*. Localization can also be restricted to specific cell organelles. In addition to the ER, localized translation have also been shown in mitochondria. The mitochondrial TOM complex, localized at the outer membrane, represents a translation platform. In yeast, the mRNA was able to interact directly with the mitochondrial-associated Puf3. Together with the nascent polypeptide-associated complex (NAC) and OM14, a mitochondrial receptor for cytosolic ribosomes, sufficient anchoring is achieved for translation and subsequent import of the protein by the TOM complex.

Zipcodes are another motif associated with mRNA targeting. In this case, it's not a conserved sequence but a structural element whose secondary structure leads to a similar kind of stem loop. These are then recognized by specific RNA binding proteins (RBPs). An example of this type of mRNA targeting is the *ASH1* mRNA. The RBP She2 is able to recognize the four zipcode elements of the *ASH1* RNA and facilitates smooth mRNA targeting (Olivier et al., 2005).

In addition to the RNA sequence-based localization markers, the RBPs mentioned above are the most important mechanisms associated with mRNA targeting. A vast localization network regulating mRNA-protein interactions has been identified. This network controls the spatial distribution of RNA expression within cells (Hogan et al., 2008; Oeffinger et al., 2007). Therefore, the mRNA localization of ER-bound RNAs is most likely to be by active transport rather than passive diffusion (Brangwynne et al., 2009).

RNA binding Proteins contain various structural motifs

RNA binding proteins (RBPs) are a group of proteins responsible for post-translational regulation throughout the RNA life cycle. The association of RBPs with nascent RNA results in the formation of ribonucleoprotein (RNP) complexes. The binding of RBPs to RNAs is often, but not always, dependent on structurally conserved binding domains such as RNA recognition motif (RRM), double-stranded RNA binding (dsRBD), DEAD-box helicase domains (Linder 1989) and KH domains (Siomi et al., 1993). In the following, I will briefly characterize these major RNA-binding domains.

RRM: The RRM is the best studied and most abundant binding domain in this list. Estimated to be present in 1% of all human proteins (Cléry and Allain, 2012), the RRM contains approximately 90 amino acids arranged in a β1α1β2β3α2β4 composition. Between β1 and β3, the two α-helices form the RNA-binding motif RNP1 and RNP2 against an antiparallel β-sheet (Cléry and Allain, 2012). Binding to RNA occurs for single-stranded RNA (ssRNA) through subsequent stacking interactions and hydrogen bonding with the RNP motif (Auweter et al., 2006).

dsRBD: As the name suggests, the dsRBD only binds double-stranded RNA (dsRNA) and consists of approximately 65-70 amino acids (Masliah et al., 2013). dsRNAs are found in proteins responsible for RNAi, viral protection and cellular transport. The dsRBD is often found in tandem repeats or together with other RNA binding domains (Cléry and Allain, 2012; Ranji et al., 2011). The dsRBD is composed as α1β1β2β3α2, with the two α-helices flanking the antiparallel β-sheets (Cléry and Allain, 2012; Masliah et al., 2013). The A-formed RNA recognition helix is specifically recognized by the dsRBD with low affinity (Cléry and Allain, 2012).

Helicase domains: The unwinding of polynucleotides is contingent upon the binding of both DNA and dsRNA. In comparison to the other RNA-binding domains, which typically comprise around 100 amino acids, the helicase domains are relatively large, comprising approximately 350-400 amino acids (Gai et al., 2004). These are once more organized into two "recombinase A (RecA)-like" subdomains, which both include an ATP-catalytic core, a nucleic acid-binding region, and smaller subdomains responsible for coordination (Gai et al., 2004; Kainov et al., 2008). The binding of RNA is surrounded by the recA-like domains, or, in the case of multimeric helicases, pulled through the center of the ring (Jankowsky, 2011; Kainov et al., 2008; Linder and Jankowsky, 2011). The helicase domain interacts with multiple nucleotides simultaneously. DEAD-box helicases have the capacity to cover at least five single-stranded or base-paired nucleotides simultaneously (Jiang et al., 2011; Linder and Jankowsky, 2011; Weir et al., 2010).

Heterogeneous nuclear ribonucleoprotein particle (hnRNP) K homologue domains represent the RNA binding domain of vigilin proteins

In addition to the RRM and dsRBD, the heterogeneous nuclear ribonucleoprotein particle (hnRNP) K homologue domain (KH domain) is one of the most prevalent RNA recognition domains utilized for interactions with RNAs (Corley et al., 2020). The designation "KH domain" is derived from the eukaryotic heterogeneous nuclear ribonucleoprotein K, which is involved in pre-mRNA processing in the nucleus. KH domains are present in all three domains of life (Archaea, Bacteria, and Eukarya) (Siomi et al., 1993), suggesting that they may have originated early in the evolutionary process (Grishin, 2001). Examples of bacterial or archaeal KH domain-containing proteins include the phosphate-dependent exonuclease PNP in *Escherichia coli (E. coli)* and the archaeal ribosomal protein S3 (Siomi et al., 1993). KH domains can be classified into two main subgroups, Type I and Type II. Both subgroups share a common structural motif comprising two beta sheets separated by two alpha helices (βααβ), with an additional αβ component located at the N- or C-terminus, depending on the subgroup. In both subgroups, the antiparallel beta sheets are arranged in opposition to the alpha helices (Figure 3). The organization of the type I subgroup is of a $\beta_1\alpha_1\alpha 2\beta_2\beta_3\alpha_3$composition while the type II consists of an $\alpha_1'\beta_1'\beta_1\alpha_1\alpha_2\beta_2$ structure (Grishin, 2001). A conserved glycine-X-X-glycine segment is located between the α-helices (A and B). This segment is only absent in a few known KH-domain-containing proteins, including KH domains 1, 3-7, and 13 of the yeast vigilin protein Scp160 (M. A. Brykailo et al., 2007). It has been proposed that these GXXG motifs are responsible for binding RNA and ssDNA with a relatively low affinity (Liu et al., 2001; Valverde et al., 2008). The absence of the GXXG motif has been observed in KH-domain-containing proteins, including vigilin proteins. In contrast to the classical GXXG motifs, these are referred to as "diverged." The number and position of diverged and classical KH domains exhibit variation between the vigilin proteins of different organism (Figure 4 a+b). The significance of diverged KH domains was assessed. The ability of Scp160 variants, in which the final two KH domains were replaced with alternative KH domains, to rescue a lethal *scp160Δ/eap1Δ* double deletion was evaluated. The results demonstrated that all of the swapped KH domains were capable of partially saving the double deletion mutant. This result indicates that both diverged and conserved motifs are necessary for the proper functionality of the protein (M. A. Brykailo et al., 2007; Li, 2003).

The number of KH domains present in a protein can vary considerably, from a single domain (as observed in PNP) to three domains in hnRNP K (Siomi et al., 1993) and up to 15 domains in human HDLBP (Xia et al., 1993). These individual KH domains can serve distinct functional roles (Vollbrandt et al., 2004). An illustrative example of this particular function of individual KH domains can be observed

in KH domains 13 and 14 of the yeast vigilin homologue Scp160. These have been demonstrated to be involved in RNA binding and polysome binding (Baum et al., 2004; Gelin-Licht et al., 2012a; Hirschmann et al., 2014; Vollbrandt et al., 2004).

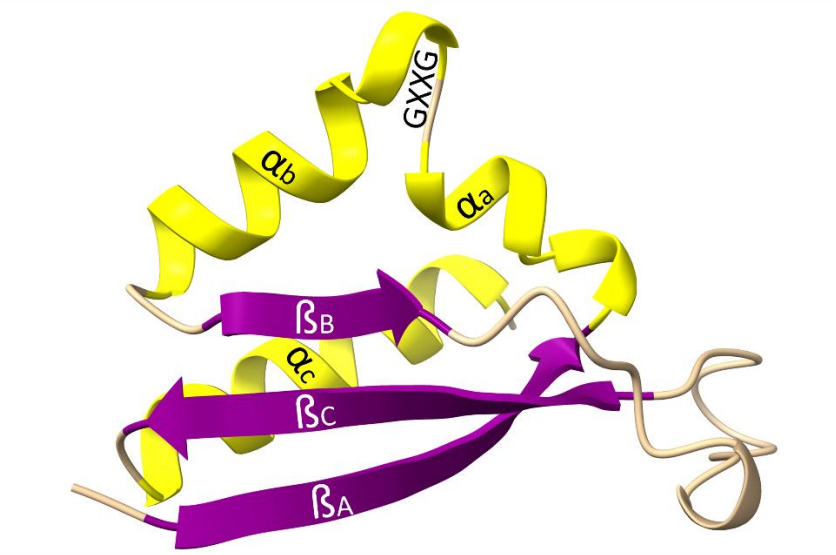


Figure 3: AlphaFold2 prediction of the second KH-domain of Saccharomyces cerevisiae vigilin protein Scp160; α: alpha-helices; ß: beta-sheets; upper- and lower-case a-c: enumeration of alpha helices and beta sheets.

Vigilin proteins are a family of multi-KH domain RNA-binding proteins

As previously stated, the vigilin protein family is a member of the KH-domain-containing group of RNA-binding proteins. The human family member was identified as a high-density lipoprotein (HDL) binding protein, which led to the designation of the human vigilin protein as "HDLBP," or HDL binding protein (McKnight et al., 1992). The first vigilin protein to be identified was that of the chicken, which was discovered in 1987. The protein was subsequently named "vigilin" in 1993, following the identification of its N-terminal conserved valine-isoleucine-glycine motif (Graham und Oram, 1987; Siomi et al., 1993).

In addition to the previously mentioned human and chicken proteins, several other vigilin homologs have been identified (Figure 4 a)). The namesake of HDLBP in *Mus musculus* (Graham und Oram, 1987), Vigilin in *Gallus gallus* (Schmidt et al., 1992), hdlbp from *Xenopus laevis* (Dodson and Shapiro, 1997), DDP1 in *Drosophila melanogaster* (Cortes, 1999), and vgln-1 in *Caenorhabditis elegans* (Weber et al., 1997), hdlbp was identified in *Danio rerio* (Chen et al., 2003), and in 2010, vgl1 was identified in

Schizosaccharomyces pombe (Wen et al., 2010). Additionally, Scp160 was identified in *Saccharomyces cerevisiae* (Weber et al., 1997).

The functionality of these proteins exhibits a high degree of diversity across different organisms. Human HDLBP has been linked to several processes, including the binding of HDL (McKnight et al., 1992), shuttling of tRNA from the nucleus to the cytoplasm (Kruse et al., 1998), ribosome binding (Kruse et al., 1996), the regulation of heterochromatin (Graham und Oram, 1987), gene imprinting (Batlle et al., 2011; Marsellach et al., 2006), and translation regulation (Cannarozzi et al., 2010). The number of known functions of HDLBP continues to expand. The depletion of HDLBP resulted in a reduction of secretion, thereby substantiating the importance of the protein (Zinnall et al., 2022). Additionally, the binding of HDLBP to C/U-rich motifs was linked, further substantiating the hypothesis that HDLBP exhibits a preference for binding to C/U-rich motifs. Additionally, recent findings indicate that in addition to gene silencing, HDLBP plays a role in maintaining genome stability and may contribute to the pathogenesis of autism spectrum disorders (Mushtaq et al., 2023). In addition to autism, HDLBP has been linked to a number of other health-relevant topics, including viral infection, lipid metabolism, and cancer (Feicht, 2024). The proto-oncogene *c-fms* is a direct target of HDLBP. A 69-nucleotide-long element in the 3'-UTR of *c-fms* plays a pivotal role in its regulation. This element contains five "CUU" triplets, which represent the binding motif of HDLBP. Furthermore, this binding motif is also bound by the RNA-binding protein HuR (Brennan et al., 2001). HuR functions to stabilize mRNA, whereas HDLBP promotes mRNA decay and downregulates *c-fms* translation. In this case, HDLBP has been demonstrated to function as a tumor repressor (Woo et al., 2011). Additionally, studies have shown that hdlbp (*X. laevis*) is capable of stabilizing vitellogenin mRNA (Dodson and Shapiro, 2002). Vitellogenin is a precursor of egg yolk. Similarly, as observed in the human vigilin HDLBP, DPP1 has been demonstrated to regulate heterochromatin and chromosome segregation (Cortes, 1999; Huertas et al., 2004).

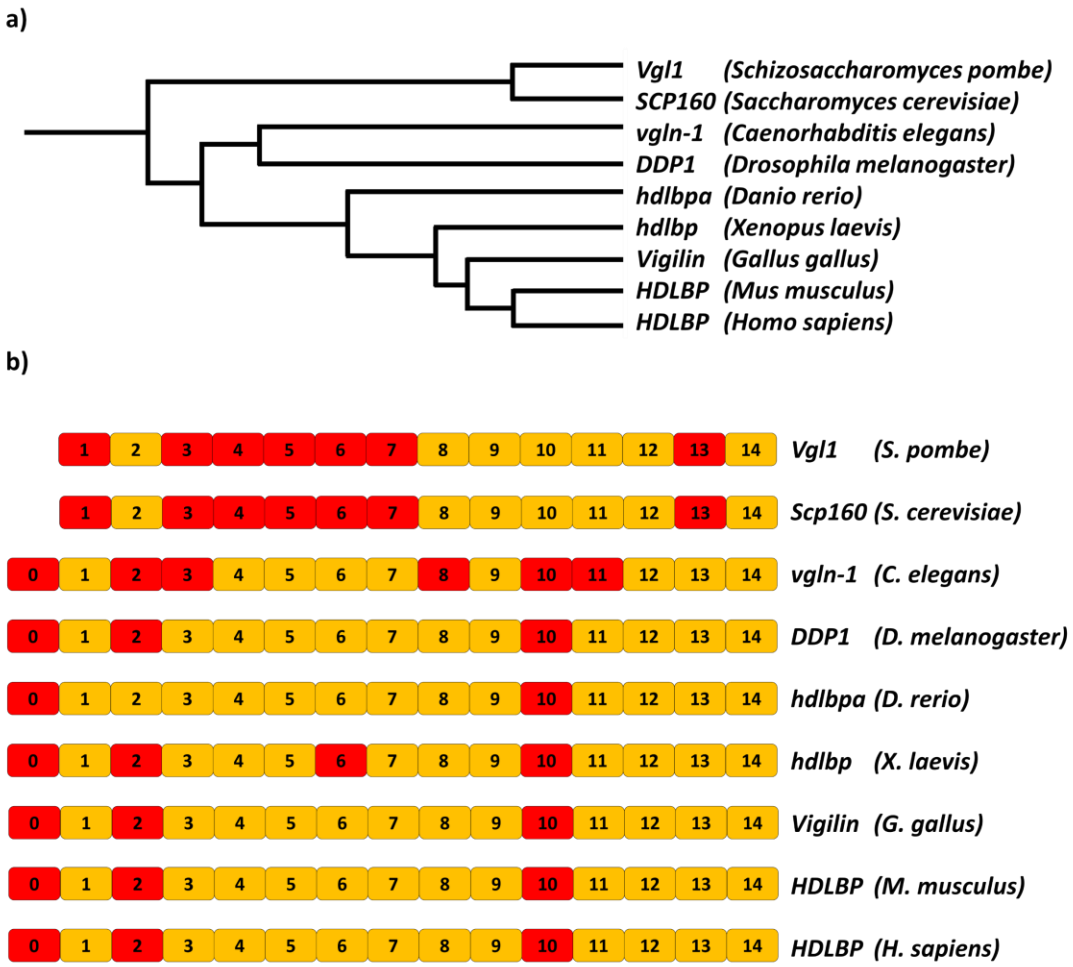


Figure 4: Evolutionary relationship and KH domain arrangement of viglin proteins; a) Phylogenetic tree of vigilin proteins (Information of was obtained by PhyloT V2 and phylogenetic tree was designed accordingly in GIMP V2.10.38) b) Overview of KH domains status of vigilin proteins: classical (orange) and diverged (red) domains; protein size and KH-domain size were not taken into aspect.

Vigilin proteins are typically located in the cytoplasm, however they frequently accumulate in close proximity to the ER or the cytoplasmic face of the outer nuclear membrane (Batlle et al., 2011; Klinger and Kruse, 1996; Wen et al., 2010). Additionally, metazoan homologues have been observed in the nucleus, whereas Scp160 of *S. cerevisiae* is exclusively present in the cytoplasm (Brykailo et al., 2007; Klinger and Kruse, 1996; Kruse et al., 2000; Zhou et al., 2008). In the nucleus, vigilin appears to be involved in the transport of tRNA and translation factor EF-1α within two complexes, designated VCCN and VCCC (vigilin-containing complexes). Furthermore, the complex VCCN contains the nuclear export receptor exportin-t, which suggests a role in the export of tRNA into the cytoplasm. The presence of a functional nuclear localization signal (NLS) in vigilin also implies a shuttling role for the RBP (Kruse et al., 2000, 1998). It should be noted that the binding of tRNA is not a universal phenomenon among metazoan vigilin homologues. It has been demonstrated that human vigilin and *Xenopus* vigilin exhibit differential binding preferences for distinct RNA species (Kanamori et al., 1998; Kruse et al., 1998).

Vigilin proteins also play a role in maintaining chromosomal integrity, which is essential for ensuring the unobstructed segregation of chromosomes during cell division. The deletion of Scp160 has been observed to result in an increase in the ploidy of the affected cells (Wintersberger et al., 1995). Ploidy is defined as the set number of chromosomes in a cell. While no direct association between vigilin and chromatin has been reported for *S. cerevisiae*, there is evidence for an interaction between the protein and the chromatin in metazoans. In the case of *D. melanogaster* vigilin, binding to a repetitive DNA sequence in the vicinity of the centromeres has been proven. It has been demonstrated that the protein binds exclusively to the C-rich strand of the dodeca satellite, which is a telomere-like centromeric satellite DNA. Conversely, no binding was observed for the complementary G-rich strand (Cortes, 1999). A comparable binding phenomenon was observed for human vigilin, wherein a high degree of affinity was demonstrated for α- and β-satellites (Wang et al., 2005). Satellite DNA is a type of tandem repeating, non-coding DNA that constitutes the core component of centromeres (Lohe et al., 1993). α-satellites are present in all chromosomes, whereas β-satellites are exclusively found in the centromeres of the following chromosomes: As reported by Altemose et al. (2022), the aforementioned satellite DNA is present in the following chromosomes: 1, 9, 13, 14, 15, 21, 22, and the Y chromosome.

It is noteworthy that both binding sequences exhibit attributes that are characteristic of the RNA targets that are most preferred by vigilin. These preferences include G-poor, single-stranded, and unstructured targets. Furthermore, the similarities in sequence preferences have been demonstrated in other species, including *Xenopus laevis, Saccharomyces cerevisiae, Mus musculus*, and *Homo sapiens* (Hogan et al., 2008; Mobin et al., 2016; Wang et al., 2005; Zinnall et al., 2022).

The reported functions of vigilins in the cytoplasm are starkly divergent from those observed in the nucleus (Figure 5 and Table 1). In mice, vigilin has been demonstrated to positively affect the translation of secreted liver proteins, which in turn modulate the blood plasma profile (Mobin et al., 2016). Scp160 is exclusively present in the cytoplasm, and its precise full function remains unknown. An overview of the localization and function of vigilin proteins can be found here (Cheng and Jansen, 2017).

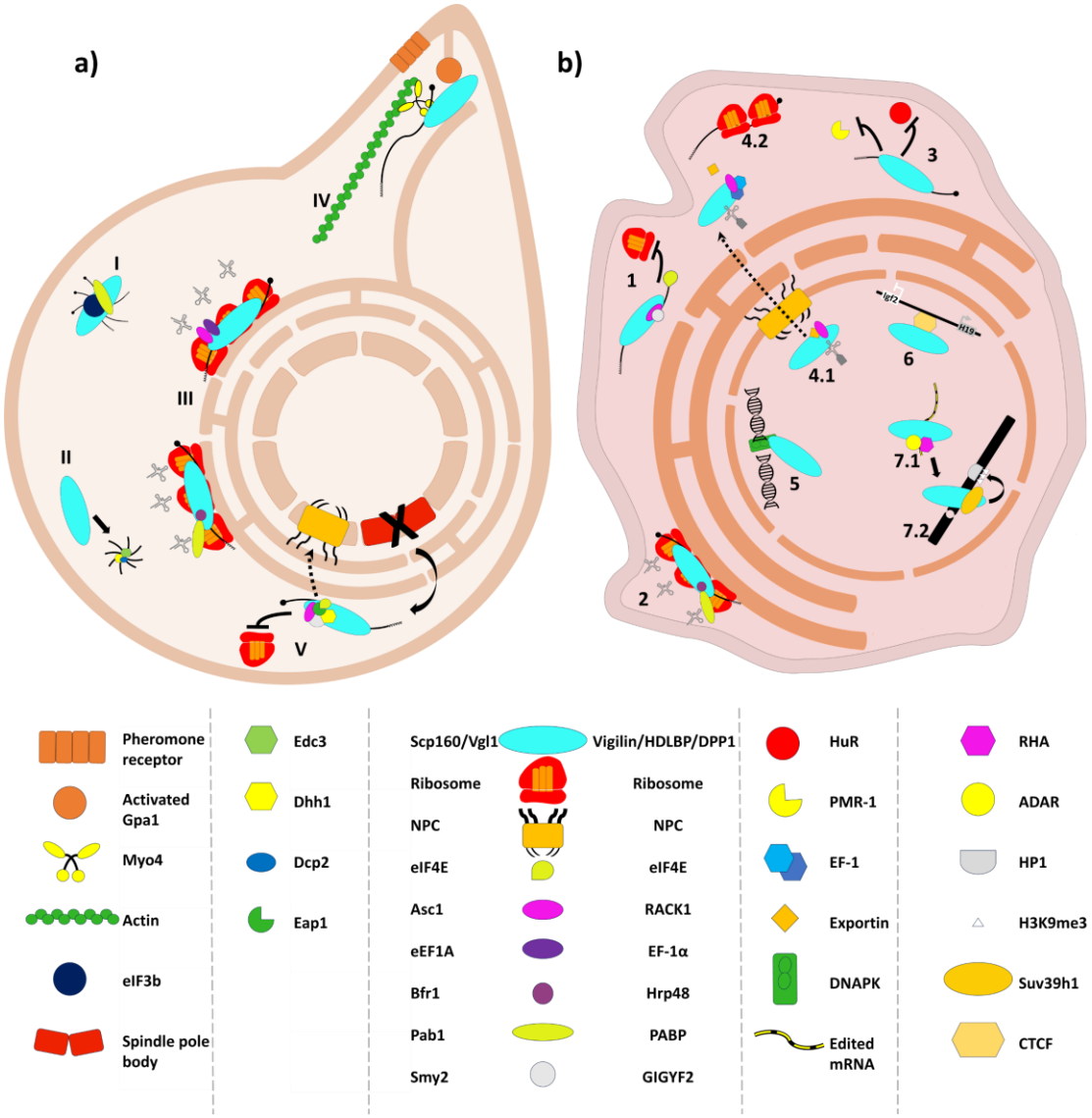


Figure 5: Vigilin functions in lower eucaryotes (a) and higher eukaryotes (b); vigilin-associated complexes are indicated by roman (lower eucaryotes) and arabic (higher eucaryotes) numbers: I: stress granules, II: P-bodies, III: translation complexes, IV: mating response complex, V: SESA; 1: repression complex, 2: translation complex, 3: mRNA stability regulator, 4.1: VCC_N, 4.2: VCC_C, 5: DNAPK-associated complex, 6: gene imprinting, 7.1: ADAR-associated complex, 7.2: Heterochromatin forming complex; proteins with homologous roles are displayed in the middle; proteins that are characterized by roles that are specific to either lower or higher eukaryotes are listed under the respective cell; Illustration modified from "A jack of all trades: the RNA-binding protein vigilin; figure 3" (Cheng and Jansen, 2017).

Table 1: Summary of Vigilin protein function and localization.

Organism	gene name	function	ref.	localization
Schizosaccharomyces pombe	vgl1	stress granule	Wen	ER, cytoplasm
Saccharomyces cerevisiae	*SCP160*	translation regulation	Sezen	ER, cytoplasm
		ploidy maintenance	Wintersberger	
		polysome association	Weidner	
		pheromone sensing	Gelin-Licht	
		Telomeric silencing	Marsellach	
		mRNA localization	Irie, Böhl	
		p-body localization	Weidner	
Caenorhabditis elegans	vgln-1	translation regulation	Zabinsky	cytoplasm
Drosophila melanogaster	DDP1	centromeric silencing	Huertas	ER, cytoplasm
		heterochromatin regulation		
Danio rerio	hdlbpa	(lipid metabolism)	Bruneau	?
Xenopus laevis	hdlbp	mRNA stabilization	Dodson	?
Gallus gallus	Vigilin	?		?
Mus musculus	Hdlbp	translation regulation	Mobin	?
Homo sapiens	HDLBP	translation regulation	Neu-Yilik, Baum, Klinger, Lang	ER, nucleus, cytoplasm
		lipid metabolism	McKnight	
		tRNA shuttling	Hirschmann	
		ribosome association	Vollbrandt	
		heterochromatin regulation	Klinger	
		gene imprinting	Liu Q	
		mRNA stabilization	McKnight	

Scp160 is characterized by its role in ploidy control

In 1995, the budding yeast protein vigilin, designated Scp160 or *S. cerevisiae* protein controlling the ploidy, was first described. The designation was selected based on the protein's molecular weight (approximately 160 kDa) and its distinctive phenotype, characterized by a loss of ploidy maintenance.

The protein in *S. cerevisiae* that controls ploidy was first identified by Wintersberger et al. (1995). Two years after its initial characterization, it was identified as a member of the vigilin protein family (Weber et al., 1997). Scp160 is comprised of 14 KH domains. Seven KH domains (domains KH2, KH8-12, and KH14) contain conserved GXXG motifs and thus represent classical KH domains. No GXXG motif is present in KH domains 1 and 3, while those in 4-7 and 13 exhibit incomplete GXXG motifs. Incomplete GXXG motifs, which lack one of the glycine residues, result in either an incomplete GXXX or XXXG motif. The absence of a GXXG motif, as well as the presence of an incomplete GXXG motif, can be classified as a diverged KH domain (Weber et al., 1997). The most C-terminal KH domains have been demonstrated to be indispensable for the protein's functionality. Specific functions and interactions have been identified for KH domains 13 and 14. The results of photo-cross-linking experiments demonstrated that there was an accumulation of RNA binding to the C-terminus of Scp160 (Kramer et al., 2014). The deletion of the last two KH domains resulted in the abolishment of ribosome association and a loss of specific mRNAs, including those encoding POL factors (mRNAs involved in polarity and secretion) and several mRNAs encoding cell wall and cell membrane components (Baum et al., 2004; Gelin-Licht et al., 2012a, 2012b). Similar observations were made in the case of the human vigilin, where the 14th KH domain was deleted, resulting in a loss of ribosome interaction (Kruse et al., 2000; Vollbrandt et al., 2004).

The sequence of the *SCP160* gene also indicates the presence of a NES and NLS in the N-terminus of the protein. However, mutation of the NES did not result in nuclear accumulation of Scp160, and a mutation of the NLS sequence showed no detectable effect. Therefore, both the NES and NLS appear to be non-functional with regard to the general protein localization, which encompasses nuclear envelopment and the rough ER (Wintersberger et al., 1995). These findings align with the identified mRNA binding partners of Scp160, which encode cell wall and secreted proteins. These classes of mRNAs are typically translated at the rough endoplasmic reticulum.

The yeast vigilin “Scp160” displays a variety of functions

The scope of action of Scp160 is outside the nucleus

The functions of Scp160 appear to be numerous and diverse. This chapter will provide a concise overview of the diverse functions of the yeast vigilin protein. Chromosome duplication and segregation are essential processes for all living organisms. It is therefore essential to ensure that this process is subject to rigorous and precise control, as any errors could prove lethal for the dividing cell. However, for *S. cerevisiae*, an uneven distribution of chromosomes is not necessarily fatal (Wintersberger et al., 1995). The genetic elements and mechanisms, as well as the interplay between the segregation apparatus, have been well characterized (Newlon 1988, Marston 2014). In the budding yeast, chromosome segregation is an intranuclear process, as the nuclear envelope does not break down during mitosis (Marston, 2014). Nevertheless, there is no evidence that Scp160 is located in the nucleus, and thus it is unlikely that it interferes physically with this process. However, Scp160 is localized at the ER membrane and thus at the nuclear envelope, which suggests the possibility of interference (Wintersberger et al., 1995). The outer nuclear membrane is continuous with the ER, and it has been suggested that both membranes accommodate a similar set of membrane proteins (Gerace and Burke 1988; Newport and Forbes 1987). It is therefore reasonable to posit that Scp160 may be in contact with the nuclear envelope.

Polysome associated Scp160 is localized at the ER

In cell fractionation experiments, the distribution of cytosolic and membrane-bound associated Scp160 to polysomes was analyzed (Frey et al., 2001). Furthermore, it was demonstrated that this interaction is dependent on mRNA (Frey et al., 2001). The use of S1 nuclease at a low concentration to degrade mRNA but not rRNA resulted in a loss of co-localization of Scp160 and the membrane-bound ribosomes. This indicates that mRNA, which is being translated, is the crucial factor in ER localization. It was demonstrated that co-localization to the ER was dependent not only on mRNA, but also on ribosomes. The release of membrane-bound ribosomes triggered by EDTA exhibited a lack of Scp160 signal at the ER, indicating that Scp160 is associated with polysomes at the ER in an mRNA-dependent manner. Subsequently, this association was found to be linked to microtubules. It was demonstrated that the absence of microtubules resulted in the delocalization of Scp160 from the ER. The transport of mRNAs in yeast is dependent on actin and microtubules. Therefore, the delocalization of Scp160 is a logical consequence of this dependency (Heym and Niessing, 2012). The hypothesis was put forth that the Scp160-mRNA-ribosome complexes are transported from the cytosol to the ER as their final destination (Frey et al., 2001). The binding of Scp160 to polysomes is dependent on the

presence of the two proteins, Asc1 and Bfr1. The loss of either of these two proteins results in the separation of Scp160 and the ribosomes. Asc1 is a core component of the 40S subunit of eukaryotic ribosomes. Bfr1 is directly associated with Scp160 in a mRNA-dependent manner and is known to be a nonessential RNA-binding protein (Baum et al., 2004; Lang, 2001).

The C-terminus of Scp160 plays a significant role in its functionality

The most C-terminal KH domains appear to be indispensable for Scp160-polysome association (Li et al., 2004). This discovery is consistent with the identification of the vigilin protein in humans, where it was demonstrated that the C-terminal domain is essential for ribosome binding (Vollbrandt et al., 2004). The last two KH domains of Scp160 were subjected to comprehensive investigation with regard to their impact on the functionality of Scp160 and vigilin (Baum et al., 2004; Melissa A. Brykailo et al., 2007; Gelin-Licht et al., 2008). However, subsequent findings indicated that the loss of KH domains 13 and 14 does not inevitably result in Scp160 delocalization from the ER (Schreck, 2010). Nevertheless, mutants lacking KH domains 13 and 14 exhibit a number of phenotypes. Initially, the overall fitness of the cells was observed to decline. An increased sensitivity to hygromycin B and cycloheximide was observed (Baum et al., 2004). Cycloheximide has been demonstrated to impede protein synthesis by interfering with the translocation process (Siegel and Sisler, 1963). Hygromycin B, on the other hand, has been shown to inhibit protein synthesis by retaining tRNA at the ribosomal A-site (González et al., 1978). Both processes result in further impairment of translation. If the increased delocalization of Scp160 from the ER is considered, the additional translation deficiency aligns with the increased sensitivity. Growth at elevated temperatures (37°C) was more severely inhibited compared to even the complete deletion of *SCP160*. The additional deletion of KH domain 11 and KH domain 12 further exacerbated the present phenotypes (Baum et al., 2004).

The presence of *SCP160* is necessary for the maintenance of consistent ploidy

Initial experiments demonstrated that the deletion of the *SCP160* gene resulted in alterations to the ploidy of the deletion strains. The haploid cells displayed a pseudo-diploid FACS profile following the truncation of the C-terminal KH domains (Frey, 2002). In the complete deletion of *SCP160*, the observation was made that pseudo-diploid and subsequently even pseudo-tetraploid cells could be identified. This imbalance in chromosome number is not the result of an error in chromosome duplication; rather, it is caused by mis segregation of chromosomes. This conclusion was reached based on the observation that in mother-daughter cell pairs resulting from cell division, one of the cells lacked any nuclear material (Wintersberger et al., 1995). One potential explanation for the absence of nuclear material is the disruption of the yeast microtubule organizing center, the spindle pole body (Sezen et al., 2009). A potential additional phenotype of *scp160Δ* cells is an increase in size.

This change in size can be attributed to an increase in DNA content and/or cell size resulting from stress (Tibayrenc et al., 2010; Wintersberger et al., 1995).

Scp160 is part of the SESA complex

The defect in chromosome segregation observed in cells lacking functional Scp160 has been postulated to originate from Scp160's involvement in the SESA (Smy2, Eap1, Scp160, Asc1) complex (Sezen et al., 2009). The complex is composed of Smy2, Eap1, Scp160, Asc1, Bfr1, and Dhh1, with some members serving as direct interactors of Scp160. Smy2, which was discovered and named after its ability to compensate for the temperature sensitivity of myo2 mutants, contains a glycine-tyrosine-phenylalanine (GYF) domain that is capable of binding to polyproline domains, which are found, for example, in Eap1. Thus far, no direct interaction between Scp160 and Smy2 has been observed (Lillie und Brown 1994; Kofler et al. 2005). Eap1 is an eIF4E-binding protein that interacts with Smy2 and Dhh1 (Chial et al., 2000; Gingras et al., 1999). Asc1 and Bfr1 have been demonstrated to bind Scp160 directly, as previously described (Sezen et al., 2009). Dhh1, a DExD-box ATPase and mRNA de-capping factor, has also been identified as an interaction partner of Scp160 (Carroll et al., 2011). The SESA complex has been demonstrated to bind the mRNA of the nuclear pore complex (NPC) protein Pom34. This protein plays a role in the process of spindle pole body duplication. The binding of the *POM34* mRNA may represent a potential mechanism for quality control in chromosome segregation. It is postulated that the SESA complex functions to impede the initiation of translation of *POM34* mRNA, thereby reducing Pom34 protein levels (Sezen et al., 2009).

The impact of Scp160 on bound RNAs is diverse. As observed in the case of *POM34*, the presence of Scp160 has been demonstrated to result in the suppression of mRNA (Sezen et al., 2009). Together with its presumptive partner Bfr1, Scp160 has been demonstrated to be essential for the formation of proper P-bodies. The P-bodies affect the turnover of RNA (Teixeira et al., 2005). In this context, Scp160 appears to maintain the P-body components in a state of readiness, preventing mRNA decay under non-stress conditions (Weidner et al., 2014)

It seems reasonable to posit that the function of the SESA complex is not solely focused on spindle body control. Other known bound mRNAs include *CCW14*, which encodes a cell wall glycoprotein; *AGA1*, which codes for a subunit α-agglutinin in the cell wall; *PRY3*, a glycosyl-phosphatidyl-inositol (GPI)-anchored cell wall protein; and *MSB2*, which encodes an integral plasma membrane mucin protein linked to osmosensing (Hirschmann et al., 2014). It is established that Scp160 exerts an influence on the composition of the cell wall, although the precise nature of this effect remains unclear. The deletion of the *SCP160* gene has been observed to result in a pronounced emulsifying

capacity of yeast cells (Nerome et al., 2020), a property that has been linked to alterations in the cell wall composition of the affected yeast strains.

Scp160 is binding different types of RNA

In addition to its role in mRNA binding, Scp160 has been observed to interact with a range of other nucleic acids, including ssDNA, dsDNA, rRNA, and tRNA (Weber et al., 1997; Hirschmann et al., 2014). In the initial attempt to assign specific mRNAs to Scp160, a relatively incoherent set of mRNAs was identified. This includes genes that code for microtubule-associated proteins, sporulation-promoting proteins, meiosis-specific DNA helicase, and other proteins (Li et al., 2003). Subsequent studies revealed that Scp160 binds to a vast array of mRNAs, particularly those encoding cell wall proteins, membrane proteins, and secreted proteins (Sezen et al., 2009). It is noteworthy that RNA processing and ribosome biogenesis, which are localized at the nucleolus, may also represent a potential target for Scp160 (Hogan et al., 2008). These findings can be corroborated. Two microarray-based studies have identified a specific set of mRNAs that bind to the protein. As previously observed, the majority of these mRNAs were found to encode proteins with established links to the cell wall or plasma membrane (Hirschmann et al., 2014). An analysis conducted using the MIPS Functional Catalogue Database revealed an enrichment of mRNA coding for extracellular, cell wall proteins and proteins involved in ER-Golgi transport (Hirschmann et al., 2014; Ruepp, 2004).

Scp160 is involved in tRNA recycling and autocorrelation

Furthermore, a function of Scp160 regarding tRNA recycling was put forth. tRNA was observed to form a complex with the human vigilin protein HDLBP and eEF1A. A general interaction between tRNA and vigilin proteins was proposed as a potential explanation (Kruse et al., 1996; Kruse et al., 1998; Baum et al., 2004). Subsequently, it was postulated that Scp160 inhibits the dissociation of tRNA from the ribosome, thereby ensuring a robust association of the ribosome, tRNA, and the corresponding aminoacyl-tRNA synthetase (Hirschmann et al., 2014). The hypothesis posited that Scp160 retained empty t-RNAs in the vicinity of the translation ribosomes, while its interaction partner Asc1 recruited aminoacyl-tRNA synthetases (Hirschmann et al., 2014). Therefor It is conceivable that both proteins function collectively to facilitate the reloading of tRNAs with the corresponding amino acids via nearby aminoacyl-tRNA synthetases resulting in an improvement of translation elongation. The binding of tRNA may be linked to autocorrelation, which refers to the reuse of tRNAs by different synonymous codons (Hirschmann et al., 2014). The group of mRNAs that exhibited a decrease upon *SCP160* knockout exhibited a common characteristic: all were significantly above average in terms of autocorrelation (Hirschmann et al., 2014).

The potential function of increasing translation elongation efficiency of Scp160 was analyzed with reporter proteins derived from the first exon of the huntingtin protein (mHtt). The translation of autocorrelated and non-autocorrelated codons was further investigated using different polyglutamine containing reporters (Cheng et al., 2018). A polyQ-containing reporter protein allows not only for the analysis of translation in a region with synonymous codons but also permits the investigation of the impact of this well-studied low complexity region (LCR) on protein folding, as it tends to form large aggregates. The aggregation patterns of various polyQ reporters in *SCP160* mutant strains were assessed, and it was demonstrated that the absence of Scp160 affects the aggregation and toxicity of the reporters, as well as the aggregation of two additional proteins, namely Nab3 and Cyc8, that naturally contain LCRs (Cheng et al., 2018).

rRNA binding has yet to be proven for Scp160

The binding of rRNA has been proposed, but thus far has not been demonstrated *in vivo*. *In vitro* experiments have demonstrated the ability of Scp160 to bind to rRNA (Weber et al., 1997). Additionally, it is conceivable that Scp160 may bind rRNA indirectly through its interaction with Asc1. The human Asc1 homologue, Rack1 has been demonstrated to bind rRNA, thereby rendering the assumption of Asc1 binding rRNA a probable outcome (Coyle et al., 2009).

Scp160 is involved in pheromone sensing and chemotropism

Scp160 was identified as a Gpa1 binding partner (Guo et al., 2003). Gpa1, the α-subunit of protein G, has been demonstrated to play a pivotal role in the pheromone response. This finding has subsequently led to the identification of Scp160 as a constituent of the mating response pathway. Scp160 is indispensable for the localization of Sro7 mRNA to the apex of budding yeast, in conjunction with another RBP, She2, and a motor protein, Myo4p (Gelin-Licht et al., 2012a). During the mating process, the presence of Scp160 is crucial, whereas She2 is not necessary for the transportation of RNA to the polarization site, the shmoo tip. During the formation of the shmoo tip, Scp160 was observed to bind to Myo4, which subsequently localized to the cortical endoplasmic reticulum. The binding of mRNA is enhanced by the activation of Gpa1 and the addition of mating factors. It is remarkable that this mRNA plays a pivotal role in regulating cell polarity. In light of these findings, it appears that Scp160 plays a role in the binding and transport of polarity and secretion factor (POL) mRNAs towards the region of active growth and reproduction. Furthermore, it was demonstrated that the deletion of the C-terminal KH domain alone can result in a phenotype similar to that observed in a full knockout (Gelin-Licht et al., 2012). Subsequently, it was demonstrated that Asc1 interacts with Scp160 for pheromone sensing and Gpa1 activation. The mitogen-activated protein kinase (MAPK)

pathways involved in filamentous growth and mating exhibit a shared dependence on Scp160/Asc1. Asc1 binding is facilitated by pheromone-activated Scp160, which regulates the two pathways at the translational level. The interaction of both proteins results in the dispersion of mating signals, which consequently leads to a reduction in filamentous growth and nutrient sensing signals (Gelin-Licht et al., 2020).

Scp160 affects telomeric silencing

The deletion of Scp160 resulted in the relaxation of telomer silencing and the mating type locus (HMR). The telomere position effect is associated with the inhibition of the silencing protein Sir3 at the telomeres. While Sir3 is essential for HMR, telomere formation, and heterochromatin formation, it is not required at the ribosomal locus (rDNA) (Marsellach et al., 2006). It is postulated that its function may originate from its localization at the endoplasmic reticulum, which is also the cytoplasmic face of the nuclear envelope. Telomere clustering is known to occur at this precise location, specifically at the nuclear face of the envelope (Cortes, 1999; Marsellach et al., 2006).

The cell wall composition is affected by Scp160 deletion

As previously stated, Scp160 has been demonstrated to interact with mRNAs for proteins involved in the formation of the cell membrane and the cell wall. In a recent study, 5076 strains were screened for an emulsifying phenotype, which is defined as the ability to mix two liquids that are typically immiscible. The deletion of Scp160 was identified as one of eight proteins that resulted in the desired outcome. It is of note that Asc1 was another of the eight findings (Nerome et al., 2020). The alteration in their emulsifying capacity was attributed to modifications in their cell wall structure. It is established that Asc1 deletion results in a defective synthesis of the chitin synthase Chs2, which consequently affects the chitin fraction of the cell wall. A protein with a more extensive history of study within the same phenotypic category, exhibiting enhanced emulsifying properties, is Mcd4. The deletion of Mcd4 has been observed to yield a similar outcome to that of Scp160, namely, a defective synthesis of GPI anchors. In the absence of these anchors, the linkage of mannoproteins and β-glycans is disrupted. The reduction in the mannose-to-glucose ratio observed in *scp160Δ* could not be demonstrated. The precise alterations induced by the deletion of Scp160 in the cell wall remain to be elucidated (Nerome et al., 2020).

Closing remarks

As previously stated, vigilins are composed of either 14 or 15 KH domains, as illustrated in Figure 4 b. It should be noted that the KH domains in question are not identical. These domains can be classified into two main categories: classical and divergent KH domains. Divergent KH domains can be identified by the absence of the GXXG motif, as observed by Weber et al. (1997). However, this alteration does not necessarily indicate a loss of function, as it has been demonstrated that diverged domains can functionally replace conserved KH domains (Brykailo et al., 2007a). The arrangement of classical and divergent domains varies among different vigilin homologues. The vigilin proteins of *S. cerevisiae*, *S. pombe*, and *A. gossypii* exhibit a comparable pattern in their 14 KH domains to that observed in the homologues of *H. sapiens*, *X. laevis, G. gallus, C. elegans, D. rerio,* and *D. melanogaster* in their 15 KH domains. HDLBP has only three diverged KH domains, in contrast to the eight observed in Scp160 (Currie and Brown, 1999; Weber et al., 1997). To date, the C-terminal KH domains have been demonstrated to be of the utmost importance for vigilin function. The binding of RNA and polysomes is impaired following the truncation of KH domains 13-14 (Baum et al., 2004; Gelin-Licht et al., 2012a; Hirschmann et al., 2014; Vollbrandt et al., 2004).

Aim of this thesis

It has been demonstrated that the KH domains of vigilins, including Scp160, are indispensable structural and functional components. Thus far, the predominant focus of analysis has been on the C-terminal KH domains, which have been postulated to be of paramount importance for Scp160 function (Baum et al., 2004; Brykailo et al., 2007; Gelin-Licht et al., 2012b; Hirschmann et al., 2014). Furthermore, it is important to note that evidence exists demonstrating the significance of both classical and divergent KH domains. In contrast, aside from the interchanging of KH domains (Brykailo et al., 2007), there is currently a dearth of data concerning the role of the vast majority of N-terminal and central KH domains of Scp160. The objective of this thesis was to draw attention to the N-terminal KH domains of Scp160 that have not been previously investigated. We sought to identify specific regions of interest, whether single KH domains or KH domain clusters, and, if possible, to ascribe distinctive properties, functions, and interaction partners to them.

Material and Methods

Methods

Standard Cloning procedure

Restriction and Ligation of plasmids and PCR-products were performed according to the manuals of the enzymes used for the corresponding cloning experiment. This normally includes a 1h restriction digestion on 37°C with either heat inactivation or purifying of the product and a subsequent ligation overnight. Plasmid used are displayed in Table 2. Primer used in this thesis are displayed in Table 3.

Standard methods

Many of the used molecular biological methods are based on standard techniques (Ausubel, 2000; Sambrook, 2001) and were used accordingly.

Yeast culture

All yeast strains used or generated during this work are listed in Table 4. Yeast Peptone Adenine Dextrose (YPAD) was used as full medium for yeast cultivation. Full medium contains 10g yeast extract, 20g peptone, 20g dextrose and 0.4g/l adenine sulfate. Synthetic complete media (SC-medium) was used for selective cultivation. SC-medium consists of 6.7g/l yeast nitrogen base, and a particular mixture of amino acids shown in Table 5.

Yeast cell lysis for protein extracts

20 OD_{600} units (equivalent to approx. 4 x 10^8 cells) were harvested in a 2 ml Eppendorf tube. The cell pellet was dissolved in 300µl of lysis buffer (25mM Tris-HCl pH 7.5, 50mM KCl, 10mM $MgCl_2$, 1mM EDTA, 5% glycerol, 0.5% Triton X-100), 200 to 250mg glass beads (0.25-0.5 mm) were added, and the mix vortexed on a BV1005 Benchmixer™ (Benchmark) for 10min at 4°C. The glass beads and the cell debris were spun down for 1min at 1.000g at 4°C. The supernatant was transferred to a new Safe-Seal® Eppendorf tube and stored at -20°C until use.

SDS PAGE and western blotting

The SDS-PAGE gel electrophoresis procedure was conducted using the Bio-Rad MP system in accordance with the following methodology. The separating gel (Table 6) was casted in adequate glass frames and overlayed with 200µl isopropanol. After polymerization the isopropanol was discarded and a stacking gel (Table 6) was prepared with desirable combs. Protein concentration in the cell lysate

was measured by a Bradford assay. 20µg protein were loaded per lane The SDS-PAGE was conducted in accordance with the methodology outlined by Laemmli (1970). The proteins were subsequently transferred onto a PVDF membrane (Amersham Hybind P 0.45µm, GE Healthcare) (Amersham Hybind P 0.45µm, GE Healthcare) via a BioRad tank blotting system (350mA for 60-90min at 4°C). The membrane was blocked in 1x PBS+5% dried milk for 20min at room temperature, followed by probing the membrane with a primary antibody in PBS-T over night at 4°C. The antibody concentration was determined in accordance with the recommended values provided by the corresponding manufacturer. The membrane was washed with PBS-T three times for 10min and incubated with a suitable secondary antibody coupled to a HRP conjugate for 1h at room temperature. The membrane was again washed three times with PBS-T and afterwards incubated for 1min with ECL substrate. Immunofluorescence was detected with a ChemiDoc MP Imaging System (BioRad).

PCR

PCR was performed according to the Agilent Inc. Herculase II Fusion DNA Polymerases protocol (see official user manual). High efficiency PCR was performed according to (Janke et al., 2004); (hot start was not performed).

Transformation in Yeast

High efficiency transformation was adapted from the original protocol according to Gietz and Schiestl (2007).

Single colonies are inoculated in 5ml YAPD medium and are grown overnight. A 50ml of YAPD were inoculated with the overnight culture to an end OD_{600} of 0.25. The suspension was allowed to grow for approximately four hours, reaching an OD_{600} of 1. 10 OD_{600} are harvested and washed twice with 20ml of sterile water. Cell pellets were dissolved in 360µl of the previously prepared transformation mixture. The transformation mix is composed of 240µl PEG3350 (50%(w/v), 35µl 1M lithium acetate, 50µl salmon sperm DNA (2mg ml^{-1}), 34µl of diluted plasmid DNA (1-5µg) in sterile water. After heavy mixing, the transformation mix is incubated at 42°C for 30min. After incubation of the tubes are pelleted for 30s at 13.000g. Supernatant is discarded and pellets are dissolved in 1ml sterile water. The cell suspension was plated in volumes between 20-200µl on corresponding selective agar plates. Colonies appear after around two to seven days.

High efficiency transformation was used for CRISPR/Cas9, genomic tagging and genomic deletion experiments.

One step transformation was performed according Chen et al. (1992). For plasmid transformation, freshly grown yeast cells were scratched with an inoculation loop from an agar plate and resuspended

in 100µl one-step-buffer (0.2 M lithium acetate, 40% (w/v) PEG 3350, 100 mM DTT). 50µg salmon sperm DNA and 10-5000ng plasmid DNA were added to the cell suspension and incubated at 45°C for 30min. The cells were pelleted for 10s at 10.000g and plated on selective plates. Colonies appear after around two days.

CRISPR/Cas9 for yeast genome editing

For the deletion of specific KH domains of *SCP160*, a CRISPR/Cas9 approach was used. Suitable guide RNAs for cuts 3´and 5´ of each KH domain were determined by the online tool ChopChop (Labun et al., 2019). For integration in the backbone of the Cas9-plasmids a CTTT-overhang for the forward and an AAAC-overhang for reverse primer was added. The guide RNA sequences to be included into the CRISPR/Cas9-plasmid were ordered as complementary DNA-primers and annealed by heating to 95°C and an increment of -1°C per minute. The reaction took place in a reaction mixture of 1µl reverse, 1µl forward primer (100µM) and 18µl of water. The 3´gRNAs were cloned into pCAS-HIS3-BsaI and the 5´gRNAs likewise in pCAS-URA3-BsaI, provided by Rapaport group (Ryan et al. 2014). For a scarless deletion, a repair template was designed, containing 50bp of homology to each side of the deletion. The transformation protocol used for the successful truncations/deletions of *SCP160* was the one for high efficiency (see above), but with due to a much lower transformation rate. Positive clones were screened by PCR using primers in Table 3.

Tagging and knock out of yeast genes

Tagging or knock out cassettes were planed using public domain ApE (https://jorgensen.biology.utah.edu/wayned/Ape/) or commercial SnapGene software (GSL Biotech LLC, Boston) and the corresponding PCR products generated using high efficiency PCR and primers listed in Table 3. The cassettes consist of a desired tag and a selection marker for a tagging cassette or a selection marker only for deletion cassettes (Janke et al., 2004; Knop et al., 1999; Gauss et al., 2005). Either is flanked with 40bp of homologue sequences C-terminal and N-terminal of the to delete/tag area of interest.

Tagging or knock out cassettes were transformed using the high efficiency transformation protocol.

Transformation in *E. coli*

For multiplying of plasmids they were transformed in *E. coli* TOP10 cells. Here for 100µg of plasmid were mixed with TOP10 cells and incubated for 10min at 4°C. The cells were heat-shocked at 42°C for 1 min and immediately plated onto a LB-plate with the corresponding antibiotic (usually ampicillin or

kanamycin). After 24h at 37°C the colonies were screened for positive clones: Plasmids were isolated and tested via PCR for required sequence. New plasmids were sent for sequencing (GATC Biotech).

Ploidy analysis: propidium iodide staining and flow cytometry

An overnight culture of the yeast strains of interest was inoculated in YPAD. The overnight culture was expanded with a starting OD_{600} of 0.25/ml. The culture was then harvested at an OD_{600}<1/ml. 2 OD_{600} units (equivalent to $4x10^7$ cells) of logarithmic growing cells were harvested and washed twice with 50mM Tris-HCl pH 8.0. The cells were resuspended in 1mL 50mM Tris-HCl pH 8.0 containing 70% Ethanol and incubated at 4°C overnight in an overhead shaker for fixation. The samples were vortexed and 150 to 250µl of fixed cells are spun down at full speed. The cells were again washed twice with 1mL 50mM Tris-HCl pH 8.0, resuspended in 0.5ml 50mM Tris-HCl pH 8.0 containing 20µg/mL heat-treated RNase A, and incubated for 2h at 50°C. Afterward 20 µl of 20mg/mL Proteinase K was added and incubated for another hour at 50°C. For staining, the cells were pelleted, washed twice with 1ml FACS buffer (Table 7), resuspend in 500µl FACS-PI solution (Table 8), and incubated for 2h in the dark at room temperature. Before measuring, the samples were again vortexed for 5 to 10s. The acquisition was made on a Beckman Coulter CytoFLEX S flow cytometer. Fluorescence signal was measured by excitation with a 638nm laser and a bandpass 610/20 mCherry filter, with a gain of 163nm. 30,000 events were recorded per sample. For exclusion of cell debris an FSC (forward scatter) vs SSC (side scatter) density plot was gated (population P1). Aggregated cells and doublets were excluded from analysis by gating P1 population using an ECD-Width (x-axis) vs ECD-Area (y-axis) density plot (population P2).

Water-kerosene emulsion assays

Emulsion assays were basically performed according to Nerome et al. (Nerome et. al. 2020). Yeast strains were grown in YPAD medium for 48h at 30°C. The cells were harvested by centrifugation at 3000g for 5min. The OD_{600} was adjusted to 0.7 in PBS supplemented with 0.6M sorbitol (PBSS). 5ml suspension was centrifuged and resuspended in 5ml PBS and 1ml of Kerosene was added. The solution was vortexed for 30s and incubated for 48h. The emulsified phases were documented by photography after 1h, 24h and 48h.

Yeast genomic DNA (gDNA) extraction

A colony of interest was resuspended in 100µl of a solution of 200mM LiAc and 1% SDS. The solution was incubated for 15min at 70°C. 300µl of 96-100% ethanol was added and vortexed. The DNA and cell debris was spun down at 15.000g for 3min. The pellet was washed 400µl of 70% ethanol and again

spun down at 15.000g for 1min. Any remaining ethanol was evaporated at 60-70°C. After dissolving the pellet in 100µl ddH_2O the cell debris was spun down a last time for 15s at 15.000g.

Fluorescence Microscopy

Fluorescence microscopy was performed with a Zeiss CellObserver microscopy controlled by Axiovison ZEN 2.6 (blue edition) software. Yeast cultures were grown with a starting OD_{600} of 0.25 in the corresponding medium. 2 OD_{600} of the culture was then harvested at an OD_{600} < 1,2. The cells were resuspended in ~40-50µl of reflow. 3.5µl of the cell suspension was transferred onto a glass slide and sealed with a cover slip. The signal of the Scp160-GFP fusion protein was detected with a 475nm LED/filter and that of the reporter-mCherry fusion protein with a 590nm LED/filter.

Immunoprecipitation

Overnight cultures were again re-inoculated in an overnight culture. The main cultures were inoculated in YAPD with a start OD_{600} of 0.25. 50 OD_{600} of cells were harvested in their logarithmic phase for immunoprecipitation. Lysis was performed in lysisbuffer II (Table 10) with 200µl 0.25mm-0.5mm glass beads in a 2ml Eppendorf tube. Tubes were vortexed on a BV1005 Benchmixer™ (Benchmark) at 4°C for three times 4min. After centrifugation at 4°C for 3 minutes at 1000g the lysates were transferred to new tubes and cleared at 20.000g for 30min at 4°C. Lysates were mixed 2:3 with dilution buffer (Table 11) and incubated with 25µl of GFP-Trap® Magnetic Particles (MP) M-270 at 4°C for 1h. Beads were afterwards washed with washing buffer (Table 12) four times. Finally, MP M-270 were diluted with 2x Laemmli buffer and boiled for 10min. MP M-270 were removed and the supernatants were used for normalization via westernblot. Here for the band intensity was measured via FIJI and samples were diluted accordingly. Normalized samples were used for mass spectrometry.

Mass spectrometry

Mass spectrometry samples were sent to the Proteome Center Tuebingen (Contact person: Dr. Mirita Franz).

In-gel digestion of proteins:

Proteins were loaded on a NuPAGE 12% Bis-Tris Gel (Thermo Fisher Scientific) for a short SDS based gel electrophoresis and stained with colloidal Coomassie using the ReadyBlue Protein Gel Stain (Merck Millipore). Gel sections containing proteins were excised and cut into smaller parts. For in-gel digestion of proteins, gel pieces were destained by washing three times with 5mM ammonium bicarbonate (ABC) in acetonitrile (ACN) (1:1, v/v) for 20min. After a dehydration step with 100% ACN for 10min, disulfide bonds were reduced with 10mM dithiothreitol (DTT) in 20mM ABC for 45min at

56°C, and thiol groups of cysteine residues were prevented from reoxidation by carbamidomethylation with 55mM iodoacetamide (IAA) in 20mM ABC for 45min in the dark. Gel pieces were then washed two times with 5 mM ABC in ACN (1:1, v/v) for 20min and dehydrated with 100% ACN for 15min. After evaporation of the liquid in a vacuum centrifuge for 10min, gel pieces were soaked in a solution of 12.5ng/µl sequencing grade trypsin (Promega) in 20mM ABC, pH 8.0 for 10min at room temperature (RT), and then covered with 20mM ABC. After in-gel digestion of proteins at 37°C overnight, peptides were extracted in three consecutive steps with different extraction buffers for 30min: first 3% (v/v) trifluoroacetic acid (TFA) in 30% (v/v) ACN was added, followed by 0.5% (v/v) formic acid (FA) in 80% (v/v) ACN, and finally by 100% ACN. ACN was evaporated from pooled supernatants by vacuum centrifugation. In the course of the digestion protocol all incubation steps were carried out under shaking.

LC MS/MS:

Peptides were desalted with C_{18} StageTips (Rappsilber et al. 2007) and analyzed on an EASY-nLC 1200 UHPLC coupled to a Q Exactive HF mass spectrometer (both Thermo Fisher Scientific) as described previously (Semanjski et al. 2018) with little modification: peptides were separated on the analytical column using a 46min segmented gradient of 10-33-50% of HPLC solvent B (80% acetonitrile in 0.1% formic acid) at a flow rate of 200nl/min.

In the mass spectrometer, MS and MS/MS spectra were generated at resolution 60k. Full MS target value and maximum IT were set to $3x10^6$ and 25ms, respectively. In each scan cycle, the 7 most intense precursor ions were picked up. The MS/MS target value was set to 10^5 charges with a maximum IT of 220ms.

MS data processing:

MS data were searched against a target-decoy (Elias, Gygi 2007) database of *Saccharomyces cerevisiae* (6,089 entries, downloaded on 16th of December 2022), the sequence of the yeGFP fragment starting at position 16, and commonly observed contaminants using the Andromeda search engine integrated into the MaxQuant software (v1.6.14.0) (Cox et al. 2008, Cox et al. 2011). Search parameters were kept to default. In addition, the iBAQ (Intensity Based Absolute Quantification) and LFQ (Label-Free Quantification) algorithms were enabled, as was the “match between runs” option (Schwanhäusser et al. 2011, Luber et al. 2010).

Resulting data was processed by the QbiC (Zentrum für Quantitative Biologie) (Contact person: Francesca Barletta):

Bioinformatic Analysis:

In order to determine the impact of mutations at the N-terminal KH domains on the binding behavior of Scp160 in *S. cerevisiae*, the six mutations: 1) ΔKH1-2yeGFP, 2) ΔKH1-5yeGFP, 3) ΔKH4-5yeGFP, 4) ΔKH5yeGFP, 5) ΔKH13-14yeGFP (positive control/change of interactome) and 6) SCPp-yeGFP-SCPt (negative control/GFP under promotor and terminator); were compared against the WT Scp160yeGFP (N-terminal KH domains full-length).

Downstream statistical analysis of proteomics was performed in R (version 4.2.2). The R package proteus (version 0.2.16) was used to analyze MaxQuant's Proteomics output file "proteinGroups.txt". As part of the analysis with proteus, data was quantile normalized to account for variation of intensity between samples followed by log2 transformation. Differential expression (DE) analysis was performed with the R package Limma (version 3.46.0) outside of the package Proteus. As cut-off for statistical significance a multiple adjusted *p* value (adj. P. Val) < 0.05 was chosen, which is corrected for multiple testing to control the false discovery rate (FDR).

In order to identify differentially expressed proteins a linear model was then fitted to each protein as follows: exp = ~ condition with "exp" representing expression of a protein and condition representing type of culture with the seven levels: WT, Scp160ΔKH1-2, Scp160ΔKH1-5, Scp160ΔKH4-5, Scp160ΔKH5, Scp160ΔKH13-14 and the negative control Scp160p-yeGFP-Scp160t (Scp160p: Scp160 promotor and Scp160t: Scp160 terminator). For graphical visualization heatmaps and a volcano plots showing statistical significance -log10(*p*-value) versus log2 FC were produced.

Contributed Figures: 30-38; Supplementary Figures: 3-7 (Figures have been adapted accordingly).

CRAC experiments

Cross-linking and analysis of cDNAs (CRAC) is a powerful technique used to study RNA-protein interactions at high resolution (Granneman et al. 2009). By using ultraviolet (UV) light, RNA is covalently cross-linked to its interacting proteins in living cells, preserving their native interactions. This allows for precise mapping of RNA-binding sites without the usage of antibodies and provides insights into the function of RNA-binding proteins. The method combines affinity purification, enzymatic processing, and high-throughput sequencing to generate comprehensive RNA-protein interaction profiles. CRAC is widely used in molecular biology to investigate RNA processing, stability, and regulation (Granneman et al. 2009, Bohnsack et al. 2012).

In the preliminary stage, the proteins of interest were tagged with a TAP-tag for tandem affinity purification. To induce covalent bonds between RNA and associated proteins, preserving their interactions, the samples were exposed to ultraviolet (UV) light for crosslinking. Subsequent to cell lysis, a mild RNase digestion was performed to facilitate downstream processing. The denaturing

affinity-purification on nickel beads ensured that only RNAs covalently linked to proteins were purified. Subsequent to adapter ligation and radioactive labeling, the samples were resolved on an SDS-PAGE and transferred to a nitrocellulose membrane. RNA was detected by autoradiography and isolated. The isolated RNA was reverse-transcribed and amplified by PCR, and the PCR products were sequenced and analyzed (Bohnsack et al. 2012).

The strains Scp160 wildtype, Scp160ΔKH1-2, Scp160ΔKH4-5, Scp160ΔKH5 and Scp160ΔKH13-14 were sent to the Dept. of Molecular Biology at the University Medical Center Göttingen in collaboration with Dr. Kathrine Bohnsack and Dr. Gustavo Nicolás Lemus Díaz.

Contributed Figures: 39-51; Supplementary Figure: 8 (Figures have been adapted accordingly).

Additional materials

Lists of used Enzymes, antibodies, chemicals, consumables and lab equipment can be found in Table 13-Table 17.

Material

Table 2: Plasmids used in this thesis.

Plasmid		
Name	genotype	RJP#
pRS316	URA3, CEN6	148
pYM25	hphNT1, yeGFP	1226
pYM26	klTRP1, yeGFP	1227
pYM29	HIS3MX6, yeGFP	1243
SEC63-RFP-URA3	SEC63-RFP, 2µ plasmid, URA	1398
pFA6a-mCherry-kanMX6	KNOP, for RFP Tagging	1422
pFA6a-mCherry-natMX6	KNOP, for RFP Tagging	1423
pFA6a-mCherry-URA3	KNOP, for RFP Tagging	2344
pFA6a-DSred-URA3	KNOP, for RFP Tagging	2342
pCas-URA3-BsaI	gRNA insertion between two BsaI restriction sites	2222
pCas-HIS3-BsaI	gRNA insertion between two BsaI restriction sites	2223
pFA6a-TRP1	KNOP, For gene deletion with TRP1	2242
pCas-URA3-KH1/5´-BsaI	CRISPR cut at KH domain 1 (5 prime terminus)	2368
pCas-URA3-KH2/5´-BsaI	CRISPR cut at KH domain 2 (5 prime terminus)	2369
pCas-URA3-KH3/5´-BsaI	CRISPR cut at KH domain 3 (5 prime terminus)	2370
pCas-URA3-KH4/5´-BsaI	CRISPR cut at KH domain 4 (5 prime terminus)	2371
pCas-URA3-KH5/5´-BsaI	CRISPR cut at KH domain 5 (5 prime terminus)	2372
pCas-URA3-KH6/5´-BsaI	CRISPR cut at KH domain 6 (5 prime terminus)	2373
pCas-HIS3-KH1/3´BsaI	CRISPR cut at KH domain 1 (3 prime terminus)	2375
pCas-HIS3-KH2/3´BsaI	CRISPR cut at KH domain 2 (3 prime terminus)	2376
pCas-HIS3-KH3/3´BsaI	CRISPR cut at KH domain 3 (3 prime terminus)	2377
pCas-HIS3-KH4/3´BsaI	CRISPR cut at KH domain 4 (3 prime terminus)	2378
pCas-HIS3-KH5/3´BsaI	CRISPR cut at KH domain 5 (3 prime terminus)	2379
pCas-HIS3-KH6/3´BsaI	CRISPR cut at KH domain 6 (3 prime terminus)	2380

Table 3: Primers used in this thesis.

Primer		
name	sequence	RJO#
scp160_check	GAACGTCTAAGTACACAACAGC	2356
scp160_ko_HIS3MX6 fw S1	TAAAATATACTTCCCACACCCCCTCCTTCCATTATAACTGCAcgtacgctgcaggtcgac	2509
scp160_ko_HIS3MX6 rev S2	GCCAAAATCTATATTGAAAAAAATTGGTTTCAAAGAGCTTGTatcgatgaattcgagctc	2510
SCP160 3´-UTR rev	GTTCGTGTTACTTCAACTTCA	2511
SCP160_HAtag_oligo1	AAGCTGGTGTCGAAAAGGCCGGTGAAATGGTTTTGAAATCCTTAAGAAGAcgtacgctgcaggtcga	2566
SCP160_HAtag_oligo2	GTAAAAGCCAAAATCTATATTGAAAAAAATTGGTTTCAAAGAGCTTGTatcgatgaattcgagctcg	2567
Scp160_select_2	CTCGCTGTTTCGTCGTCTTCT	2579
Scp160_seq4	AGGTGGCCTTCTTGTTCAG	2668
Scp160_seq5	AGATTTTTGGCATGGGTCAG	2671
Scp160_seq7	TCACGAATAGTTCTACCACCTG	2673
SCP160F1	GAAGCTCTCGATACTGCTGTTA	2863
Scp160deltaKH13/14_Knop_S3	AATGTCGAGAAAGCTGAGAAGAAAATCTTGAATGAAATAATCAGGGAAcgtacgctgcaggtcgac	3342
N-term mCherry rev	GTCTGGGTGCCCTCGTAG	4927
In EGFP Seq primer	CACATGGTCCTGCTGGAGTTCGTG	5623
yeGFP-Det-R	ATCACCTTCACCTTCACCGGAG	6736
scp160tag_fwd	GGTGTCGAAAAGGCCGGTGAAATGGTTTTGAAATCCTTAAGAAGAcgtacgctgcaggtcgac	7108
scp160tag_rev	AAAGCCAAAATCTATATTGAAAAAAATTGGTTTCAAAGAGCTTGTatcgatgaattcgagctc	7109
SCP160_KH1	GGCTTTCACTTTGGACCTGC	7164
SCP160_KH2	CCCATCCAAATGTAAAGCCAG	7165
SCP160_KH3	GATGTTTCCGAGTTTGCCTCC	7166
SCP160_KH4	GGTCAAGATTCCATCCAAGTTC	7167
SCP160_KH5	GGTCTTGAAGAATCTCATCC	7168

BamHI_P_Scp160_1	CGCAAAGGATCCCATTGATCCTTCTTTCTCATTCC	7187
Scp160_T_HindIII_2	CGCAAAaagcttAGTTATGCTTTTCACCGCC	7188
scp_tag_dupl_fwd	GCTGGTGTCGAAAAGGCC	7232
scp_tag_dupl_rev	GCTGTTTCGTCGTCTTCTCC	7233
scp_KNOP_alt_rev	AAAGCCAAAATCTATATTGAAAAAAATTGGTTTCAAAGAGCTTGTtcatcgatgaattcgagctcg	7251
SCP160-HTP fwd	GGTGTCGAAAAGGCCGGTGAAATGGTTTTGAAATCCTTAAGAAGAcaccatcaccatcaccatga	7668
SCP160-HTP rev	AAAGCCAAAATCTATATTGAAAAAAATTGGTTTCAAAGAGCTTGTtacgactcactatagggcga	7669
CDC33-mAID fwd	CATTCCAGTGCCAATGGTAGACACCCTCAACCATCAATCACCTTGCGTACGCTGCAGgtcgac	7670
CDC33-mAID rev	TTTTGATTAAAATAACAATTATCTTAAGAAAAATTCAGACTATCATCGATGAATTCGAGCTCG	7671
HISx6 fwd	tttaaaCCCGGGCACCATCACCATCACCATGAATTCtttaaa	7672
HISx6 rev	tttaaaGAATTCATGGTGATGGTGATGGTGCCCGGGtttaaa	7673
cbk1tag_fwd2	TACACTTACTCCAGATTTGACTATTTGACAAGAAAAAATGCGTTGcgtacgctgcaggtcga	7675
cbk1tag_rev2	CCATAGATAAATACTTGAATAAAGAGGAATGTCCTTAACGCGTCCatcgatgaattcgagctcg	7676
KH1 frw 5´ bRA66	GCGTTCAAAGAACATCCAAGGTTTT	7677
KH1 rev 5´ bRA66	CTTGGATGTTCTTTGAACGCGATCA	7678
KH2 frw 3´ bRA66	TAAAATCTTGGCCATTGTTAGTTTT	7679
KH2 rev 3´ bRA66	TAACAATGGCCAAGATTTTAGATCA	7680
AID-GCN4_fwd	TTGGAAAATGAGGTTGCCAGATTAAAGAAATTAGTTGGCGAACGCcggatccccgggttaattaa	7681
AID-GCN4_rev	TACACGAGAATGAAATAAAAAATATAAAATAAAAGGTAAATGAAAgaattcgagctcgtttaaac	7682
AID-CDC33_fwd	CATTCCAGTGCCAATGGTAGACACCCTCAACCATCAATCACCTTGcggatccccgggttaattaa	7683
AID-CDC33_rev	ATTTTGATTAAAATAACAATTATCTTAAGAAAAATTCAGACTATCgaattcgagctcgtttaaac	7684
GCN4_contr_fwd	ACCGTTACGGAAACATCTTG	7685
GCN4_contr_rev	AATACCAGAACATACGGCAG	7686

myc_XmaI_fwd	ccggGAAGCTTATTTCTGAAGAAGACTTGTAA	7687
myc_XhoI_rev	tcgaTTACAAGTCTTCTTCAGAAATAAGCTTC	7688
Trm2_BamHI_fwd	CGCAAAGGATCCATGTACGAACAGTTTGAATT	7689
Trm2_PstI_rev	CGCAAACTGCAGTTAGATTCTCTTCATTATACACAC	7690

Table 4: Strains created/used in this thesis.

background	**deletion**	**RJY#**	**origin**	**Scp160-tag**
BY4741	KH1-2	5842	2049	
BY4741	KH1-2	5807	2049	
BY4741	KH1-2	5808	2049	
BY4741	KH1-3	5843	2049	
BY4741	KH1-3	5957	2049	
BY4741	KH1-3	5958	2049	
BY4741	KH1-4	5844	2049	
BY4741	KH1-4	5959	2049	
BY4741	KH1-4	5960	2049	
BY4741	KH1-5	5845	2049	
BY4741	KH1-5	5962	2049	
BY4741	KH1-5	5963	2049	
BY4741	KH2-3	6061	2049	
BY4741	KH2-3	6062	2049	
BY4741	KH2-3	6063	2049	
BY4741	KH3-4	5846	2049	
BY4741	KH3-4	6073	2049	
BY4741	KH3-4	6074	2049	
BY4741	KH4-5	5847	2049	
BY4741	KH4-5	5980	2049	
BY4741	KH4-5	5981	2049	
BY4741	KH5-6	5848	2049	
BY4741	KH5-6	5982	2049	

BY4741	KH5-6	5983	2049	
BY4741	KH3-5	5849	2049	
BY4741	KH3-5	6071	2049	
BY4741	KH3-5	6072	2049	
BY4741	KH2	6065	2049	
BY4741	KH2	6066	2049	
BY4741	KH2	6067	2049	
BY4741	KH5	6064	2049	
BY4741	KH5	5978	2049	
BY4741	KH5	5979	2049	
BY4741	KH6	6068	2049	
BY4741	KH6	6069	2049	
BY4741	KH6	6070	2049	
BY4741		5850	2049	yeGFP
BY4741	KH13-14	5851	2049	yeGFP
BY4741	KH1-2	5852	5842	yeGFP
BY4741	KH1-3	5853	5843	yeGFP
BY4741	KH1-4	5854	5844	yeGFP
BY4741	KH1-5	5855	5845	yeGFP
BY4741	KH2-3	5987	6061	yeGFP
BY4741	KH3-4	5856	5846	yeGFP
BY4741	KH4-5	5857	5847	yeGFP
BY4741	KH5-6	5858	5848	yeGFP
BY4741	KH3-5	5859	5849	yeGFP
BY4741	KH2	6075	6065	yeGFP
BY4741	KH5	6076	6066	yeGFP
BY4741	KH6	6077	6067	yeGFP
BY4741		5906	2049	HTP
BY4741	KH1-2	5907	5842	HTP
BY4741	KH4-5	5912	5847	HTP

BY4741	KH5	6078	6064	HTP
BY4741	KH13-14	5915	2049	HTP

Table 5: List of amino acids used in this thesis.

Amino Acids	per 600ml
Adenine hemisufate	20mg
Arginine HCl	20mg
Histidin HCl*	20mg
Isoleucine	20mg
Leucine	40mg
Lysine HCl	40mg
Methionine*	40mg
Phenylalanine	40mg
Serine	40mg
Threonine	40mg
Tryptophan*	30mg
Tyrosine	30mg
Uracil*	30mg
Valine	12mg

Table 6: Recipe of polyacrylamide gels used in this thesis.

Seperating gel:		# of gels				*Unit*
		1	2	3	4	
8% polyacrylamid gels	miliQ-H_2O	3,25	6,5	9,75	13	*ml*
	1,5M TRIS pH 8,8	1,75	3,5	5,25	7	*ml*
	10% SDS	70	140	210	280	*µl*
	30% Acrylamide	1,88	3,76	5,64	7,52	*ml*
	10% APS	70	140	210	280	*µl*
	TEMED	7	14	21	28	*µl*

Stacking gel:		# of gels				*Unit*
		1	2	3	4	
6% polyacrylamid gels	miliQ-H_2O	2,65	5,3	7,95	10,6	*ml*
	0,5M TRIS pH 6,8	1,25	2,5	3,75	5	*ml*
	10% SDS	50	100	150	200	*µl*
	30% Acrylamide	1	2	3	4	*ml*
	10% APS	50	100	150	200	*µl*
	TEMED	5	10	15	20	*µl*

Table 7: Recipe of FACS buffer.

FACS buffer	Tris pH 7,5	200mM
	NaCl	211mM
	$MgCl_2$	78mM

Table 8: Recipe of FACS-PI buffer.

FACS-PI	Tris pH 7,5	180mM
	NaCl	190mM
	$MgCl_2$	70mM
	propidium iodide	100µg/ml

Table 9: Recipe of Laemmli buffer (5x).

Laemmli 5x	Tris-HCl pH 6.8	250mM
	Glycerol	25%
	SDS	10%
	bromophenol blue	0.25%
	DTT	0.5M

Table 10: Recipe of lysis buffer II.

Lysis buffer II	Tris pH 7,5	50mM
	NaCl	100mM
	$MgCl_2$	1,5mM
	NP40	0,15%
	DTT	1mM

Table 11: Recipe of Dilution buffer.

Dilution buffer	Tris pH 7,5	10mM
	NaCl	150mM
	EDTA	0,5mM

Table 12: Recipe of wash buffer.

Wash buffer	Tris pH 7,5	10mM
	NaCl	150mM
	EDTA	0,5mM
	NP40	0,05%

Table 13: List of Enzymes used in this thesis.

Enzyme	Company
Herculase® II Fusion DNA Polymerase	Agilent
RNase A	Roche
T4 DNA Ligase	Thermo
Restriction endonucleases	NEB/Thermo
RQ1 DNaseI	Quiagen

Table 14: List of antibodies used in this thesis.

Antibody	organism	Dilution	Company
Scp160	rabbit	1:3.000	gift from Matthias Seedorf
HIS	mouse	1:3.000	Qiagen/Sigma
GFP	chicken	1:5.000	Abcam
GFP (IgG1κ)	mouse	1:1.000	Fisher Scientific
anti-mouse-IgG-HRP	sheep	1:5.000	Dianova Jackson
anti-rabbit-IgG-HRP	goat	1:5.000	Millipore/Amersham
anti-chicken-IgG-HRP	goat	1:5.000	Abcam

Table 15: List of chemicals used in this thesis.

Chemicals	Company
2-mercaptoethanol	Roth
Acetic acid	Roth
Agarose	Roth
Bromophenol blue	Roth
BSA	Roche
CHX	Roth
DNA ladder	Fermentas
DTT	Merck
EDTA	Merck
Ethanol	Merck
Gel red	Merck
Glycerin	Roth
Glycine	Merck
HCl	Merck
Isopropanol	Merck
KCl	Merck
KH2PO4	Roth
LiAc	Roth
$MgCl_2$	Roth
Milk powder	Roth
Na_2HPO_4	Roth

NaCl	Merck
NP40	Applichem
PEG 4000	Roth
Propidium iodide	Merck
SDS	Merck
ssDNA	Biomol
TEMED	PanReacApplich
TRIS	Roth
Triton X-100	Roth

Table 16: List of consumables used in this thesis.

Consumables	Company
Pipette tips 1-10µl	Sarstedt
Pipette tips 2-200µl	Sarstedt
Pipette tips 100-1000µl	Sarstedt
Pipette tips 1-10µl (filter tip)	Nerbe plus
Pipette tips 2-200µl (filter tip)	Nerbe plus
Pipette tips 100-1000µl (filter tip)	Nerbe plus
PCR tubes	Nerbe plus/Quiagen
1 ,5ml microcentrifuge cups	Sarstedt
2 ml microcentrifuge cups	Sarstedt
1 ,5ml microcentrifuge cups (safeseal)	Sarstedt
2 ml microcentrifuge cups (safeseal)	Sarstedt
15ml falcons	Greiner
50ml falcons	Sarstedt
Glass beads 0,25-0,5mm	Roth
PageRuler Pre-Stained Protein Ladder	Li-Cor
PVDF membranes	Cytiva
Plastic cuvettes	Sarstedt
Whatman papers	Roth
GFP-Trap® Magnetic Agarose	ChromoTek
Inoculation loops	Sarstedt
Gel and PCR clean up kit	Macherey-Nagel/Quiagen

Table 17: List of equipment used in this thesis.

Equipment	Company
1-10µl pipette	Eppendorf
20-200µl pipette	Eppendorf
100-1000µl pipette	Eppendorf
Centrifuge 5415 R	Eppendorf
Pipetboy acu	IBS
Cell density meter model 40	Fisher Scientific
Mini PROTEANT Tetra Cell	BioRad
Shaker gyro rocker SSL3	Stuart
Photometer Genesys 10 Bio	Thermo
ChemiDoc MP	BioRad
Nanodrop one	Thermo scientific
PCR cycler	Bio-Center/BioRad
Decon DeVision DB0X gel documentation	Decon
Gel electrophoresis chamber	Peqlab
Tube roller RS-TR05	Phoenix
Power supply	BioRad
Vortex mixer	Benchmark
pH meter	Mettler Toledo
Heat mixing block	Bioer
Kimble Kimax media bottle	Merck
Micro Balance	Sartorius
Balance PLS 2100-2	Kern

Results

Establishing a set of yeast strains expressing truncated or deletion variants of Scp160

The manipulation or deletion of the KH domains in vigilin to analyze the proteins' function is a long-established method (Baum et al., 2004; Brykailo et al., 2007; Hirschmann et al., 2014; Zhou et al., 2008). However, the majority of the KH domains utilized in these experiments were located at the C-terminal domains, with minimal inclusion of the N-terminal KH domains. It should be noted that not all experiments concerning KH domain deletion were conducted under optimal conditions. In some cases, vigilin proteins were expressed in the genome under the endogenous promoter (Zhou et al., 2008). The objective of subsequent experiments was to express all modified proteins in the endogenous locus and under the endogenous promoter. To analyze the N-terminal KH domains of Scp160, new strains lacking specific KH domains were created (Table 18). The deletion of central KH domains has historically been a challenging undertaking due to the lack of suitable methodologies. The advent of the CRISPR/Cas9 system as a means of freely and easily modifying a gene sequence has opened the possibility of generating new kinds of Scp160 deletions and truncations. Plasmids for CRISPR/Cas9 editing were cloned to enable the truncation/deletion of all necessary combinations, as detailed in the Material and Methods section. To ensure efficient genome editing, two plasmids were created for each truncation/deletion: one plasmid for 5' cuts and one plasmid for 3' cuts of the target deletion area. Yeast strains expressing Scp160 versions lacking KH domains were tested for phenotypes previously associated with loss of Scp160, including ploidy change (Cheng et al., 2018; Wintersberger et al., 1995). As observed by various researchers, the effects of KH domain deletion include alterations in cell size (Li, 2004; Wintersberger et al., 1995), changes in cell wall properties (Nerome et al., 2020), and localization at the endoplasmic reticulum (Baum et al., 2004; Li, 2004; Wintersberger et al., 1995). The strategy employed to identify the functionally important KH domains was as follows: initially, truncations were made, commencing with the first two and concluding with the deletion of the first KH domains. These truncations were employed primarily to indicate areas of interest. Subsequently, the deletion areas were reduced to smaller regions, thereby facilitating the identification of more specific areas of interest. For further analysis, the strains were tagged with yeGFP for microscopy and His6-TEV-Protein A (HTP) for studies on RNA binding. C-terminal tagging of Scp160 with GFP has been performed on numerous occasions over the past few decades, thereby proving the functionality of the tagged protein (Brykailo et al., 2007; Frey et al., 2001; Gelin-Licht et al., 2012; Li, 2004; Weidner et al., 2014). The tagged Scp160 variants were subsequently employed in

a series of further experiments, including ploidy analysis, the localization of mutant proteins, emulsion assays, pulldown assays to analyze protein-protein interaction (PPI), and protein-RNA interaction.

Table 18: Overview of relevant strains; column one: Scp160 lacking specific KH domains; column two: Yeast strains expressing Scp160 variants in the yeast background BY4741; column three and four: tagged versions of the corresponding strain with yeast enhanced GFP (yeGFP) and His6-TEV-Protein A (HTP) tag, if available.

strain	expression	Variant with yeGFP tag	Variant with HTP tag
Scp160ΔKH1-2	✓	✓	✓
Scp160ΔKH1-3	✓	✓	✓
Scp160ΔKH1-4	✓	✓	✓
Scp160ΔKH1-5	✓	✓	✓
Scp160ΔKH2-3	✓	✓	-
Scp160ΔKH3-4	✓	✓	✓
Scp160ΔKH4-5	✓	✓	✓
Scp160ΔKH5-6	✓	✓	✓
Scp160ΔKH3-5	✓	✓	✓
Scp160ΔKH2	✓	✓	-
Scp160ΔKH5	✓	✓	✓
Scp160ΔKH6	✓	✓	-
Scp160ΔKH13-14	✓	✓	✓

All strains listed were produced using a duo-plasmid CRISPR/Cas9 system, as detailed in the Materials and Methods section. The efficacy of successful genomic modification exhibited variability contingent on the utilized gRNAs. In theory, a single CRISPR/Cas9 plasmid could have been sufficient to achieve the required deletion of a specific DNA area. The decision to utilize two plasmids was driven by the limited availability of suitable gRNAs for specific KH domains. In the case of certain domains, such as KH domains 3 and 4, the efficacy of the available gRNAs was found to be relatively low, resulting in suboptimal outcomes with regard to successful deletion events. The combination of two plasmids containing two different gRNAs has been demonstrated to enhance the efficiency of the system for the majority of plasmid pairings. In a few cases, the efficiency of the gRNAs was inadequate for

successful deletion. Even with two plasmids, the required KH domain deletion was rarely achieved or not possible at all. The KH domains three and four proved particularly problematic in this regard.

The success of truncations or deletions was initially validated through the extraction of genomic DNA and PCR analysis, employing primers that spanned the targeted deletion region. Furthermore, a second PCR was conducted with one primer within and one primer outside the area to be deleted (Figure 6).

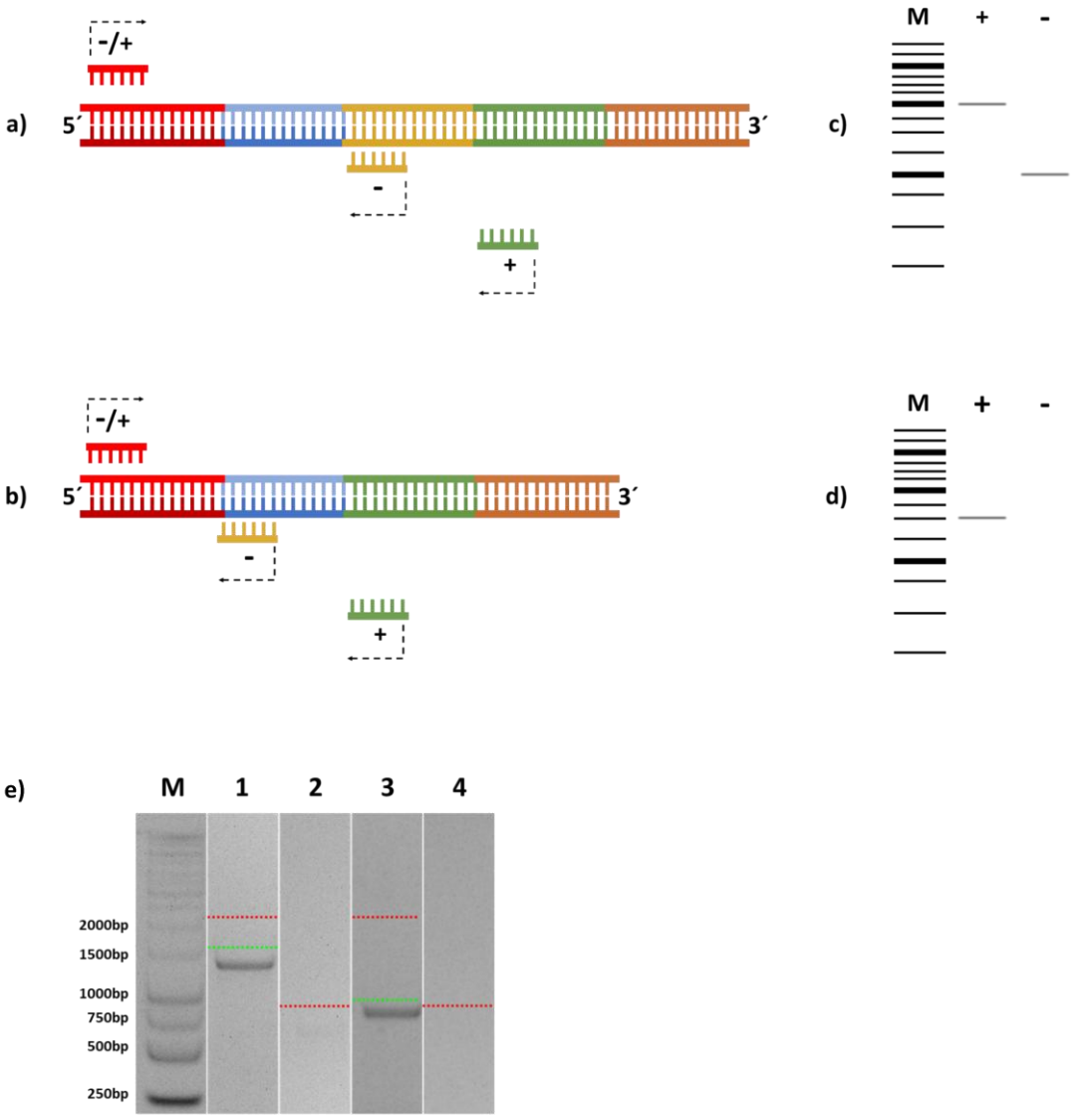


Figure 6: Verification of successful deletion; a) and b) DNA fragments of interest; Short ssDNA fragments: Primer; red: N-terminal flanking area and primer binding in this domain, +/-: primer used for both positive control and success control; yellow: area to be deleted and primer binding in this domain, -: complementary primer used for success control; green C-terminal flanking area and primer binding in this domain,+: complementary primer used for positive control; c) and d) band pattern of PCR: M: marker, +: band for successful truncation/deletion event; -: band for unsuccessful truncation/deletion event; c) example of band pattern for wildtype/unsuccessful truncation/deletion; d) example of band pattern for successful truncation/deletion; e) Agarose gel of exemplary truncations/deletions; dashed line: green: band height belonging to successful truncation/deletion event; dashed line: red: band height belonging to unsuccessful truncation/deletion event 1 and 2: Scp160ΔKH1-2; line 1: band corresponding to an successful truncation/deletion runs at approx. 1500bp; wildtype allele runs at approx. 2000bp, line 2: control with primer annealing in the deleted area which would result in an approx. 750bp fragment. Line 3 and 4: Scp160ΔKH1-5, line 3: band corresponding to a successful truncation/deletion runs at approx. 800bp; wildtype allele runs at approx. 2000bp, line 4: control with primer annealing in the deleted area which would result in an approx. 750bp fragment. DNA bands were stained by GelRed©.

To ascertain whether truncation or deletion of KH domains would result in increased protein degradation or loss of translation, the expression of mutant proteins was analyzed by western blot analysis (Figure 7). All strains created in this thesis express either the full-length or a variant of Scp160 lacking specific KH domains. The expression of all truncated/KH-domain deleted proteins was

observed to occur at a similar rate as that of the full-length Scp160p. The discrepancies in the observed molecular weight shifts are indicative of the loss of KH domains in the corresponding protein. Subsequently, strains expressing Scp160 variants with KH domain deletions were tagged with either a yeast green fluorescent protein (yeGFP) tag for fluorescence microscopy or an HTP tag for RNA-protein crosslinking analysis. The tagging was confirmed via genomic DNA extraction and PCR. Subsequently, the yeGFP-tagged colonies were examined under a microscope to ascertain the presence of a GFP signal.

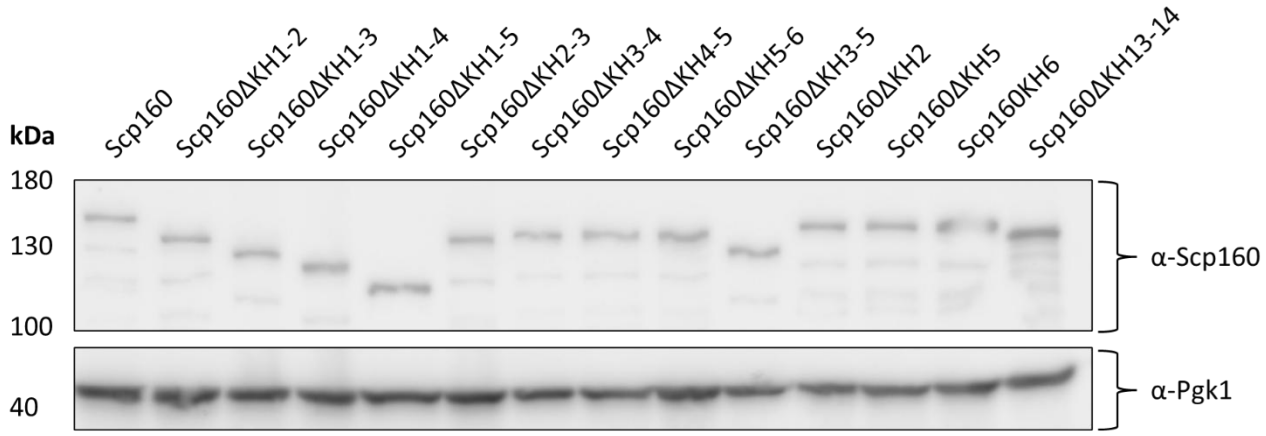


Figure 7: Western blot of Scp160 and KH domain deleted Scp160 proteins (total cell lysate); upper part was incubated with α-Scp160 antibody, lower part was incubated with α-Pgk1 antibody; (see material and methods).

KH domain truncation/deletion in Scp160 affects the ploidy profiles in a nonlinear manner

It is established that yeast cells lacking *SCP160* exhibit a known phenotype, namely an aberration of DNA content or ploidy. (Wintersberger et al., 1995). Following the loss of *SCP160*, a population of *scp160Δ* cells exhibit an increasing accumulation of (pseudo-)diploid or polyploid cells (Cheng et al., 2018; Hirschmann et al., 2014; Weber et al., 1997; Wintersberger et al., 1995). As DNA content or ploidy can be readily analyzed by fluorescence-activated cell sorting (FACS), it serves as an optimal initial analytical tool for evaluating the loss of function of the generated truncation/deletion mutants. The initial objective was to identify KH domains of interest. These areas would be distinguished by a change in their ploidy profile, which would reflect that observed in the deletion of the entire *SCP160* gene. The initial strains subjected to analysis were KH domain truncations with progressively increasing accumulating deletions in the genetic background of W303 (https://www.yeastgenome.org/strain/w303). As previously described, the impairment of Scp160 can

result in ploidy changes, leading to the formation of aneuploid cells (Chial et al., 1999; Mendelsohn, 2003; Wintersberger et al., 1995). Accordingly, the term "pseudo" is employed for the purpose of describing ploidy changes. However, inconsistencies were observed in the ploidy changes exhibited by individual mutants. For example, while certain clones of a mutant expressing Scp160ΔKH5-6 exhibited a pseudo-tetraploid genome, others displayed a (pseudo-) diploid phenotype (Supplementary Figure 1). In other instances, the resulting ploidy of a mutant exhibited temporal alterations, as observed in cells expressing Scp160ΔKH1-3. Given the potential influence of the W303 genetic background on the observed instability, a switch to the BY4741 genetic background (Brachmann et al., 1998) was undertaken. The BY4741 strains expressing truncated/KH domain deleted versions of Scp160 demonstrated a stable ploidy profile, with no observable changes over time. To ensure the reliability of the FACS results, a haploid wild-type, a diploid wild-type, and the scp160Δ strains were employed as controls. All strains were analyzed in triplicate.

In the series of Scp160 variants lacking KH domains, the deletion of the first two (Scp160ΔKH1-2), two to three (Scp160ΔKH2-3) KH domains, as well as KH domain two alone (Scp160ΔKH2) resulted in no alteration of the originally observed haploid phenotype (Supplementary Figure 1). The deletion of the first three KH domains (Scp160ΔKH1-3) resulted in the emergence of a (pseudo-)diploid phenotype (Supplementary Figure 1). The same observation was made with the deletion of KH domains one to four (Scp160ΔKH1-4), three to four (Scp160ΔKH3-4), and KH domain six (Scp160ΔKH6) alone. The third observed ploidy change, designated as (pseudo-)tetraploid, was observed in strains lacking the KH domains 1-5 (Scp160ΔKH1-5), 3-5 (Scp160ΔKH3-5), and 4-5 (Scp160ΔKH4-5), as well as in the complete deletion of *SCP160* (Supplementary Figure 1).

Two strains exhibited a ploidy profile that underwent alterations over time or between individual clones. The strains expressing a Scp160 lacking the KH domains five to six (Scp160ΔKH5-6) and five (Scp160ΔKH5) both exhibited a ploidy profile comprising two (pseudo-)tetraploid and one (pseudo-) diploid profiles (Supplementary Figure 1).

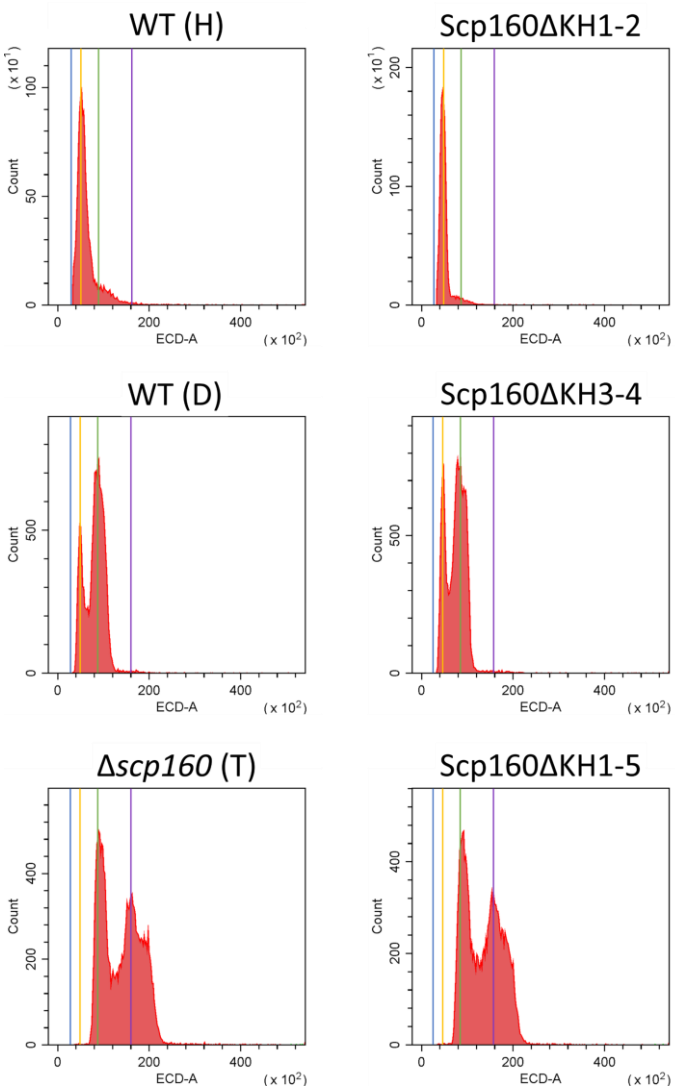


Figure 8: Example for profiles assigned to specific ploidy; blue line: haploid peak, yellow: diploid peak, green: tetraploid peak; violet: octoploid; (H): haploid, (D): diploid, (T): tetraploid; cells were fixed while in exponential growth (in mitosis) therefor chromosome amount is doubled.

For the sake of convenience, the ploidy has heretofore been designated as pseudo-haploid, -diploid, or -tetraploid. However, it was not always possible to assign the ploidy profiles to a specific ploidy type, such as diploid or tetraploid. The deletion of the KH domains one to five (Scp160ΔKH1-5) resulted in a change of the ploidy profile from haploid to a phenotype that is definitively higher than diploid but does not reflect the complete tetraploid phenotype of *scp160Δ* (Figure 9).

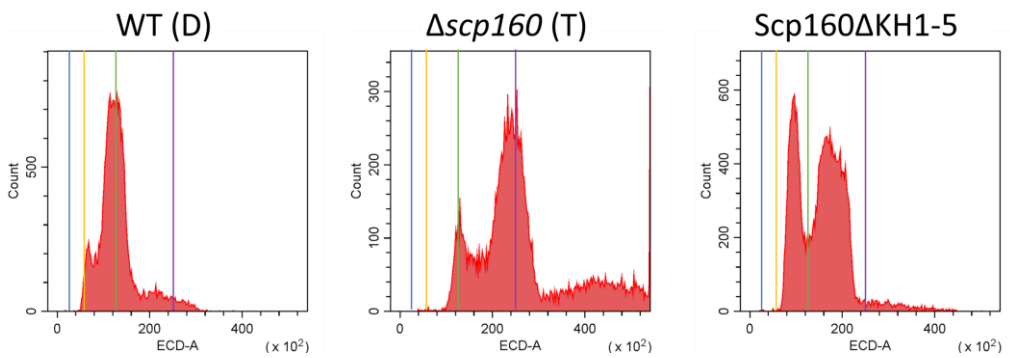


Figure 9: Example for profile not attributed to specific ploidy; blue line: haploid peak, yellow: diploid peak, green: tetraploid peak; violet: octoploid; cells were fixed while in exponential growth (in mitosis) therefor chromosome amount is doubled.

Thus far, I have demonstrated that ploidy changes occur upon deletion of N-terminal KH domains, thereby confirming the validity of this method for a preliminary screening of the generated strains. In the W303 background, shifts from a haploid to a pseudo-diploid phenotype were observed (Supplementary Figure 2).

Alteration of KH domain composition affects the localization of Scp160 proteins at the endoplasmic reticulum

As previously described in Frey et al. (2001), Gelin-Licht et al. (2012b), Lang (2001), Weber et al. (1997), and Wintersberger et al. (1995), Scp160 is localized at the ER. The localization of Scp160 to the ER has been proposed to be dependent on the presence of mRNA (Lang, 2001) and microtubules (Frey et al., 2001). The deletion of the C-terminal KH domains 13 and 14 has been demonstrated to abolish the localization to polysomes, including both ER-bound and cytosolic ones (Baum et al., 2004; Brykailo et al., 2007). Furthermore, the loss of KH domain 14 alone has been shown to result in delocalization of the protein (Li, 2004). A similar outcome was observed in an N-terminal truncation where the initial 74 amino acids were absent (Li, 2004). The 74 amino acids in question contain a proposed NES and are located before the first KH domain. This prompted me to investigate the impact of the KH deletion and truncation mutants I had generated on ER localization. To this end, the Scp160 variants lacking KH domains were tagged at the C-terminus with the fluorescent reporter protein yeast enhanced green fluorescent protein (yeGFP). Each strain carrying a tagged deletion/truncation variant of Scp160 was analyzed by fluorescence microscopy, employing a mCherry-tagged Sec63 or Elo3 protein as an ER marker (David et al., 1998; Young, 2001). The localization patterns of the Scp160 mutant proteins could be classified into three groups. The first group comprises those proteins for which the localization of Scp160 to the ER is unquestionably visible (Figure 10 a). In the second group there was no discernible

localization of Scp160 to the ER (Figure 10 b)), and in the third group the frequency of localization to the ER was low.

Mutant proteins that fall into group 1 and share the same localization pattern as the wildtype protein (Figure 11) include Scp160ΔKH1-2 (Figure 12), Scp160ΔKH1-3 (Figure 13) and Scp160ΔKH2-3 (Figure 16).

Strains showing a group 2 pattern (cytosolic localization) are Scp160ΔKH1-4 (Figure 14), Scp160ΔKH1-5 (Figure 15), Scp160ΔKH4-5 (Figure 18) and Scp160ΔKH3-5 (Figure 20).

The strains with a low frequency of ER localization (group 3), in comparison to the wildtype, are Scp160ΔKH3-4 (Figure 17), Scp160ΔKH5-6 (Figure 19), Scp160ΔKH5 (Figure 23), and Scp160ΔKH13-14 (Figure 21). It should be noted that the frequency observed in the four Scp160 variants is not uniform. The Scp160ΔKH3-4 variant exhibited a localization rate of approximately 30%. The removal of the KH domains 5 and 6 resulted in the delocalization of the protein in approximately 90% of the cells that were counted. The Scp160ΔKH5 variant exhibited an even lower localization rate, with only ~5% of the cells displaying this behavior. The Scp160 variant of Scp160ΔKH13-14 exhibited a localization rate of approximately 18%.

The subsequent figures (Figure 11-Figure 24) illustrate representative fluorescence images of yeast cells expressing various Scp160p variants.

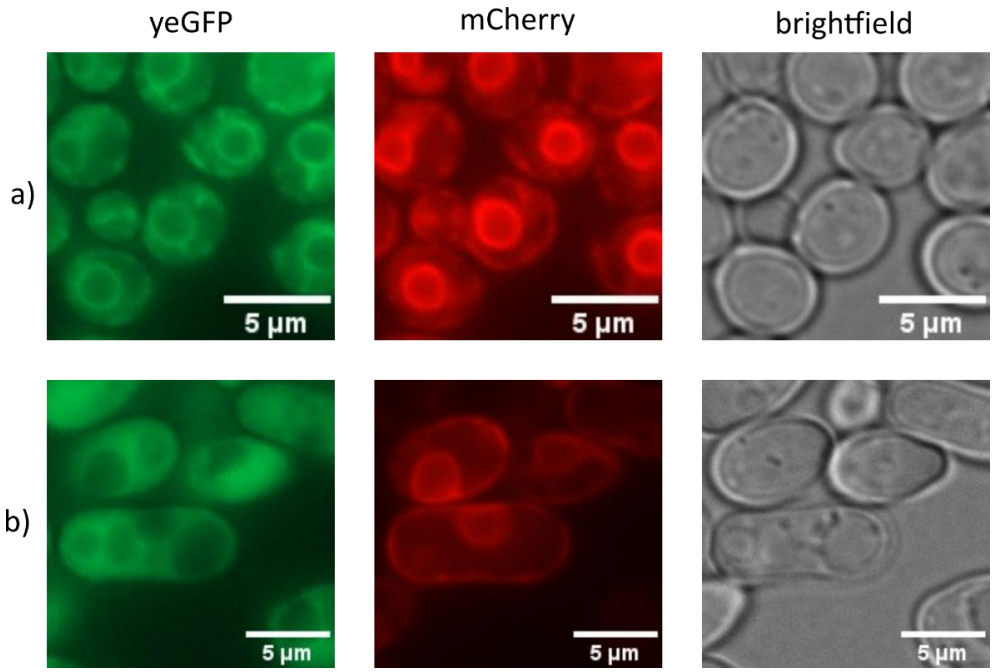


Figure 10: Exemplary microscopy images of Scp160yeGFP protein showing ER (a); group 1 pattern) or cytosolic (b); group 2 pattern) localization; green: GFP channel, red: mCherry channel, grey: brightfield.

The localization pattern for full-length Scp160yeGFP proteins display the typical ring-shaped distribution at the endoplasmic reticulum. This distribution is regarded as the standard against which all Scp160 variants lacking KH domains are evaluated.

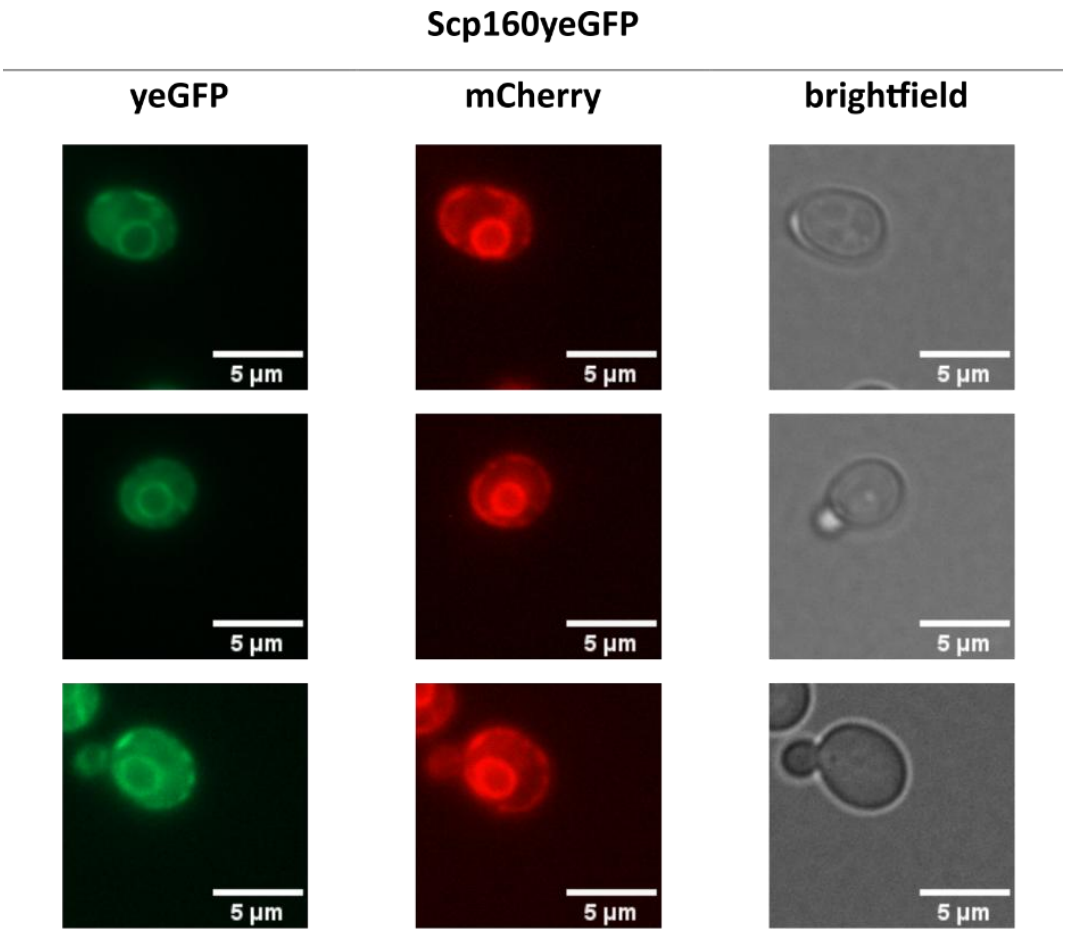


Figure 11: Microscopy images of full-length Scp160yeGFP; green: GFP channel, red: mCherry channel, grey: brightfield.

The ER localization of the truncated Scp160ΔKH1-2yeGFP was not affected (Figure 12). The truncated protein was observed to be localized at the endoplasmic reticulum.

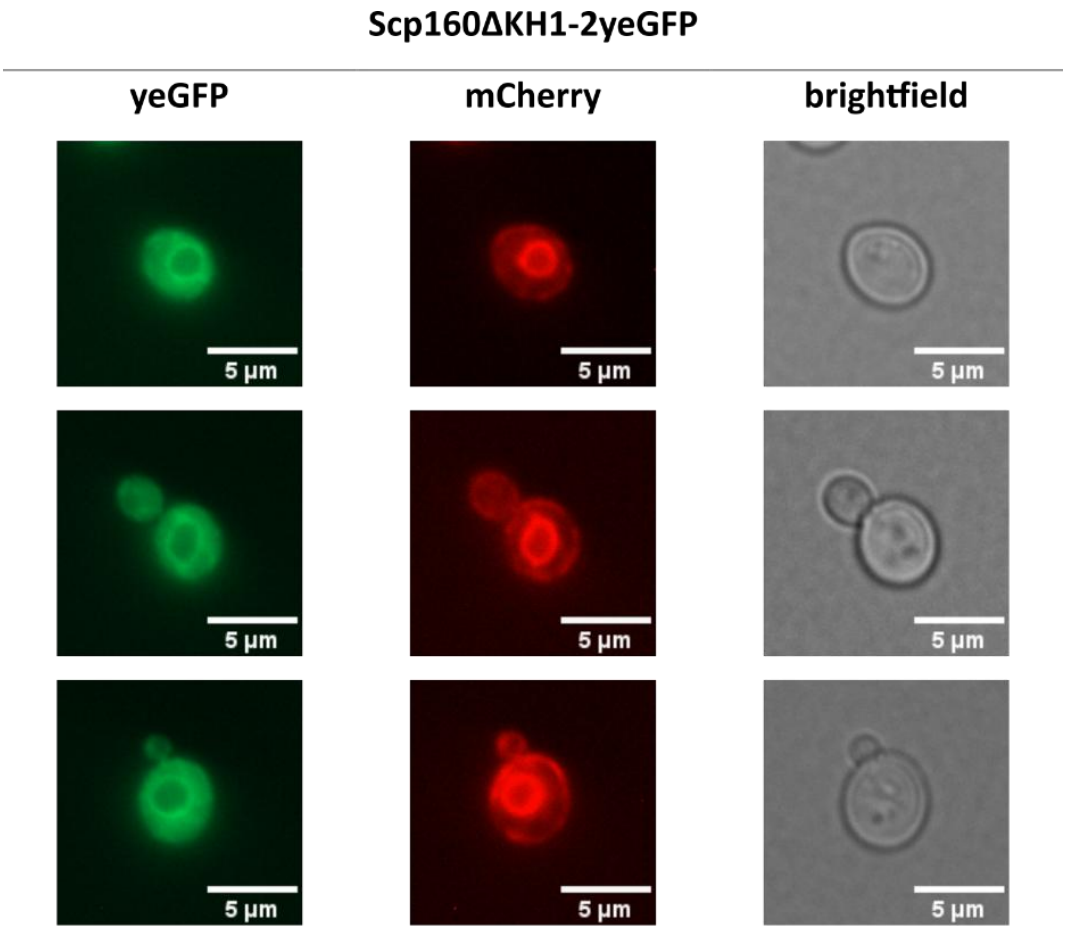


Figure 12: Microscopy images of Scp160yeGFP lacking KH domains one and two; green: GFP channel, red: mCherry channel, grey: brightfield.

The localization of the truncated Scp160ΔKH1-3yeGFP to the endoplasmic reticulum was incomplete (Figure 13). An increase in cytoplasmic staining was observed. Nevertheless, the truncated protein has been demonstrated to be capable of localizing at the ER.

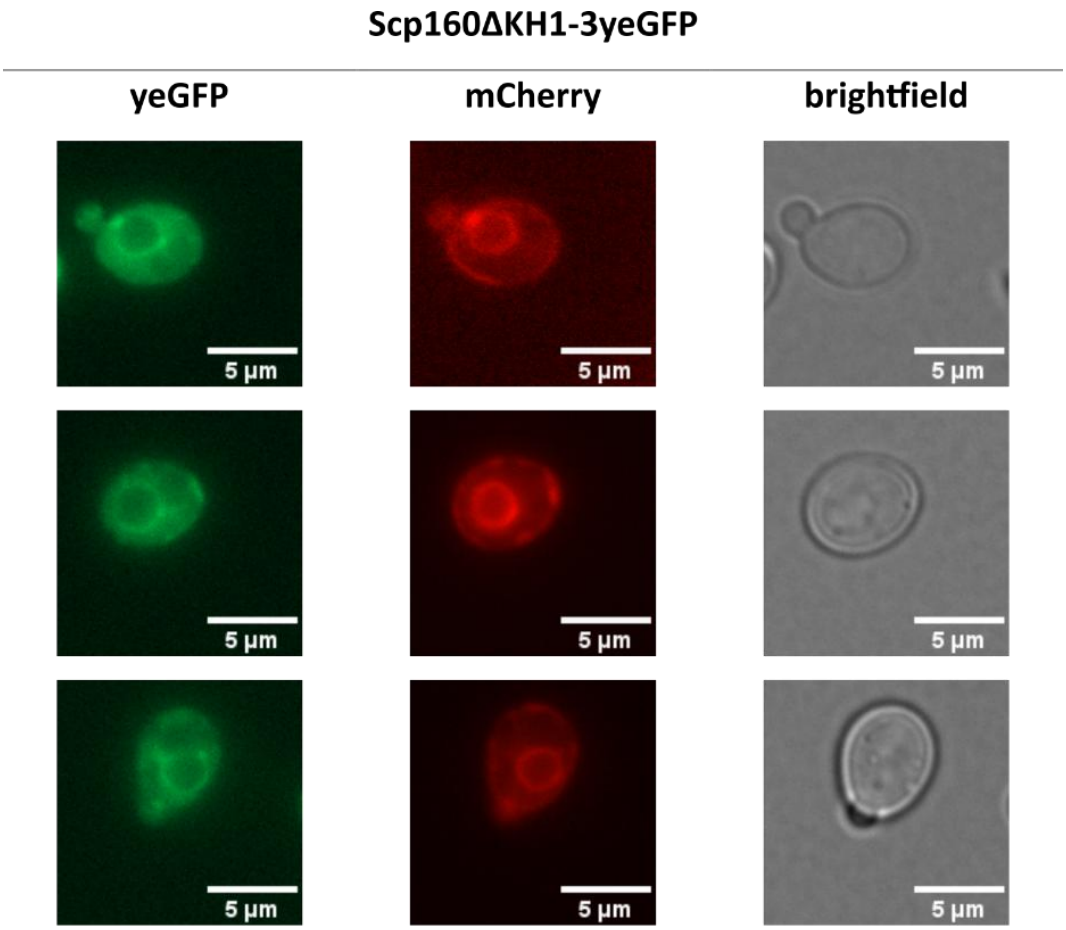


Figure 13: Microscopy images of Scp160yeGFP lacking KH domains one to three; green: GFP channel, red: mCherry channel, grey: brightfield.

The truncation of Scp160ΔKH1-4yeGFP resulted in the mislocalization of Scp160 to the cytosol (Figure 14). The GFP signal was observed to be distributed throughout the cytosol.

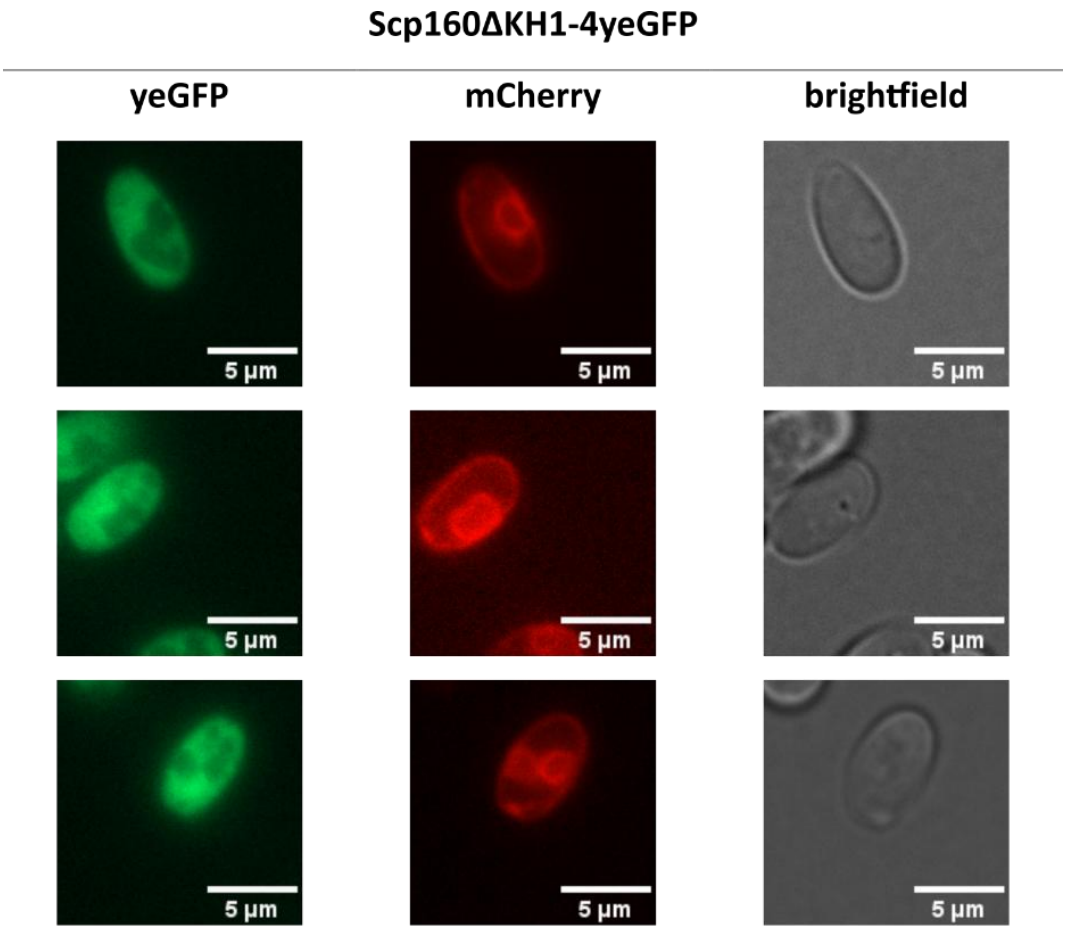


Figure 14: Microscopy images of Scp160yeGFP lacking KH domains one to four; green: GFP channel, red: mCherry channel, grey: brightfield.

The Scp160ΔKH1-5yeGFP mutant protein was observed to lack ER localization (Figure 15). The truncated protein was found to be incapable of localizing to the endoplasmic reticulum. The GFP signal was distributed throughout the cytosol.

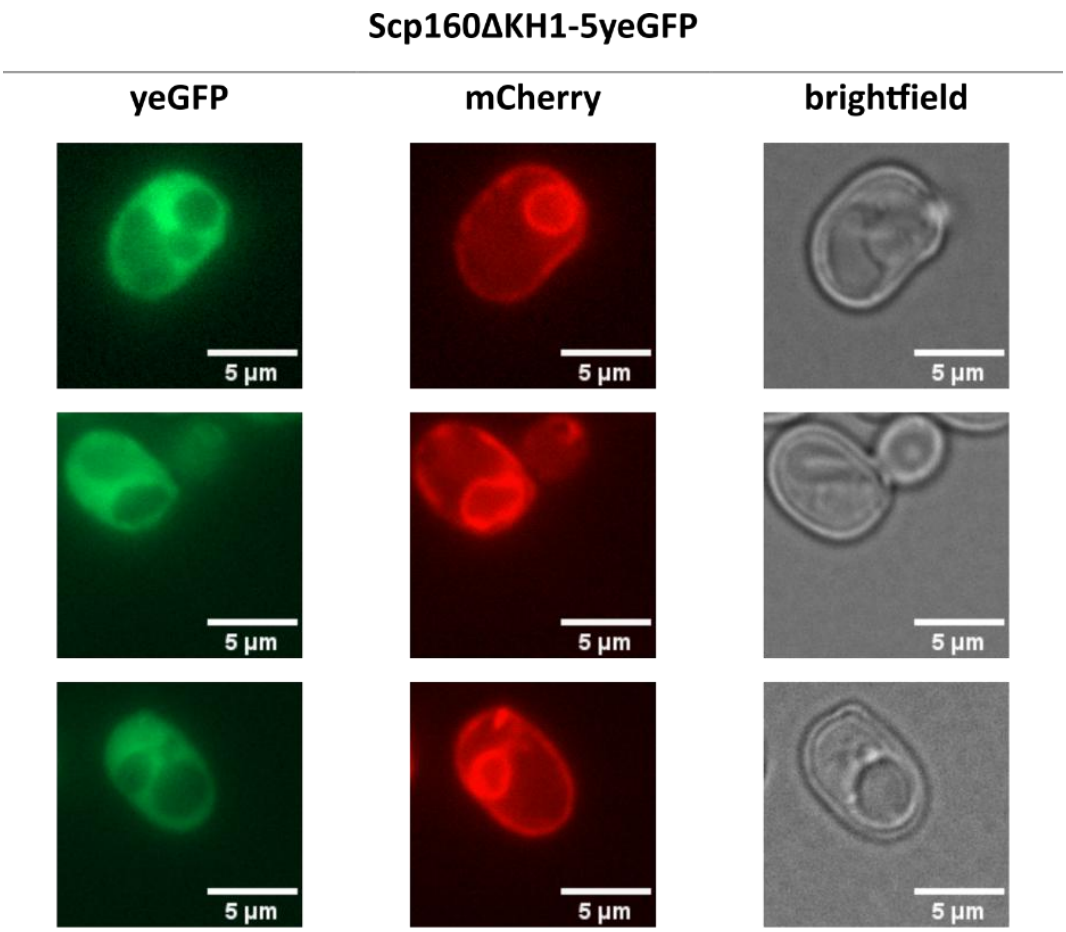


Figure 15: Microscopy images of Scp160yeGFP lacking KH domains one to five; green: GFP channel, red: mCherry channel, grey: brightfield.

As illustrated in Figure 16, ER localization was observed for the Scp160ΔKH2-3yeGFP protein.

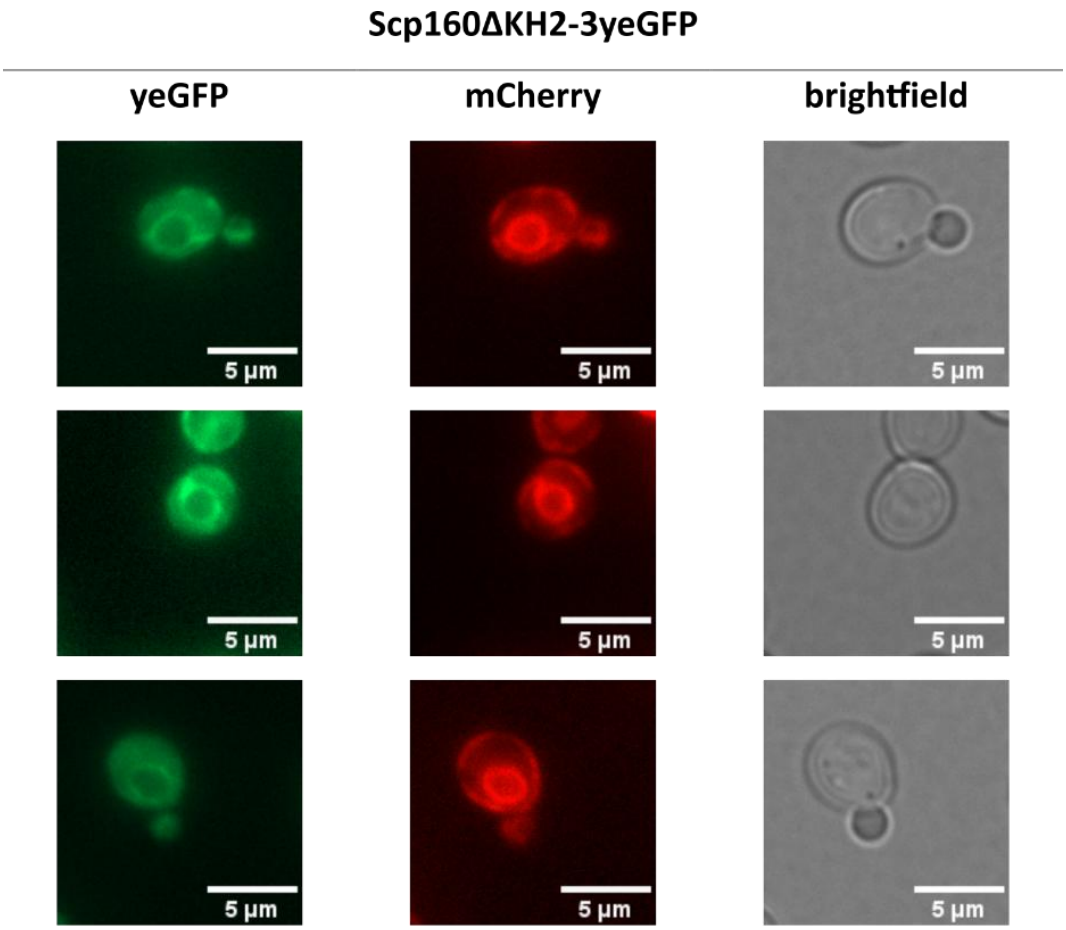


Figure 16: Microscopy images of Scp160yeGFP lacking KH domains two to three; green: GFP channel, red: mCherry channel, grey: brightfield.

Scp160ΔKH3-4yeGFP has been observed to localize to the endoplasmic reticulum (Figure 17).

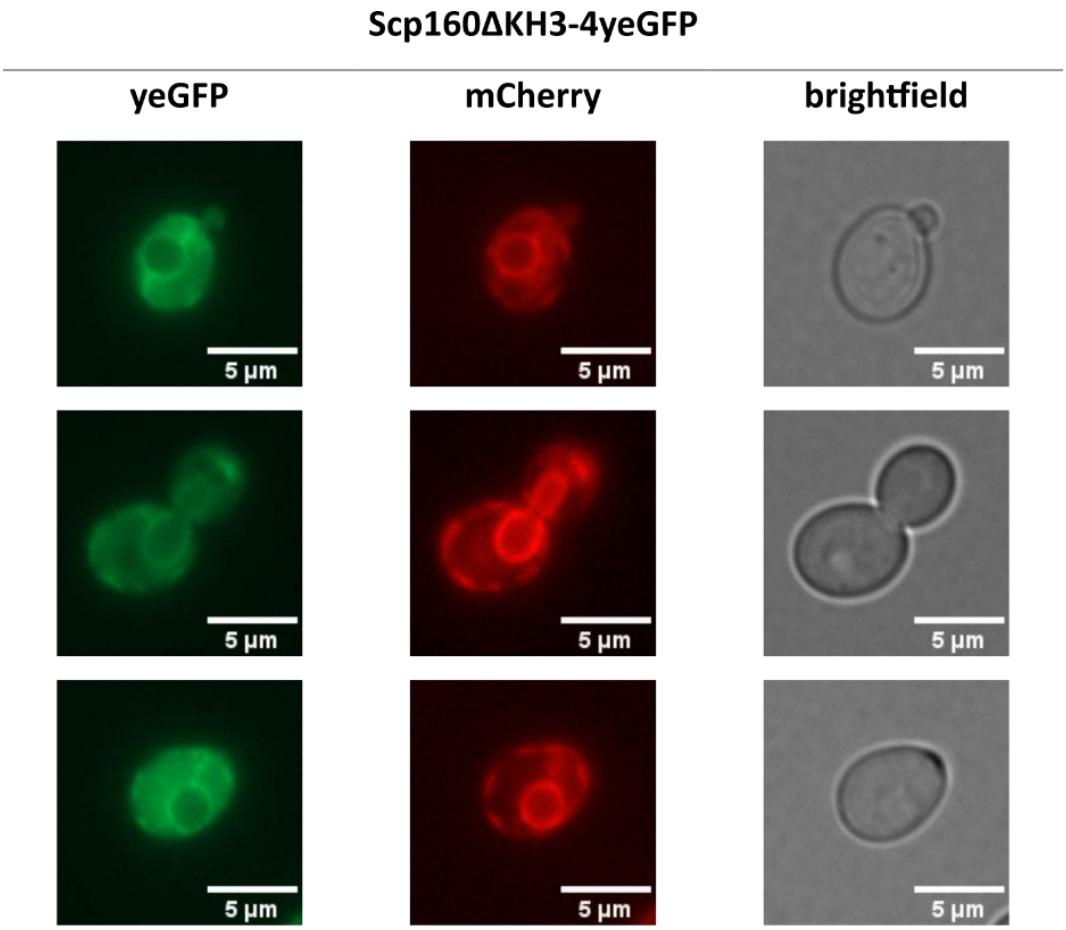


Figure 17: Microscopy images of Scp160yeGFP lacking KH domains three to four; green: GFP channel, red: mCherry channel, grey: brightfield.

Localization of Scp160ΔKH4-5yeGFP was abolished (Figure 18). The protein lacking KH domain four and five has been found not to be able to localize at the endoplasmic reticulum. The GFP signal was distributed in the cytosol.

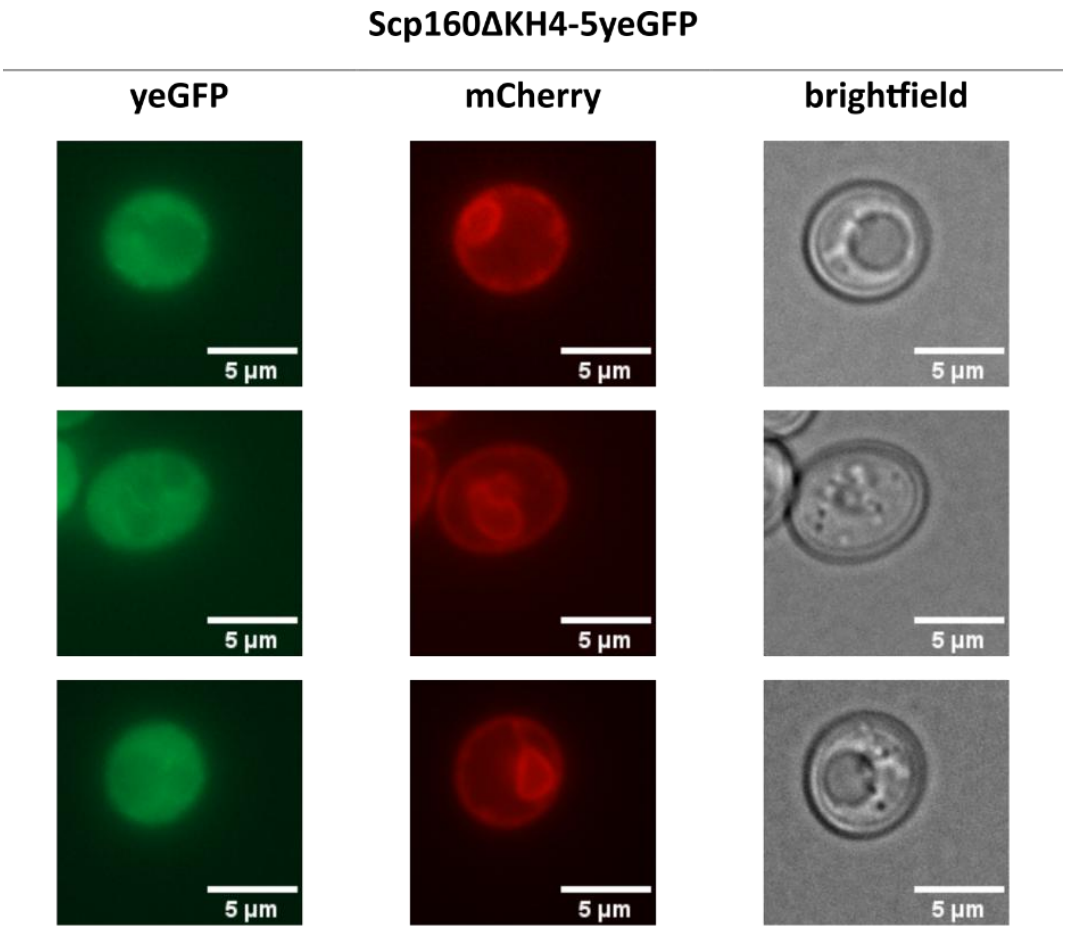


Figure 18: Microscopy images of Scp160yeGFP lacking KH domains four to five; green: GFP channel, red: mCherry channel, grey: brightfield.

As with the Scp160ΔKH4-5yeGFP construct, the GFP fusion of Scp160ΔKH5-6yeGFP was unable to localize to the ER in the majority of cells (Figure 19). However, in a small number of cases, there was evidence of potential ER localization. The GFP signal was primarily distributed throughout the cytosol.

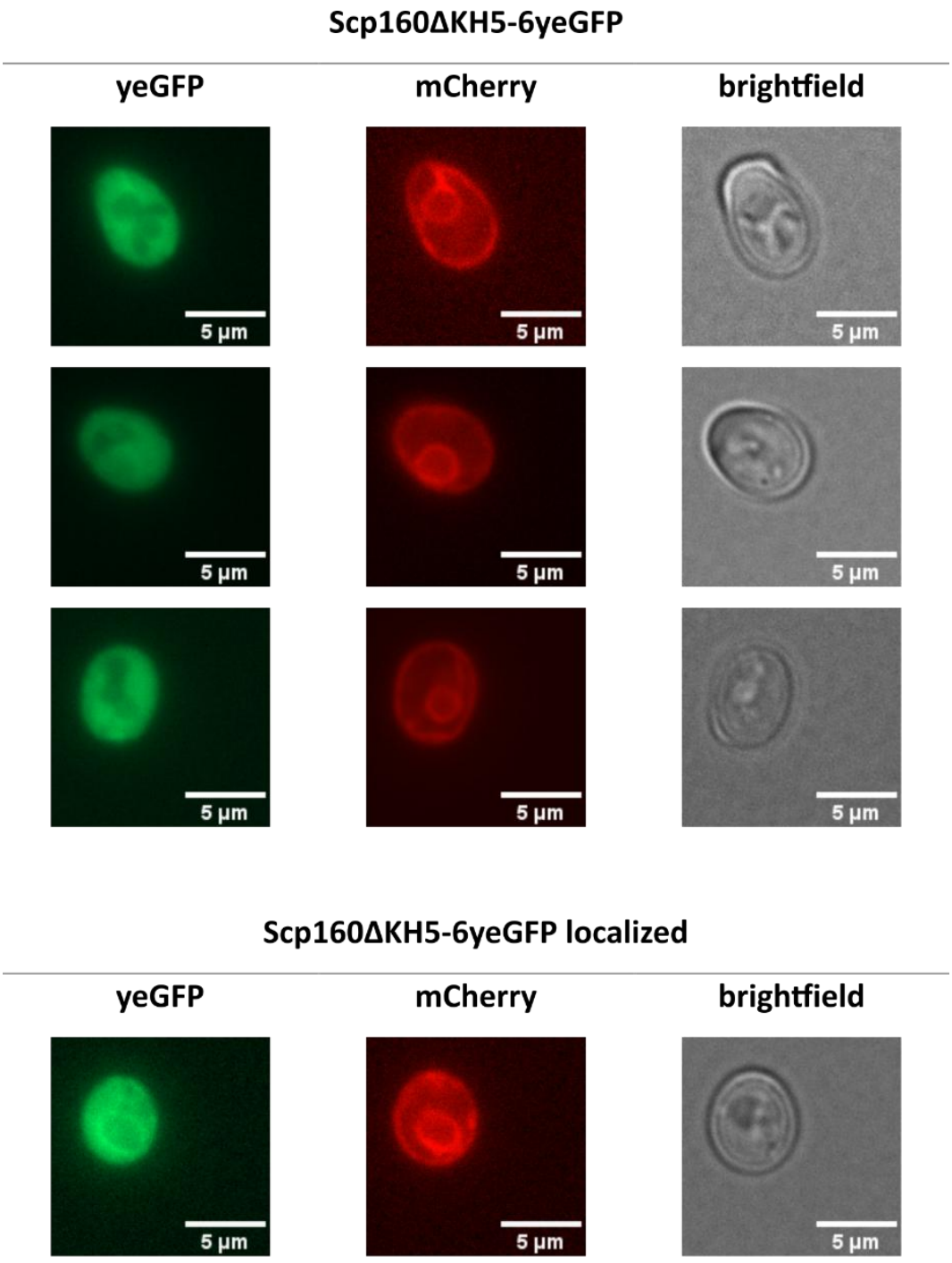


Figure 19: Microscopy images of Scp160yeGFP lacking KH domains five to six; green: GFP channel, red: mCherry channel, grey: brightfield.

The results demonstrated that Scp160ΔKH3-5yeGFP was unable to localize at the endoplasmic reticulum (Figure 20). Instead, the GFP signal was distributed throughout the cytosol.

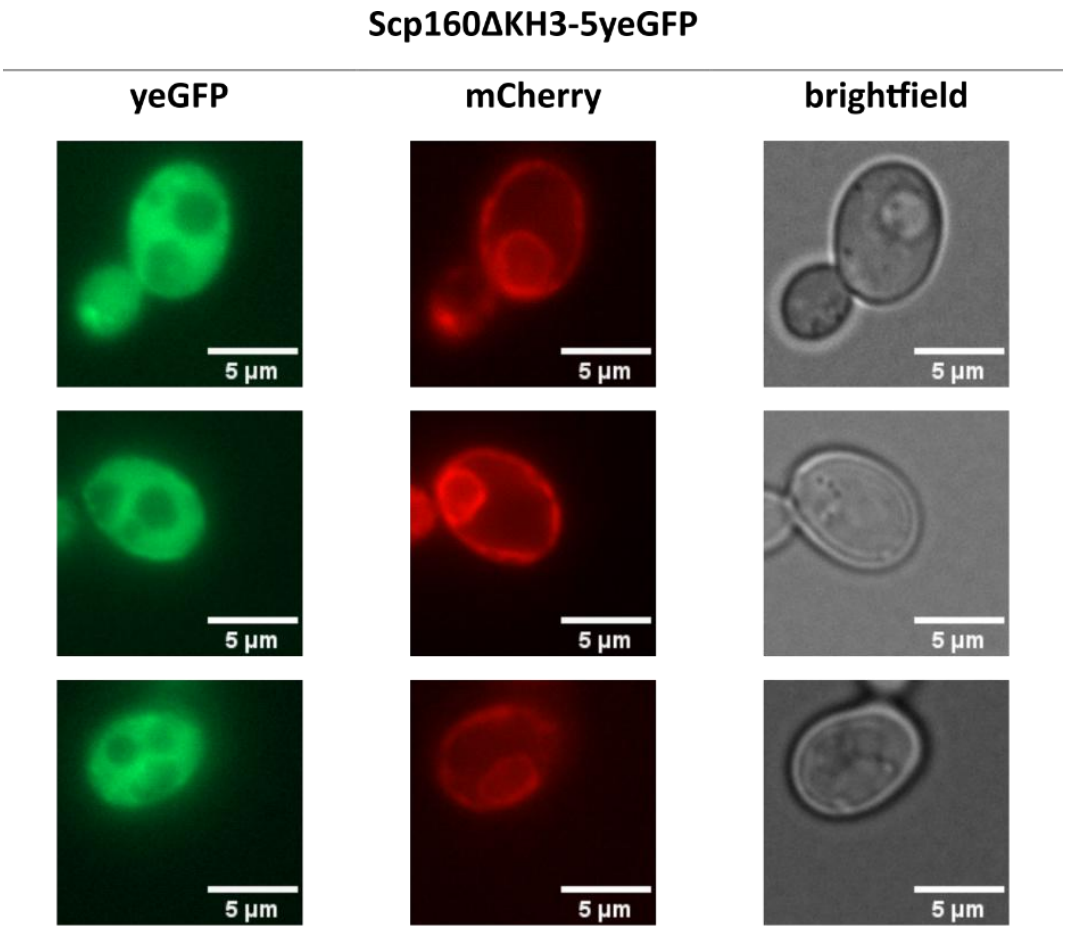


Figure 20: Microscopy images of Scp160yeGFP lacking KH domains three to five; green: GFP channel, red: mCherry channel, grey: brightfield.

The truncated Scp160ΔKH13-14yeGFP protein exhibited incomplete ER localization (Figure 21). An increase in cytoplasmic staining was observed. The truncated protein has been demonstrated to localize at the endoplasmic reticulum, albeit to a lesser degree compared to the wildtype. Additionally, the GFP signal was highly distributed in the cytosol.

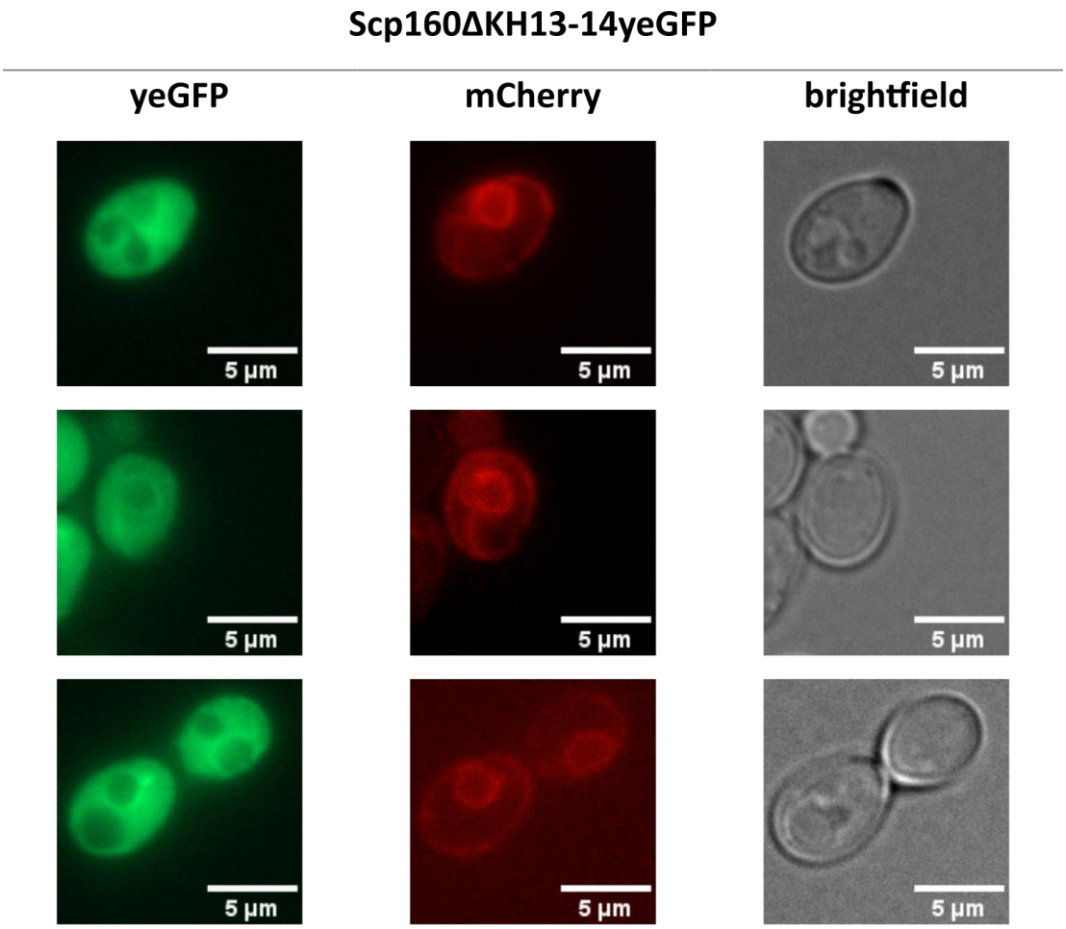


Figure 21: Microscopy images of Scp160yeGFP lacking KH domains 13 to 14; green: GFP channel, red: mCherry channel, grey: brightfield.

The ER localization of Scp160ΔKH2yeGFP was not affected (Figure 22). The protein only lacking KH domain two has been demonstrated to localize at the endoplasmic reticulum.

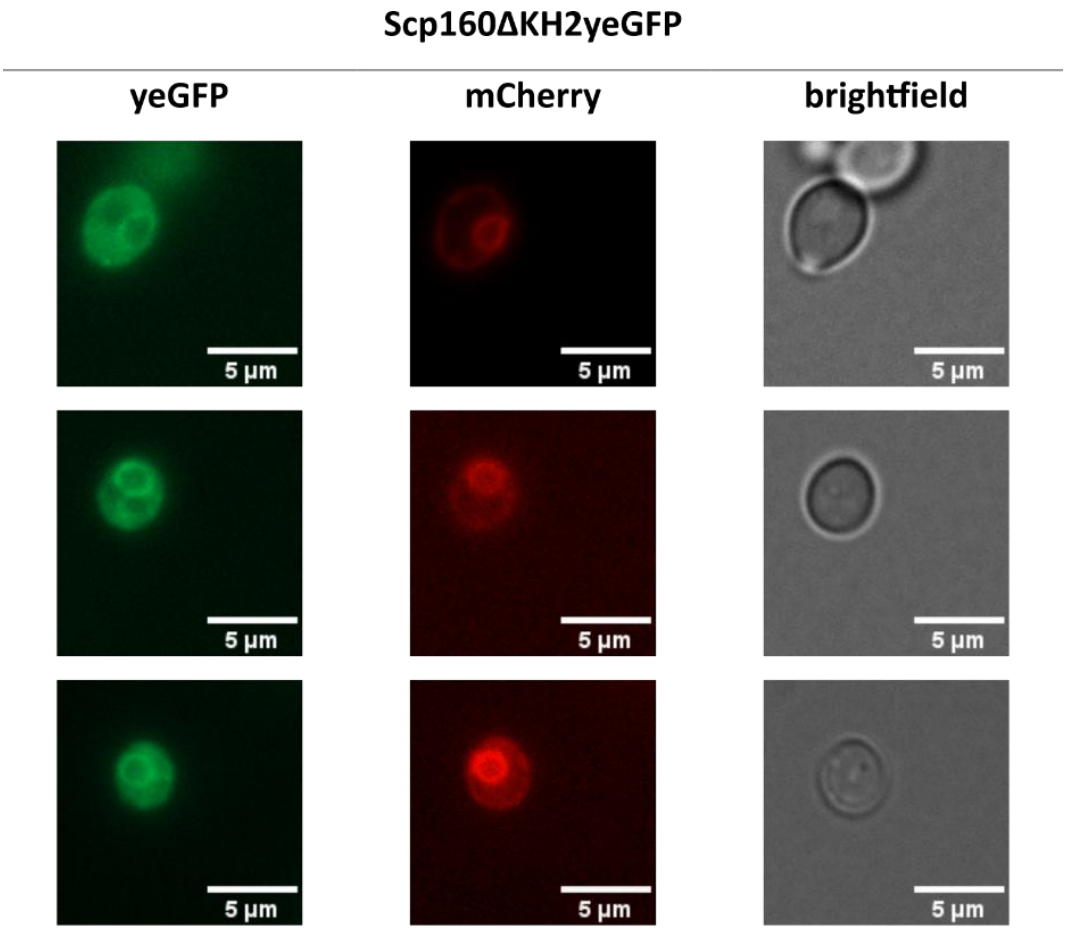


Figure 22: Microscopy images of Scp160yeGFP lacking KH domain two; green: GFP channel, red: mCherry channel, grey: brightfield.

Scp160ΔKH5yeGFP has been observed to be absent from the endoplasmic reticulum (Figure 23). However, in a limited number of instances, evidence suggests the potential for localization at this site. The GFP signal was predominantly distributed throughout the cytosol.

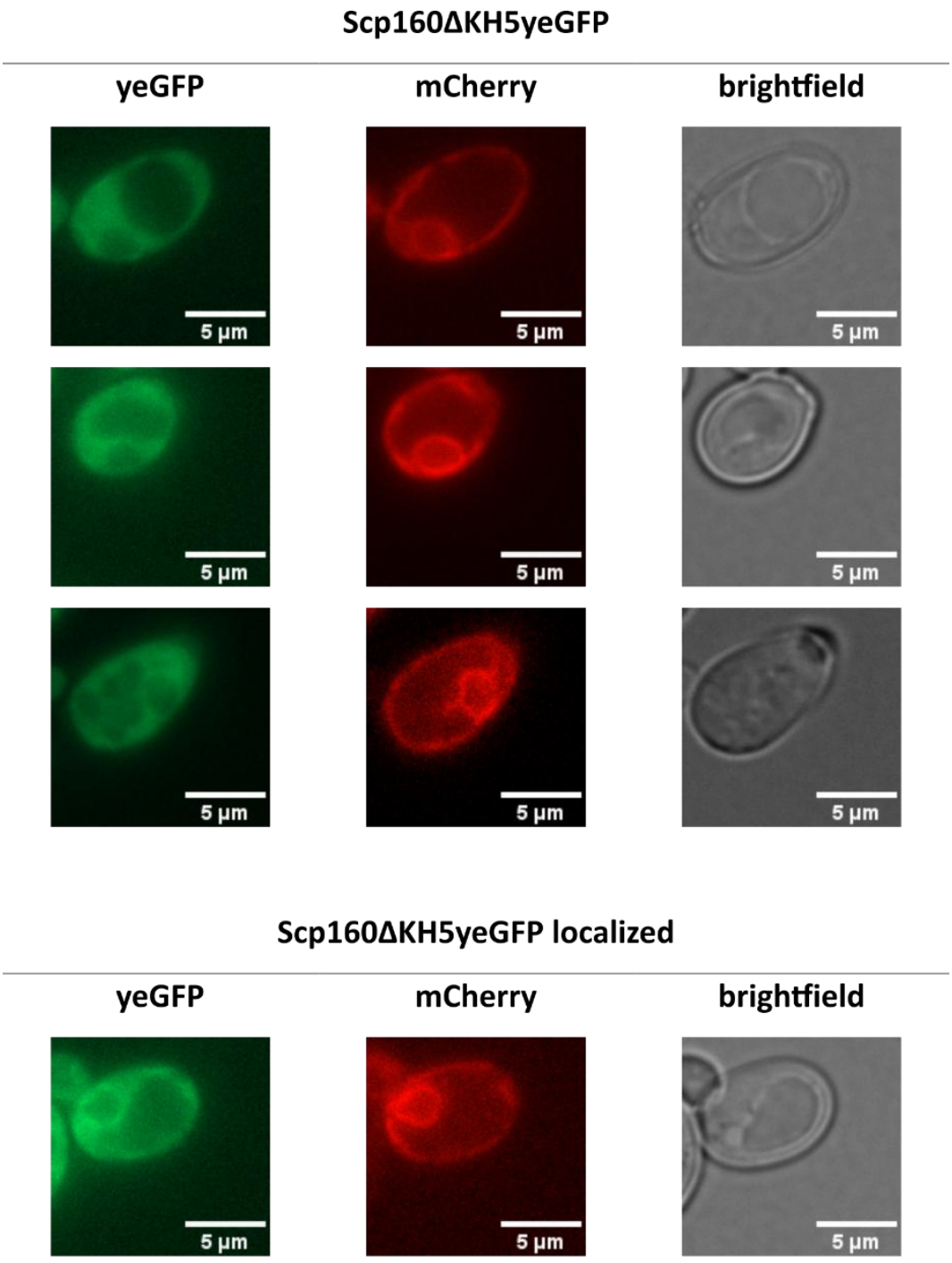


Figure 23: Microscopy images of Scp160yeGFP lacking KH domain five; green: GFP channel, red: mCherry channel, grey: brightfield.

An attempt was made to quantify the ER localization using the FIJI tool "peak finder". In this method, the intensity peaks along a drawn line are represented as a vertex graph, which could, in theory, be quantified. However, the cytosolic signal, the vacuole(s), and the general fluctuation of signal distribution resulted in inconsistent and frequently erroneous positive results. As a result, the quantitative analysis of ER localization in different strains expressing Scp160 variants lacking KH domains was deemed to be unsuitable for further investigation. The analysis of single plane images was hindered by the inability to represent the entirety of the protein signal, a consequence of the three-dimensional nature of the ER (Figure 24). It was not possible to identify general Z-stack settings that would illustrate all localization events due to the presence of varying cell sizes per culture and between strains. If a sufficient number of stacks was used in general, the resulting overlay signal was often saturated. Another ongoing issue was the movement of cells during imaging, which prevented the overlay of a series of pictures in all three channels. Despite employing a range of immobilization techniques, the cells remained mobile, preventing the desired outcome. The addition of glycerol to the immobilization buffer resulted in the rapid delocalization of Scp160. One promising approach was the use of agarose to coat microscopy glass slides. However, minor issues persisted due to irregularities in the agarose coating and light scattering. These included cells outside the general focus level (Figure 24) and light contamination, which impaired the fluorescence signal. The localization of Scp160 at the ER was a particularly challenging aspect. Even under optimal conditions with regard to growth time, nutrient availability, and temperature, the quality of the sample remained unpredictable. This manifested as the potential presence of enlarged vacuoles, which could result in cell enlargement and/or rapid delocalization of Scp160 from the ER.

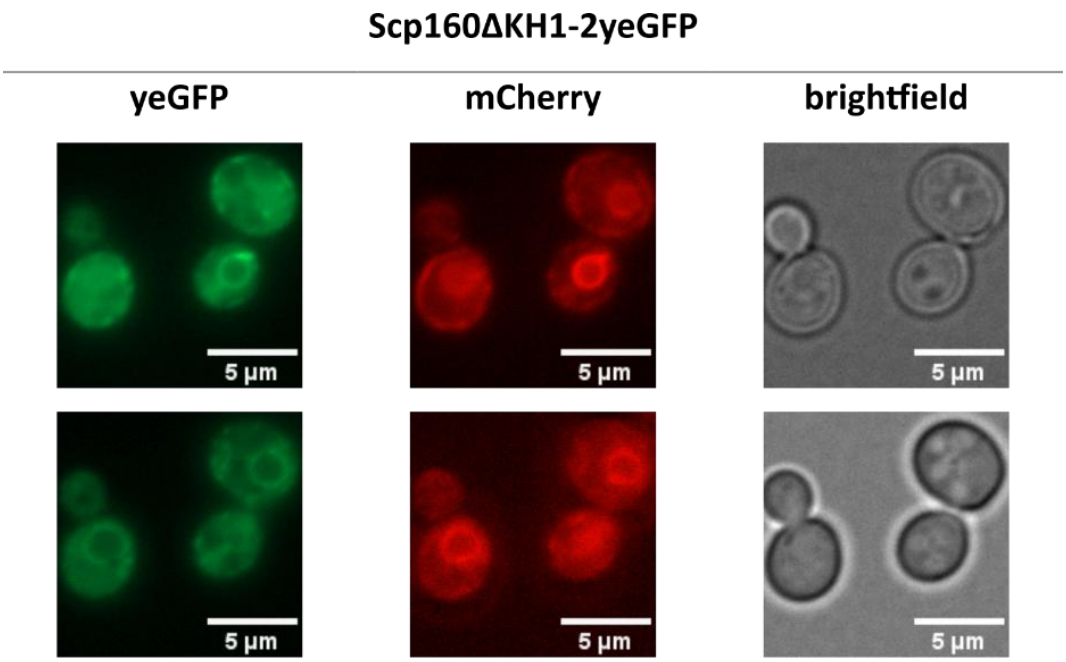


Figure 24: Signal intensity vary in different focus levels. Microscopy images of Scp160yeGFP lacking KH domains one and two, shown in two different focus levels; green: GFP channel, red: mCherry channel, grey: brightfield.

In conclusion, the deletion of N-terminal KH domains has been demonstrated to influence the localization of the protein to the ER (Table 19). The extent of delocalization varies, ranging from complete delocalization, as observed in Scp160ΔKH1-5, to partial delocalization, as seen in Scp160ΔKH5-6, and finally, to the absence of delocalization, as evidenced in Scp160ΔKH1-2.

Table 19: Summary of likelihood of Scp160 variants lacking N-terminal KH domains to localize at the ER.

	(partial) ER localization
Scp160ΔKH1-2	yes
Scp160ΔKH1-3	yes
Scp160ΔKH1-4	no
Scp160ΔKH1-5	no
Scp160ΔKH2-3	yes
Scp160ΔKH3-4	yes
Scp160ΔKH4-5	no
Scp160ΔKH5-6	yes
Scp160ΔKH3-5	no
Scp160ΔKH2	yes
Scp160ΔKH5	yes

The deletion of KH domains has been observed to induce alterations in cell size and morphology

It has been demonstrated that the removal of *SCP160* leads to an increase in cellular size (Wintersberger et al., 1995). To ascertain whether the deletion of distinct KH domains gives rise to analogous alterations or modifications in cell morphology, yeast cells expressing the truncation/deletion variants were examined using bright-field microscopy. The cells were analyzed under conditions that optimized their growth and viability. The cells were cultivated in YPD +2% glucose until an OD_{600} of 1/mL was reached, after which samples were examined with the assistance of a Beckman Coulter CytoFLEX S flow cytometer. Brightfield images of yeast were obtained, and the individual cell area was mapped for 100 cells per strain using the Fiji software. The resulting measurements are presented in Table 20-Table 22 and Figure 25. Given the correlation between cell size and DNA content in wild-type yeast, it was anticipated that a similar correlation would be observed in relation to changes in size depending on the ploidy type. However, the results do not consistently demonstrate a correlation for the various KH truncation/deletion mutants. For instance, the cell size of a haploid Scp160ΔKH1-3 strain is more akin to that of the diploid wildtype than the haploid wildtype. Similarly, cells expressing Scp160ΔKH5 or Scp160ΔKH6 have a similar size as the tetraploid *scp160Δ* mutant, despite exhibiting a pseudo-diploid phenotype. It should be noted that the DNA content of the different strains lacking KH domains, and thus the observed increase in cell size, is not limited to the defined ploidy types of haploid, diploid, and tetraploid. Consequently, it was not feasible to assign a ploidy type to a specific cell size. The diameter of the numerous Scp160 KH-domain deletion strains associated with a specific ploidy overlapped with the diameter of other ploidy types. For example, the diameter of the pseudo-diploid Scp160ΔKH6 strain is within the pseudo-tetraploid range.

(pseudo-)Haploid:

Table 20: Average area of 100 cells per strain.

	Average in cm^2	size in comparison to Scp160 (H)
WT (H)	17,1	-
Scp160ΔKH1-2	17,7	+3,5%
Scp160ΔKH2-3	20,0	+16,9%
Scp160ΔKH2	19,6	+14,6%

(pseudo-)Diploid:

Table 21: Average area of 100 cells per strain.

	Average in cm^2	size in comparison to Scp160 (D)
WT (D)	24,2	-
Scp160ΔKH1-3	26,0	+6,5%
Scp160ΔKH1-4	23,9	-2,0%
Scp160ΔKH3-4	28,4	+16,4%
Scp160ΔKH5	43,7	+42,2%
Scp160ΔKH6	39,0	+59,8%

(pseudo-)Tetraploid:

Table 22: Average area of 100 cells per strain.

	Average in cm^2	size in comparison to *scp160Δ* (T)
scp160Δ(T)	38,7	-
Scp160ΔKH1-5	38,3	-1,0%
Scp160ΔKH4-5	40,0	+3,4%
Scp160ΔKH5-6	32,3	-16,5%
Scp160ΔKH3-5	42,2	+9,0%

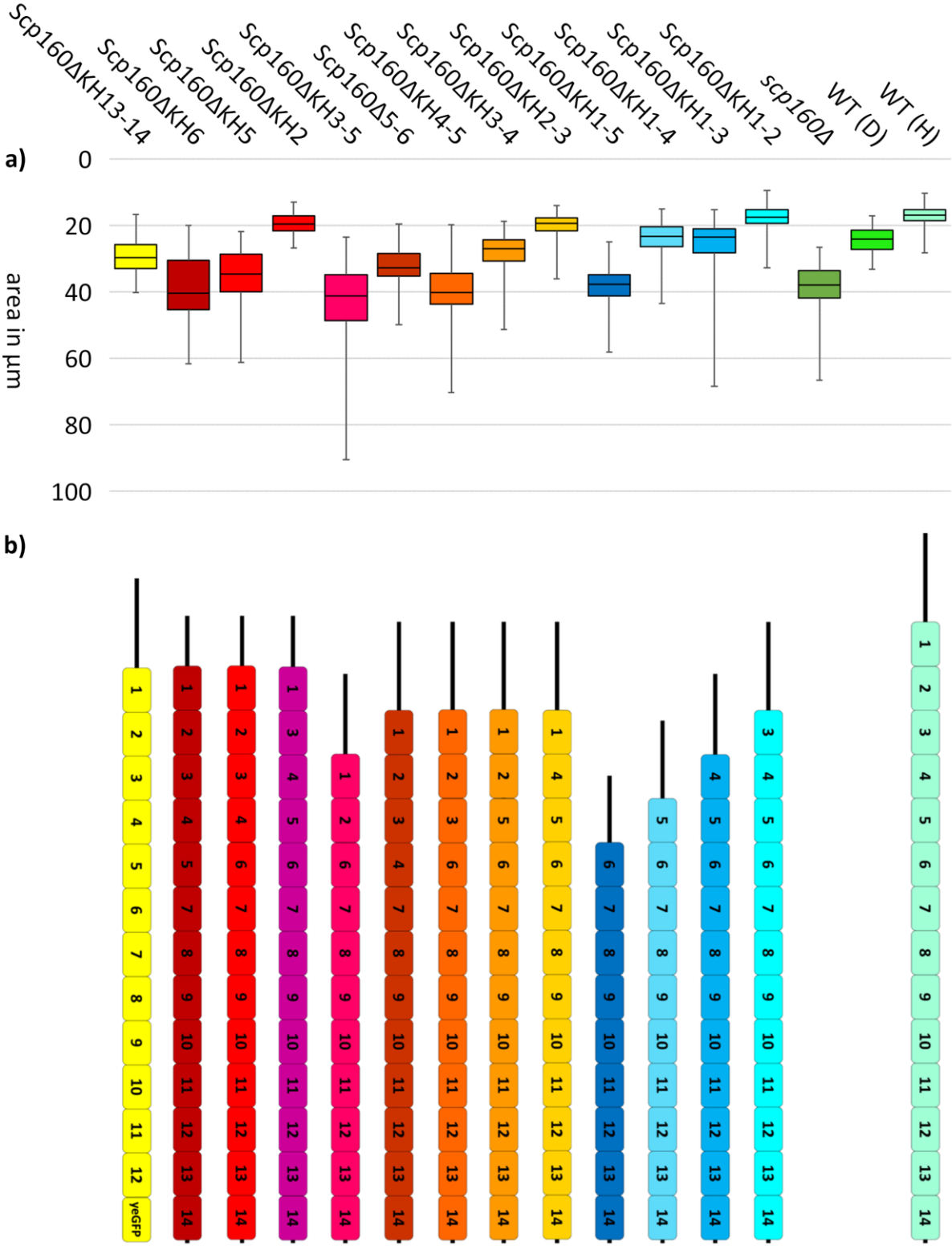


Figure 25: Size comparison between different cells expressing Scp160 and Scp160 lacking specific KH domains; a); area of 100 cells illustrated in boxplot assays x-axis: Scp160 strains, y-axis: area in µm; b) corresponding images of Scp160 variants lacking KH domains; Scp160 in WT (D) equates to Scp160 in WT (H).

The alterations in cell wall characteristics that arise from KH domain deletions are associated with changes in emulsification

In a recent publication (Nerome et al., 2020), a new phenotype of *scp160Δ* was identified. In an effort to identify yeast strains with altered cell wall composition that could be utilized as food emulsifiers, a yeast knockout collection was screened for potential candidates. The search was conducted on strains that exhibited an increase in their emulsifying properties. The wild-type BY4741 strain, which was used in this screening process, did not exhibit any observable emulsifying effects. Of the 5,076 gene deletion strains tested, the deletion of *SCP160* was identified as one of eight hits for the desired phenotype. It is noteworthy that the deletion of the gene encoding the Scp160-interacting Bfr1 protein also resulted in an increase in the emulsifying effect (Nerome et al., 2020). The precise nature of the factor responsible for the emulsifying effect observed in cells depleted of Scp160 remains unclear. Given that the protein in question is linked to the secretory pathway, there are a number of potential connections to be considered, including alterations in cell wall properties that may result in emulsifying effects. Such alterations may be the consequence of the loss of an enzyme involved in cell wall biosynthesis. To ascertain whether the generated truncation/deletion mutants exert an influence on the emulsion properties of the corresponding yeast strain, emulsification assays were conducted in accordance with the methodology outlined by Nerome et al. (2020). In summary, the assay entails the preparation of a logarithmic yeast culture with defined parameters, which is then mixed with a non-polar solution (kerosene). The mixture is then vortexed, and if the yeast strain exhibits an emulsifying effect, the kerosene will form a foamy phase (Figure 26). Of the tested strains expressing Scp160 variants lacking KH domains, three (Scp160ΔKH1-5, Scp160ΔKH3-5, and Scp160ΔKH4-5) exhibited a similar emulsifying effect as the full deletion of *SCP160* (Figures 27-29). No other tested strain exhibited an emulsifying effect. The deletion of the KH domain 5 alone did not result in an emulsifying effect.

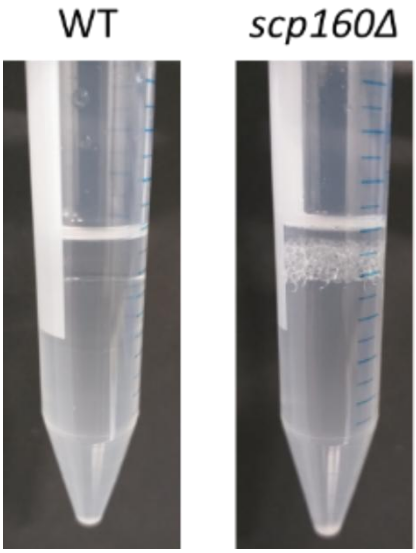


Figure 26: Comparison of emulsifying properties of the scp160Δ mutant in comparison to the wildtype strain; lower phase: water with 10 OD_{600} yeast cells; upper phase: kerosene.

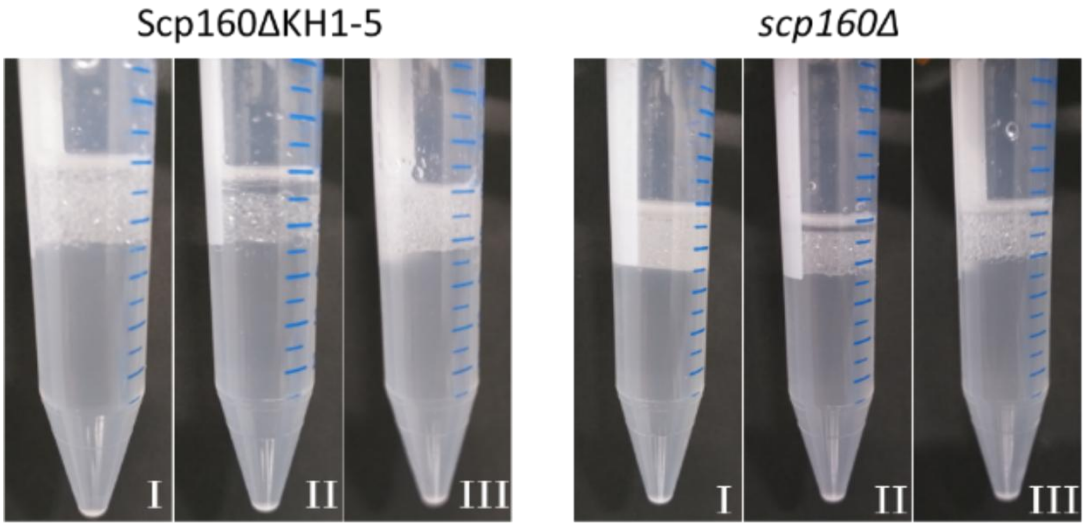


Figure 27: Emulsion properties of Scp160ΔKH1-5 in comparison to the scp160Δ control.

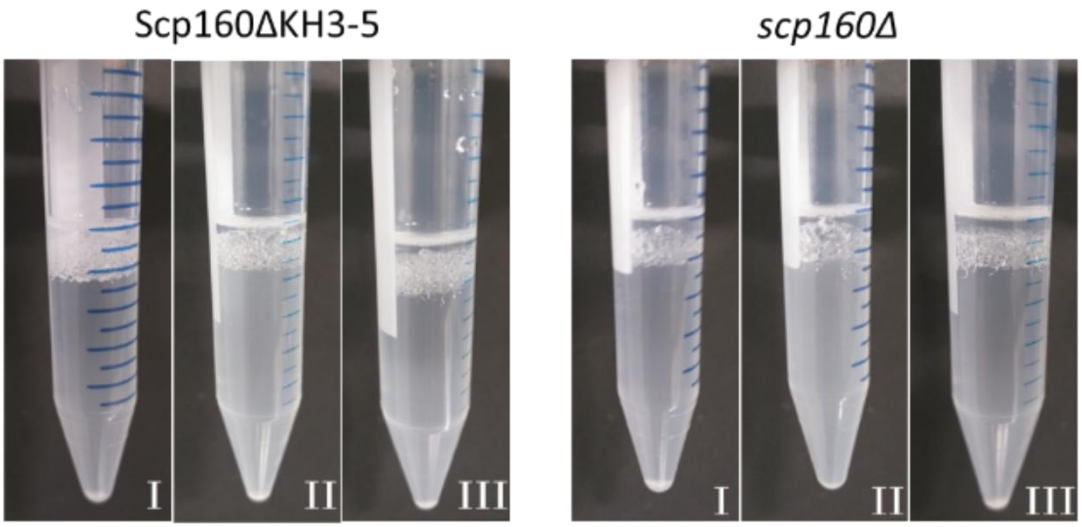


Figure 28: Emulsion properties of Scp160ΔKH3-5 in comparison to the scp160Δ control.

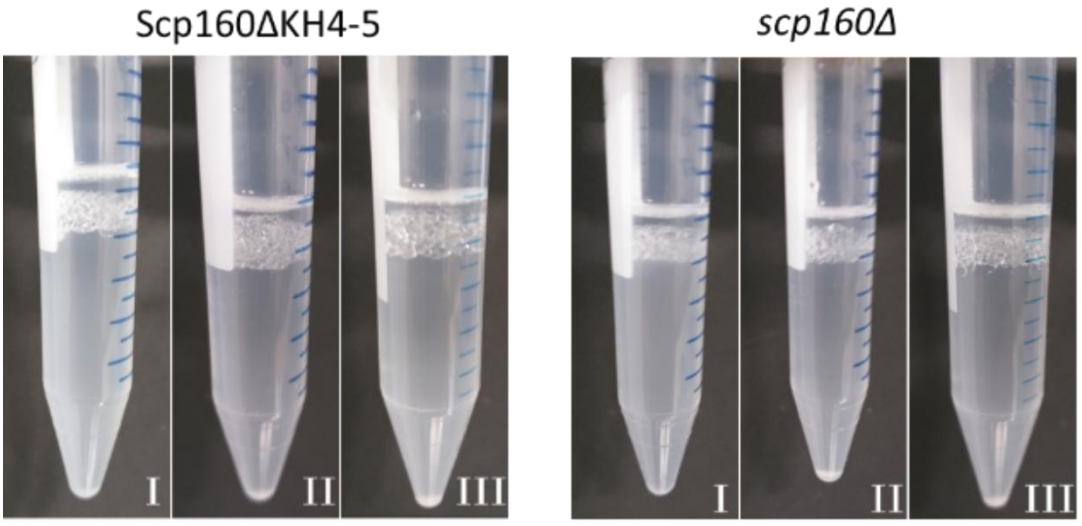


Figure 29: Emulsion properties of Scp160ΔKH4-5 in comparison to the scp160Δ control.

Summary of phenotypes observed for Scp160 truncations or deletions

The presented data collectively indicate the existence of a number of promising KH domain truncation/deletion strains, including Scp160ΔKH5, Scp160ΔKH4-5, Scp160ΔKH3-5, and Scp160ΔKH1-5. The four strains exhibit discernible phenotypes in ploidy change, cell size, localization, and emulsifying properties. A unifying characteristic of these strains is the loss of the fifth KH domain. These strains display phenotypes similar to those observed in the complete deletion of *SCP160*, which suggests that KH domain 5 or KH domains 3-5 play a crucial role. Given the limitations in available

resources and time frame, it was necessary to reduce the number of strains to be tested in further experiments regarding proteomic and transcriptomic analysis. Accordingly, the following strains were selected for further investigation: The strains under consideration are Scp160ΔKH1-5, Scp160ΔKH4-5, and Scp160ΔKH5. The Scp160ΔKH4-5 strain was selected for analysis due to the observation that the deletion of KH domains four and five resulted in comparable severe phenotypes to those observed in the Scp160ΔKH1-5 strain, albeit with a lesser impact on the overall primary structure of the protein. Although it exhibited a relatively weaker phenotype thus far, the strain Scp160ΔKH5 was selected as the final sample since it represents the sole strain with a single KH domain deletion within the identified region that demonstrated a phenotype in my assays. Consequently, I proceeded to conduct further molecular analysis, focusing on alterations in protein-protein and protein-RNA interactions.

The loss of KH domains and its implication for protein-protein interaction

The phenotypes described above are likely the result of a loss of protein function or a loss of interaction with partner proteins, including ribosomes. A number of partners have been identified in recent years, including Bfr1 (Lang, 2000), Asc1 (Baum et al., 2004), and Pab1 (Lang, 2000). In light of the observation that the loss of specific protein-protein interactions in the case of C-terminal KH domains, such as Asc1, has been documented (Baum et al., 2004), it is imperative to determine whether N-terminal deletions also result in the loss of interaction. Finally, it would be beneficial to ascertain whether the deletion of KH domains affects protein-protein interactions and RNA binding in a comparable manner or if there is no correlation between the two. To identify potential alterations in the protein-protein interactions of Scp160 variants on a global scale, the protein interactome of Scp160 in selected strains with KH domain truncations or deletions was compared to that of wild-type cells. Furthermore, it was my intention to ascertain whether the observed phenotypes have their origin in the absence of interacting proteins resulting from the loss of binding domains in the form of KH domains, or alternatively, from structural alterations to the protein. It is also possible that the deletion process has indirectly impaired interactions with other proteins. Consequently, alterations in protein-protein interactions were examined. The strains Scp160ΔKH1-2, Scp160ΔKH1-5, Scp160ΔKH4-5, Scp160ΔKH5, and Scp160ΔKH13-14 were selected as candidates for screening for changes in interaction partners for the initial attempt. To monitor changes in the interactome of Scp160, co-immunoprecipitation assays were employed. Immunoprecipitation was conducted with ChromoTek GFP-Trap® Magnetic Agarose beads, with all strains tagged with yeGFP.

Scp160ΔKH1-2 was selected as a representative example of a truncated protein for which no discernible mutant phenotype could be identified. To illustrate the opposite end of the spectrum of

mutant phenotypes, the strain with the most severe truncated sequence, Scp160ΔKH1-5, was selected as an example. The two strains, Scp160ΔKH4-5 and Scp160ΔKH5, which carry relatively small deletions of KH domain regions and display similar phenotypes, were selected based on the relatively strong defects observed in previous experiments. As a control for the experimental procedure, Scp160ΔKH13-14 was selected. Previous research has demonstrated that this strain exhibits alterations in mRNA binding (Gelin-Licht et al., 2012), ribosome localization (Li et al., 2004), and protein-protein interactions when compared to the wild-type strain (Baum et al., 2004). Furthermore, yeGFP was expressed under the control of the Scp160 promoter and transcriptional terminator to serve as a control for nonspecific binding to the GFP tag. Co-immunoprecipitation was performed in triplicates, yielding a total of 21 samples. The protein samples were subjected to analysis by a label-free mass spectrometry experiment, which was conducted at the Interfaculty Institute for Cell Biology (IFIZ) in Tübingen by Mirita Franz. Data analysis was conducted by Francesca Barletta at the University of Tübingen's Center for Quantitative Biology (QBiC) core facility.

The overall variation between the interactome datasets is illustrated in Figure 30. The principal component analysis (PCA) plot demonstrates a low to very low divergence between the triplicates, indicating a high degree of comparability for the triplicates. The only outlier is found in the strain expressing Scp160ΔKH4-5yeGFP, specifically sample #3 (green x). Upon examination of the raw data for this sample, it was determined that neither technical irregularity nor flawed sample was indicated. This is evidenced by the similar clustering of the triplicates, as seen in Figure 31. Therefore, it was assumed that the observed variance was biological in nature and thus not included for further analysis.

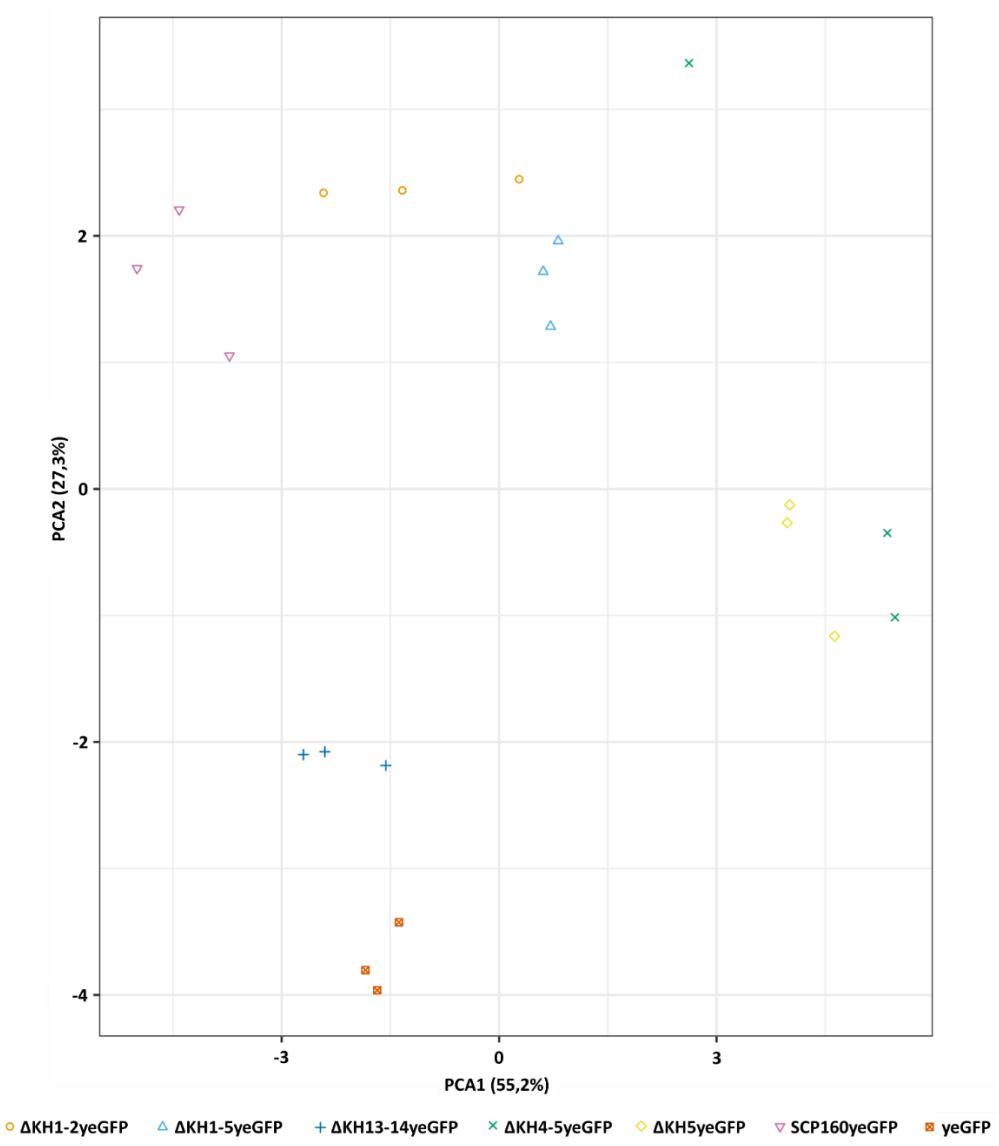


Figure 30: Principal component analysis (PCA) of the yeast strains expressing Scp160 variants lacking KH domains and control triplicates; orange circle: Scp160ΔKH1-2-yeGFP; bright blue triangle: Scp160ΔKH1-5-yeGFP; dark blue plus: Scp160ΔKH13-14-yeGFP; green x: Scp160ΔKH4-5-yeGFP; yellow square: Scp160ΔKH5-yeGFP; pink triangle: Scp160-yeGFP; orange square (crossed): yeGFP under Scp160 promoter and terminator.

Figure 31 presents a heatmap that offers a comprehensive overview of the identified interaction partners of Scp160, the KH domain deleted versions of Scp160, and yeGFP in each sample. As can be observed from the distribution, the three replicates from each strain exhibit highly similar interaction patterns, thereby indicating the low degree of variation between the replicates. The heatmap indicates a strikingly similar pattern of Scp160-interacting proteins in the KH deletion strains Scp160ΔKH1-2, Scp160ΔKH1-5, Scp160ΔKH4-5, and Scp160ΔKH5, with only minor variations discernible when comparing these strains to the wildtype. The controls for Scp160ΔKH13-14 and yeGFP both exhibit a similar pattern, yet do not display a markedly different picture in comparison to the other samples. For samples exhibiting a higher enrichment level (dark blue), the corresponding enrichment levels of the other strains often also display an elevated enrichment level (light to medium blue). However, a few strains display slight discrepancies in their binding patterns when compared to the other samples. One illustrative example is the protein-protein interaction pattern (a) in Figure 31. Such discrepancies could also be attributed to inherent variabilities in the samples. Irregularities are exemplified by some of the triplicates, as evidenced by the protein-protein interaction patterns (b) and (c) (Figure 31). The controls of Scp160ΔKH13-14 and yeGFP exhibit a slightly different cluster pattern, which demonstrates the distinctions between the deletion of N-terminal KH domains and that of C-terminal KH domains. To provide a more detailed view of the common and different interactors of the various Scp160 truncation/deletion variants, a second heatmap was generated (Figure 32). Additionally, quantitative assessment of enrichment or depletion of specific interacting proteins in the various strains expressing Scp160 variants lacking KH domains is presented in volcano plots (Figure 33-Figure 38).

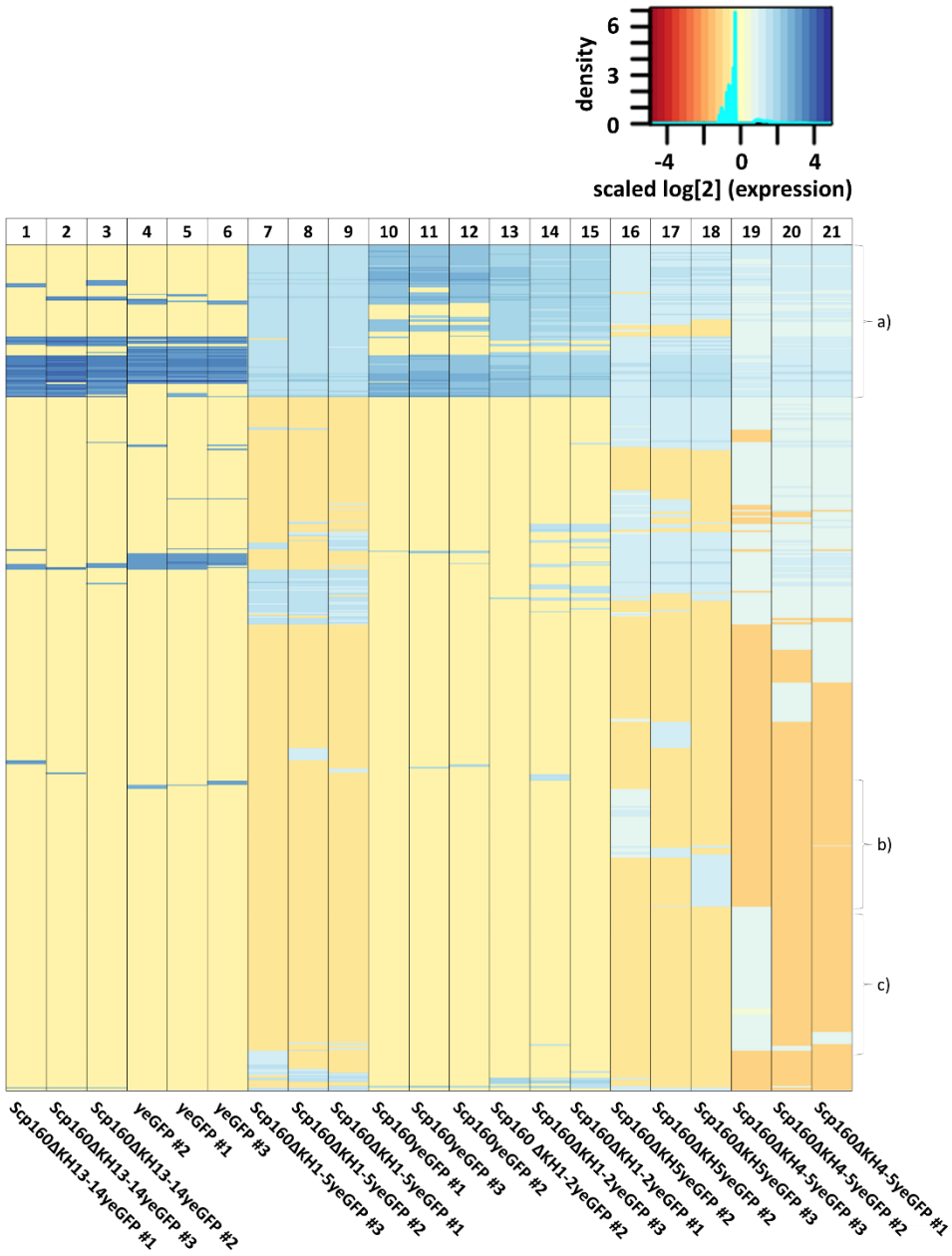


Figure 31: Heatmap showing the comparison of found Scp160 and Scp160 variant interactors in between all chosen strains expressing versions of Scp160 lacking KH domains, full-length Scp160 and yeGFP. Color scale: orange to red: depletion of interacting proteins; light to dark blue: enrichment of interacting proteins; (a)-(c): areas of interest: see text above.

Figure 32 depicts the identified interaction partners of the full-length Scp160, the KH domain deleted versions of Scp160, and the negative control yeGFP, with a particular emphasis on the Scp160ΔKH4-5 variant. A total of 629 significant interactors were identified across the 21 samples. Of the 629 samples, 96 exhibited significant enrichment or depletion in abundance when compared across all samples (Scp160, Scp160ΔKH1-2, Scp160ΔKH1-5, Scp160ΔKH4-5, Scp160ΔKH5, and Scp160ΔKH13-14, as well as yeGFP).

In comparison to the full-length Scp160, the negative control yeGFP did not interact with the same proteins shown in columns 10 to 12 (yeGFP). This is in contrast to the interaction profile observed for the full-length Scp160 in columns 19 to 21. As an illustration, the known Scp160 interaction partner Bfr1 (first row) is absent in yeGFP and enriched for Scp160. Of the 96 significant interactors, 63 (out of 79 total) are ribosomal proteins of both the small and large subunit. This indicates that the binding of ribosomal proteins represents the largest group of interacting proteins, with a frequency of approximately 66%. This finding aligns with the description of Scp160 as a polyribosome-binding protein (Frey et al., 2001). For this triplicate, 86 out of the 96 proteins are represented. The remaining ten proteins are represented in the other six heatmaps (Supplementary Figure 3-7). The strains Scp160ΔKH1-5, Scp160ΔKH4-5, and Scp160ΔKH5, which are shown in the heatmap focusing on Scp160ΔKH4-5 (Figure 32), exhibit a similar clustering for the 86 shown interactors. The control strains Scp160ΔKH13-14 and yeGFP also exhibit comparable characteristics. In the case of the wildtype, the pattern is comparable to that observed in the strain Scp160ΔKH1-2, exhibiting similar patterns of enrichment and decline in protein-protein interactions. The Scp160ΔKH1-2 strain exhibits characteristics analogous to those observed in the Scp160ΔKH1-5, Scp160ΔKH4-5, and Scp160ΔKH5 strains.

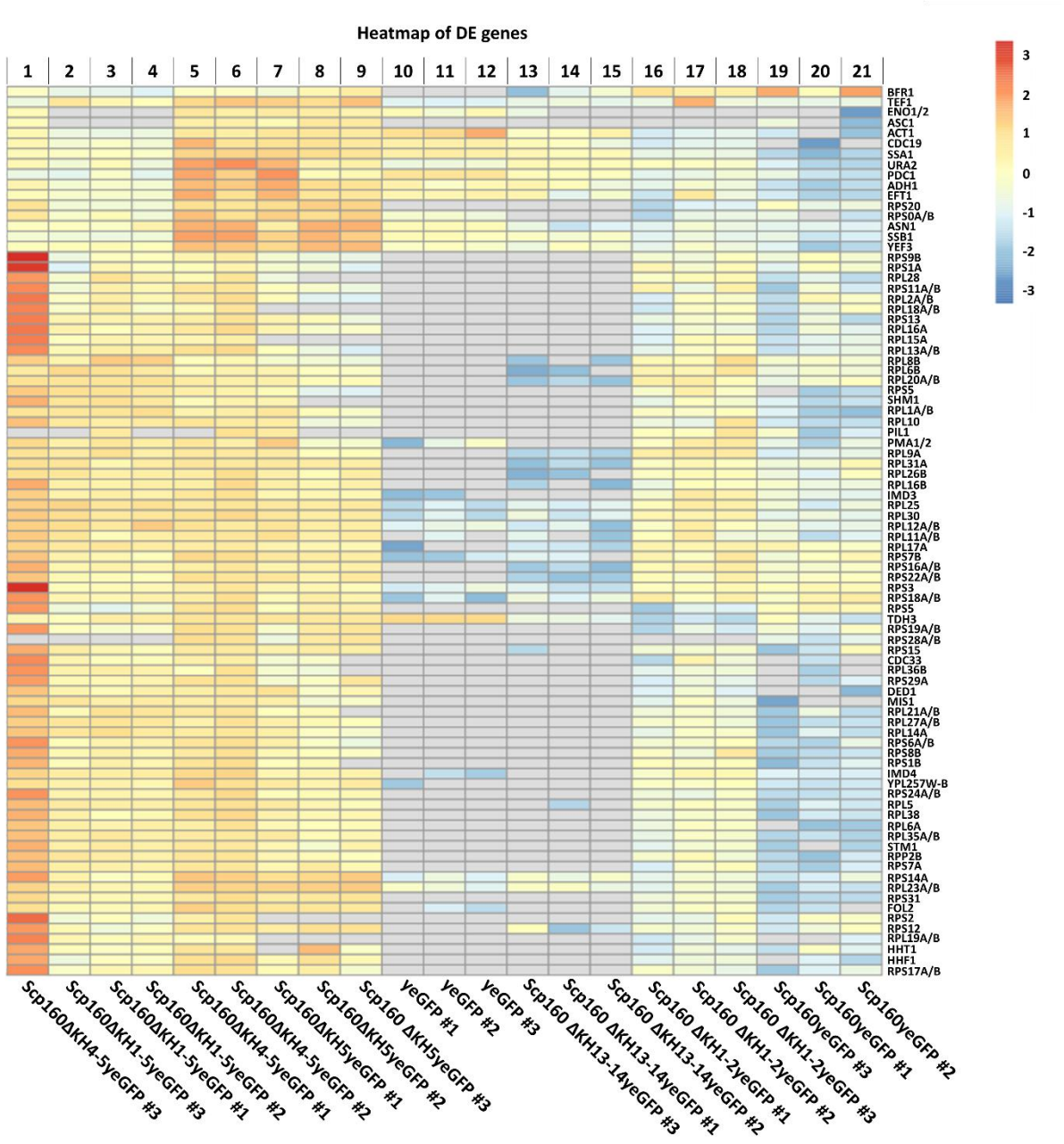


Figure 32: Heatmap *of the differentially enriched interaction partners ('DE genes', top): Comparison of Scp160ΔKH4-5 to WT; Color scale: orange to red: depletion of interacting proteins; light to dark blue: enrichment of interacting proteins; gene name of the interactors (rows) depicted on the right side of the heatmap.*

The negative control of yeGFP under the *SCP160* promoter/terminator exhibits Scp160 as the highest reduced-associated protein in comparison to Scp160. It should be noted that the plot illustrates a loss of interaction with proteins on the left side (Figure 33).

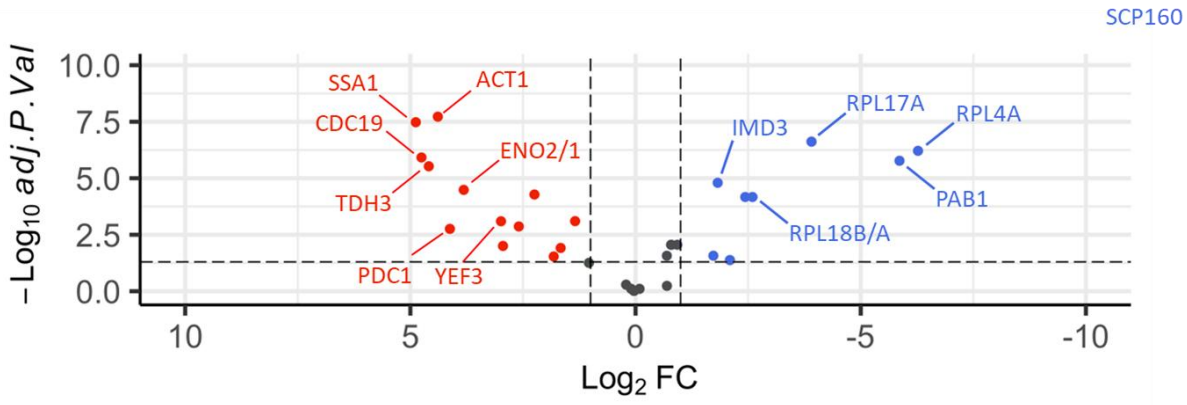


Figure 33: Volcano plot showing proteins significantly enriched or depleted in an immunoprecipitation of yeGFP under Scp160 promoter and terminator compared to a control (unaltered BY4741 wildtype); blue: reduced associated proteins, red: increased associated proteins; black: no significant change in interaction.

In the case of the KH domain deletion mutants, these are the more significant, and thus the presentation of the data has been chosen in this manner because in the subsequent plots, the strains of interest are plotted against the full-length protein. This approach ensures a rigorous and consistent evaluation of the results. The adjusted p-value for Scp160 is so high that it is not possible to map it in Figure 33.

Another protein with a reduced detection rate in these samples is Pab1. Given that Pab1 is a known interaction partner, the presence of Pab1 in the higher-up reduced associated proteins also indicates that the control is functioning as intended. It should be noted that another direct interaction partner frequently observed in conjunction with Scp160, Bfr1, was not identified in this control. In addition to the ribosomal proteins Rpl4A, Rpl17A, and Rps18B/A, the Inosine monophosphate dehydrogenase Imd3 is among the proteins with decreased association with the negative control.

The increased associated proteins exhibit partial overlap with the findings observed in the samples of interest. If this is due to unspecific findings, the specific reasons will be described in the discussion section. Proteins identified as being associated with the negative control include highly abundant proteins such as actin (Act1) and enolase (Eno2/1), as well as Ssa1, Cdc19, Tdh3, Yef3, and Pdc1 (Figure 33). It is necessary to consider these as unspecific binding in the subsequent graphs.

The deletion of the initial two KH domains results in a change of the protein-protein interaction in comparison to the full-length protein (Figure 34).

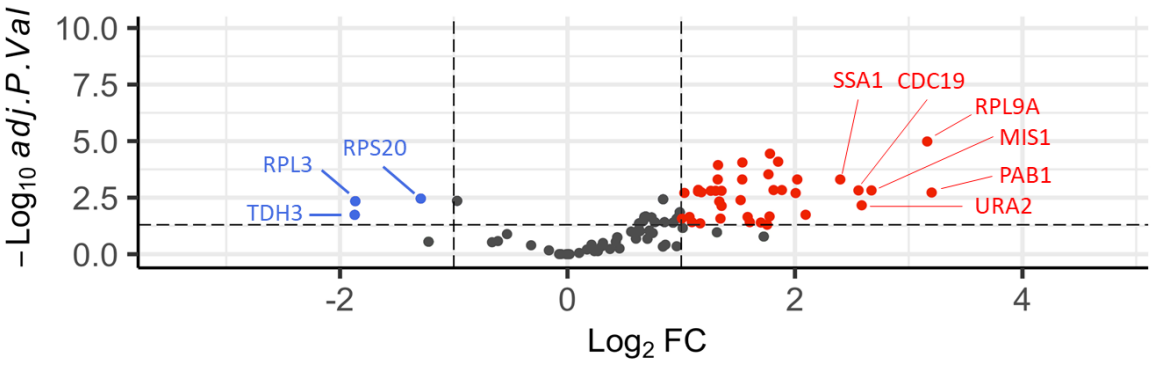


Figure 34: Volcano plot showing protein-protein interaction comparison of Scp160ΔKH1-2 to full-length Scp160; blue: reduced associated proteins, red: increased associated proteins; black: no significant change in interaction.

Three proteins are significantly depleted: the ribosomal proteins Rpl3 and Rps20, and Tdh3. Tdh3, a glyceraldehyde-3-phosphate dehydrogenase (GAPDH), has been observed in the cytoplasm and cell wall (Delgado et al., 2003, 2001; Silva et al., 2011). It was unexpected that a known interaction partner of Scp160, Pab1, was found to be more associated with Scp160 in the deletion mutant lacking KH domains 1 and 2. Pab1 binds to Scp160 in an mRNA-dependent manner (Lang, 2000). Additionally, both proteins were found to associate with Bfr1 in a polyribosome-associated manner (Lang, 2001). Additionally, the association of other proteins was found to be increased, including Mis1 (mitochondrial tetrahydrofolate synthase), Ura2 (bifunctional carbamoylphosphate synthetase/aspartate transcarbamylase), Cdc19 (pyruvate kinase), and Ssa1 (heat shock protein).

The deletion of the KH domains one to five results in a change of the protein-protein interaction in comparison to the full-length protein (Figure 35).

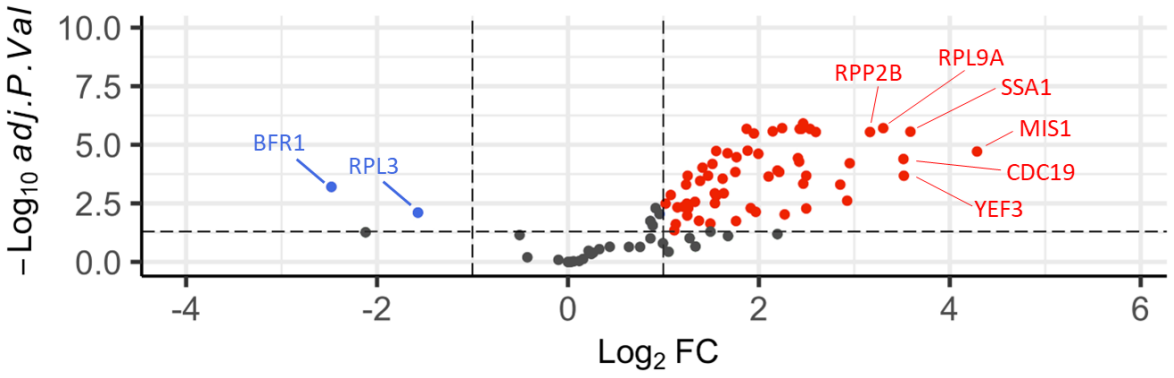


Figure 35: Volcano plot showing protein-protein interaction comparison of Scp160ΔKH1-5 to full-length Scp160; blue: reduced associated proteins, red: increased associated proteins; black: no significant change in interaction.

Two proteins that interact with the full-length Scp160 exhibited a notable reduction in this mutant. The ribosomal protein Rpl3, which was already found to be reduced in Scp160ΔKH1-2, and, more interestingly, the known Scp160 interactor Bfr1. Furthermore, the increased associated proteins Mis1, Ssa1, and Cdc19 were also identified in Scp160ΔKH1-2 (Figure 34). Likewise, the ribosomal proteins Rpl9A, Rrp2B, and Yef3, an elongation factor, were identified among the enriched proteins.

A comparison of the interactome of Scp160ΔKH4-5 with that of the full-length Scp160 (Figure 36) revealed a significant reduction in the association of only one protein, namely Bfr1.

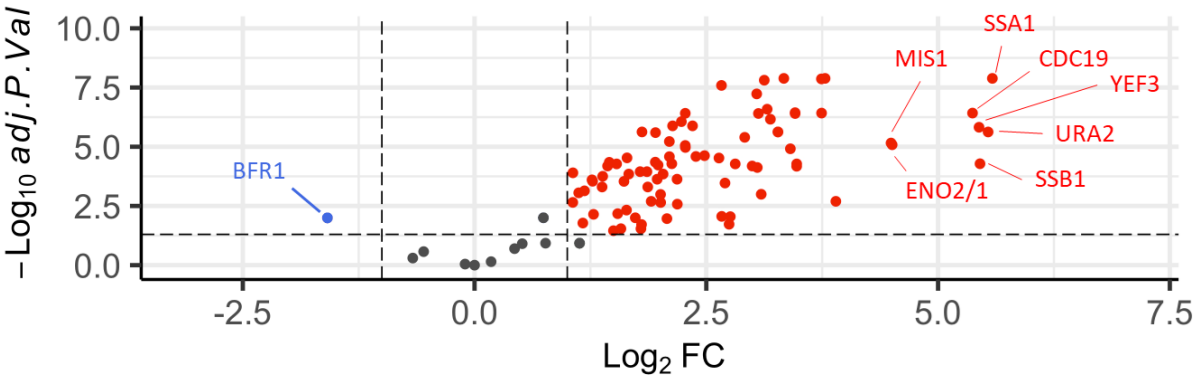


Figure 36: Volcano plot showing protein-protein interaction comparison of Scp160ΔKH4-5 to full-length Scp160; blue: reduced associated proteins, red: increased associated proteins; black: no significant change in interaction.

This protein has also been identified in the truncation mutant lacking the first five KH domains (Figure 35). The proteins that demonstrate an elevated level of association with the Scp160ΔKH4-5 protein exhibit a pattern that is consistent with that observed in Scp160KHΔ1-5 (Figure 35).

The loss or gain of associated proteins in a Scp160 mutant, which lacks the KH domain five (Figure 37), reflects the pattern observed in the previous two mutants (Figure 35 and Figure 36).

The levels of three proteins were significantly reduced in comparison to the full-length. In addition to Bfr1, two more ribosomal proteins, including Rpl3, exhibited a reduction in association compared to the wildtype. Moreover, the proteins that were increased in association were analogous to those identified in the preceding Scp160 variants. An additional heat shock protein, Ssa1, was identified (Figure 37).

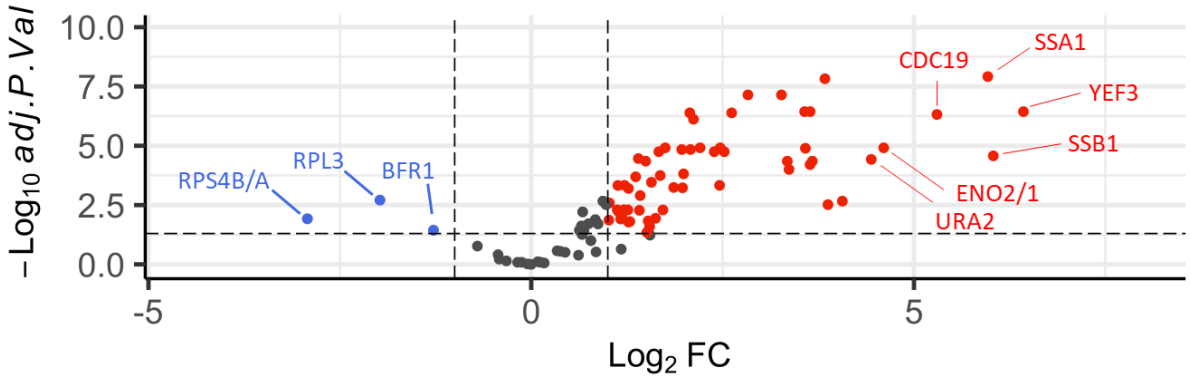


Figure 37: Volcano plot showing protein-protein interaction comparison of Scp160ΔKH5 to full-length Scp160; blue: reduced associated proteins, red: increased associated proteins; black: no significant change in interaction.

In conclusion, the alterations in protein-protein interactions in terms of loss of interaction partners thus far appear to be relatively mild. It is noteworthy that the number of increased interactions significantly exceeds that of reduced associations. Of the few significant protein-protein interaction changes, there are numerous instances of overlap, such as those observed with Mis1 or Bfr1, which indicate that the effects of N-terminal KH deletion are similar.

It has been demonstrated that the removal of the final two KH domains (Figure 38) results in a reduction in the protein's capacity to bind to polyribosomes (Lang, 2000). The results of my analysis of the Co-IP for this mutant are consistent with this finding.

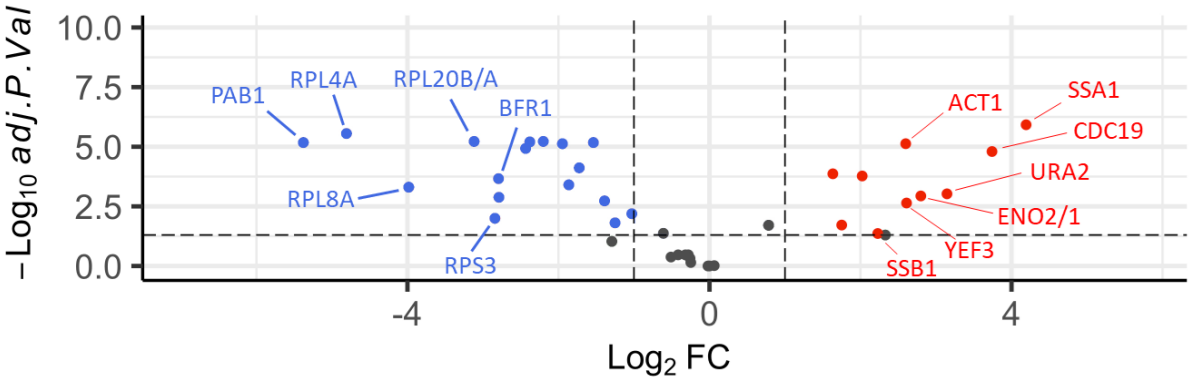


Figure 38: Volcano plot showing protein-protein interaction comparison of Scp160ΔKH13-14 to full-length Scp160; blue: reduced associated proteins, red: increased associated proteins; black: no significant change in interaction.

The levels of Bfr1, Pab1, and ribosomal proteins are markedly diminished in comparison to the wildtype, suggesting that the entire series of IP experiments has yielded meaningful outcomes. The highest increased-associated proteins for this truncation have already been identified in previous Scp160 variants.

A summary of the observed increases and decreases in associated proteins is provided in Table 23. The data are presented in descending order, corresponding to the ranks of the proteins associated with Scp160 variants lacking KH domains. The rank represents the position of the protein in terms of total abundance in the corresponding dataset and, as such, can be interpreted as an indirect indicator of binding preference. In the case of the red numbers, the higher the rank, the more abundant the corresponding protein was in the strain expressing the KH domain-deleted Scp160 protein in comparison to the full-length Scp160 protein. In contrast, the opposite is true for blue numbers. The grey rows indicate data that should be disregarded due to the presence of similar values in the negative control.

Table 23: Summary in increased and decreased protein association in comparison to the wildtype Scp160; red: protein has a higher association to the Scp160 variants in comparison to the wildtype Scp160; blue: protein has a lower association to the Scp160 variants lacking KH domains in comparison to the wildtype Scp160; ranked positions sorted in terms of higher/lower association with Scp160.

	Scp160-					
	ΔKH1-2	**ΔKH1-5**	**ΔKH4-5**	**ΔKH5**	**ΔKH13-14**	**yeGFP**
protein	**Ranking position sorted in terms of higher/lower association with Scp160**					
Act1	38	41	20	14	6	4
Adh1	23	36	14	11	8	10
Bfr1		1	1	3	6	
Ccd19	5	4	5	4	2	2
Dbp2	9	7	29	47		
Eft1	18	33	13	12	9	9
Eno1/2			6	5	4	6
Imd3	10	22	18	31		5
Mis1	3	1	7	15		
Pab1	1				1	2
Pdc1			8	8		5
Scp160						1
Ssa1	6	2	1	3	1	1
Ssb1			3	2	7	8
Tdh3	1	9	12	10		3
Tef1			31	7		
Ura2	4	20	2	6	3	
Yef3	7	3	4	1	5	7
Rpl17A		59	80		10	4
Rpl20A/B	22	18	50	53	4	
Rpl23A/B	24	11	9	9	10	13
Rpl25	21	17	38	30		10
Rpl3	2	2		2		
Rpl4A		13			2	2
Rpl6A	15	10	24	24		
Rpl6B	31	32	82	68	9	
Rpl7A					7	
Rpl8B	16	8	62		3	
Rpl9A	2	5	15	23	15	
Rpp0	17					9
Rpp2B	8	6	10	20		
Rps16A/B	40	62	59	59	8	
Rps18A/B			81		16	6
Rps20	3		47			
Rps3			30		5	7
Rps4A/B				1		
Rps5	4		85			
Rps7B			75	64	17	8

An analysis of the unique protein-protein interactions exhibited by the Scp160 variants lacking KH domains reveals the presence of a number of additional proteins not found in the full-length Scp160 (Table 24). The sole exception to this observation is Scp160ΔKH13-14, wherein the full-length Scp160 exhibits a higher number of PPIs.

The categorization of these unique PPIs was undertaken to facilitate the identification of their potential functional domains. The categorization of RNA is based on three overarching classifications: "molecular binding", "cellular components" and "biological processes" and was summarized in Table 24. The classification was achieved using g:profiler (Kolberg et al., 2023). The lowest number of unique PPIs is assigned to Scp160ΔKH1-2, where only 32 unique interactions were found. These interactions do not appear to stem from distinct classifications or functions. In contrast, Scp160 lacking the first five KH domains exhibits 76 unique PPIs, although only a limited number can be ascribed to specific functions. A small number of these can be attributed to RNA binding, ribosome biogenesis and membraneless organelles. Furthermore, Scp160ΔKH5 has been observed to bind an additional 111 proteins in comparison to the wildtype Scp160. A significant proportion of these proteins are involved in binding to other proteins and to the cytoplasm. The highest number of unique PPIs is exhibited by Scp160 lacking KH domain number four and five, with a total of 192. These interactions are attributable to the binding of diverse RNA species, ribosomal proteins, proteins localized to the nucleus and cytoplasm, and numerous additional factors.

The C-terminal truncation variant, lacking KH domain 13 and 14, differs from the other KH deleted variants (Table 24). In this case, the full-length protein displays a higher number of unique PPIs in comparison to Scp160ΔKH13-14. These unique proteins can be attributed, among other factors, to RNA binding, ribosomes and membraneless organelles.

Table 24: Number of unique protein protein interactions found for each Scp160 variant lacking KH domains in comparison to the full-length Scp160; categorization of RNA in regard to metabolic components, cellular components and biological processes; used program: g:profiler; only relevant categories were displayed in this table, thus not all unique PPI were assigned to a relevant category.

		Scp160ΔKH1-2	**Scp160**	**Scp160ΔKH1-5**	**Scp160**	**Scp160ΔKH4-5**	**Scp160**	**Scp160ΔKH5**	**Scp160**	**Scp160ΔKH13-14**	**Scp160**
	total number of unique PPIs	**32**	**0**	**78**	**2**	**192**	**0**	**111**	**3**	**3**	**51**
metabolic functions	structural constituent of ribosome					29					17
	Protein binding					122		61			
	RNA binding			13		64					16
	mRNA binding			7		27					6
	tRNA binding			4							
	rRNA binding					14					
	snoRNA binding					11					
	ATP binding					27					
biological processes	ribosome biogenesis			10		54		18			13
	gene expression					119					22
	translation					52					19
	RNA processing					47					
	nuclear transport					22					
	RNA biosynthetic process					67					
Cellular components	ribosome			7		34					18
	membraneless organelle			17		109		44			22
	nucleus					101					
	nucleolus					40					7
	cytoplasm					159		84			
	cytosol					65		33			18

Deletion of N-terminal KH domains impact the RNA interaction of Scp160

Detection of Scp160p-bound RNAs via Cross-Linking Analysis of cDNAs (CRAC)

Scp160 has been classified as an RNA-binding protein with multiple target RNAs (Hogan et al., 2008). It is thought to act as a nonessential factor of the translational machinery (Baum et al., 2004). The mechanism with which Scp160 binds RNA is not yet understood. A deletion of individual KH domains could potentially impact the RNA-binding behavior of the protein. In collaboration with the Bohnsack laboratory of the Department of Molecular Biology at Göttingen University, CRAC experiments were conducted (Granneman et al., 2009). In these experiments, protein-RNA crosslinks are introduced by ultraviolet light, the corresponding RNA-protein complexes are purified under denaturing conditions, and the RNAs are subsequently released and identified via high-throughput sequencing. In order to conduct the experiment, the strains were tagged with an His6-TEV-Protein A (HTP) tag. This tag was utilized in the development of the method to ensure that only RNAs that were covalently linked to the bait proteins were purified. In order to analyze the interaction between RNA and the protein in question (Scp160), a strain expressing the full-length Scp160 was compared with four different strains in which the protein has been truncated or KH domains have been deleted. The changes in RNA types binding to Scp160 and its KH deleted variants is displayed in Figure 39. The RNAs found to interact with the full-length Scp160 correspond to the previously described types of Scp160 binding RNAs (Figure 40): A significant proportion of mRNAs were found to be associated with components of the secretory pathway, ER-Golgi, and cell wall (Hirschmann et al., 2014; Sezen et al., 2009). Furthermore, the presence of the current ribosomal RNAs is not unexpected, given the close proximity of Scp160 and ribosomes. Prior experiments have demonstrated the binding of Scp160 to rRNA *in vitro* (Weber et al., 1997). Furthermore, smaller RNA clusters of vacuolar, granule, peptidase, and mitochondrial RNAs were identified. The involvement of vigilin in the translation of secreted proteins was not only assigned to the yeast vigilin Scp160 but also to human HDLBP (Zinnall et al., 2022). Additionally, a weak binding to mitochondrial RNAs was observed (Figure 40). For the wildtype Scp160, a high number of mRNAs were identified, with a lower number of ncRNA, snoRNA, and transposable elements (Figure 39).

The overall binding of Scp160 and its KH deletion versions exhibits notable discrepancies in the quantity of detected RNA, as illustrated in (Figure 39). The full-length Scp160 yielded approximately 30,000 identified RNA sequences encoding proteins (Figure 39; orange bars). The deletion of the KH domain 1 and 2 results in a slight decrease in protein-coding RNAs. However, in comparison to the other samples, the number of protein-coding RNAs in Scp160ΔKH1-2 is comparable to that of full-length. The number of protein-coding RNAs identified in the KH domain deletion 13-14 was reduced

by approximately one-third. Furthermore, Scp160ΔKH4-5 and Scp160ΔKH5 exhibited even greater impairment in this regard. The number of protein-coding RNAs bound to the protein was reduced by approximately one-third in comparison to the wild-type Scp160 protein. In the case of the full-length Scp160, the mRNAs constitute the most prevalent RNA type, albeit in low numbers (Figure 39). Additionally, ncRNA, snoRNA, and transposable elements have also been identified as being bound by the protein. It is noteworthy that the Scp160 protein lacking KH domains one and two exhibited a higher number of ncRNA, rRNA, snoRNA, snRNAs, transposable elements, and tRNAs. It is evident from the CRAC data that the level of rRNA is elevated to a greater extent in the Scp160ΔKH1-2 sample than in the other Scp160 variants (Figure 39). The Scp160ΔKH13-14 protein displays a higher number of snRNAs and tRNAs, as does the Scp160ΔKH5 protein. The highest number of small nucleolar RNAs (snoRNAs) was identified in the Scp160ΔKH4-5 sample (Figure 39).

The results of the CRAC experiments are presented in the accompanying Figure 40-44, which depicts cell component cluster analysis. Cell component clusters demonstrate the alteration or loss of RNAs encoding proteins that are associated with a particular cellular compartment or perform analogous functions. These clusters comprise RNAs that encode specific components of compartments and complexes, which interact with Scp160. Furthermore, these clusters can be linked to other clusters to form superclusters. The super-clusters collectively constitute the RNA interactome of Scp160. In general, the Scp160ΔKH1-2 strain exhibited minimal to no functional deficits relative to the wild-type strain, which aligns with the initial observation that this truncation appears to have a limited impact on the overall functionality of the protein. The deletion strains Scp160ΔKH4-5 and Scp160ΔKH5 were selected for analysis due to their phenotypic similarities to the full deletion of Scp160. The Scp160ΔKH13-14 was selected as a control, given its reported impact on the protein's function (Li et al., 2004; Baum et al., 2004; Zhou et al., 2008; Gelin-Licht, 2012). All Scp160 mutants were previously examined for alterations in protein-protein interactions, thus enabling a comparison between changes in RNA- and protein interactions. The CRAC experiments were conducted in duplicate. The second attempt yielded results that were of lesser quality than those of the first experiment, resulting in a decline in the number of identified RNAs.

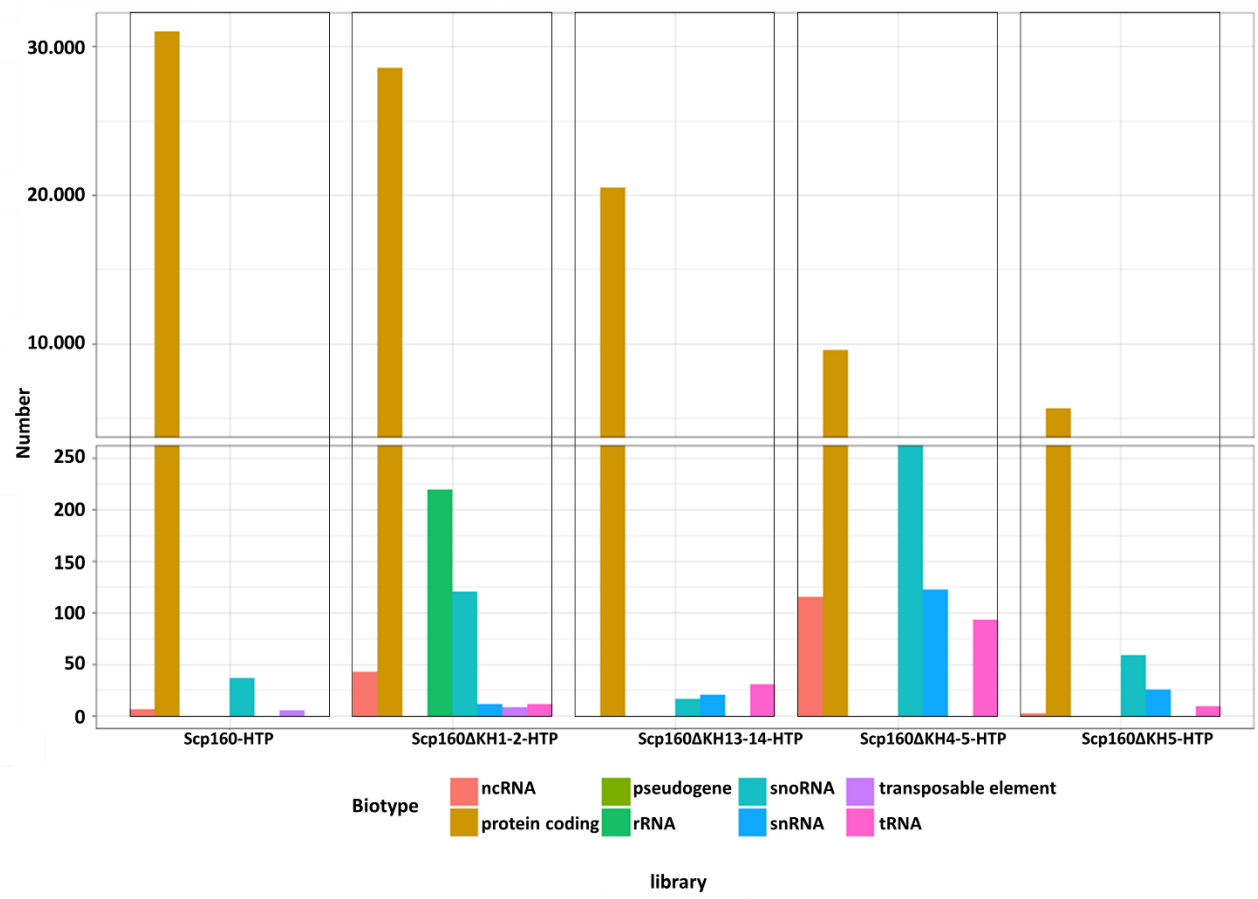


Figure 39: Biotype fractions of RNAs for Scp160-HTP and the KH domain truncated/deleted Scp160 proteins; RNAs were identified in CRAC attempt I; Y-axis: number of found RNAs; X-axis: HTP tagged Scp160 and KH truncated/deleted Scp160 variants analyzed.

In order to establish a baseline for comparison, a CRAC experiment was conducted initially on the full-length Scp160. In accordance with prior findings (Hirschmann et al., 2014; Nerome et al., 2020), an examination of the results indicated that mRNAs encoding proteins of the secretory pathway constituted the most prominent cluster (Figure 40, supercluster a). This encompasses mRNAs that code for proteins of the endoplasmic reticulum and the Golgi apparatus. Subsequently, mRNAs encoding vacuolar proteins (supercluster k; seven RNA clusters identified) were identified. Moreover, a cluster of mRNAs encoding components of the pre-initiation complex, including pre-ribosomes and polysome-associated proteins, was identified (superclusters d, h, and g) (Cheng et al., 2018). Furthermore, the association to RNAs connected to stress granule formation, was also observed (superclusters j). The binding of mRNAs encoding mitochondrial proteins (supercluster b) and

members of small nucleolar RNAs (supercluster d) was already observed in hepatic mouse vigilin (Mobin et al., 2016).

It has been demonstrated that vigilin interacts with signal peptide peptidases (SPPs) in the context of the proteasome (Lu et al., 2012). In the yeast organism, these peptidases are found within a complex designated the Signal Peptidase Complex (SPC). The complex is composed of four proteins: Spc1, Spc2, Spc3, and Sec11 (Böhni et al., 1988; YaDeau et al., 1991). The RNA of Spc1 was not identified in the CRAC analysis. It should be noted that the interaction between vigilin and the SPP has only been demonstrated at the protein-protein interaction level.

In the second CRAC experiment, the number of identified clusters for the full-length Scp160 was reduced (49 versus 77 in experiment #1), yet the cluster distribution remained comparable. However, due to the inferior quality compared to the first experiment, not only are cluster sizes smaller, but whole clusters are lost, including proteasome, lipid droplets, and chaperones. The most notable contribution of the second cluster list is the addition of a cell wall cluster present only in the Scp160ΔKH13-14 dataset (Figure 44).

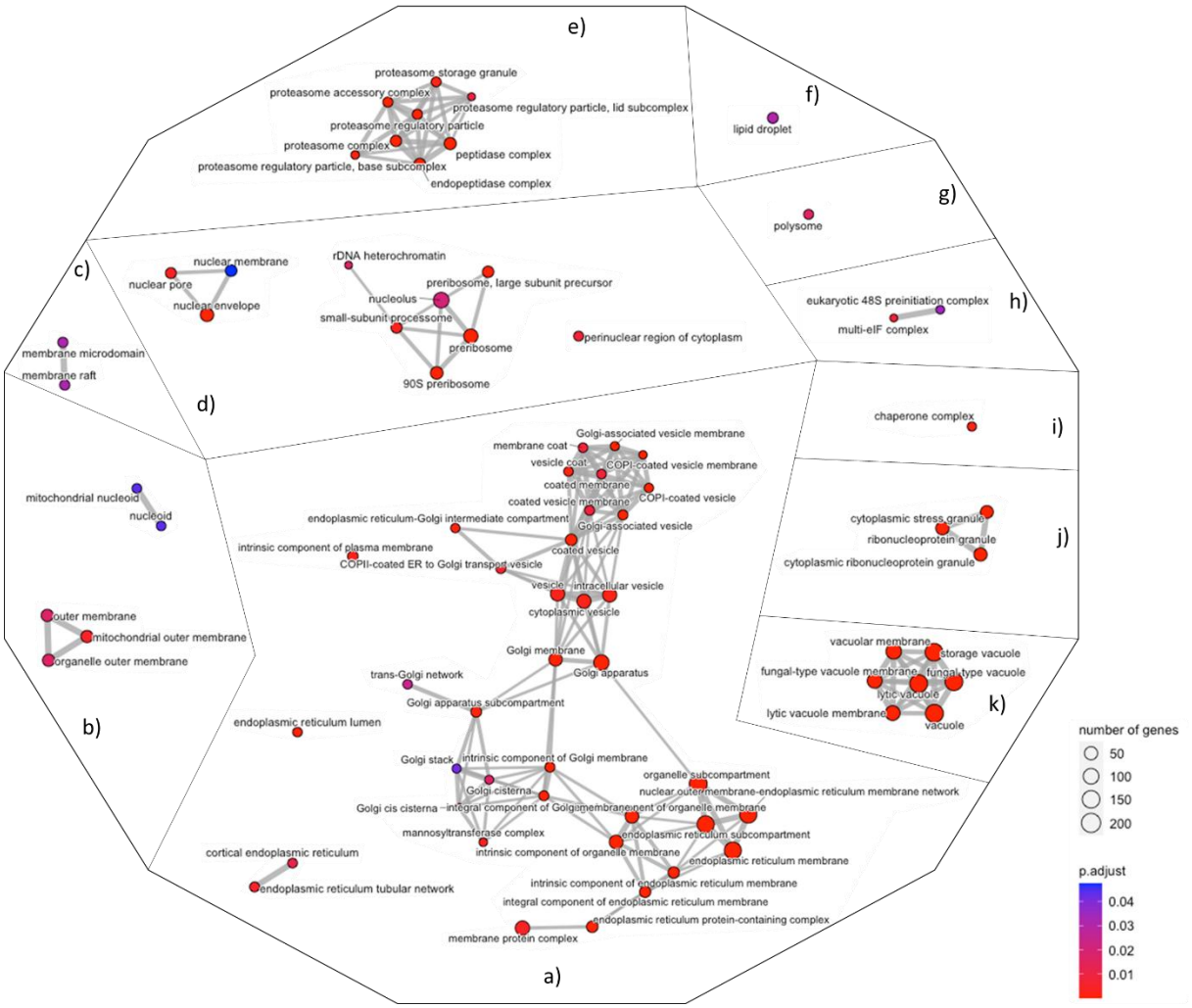


Figure 40: RNA interactome of full-length Scp160; Each circle represents an RNA cluster (RNA with related function or similar location). Connected clusters form a supercluster. All superclusters form the total RNA interactome. Circle size displays the number of genes in a cluster. Color reflects adjusted p-value of the cluster, ranging from red (0.01) to blue (0.04). The RNA-Interactome displayed here only includes CRAC experiment #1.

An overview of identified RNA clusters in the full-length Scp160 (for both CRAC experiments) is summarized in Table 25.

Table 25: Clustered RNAs for the full-length Scp160 as seen in Figure 40; Comparison between both CRAC attempts in terms of found RNA clusters.

		CRAC I	CRAC II
	RNA cluster	cluster count	
a)	Endoplasmic reticulum and Golgi network	37	26
b)	mitochondrial	5	1
c)	Membrane	2	2
d)	nuclear region	10	2
e)	Proteasome	8	0
f)	lipid droplet	1	0
g)	Ribosome	1	1
h)	translation initiation	2	1
i)	Chaperone	1	0
j)	stress granule	3	3
k)	Vacuole	7	7
	cell wall	0	6

The cluster distribution of mRNAs crosslinked to Scp160ΔKH1-2 is comparable to that of the wildtype, although not identical (Figure 41). Three notable discrepancies are evident: The loss of lipid droplet clusters (Figure 40, f) and the presence of a specific set of mRNAs encoding components of the ATPase complex (Figure 41, f) are notable observations. As evidenced by the second replica, several clusters are also lost, including those encoding mitochondrial, nuclear region, proteasome, ATPase complex, ribosome, translation initiation, chaperone, and different kinds of granules (cytoplasmic stress granule, ribonucleoprotein granule, and cytoplasmic granule ribonucleoprotein). However, the list also includes three new clusters: actin cytoskeleton, mannosyltransferase, and the extracellular region. The mannosyltransferase is typically included within the cluster of the endoplasmic reticulum and the Golgi apparatus.

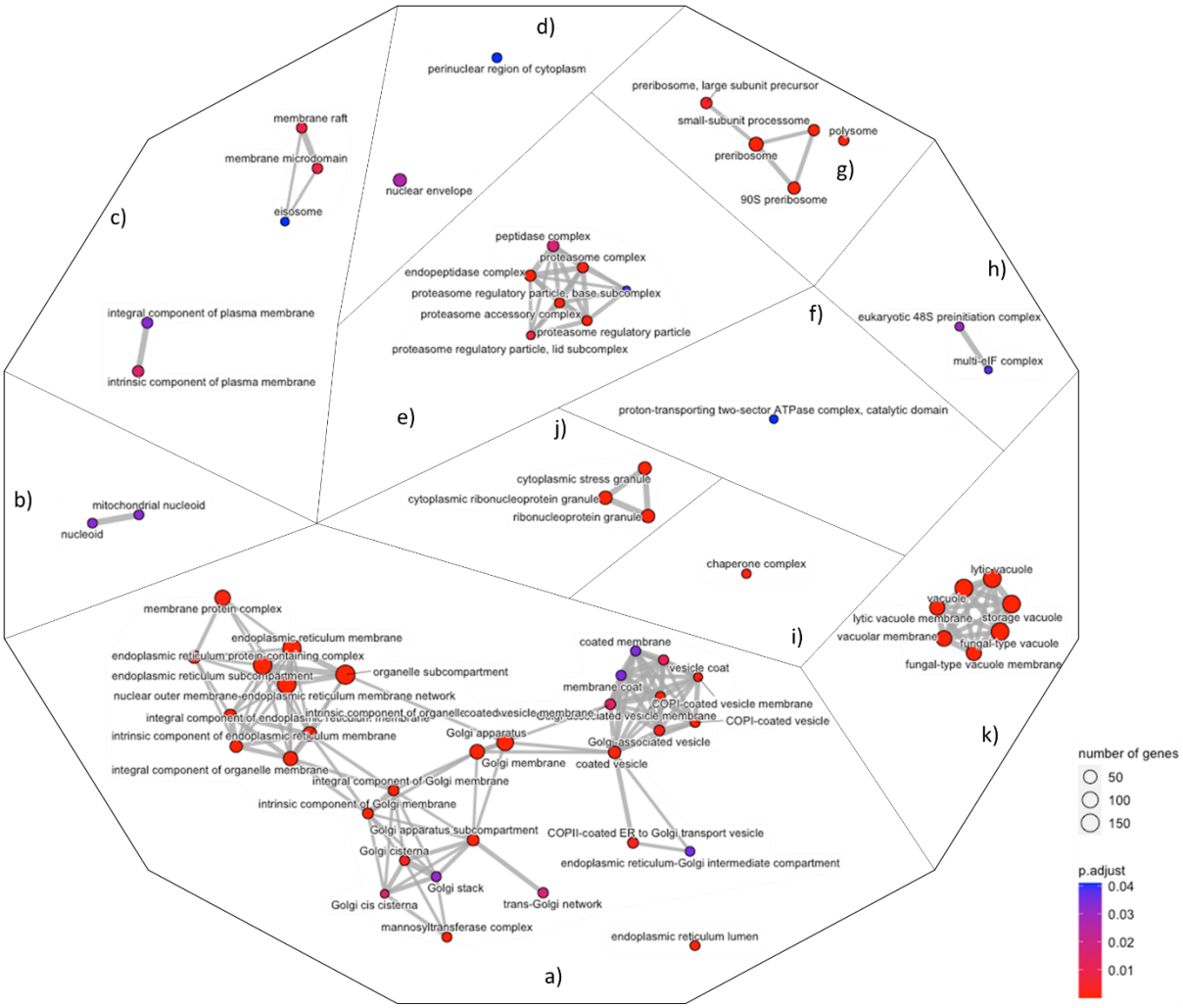


Figure 41: RNA interactome of Scp160ΔKH1-2; single circle represents an RNA cluster. Connected clusters form a supercluster. All superclusters form the total RNA interactome. Circle size displays the number of genes in the cluster. Color reflects the adjusted p-value of the cluster: from red 0.01 to blue 0.04. RNA-Interactome displayed here only includes CRAC experiment #1.

An overview of identified RNA clusters in the Scp160 variant lacking KH domains one and two (for both CRAC experiments) is summarized in Table 26.

Table 26: Interacting RNAs with Scp160ΔKH1-2 as seen in Figure 41; Comparison between both CRAC attempts in terms of found RNA clusters.

		CRAC I	CRAC II
	RNA cluster	cluster count	
a)	Endoplasmic reticulum and Golgi network	32	16
b)	mitochondrial	2	0
c)	Membrane	5	2
d)	nuclear region	2	0
e)	Proteasome	7	0
f)	ATPase complex	1	0
g)	Ribosome	5	0
h)	translation initiation	2	0
i)	Chaperone	1	0
j)	Granule	3	0
k)	Vacuole	7	7
	actin cytoskeleton	0	1
	Mannosyltransferase	0	1
	extra cellular region	0	2

In comparison to the full-length Scp160 and the Scp160ΔKH1-2 strains, the number of clusters and the cluster sizes identified in the Scp160ΔKH4-5 strain are markedly diminished (Figure 42). The clusters pertaining to the endoplasmic reticulum/Golgi network, membrane, nuclear region, translation initiation, chaperone, and vacuole are all absent. The overall cluster number of 15, in comparison to the 77 clusters observed in Scp160 and 67 in Scp160ΔKH1-2, clearly demonstrates that this mutant Scp160 has lost a significant portion of its originally bound mRNAs. The complete absence of the largest interconnected cluster of the ER and Golgi network is particularly remarkable, given that this network represents almost half of all interacting mRNAs in Scp160 and Scp160ΔKH1-2 in the first CRAC attempt (Table 25 and Table 26). It is noteworthy that the ER and Golgi network cluster was present in the second CRAC experiment, as well as a new ribosomal protein mRNA cluster. In the second CRAC experiment the ER and Golgi network -/ ribosome clusters replace all clusters from the previous attempt (Table 27).

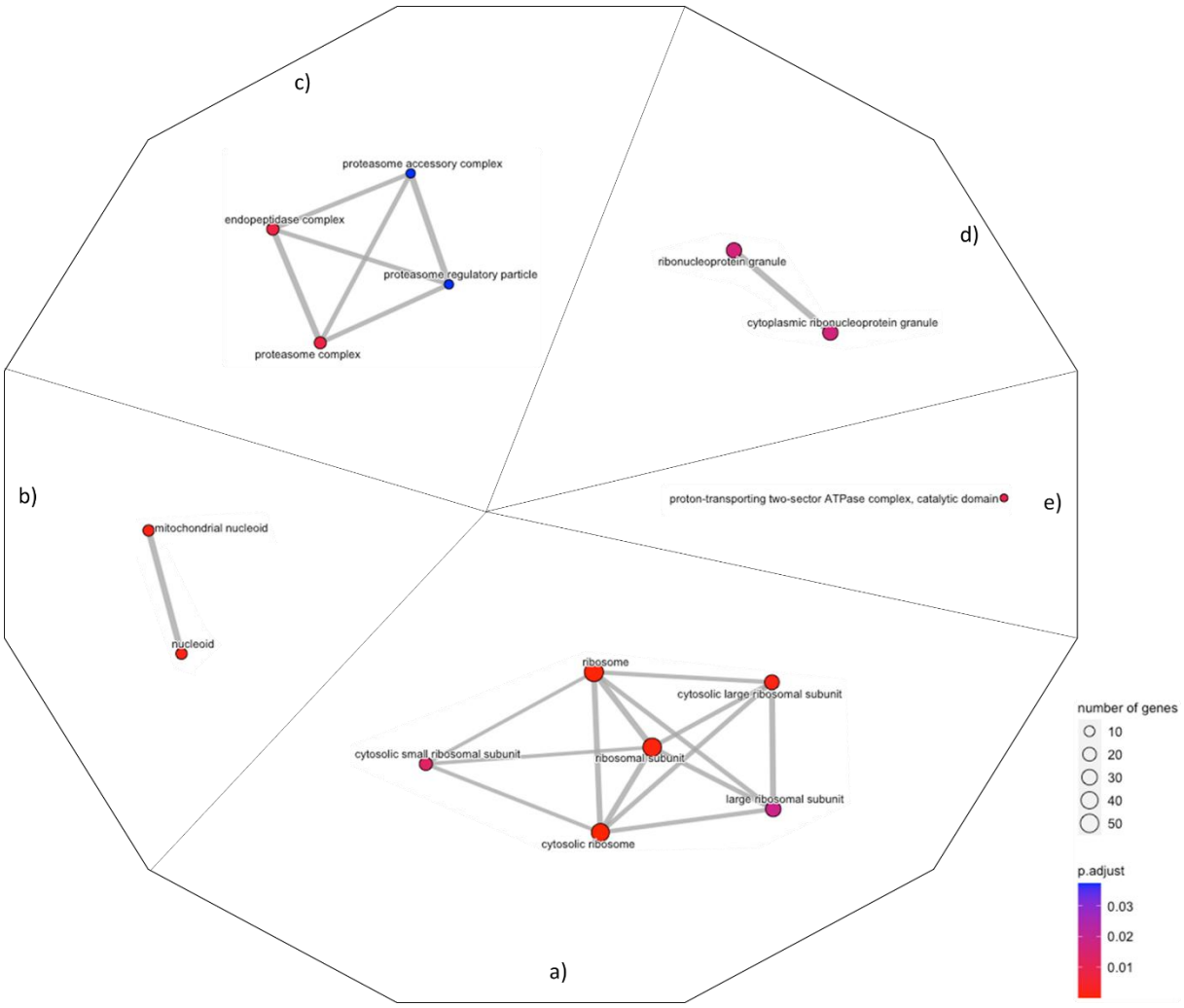


Figure 42: RNA interactome of Scp160ΔKH4-5; display of interacting RNAs with the Scp160ΔKH4-5 protein. Single circle represents an RNA cluster. Connected clusters form a supercluster. All superclusters form the total RNA interactome. Circle size displays the number of genes in the cluster. Color reflects the adjusted p-value of the cluster: from red 0.01 to blue 0.04. RNA-Interactome displayed here only includes CRAC experiment #1.

An overview of identified RNA clusters in the Scp160 variant lacking KH domains four and five (for both CRAC experiments) is summarized in Table 27.

Table 27: Interacting RNAs with Scp160ΔKH4-5 as seen in Figure 42; Comparison between both CRAC attempts in terms of found RNA clusters.

		CRAC I	CRAC II
	RNA cluster	cluster count	
a)	Ribosome	6	0
b)	Mitochondrial	2	0
c)	Proteasome	4	0
d)	Granule	2	0
e)	ATPase complex	1	0
	Endoplasmic reticulum and Golgi network	0	7
	Ribosome	0	2

The number and distribution of clusters in the strain expressing the Scp160 protein lacking the KH domain number five (Figure 43) can be visualized as an intermediate between the Scp160ΔKH4-5 mutant and Scp160ΔKH1-2/wildtype. The ER/Golgi network clusters are not absent, but they are smaller in size and lack a number of mRNAs. The CRAC data for Scp160ΔKH5 also demonstrate a loss of the translation initiation cluster, and the clusters associated with the nuclear domains. Furthermore, a transport vesicle membrane cluster (d) is observed. In the CRAC experiments conducted with the strain expressing full-length Scp160, the supercluster of the ER/Golgi network was included, but due to the loss of connecting clusters, it was isolated in Scp160ΔKH5 (Figure 43). In addition, the CRAC experiment #2 yielded a considerably shorter list of clusters, exhibiting a reduction in both the number and diversity of clusters. Six clusters are absent (ubiquitin ligase, transport vesicle membrane, proteasome, ATPase complex, nucleoid, and chaperone), yet three new clusters (translation initiation, extracellular region, and mitochondrial) are present in other cluster sets, offering a degree of compensation.

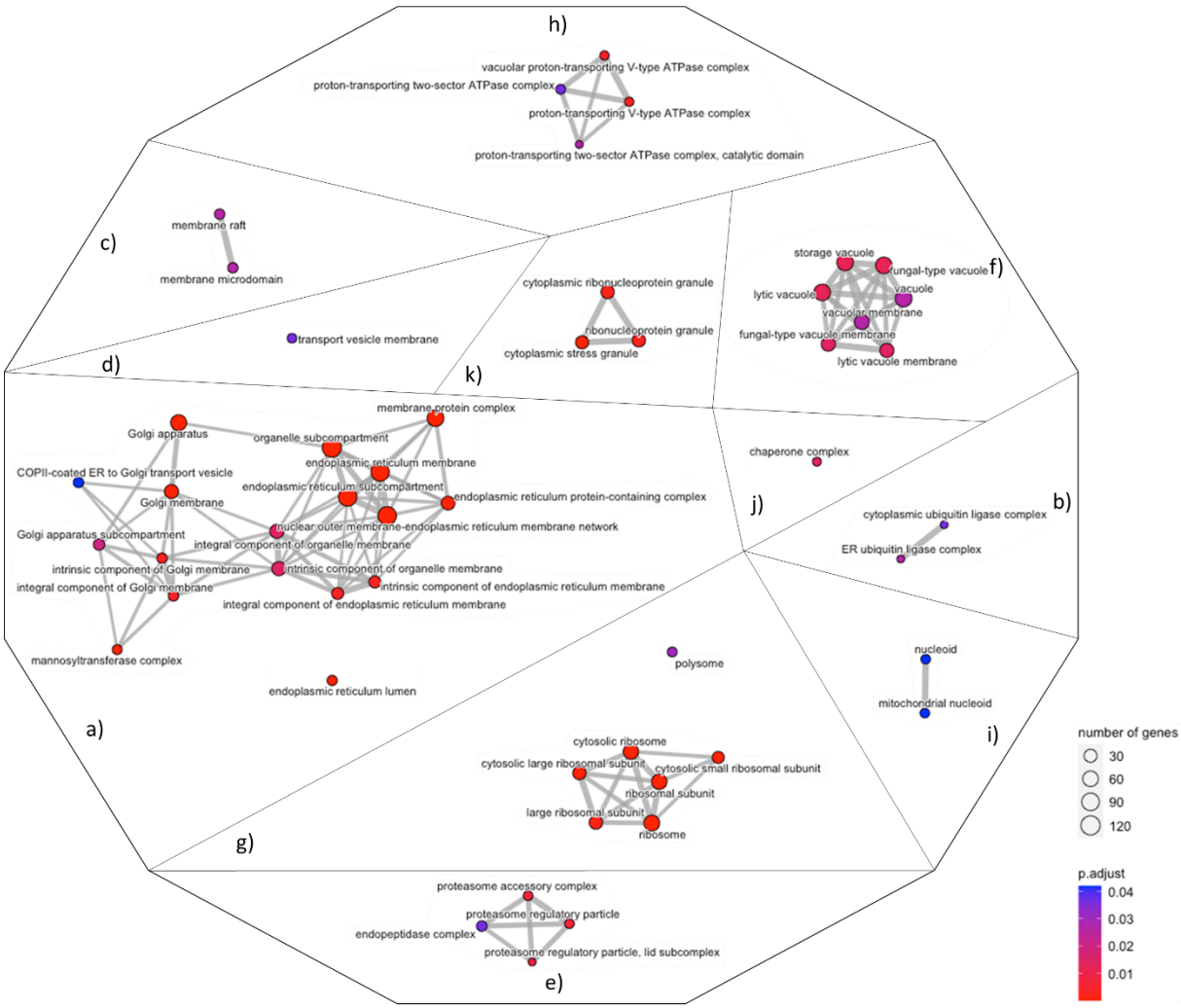


Figure 43: RNA interactome of Scp160ΔKH5; display of interacting RNAs with the Scp160ΔKH5 protein. Single circle represents an RNA cluster. Connected clusters form a supercluster. All superclusters form the total RNA interactome. Circle size displays the number of genes in the cluster. Color reflects adjusted p-value of the cluster: from red 0.01 to blue 0.04. RNA-Interactome displayed here only includes CRAC experiment #1.

An overview of identified RNA clusters in the Scp160 variant lacking KH domain five (for both CRAC experiments) is summarized in Table 28.

Table 28: Interacting RNAs with Scp160ΔKH5 as seen in Figure 43; Comparison between both CRAC attempts in terms of found RNA clusters.

		CRAC I	CRAC II
	RNA cluster	cluster count	
a)	Endoplasmic reticulum and Golgi network	18	7
b)	ubiquitin ligase	2	0
c)	Membrane	2	1
d)	transport vesicle membrane	1	0
e)	Proteasome	4	0
f)	Vacuole	7	7
g)	Ribosomal	7	2
h)	ATPase complex	4	0
i)	Nucleoid	2	0
j)	Chaperone	1	0
k)	(stress) granule	3	3
	translation initiation	0	2
	extracellular region	0	1
	Mitochondrial	0	1

In conclusion, it can be stated that the removal of particular KH domains affects the RNA binding behavior of Scp160. The influence is not contingent on the number of KH domains deleted, but rather on their position. For example the deletion of two KH domains does not harbor the same impact: To illustrate this, the deletion of KH domains one and two (Figure 41) does not have the same effect as the deletion of KH domains four and five (Figure 42).

It was unexpected that the CRAC of the Scp160ΔKH13-14 mutant (Figure 44) exhibited no significant differences compared to the full-length protein. Notably, the majority of ER-Golgi clusters remained intact, despite previous research in yeast indicating that the deletion of KH13-14 led to a disruption in vigilin capability of RNA binding (Brykailo et al., 2007; Hirschmann et al., 2014). Additionally, the mitochondrial cluster, along with the nuclear region and proteasome clusters, were absent. However, in the case of Scp160ΔKH13-14, two distinctive RNA clusters were identified: cell wall and cell septum.

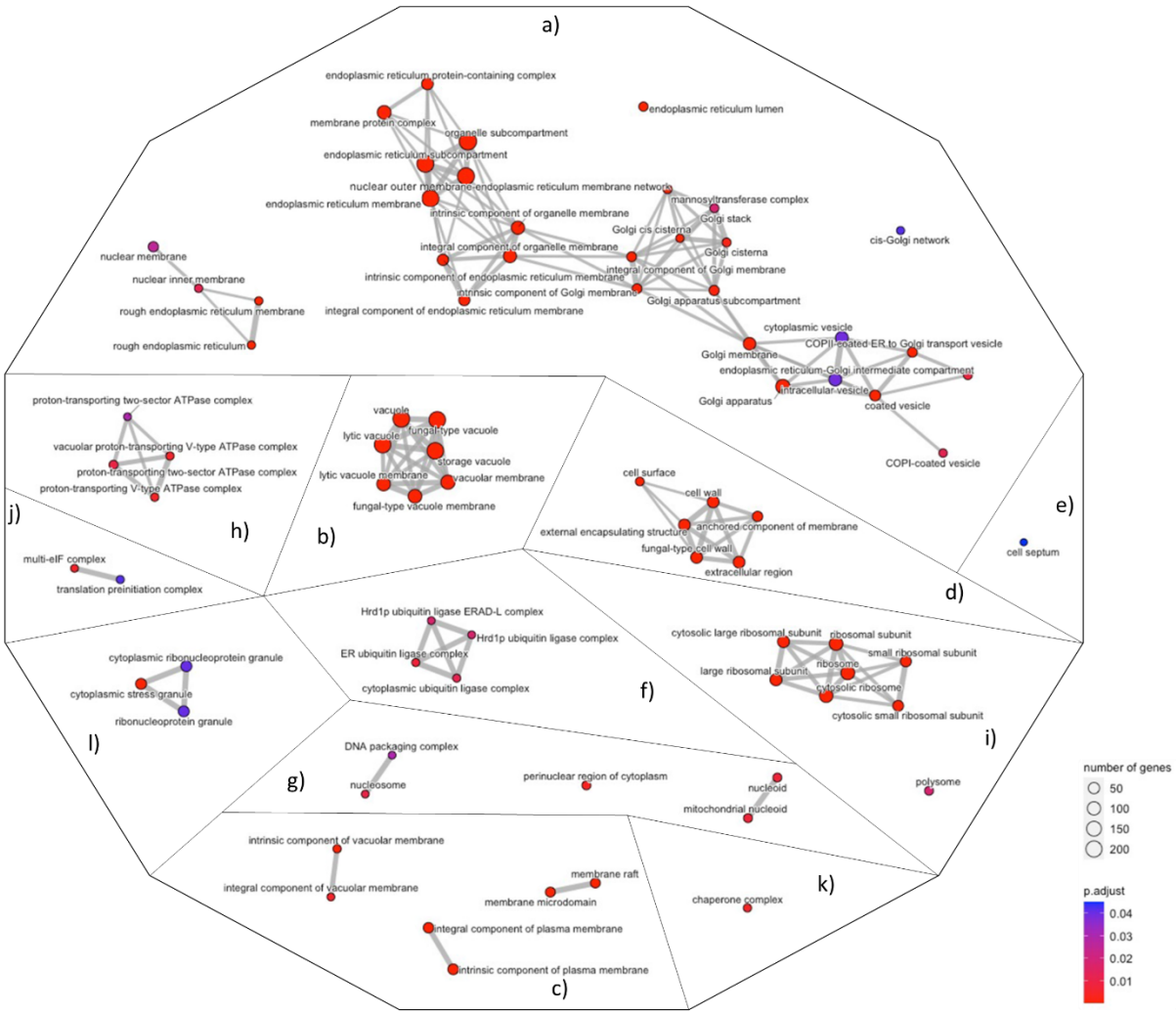


Figure 44: RNA interactome of Scp160ΔKH13-14; display of interacting RNAs with the Scp160ΔKH13-14 protein. Single circle represents an RNA cluster. Connected clusters form a supercluster. All superclusters form the total RNA interactome. Circle size displays the number of genes in the cluster. Color reflects the adjusted p-value of the cluster: from red 0.01 to blue 0.04. RNA-Interactome displayed here only includes CRAC experiment #1.

An overview of identified RNA clusters in the Scp160 variant lacking KH domains 13 and 14 (for both CRAC experiments) is summarized in Table 29.

Table 29: Interacting RNAs with Scp160ΔKH13-14 as seen in Figure 44; Comparison between both CRAC attempts in terms of found RNA clusters.

		CRAC I	CRAC II
	RNA cluster	cluster count	
a)	Endoplasmic reticulum and Golgi network	31	19
b)	Vacuole	7	7
c)	Membrane	6	4
d)	cell wall	6	6
e)	cell septum	1	0
f)	ubiquitin ligase	4	0
g)	nucleoid/perinuclear	5	1
h)	ATPase complex	4	2
i)	Ribosome	8	0
j)	translation initiation	2	0
k)	Chaperone	1	0
l)	stress granule	3	0
	Mitochondrial	0	1

The results from the various Scp160 proteins were summarized in Table 30, displaying a rather irregular pattern of RNA binding.

Table 30: RNA clusters identified in the used strains expressing the KH truncated/deleted versions of Scp160 and the wildtype Scp160 for both CRAC experiments; CRAC experiments are labeled as I and II accordingly.

	Scp160		**Scp160-**							
			ΔKH1-2		**ΔKH4-5**		**ΔKH5**		**ΔKH13-14**	
	I	II	I	II	I	II	I	II	I	II
ER/Golgi network	●	●	●	●		●	●	●	●	●
mitochondrial	●	●	●		●		●	●		●
membrane	●	●	●	●			●	●	●	●
nuclear region	●	●	●							●
proteasome	●		●		●		●			
lipid droplet	●									
ATPase complex			●		●		●		●	●
ribosome	●	●	●		●	●	●		●	
translation initiation	●	●	●					●	●	
chaperone	●		●				●		●	
granule	●	●	●		●		●	●	●	
vacuole	●	●	●	●			●	●	●	●
ubiquitin ligase							●		●	
cell septum									●	●
cell wall		●						●	●	●
transport vesicle membrane							●			

The deletion of KH domains 4 and 5 has a discernible impact on the RNA binding capability (Figure 42). Thus, I was intrigued by the specific alterations in bound RNA in comparison to other Scp160 variants lacking KH domains. An UpSet plot is a visual representation of the intersections between the data sets of Scp160 and the four Scp160 variants lacking KH domains (Figure 45). This allows for the analysis of shared RNA binding events among the different mutants. A total of 232 unique genes were identified as being specific to the Scp160ΔKH4-5 strain (i.e., RNAs that were only found to associate with Scp160 in this mutant), 160 unique genes for Scp160ΔKH5, 72 unique genes for Scp160ΔKH13-14, 116 unique genes for Scp160ΔKH1-2, and 206 unique genes for the wild-type *SCP160* strain (Figure 45). The overlap between the Scp160ΔKH4-5 and Scp160ΔKH5 variants is not as extensive as anticipated, given that both exhibited comparable phenotypes in the preliminary experiments. A total of 51 genes were identified as being present in both datasets. In comparison, the overlap between the Scp160ΔKH1-2 and wildtype datasets is as high as 242. In order to ascertain the impact of the deletion of Scp160ΔKH4-5, it is necessary to compare the overlaps between the strain lacking these two domains or Scp160ΔKH5 and the other three strains used in this experiment. The number of RNAs that are common in the full-length Scp160, as well as the variant lacking KH1-2 and the variant lacking KH domains 13-14, is 182. The addition of the strain Scp160ΔKH5 results in an RNA overlap value of 141, which remains relatively high. Inclusion of Scp160ΔKH4-5 instead of Scp160ΔKH5 results in a reduction of the value to 76. The deletion of KH domain five clearly impairs the RNA binding of the mutant, but the deletion of KH domains four and five has an even more pronounced effect. The number of deleted KH domains is not a critical factor in this example; however, the deletion of specific KH domains clearly affects RNA binding in a distinctive manner.

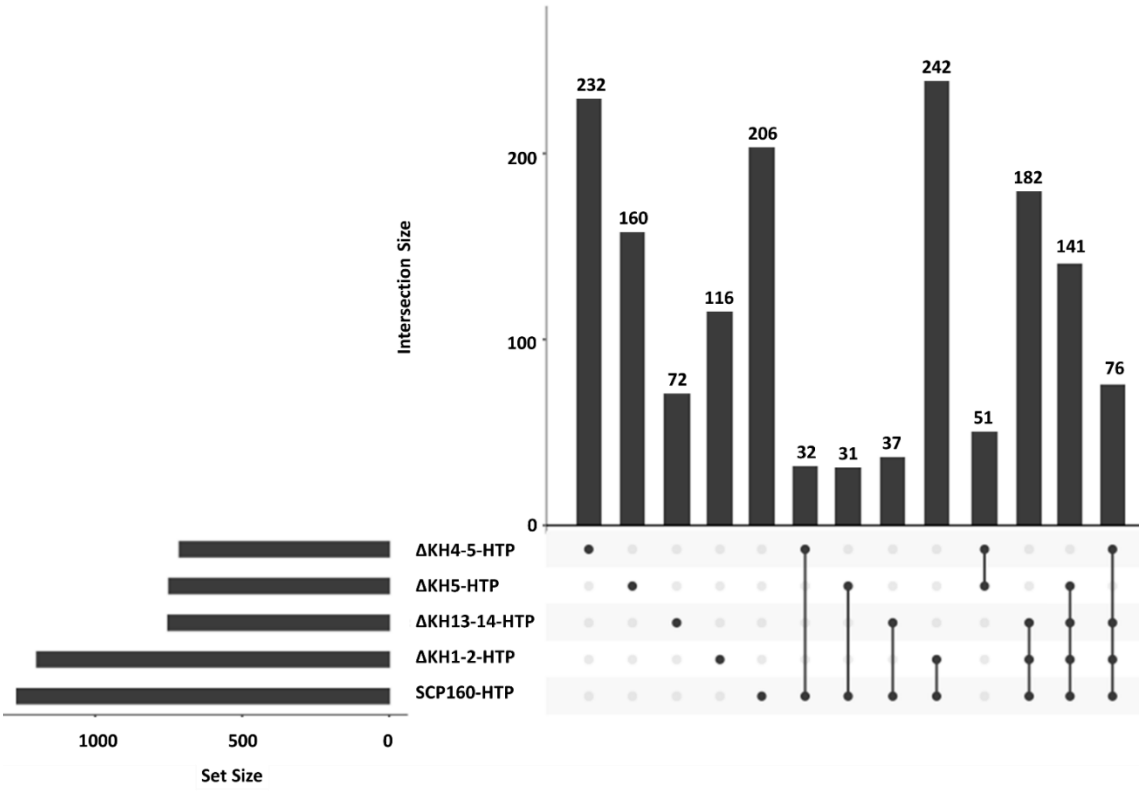


Figure 45: UpSet plot: Comparison of strains expressing KH domain truncated/deleted versions of Scp160 in terms of unique or overlapping found RNAs; upper part represent number of RNAs only found in the specific intersection; lower part depict the set size (all found RNAs in the sample) of the used data sets (left) and the intersecting data sets (Right) according to the intersecting size.

The objective of this study was to ascertain the number of lost RNAs in the Scp160ΔKH5 and SCP160ΔKH4-5 mutants relative to the full-length Scp160. A total of 1,241 RNAs were identified in the full-length Scp160. The Scp160ΔKH5 strain has 726 interactors with 317 unique RNAs, which is 409 RNA interactors in common with Scp160. In total, 681 RNA interactors were identified in Scp160ΔKH4-5, indicating a reduction in interactors compared to Scp160ΔKH5. Of these 681 interacting RNAs, 361 are unique to Scp160ΔKH4-5.

The ploidy profile of strains expressing KH deleted proteins and its implications in regard on total RNA changes by RNA-sequencing

As already mentioned (p.50 onwards), the mutants analyzed in the CRAC experiments have been observed to alter the ploidy of the cell. Given the prior observation of ploidy-dependent alterations in gene expression in yeast (Galitski et al., 1999), I sought to ascertain whether the disparate binding patterns of RNAs observed in the Scp160 mutants are attributable to fluctuations in the abundance of these RNAs. To this end, RNA sequencing was conducted by our collaborators at the University of

Göttingen. The same five strains utilized for CRAC analysis were selected for this experiment (Scp160, Scp160ΔKH1-2, Scp160ΔKH5, Scp160ΔKH4-5, Scp160ΔKH13-14).

The RNA-seq data were subjected to differential expression analysis using the DESeq2 program (Love et al., 2014). The adjusted p-value was employed for the evaluation and visualization of the data. For each gene, a statistical test was performed to determine the significance of differential expressions between the different samples. However, the aforementioned raw p-values do not account for false positive hits. The occurrence of false positives can be attributed to the simultaneous execution of numerous tests. A Benjamini-Hochberg procedure (Benjamini and Hochberg, 1995), a multiple testing correction method, was employed to control the false discovery rate (FDR), resulting in the adjusted p-value. The strains expressing the deletion variants were tested against the wildtype using the adjusted p-value, which implies greater statistical significance. With regard to the raw p-value, the Scp160ΔKH1-2 strain was found to exhibit only one negatively correlated hit in comparison to the wildtype. With respect to the adjusted p-value, no significant change could be discerned (Figure 46). In the Scp160ΔKH5 strain, five negatively correlated RNAs were identified, with only two remaining in the adjusted graph (Figure 47). A comparable pattern was observed in the Scp160ΔKH4-5 strain, with four hits identified for the raw p-value and two for the adjusted p-value. The control strain, Scp160ΔKH13-14, exhibited five negatively correlated RNAs for both the raw and adjusted p-values, as well as one positively correlated RNA for both p-values.

Volcano plots display the most significant hits, utilizing the adjusted p-value for the multicomparison correction. The significance level was determined in accordance with the decision to focus on a stringent hit filtering (Supplementary Figure 8). Significance level was chosen as follows: FC > 2 and p-value < 0.05.

There is no significant difference between the wildtype Scp160 and Scp160ΔKH1-2 in terms of binding RNA (Figure 46).

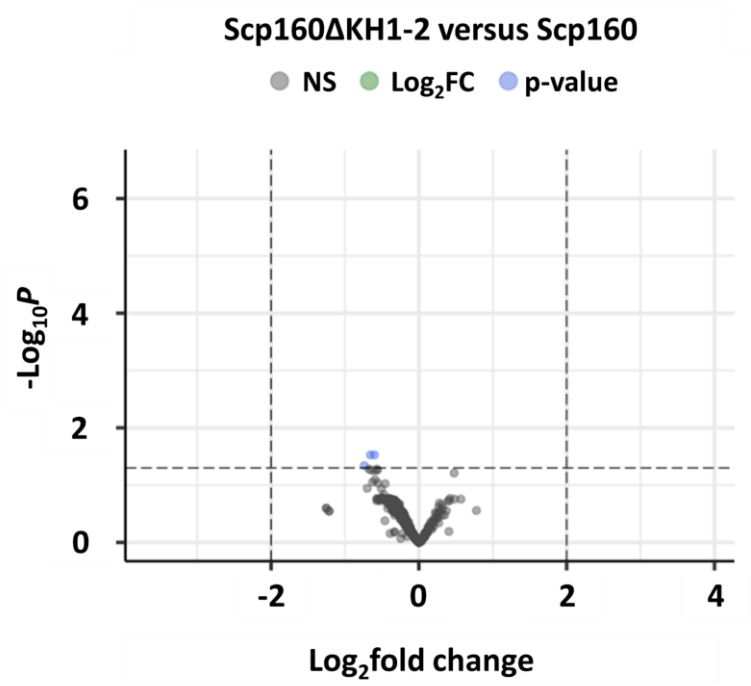


Figure 46: Volcano plot exhibiting the differences in RNA abundance between Scp160ΔKH1-2 and Scp160; blue dots: hits with FC < 2 and p-value < 0.05; green dots: hits with FC > 2 and p-value > 0.05; red dots: Significant hits with FC > 2 and p-value < 0.05; negative fold change (left) displays decreased RNA abundance in Scp160ΔKH1-2.

With only two relevant differences between the full-length Scp160 and Scp160ΔKH5, this Scp160 variant exhibits a remarkable small impact on RNA binding (Figure 47). The less bound RNAs, Dad4 (YDR320C-A) and Gim4 (YEL003W), also exhibit no striking difference in RNA binding between this strain and the full-length. Dad4 is a crucial component of the Dam1 complex, also referred to as the DASH complex. This complex plays a crucial role in the binding of microtubules to the kinetochore. It connects the kinetochores to the force generated by MT depolymerization, thereby facilitating chromosome segregation (Li et al., 2002). Gim4 is a subunit of the hetero-hexameric cochaperone prefoldin complex, which binds to cytosolic chaperonin and subsequently transfers target proteins to it (Geissler et al., 1998).

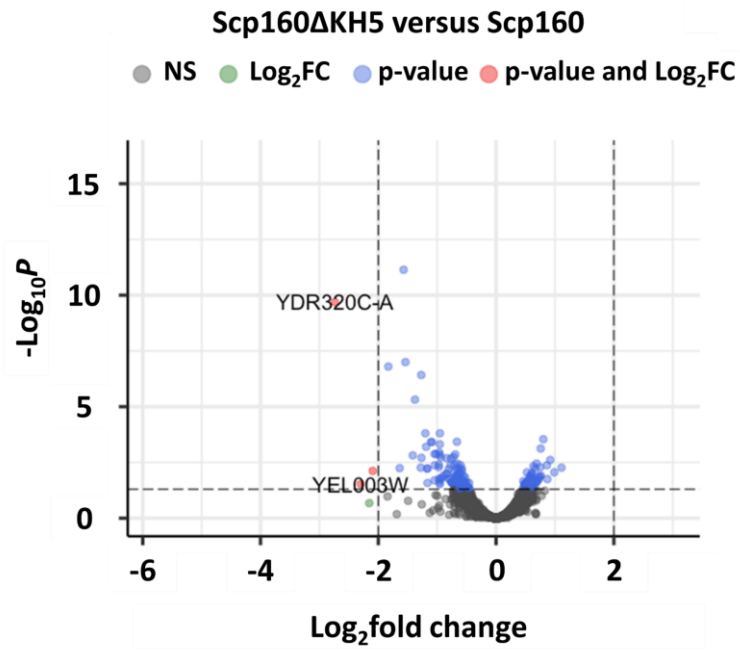


Figure 47: Volcano plot exhibiting the RNA abundance differences between Scp160ΔKH5 and Scp160; blue dots: hits with FC < 2 and p-value < 0.05; green dots: hits with FC > 2 and p-value > 0.05; red dots: Significant hits with FC > 2 and p-value < 0.05; negative fold change (left) displays decreased RNA abundance in Scp160ΔKH5.

Similarly, Scp160ΔKH4-5 (Figure 48) has two RNAs less bound by the full-length Scp160, analogous to the two observed in Scp160ΔKH5. The initial 21S rRNA (Q0158) is a mitochondrial 21S rRNA. The intron encodes the I-SceI DNA endonuclease (Jacquier and Dujon, 1985). The second is HMRA2 (YCR096C), which is a silenced copy of a2 at the mating-type cassette HMR. It displays a resemblance to Alpha2 and is indispensable when coupled with a1p for the purpose of impeding the HO endonuclease in a/α HO/HO diploid cells that possess an active mating-type interconversion system (Jensen et al., 1983).

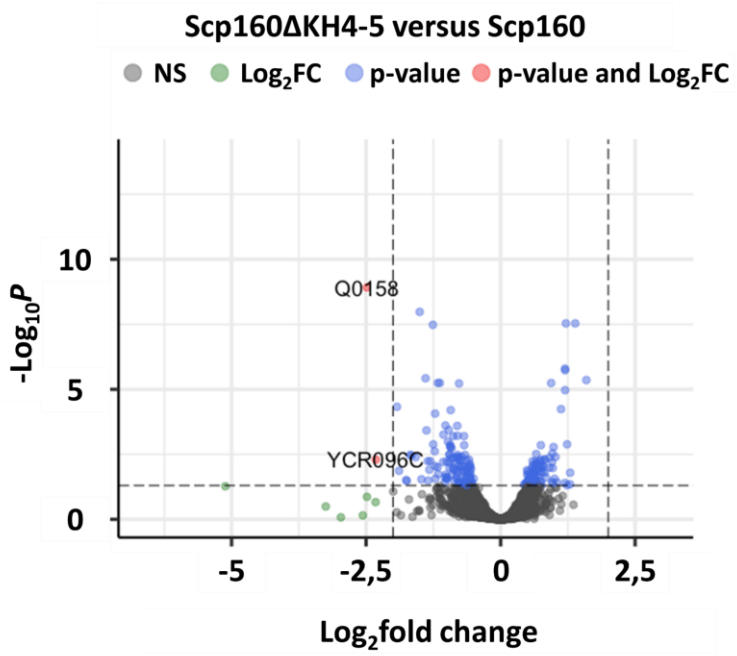


Figure 48: Volcano plot exhibiting the RNA abundance differences between Scp160ΔKH4-5 and Scp160; blue dots: hits with FC < 2 and p-value < 0.05; green dots: hits with FC > 2 and p-value > 0.05; red dots: Significant hits with FC > 2 and p-value < 0.05; negative fold change (left) displays decreased RNA abundance in Scp160ΔKH4-5.

strain exhibiting the most notable alterations in individual RNA levels is Scp160ΔKH13-14. Despite the reduction in the amount of five RNAs (YDR320C-A, YOL126C, YKL216W, YDR031W, YNCB0019C) and the increase in the amount of one RNA (YBR089W) in comparison to the full-length Scp160 protein, this strain does not exhibit an overall distinct RNA binding pattern. It is noteworthy that the number of significantly altered RNAs is identical for raw and adjusted p-values. However, the lists of RNAs are not entirely concordant when compared to each other. The RNAs that exhibited a decreased quantity were Dad4 (YDR320C-A), Mdh2 (YOL126C), Ura1 (YKL216W), Lsr1 (YNCB0019C), and Tsc3 (YBR058C-A). The protein Dad4 has already been mentioned (Figure 46). Mdh2 is a cytoplasmic malate dehydrogenase. The enzyme catalyzes the interconversion of malate to oxaloacetate and vice versa (Minard and McAlister, 1991). The dihydroorotate dehydrogenase Ura1 catalyzes the conversion of dihydroorotic acid into orotic acid. This reaction represents the fourth step in the de novo biosynthesis of pyrimidines (Lacroute, 1968). Lsr1 is a component of the spliceosome, specifically a U2 small nuclear RNA (snRNA) (Maniatis and Reed, 1987). Tsc3 has been demonstrated to enhance the activity of serine-C-palmitoyltransferase and is implicated in the biosynthesis of sphingolipids (Gable et al., 2000). In contrast, for the adjusted p-value, Tsc3 is replaced by Mix14 (YDR031W), a mitochondrial intermembrane space protein of unknown function (Longen et al., 2009). It is notable that Mix14 contains twin Cx9C (cysteine-x9-cysteine) motifs. For RNAs with an increase in their quantity, both raw and adjusted p-value, YBR089W was identified. YBR089W is designated as a potential open reading frame that is unlikely to encode a functional protein (Fisk et al., 2006).

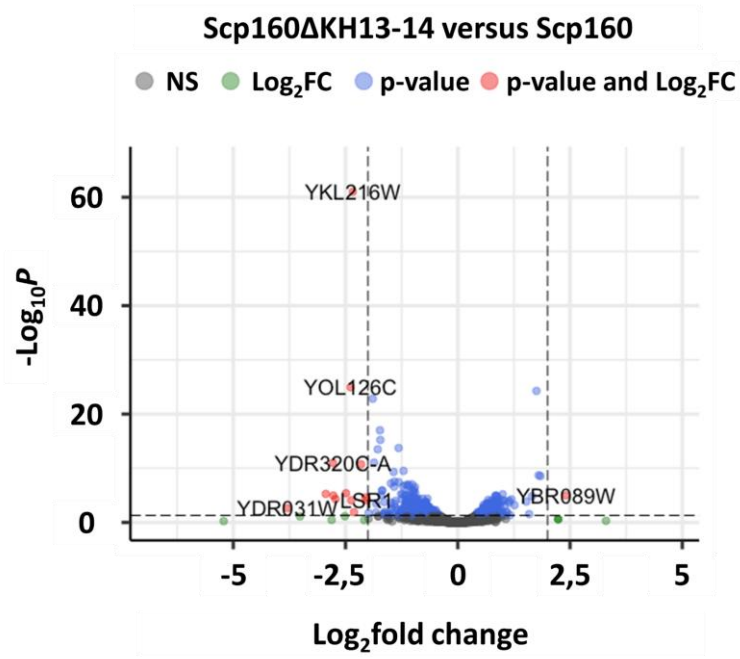


Figure 49: Volcano plot exhibiting the RNA abundance differences between Scp160ΔKH13-14 and Scp160; blue dots: hits with FC < 2 and p-value < 0.05; green dots: hits with FC > 2 and p-value > 0.05; red dots: Significant hits with FC > 2 and p-value < 0.05; negative fold change (left) displays decreased RNA abundance in Scp160ΔKH13-14.

The low number of altered RNA levels suggests that RNA expression is consistent and independent of changes in ploidy. To facilitate comparison, the results were summarized in Table 31.

Table 31: Summary of most significant enriched and decreased RNA bound by the corresponding Scp160 variants in comparison to the wildtype; Arabic number reduced associated RNA; Greek number: increased associated RNA.

condition	1	2	3	4	5	I
Scp160ΔKH1-2 vs Scp160						
Scp160ΔKH4-5 vs Scp160	Q0158 /YNCQ0006W	YCR096C				
Scp160ΔKH5 vs Scp160	YDR320C-A	YEL003W				
Scp160ΔKH13-14 vs Scp160	YDR320C-A	YOL126C	YKL216W	YNCB0019C	YDR031W	YBR089W

When considered collectively, the potential for significant transcriptome alterations resulting from KH domain deletion and the resulting increase in ploidy, appears to be either inconsequential or nonexistent. It can be concluded that the alterations in protein-protein interactions (PPIs) and RNA binding are not a consequence of changes in RNA abundance. Rather, they are more likely to be attributed to a deficiency in specific KH domains. The results of the CRAC and mass spectrometry experiments were therefor validated by this experiment.

Comparison of RNA sequencing and CRAC results

To ascertain the relationship between the RNAs bound by Scp160 or its truncation/deletion mutants and the expression levels of these RNAs in the corresponding strains, the RNA-seq data and the CRAC data were subjected to visualization in heatmaps (Figure 50 and Figure 51). In these maps, cells with similar colors reveal that the values representing expression rates are similar to each other. The band pattern in Figure 50 indicates that the RNAs recovered from CRAC samples are not symmetrical. This suggests that the RNA-binding behavior of Scp160 proteins lacking specific KH domains is divergent. The abundance of some CRAC-recovered mRNAs is subject to fluctuations, while the presence of others is entirely absent. The observed differences in clustering indicate that Scp160 and the strains lacking KH domains exhibit distinctive binding behaviors. The loss of specific KH domains, therefore, results in the loss of binding to specific RNAs. For instance, in column one/area I/Figure 50 (Scp160ΔKH5-HTP), there is a distinctive enrichment of RNAs that are not present in the other samples. The same is true for column four/area II (Scp160ΔKH4-5-HTP). This observation is only

meaningful if the overall RNA expression between the tested strains is similar to each other, as will be demonstrated in the following pages (Figure 51).

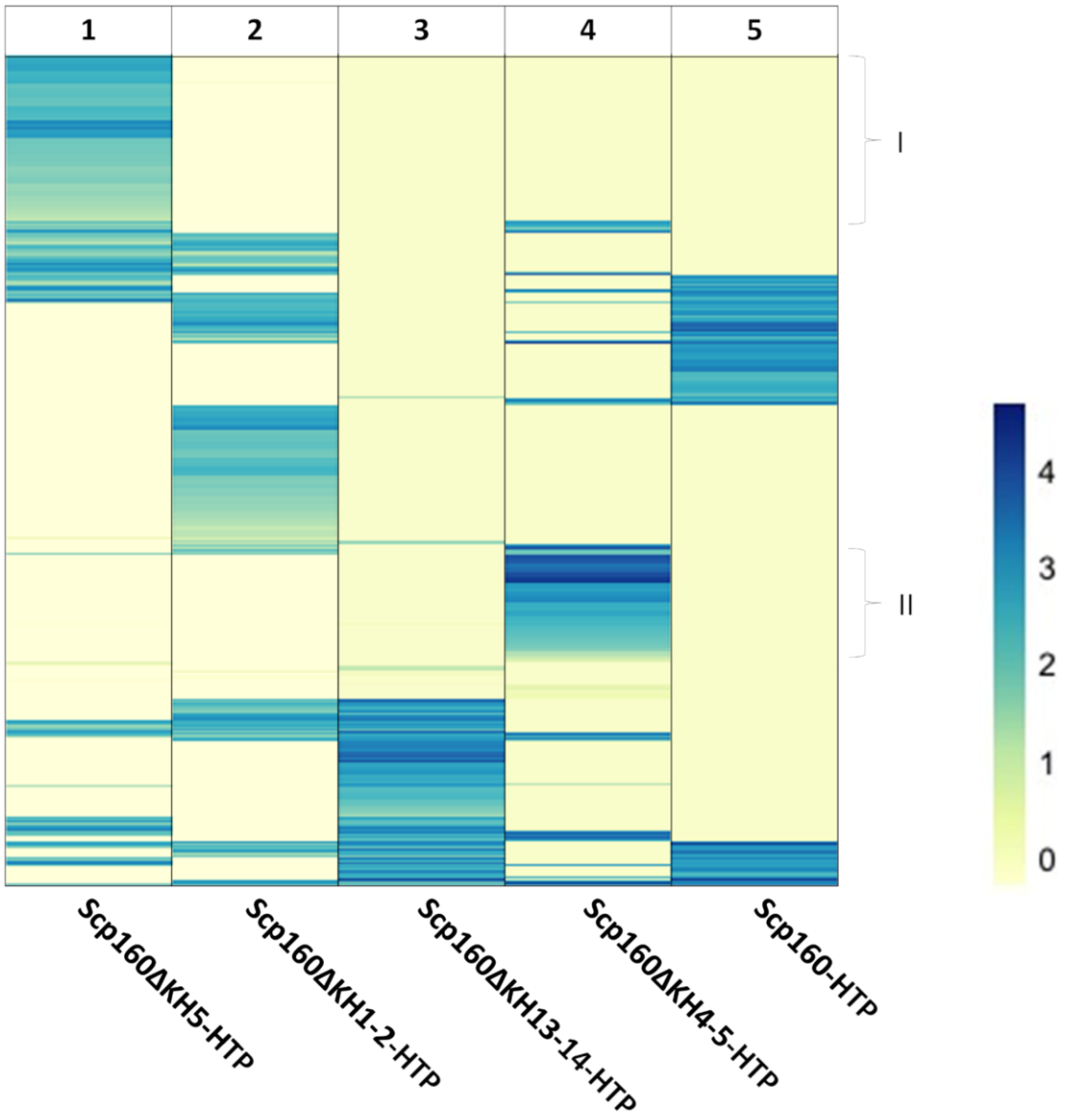


Figure 50: Heatmap of CRAC recovered mRNA; Y-axis: RNA clustered due to similarity; X-axis: RNAs recovered from binding with Scp160 protein and its KH domain lacking mutants after IP; 1: Scp160ΔKH5-HTP , 2: Scp160ΔKH1-2-HTP, 3: Scp160ΔKH13-14-HTP, 4: Scp160ΔKH4-5-HTP and 5: Scp160-HTP color-code: yellow: low presence; dark blue: high presence; I: Example for similar clustered RNA in Scp160ΔKH5-HTP; II: Example for similar clustered RNA in Scp160ΔKH4-5-HTP.

To compare the RNA binding to the Scp160 proteins and the overall abundance of these proteins in these strains, the amount of RNA identified in the RNA-seq samples was analyzed. Figure 51 illustrates the transcripts identified by CRAC within the RNA-seq data set. The band patterns are notably consistent across the various strains, suggesting a comparable expression rate of the analyzed mRNAs across all strains. Given the uniform distribution of RNA across the samples, it can be reasonably inferred that the data presented here, derived from CRAC analysis, are robust.

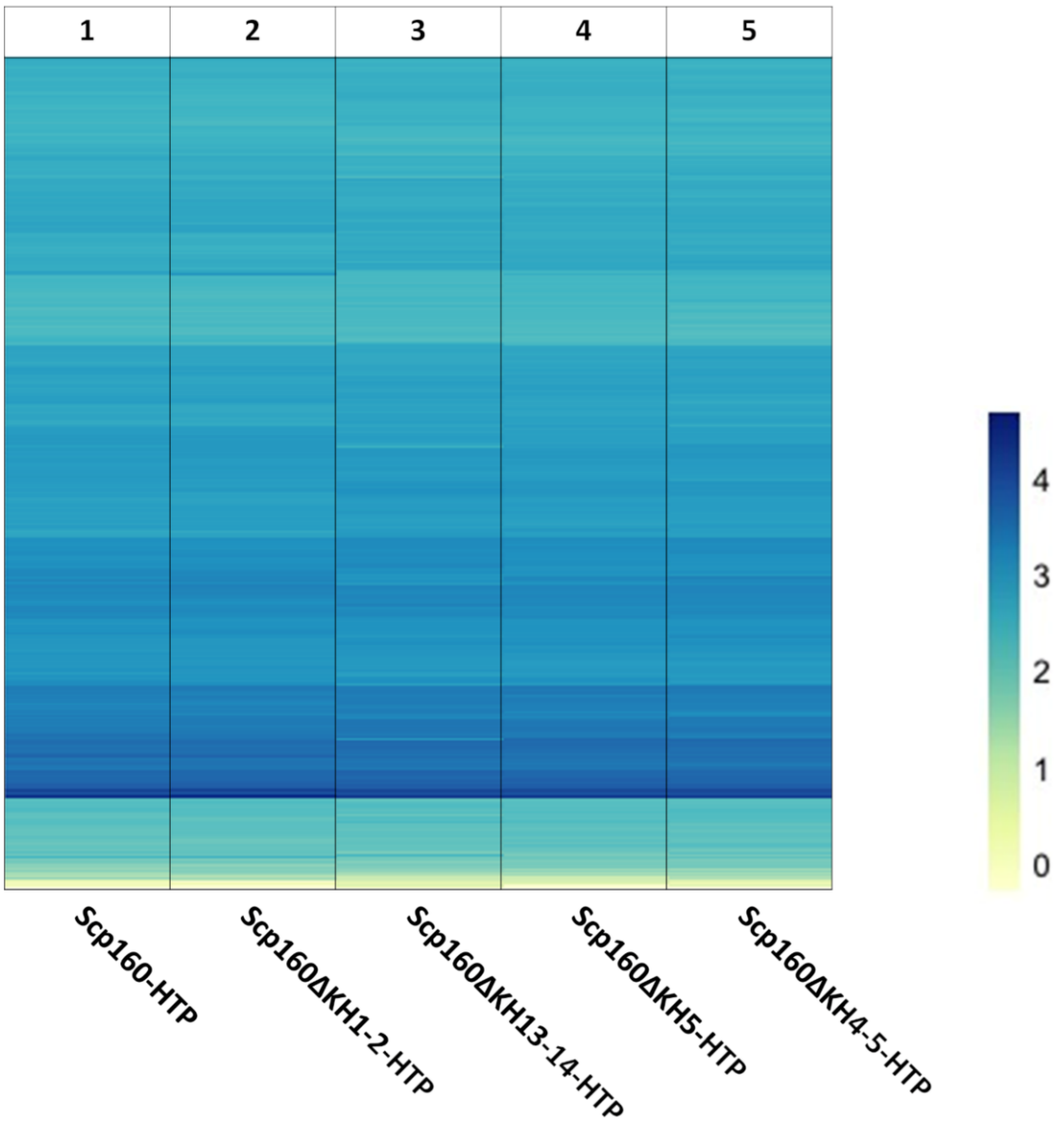


Figure 51: Heatmap of total recovered mRNA of the RNA-Seq approach; Y-axis: RNA clustered due to similarity; X-axis: recovered mRNA of the RNA-Seq approach; 1: Scp160-HTP, 2: Scp160ΔKH1-2-HTP, 3: Scp160ΔKH13-14-HTP, 4: Scp160ΔKH5-HTP and 5: Scp160ΔKH4-5-HTP; color-code: yellow: low presence; dark blue: high presence.

KH domain deletion results in a gain and loss of specific RNAs

As demonstrated in the preceding figures (Figure 50), there are discrepancies in RNA binding. However, it would be worthwhile to ascertain whether complete loss of specific RNAs occur in the event of KH domain deletion. A comparison of the bound RNAs found in the first CRAC attempt between the full-length Scp160 and the KH domain-deleted variants revealed the presence and subsequent absence of unique RNAs for either protein. The most significant findings are outlined in the following passages. The categorization of RNA is based on three overarching classifications: "molecular binding", "cellular components" and "biological processes". Classification was achieved using g:profiler (Kolberg et al., 2023). This approach is adopted to highlight the distinctions in question.

Molecular binding

The RNAs assigned to molecular binding showed no significant differences for Scp160 and Scp160ΔKH1-2. This is consistent with the previous findings. A small number of unique RNAs of structural constituent of ribosomes were identified for all tested strains except the full-length Scp160. Furthermore, unique RNAs for RNA binding were identified for the full-length Scp160 in comparison to Scp160ΔKH4-5, Scp160ΔKH5 and Scp160ΔKH13-14. Similar findings were made regarding ATP dependent activity for Scp160 variants lacking KH domain five. Notably, Scp160ΔKH5 exhibited a high number of unique RNAs associated with protein binding.

Table 32: Number of unique RNAs found for each Scp160 variant lacking KH domains in comparison to the full-length Scp160; categorization of RNA in regard to molecular functions; used program: g:profiler; only relevant categories were displayed in this table, thus not all unique RNA were assigned to a relevant category.

	Scp160ΔKH1-2	**Scp160**	**Scp160ΔKH4-5**	**Scp160**	**Scp160ΔKH5**	**Scp160**	**Scp160ΔKH13-14**	**Scp160**
total of unique RNAs	**269**	**337**	**361**	**922**	**317**	**832**	**180**	**141**
structural constituent of ribosome	21		41		40		26	
RNA binding			52	119	48	113		29
mRNA binding				58		49		17
Protein binding					157			79
ATP dependent activity				72		68		

Cellular components

In comparison to the RNAs assigned to molecular binding, which demonstrated no significant differences for Scp160 and Scp160ΔKH1-2, a considerable number of RNAs assigned to cellular components were found to be unique to Scp160ΔKH1-2 throughout all categories. It is evident that both Scp160ΔKH4-5 and Scp160ΔKH5 exhibit a conspicuous reduction of unique RNA species associated with the ER membrane and the Golgi apparatus, a feature that is exclusive to the full-length Scp160. Concurrently, both Scp160 variants contain unique RNAs associated with the cytosol and cytosolic ribosomes, while the full-length Scp160 exhibits a high number of unique RNAs connected to the cytoplasm. In general, all the Scp160 variants examined in this study exhibited distinctive RNAs that were associated with the cytoplasm. Scp160ΔKH13-14 exhibited binding of unique RNA species to the endoplasmic reticulum as well as to those located within the cytosol.

Table 33: Number of unique RNAs found for each Scp160 variant lacking KH domains in comparison to the full-length Scp160; categorization of RNA in regard to cellular components; used program: g:profiler; only relevant categories were displayed in this table, thus not all unique RNA were assigned to a relevant category.

	Scp160ΔKH1-2	**Scp160**	**Scp160ΔKH4-5**	**Scp160**	**Scp160ΔKH5**	**Scp160**	**Scp160ΔKH13-14**	**Scp160**
total of unique RNAs	**269**	**337**	**361**	**922**	**317**	**832**	**180**	**141**
ER	54	71		345	65	209	51	
ER membrane	35	48		155	46	124	33	
Golgi	29			108		96		
Golgi membrane	20			59	22	58		
cytoplasm	207	283	259	777	265	701	145	116
cytosol			89		90			
ribosome	27		48		49		30	
cytosolic ribosome	20		38		41		26	

Biological processes

As was the case with the RNAs assigned to molecular binding, the RNAs assigned to biological processes showed no significant differences for Scp160 and Scp160ΔKH1-2, except of a small number of RNAs assigned to cytoplasmic translation. The Scp160 truncation, deficient in KH domains 13 and 14, also exhibited only minimal loss of function with respect to the RNAs assigned to biological processes. However, it has been observed to bind a limited number of unique RNAs, which are responsible for translation and glycosylation. As observed in the previous two passages, the Scp160 variants lacking KH domain five (Scp160ΔKH4-5 and Scp160ΔKH5) both exhibited a reduction in a number of RNAs compared to the full-length protein. Nevertheless, both also gain a small number of unique RNAs for cytoplasmic translation and translation, as was seen for Scp160ΔKH13-14.

Table 34: Number of unique RNAs found for each Scp160 variant lacking KH domains in comparison to the full-length Scp160; categorization of RNA in regard to biological processes; used program: g:profiler; only relevant categories were displayed in this table, thus not all unique RNA were assigned to a relevant category.

	Scp160ΔKH1-2	**Scp160**	**Scp160ΔKH4-5**	**Scp160**	**Scp160ΔKH5**	**Scp160**	**Scp160ΔKH13-14**	**Scp160**
total of unique RNAs	**269**	**337**	**361**	**922**	**317**	**832**	**180**	**141**
cytoplasmic translation	24		38		43		24	
translation			55		61		34	
glycosylation					14	29	12	
rRNA processing				73		67		
GPI anchored protein biosynthesis				15		13		
ribosome assembly			15	28				
ribosome biogenesis				98		88		
intracellular transport				122		132		
protein localization		59		161				
protein transport				118		108		
Golgi vesicle transport				53		50		

The conclusion regarding alterations in RNA binding consequent to the loss of KH domains is subject to variation when viewed from different standpoints.

With regard to molecular binding, the full-length Scp160 demonstrates the highest number of unique RNAs, with a few exceptions. In terms to cellular components, full-length Scp160 again displays the highest number of unique RNAs, but the variants lacking specific KH domains also have a high number of unique RNAs for specific cellular components. Finally, when the focus shifts to biological processes, full-length Scp160 once again demonstrates a high number of unique RNAs. Taken together, the full-length Scp160 exceeds in number of unique RNAs compared to all analyzed Scp160 variants lacking KH domains, with exception to Scp160ΔKH13-14. The deletion of the last two KH domains resulted in a number of unique RNAs that was slightly higher than the full-length Scp160. Since the deleted regions are RNA binding domains, the fact that full-length Scp160 outnumbers the KH domain-deleted variants in terms of unique RNAs is expected. Why Scp160ΔKH13-14 displays a higher number of unique RNAs is yet to be analyzed.

Protein structure prediction of Scp160 lacking specific KH domains using AlphaFold2

It is plausible that the alterations in protein-protein interaction (PPI) and ribonucleic acid (RNA) interaction observed in mutants lacking KH domains are attributable to comprehensive changes in protein structure, including a potential collapse of the protein. The mean length of a KH domain in Scp160 is 73 amino acids. The executed deletion of one or more KH domains, which represent approximately 6% (one KH domain) or 29% (five KH domains) of the protein's overall size, respectively, can undoubtedly impact the protein's structure. Without access to experimentally determined protein structures, an attempt was made to gain insight into the structural changes of Scp160 caused by KH domain truncation or deletion. To this end, the protein structure prediction program AlphaFold2 was employed (Jumper et al., 2021). Of particular interest was the concept of two or more KH domains cooperatively binding to RNAs. While the existence of these kinds of RNA binding cannot be definitively proven without the presence of co-structures of the relevant RNAs, a significant disruption of the overall KH domain conformation would be evident. These findings provide a foundation for further experimentation. Accordingly, a structure of Scp160 was generated and compared with predictions of Scp160 variants lacking specific KH domains. With the exception of the unstructured N- and C-termini, the protein structures demonstrate a high degree of accuracy in their prediction, exhibiting only a few areas with a prediction quality below 70%.

The predicted structure of Scp160 is illustrated in Figure 52. The KH domains four and five, which were the focus of previous chapters, are situated in the initial lobe, which is formed by the KH domains four to six. This configuration brings KH domains two and nine into close proximity. The protein assumes a curved structure with KH domains ten to twelve, while KH domains 13 and 14 constitute a discrete unit.

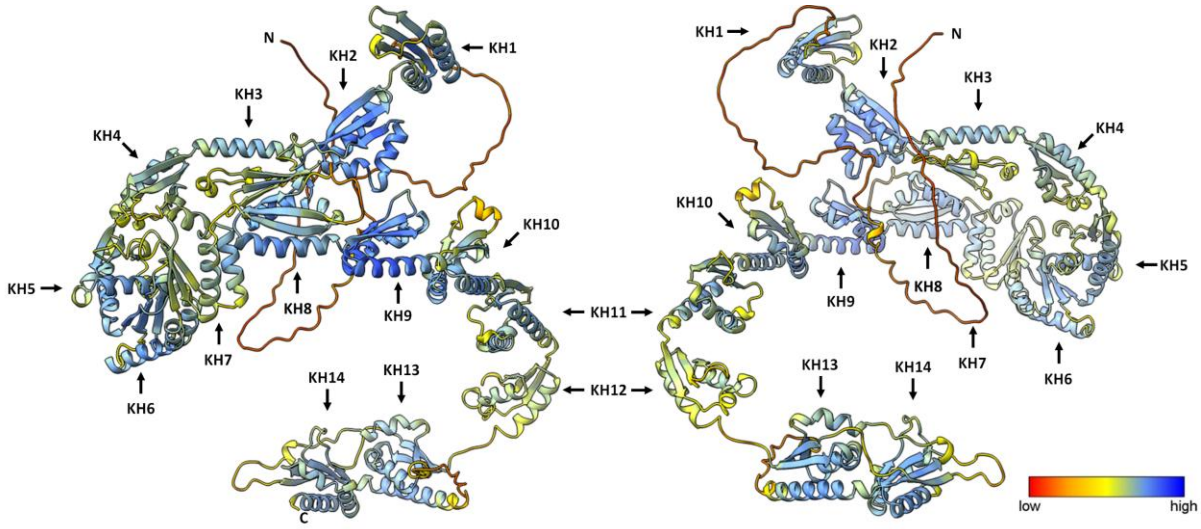


Figure 52: AlphaFold2 predicted structure of Scp160 illustrated in ChimeraX; rank 1 prediction; 180° horizontal rotation; KH domain position indicated by arrow symbol; N: N-terminus; C: C-terminus.

The removal of the first two KH domains (Scp160ΔKH1-2) does not result in significant alterations to the overall protein structure (Figure 53).

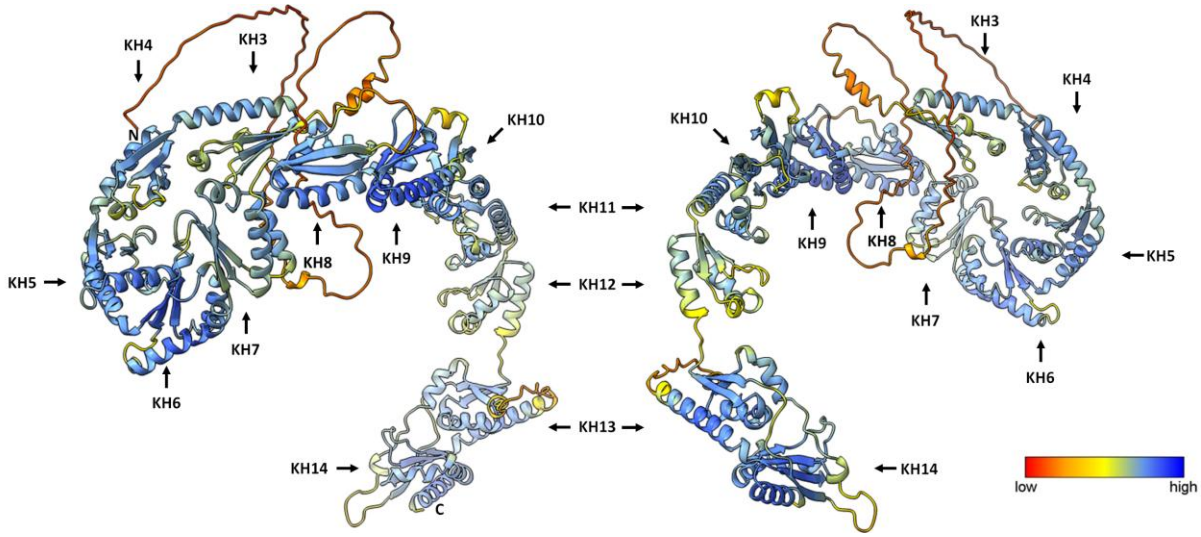


Figure 53: AlphaFold2 predicted structure of Scp160ΔKH1-2 illustrated in ChimeraX; rank 1 prediction; 180° horizontal rotation; KH domain position indicated by arrow symbol; N: N-terminus; C: C-terminus.

The removal of the KH domain five (Figure 54) has a more pronounced effect on the predicted structure than the removal of the first KH domains. The original structure's lobe remains intact, retaining KH domains four, six, and seven. However, the curved structure formed by KH domains 10-12 has undergone significant disruption. Additionally, the positioning of the KH domains relative to one another has undergone a significant alteration. In the untruncated protein, KH domain one is observed to be in close proximity to KH domains ten and eleven. KH domain two is in close proximity to KH domain ten, whereas it is originally situated in the vicinity of KH domain nine. The position of KH domain three relative to KH domain eight is no longer aligned. In conclusion, the formation of a potential new binding cluster may be facilitated by the novel protein structure.

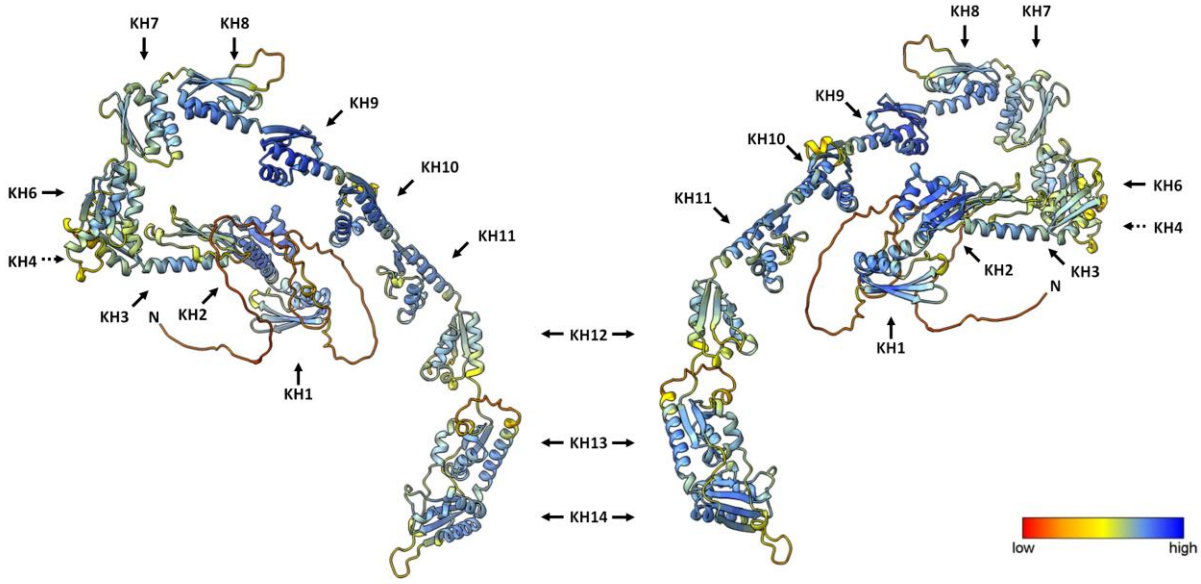


Figure 54: AlphaFold2 predicted structure of Scp160ΔKH5 illustrated in ChimeraX; rank 1; prediction; 180° horizontal rotation; KH domain position indicated by arrow symbol; N: N-terminus; C: C-terminus.

The deletion of the KH domains 4 and 5 (Figure 55) results in a protein conformation that is analogous to that observed in Scp160KHΔ5.

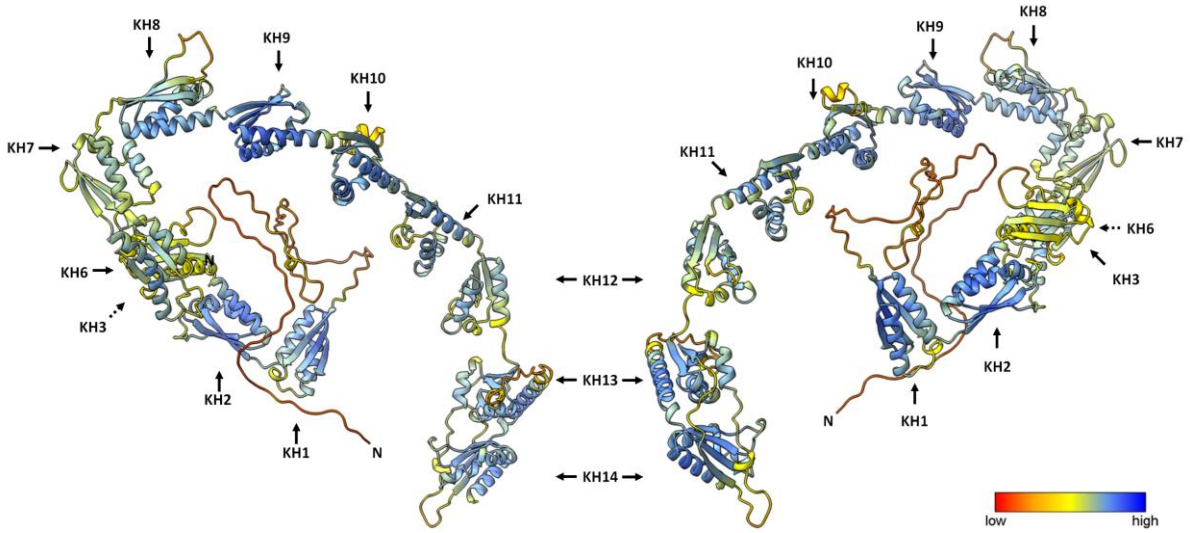


Figure 55: AlphaFold2 predicted structure of Scp160ΔKH4-5 illustrated in ChimeraX; rank 1; prediction; 180° horizontal rotation; KH domain position indicated by arrow symbol; N: N-terminus; C: C-terminus.

As anticipated based on the initial hypothesis, the removal of the final two KH domains (Scp160ΔKH13-14) did not lead to significant alterations in the overall protein structure (Figure 56).

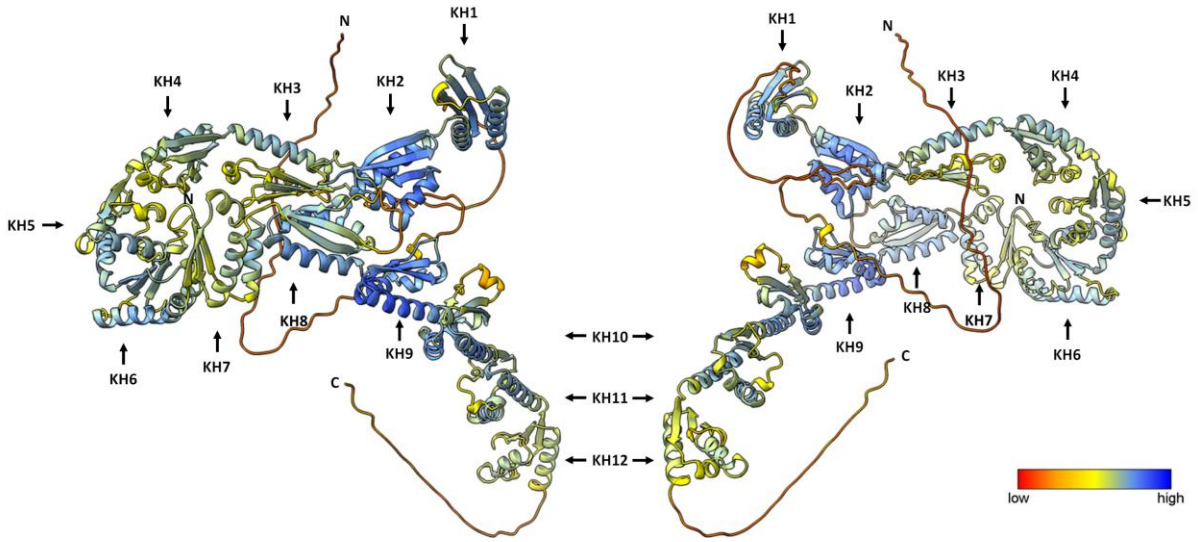


Figure 56: AlphaFold2 predicted structure of Scp160ΔKH13-14 with a C-terminal HIS-tag illustrated in ChimeraX; rank 1 prediction; 180° horizontal rotation; KH domain position indicated by arrow symbol; N: N-terminus; C: C-terminus.

The predicted AlphaFold structure of the KH domain truncated Scp160 variants indicates potential alterations in KH domain conformations. This includes the positioning of these elements in relation to one another and the potential formation of novel binding clusters. These alterations appear to have a relatively limited impact on PPI, as the overall changes in protein binding are unremarkable. It is possible that the conformational changes may influence the RNA binding behavior. It is not yet clear whether the changes in RNA binding observed in the CRAC and RNA-Seq experiments are due to the absence of particular KH domains or to altered protein structure. Further analysis and debate are required to elucidate this point.

Discussion

Fundamental thoughts

Over the past decade, the RNA-binding protein Scp160 has been the subject of extensive investigation with regard to its role in the translational process. In order to ascertain the specific areas of functional importance, the protein was frequently modified. In this regard, the C-terminal KH domains appeared to be of paramount importance. The deletion of KH domains 13 and/or 14 is a frequently observed modification in previous publications examining the functionality and role of KH domains. In order to determine the localization status and overall fitness of the protein, Baum et al. (2004) deleted the C-terminal KH domains 13-14 and 11-14. It was thus concluded that the C-terminal KH domains 13 and 14 are essential for fitness and ribosome association, as well as ploidy maintenance. In a similar attempt, KH domain 13 and the KH domain 14 were deleted or replaced in different combinations with diverged and conserved KH domains to examine the impact of these alterations on the functionality of the altered Scp160 protein (Brykailo et al., 2007). The results of these experiments indicated that diverged KH domains are crucial for the protein's functionality, even in the absence or disruption of the GXXG motif. The reintroduction of diverged KH domains may potentially restore the anticipated mRNP formation function to a certain extent. The paper presents the intriguing hypothesis that neither the overall protein size nor the total number of KH domains is a determining factor in the function of Scp160. In 2013, Hirschmann et al. analyzed the RNA-binding behavior of Scp160 in the context of KH domain loss (Hirschmann et al., 2013). The deletion of KH domains 13 and 14 resulted in a reduction in RNA binding to a number of specific mRNAs.

The collective evidence demonstrated that the C-terminal KH domains are crucial for the function of Scp160. It would be reasonable to postulate that the N-terminal KH domains may also play a role in the functionality of the protein. In accordance with the principle of parsimony, it can be posited that a protein comprising 14 KH domains will not exhibit its functionality based on a mere subset of these KH domains. The use of C-terminal truncation of Scp160 as a method of comprehending the implications of KH domain deletion was a logical choice at that time due to the limitations of efficient methods to delete interior areas of a protein. The deletion of C-terminal KH domains was also more likely to result in a significant phenotype than the N-terminal domains. This is due to the fact that the latter lack conserved KH domains, whereas the former consist, with the exception of KH domain 13, solely of conserved KH domains. The advent of CRISPR/Cas9 heralded the emerge of a novel approach to protein modification. This allowed for the deletion of any desired KH domain and the subsequent analysis of the resulting phenotype.

In the case of vigilin proteins, it is of the utmost importance to exercise caution when interpreting any kind of results, ensuring that they are contextualized correctly and presented in an accurate manner. As previously (Figure 4 b) discussed, vigilin proteins can exhibit diversity in their primary structure. For instance, the yeast vigilin Scp160 contains 14 KH domains, while the human vigilin HDLBP contains 15, with only a small overlap in the positioning of diverged/classical KH domains. Nonetheless, the number of KH domains, as well as the arrangement of conserved and diverged KH domains of vigilin proteins across especially closer related species, are to a certain extent comparable (Figure 4 b). It is important to note that these apparent discrepancies are not the only factors that require consideration. A number of phenotypes have been documented for the yeast vigilin protein Scp160. It would be erroneous to question the increase in ploidy resulting from the deletion of *SCP160*. However, it became evident that even within a single species, the observed phenotypes can vary due to inherent differences in the genetic background. This is a well-known issue that has been evaluated and determined that this issue occurs in 18.5% of experiments in which a phenotype is tested between two *Saccharomyces* strains (Galardini et al. 2019). For example, the deletion of *Saccharomyces cerevisiae SCP160* has been observed to result in an increase in ploidy, but it should be noted that this increase is not always constant. In the genetic background of W303, only shifts from haploid to pseudo-diploid phenotypes could be discerned. In contrast, the BY4741 *scp160Δ* strain displays a pseudo-tetraploid phenotype. This finding indicates that the genetic background is an additional variable that must be taken into account during the experimental planning phase and the subsequent evaluation of experimental results.

The deletion of specific KH domains in this study resulted in pseudo-diploid to pseudo-tetraploid phenotypes and ploidies that fall between these two extremes. The observed fluctuation in this, arguably the most prominent phenotype of the vigilin protein, underscores the necessity for rigorous scrutiny of any other observed phenotype and conclusions derived from experiments, both past and present. This is because any subsequent observation could be a secondary effect of an increase in chromosomes.

The selected yeast strains may influence the intertwining and significance of Scp160 in its function. The W303 background is a widely utilized strain with an uncertain origin. With 74 to 85% consistency to S288C, a notable distinction between W303 and S288C is the absence of a functional *SSD1* gene (Novačić et al., 2021). The lack of Ssd1 apparently does not affect most cellular processes and does not impair the viability of the strain. However, it does appear to bind similar RNA targets as Scp160 which are regulating the fungal cell walls (Hall and Wallace, 2022). Therefor it maybe affects the phenotype of Scp160 deletion mutants due to redundant RNA targets. Ssd1 plays a role in maintaining cell wall integrity (Kaeberlein and Guarente 2002, Reinke 2004), auto-diploidization (Tung et al. 2021),

and translation repression (Hu et al. 2018). It is established that it binds to the 5' untranslated regions of mRNAs encoding cell wall proteins (Bayne et al. 2022). It is noteworthy that it has a functional interplay with the DEAD-box helicase Dhh1 (Moriya and Isono 1999, Jansen et al. 2009). As a direct potential interaction partner (Carroll et al., 2011) of Scp160, it may serve as a possible link to Ssd1. The influence of alterations to single members of a functional chain may be enhanced or diminished by the presence of overlapping functions and redundant processes amongst certain genes.

The final decision regarding the genetic background to be utilized for the study of KH domain deletions was made in favor of the BY4741 background, since the observed changes in ploidy were found to be more stable in comparison to those observed in the W303 background.

The deletion of N-terminal KH domains is opening a new avenue for understanding the functional importance of Scp160

There are only a few documented instances of N-terminal editing of Scp160 (Li et al., 2004). The presence of a visible phenotype indicated the harboring potential of N-terminal truncation. Given that the N-terminus consists exclusively of diverged KH domains, which do not possess a classical GXXG motif, the potential for the manifestation of minuscule phenotypes was a plausible consideration. Nevertheless, the documented functionality of diverged KH domains (Brykailo et al. 2007) and the fact that it was an uncharted territory at the time led to the conclusion that the topic merited further investigation. It was expected that the deletion of multiple KH domains increases the probability of observing substantial phenotypes. However, as the number of KH domains deleted increases, so does the risk of phenotypes emerging not from the loss of specific (sets of) KH domains, but from resulting structural changes of the entire protein. Accordingly, my objective was to maintain the deletions as minimally as feasible. However, very small editing approaches like the introduction of point mutations instead of the deletion of entire KH domains, were not pursued due to the absence of clear functional motifs, such as GXXG in several KH domains of Scp160.

Ploidy alterations in *Saccharomyces cerevisiae* differ on the deletion of specific N-terminal KH domains

Given that the most conspicuous phenotype of the yeast vigilin deletion is a change in the DNA content, as indicated by its designation "Scp160" (*S. cerevisiae* protein controlling the ploidy) an examination of this aspect in the newly strains lacking specific KH domains was a logical next step. In

1995, the ploidy impairment resulting from the Scp160 deletion was confirmed (Wintersberger et al., 1995). The precise cause of this phenotype remains unclear. Nevertheless, it represents an invaluable starting point for the analysis of new Scp160 mutant strains. The initial investigation in this regard was conducted using basic microscopy techniques to examine the DNA content of *scp160Δ* cells after staining with DAPI. This experiment demonstrated that daughter cells lacking DNA content were produced (Wintersberger et al., 1995). Therefore, the mother cells that are unable to distribute surplus chromosomes must retain additional DNA. The hypothesis that the absence of functional chromosome segregation results in the formation of empty daughter cells and diploid mother cells cannot be supported by the ploidy profiles of the Scp160 variants lacking KH domains produced in this study. Firstly, if this event were to occur in every budding cell, it would result in an increase in ploidy for every generation. If a threshold of maximal ploidy existed, all strains would ultimately reach this upper limit of chromosomes. The existence of pseudo-diploid and pseudo-tetraploid strains, however, suggests that this is not the case. Furthermore, the ploidy profiles of several strains cannot be assigned to a single ploidy type, such as diploid or tetraploid. Instead, they appear to fall somewhere between these two extremes (Figure 8 and Supplementary Figure 1).

A number of potential theories have been put forth to explain the phenomenon of ploidy increase, resulting from the deletion of Scp160. An early, though not particularly advanced, theory posits the interaction of Bfr1 with the protein Bbp1, which plays an essential role in the spindle pole body in yeast. The deletion of Scp160 and the consequent change in the binding partner Bfr1 could result in an alteration of the spindle pole apparatus as a secondary effect (Lang, 2000). This theory was further developed with the addition of placing Bbp1 into a broader context. Its interaction with the protein Mps2, a membrane protein essential for the insertion of the spindle pole body formation into the nuclear envelope, provides a broader foundation for considering the impact of Scp160 in the context of ploidy control. Similarly, the interaction of Scp160 with Eap1 can be considered. The deletion of both genes is a synthetically lethal event (Mendelsohn, 2003), indicating that the two proteins play a vital role when present together.

In a separate experiment, EAP1 was identified in a synthetic lethal screen with ndc1-1 (Chial et al., 1999). ndc1-1 is an allele of NDC1, which is a subunit of the transmembrane ring of the nuclear pore complex (NPC). It plays a role in NPC biogenesis and spindle pole body duplication (Thomas and Botstein, 1986). In RT-PCR experiments employing *NDC1*-specific primers, it was demonstrated that *NDC1* mRNA is markedly enriched in Scp160p-associated complexes (Mendelsohn, 2003). It is possible that Scp160p is necessary to ensure an appropriate level of *NDC1* RNA. These findings are worth mentioning given that *NDC1* was observed to interact with the full-length Scp160 and the

Scp160ΔKH1-2 variant, but not with the other KH domain deletion strains tested in the CRAC and RNA-Seq experiments. Scp160ΔKH1-2, as the full-length Scp160, maintains a haploid phenotype.

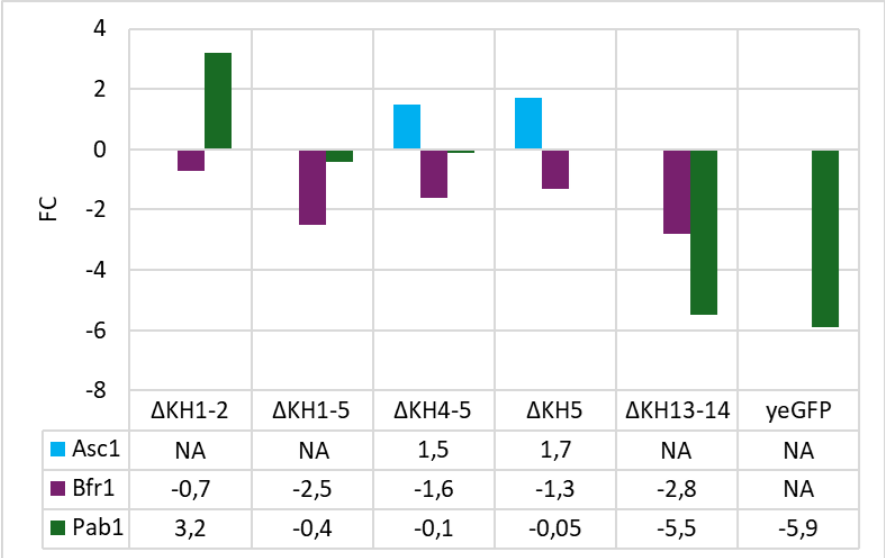
The unequal distribution of chromosomes is not solely dependent on Scp160, but rather on the disruption of a protein complex that governs the proper segregation of chromosomes. This complex appears to comprise Smy2, Eap1, Scp160, and Asc1, and was thus designated "SESA" (Sezen et al., 2009). Subsequently, Dhh1 was put forth as an additional member of the SESA complex (Ergüden, 2019). It is postulated that the SESA complex exerts control over chromosome segregation at the level of spindle pole body (SPB) maintenance. The model proposes that the proteins Pom34, Pom152, and Ndc1 form a complex during the biogenesis of the NPC. Furthermore, Ndc1 is involved in the SPB duplication process along with Mps2, Bbp1, and Nbp1. The incapacitation of *POM34*, *POM152*, or *NDC1* results in the rescue of SPB duplication defects, indicating an inhibitory effect of these genes. In the event of SPB duplication defects, the SESA complex downregulates POM34, thereby facilitating SPB duplication (Sezen et al., 2009). It was postulated that Scp160 is the binding element of *POM34*. This hypothesis was not substantiated in the present study. As neither the Scp160 protein nor its KH-deleted variants were observed to bind *POM34* RNA, it is plausible that the binding of *POM34* RNA is facilitated by Scp160, but that it is not a sufficient condition for binding. The binding of *NDC1* RNA, however, could be assigned to Scp160. *NDC1* RNA is binding to Scp160 and Scp160ΔKH1-2, which both do not show a change in ploidy. In contrast, for the other tested KH deleted Scp160 variants (Scp160ΔKH5 and Scp160ΔKH4-5) lacking *NDC1* binding, a change in ploidy could be observed (Supplementary Figure 1). The alteration of *NDC1* dosage has been demonstrated to result in aneuploidy and polyploidy, thereby confirming the potential of this RNA to affect ploidy in Scp160-impaired strains.

The localization of Scp160 is dependent on the presence of specific KH domains

It has been established that Scp160 is located at the ER. However, the factors that must be altered to disrupt this association remain to be elucidated. The localization of Scp160 from the ER to the cytosol may be attributed to alterations in its interactions with other proteins (Baum et al., 2004), upon osmotic stress (Baum et al., 2004), or alterations in its RNA-binding behavior (Frey et al., 2001). RNase treatment has been demonstrated to result in a separation between ribosomes and Scp160 and therefor the localization of Scp160 at the ER (Frey et al., 2001). Nevertheless, the question of the mechanism by which Scp160 is bound to the ER remains unresolved. Given its status as an RNA-binding protein, it is plausible that Scp160 forms a bond with ribosomes via a simultaneous mRNA/rRNA

connection. An additional possibility is that ribosomal proteins such as Asc1 may be involved, as proposed in Baum et al., 2004. Asc1 is the ortholog of human RACK1, which has been demonstrated to interact with the human vigilin protein Hdlbp (Brugier et al., 2022). Alternatively, a combination of RNA and protein binding may be responsible. The cysteine at position 1067 of Scp160 is capable of forming a crosslink with Asc1 and eEF1A, indicating that Scp160 is situated at the ribosome at a distance of approximately 8 Å from Asc1/eEF1A (Baum et al., 2004).

With regard to ER localization, the Scp160 interactors Pab1 and Bfr1 (Lang, 2000) are worthwhile consideration. Pab1 is a polyadenylate-binding protein that interacts with Scp160 in messenger ribonucleoprotein particles (mRNPs). Together with Bfr1, it colocalizes with Scp160 at the ER (Lang, 2001). It is thus not implausible that Scp160 requires these proteins to bind polysomes at the ER. A comparison of the fold change of these proteins (Asc1, Bfr1, and Pab1) in the KH domain deleted Scp160 variants reveals a reduction in the amount of Bfr1 bound in those variants that are impaired in their ability to localize to the ER (Figure 57). This reduction is not observed in the case of the KH domain deleted Scp160 variants, which still localizes to the ER. It is not possible to make a statement about Asc1, as it was not identified for Scp160ΔKH1-2, Scp160ΔKH1-5, Scp160ΔKH13-14, or the negative control. It is not possible to determine whether this is due to a lack of binding or flaws in the experiment. The fold change for Pab1 was elevated by a factor of three for Scp160ΔKH1-2 and reduced by a factor of five for Scp160ΔKH13-14.



	ΔKH1-2	ΔKH1-5	ΔKH4-5	ΔKH5	ΔKH13-14	yeGFP
Asc1	NA	NA	1,5	1,7	NA	NA
Bfr1	-0,7	-2,5	-1,6	-1,3	-2,8	NA
Pab1	3,2	-0,4	-0,1	-0,05	-5,5	-5,9

Figure 57: Corresponing column chart for the fold change of known protein interactors of Scp160 variants lacking KH domains compared to the full-length Scp160; blue: Asc1; purple: Bfr1; green: Pab1; FC: Fold change; NA: not annotated.

In addition to the three previously identified interaction partners of Scp160, it was observed to interact with 47 small ribosomal subunit proteins (RPS) and 80 large ribosomal subunit proteins (RPL). This phenomenon is likely attributable to co-precipitation rather than direct interaction. Notably, Scp160ΔKH1-5, the variant with the greatest number of deleted KH domains, did not exhibit the highest number of proteins with altered fold change; rather, this was observed in Scp160ΔKH4-5. This is detectible not only for ribosomal proteins but the entirety of the identified proteins. This may be interpreted as an indicator of a higher binding capacity of Scp160ΔKH4-5 to ribosomes, which would be limited to cytosolic ribosomes, given the compromised ER localization of Scp160ΔKH4-5. A comparison of the RNA clusters found in Scp160ΔKH4-5 and the other tested Scp160 proteins reveals that this variant has the fewest RNA clusters connected to the secretory pathway (Table 27). Scp160ΔKH1-5 and Scp160ΔKH5 are also delocalized from the ER and exhibit elevated binding to RPL/RPS proteins (44 ribosomal proteins), although to a lesser extent than Scp160ΔKH4-5 (57 ribosomal proteins). Furthermore, the increase in binding of Asc1 in Scp160ΔKH5 is also comparatively increased in Scp160ΔKH4-5 (Figure 57). The number of mRNA clusters for Scp160ΔKH5 is also decreased by approximately 50% in the initial CRAC experiment and 27% in the subsequent attempt (Figure 43). Unfortunately, no data are available for Scp160ΔKH1-5, as the sample number was limited and the Scp160 variants shown in Figure 41-44 were given priority. The findings suggest that the binding of Scp160 to ribosomes is not hindered by the deletion of KH domains. However, the impact of KH domain deletion on RNA binding is significant, especially including the deletion of KH domain number five. The absence of specific KH domains can therefore result in the relocation of Scp160 via the loss of specific RNAs.

The Scp160 dependent emulsifying properties of yeast can be achieved by simple KH domain deletion

Emulsification is a procedure that is frequently employed in the food and cosmetics industries. The increased accessibility of naturally derived emulsifiers would reduce the dependence on chemically synthesized ones. The emulsifying properties of yeast cells have therefor been a subject of scientific inquiry for a considerable period of time. The potential of yeast as a food supplement has been the subject of investigation since the 1980s (Cameron et al., 1988). Nevertheless, the knowledge of emulsifying properties associated with the expression of Scp160 is relatively recent (Nerome, 2022). Consequently, the available data regarding this particular phenotype is limited. The initial insights into the origin of yeast's emulsifying properties identified mannoproteins as a key driver (Cameron et al., 1988). However, the concentration of mannoproteins is unlikely to be the sole essential factor. *mcd4Δ*

cells exhibit a markedly enhanced emulsifying property while displaying a reduced mannose-to-glucose ratio in comparison to wild-type cells (Nerome, 2020). The deletion of *MCD4* results in a significant reduction in cell wall mannan, which is defective in the synthesis of GPI anchors. These are essential for fixing mannoproteins to β-glucans. This evidence indicates that not only mannoproteins, but also other components of the cell wall, may be involved in the alteration of emulsifying properties. It was demonstrated that the washing of yeast cells with Milli-Q grade water could diminish the emulsifying effect, whereas the washing water utilized in the experiment exhibited a notable enhancement in its emulsifying properties (Moreira et al., 2016). These results indicate that the emulsifying properties are, to some extent, dependent on secretion (Moreira et al., 2016). Given that Scp160 is likely to be involved in the translation of secreted proteins, the observed phenotype may be caused by a change in cell wall properties or a change in the composition of secreted proteins. The cluster analysis of RNA changes (Figures 40-44) demonstrated the presence of ER-Golgi network RNAs in the full-length Scp160, Scp160ΔKH1-2, Scp160ΔKH5, and Scp160ΔKH13-14. However, no tested Scp160 variant displayed binding of mRNAs encoding cell wall proteins in the first CRAC attempt. The sole exception is the Scp160KH13-14 truncation (Figure 44). In the second experiment, RNAs encoding proteins related to cell wall function were identified in the full-length Scp160, Scp160ΔKH1-2, and Scp160ΔKH13-14. The RNA-seq data indicated a relatively uniform expression rate of mRNAs (Figure 51). Regarding the protein-protein interaction of cell wall components and Scp160, there was no discernible increase in the number of cell wall proteins involved in the identified interactions. The alteration in cell wall proteins interaction would have been an explanation for the arise of an emulsifying effect after deletion of KH domains one to five, three to five and four to five. These findings consequently diminish the potential of cell wall components as a catalyst of changes in emulsifying properties for KH domain-deleted variants of Scp160.

The deletion of KH domains exerts a mild effect on the protein-protein interactions of Scp160

The removal of KH domains may result in alterations to the proteins structure, potentially leading to a disruption in PPI. It is reasonable to hypothesize that the observed phenotypes in previous experiments are caused by secondary effects and not by deleting different KH domains. Alternatively, there is a possibility that specific KH domains bind with a particular subset of interaction partners. Regarding alterations in protein-protein interactions, all Scp160 variants with a reduced number of N-terminal KH domains exhibited a notable number of proteins increased in binding, rather than a reduction in protein-protein interactions, in comparison to the full-length protein (Table 23). The

negative control of yeGFP and the C-terminal truncation Scp160ΔKH13-14 exhibit comparable alterations in protein-protein interactions to each other (Table 23). Following the normalization of the datasets to the wildtype, a comparison was made between the Scp160ΔKH13-14 PPI and the negative control (SCPp-yeGFP-SCPt). This analysis revealed a high degree of overlap between significantly changed proteins. Therefore, it was determined that these identified proteins are considered background and must be excluded (Table 23; grey coloring).

Scp160 interacts with ribosomal proteins

One readily apparent feature of the identified interacting proteins is the considerable number of ribosomal proteins that interact with Scp160 and its KH domain deleted variants. Of the total 1,568 protein-protein interactions identified, 115 were with ribosomal proteins. Additionally, 61 were mitochondrial ribosomal proteins and 29 were proteins involved in ribosomal biogenesis, ribosome assembly, and ribosomal quality control. The list of proteins displaying changes in binding due to Kh domain deletion includes a considerable number of ribosomal proteins. Scp160 has been demonstrated to interact with Asc1, a well-established ribosomal protein interactor of Scp160 (Baum et al., 2004). The elevated quantity of ribosomal proteins can be attributed to the co-immunoprecipitation of the complete ribosome rather than to their direct interaction with Scp160. Nonetheless, the presence of ribosomal proteins or proteins involved in ribosome regulation interacting with Scp160 is not a novel observation when compared to other data sets. A comparison of my and published data sets revealed an overlap of 57 out of 167 interactions (Delaveau et al., 2016), 21 out of 21 (total verified PPI) (Rössler et al., 2019), and 9 out of 14 (Gavin et al., 2006) ribosomal proteins and proteins of ribosome regulation.

With regard to the observed alterations in protein abundance, the ribosomal proteins merit particular attention. Rpl3 exhibited a substantial loss in Scp160ΔKH1-2, Scp160ΔKH1-5 and Scp160ΔKH5. In all three strains Rpl3 ranked second in terms of loss in association to Scp160 (Table 23). Furthermore, Rps20 and Rps5 were identified as being amongst the top five proteins for Scp160ΔKH1-2 to bind less in comparison to the full-length Scp160 (Table 23). In Scp160ΔKH5, the most significant protein to be reduced is the ribosomal protein Rps4A/B (Table 23). While these findings suggest a potential direct interaction between these ribosomal proteins and Scp160, further investigation is necessary to substantiate this hypothesis. However, the overall pattern of increased and decreased fold changes for ribosomal proteins is rather inconclusive, exhibiting numerous inconsistent patterns. For instance, no observed decrease in the fold change of Scp160ΔKH4-5 for Rpl3 is evident. The marked reduction in the binding affinity of Scp160ΔKH1-5 and Scp160ΔKH5 for this protein would suggest that a similar

effect may be observed in the case of a KH domain deletion between these two numbers of deleted KH domains. Another unexpected pattern observed in the data is the similarity of Scp160ΔKH1-2, a Scp160 variant with unremarkable phenotypes, to the variants Scp160ΔKH1-5, Scp160ΔKH4-5, and Scp160ΔKH5, which all display more pronounced alterations in their overall fitness (Table 23).

Some proteins harbor the potential to be direct interaction partners of Scp160

A number of proteins have been identified as interacting with Scp160 on multiple occasions. The list includes Asc1, Rps2, Arb1, Pat1, Sis1 and Zuo1 (Supplementary Table 3). Bfr1 and Asc1, which have been previously mentioned and have been shown to interact with Scp160 on seven (Bfr1) and four (Asc1) external experiments, respectively. Similarly, the ribosomal protein Rps2 is also likely to interact with Scp160 to a similar extent as Asc1, as it was found in four independent experiments. In addition to the aforementioned interactors, the proteins Arb1, Zuo1, Pat1, and Sis1 were identified as interacting with Scp160 in three separate external experiments each. Arb1 is an ATPase belonging to the ATP-binding cassette (ABC) family and plays a role in the biogenesis of the 40S and 60S ribosomes (Delaveau et al., 2016; Dong et al., 2005). Zuo1, a ribosome-associated co-chaperone, is also involved in ribosome biogenesis, which aligns with the high number of proteins associated with this process that are bound by Scp160 (Delaveau et al., 2016; Gavin et al., 2002; Gavin et al., 2006). The protein known as Sis1 has been categorized as a type II co-chaperone, which interacts with the Hsp70 protein Ssa1. The interaction of Ssa1 has been observed in a number of strains in which the KH domain has been deleted (Figure 34-38). It has also been observed in the negative control (Figure 33), which could be interpreted as a possible reaction to misfolding of the proteins due to the changes in their sequence (Ali et al., 2023; Delaveau et al., 2016; Feder et al., 2021). Lastly, the deadenylation-dependent mRNA-decapping factor Pat1, which connects to Scp160 in its function of ensuring reliable chromosome transmission and P-body formation (Delaveau et al., 2016; Mitchell et al., 2013; Weidner et al., 2014). The presence of Pat1 and Sis1 could not be detected in either the Scp160 or the KH domain-deleted Scp160 versions. The remaining proteins of interest (POI) in this paragraph demonstrate a binding pattern that is rather inconclusive when compared across the Scp160 variants. Asc1 is lost for Scp160ΔKH1-2, Scp160ΔKH1-5 and Scp160ΔKH13-14, which have only one thing in common: the loss of either the N- or C-terminus. This finding suggests the potential requirement for both termini for Asc1 binding. For Scp160ΔKH5 and Scp160ΔKH4-5, a reduction in Asc1 binding was observed. An even more random finding was made for Rps2, where Scp160ΔKH5 and Scp160ΔKH13-14 demonstrated an inability to bind the protein, Scp160ΔKH4-5 exhibited a reduced binding capacity, and Scp160ΔKH1-2 and Scp160ΔKH1-5 displayed a binding profile to Rps2 analogous to that of the full-length Scp160. The

underlying rationale behind this finding, if indeed it exists, does not emerge from the limited sample size employed. One of the POIs was only found in Scp160ΔKH5 and Scp160ΔKH4-5, namely Zuo1. The reasons for the absence of this protein in the wildtype remain to be elucidated.

These findings suggest that the aforementioned protein-protein interactions seem to be largely independent of the KH domain deleted. Furthermore, it is more likely that the changes in protein structure could inhibit the protein binding capacity of Scp160.

Protein-protein interactions changes accompanied by KH domain deletion

The number of significant losses in PPI for Scp160 variants lacking specific KH domains is quite limited. As evidence of the validity of these changes, the binding behavior of a well-known interactor can be illuminating. Bfr1, which has previously been identified as a key interaction partner of Scp160, (Lang et al., 2001) exhibited significantly reduced binding by Scp160 variants lacking KH domains 1 to 5, 4 to 5, 5, as well as 13 to 14 (Table 23). Only the Scp160ΔKH1-2 variant demonstrated no impairment in its ability to bind Bfr1. This pattern is consistent with the results of my other experiments, showing no phenotypic changes upon deletion of the first two KH domains. It also suggests that the ability of the other KH-deleted Scp160 variants to bind Bfr1 is specifically altered.

The proteins exhibiting an increase in binding outside the ribosomal proteins also demonstrated inconsistency in their function. All variants of Scp160 lacking KH domains show enrichment for a specific protein over the full-length Scp160 protein, termed Ura2. This enzyme plays a pivotal role in the initial two enzymatic steps of pyrimidine biosynthesis. Another protein, which was more abundant in the N-terminal KH domain deletion variants, but not in those with a KH domain 13 to 14 truncation, is Mis1. This C1-tetrahydrofolate synthase, which is located in the mitochondria, is responsible for the interconversion between different oxidation states of the folic acid derivative tetrahydrofolate (Shannon and Rabinowitz, 1988). Together with the RNA helicase Dbp2, which is involved in mRNA decay and rRNA processing, and the inosine monophosphate dehydrogenase Imd3, these four proteins show a higher abundance in the N-terminal Scp160 variants compared to the full-length Scp160. Whether or not these proteins have a similar property that leads to this increased binding has yet to be analyzed. If the higher abundance of these four proteins is valid, it is imperative that this findings are validated through additional experimentation.

Another interactor is Tdh3, the budding yeast GAPDH. This protein exhibits the most pronounced partial loss (negative fold-change -1.87) in the Scp160 variant lacking KH domains one and two, but no negative fold-change is observed for any other tested sample. It remains to be seen whether further experiments will reveal why this change in PPI is so specific to Scp160ΔKH1-2, while Scp160ΔKH1-5,

which also lacks the first two KH domains, is not influenced in the same manner with respect to this interactor.

Finally, the known Scp160 interactor Pab1 has to be discussed. This protein is nearly absent in the negative control of yeGFP and in the KH domain 13 to 14 truncated variant of Scp160. Given its status as a binding partner of Scp160, the low abundance of Pab1 can be attributed to the absence of Scp160 as a bait protein. The loss of KH domains 13 and 14 in the C-terminus of Scp160 appears to have a specific impact in its ability to bind Pab1. This contrasts with the other N-terminal KH domain deletion variants, which did not result in a loss of Pab1.

The comparison of the unique binding of proteins between Scp160 and its KH domain-lacking variants demonstrates an increase in unique binding for the Scp160 variants lacking N-terminal KH domains. The highest number of unique PPIs was found for Scp160ΔKH4-5 (Table 24), with proteins attributed to a number of different functions. This is surprising because the Scp160 variant lacking the first five KH domains was also tested and only reached ~40% of the number of Scp160ΔKH4-5. The underlying reasons for the propensity of Scp160ΔKH4-5 to exhibit a distinct array of unique PPIs remain to be elucidated. One hypothesis is that the impairment of Scp160ΔKH4-5 folding is more severe in comparison to Scp160ΔKH1-5. This may result in a more unspecific protein binding, which is a plausible consequence of this impairment. It is interesting to note that, in the case of Scp160ΔKH13-14, the full-length Scp160 exhibits a higher number of unique protein-protein interactions (PPIs). When the suggested protein structures of AlphaFold are taken into consideration (Figure 56), it appears that the loss of the last two KH domains is not accompanied by an overall structural change. Consequently, it can be hypothesized that the lost KH domains are responsible for binding the lost unique proteins found only in full-length Scp160. This finding could be used as a basis for further investigations of C-terminal binding of proteins.

The overall alteration in the interactome is rather minor (Figure 31). The interaction pattern for the triplicates is largely consistent, with the exception of Scp160ΔKH4-5 (Figure 31 c) and Scp160ΔKH5 (Figure 31 b), which exhibit notable divergence. This suggests that the phenotypes observed in this thesis are not necessarily the result of only changing protein-protein interactions. Conversely, the probability of changes in the RNA interactome (Figure 50) having a more substantial impact on the appearance of phenotypes is considerably higher.

AI based protein structure prediction tools aid in understanding functional alterations that occur in structurally impaired proteins

Understanding the structure of Scp160 might help to explain the observed changes in protein and RNA binding. The loss of KH domains and their implications with regard to conformational changes may serve to indicate the connection between structural changes and resulting phenotypes. The software AlphaFold3, represents a promising tool for rapid protein structure predictions. However, it necessitates a degree of skepticism when interpreting the structures, since all AI based predictions are not infallible. Nevertheless, the overall predicted structure is likely to reflect the actual protein, with possible minor deviations. It has to be noted, that the absence of RNA, the potential consequences of KH domain deletion remain inconclusive. The comprehension of Scp160 and its function in RNA binding may be augmented through the implementation of artificial intelligence (AI). This necessitates the ability to simulate RNA binding to the protein which is not yet possible.

The cooperative binding of RNA to multiple KH domains has been demonstrated in studies involving two and three cooperative KH domains (Adinolfi et al., 1999; Git and Standart, 2002; Paziewska et al., 2004). It is therefore reasonable to suggest that cooperative RNA binding may occur in an RNA-binding protein comprising 14 KH domains.

The structural analysis provided insight into protein folding in the event of the removal of distinct KH domains. The analysis revealed positions of particular significance, which appear to be more important to the overall structure of the protein than others. For example, the lobes surrounding KH domains four to five, five to six, and nine to ten (Figure 52). The location of KH domain five at the vertex underscores its pivotal role in the protein's overall structure. This may provide an explanation for why the deletion of KH domain number five exhibits a more severe phenotype in comparison to the deletion of KH domains one and two, which, in addition to the higher number of deleted domains, also includes a classical KH domain (Fig. 4). This suggests that the deletion of KH domains four and six alone should also have a greater impact on the function of Scp160 than the deletion of one and two, due to their positioning.

An intriguing investigation would be the reconstruction of a vigilin protein comprising 14 conserved KH domains. The introduction of the optimal KH domain 8 on position of KH domain 5 resulted in a profound alteration to the protein structure. However, the introduction of a conserved GXXG motif into assumed position of the diverged GXXG motif in number five resulted in no alteration to the protein structure at all (Supplementary Figure 10). While the GXXG motif of KH domains may be the most conserved part, it is still only a single segment of the whole functional domain. In order to

"rescue" an aberrant KH domain and transform it into a conserved KH domain, a number of further modifications are necessary, in addition to the restoration of the GXXG motif. The correct positioning of hydrophobic amino acids, the appropriate residues for side chains, and the maintenance of the β1α1α2β2β3α3 structure are equally crucial (Grishin, 2001). The concept of exchanging KH domains has already been demonstrated (Brykailo et al., 2007). The exchange of KH domains may be employed to analyze the impact of classical KH domains in the position of diverged KH domains, and vice versa. This approach may elucidate the potential of exchangeability of KH domains in general. It is possible that the exchange of different KH domains may not restore 100% of the protein's functionality. However, the improvement in function of the rescued Scp160 is nevertheless noteworthy (Brykailo et al., 2007). It remains to be seen whether a similar outcome would be achieved with N-terminal KH domains, where a structural interruption would be more significant.

KH domains represent a platform for specific or general RNA binding

As an RNA-binding protein comprising 14 KH domains, the binding sites of RNAs can be situated at any point on the protein. The precise manner in which Scp160 binds specific RNAs with particular KH domains remains to be fully elucidated. The deletion of KH domains 13 and 14 resulted in a reduction of binding to specific RNAs, including *AGA1*, *PRY3*, and *CCW14* (Hirschmann et al., 2014). Other researchers have employed analogous methodologies, either through the deletion of particular KH domains (Baum et al., 2004; Brykailo et al., 2007; Hirschmann et al., 2014; Zhou et al., 2008) or the inverse approach of expressing specific KH domains to examine their RNA binding characteristics (Zinnall et al., 2022). An illustrative example is the analysis of three to five consecutive KH domains of the vigilin protein HDLBP, which displayed differences in their capability of binding different RNA oligomers (Zinnall et al., 2022). However, the process of mapping KH domains and their RNA targets is still in its infancy. This work was intended to be a step ahead towards this aim.

The initial and most evident outcome of the comparison between the full-length Scp160 protein and the investigated KH domain-deleted variants is the notable reduction in RNA binding capacity observed in the Scp160 proteins lacking KH domains five, four and five and 13-14 (Table 30). The deletion of KH domains four and five results in a loss of approximately 80% and in the case of domains five approximately 70% (Figure 39). The removal of the C-terminal KH domains (KH13-14) resulted in a reduction of mRNA binding by approximately 30%. However, the deletion of KH domains one and two led to a notable preservation of mRNA binding, with a remaining level of approximately 90%. This leads to the conclusion that KH domains four and five are essential for mRNA binding. It is yet to be

determined whether this is due to their RNA binding capability or due to structural changes. The more pronounced reduction in mRNA binding capacity associated with the middle section of the protein, in comparison to the first two or last two KH domains, is reflected in the more significant structural alterations observed in the AlphaFold-predicted structures (Figures 53-56). The removal of KH domains further away from the protein's C-terminus could potentially disrupt either cooperative KH domain binding clusters or the overall protein structure, leading to loss of RNA binding. It is remarkable that a HDLBP fragment, comprising only KH domains five to nine, exhibits a three- to five-fold higher affinity for RNAs with elevated multivalent potential and for the tested mRNA and 18S RNA, which were bound with similar affinities compared to the full-length HDLBP (Zinnall et al., 2022). Given the unlikelihood of the structure of this fragment being comparable to that of the full-length protein (Supplementary Figure 9), it is necessary to consider the possibility of RNA binding without structural dependence. The deletion of KH domains exhibited other irregularities (Figure 39). In comparison to the full-length protein, the Scp160ΔKH1-2 variant displays a markedly elevated capacity for rRNA and snoRNA binding. Scp160ΔKH4-5 exhibits a comparatively higher level of ncRNA, snoRNA, snRNA, and tRNA (Figure 39). The absence of tRNA in the full-length protein (Figure 39), which Scp160 was previously proposed to be part of the recycling machinery (Hirschmann et al., 2014), is particularly difficult to reconcile. As it is with all experiments, it is preferable to conduct the experiment more than once to ensure the validity of the obtained data.

The analysis of the various RNA clusters identified in the Scp160/RNA interactome reveals a similar pattern of specific RNA binding loss in Scp160ΔKH4-5 and Scp160ΔKH5 (Figure 42-43). The number of RNA clusters found in Scp160ΔKH4-5 is significantly reduced in comparison to the full-length Scp160, with a loss of approximately 81% in the first and second CRAC attempt. In the case of Scp160ΔKH5, the number of RNA clusters lost in comparison to the full-length protein is not as severe as for Scp160ΔKH4-5. Exhibiting a loss of approximately 35% in the first and 51% in the second CRAC approach if compared to the wildtype. As stated above (see Detection of Scp160p-bound RNAs via Cross-Linking Analysis of cDNAs (CRAC)), the most prominent RNA clusters lost for Scp160ΔKH4-5 and Scp160ΔKH5 encode for proteins of the ER, Golgi network and the membrane, which is consistent with a proposed function of the protein in the synthesis of secreted proteins and those of the endomembrane systems.

The alterations in RNA binding seen in the different mutants may have originated from a number of potential sources. It is possible that structural changes and the loss of RNA binding domains may have occurred. However, another factor that required consideration was the impact of ploidy changes that occur concurrently with KH domain deletion. A higher ploidy could potentially distort the results, therefore the overall RNA expression was assessed and compared between the strains expressing full-

length Scp160, Scp160ΔKH1-2, Scp160ΔKH4-5, Scp160ΔKH5, and Scp160ΔKH13-14 (Figure 51). The expression levels of the RNAs identified in the CRAC analysis were subsequently analyzed and compared. The resulting heatmap demonstrates that the relative expression of RNAs is comparable between Scp160 and KH domain-deleted/truncated Scp160 variants. It is therefore highly unlikely that the observed phenotypes, including ER delocalization, emulsion property variations, alterations in protein-protein interactions, and changes in RNA binding, can be attributed to ploidy changes and the subsequent increase in chromosomes. The corresponding heatmap, which illustrates the bound RNAs by the aforementioned Scp160 proteins, suggests a change in RNA targets for Scp160 proteins lacking specific KH domains (Figure 50). The absence of RNAs in the marked areas (I) and (II) of Figure 50, in which RNAs are bound by Scp160ΔKH5 (I) or Scp160ΔKH4-5 (II), but not by any other Scp160 variant, suggests not a loss but a shift in RNA targets with alteration of the accessibility of KH domains.

It is notable that the number for the ER-Golgi-network clusters in the C-terminal KH domain deletion Scp160ΔKH13-14 is nearly equivalent to that observed in the full-length protein (Table 29). This result suggests that Scp160ΔKH13-14 may still be involved in interactions with ER-localized ribosomes, which would contradict the assumption that it is not localized at the ER (Baum et al., 2004). A notable discovery is the relatively limited number of RNA clusters encoding proteins associated with the cell wall and plasma membrane, as illustrated in Table 30. The majority of RNAs encoding membrane proteins are associated with the endoplasmic reticulum or Golgi apparatus. This may indicate a primary role in maintaining the ER-Golgi network.

The RNA clusters identified for the full-length protein align with the broader conceptualization of RNAs that are hypothesized to interact with Scp160. These clusters were found to contain RNAs from a number of different cellular locations, including the ER-Golgi network (Wintersberger et al., 1995; Zinnall et al., 2022), ribosomes (Frey et al., 2001), translation initiation (Sezen et al., 2009), and membrane and cell wall proteins (Hirschmann et al., 2014). The clusters of the ER-Golgi network displayed the highest number for most of the tested strains (Figure 40-Figure 44). The sole exception is the Scp160 variant lacking KH domains four and five. In the initial CRAC attempt, no RNA clusters of ER-Golgi network could be identified. In the second experiment, which was of a lesser quality than the first in terms of recovered RNAs, RNA clusters of the ER-Golgi network were identified. The absence of the most crucial RNA clusters further underscores the severe functional deficit of the protein resulting from the loss of KH domains four to five.

Of the analyzed Scp160 proteins lacking N-terminal or C-terminal KH domains only a limited number of RNAs that exhibited a significant decrease in binding to Scp160 were found (Figure 46-Figure 49). Four of these RNAs are coding for proteins located at the mitochondria and are exclusively present in

the N-terminal KH deletions. If the 13 found RNAs are of a substantial importance for the overall purpose of Scp160 and its function as a RBP remains unclear. Notably, the RNA *DAD4* represents a prominent exception. It is evident that, in its capacity as a subunit of the Dam1 complex, the function of Dad4 in chromosome segregation can be regarded as being closely related to the function of Scp160 (Westermann et al., 2005). Furthermore, it is the sole RNA observed in the dataset of two KH deletion strains (Figure 47 and Figure 49). Specifically, these RNAs were identified in Scp160ΔKH5 and Scp160ΔKH13-14.

To truly understand the impact of KH domain deletion on RNA binding, the data collected must be subjected to further in-depth bioinformatic analysis. This could include the study of binding partners, where the significance level is lowered, and a wider range of reduced RNA interactions is considered.

In comparison to other attempts to identify Scp160 mRNA targets, there is always the possibility of evaluating the informative value of the findings produced. In recent decades, not many studies were conducted to ascertain the binding specifications of Scp160 to RNA. The only comprehensive lists available are those provided by Hogan et al. (2008) and Delaveau et al. (2016). Of the total approximately 6,000 identified RNAs and potential open reading frames in the yeast *S. cerevisiae* (Goffeau et al., 1996), 1242 were present in the first CRAC, which was deemed as of better quality compared to the second CRAC attempt. This represents an overlap of approximately 21%, meaning that Scp160 is capable of binding a fifth of all known RNAs in *Saccharomyces cerevisiae*. In the Delaveau publication, 46% (278 out of 369) of the identified RNAs were also present in the CRAC dataset (Delaveau et al. 2016). 22% of the RNAs found in the first CRAC analysis match the third set of data, where it was assumed that Scp160 was a RBP binding nearly all mRNA in yeast (Hogan et al. 2008).

It was postulated that mammalian vigilin Hdlbp exhibit a preference for binding to CU-rich motifs in RNA. Hdlbp was identified as a primary regulator of apoB translation and VLDL secretion in hepatocytes (Mobin et al., 2016). The motif enriched in the bound RNAs showed this characteristic CU-rich sequence, which was also present in a number of proatherogenic secreted proteins in addition to *Apob* (Mobin et al., 2016). Subsequent experiments indicated a distinction between all ER-localized mRNAs and those localized to the cytosol. The sequence composition of these two groups of mRNAs differs in the frequency of CU-rich binding sites, which are bound by HDLBP (Zinnall et al., 2022). Importantly, an analysis of enriched motifs within the Scp160-bound RNAs in the data sets of this thesis did not result in a distinct motif (Supplementary Table 2). Furthermore, two motif analysis tools, MEME (Machanick and Bailey, 2011) and EXSTREME (Grant and Baily, 2021), were employed to ascertain whether the deletion of KH domains resulted in a loss or gain of a prominent binding motif.

This endeavor did not yield a compelling motif, either. Additionally, a preference of Scp160 for CU-rich domains could not be substantiated. This does not necessarily disprove the CU-rich binding motif, as the data was not separated into cytosolic and ER-localized RNAs. However, it is feasible to assume that the CU-rich is rather a vigilin binding motif of higher eukaryotes.

As discussed, a comparison of protein-protein interactions in the absence of KH domains reveals a gain in unique protein binding for the KH domain deletion variants. However, a contrasting pattern is observed in the loss of RNA. While Scp160ΔKH1-2 and Scp160ΔKH13-14 appear to be relatively balanced in terms of unique RNA binding partners, Scp160ΔKH4-5 and Scp160ΔKH5 exhibit a substantial loss of RNA compared to the wildtype (Table 24). A comparison of the loss and gain of unique RNAs for Scp160ΔKH4-5 and Scp160ΔKH5 reveals that they are comparable if categorization of RNAs is considered: For instance, the loss of RNAs attributed to cellular components (Table 33) can be assigned to ER and Golgi components, while the gain in unique RNAs is attributed to ribosomal and cytosolic RNAs. The comparable pattern of RNA loss, as well as the number of RNAs, is attributable to the substantial overlap between lost unique RNAS. The proportion of overlap of lost RNAs between these two groups is significant (approximately 70%). When compared to the other Scp160 variants lacking KH domains, all RNAs found to be absent in Scp160ΔKH1-2 and Scp160ΔKH13-14 were also found to be absent in Scp160ΔKH4-5 and Scp160ΔKH5. This finding indicates that the RNA binding properties of these specific RNAs are not exclusively attributed to particular KH domains, which was already hypothesized in previous research (Valverde et al. 2008). It is noteworthy that Scp160ΔKH5 displays a number of absent unique RNAs, although it remains uncertain whether this is attributable to the absence of the corresponding binding domain. This issue is further compounded by the observation that Scp160ΔKH4-5 should also be expected to lack these particular RNAs, which it does not. One potential explanation for this phenomenon is compensatory structural changes of KH domains. This finding suggests the possibility of alternative binding behavior, indicating a deficiency in our understanding of the impact of KH domains on each other. The binding of RNA through multiple KH domains has been demonstrated, and this could provide a fundamental understanding of the binding of RNA from the Scp160 protein and vigilin proteins in general (Paziewska et al. 2004, Schneider et al. 2019). It is not yet known whether the absence of KH domains can be compensated for by other present KH domains. Consequently, it is imperative that the unique absent RNAs identified in Scp160ΔKH4-5 are validated through additional experiments before any definitive conclusions can be drawn.

Taking together the relationship between the loss of RNA and the absence of KH domains in Scp160 variants are partially unrelated. The results demonstrate that the loss of many RNA is not directly dependent on the deleted KH domains, since the absence of specific RNAs is observed in different KH

domain deletions. The collected data, however, is inconclusive regarding whether the loss of unique RNAs is attributable to the absence of specific KH domains or other concomitant changes in the proteins' function.

Summary

The investigation of Scp160 and its various mutants provides insights into its cellular functions, particularly its role in gene expression regulation and cell division.

Microscopic analysis of Scp160 mutants lacking the KH domains 1 and 2 (Scp160ΔKH1-2) reveals that these mutants remain localized at the ER as the wildtype protein, indicating that these domains are not essential for ER targeting. Furthermore, strains expressing Scp160ΔKH1-2 do not show phenotypic alterations in ploidy or emulsifying effects. Mass spectrometry shows that Scp160ΔKH1-2 binds more PAB1 compared to the wildtype, but significantly less TDH3, indicating altered protein interactions for these two proteins. RNA sequencing reveals an increased binding affinity for rRNA and snoRNA, which suggests a shift in RNA interaction profiles compared to the wildtype. The Scp160 mutants lacking the KH domain 5 (Scp160ΔKH5) have however a reduced tendency to localize at the ER, even with only one deleted KH domain. Loss of this particular KH domain results in increased diploid and tetraploid phenotypes, implicating that it is important for ploidy regulation. In addition, cells expressing this variant of Scp160 show an emulsifying effect in polar and nonpolar mixtures. Mass spectrometry indicates a significant reduction in the binding of the known interaction partner Bfr1, while overall protein binding remains similar to the wildtype. RNA sequencing displays an overall decrease in RNA binding in comparison to the wildtype Scp160. If in addition the adjacent KH4 domain is deleted, the phenotypes resemble those of Scp160ΔKH5, regarding loss of ER localization, increase in diploid or tetraploid phenotypes, and an elevated emulsifying effect. Bfr1 binding is significantly reduced, similar to Scp160ΔKH5. The overall protein binding pattern is comparable to the wildtype. RNA sequencing of Scp160-associated RNAs indicates higher levels of non-coding RNA, snoRNA, snRNA, and tRNA, with mRNA clusters resembling those also found for the Scp160ΔKH5 mutant, but with an even lower amount of mRNA clusters compared to the wildtype. The deletion of all five N-terminal KH domains (Scp160ΔKH1-5) accordingly display the severe phenotypes as seen in Scp160ΔKH5 and Scp160ΔKH4-5 expressing cells. The RNA binding behavior has yet to be analyzed to confirm the consistency of RNA binding loss in comparison to the other strains lacking KH domain five. The strain to compare deletion of N-terminal and middle KH domains was Scp160ΔKH13-14. This strain was analyzed likewise to the already mentioned strains. Microscopic examination of Scp160 mutants

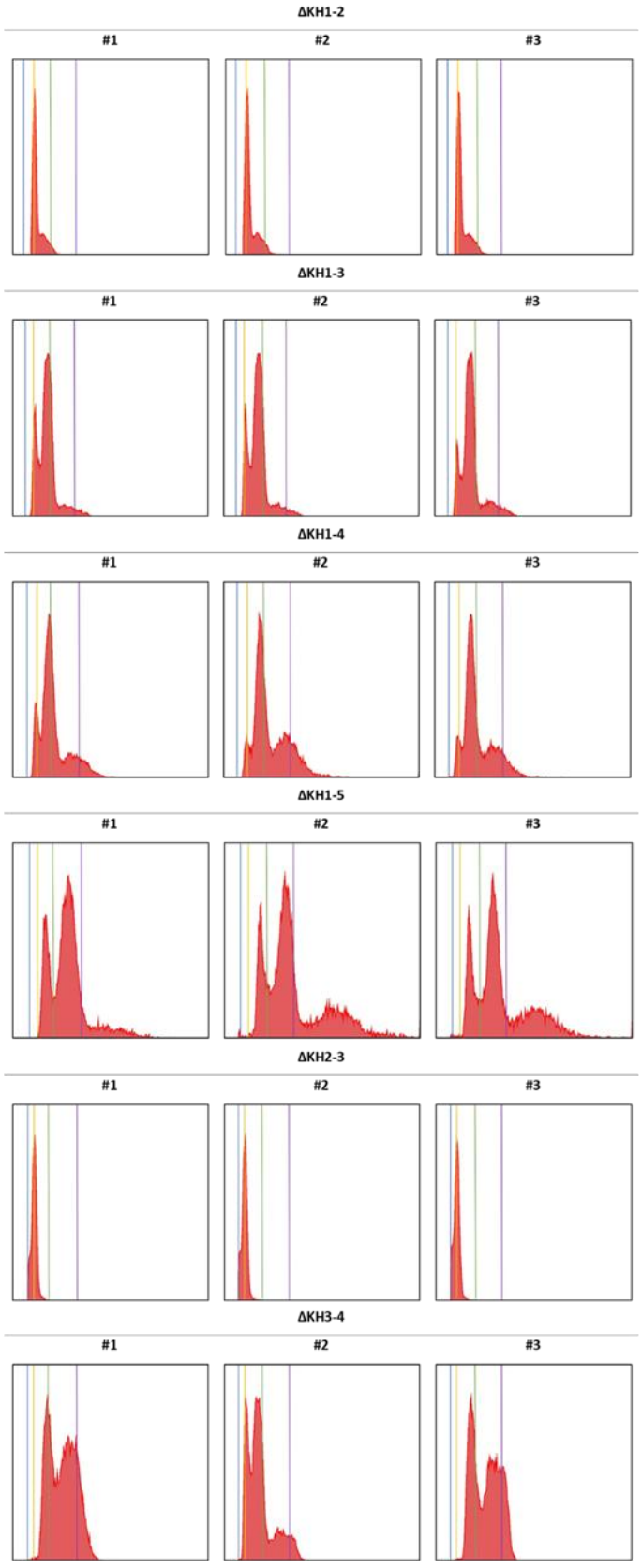
lacking KH domains 13 and 14 show occasional ER localization. These strains exhibit an increased diploid phenotype but do not show an emulsifying effect in polar and nonpolar mixtures. Mass spectrometry indicates a significant reduction in Bfr1 and Pab1 binding, while overall protein binding remains similar to the wildtype. This mutant was the only one with impairment of binding the known Scp160 interaction partner Pab1. RNA sequencing shows that the number of RNA clusters is similar to the wildtype, although the specific clusters differ, such as RNAs coding for cell wall proteins.

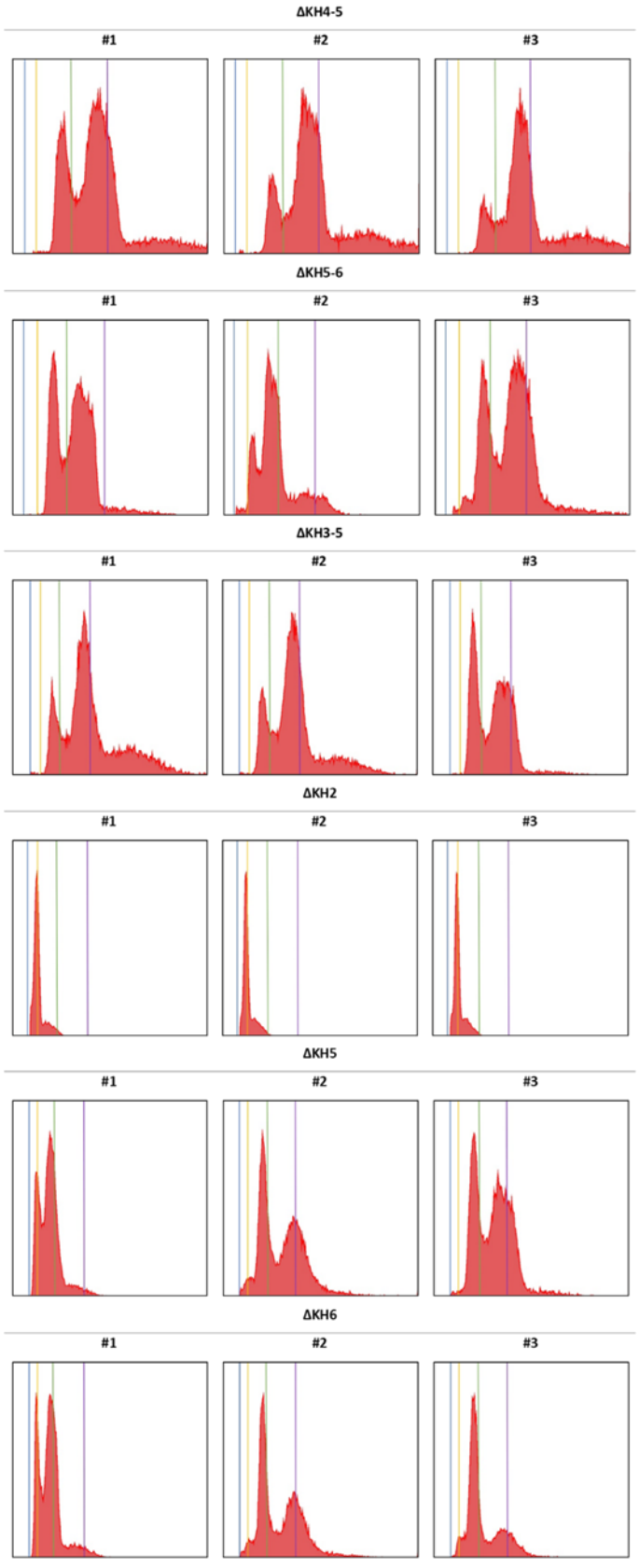
These findings collectively highlight the multifaceted roles of Scp160 and its KH domains in regulating ploidy, protein interactions, and RNA metabolism, as well as in maintaining ER localization.

Outlook

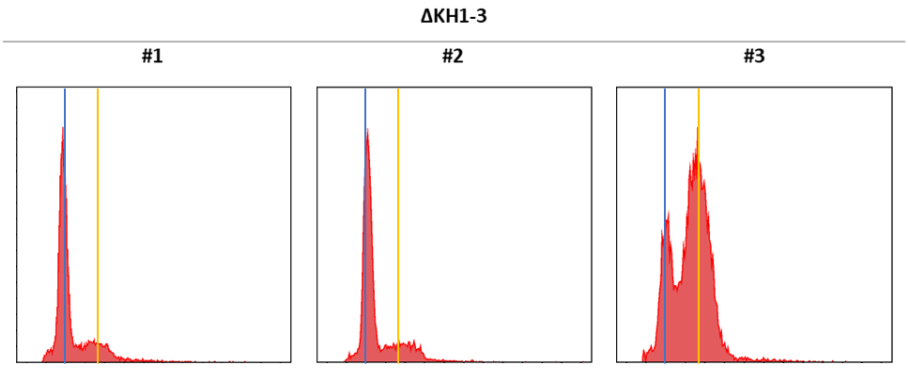
Recent studies have highlighted the importance of vigilin proteins in binding specific subsets of mRNAs, suggesting a role in selective mRNA transport and localized translation. Recently, the question of the purpose of the multiple KH domains of vigilin proteins has been raised. Deletion of KH domains provided a first starting point for investigating the relationship between different KH domains. The absence of specific KH domains has been associated with different phenotypes such as ER delocalization, ploidy change and alteration of protein and RNA interaction partners. However, there are still several uncertainties regarding the impact of deletion of KH domains, the resulting loss of RNAs and the observed phenotypes. The link between the observed phenotypes and the absence of specific RNAs could not be finally proven in this thesis. This relationship needs to be established in further experiments using a more targeted approach. Moreover, it is imperative to extend the number of experiments analyzing the effect of deleting specific KH domains and the resulting changes in RNA binding. This should encompass an increase in replicates and the incorporation of additional Scp160 variants lacking different KH domains. Due to lack of resources, only the most promising of the Scp160 variants lacking specific KH domains could be studied. The inclusion of more combinations of deleted KH domains may improve the understanding of the potential of KH domains to bind RNA, as singular units or as cooperative clusters. Although the loss of KH5 showed a clear impact on RNA binding, it still needs to be shown if other single KH domains, or combinations thereof are responsible for binding specific (or specific sets of) RNAs. In addition, the importance of the presence of the classical GXXG motif for RNA binding also needs to be analyzed, since deletion of diverged KH domains lacking this motif (for example domain four and five) still results in observable phenotypes. To understand the evolution of KH domain-dependent RNA binding, knowledge of KH domain specific RNA targets could in the future be compared with other higher eukaryotic vigilin proteins. This is of specific importance since vigilin proteins in higher eukaryotes, such as humans (HDLBP), have been linked to cancer progression, progression of viral infection, and metabolic regulation. Comparative studies of Scp160 and its mammalian homologues may reveal conserved mechanisms and provide a model for understanding how dysregulation or modulation of vigilin proteins contributes to disease. In addition, this line of research may also identify potential therapeutic targets for modulating vigilin protein activity in human disease. Emerging technologies may accelerate the understanding of RNA and protein binding. The continued development of AI based protein structure prediction tools, such as AlphaFold 3, has the potential to facilitate the study and understanding of protein structure, protein-protein interactions and, in particular, the interaction of RNA-binding proteins with RNAs to an even greater extent than is currently possible. These structural insights would be invaluable in elucidating the molecular basis of RNA recognition and binding specificity, potentially leading to the development of small molecules or peptides that can modulate Scp160 function.

Supplement





Supplementary Figure 1: Ploidy profiles of three independent Scp160 strains lacking specific KH domains; blue: haploid peak, yellow: diploid peak, green: tetraploid peak, violet: octoploid peak; strain BY4741.



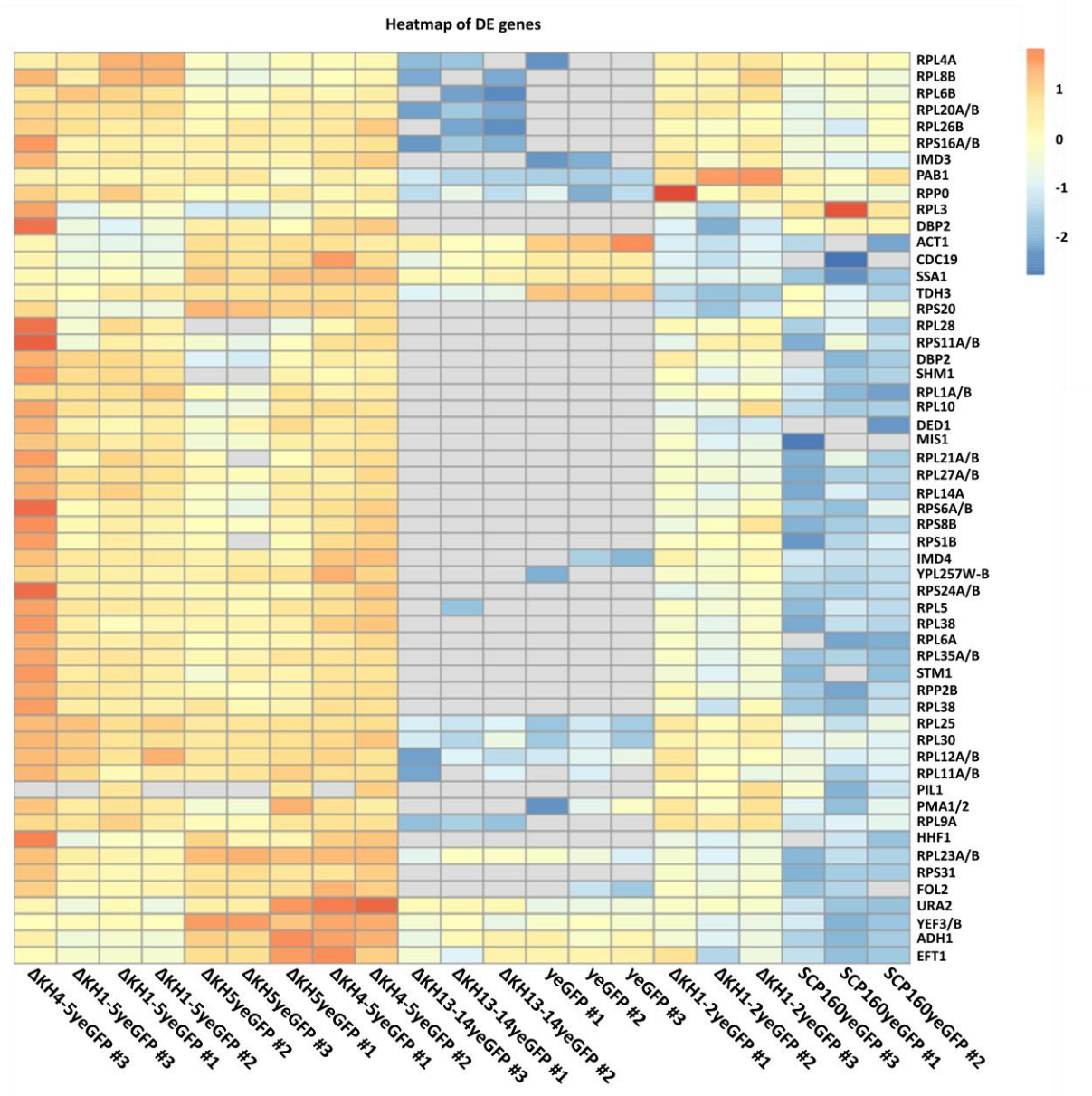
Supplementary Figure 2: Ploidy profile of one Scp160 strain lacking KH domains one to three; ploidy shift in the third measuring; blue: haploid peak, yellow: diploid peak, green: tetraploid peak, violet: octoploid peak; strain W303.

Supplementary Table 1: complete list of measured cell area of 100 cells per strain.

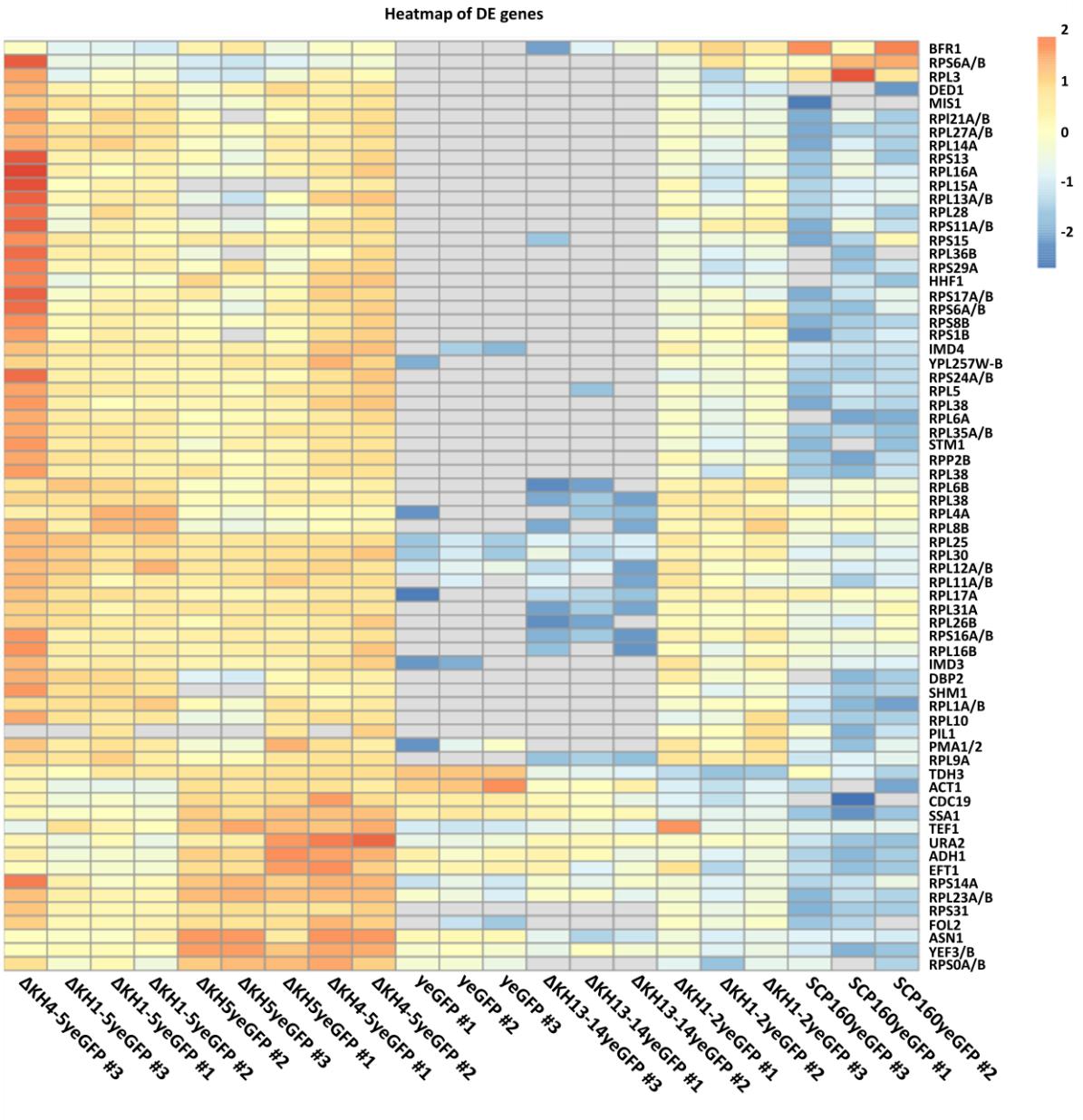
Scp160															
haploid	diploid	*scp160Δ*	ΔKH1-2	ΔKH1-3	ΔKH1-4	ΔKH1-5	ΔKH2-3	ΔKH3-4	ΔKH4-5	ΔKH5-6	ΔKH3-5	ΔKH2	ΔKH5	ΔKH6	ΔKH13-14
17,6	25,4	35,0	25,0	29,1	26,5	45,6	28,1	40,9	46,1	31,9	44,6	19,6	43,3	32,7	30,7
18,5	25,4	52,2	17,3	27,0	22,6	54,4	23,8	48,0	43,2	24,3	41,8	22,1	41,8	45,7	23,6
10,7	23,7	34,6	15,2	28,2	33,2	40,1	19,0	31,5	41,1	30,4	53,0	25,3	38,7	42,2	23,7
12,8	27,9	29,7	17,5	36,1	23,8	49,2	16,9	30,9	43,0	26,1	60,5	20,9	27,3	30,2	30,2
14,9	26,9	64,0	19,7	26,7	23,9	40,3	19,4	34,3	44,8	35,7	37,6	16,8	25,4	43,1	37,5
15,2	25,8	58,3	18,0	27,7	22,6	40,9	14,3	41,5	29,6	35,3	41,4	14,1	26,9	41,8	33,8
12,6	25,1	43,8	19,7	34,9	19,1	37,7	19,2	40,7	41,5	33,0	27,0	20,4	27,7	43,6	31,7
10,6	26,2	45,5	19,7	29,1	29,9	36,0	17,4	35,8	39,4	35,0	45,1	21,7	29,5	40,3	34,6
16,2	32,2	36,6	17,7	26,4	30,3	45,6	15,3	27,4	41,1	34,4	39,9	20,5	31,8	45,4	25,0
15,1	28,7	65,4	20,6	24,5	16,4	41,1	16,6	32,5	51,0	35,0	27,9	24,3	41,7	52,5	16,7
18,3	29,7	53,0	18,6	23,3	26,2	26,2	15,3	26,7	43,5	36,5	44,6	19,4	58,5	31,3	33,0
17,1	28,8	66,6	19,4	27,9	23,7	45,0	17,8	31,7	41,4	36,3	33,3	22,2	28,9	36,3	28,9
15,5	27,7	36,1	13,3	34,7	22,3	33,5	15,7	29,7	47,2	38,6	50,8	16,8	61,3	41,7	32,1
14,8	27,5	42,4	18,4	26,9	25,7	43,9	20,6	24,6	36,4	30,1	41,4	21,6	40,6	47,5	38,8
18,5	23,5	34,4	17,8	24,9	21,3	36,4	20,3	28,0	37,2	30,1	48,8	18,5	32,2	40,7	39,3
17,2	25,4	41,4	19,1	27,1	26,2	35,0	18,7	29,2	31,1	35,0	54,9	23,1	42,9	43,3	40,2
17,0	28,1	51,1	16,7	23,7	19,4	35,5	20,8	51,2	38,3	31,5	56,3	22,4	41,0	50,1	33,3
15,3	23,7	28,6	22,4	22,0	21,7	36,5	29,0	32,9	40,4	34,7	36,4	16,5	27,7	39,3	34,6
15,3	27,9	34,5	32,8	38,3	24,3	36,6	18,7	34,8	46,7	28,4	39,6	22,5	25,5	57,0	32,2
16,5	26,9	34,0	18,1	25,5	22,4	34,7	15,3	26,7	40,2	28,4	51,4	17,2	24,6	40,1	29,4
18,0	25,8	44,0	21,5	22,4	20,4	31,1	19,7	23,2	40,7	35,1	38,6	21,4	22,8	46,7	39,0
21,7	25,1	41,8	23,5	27,2	24,4	37,1	20,1	42,3	46,7	30,9	24,3	17,9	44,4	35,0	30,9
17,9	26,2	38,3	17,8	22,1	15,4	38,6	18,1	27,4	41,1	31,5	32,6	18,8	30,9	39,2	35,3
20,1	32,2	33,4	17,8	24,8	28,6	29,9	18,4	26,9	37,2	23,1	41,3	16,3	24,6	41,2	28,8
18,0	28,7	40,6	19,5	32,6	19,7	37,3	23,2	23,5	32,0	33,4	62,4	19,9	24,9	40,4	29,7
22,1	29,7	39,4	14,4	29,2	19,5	38,2	27,6	27,4	36,8	25,8	40,7	18,4	30,0	44,5	32,8
16,6	28,8	41,8	13,8	23,4	15,9	25,3	18,7	25,3	40,4	31,2	27,1	25,4	29,6	45,1	32,1
15,7	27,7	40,6	19,2	22,3	19,1	34,7	21,7	18,8	40,0	25,7	39,9	23,5	24,7	35,1	32,4

20,3	27,5	35,7	17,3	18,5	18,6	34,2	20,3	27,8	35,9	32,7	31,8	23,9	37,2	43,2	31,3
16,5	23,5	42,3	12,9	17,6	27,4	32,0	20,9	26,9	36,0	41,1	42,0	19,3	24,6	27,5	27,0
16,9	33,1	47,8	14,7	20,3	28,6	36,1	18,8	29,1	34,5	43,8	46,4	16,7	26,2	45,3	30,8
16,5	22,1	43,3	12,9	24,2	38,0	41,1	17,8	28,0	42,7	35,2	23,4	22,3	25,7	41,8	31,0
15,7	27,7	35,6	15,1	22,0	24,2	33,5	14,0	29,4	41,2	33,0	34,8	19,4	25,0	38,1	27,1
18,0	21,8	31,8	19,1	22,2	21,0	35,4	19,6	25,0	42,0	32,4	39,8	12,9	35,7	39,9	30,8
18,1	20,0	39,0	14,6	25,0	22,1	53,5	16,4	30,9	43,9	32,8	32,3	26,7	28,8	40,1	30,1
19,0	18,8	29,8	18,8	22,5	18,9	42,6	14,4	28,0	41,5	25,7	32,4	16,9	38,9	44,8	40,0
18,8	20,4	31,9	16,6	21,8	27,7	52,5	14,1	27,3	40,4	34,7	30,2	16,3	35,8	47,3	30,0
16,3	24,1	31,0	20,3	29,8	21,6	43,1	24,9	31,1	40,5	33,5	73,5	14,4	34,7	33,8	24,0
17,9	18,3	47,8	16,4	31,3	26,6	45,1	19,5	27,1	45,5	26,3	30,1	23,7	26,2	50,5	26,8
14,1	20,2	35,6	14,8	22,7	17,4	42,5	24,8	24,6	21,1	38,2	23,9	24,4	22,6	37,0	24,1
14,6	20,3	37,9	16,4	21,9	14,9	40,4	17,0	25,6	45,1	28,0	29,7	19,4	32,2	46,6	28,1
13,1	22,7	32,7	13,3	18,9	21,8	58,2	26,7	26,5	43,7	49,1	55,0	16,6	25,4	42,9	27,7
17,4	22,3	32,3	13,0	18,8	26,4	43,5	36,1	23,5	34,2	44,0	42,5	19,0	30,6	35,5	33,4
10,7	26,8	35,0	17,8	23,0	26,3	49,0	20,0	25,7	37,7	29,2	45,6	20,8	25,2	45,2	27,3
16,6	21,8	41,0	15,9	21,2	25,4	26,9	19,9	23,3	34,6	30,4	43,5	19,8	22,5	45,7	29,6
17,7	21,0	36,9	11,8	30,3	25,5	28,7	24,1	20,4	36,7	31,9	54,0	19,1	23,5	44,9	23,3
14,7	21,4	39,0	9,4	20,6	35,9	26,3	30,1	28,3	47,8	31,0	63,1	18,0	23,3	58,2	24,8
14,6	23,9	38,0	15,0	28,1	26,4	39,5	22,7	27,4	45,4	23,0	39,7	21,4	21,7	46,4	35,3
17,1	22,7	35,7	12,7	22,6	22,7	35,0	19,9	28,1	49,8	33,3	55,5	19,7	22,6	56,4	34,6
11,9	22,1	33,8	16,3	18,7	19,4	35,0	17,5	24,8	39,4	34,4	36,2	16,7	28,8	56,3	35,0
18,6	25,5	41,8	12,6	22,8	23,2	26,9	21,7	27,7	33,6	33,4	43,1	20,4	38,9	50,9	26,3
13,8	18,7	37,9	16,1	26,3	24,7	38,6	25,9	26,3	27,4	42,3	29,2	16,9	33,7	57,3	23,7
16,1	17,6	42,7	12,7	19,3	21,2	33,1	19,6	32,9	39,1	38,3	33,7	22,6	34,2	52,1	31,8
12,6	26,0	38,3	14,3	68,5	20,5	35,4	19,4	25,8	31,5	26,7	40,9	23,8	28,4	50,6	25,9
11,6	21,3	38,0	12,3	34,7	18,5	32,0	17,6	43,7	34,2	30,6	35,0	19,5	28,9	41,4	25,5
17,2	28,6	35,9	14,3	20,2	23,0	35,0	24,1	35,0	35,1	29,3	36,8	14,7	38,8	57,9	30,2
10,5	31,3	35,8	12,3	23,7	33,1	33,8	22,3	33,4	32,9	43,2	34,6	17,2	38,1	61,7	25,5
16,9	21,2	33,2	18,1	21,0	34,7	39,9	23,5	26,9	27,9	33,3	34,7	13,5	43,9	52,4	30,0
16,8	24,0	43,1	18,9	20,4	29,3	37,7	25,0	24,5	58,5	49,8	42,6	15,9	34,5	52,1	27,6
22,4	29,3	42,6	17,8	22,9	25,2	39,1	21,1	24,7	58,5	35,7	42,9	21,5	41,2	36,0	21,1
16,1	23,4	39,0	19,3	17,5	31,9	38,0	23,9	24,2	27,4	30,7	28,4	19,7	38,9	52,2	24,1
16,8	28,3	37,3	25,0	25,1	43,4	32,6	19,0	22,7	45,0	37,5	33,5	16,6	40,6	54,4	24,0
17,5	22,8	34,3	19,4	20,0	28,5	44,9	18,5	24,4	31,3	33,0	45,0	17,3	35,4	44,9	29,3
18,8	20,3	42,2	23,2	19,4	22,6	37,8	20,1	21,8	39,6	32,0	35,4	18,2	32,5	44,1	20,7
15,2	18,3	28,1	18,5	23,2	17,2	31,8	16,4	27,3	56,2	28,8	44,2	12,8	38,4	43,0	32,2
16,7	27,1	30,4	19,9	22,1	21,6	34,5	18,7	23,9	28,8	31,7	37,0	17,0	37,7	42,7	33,1
12,7	28,5	31,6	22,3	20,5	23,5	41,0	22,1	26,2	54,7	40,2	39,9	19,5	37,0	27,6	27,2
10,3	23,4	34,3	15,4	23,3	24,5	37,1	17,0	28,0	26,7	46,9	38,7	26,7	42,2	25,1	27,7
15,4	23,5	42,2	22,0	20,7	29,1	39,0	19,3	23,8	51,9	28,9	33,5	20,9	45,0	31,2	21,7
11,7	21,1	32,9	24,4	18,9	27,3	39,4	20,5	26,3	40,9	45,0	38,8	19,4	33,6	29,9	18,2
18,7	22,6	37,1	16,2	18,6	18,0	36,2	21,4	24,4	49,8	26,7	28,6	17,8	59,3	41,5	17,8
19,1	22,5	41,1	19,5	19,6	21,9	43,6	18,1	24,2	62,3	25,8	36,1	21,4	38,8	30,5	23,8
15,9	22,2	30,0	15,3	24,9	20,3	43,6	26,2	27,1	41,7	20,6	36,7	19,7	43,3	27,1	25,6

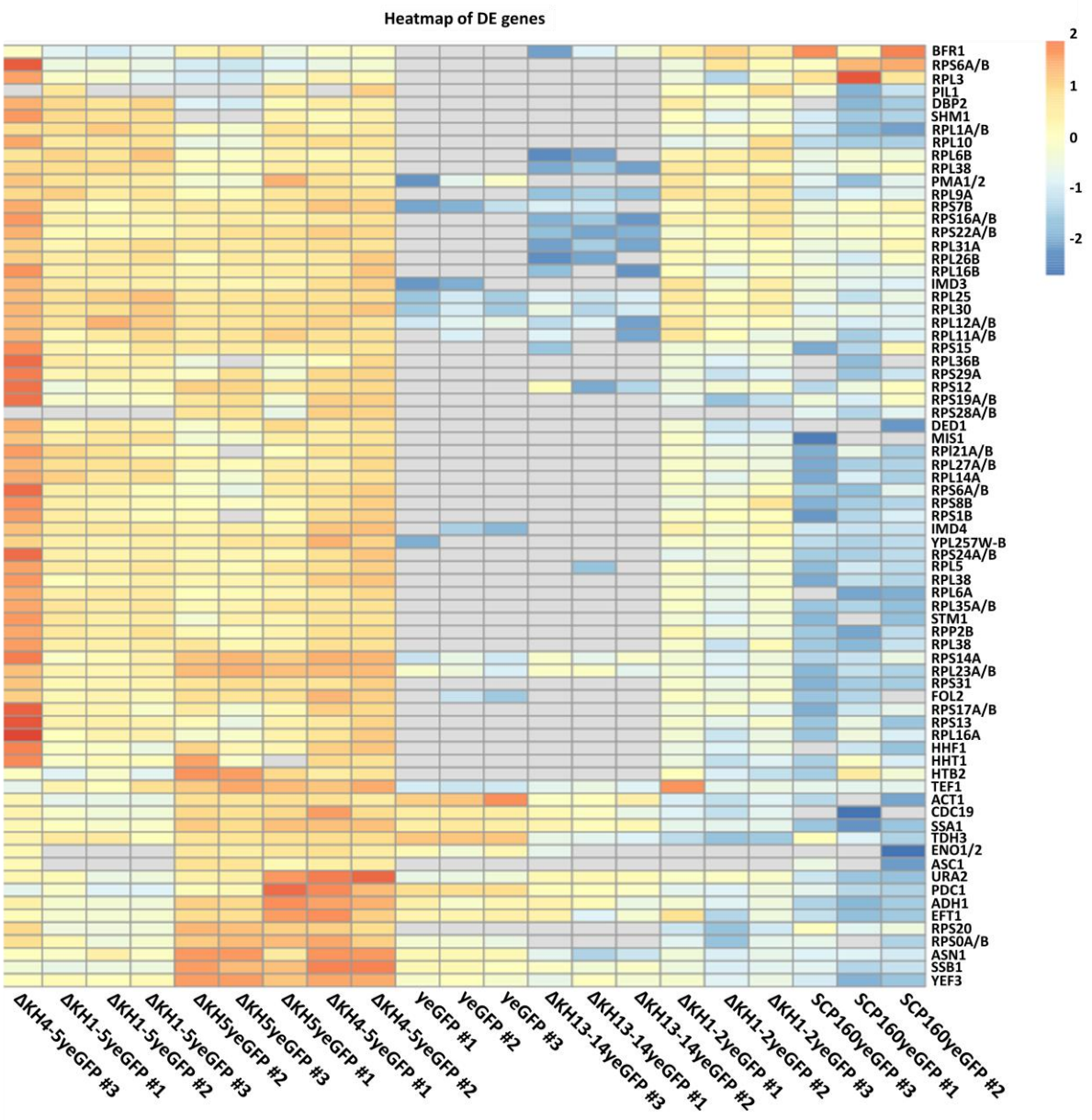
19,1	20,8	32,7	17,7	17,1	23,1	39,8	17,6	23,7	40,3	38,2	30,3	19,3	35,0	34,4	35,9
13,6	22,3	32,3	17,1	20,7	18,8	40,1	19,1	19,0	29,6	37,1	38,6	17,6	44,8	26,3	27,4
20,3	20,9	38,1	19,8	15,2	16,8	37,0	20,8	22,2	36,5	34,6	38,2	22,3	33,8	27,4	30,1
23,7	20,8	26,6	15,2	21,0	20,7	34,8	23,0	24,2	40,3	33,2	90,6	21,1	37,0	22,7	17,6
19,5	22,4	32,3	15,4	20,7	26,1	31,2	18,0	23,8	70,2	34,1	48,1	16,0	38,8	26,4	24,4
16,4	24,0	33,3	16,8	25,2	20,0	30,8	18,4	25,0	38,8	19,7	29,3	22,3	41,6	31,0	29,0
18,3	19,2	41,1	24,5	22,7	25,4	46,5	21,6	20,4	68,4	35,7	45,9	21,6	41,5	25,6	27,1
23,2	26,4	40,5	28,1	25,1	27,7	37,2	20,5	22,0	56,5	21,4	42,2	25,3	51,4	30,3	29,6
13,4	20,6	40,5	15,7	29,1	27,7	45,6	16,9	21,8	45,8	23,4	54,7	22,2	41,4	27,0	27,1
18,0	20,5	33,2	14,7	21,8	19,0	37,5	17,4	28,9	41,8	24,2	41,2	19,9	34,4	25,5	28,9
17,4	19,4	30,9	18,8	23,2	27,3	25,0	18,0	20,0	31,2	40,8	43,5	19,3	43,3	27,3	27,1
20,0	19,7	41,7	14,4	21,0	28,7	39,2	19,0	26,2	41,5	27,1	32,4	16,5	40,2	27,6	38,9
16,3	20,9	32,9	17,5	20,0	21,0	47,8	16,6	19,0	36,9	31,4	44,6	15,5	38,4	33,4	31,6
14,9	17,0	33,0	18,0	28,3	24,7	39,1	16,8	26,5	34,8	24,6	37,6	26,7	29,3	22,4	26,8
17,8	24,6	40,2	14,4	42,7	22,8	46,0	18,4	31,2	37,0	33,3	50,1	21,2	34,4	45,1	22,9
16,7	30,2	38,8	16,4	33,8	20,8	35,8	15,3	26,3	25,3	23,8	49,2	20,7	41,9	35,5	26,6
17,5	26,8	47,2	14,6	32,9	30,0	36,9	15,3	28,9	39,7	32,7	56,0	15,6	46,1	20,0	39,5
19,6	22,8	34,1	20,3	37,5	26,2	41,3	20,6	24,3	33,8	33,3	61,7	19,6	31,1	24,2	34,4
15,4	21,5	40,3	18,0	53,8	26,1	37,4	19,4	45,1	61,5	29,8	46,7	16,5	42,8	23,8	33,4
24,0	26,9	35,0	19,2	23,4	24,5	50,3	20,5	31,1	25,0	39,6	45,6	17,9	35,6	25,8	33,4
28,2	25,8	38,3	24,8	31,9	23,3	40,5	20,3	35,4	19,7	32,8	54,1	16,6	30,6	20,5	27,7
24,6	27,2	37,6	23,8	44,6	15,4	36,5	18,2	30,4	28,1	20,8	51,5	18,3	31,8	30,8	30,3
23,8	26,7	51,2	17,0	39,3	24,8	38,3	23,5	34,4	28,8	19,6	52,0	21,1	36,0	23,4	37,4
28,3	28,6	42,3	17,2	34,7	20,2	32,6	15,8	22,6	28,4	27,5	51,3	17,4	38,9	30,6	36,2
21,4	26,3	28,8	16,0	35,2	19,0	34,3	22,4	23,0	39,3	26,2	49,2	18,4	32,0	26,2	22,8
20,4	21,1	35,3	15,9	29,7	18,8	41,1	19,9	48,3	30,7	31,0	56,0	24,1	35,0	23,0	23,0



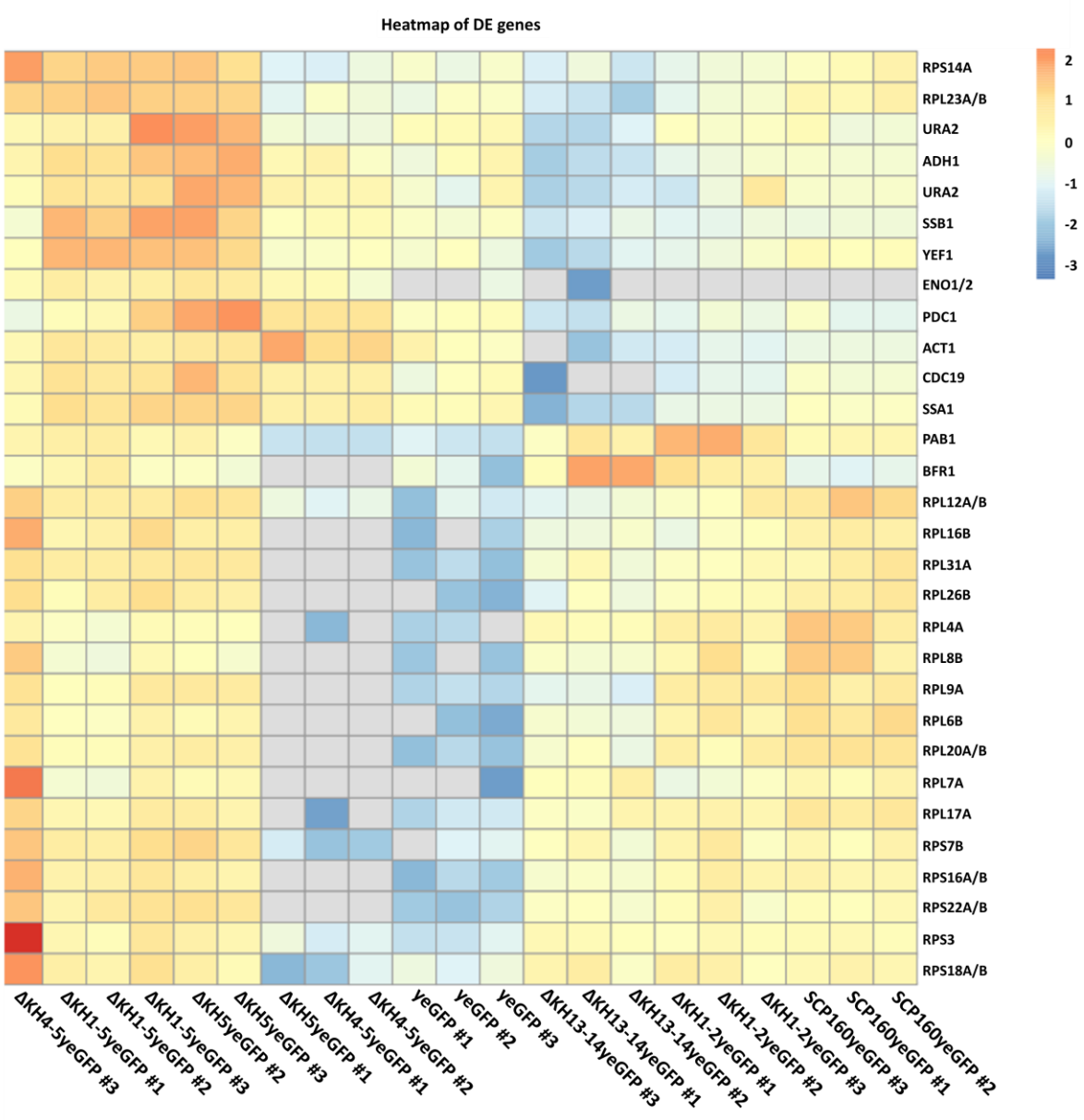
Supplementary Figure 3: Heatmap of the differentially enriched protein interaction partners ('DE genes', top): Comparison of Scp160ΔKH1-2 to WT; Color scale: orange to red: depletion of interacting proteins; light to dark blue: enrichment of interacting proteins; gene name of the interactors (rows) depicted on the right side of the heatmap.



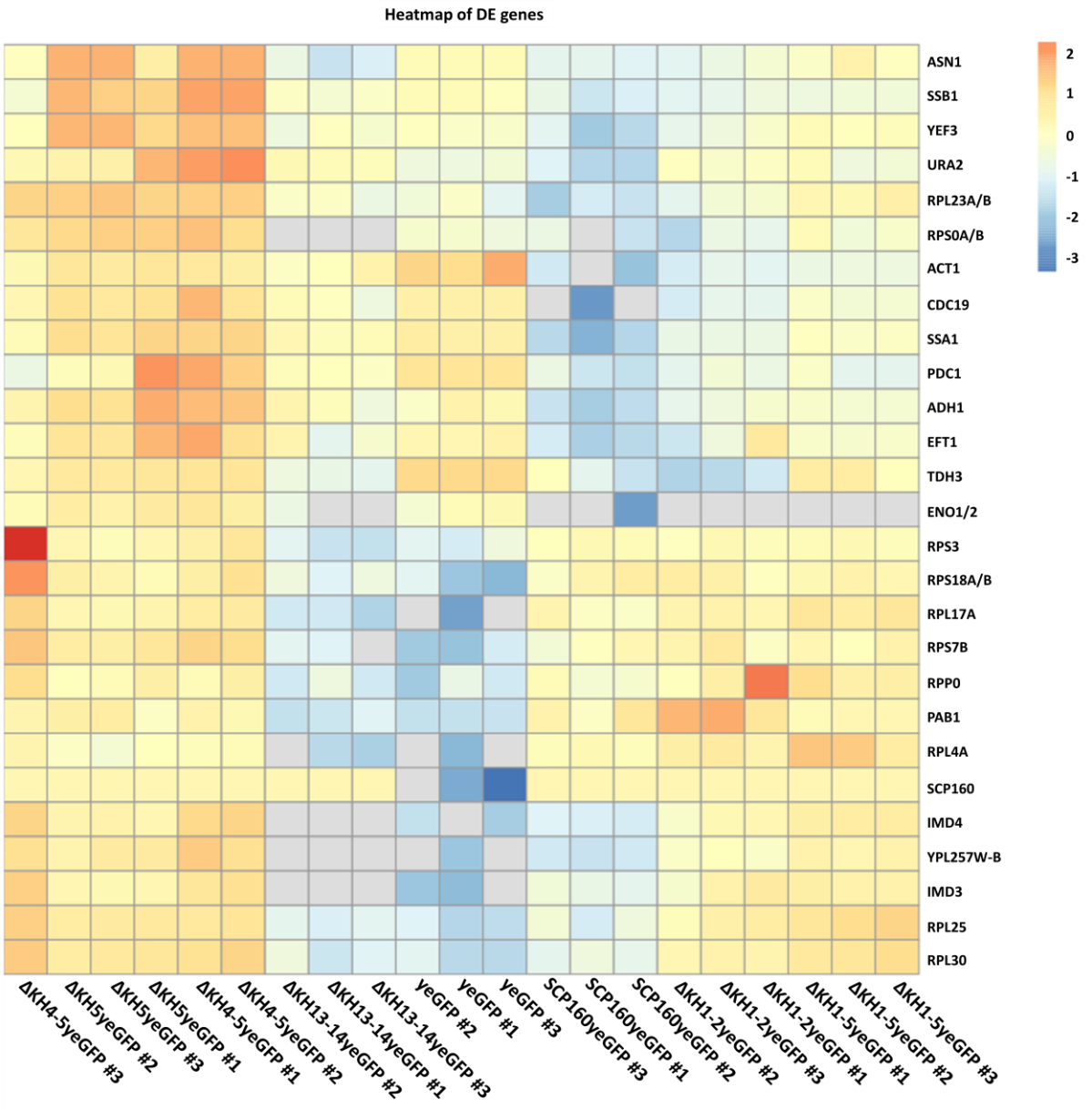
Supplementary Figure 4: Heatmap of the differentially enriched protein interaction partners ('DE genes', top): Comparison of Scp160ΔKH1-5 to WT; Color scale: orange to red: depletion of interacting proteins; light to dark blue: enrichment of interacting proteins; gene name of the interactors (rows) depicted on the right side of the heatmap.



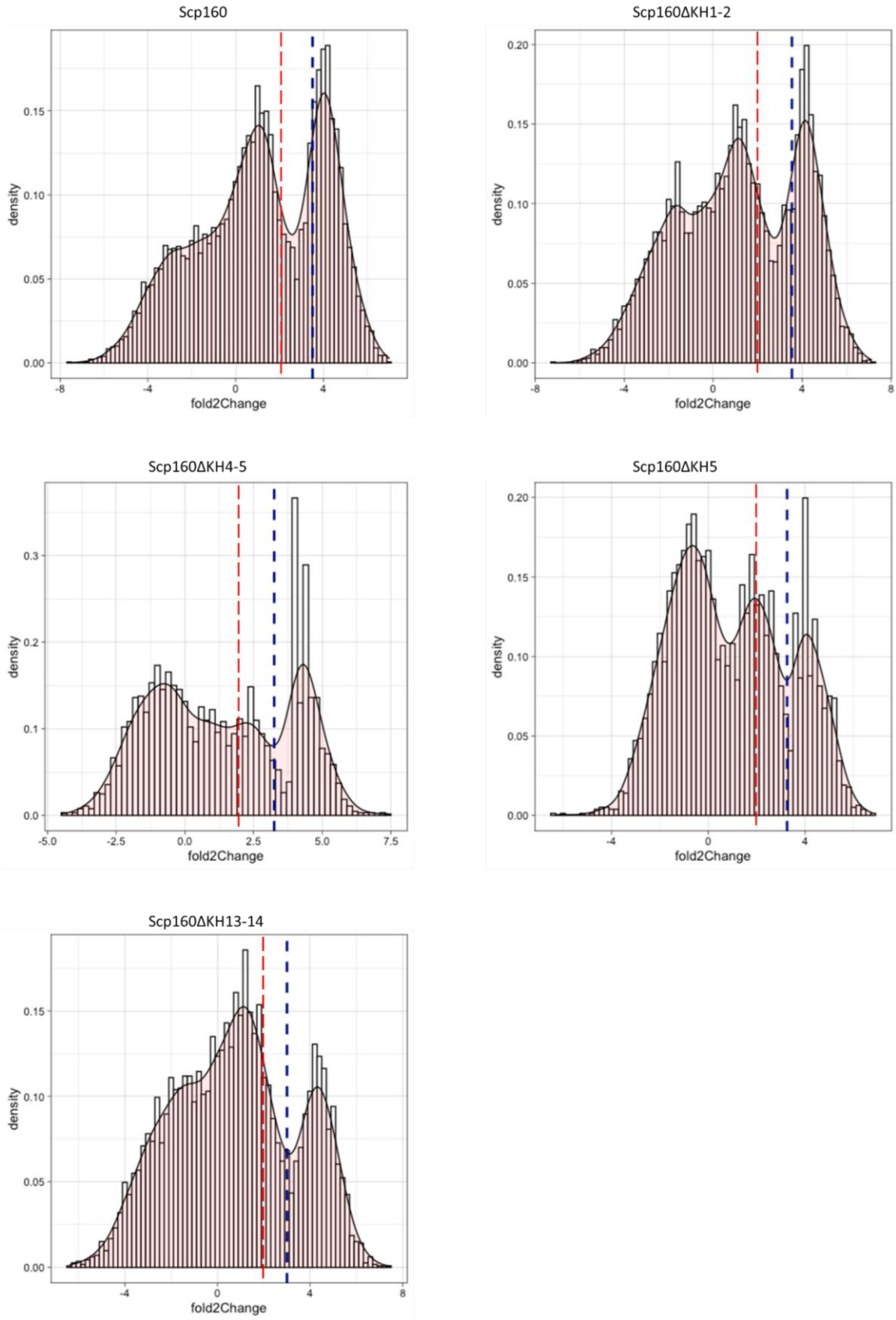
Supplementary Figure 5: Heatmap of the differentially enriched protein interaction partners ('DE genes', top): Comparison of Scp160ΔKH5 to WT; Color scale: orange to red: depletion of interacting proteins; light to dark blue: enrichment of interacting proteins; gene name of the interactors (rows) depicted on the right side of the heatmap.



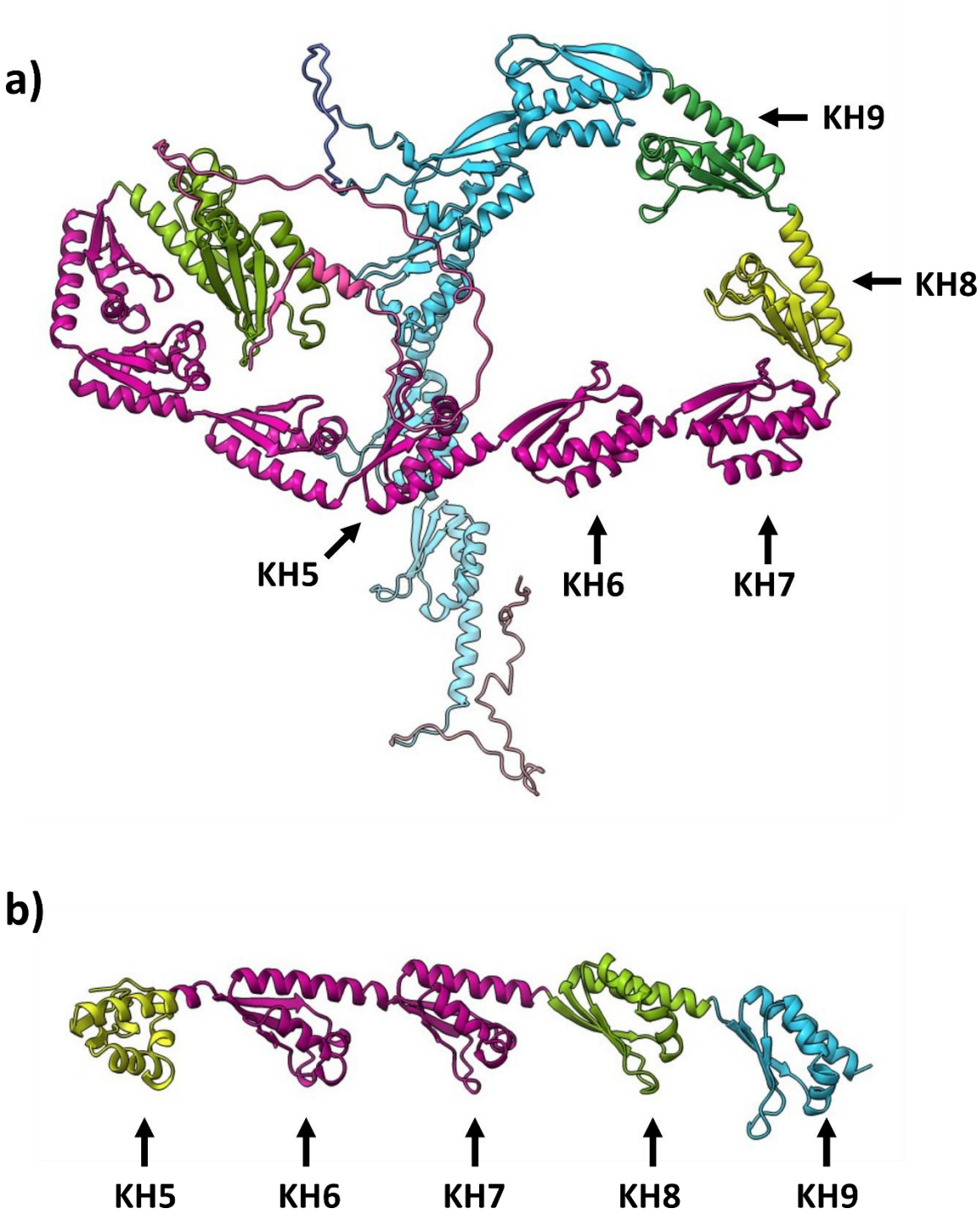
Supplementary Figure 6: Heatmap of the differentially enriched protein interaction partners ('DE genes', top): Comparison of Scp160ΔKH13-14 to WT; Color scale: orange to red: depletion of interacting proteins; light to dark blue: enrichment of interacting proteins; gene name of the interactors (rows) depicted on the right side of the heatmap.



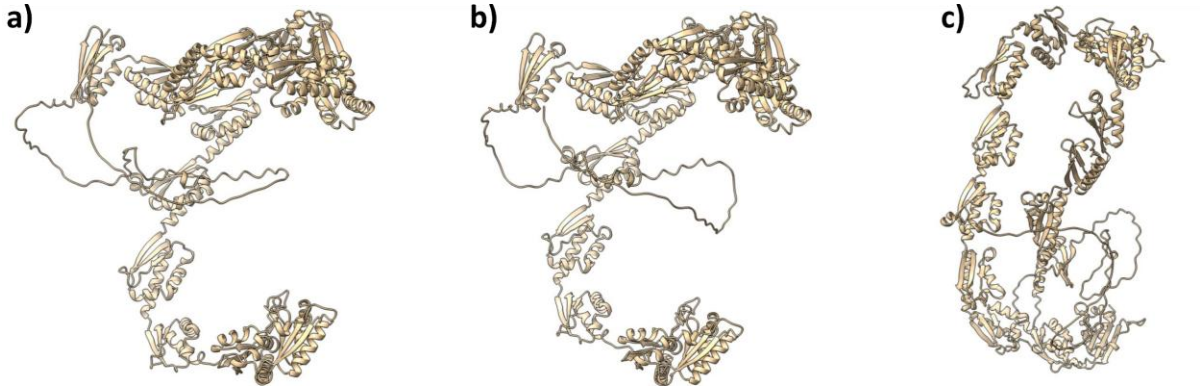
Supplementary Figure 7: Heatmap of the differentially enriched protein interaction partners ('DE genes', top): Comparison of yeGFP to Scp160; Color scale: orange to red: depletion of interacting proteins; light to dark blue: enrichment of interacting proteins; gene name of the interactors (rows) depicted on the right side of the heatmap.



Supplementary Figure 8: Fold change analysis in different Scp160 variants; shoulder peaks were used for determination of significance level applied in the volcano plots (Figure 46-Figure 49); red dotted line: chosen fold change (Log_2FC: 2); blue dotted line: alternative cut off for more stringent analysis.



Supplementary Figure 9: AlphaFold3 predicted structure of human vigilin (Hdlbp) illustrated in ChimeraX; KH domain position indicated by arrow symbol; a) full-length Hdlbp; b) Hdlbp fragment consisting only of KH domains 5-9; Predicted Aligned Error (PAE) coloration: identical coloration suggest correct fold and relative positioning.



Supplementary Figure 10: Comparison of rescue attempts of diverged KH domains: a) wild-type Scp160; b) Scp160 with diverged GXXG motif of KH domain five (HAKN) replaced with conserved GXXG motif (GKGG); c) Scp160 with diverged KH domain five replaced by conserved KH domain number eight.

Supplementary Table 2: Motif Analysis of Scp160 and KH deleted Scp160 variants; A= adenine, T= thymine, C= cytosine, G= guanine, N= variable nucleobase.

method	Scp160	Scp160ΔKH1-2	Scp160ΔKH4-5	Scp160ΔKH5	Scp160ΔKH13-14
CHIP-MEME	CAUUGAUGN	GNUGNNNUAUCG	GGCACUGUAC	UGGUGG	AGCAGN
		ACUGGU	GACUUUUACUGGG		
			CUUCCAUC		
XSTREME	GNUGGUGN	GCUNNUGNUGNNGAN	GUACUUGCUUCUAUC	UGGUGG	GGUGCNGUUAUUGNU
	AUGCCA	ACUGGU	UGAAGGCUUUGACGG	GGAACU	GAUGGNGAAUNUGUC
	GGUUCC	AGGCAAANUU	ACAUCCCUGUGCCUU	CCAUUUCUNA	GGAGAN
	AUUCUGGA	GAGUACACNU	GGUAUCUAAAA	AUUUGUUG	ACCACAGACCG
		AUUUGACCNAC	AAAUAUCUCCG		AAUGGGCNC
			UGCAAUCCAAUACCG		GCGCAAG
			GCAUUAGACAAUAAC		
			CGUUGAUGAUUG		
			CGGCACUG		

Supplementary Table 3: Number of independent (external) experiments in which specific proteins were found to bind Scp160; #: number.

Gene names	# Experiments	protein IDs	Publication
BFR1	7	P38934	Lang 2000, Lang 2001, Li 2004, Sezen 2009, Weidner 2014, Lapointe 2015, Michaelis 2023
ASC1	4	P38011	Gavin 2002, Gavin 2006, Opitz 2017, Rössler 2019
RPS2	4	P25443	Delaveau 2016, Schmitt 2019, Rössler 2019, Schmitt 2021
ARB1	3	P40024	Dong 2005 (2x), Delaveau 2016
PAT1	3	P25644	Mitchel 2013, Weidner 2014, Delaveau 2016
SIS1	3	P25294	Delaveau 2016, Feder 2021, Ali 2023
ZUO1	3	P32527	Gavin 2002, Gavin 2006, Delaveau 2016

References

Ali, A., Garde, R., Schaffer, O.C., Bard, J.A.M., Husain, K., Kik, S.K., Davis, K.A., Luengo-Woods, S., Igarashi, M.G., Drummond, D.A., Squires, A.H., Pincus, D., 2023. Adaptive preservation of orphan ribosomal proteins in chaperone-dispersed condensates. Nat. Cell Biol. 25, 1691–1703. https://doi.org/10.1038/s41556-023-01253-2

Altemose, N., Logsdon, G.A., Bzikadze, A.V., Sidhwani, P., Langley, S.A., Caldas, G.V., Hoyt, S.J., Uralsky, L., Ryabov, F.D., Shew, C.J., Sauria, M.E.G., Borchers, M., Gershman, A., Mikheenko, A., Shepelev, V.A., Dvorkina, T., Kunyavskaya, O., Vollger, M.R., Rhie, A., McCartney, A.M., Asri, M., Lorig-Roach, R., Shafin, K., Lucas, J.K., Aganezov, S., Olson, D., De Lima, L.G., Potapova, T., Hartley, G.A., Haukness, M., Kerpedjiev, P., Gusev, F., Tigyi, K., Brooks, S., Young, A., Nurk, S., Koren, S., Salama, S.R., Paten, B., Rogaev, E.I., Streets, A., Karpen, G.H., Dernburg, A.F., Sullivan, B.A., Straight, A.F., Wheeler, T.J., Gerton, J.L., Eichler, E.E., Phillippy, A.M., Timp, W., Dennis, M.Y., O'Neill, R.J., Zook, J.M., Schatz, M.C., Pevzner, P.A., Diekhans, M., Langley, C.H., Alexandrov, I.A., Miga, K.H., 2022. Complete genomic and epigenetic maps of human centromeres. Science 376, eabl4178. https://doi.org/10.1126/science.abl4178

Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., & Struhl, K. (1992). Short protocols in molecular biology. *New York*, *275*, 28764-28773.

Auweter, S.D., Oberstrass, F.C., Allain, F.H.-T., 2006. Sequence-specific binding of single-stranded RNA: is there a code for recognition? Nucleic Acids Res. 34, 4943–4959. https://doi.org/10.1093/nar/gkl620

Aviram, N., Ast, T., Costa, E.A., Arakel, E.C., Chuartzman, S.G., Jan, C.H., Haßdenteufel, S., Dudek, J., Jung, M., Schorr, S., Zimmermann, R., Schwappach, B., Weissman, J.S., Schuldiner, M., 2016. The SND proteins constitute an alternative targeting route to the endoplasmic reticulum. Nature 540, 134–138. https://doi.org/10.1038/nature20169

Batlle, M., Marsellach, F.-X., Huertas, D., Azorín, F., 2011. Drosophila vigilin, DDP1, localises to the cytoplasm and associates to the rough endoplasmic reticulum. Biochim. Biophys. Acta BBA - Gene Regul. Mech. 1809, 46–55. https://doi.org/10.1016/j.bbagrm.2010.10.005

Baum, S., Bittins, M., Frey, S., Seedorf, M., 2004. Asc1p, a WD40-domain containing adaptor protein, is required for the interaction of the RNA-binding protein Scp160p with polysomes. Biochem. J. 380, 823–830. https://doi.org/10.1042/bj20031962

Bayne, R. A., Jayachandran, U., Kasprowicz, A., Bresson, S., Tollervey, D., Wallace, E. W., & Cook, A. G. (2022). Yeast Ssd1 is a non-enzymatic member of the RNase II family with an alternative RNA recognition site. Nucleic Acids Research, 50(5), 2923-2937

Benjamini, Y., Hochberg, Y., 1995. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J. R. Stat. Soc. Ser. B Methodol. 57, 289–300.

Böhni, P.C., Deshaies, R.J., Schekman, R.W., 1988. SEC11 is required for signal peptide processing and yeast cell growth. J. Cell Biol. 106, 1035–1042. https://doi.org/10.1083/jcb.106.4.1035

Bohnsack, M. T., Tollervey, D., & Granneman, S. (2012). Identification of RNA helicase target sites by UV cross-linking and analysis of cDNA. In *Methods in enzymology* (Vol. 511, pp. 275-288). Academic Press.

Brachmann, B., C., Davies, A., Cost, G. J., Caputo, E., Li, J., Hieter, P., & Boeke, J. D. (1998). Designer deletion strains derived from Saccharomyces cerevisiae S288C: a useful set of strains and plasmids for PCR-mediated gene disruption and other applications. *Yeast*, *14*(2), 115-132.

Brangwynne, C.P., Koenderink, G.H., MacKintosh, F.C., Weitz, D.A., 2009. Intracellular transport by active diffusion. Trends Cell Biol. 19, 423–427. https://doi.org/10.1016/j.tcb.2009.04.004

Brennan, C. M., & Steitz*, J. A. (2001). HuR and mRNA stability. Cellular and Molecular Life Sciences CMLS, 58, 266-277

Brugier, A., Hafirrassou, M.L., Pourcelot, M., Baldaccini, M., Kril, V., Couture, L., Kümmerer, B.M., Gallois-Montbrun, S., Bonnet-Madin, L., Vidalain, P.-O., Delaugerre, C., Pfeffer, S., Meertens, L., Amara, A., 2022. RACK1 Associates with RNA-Binding Proteins Vigilin and SERBP1 to Facilitate Dengue Virus Replication. J. Virol. 96, e01962-21. https://doi.org/10.1128/jvi.01962-21

Brykailo, Melissa A., Corbett, A.H., Fridovich-Keil, J.L., 2007. Functional overlap between conserved and diverged KH domains in Saccharomyces cerevisiae SCP160. Nucleic Acids Res. 35, 1108–1118. https://doi.org/10.1093/nar/gkl1160

Brykailo, M. A., McLane, L.M., Fridovich-Keil, J., Corbett, A.H., 2007. Analysis of a predicted nuclear localization signal: implications for the intracellular localization and function of the Saccharomyces cerevisiae RNA-binding protein Scp160. Nucleic Acids Res. 35, 6862–6869. https://doi.org/10.1093/nar/gkm776

Cameron, D.R., Cooper, D.G., Neufeld, R.J., 1988. The mannoprotein of Saccharomyces cerevisiae is an effective bioemulsifier. Appl. Environ. Microbiol. 54, 1420–1425. https://doi.org/10.1128/aem.54.6.1420-1425.1988

Caro, L.G., Palade, G.E., 1964. PROTEIN SYNTHESIS, STORAGE, AND DISCHARGE IN THE PANCREATIC EXOCRINE CELL. J. Cell Biol. 20, 473–495. https://doi.org/10.1083/jcb.20.3.473

Chen, D.-C., Yang, B.-C., Kuo, T.-T., 1992. One-step transformation of yeast in stationary phase. Curr. Genet. 21, 83–84. https://doi.org/10.1007/BF00318659

Chen, J.-Y., Chen, J.-C., Wu, J.-L., 2003. Molecular cloning and functional analysis of zebrafish high-density lipoprotein-binding protein. Comp. Biochem. Physiol. B Biochem. Mol. Biol. 136, 117–130. https://doi.org/10.1016/S1096-4959(03)00181-7

Cheng, M.H.K., Hoffmann, P.C., Franz-Wachtel, M., Sparn, C., Seng, C., Maček, B., Jansen, R.-P., 2018. The RNA-Binding Protein Scp160p Facilitates Aggregation of Many Endogenous Q/N-Rich Proteins. Cell Rep. 24, 20–26. https://doi.org/10.1016/j.celrep.2018.06.015

Chial, H.J., Giddings, T.H., Siewert, E.A., Hoyt, M.A., Winey, M., 1999. Altered dosage of the Saccharomyces cerevisiae spindle pole body duplication gene, NDC1 , leads to aneuploidy and polyploidy. Proc. Natl. Acad. Sci. 96, 10200–10205. https://doi.org/10.1073/pnas.96.18.10200

Cho, P.F., Poulin, F., Cho-Park, Y.A., Cho-Park, I.B., Chicoine, J.D., Lasko, P., Sonenberg, N., 2005. A New Paradigm for Translational Control: Inhibition via 5'-3' mRNA Tethering by Bicoid and the eIF4E Cognate 4EHP. Cell 121, 411–423. https://doi.org/10.1016/j.cell.2005.02.024

Cléry, Antoine, and H T. Allain Frédéric. "From Structure to Function of RNA Binding Domains." (2012).

Cohen-Zontag, O., Baez, C., Lim, L.Q.J., Olender, T., Schirman, D., Dahary, D., Pilpel, Y., Gerst, J.E., 2019. A secretion-enhancing cis regulatory targeting element (SECReTE) involved in mRNA localization and protein synthesis. PLOS Genet. 15, e1008248. https://doi.org/10.1371/journal.pgen.1008248

Corley, M., Burns, M.C., Yeo, G.W., 2020. How RNA-Binding Proteins Interact with RNA: Molecules and Mechanisms. Mol. Cell 78, 9–29. https://doi.org/10.1016/j.molcel.2020.03.011

Cortes, A., 1999. DDP1, a single-stranded nucleic acid-binding protein of Drosophila, associates with pericentric heterochromatin and is functionally homologous to the yeast Scp160p, which is involved in the control of cell ploidy. EMBO J. 18, 3820–3833. https://doi.org/10.1093/emboj/18.13.3820

Costa, E.A., Subramanian, K., Nunnari, J., Weissman, J.S., 2018. Defining the physiological role of SRP in protein-targeting efficiency and specificity. Science 359, 689–692. https://doi.org/10.1126/science.aar3607

Cox J, Mann M (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol.* 26(12): 1367

Cox J, Neuhauser N, Michalski A, Scheltema RA, Olsen JV, Mann M (2011). Andromeda: a peptide search engine integrated into the MaxQuant environment. *J Proteome Res.* 1;10(4): 1794

Currie, J.R., Brown, W.T., 1999. KH domain-containing proteins of yeast: Absence of a fragile X gene homologue. Am. J. Med. Genet. 84, 272–276. https://doi.org/10.1002/(SICI)1096-8628(19990528)84:3<272::AID-AJMG21>3.0.CO;2-D

David, D., Sundararabu, S., Gerst, J.E., 1998. Involvement of Long Chain Fatty Acid Elongation in the Trafficking of Secretory Vesicles in Yeast. J. Cell Biol. 143, 1167–1182.

Delaveau, T., Davoine, D., Jolly, A., Vallot, A., Rouvière, J.O., Gerber, A., Brochet, S., Plessis, M., Roquigny, R., Merhej, J., Leger, T., Garcia, C., Lelandais, G., Laine, E., Palancade, B., Devaux, F., Garcia, M., 2016. Tma108, a putative M1 aminopeptidase, is a specific nascent chain-associated protein in Saccharomyces cerevisiae. Nucleic Acids Res. 44, 8826–8841. https://doi.org/10.1093/nar/gkw732

Delgado, M., Gil, M., Gozalbo, D., 2003. Starvation and temperature upshift cause an increase in the enzymatically active cell wall-associated glyceraldehyde-3-phosphate dehydrogenase protein in yeast. FEMS Yeast Res. 4, 297–303. https://doi.org/10.1016/S1567-1356(03)00159-4

Delgado, M.L., O'Connor, J.E., Azorín, I., Renau-Piqueras, J., Gil, M.L., Gozalbo, D., 2001. The glyceraldehyde-3-phosphate dehydrogenase polypeptides encoded by the Saccharomyces cerevisiae TDH1, TDH2 and TDH3 genes are also cell wall proteins. Microbiology 147, 411–417. https://doi.org/10.1099/00221287-147-2-411

Dodson, R.E., Shapiro, D.J., 1997. Vigilin, a Ubiquitous Protein with 14 K Homology Domains, Is the Estrogen-inducible Vitellogenin mRNA 3'-Untranslated Region-binding Protein. J. Biol. Chem. 272, 12249–12252. https://doi.org/10.1074/jbc.272.19.12249

Dodson, R.E., Shapiro, D.J., n.d. Regulationof Pathways of mRNA Destabilization and Stabilization.

Dong, J., Lai, R., Jennings, J.L., Link, A.J., Hinnebusch, A.G., 2005. The Novel ATP-Binding Cassette Protein ARB1 Is a Shuttling Factor That Stimulates 40S and 60S Ribosome Biogenesis. Mol. Cell. Biol. 25, 9859–9873. https://doi.org/10.1128/MCB.25.22.9859-9873.2005

Egea, P.F., Stroud, R.M., Walter, P., 2005. Targeting proteins to membranes: structure of the signal recognition particle. Curr. Opin. Struct. Biol. 15, 213–220. https://doi.org/10.1016/j.sbi.2005.03.007

Ergüden, B., 2019. Dhh1 is a member of the SESA network. Yeast 36, 99–105. https://doi.org/10.1002/yea.3363

Feder, Z.A., Ali, A., Singh, A., Krakowiak, J., Zheng, X., Bindokas, V.P., Wolfgeher, D., Kron, S.J., Pincus, D., 2021. Subcellular localization of the J-protein Sis1 regulates the heat shock response. J. Cell Biol. 220, e202005165. https://doi.org/10.1083/jcb.202005165

Fisk, D. G., Ball, C. A., Dolinski, K., Engel, S. R., Hong, E. L., Issel-Tarver, L., ... & Michael Cherry, J. (2006). Saccharomyces cerevisiae S288C genome annotation: a working hypothesis. Yeast, 23(12), 857-865

Frey, S., Pool, M., Seedorf, M., 2001. Scp160p, an RNA-binding, Polysome-associated Protein, Localizes to the Endoplasmic Reticulum of Saccharomyces cerevisiae in a Microtubule-dependent Manner. J. Biol. Chem. 276, 15905–15912. https://doi.org/10.1074/jbc.M009430200

Gable, K., Slife, H., Bacikova, D., Monaghan, E., & Dunn, T. M. (2000). Tsc3p is an 80-amino acid protein associated with serine palmitoyltransferase and required for optimal enzyme activity. Journal of Biological Chemistry, 275(11), 7597-7603.

Gai, D., Li, D., Finkielstein, C.V., Ott, R.D., Taneja, P., Fanning, E., Chen, X.S., 2004. Insights into the Oligomeric States, Conformational Changes, and Helicase Activities of SV40 Large Tumor Antigen. J. Biol. Chem. 279, 38952–38959. https://doi.org/10.1074/jbc.M406160200

Galardini, M., Busby, B. P., Vieitez, C., Dunham, A. S., Typas, A., & Beltrao, P. (2019). The impact of the genetic background on gene deletion phenotypes in Saccharomyces cerevisiae. *Molecular Systems Biology*, *15*(12), e8831.

Galitski, T., Saldanha, A.J., Styles, C.A., Lander, E.S., Fink, G.R., 1999. Ploidy Regulation of Gene Expression. Science 285, 251–254. https://doi.org/10.1126/science.285.5425.251

Gavin, A.-C., Aloy, P., Grandi, P., Krause, R., Boesche, M., Marzioch, M., Rau, C., Jensen, L.J., Bastuck, S., Dümpelfeld, B., Edelmann, A., Heurtier, M.-A., Hoffman, V., Hoefert, C., Klein, K., Hudak,

M., Michon, A.-M., Schelder, M., Schirle, M., Remor, M., Rudi, T., Hooper, S., Bauer, A., Bouwmeester, T., Casari, G., Drewes, G., Neubauer, G., Rick, J.M., Kuster, B., Bork, P., Russell, R.B., Superti-Furga, G., 2006. Proteome survey reveals modularity of the yeast cell machinery. Nature 440, 631–636. https://doi.org/10.1038/nature04532

Gavin, A.-C., Bösche, M., Krause, R., Grandi, P., Marzioch, M., Bauer, A., Schultz, J., Rick, J.M., Michon, A.-M., Cruciat, C.-M., Remor, M., Höfert, C., Schelder, M., Brajenovic, M., Ruffner, H., Merino, A., Klein, K., Hudak, M., Dickson, D., Rudi, T., Gnau, V., Bauch, A., Bastuck, S., Huhse, B., Leutwein, C., Heurtier, M.-A., Copley, R.R., Edelmann, A., Querfurth, E., Rybin, V., Drewes, G., Raida, M., Bouwmeester, T., Bork, P., Seraphin, B., Kuster, B., Neubauer, G., Superti-Furga, G., 2002. Functional organization of the yeast proteome by systematic analysis of protein complexes. Nature 415, 141–147. https://doi.org/10.1038/415141a

Geissler, S., Siegers, K., & Schiebel, E. (1998). A novel protein complex promoting formation of functional α-and γ-tubulin. *The EMBO journal*.

Gelin-Licht, R., Conlon, P.J., Singh, R., Nair, R.R., Haimovich, G., Baez, C., Gal, L., Schuldiner, M., Levchenko, A., Gerst, J.E., 2020. Translational control as a novel regulator of gradient sensing and chemotropism in yeast. https://doi.org/10.1101/2020.12.13.422562

Gelin-Licht, R., Paliwal, S., Conlon, P., Levchenko, A., Gerst, J.E., 2012a. Scp160-Dependent mRNA Trafficking Mediates Pheromone Gradient Sensing and Chemotropism in Yeast. Cell Rep. 1, 483–494. https://doi.org/10.1016/j.celrep.2012.03.004

Gelin-Licht, R., Paliwal, S., Conlon, P., Levchenko, A., Gerst, J.E., 2012b. Scp160-Dependent mRNA Trafficking Mediates Pheromone Gradient Sensing and Chemotropism in Yeast. Cell Rep. 1, 483–494. https://doi.org/10.1016/j.celrep.2012.03.004

Genuth, N.R., Barna, M., 2018. The Discovery of Ribosome Heterogeneity and Its Implications for Gene Regulation and Organismal Life. Mol. Cell 71, 364–374. https://doi.org/10.1016/j.molcel.2018.07.018

Goffeau, A., Barrell, B. G., Bussey, H., Davis, R. W., Dujon, B., Feldmann, H., ... & Oliver, S. G. (1996). Life with 6000 genes. Science, 274(5287), 546-567.

Gietz, R.D., Schiestl, R.H., 2007. High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. Nat. Protoc. 2, 31–34. https://doi.org/10.1038/nprot.2007.13

Git, A., Standart, N., 2002. The KH domains of Xenopus Vg1RBP mediate RNA binding and self-association. RNA 8, 1319–1333. https://doi.org/10.1017/S135583820202705X

González, A., Jiménez, A., Vázquez, D., Davies, J.E., Schindler, D., 1978. Studies on the mode of action of hygromycin B, an inhibitor of translocation in eukaryotes. Biochim. Biophys. Acta BBA - Nucleic Acids Protein Synth. 521, 459–469. https://doi.org/10.1016/0005-2787(78)90287-3

Graham, D. L., & Oram, J. F. (1987). Identification and characterization of a high density lipoprotein-binding protein in cell membranes by ligand blotting. Journal of Biological Chemistry, 262(16), 7439-7442

Granneman, S., Kudla, G., Petfalski, E., Tollervey, D., 2009. Identification of protein binding sites on U3 snoRNA and pre-rRNA by UV cross-linking and high-throughput analysis of cDNAs. Proc. Natl. Acad. Sci. 106, 9613–9618. https://doi.org/10.1073/pnas.0901997106

Grant, C. E., & Bailey, T. L. (2021). XSTREME: Comprehensive motif analysis of biological sequence datasets. *BioRxiv*, 2021-09.

Grishin, N.V., 2001. KH domain: one motif, two folds. Nucleic Acids Res. 29, 638–643. https://doi.org/10.1093/nar/29.3.638

Gygi SP, Elias JE (2007). Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. *Nat Methods* **4**(3): 207

Halic, M., Beckmann, R., 2005. The signal recognition particle and its interactions during protein targeting. Curr. Opin. Struct. Biol. 15, 116–125. https://doi.org/10.1016/j.sbi.2005.01.013

Hall, R. A., & Wallace, E. W. (2022). Post-transcriptional control of fungal cell wall synthesis. *The Cell Surface*, *8*, 100074.

Heym, R.G., Niessing, D., 2012. Principles of mRNA transport in yeast. Cell. Mol. Life Sci. 69, 1843–1853. https://doi.org/10.1007/s00018-011-0902-4

Hinnebusch, A.G., 2014. The Scanning Mechanism of Eukaryotic Translation Initiation. Annu. Rev. Biochem. 83, 779–812. https://doi.org/10.1146/annurev-biochem-060713-035802

Hirschmann, W.D., Westendorf, H., Mayer, A., Cannarozzi, G., Cramer, P., Jansen, R.-P., 2014. Scp160p is required for translational efficiency of codon-optimized mRNAs in yeast. Nucleic Acids Res. 42, 4043–4055. https://doi.org/10.1093/nar/gkt1392

Hogan, D.J., Riordan, D.P., Gerber, A.P., Herschlag, D., Brown, P.O., 2008. Diverse RNA-Binding Proteins Interact with Functionally Related Sets of RNAs, Suggesting an Extensive Regulatory System. PLoS Biol. 6, e255. https://doi.org/10.1371/journal.pbio.0060255

Hu, Z., Xia, B., Postnikoff, S. D., Shen, Z. J., Tomoiaga, A. S., Harkness, T. A., ... & Tyler, J. K. (2018). Ssd1 and Gcn2 suppress global translation efficiency in replicatively aged yeast while their activation extends lifespan. *Elife*, *7*, e35551.

Huertas, D., Cortés, A., Casanova, J., Azorín, F., 2004. Drosophila DDP1, a Multi-KH-Domain Protein, Contributes to Centromeric Silencing and Chromosome Segregation. Curr. Biol. 14, 1611–1620. https://doi.org/10.1016/j.cub.2004.09.024

Jackson, R.J., Hellen, C.U.T., Pestova, T.V., 2010. The mechanism of eukaryotic translation initiation and principles of its regulation. Nat. Rev. Mol. Cell Biol. 11, 113–127. https://doi.org/10.1038/nrm2838

Jacquier, A., & Dujon, B. (1985). An intron-encoded protein is active in a gene conversion process that spreads an intron into a mitochondrial gene. Cell, 41(2), 383-394.

Janke, C., Magiera, M.M., Rathfelder, N., Taxis, C., Reber, S., Maekawa, H., Moreno-Borchart, A., Doenges, G., Schwob, E., Schiebel, E., Knop, M., 2004. A versatile toolbox for PCR-based tagging of yeast genes: new fluorescent proteins, more markers and promoter substitution cassettes. Yeast 21, 947–962. https://doi.org/10.1002/yea.1142

Jankowsky, E., 2011. RNA helicases at work: binding and rearranging. Trends Biochem. Sci. 36, 19–29. https://doi.org/10.1016/j.tibs.2010.07.008

Jansen, J. M., Wanless, A. G., Seidel, C. W., & Weiss, E. L. (2009). Cbk1 regulation of the RNA-binding protein Ssd1 integrates cell fate with translational control. *Current Biology*, *19*(24), 2114-2120.

Jensen, R., Sprague Jr, G. F., & Herskowitz, I. (1983). Regulation of yeast mating-type interconversion: feedback control of HO gene expression by the mating-type locus. Proceedings of the National Academy of Sciences, 80(10), 3035-3039

Jiang, L., Lin, N., Cheng, Z., Yang, B., Ou, Q., 2011. The effect of the expression of virulence-associated DEAD-box RNA helicase mRNA on the imbalance of Th1-Th2 cytokines in the CSF of patients with Cryptococcus neoformans meningitis. Eur. J. Clin. Microbiol. Infect. Dis. 30, 1483–1487. https://doi.org/10.1007/s10096-011-1245-7

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., ... & Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. *nature*, *596*(7873), 583-589.

Kaeberlein, M., & Guarente, L. (2002). Saccharomyces cerevisiae MPT5 and SSD1 function in parallel pathways to promote cell wall integrity. *Genetics*, *160*(1), 83-95.

Kainov, D.E., Mancini, E.J., Telenius, J., Lísal, J., Grimes, J.M., Bamford, D.H., Stuart, D.I., Tuma, R., 2008. Structural Basis of Mechanochemical Coupling in a Hexameric Molecular Motor. J. Biol. Chem. 283, 3607–3617. https://doi.org/10.1074/jbc.M706366200

Kanamori, H., Dodson, R.E., Shapiro, D.J., 1998. In Vitro Genetic Analysis of the RNA Binding Site of Vigilin, a Multi-KH-Domain Protein. Mol. Cell. Biol. 18, 3991–4003. https://doi.org/10.1128/MCB.18.7.3991

Klinger, M.H.F., Kruse, C., 1996. Immunocytochemical localization of vigilin, a tRNA-binding protein, after cell fractionation and within the exocrine pancreatic cell of the rat. Ann. Anat. - Anat. Anz. 178, 331–335. https://doi.org/10.1016/S0940-9602(96)80086-0

Kolberg, L., Raudvere, U., Kuzmin, I., Adler, P., Vilo, J., & Peterson, H. (2023). g: Profiler—interoperable web service for functional enrichment analysis and gene identifier mapping (2023 update). *Nucleic acids research*, *51*(W1), W207-W212.

Koonin, E.V., Novozhilov, A.S., 2009. Origin and evolution of the genetic code: The universal enigma. IUBMB Life 61, 99–111. https://doi.org/10.1002/iub.146

Kramer, K., Sachsenberg, T., Beckmann, B.M., Qamar, S., Boon, K.-L., Hentze, M.W., Kohlbacher, O., Urlaub, H., 2014. Photo-cross-linking and high-resolution mass spectrometry for assignment of RNA-binding sites in RNA-binding proteins. Nat. Methods 11, 1064–1070. https://doi.org/10.1038/nmeth.3092

Kruse, C., Grünweller, A., Notbohm, H., Kügler, S., Purschke, W.G., Müller, P.K., 1996. Evidence for a novel cytoplasmic tRNA-protein complex containing the KH-multidomain protein vigilin. Biochem. J. 320, 247–252. https://doi.org/10.1042/bj3200247

Kruse, C., Grünweller, A., Willkomm, K.D., Pfeiffer, T., Hartmann, K.R., Müller, K.P., 1998. tRNA is entrapped in similar, but distinct, nuclear and cytoplasmic ribonucleoprotein complexes, both of which contain vigilin and elongation factor 1α. Biochem. J. 329, 615–621. https://doi.org/10.1042/bj3290615

Kruse, C., Willkomm, D.K., Gru, A., Vollbrandt, T., Sommer, S., Busch, S., Pfeiffer, T., 2000. Export and transport of tRNA are coupled to a multi-protein complex.

Lacroute, F. (1968). Regulation of pyrimidine biosynthesis in Saccharomyces cerevisiae. Journal of bacteriology, 95(3), 824-832.

Labun, K., Montague, T.G., Krause, M., Torres Cleuren, Y.N., Tjeldnes, H., Valen, E., 2019. CHOPCHOP v3: expanding the CRISPR web toolbox beyond genome editing. Nucleic Acids Res. 47, W171–W174. https://doi.org/10.1093/nar/gkz365

Lakkaraju, A.K.K., Luyet, P.-P., Parone, P., Falguières, T., Strub, K., 2007. Inefficient targeting to the endoplasmic reticulum by the signal recognition particle elicits selective defects in post-ER membrane trafficking. Exp. Cell Res. 313, 834–847. https://doi.org/10.1016/j.yexcr.2006.12.003

Lang, B.D., 2001. The brefeldin A resistance protein Bfr1p is a component of polyribosome-associated mRNP complexes in yeast. Nucleic Acids Res. 29, 2567–2574. https://doi.org/10.1093/nar/29.12.2567

Lang, B.D., 2000. Scp160p, a multiple KH-domain protein, is a component of mRNP complexes in yeast. Nucleic Acids Res. 28, 1576–1584. https://doi.org/10.1093/nar/28.7.1576

Lapointe, C. P., Wilinski, D., Saunders, H. A., & Wickens, M. (2015). Protein-RNA networks revealed through covalent RNA marks. *Nature methods*, *12*(12), 1163-1170.

Li, A. -m., 2004. Both KH and non-KH domain sequences are required for polyribosome association of Scp160p in yeast. Nucleic Acids Res. 32, 4768–4775. https://doi.org/10.1093/nar/gkh812

Li, A.-M., 2003. Scp160p associates with specific mRNAs in yeast. Nucleic Acids Res. 31, 1830–1837. https://doi.org/10.1093/nar/gkg284

Li, Y., Bachant, J., Alcasabas, A. A., Wang, Y., Qin, J., & Elledge, S. J. (2002). The mitotic spindle is required for loading of the DASH complex onto the kinetochore. Genes & development, 16(2), 183-197

Linder P, Lasko PF, Ashburner M, Leroy P, Nielsen PJ, Nishi K, Schnier J, Slonimski PP. Birth of the D-E-A-D box. Nature. 1989 Jan 12;337(6203):121-2. doi: 10.1038/337121a0. Erratum in: Nature 1989 Jul 20;340(6230):246. PMID: 2563148

Linder, P., Jankowsky, E., 2011. From unwinding to clamping — the DEAD box RNA helicase family. Nat. Rev. Mol. Cell Biol. 12, 505–516. https://doi.org/10.1038/nrm3154

Liu, Z., Luyten, I., Bottomley, M.J., Messias, A.C., Houngninou-Molango, S., Sprangers, R., Zanier, K., Kramer, A., Sattler, M., 2001. Structural Basis for Recognition of the Intron Branch Site RNA by Splicing Factor 1 294.

Lohe, A. R., Hilliker, A. J., & Roberts, P. A. (1993). Mapping simple repeated DNA sequences in heterochromatin of Drosophila melanogaster. Genetics, 134(4), 1149-1174.

Longen, S., Bien, M., Bihlmaier, K., Kloeppel, C., Kauff, F., Hammermeister, M., ... & Riemer, J. (2009). Systematic analysis of the twin Cx9C protein family. Journal of molecular biology, 393(2), 356-368.

Love, M.I., Huber, W., Anders, S., 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15, 550. https://doi.org/10.1186/s13059-014-0550-8

Lu, S.H.-J., Jeon, A.H.W., Schmitt-Ulms, G., Qamar, S., Dodd, R., McDonald, B., Li, Y., Meadows, W., Cox, K., Bohm, C., Chen, F., Fraser, P., George-Hyslop, P.S., 2012. Vigilin interacts with signal peptide peptidase. Proteome Sci. 10, 33. https://doi.org/10.1186/1477-5956-10-33

Luber CA, Cox J, Lauterbach H, Fancke B, Selbach, M., Tschopp, J., Akira, S., Wiegand, M., Hochrein, H., O'Keeffe, M., and Mann, M. **(**2010). Quantitative proteomics reveals subset-specific viral recognition in dendritic cells. Immunity 32, 279-89

Machanick, P., & Bailey, T. L. (2011). MEME-ChIP: motif analysis of large DNA datasets. *Bioinformatics*, *27*(12), 1696-1697.

Maniatis, T., & Reed, R. (1987). The role of small nuclear ribonucleoprotein particles in pre-mRNA splicing. Nature, 325(6106), 673-678.

Marsellach, F.-X., Huertas, D., Azorín, F., 2006. The Multi-KH Domain Protein of Saccharomyces cerevisiae Scp160p Contributes to the Regulation of Telomeric Silencing. J. Biol. Chem. 281, 18227–18235. https://doi.org/10.1074/jbc.M601671200

Marston, A.L., 2014. Chromosome Segregation in Budding Yeast: Sister Chromatid Cohesion and Related Mechanisms. Genetics 196, 31–63. https://doi.org/10.1534/genetics.112.145144

Masliah, Grégoire, Pierre Barraud, and Frédéric H -T. Allain. "RNA recognition by double-stranded RNA binding domains: a matter of shape and sequence." Cellular and molecular life sciences 70 (2013): 1875-1895

McKnightS, G.L., Reasoner, J., Gilbert, T., Champagne, J., Johnson, C.J., Bailey, M.C., Holly, R., O'HaraS, P.J., 1992. Cloning and Expression of a Cellular High Density Lipoprotein- binding Protein ThaIts Up-regulated by Cholesterol Loadingof Cells. Bind. Protein.

Mendelsohn, B.A., 2003. Genetic and biochemical interactions between SCP160 and EAP1 in yeast. Nucleic Acids Res. 31, 5838–5847. https://doi.org/10.1093/nar/gkg810

Michaelis, A. C., Brunner, A. D., Zwiebel, M., Meier, F., Strauss, M. T., Bludau, I., & Mann, M. (2023). The social and structural architecture of the yeast protein interactome. *Nature*, *624*(7990), 192-200

Minard, K. I., & McAlister-Henn, L. (1991). Isolation, nucleotide sequence analysis, and disruption of the MDH2 gene from Saccharomyces cerevisiae: evidence for three isozymes of yeast malate dehydrogenase. Molecular and cellular biology, 11(1), 370-380.

Mitchell, S.F., Jain, S., She, M., Parker, R., 2013. Global analysis of yeast mRNPs. Nat. Struct. Mol. Biol. 20, 127–133. https://doi.org/10.1038/nsmb.2468

Mobin, M.B., Gerstberger, S., Teupser, D., Campana, B., Charisse, K., Heim, M.H., Manoharan, M., Tuschl, T., Stoffel, M., 2016. The RNA-binding protein vigilin regulates VLDL secretion through modulation of Apob mRNA translation. Nat. Commun. 7, 12848. https://doi.org/10.1038/ncomms12848

Moreira, T.C.P., Da Silva, V.M., Gombert, A.K., Da Cunha, R.L., 2016. Stabilization mechanisms of oil-in-water emulsions by Saccharomyces cerevisiae. Colloids Surf. B Biointerfaces 143, 399–405. https://doi.org/10.1016/j.colsurfb.2016.03.043

Moriya, H., & Isono, K. (1999). Analysis of genetic interactions between DHH1, SSD1 and ELM1 indicates their involvement in cellular morphology determination in Saccharomyces cerevisiae. *Yeast*, *15*(6), 481-496.

Mutka, S.C., Walter, P., 2001. Multifaceted Physiological Response Allows Yeast to Adapt to the Loss of the Signal Recognition Particle- dependent Protein-targeting Pathway□D. Mol. Biol. Cell 12.

Nerome, S., Onishi, M., Saito, D., Mizobuchi, A., Ando, T., Daira, Y., Matsumoto, A., Ojima, Y., Azuma, M., 2020. Cell surface changes that advance the application of using yeast as a food emulsifier. Food Chem. 315, 126264. https://doi.org/10.1016/j.foodchem.2020.126264

Neuhof, A., Rolls, M.M., Jungnickel, B., Kalies, K.-U., Rapoport, T.A., 1998. Binding of Signal Recognition Particle Gives Ribosome/Nascent Chain Complexes a Competitive Advantage in Endoplasmic Reticulum Membrane Interaction. Mol. Biol. Cell 9, 103–115. https://doi.org/10.1091/mbc.9.1.103

Novačić, A., Šupljika, N., Bekavac, N., Žunar, B., & Stuparević, I. (2021). Interplay of the RNA Exosome Complex and RNA-Binding Protein Ssd1 in Maintaining Cell Wall Stability in Yeast. *Microbiology spectrum*, *9*(1), 10-1128.

Oeffinger, M., Wei, K.E., Rogers, R., DeGrasse, J.A., Chait, B.T., Aitchison, J.D., Rout, M.P., 2007. Comprehensive analysis of diverse ribonucleoprotein complexes. Nat. Methods 4, 951–956. https://doi.org/10.1038/nmeth1101

Olivier, C., Poirier, G., Gendron, P., Boisgontier, A., Major, F., Chartrand, P., 2005. Identification of a Conserved RNA Motif Essential for She2p Recognition and mRNA Localization to the Yeast Bud. Mol. Cell. Biol. 25, 4752–4766. https://doi.org/10.1128/MCB.25.11.4752-4766.2005

Opitz, N., Schmitt, K., Hofer-Pretz, V., Neumann, B., Krebber, H., Braus, G. H., & Valerius, O. (2017). Capturing the Asc1p/receptor for activated C kinase 1 (RACK1) microenvironment at the head region of the 40S ribosome with quantitative BioID in yeast. *Molecular & Cellular Proteomics*, *16*(12), 2199-221

Oughtred, R., Rust, J., Chang, C., Breitkreutz, B., Stark, C., Willems, A., Boucher, L., Leung, G., Kolas, N., Zhang, F., Dolma, S., Coulombe-Huntington, J., Chatr-aryamontri, A., Dolinski, K., Tyers, M., 2021. The BIOGRID database: A comprehensive biomedical resource of curated protein, genetic, and chemical interactions. Protein Sci. 30, 187–200. https://doi.org/10.1002/pro.3978

Palazzo, A.F., Springer, M., Shibata, Y., Lee, C.-S., Dias, A.P., Rapoport, T.A., 2007. The Signal Sequence Coding Region Promotes Nuclear Export of mRNA. PLoS Biol. 5, e322. https://doi.org/10.1371/journal.pbio.0050322

Paz, I., Choder, M., 2001. Eukaryotic Translation Initiation Factor 4E-Dependent Translation Is Not Essential for Survival of Starved Yeast Cells. J. Bacteriol. 183, 4477–4483. https://doi.org/10.1128/JB.183.15.4477-4483.2001

Paziewska, A., Wyrwicz, L.S., Bujnicki, J.M., Bomsztyk, K., Ostrowski, J., 2004. Cooperative binding of the hnRNP K three KH domains to mRNA targets. FEBS Lett. 577, 134–140. https://doi.org/10.1016/j.febslet.2004.08.086

Pfeffer, S., Brandt, F., Hrabe, T., Lang, S., Eibauer, M., Zimmermann, R., Förster, F., 2012. Structure and 3D Arrangement of Endoplasmic Reticulum Membrane-Associated Ribosomes. Structure 20, 1508–1518. https://doi.org/10.1016/j.str.2012.06.010

Rapoport, T.A., 2007. Protein translocation across the eukaryotic endoplasmic reticulum and bacterial plasma membranes. Nature 450, 663–669. https://doi.org/10.1038/nature06384

Rappsilber J, Mann M, Ishihama Y (2007). Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat. Protoc*. 2: 1896–1906

Reid, D.W., Nicchitta, C.V., 2015. Diversity and selectivity in mRNA translation on the endoplasmic reticulum. Nat. Rev. Mol. Cell Biol. 16, 221–231. https://doi.org/10.1038/nrm3958

Reid, D.W., Nicchitta, C.V., 2012. Primary Role for Endoplasmic Reticulum-bound Ribosomes in Cellular Translation Identified by Ribosome Profiling. J. Biol. Chem. 287, 5518–5527. https://doi.org/10.1074/jbc.M111.312280

Reinke, A., Anderson, S., McCaffery, J. M., Yates, J., Aronova, S., Chu, S., ... & Powers, T. (2004). TOR complex 1 includes a novel component, Tco89p (YPL180w), and cooperates with Ssd1p to maintain cellular integrity in Saccharomyces cerevisiae. *Journal of Biological Chemistry*, *279*(15), 14752-14762.

Rössler, I., Embacher, J., Pillet, B., Murat, G., Liesinger, L., Hafner, J., Unterluggauer, J.J., Birner-Gruenberger, R., Kressler, D., Pertschy, B., 2019. Tsr4 and Nap1, two novel members of the ribosomal protein chaperOME. Nucleic Acids Res. 47, 6984–7002. https://doi.org/10.1093/nar/gkz317

Ruepp, A., 2004. The FunCat, a functional annotation scheme for systematic classification of proteins from whole genomes. Nucleic Acids Res. 32, 5539–5545. https://doi.org/10.1093/nar/gkh894

Ryan, O. W., Skerker, J. M., Maurer, M. J., Li, X., Tsai, J. C., Poddar, S., ... & Cate, J. H. (2014). Selection of chromosomal DNA libraries using a multiplex CRISPR system. *elife*, *3*, e03703

Sambrook, J., Fritsch, E. F., & Maniatis, T. (1989). Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Pressn New York.

Schmidt, C., Henkel, B., Pöschl, E., Zorbas, H., Purschke, W.G., Gloe, T.R., Müller, P.K., 1992. Complete cDNA sequence of chicken vigilin, a novel protein with amplified and evolutionary conserved domains. Eur. J. Biochem. 206, 625–634. https://doi.org/10.1111/j.1432-1033.1992.tb16967.x

Schmitt, K., Kraft, A. A., & Valerius, O. (2021). A Multi-Perspective Proximity View on the Dynamic Head Region of the Ribosomal 40S Subunit. *International Journal of Molecular Sciences*, *22*(21), 11653.

Schneider, T., Hung, L. H., Aziz, M., Wilmen, A., Thaum, S., Wagner, J., ... & Bindereif, A. (2019). Combinatorial recognition of clustered RNA elements by the multidomain RNA-binding protein IMP3. *Nature Communications*, *10*(1), 2266.

Schreck, H. (2010). Translational control by the multi-KH domain protein Scp160 (Doctoral dissertation, lmu)

Schuller, A.P., Green, R., 2018. Roadblocks and resolutions in eukaryotic translation. Nat. Rev. Mol. Cell Biol. 19, 526–541. https://doi.org/10.1038/s41580-018-0011-4

Schwanhäusser B, Busse D, Li N, Dittmar G, Schuchhardt J, Wolf J, Chen W, Selbach M (2011). Global quantification of mammalian gene expression control. *Nature*, 473(7347): 337

Semanjski M, Germain E, Bratl K, Kiessling A, Gerdes K, Macek B. The kinases HipA and HipA7 phosphorylate different substrate pools in *Escherichia coli* to promote multidrug tolerance. Sci Signal. 2018 Sep 11.

Sezen, B., Seedorf, M., Schiebel, E., 2009. The SESA network links duplication of the yeast centrosome with the protein translation machinery. Genes Dev. 23, 1559–1570. https://doi.org/10.1101/gad.524209

Shannon, K.W., Rabinowitz, J.C., 1988. Isolation and characterization of the Saccharomyces cerevisiae MIS1 gene encoding mitochondrial C1-tetrahydrofolate synthase. J. Biol. Chem. 263, 7717–7725. https://doi.org/10.1016/S0021-9258(18)68558-8

Siegel, M.R., Sisler, H.D., 1963. Inhibition of Protein Synthesis in vitro by Cycloheximide. Nature 200, 675–676. https://doi.org/10.1038/200675a0

Silva, A., Almeida, B., Sampaio-Marques, B., Reis, M.I.R., Ohlmeier, S., Rodrigues, F., Vale, A.D., Ludovico, P., 2011. Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) is a specific substrate of yeast metacaspase. Biochim. Biophys. Acta BBA - Mol. Cell Res. 1813, 2044–2049. https://doi.org/10.1016/j.bbamcr.2011.09.010

Siomi, H., Matunis, M.J., Michael, W.M., Dreyfuss, G., 1993. The pre-mRNA binding K protein contains a novel evolutionary conserved motif. Nucleic Acids Res. 21, 1193–1198. https://doi.org/10.1093/nar/21.5.1193

Smirnova, J.B., Selley, J.N., Sanchez-Cabo, F., Carroll, K., Eddy, A.A., McCarthy, J.E.G., Hubbard, S.J., Pavitt, G.D., Grant, C.M., Ashe, M.P., 2005. Global Gene Expression Profiling Reveals Widespread yet Distinctive Translational Responses to Different Eukaryotic Translation Initiation Factor 2B-Targeting Stress Pathways. Mol. Cell. Biol. 25, 9340–9349. https://doi.org/10.1128/MCB.25.21.9340-9349.2005

Sonenberg, N., Hinnebusch, A.G., 2009. Regulation of Translation Initiation in Eukaryotes: Mechanisms and Biological Targets. Cell 136, 731–745. https://doi.org/10.1016/j.cell.2009.01.042

Svitkin, Y.V., Herdy, B., Costa-Mattioli, M., Gingras, A.-C., Raught, B., Sonenberg, N., 2005. Eukaryotic Translation Initiation Factor 4E Availability Controls the Switch between Cap-Dependent andInternal Ribosomal Entry Site-Mediated Translation. Mol. Cell. Biol. 25, 10556–10565. https://doi.org/10.1128/MCB.25.23.10556-10565.2005

Teixeira, D., Sheth, U., Valencia-Sanchez, M.A., Brengues, M., Parker, R., 2005. Processing bodies require RNA for assembly and contain nontranslating mRNAs. RNA 11, 371–382. https://doi.org/10.1261/rna.7258505

Thomas, J. H., & Botstein, D. (1986). A gene required for the separation of chromosomes on the spindle apparatus in yeast. Cell, 44(1), 65-76.

Tibayrenc, P., Preziosi-Belloy, L., Roger, J.-M., Ghommidh, C., 2010. Assessing yeast viability from cell size measurements? J. Biotechnol. 149, 74–80. https://doi.org/10.1016/j.jbiotec.2010.06.019

Tung, S., Bakerlee, C. W., Phillips, A. M., Nguyen Ba, A. N., & Desai, M. M. (2021). The genetic basis of differential autodiploidization in evolving yeast populations. *G3*, *11*(8), jkab192.

Valverde, R., Edwards, L., Regan, L., 2008. Structure and function of KH domains. FEBS J. 275, 2712–2726. https://doi.org/10.1111/j.1742-4658.2008.06411.x

Viotti, C. (2016). Unconventional protein secretion, methods and protocols. Methods Mol Biol, 1459, 3-29.

Voigt, F., Zhang, H., Cui, X.A., Triebold, D., Liu, A.X., Eglinger, J., Lee, E.S., Chao, J.A., Palazzo, A.F., 2017. Single-Molecule Quantification of Translation-Dependent Association of mRNAs with the Endoplasmic Reticulum. Cell Rep. 21, 3740–3753. https://doi.org/10.1016/j.celrep.2017.12.008

Vollbrandt, T., Willkomm, D., Stossberg, H., Kruse, C., 2004. Vigilin is co-localized with 80S ribosomes and binds to the ribosomal complex through its C-terminal domain. Int. J. Biochem. Cell Biol. 36, 1306–1318. https://doi.org/10.1016/j.biocel.2003.11.006

Walter, P., Blobel, G., 1981. Translocation of proteins across the endoplasmic reticulum. II. Signal recognition protein (SRP) mediates the selective binding to microsomal membranes of in-vitro-assembled polysomes synthesizing secretory protein. J. Cell Biol. 91, 551–556. https://doi.org/10.1083/jcb.91.2.551

Wang, Q., Zhang, Z., Blackwell, K., Carmichael, G.G., 2005. Vigilins Bind to Promiscuously A-to-I-Edited RNAs and Are Involved in the Formation of Heterochromatin. Curr. Biol. 15, 384–391. https://doi.org/10.1016/j.cub.2005.01.046

Warner, J.R., 1999. The economics of ribosome biosynthesis in yeast. Trends Biochem. Sci. 24, 437–440. https://doi.org/10.1016/S0968-0004(99)01460-7

Weber, V., Wernitznig, A., Hager, G., Harata, M., Frank, P., Wintersberger, U., 1997. Purification and Nucleic-Acid-Binding Properties of a Saccharomyces Cerevisiae Protein Involved in the Control of Ploidy. Eur. J. Biochem. 249, 309–317. https://doi.org/10.1111/j.1432-1033.1997.00309.x

Weidner, J., Wang, C., Prescianotto-Baschong, C., Estrada, A.F., Spang, A., 2014. The polysome-associated proteins Scp160 and Bfr1 prevent P body formation under normal growth conditions. J. Cell Sci. jcs.142083. https://doi.org/10.1242/jcs.142083

Weir, J.R., Bonneau, F., Hentschel, J., Conti, E., 2010. Structural analysis reveals the characteristic features of Mtr4, a DExH helicase involved in nuclear RNA processing and surveillance. Proc. Natl. Acad. Sci. 107, 12139–12144. https://doi.org/10.1073/pnas.1004953107

Wen, W.-L., Stevenson, A.L., Wang, C.-Y., Chen, H.-J., Kearsey, S.E., Norbury, C.J., Watt, S., Bahler, J., Wang, S.-W., 2010. Vgl1, a multi-KH domain protein, is a novel component of the fission yeast stress granules required for cell survival under thermal stress. Nucleic Acids Res. 38, 6555–6566. https://doi.org/10.1093/nar/gkq555

Westermann, S., Avila-Sakar, A., Wang, H.-W., Niederstrasser, H., Wong, J., Drubin, D.G., Nogales, E., Barnes, G., 2005. Formation of a Dynamic Kinetochore- Microtubule Interface through Assembly of the Dam1 Ring Complex. Mol. Cell 17, 277–290. https://doi.org/10.1016/j.molcel.2004.12.019

Wintersberger, U., Kühne, C., Karwan, A., 1995. Scp160p, a new yeast protein associated with the nuclear membrane and the endoplasmic reticulum, is necessary for maintenance of exact ploidy. Yeast 11, 929–944. https://doi.org/10.1002/yea.320111004

Woo, H.-H., Yi, X., Lamb, T., Menzl, I., Baker, T., Shapiro, D.J., Chambers, S.K., 2011. Posttranscriptional Suppression of Proto-Oncogene c -fms Expression by Vigilin in Breast Cancer. Mol. Cell. Biol. 31, 215–225. https://doi.org/10.1128/MCB.01031-10

Xia, Y. R., Klisak, I., Sparkes, R. S., Oram, J., & Lusis, A. J. (1993). Localization of the gene for high-density lipoprotein binding protein (HDLBP) to human chromosome 2q37. Genomics, 16(2), 524-525.

YaDeau, J.T., Klein, C., Blobel, G., 1991. Yeast signal peptidase contains a glycoprotein and the Sec11 gene product. Proc. Natl. Acad. Sci. 88, 517–521. https://doi.org/10.1073/pnas.88.2.517

Young, B.P., 2001. Sec63p and Kar2p are required for the translocation of SRP-dependent precursors into the yeast endoplasmic reticulum in vivo. EMBO J. 20, 262–271. https://doi.org/10.1093/emboj/20.1.262

Zhang, X., Shan, S., 2014. Fidelity of Cotranslational Protein Targeting by the Signal Recognition Particle. Annu. Rev. Biophys. 43, 381–408. https://doi.org/10.1146/annurev-biophys-051013-022653

Zhou, J., Wang, Q., Chen, L.-L., Carmichael, G.G., 2008. On the mechanism of induction of heterochromatin by the RNA-binding protein vigilin. RNA 14, 1773–1781. https://doi.org/10.1261/rna.1036308

Zinnall, U., Milek, M., Minia, I., Vieira-Vieira, C.H., Müller, S., Mastrobuoni, G., Hazapis, O.-G., Del Giudice, S., Schwefel, D., Bley, N., Voigt, F., Chao, J.A., Kempa, S., Hüttelmaier, S., Selbach, M., Landthaler, M., 2022. HDLBP binds ER-targeted mRNAs by multivalent interactions to promote protein synthesis of transmembrane and secreted proteins. Nat. Commun. 13, 2727. https://doi.org/10.1038/s41467-022-30322-7

Acknowledgements

Ich möchte mich bei allen bedanken, denen ich diese Arbeit zu verdanken habe:

Meinem Professor:

Ralf, danke, dass du einen Platz in deiner Arbeitsgruppe für mich gefunden hast, auch als ich nicht auf die eigentlich beworbene Stelle gepasst habe. Ich bin dankbar für die vielen Freiheiten, die du mir gelassen hast, aber gleichwohl auch dafür, dass deine Türe immer offen war, um die Ergebnisse und Probleme zu besprechen (Verhältnis eher 10:90). Auch wenn wir im Projekt und in der Arbeitsgruppe so manche schwierige Zeit hatten, werde ich mich doch immer mit einem Lächeln zurückerinnern.

Meiner Arbeitsgruppe und akademischen Helfern:

Ich hätte es allein nicht geschafft diese Arbeit zu einem Abschluss zu bringen. Ich danke der ganzen AG Jansen für ihre Unterstützung und ihren Input. Mein besonderer Dank gilt Ronja Weissinger, Iliana Nikolou und Orit Hermesh-Raue, die mir nicht nur akademische Einblicke gewährt haben, sondern auch ihre Freundschaft. Selbstverständlich wäre ich genauso aufgeschmissen gewesen ohne den Rat und Tat unserer tatkräftigen TAs: Ingrid, Ulrike und Ruth, danke für eure Hilfe. Zuletzt danke ich auch meinen Studenten, die mich stets eifrig unterstützt haben.

Meinen Kolaborationspartnern: ich danke Dr. Mirita Franz und Dr. Francesca Barletta für ihre Hilfe zu den Themen Massenspektrometrie und Datenanalyse. Dr. Kathrine Bohnsack und Dr. Nicolas Lemus danke ich für ihre Hilfe bei meinen RNA-experimenten.

Zudem bedanke ich mich bei meinem zweiten Betreuer Prof. Dr. Thorsten Stafforst und meinem Prüfungskomitee, ohne die ich die finalen Meter nicht beschreiten könnte.

Meiner Familie:

Dafür, dass ihr immer hinter mir steht. Frederic, Sinje und Julia und Klaus: ich danke euch dafür, dass ihr immer für mich da seid.

Meinen Freunden:

Einen besonderen Dank widme ich den wohl besten Menschen, die mir je begegnet sind und denen kein Wort, welches ich hier anbringen könnte, gerecht werden kann: Olli, Katja, Hannes, Meli und Tobi. Ohne euch wäre meine Leben nicht das gleiche.

Meiner Sonne:

Das Wertvollste fällt einem oft in die Hände, wenn man aufhört danach zu suchen. Ich hätte niemals erwartet, dass das Wichtigste, was ich aus meiner Zeit als Doktorand mitnehme, nicht mein Doktortitel ist, sondern die Sonne meines Lebens: Lisa, ich liebe dich über Alles und ich kann nicht in Worte fassen, was du mir bedeutest. Danke, dass du dich für mich entschieden hast.

Ich widme diese Doktorarbeit meinen Eltern: Gabriele und Leo Feicht, ohne die ich nicht an diesem Punkt stehen würde. Ihr seid die besten Eltern, die man sich wünschen kann und ich verdanke euch alles.
Ich liebe Euch.