Deciphering the MademoiseLLE-binding code in endosomal mRNA transport

Inaugural dissertation

for the attainment of the title of doctor
in the Faculty of Mathematics and Natural Sciences
at the Heinrich Heine University Düsseldorf

presented by

Senthil Kumar Devan

from Devakottai, Tamil Nadu, India

Heinrich Heine Universität Düsseldorf

Düsseldorf, December 2022

From the Institute of Microbiology
at the Heinrich-Heine-University Düsseldorf

Published with the permission of the
Faculty of Mathematics and Natural Sciences at
Heinrich Heine University Düsseldorf

Supervisor: Prof. Dr. Michael Feldbrügge

Co-Supervisor: Prof. Dr. Holger Gohlke

Date of the oral examination: 17.03.2023

Statutory declaration

This dissertation is the result of my research, considering the "Rules and principles for safeguarding and good scientific practice at the Heinrich Heine University Düsseldorf". No other person's work has been used without due acknowledgement. I did not use literature other than the ones cited in the text and listed in the reference section. This dissertation has not been submitted or published in the same or similar form elsewhere. I have not failed a doctoral examination procedure previously.

_____

City: Düsseldorf

Date: 15.12.2022

Senthil Kumar Devan

Matriculation number: 2566212

The research detailed in this dissertation was conducted from August 2016 until May 2022 at the Heinrich Heine University Düsseldorf in the Institute for Microbiology under Prof. Dr. Michael Feldbrügge.

To Keerthi

Summary

mRNA transport and local translation ensure the spatiotemporal regulation of gene expression. Endosome-mediated mRNA transport is a conserved biological process from fungal hyphae to plant endosperms and neuronal cells. *Ustilago maydis*, a plant pathogen and filamentous fungus, is the best-studied model organism for endosome-mediated mRNA transport. Numerous components of endosomal mRNA transport are already characterized in *U. maydis*. Previous studies have shown that the key RNA-binding protein, Rrm4, carried out the mRNA transport in the fungal hyphae. Rrm4, along with the accessory mRNA poly(A)-binding protein, Pab1, interacts with the endosomal adaptor protein Upa1. Rrm4 and Pab1 have MademoiseLLE (MLLE) domains in their C-terminus, which are crucial for endosomal attachment. Loss of the C-terminus of Rrm4 impairs the endosomal attachment and results in severe growth defects similar to the loss of full-length protein. MLLE domains generally consist of five helices and facilitate protein-protein interaction by forming a peptide binding pocket for binding partners containing the poly(A)-binding protein associated motif (PAM2).

Interestingly, MLLE of Rrm4 recognized not PAM2 but a PAM2-like motif in Upa1 ($PAM2L^{Upa1}$). The mechanistic understanding of how Rrm4 and Pab1 are connected to the endosomes through their MLLE domain was not clear at the beginning of this study. This dissertation addresses this question using the structural biology, biochemical, biophysical, and fungal genetic approaches.

In the first part, this study established that Rrm4 contains not one but three MLLE domains at the C-terminus, forming a protein-protein interaction platform. Importantly, they function with a strict hierarchy. The third MLLE domain ($MLLE3^{Rrm4}$) is essential, and the other two domains ($MLLE1,2^{Rrm4}$) play accessory roles. Next, this study provided evidence that the $MLLE3^{Rrm4}$ consisted of not five helices but seven helices. The additional two helices are essential for ligand recognition. $MLLE3^{Rrm4}$ recognizes the $PAM2L^{Upa1}$ by non-canonical interactions. Finally, by deciphering the $MLLE3^{Rrm4}$-binding code, several new PAM2L-containing proteins were identified by *de novo* prediction.

One of them is an endosomal CORVET complex subunit Vps8 ($PAM2L^{Vps8}$). *In vitro* experiments verified the interaction between $MLLE3^{Rrm4}$ and $PAM2L^{Vps8}$. Therefore, Vps8 is postulated as a secondary tethering molecule of Rrm4-associated mRNPs. Based on the above results, the model of the endosomal mRNA transport unit was advanced. Thus, this study provided the most detailed mechanistic description of how an mRNA-binding protein and its bound cargo mRNAs are attached to endosomes to date.

Zusammenfassung

Der mRNA-Transport und die lokale Translation gewährleisten die räumlich-zeitliche Regulierung der Genexpression. Der Endosomen-vermittelte mRNA-Transport ist ein konservierter biologischer Prozess, der von Pilzhyphen bis zu Pflanzenendospermien und neuronalen Zellen reicht. *Ustilago maydis*, ein Pflanzenpathogen und filamentös wachsender Pilz, ist der am besten untersuchte Modellorganismus für den Endosomen-vermittelten mRNA-Transport. Zahlreiche Komponenten des endosomalen mRNA-Transports sind in *U. maydis* bereits identifiziert. Führe Untersuchen habe gezeigt, dass das zentrale RNA-bindende Protein Rrm4 für den mRNA-Transport in der Pilzhyphen verantwortlich ist. Rrm4 interagiert, zusammen mit dem akzessorischen mRNA-Poly(A)-bindenden Protein Pab1, mit dem endosomalen Adaptorprotein Upa1. Rrm4 und Pab1 haben beide MademoiseLLE (MLLE)-Domänen in ihrem C-Terminus, die für die endosomale Bindung entscheidend sind. Der Verlust des C-Terminus von Rrm4 beeinträchtigt die endosomale Bindung und führt zu Wachstumsdefekten, ähnlich wie der Verlust des Volllängenproteins. MLLE-Domänen bestehen im Allgemeinen aus fünf Helices und erleichtern die Protein-Protein-Interaktion, indem sie eine Peptid-Bindungstasche für Bindungspartner bilden, die das Poly(A)-bindende Protein-assoziierte Motiv (PAM2) enthalten.

Interessanterweise hat die MLLE von Rrm4 nicht das PAM2 Motiv erkannt, sondern ein PAM2-ähnliches Motiv in Upa1 ($PAM2L^{Upa1}$). Das mechanistische Verständnis, wie Rrm4 und Pab1 über ihre MLLE-Domäne mit den Endosomen verbunden sind, war zu Beginn dieser Studie nicht klar. Die vorliegende Dissertation befasst sich mit diesen Fragen unter Verwendung der Strukturbiologie, biochemischer, biophysikalischer und pilzgenetischer Ansätze.

Im ersten Teil dieser Studie wurde festgestellt, dass Rrm4 nicht nur eine, sondern drei MLLE-Domänen am C-Terminus enthält, die eine Protein-Protein-Interaktionsplattform bilden. Wichtig ist, dass sie in einer strengen Hierarchie funktionieren. Die dritte MLLE-Domäne ($MLLE3^{Rrm4}$) ist essenziell, während die beiden anderen Domänen ($MLLE1,2^{Rrm4}$) akzessorische Funktionen haben. Die Studie hat außerdem gezeigt, dass $MLLE3^{Rrm4}$ nicht aus fünf, sondern aus sieben Helices besteht. Die zusätzlichen zwei Helices sind für die Ligandenerkennung unerlässlich. $MLLE3^{Rrm4}$ erkennt das $PAM2L^{Upa1}$ durch nicht-kanonische Wechselwirkungen. Durch die Entschlüsselung des $MLLE3^{Rrm4}$-Bindungscodes konnten mehrere neue PAM2L-enthaltende Proteine anhand von *de novo*-Vorhersagen identifiziert werden.

Eines davon ist die endosomale CORVET-Komplex-Untereinheit Vps8 ($PAM2L^{Vps8}$). *In vitro*-Experimente haben die Interaktion zwischen $MLLE3^{Rrm4}$ und $PAM2L^{Vps8}$ bestätigt. Daher wird Vps8 als sekundäres Verbindungsprotein von Rrm4-assoziierten mRNPs postuliert. Auf der Grundlage der obigen Ergebnisse wird das Modell der endosomalen mRNA-Transporteinheit weiterentwickelt. Damit liefert diese Arbeit die derzeit detaillierteste mechanische Beschreibung darüber, wie ein RNA-bindendes Protein und dessen Ziel-mRNAs an Endosomen gebunden sind

List of Abbreviations

Abbreviations	Meaning
°C	Degree centigrade
3'	three prime end
5'	five prime end
Å	Angstrom
α	Anti
αI,II	Alpha helices I and II of MLLE3 domain of Rrm4
β-ME	Beta mercapto ethanol
µM	Micrometer
µL	Microliter
AA	Amino acid
AD	Activation domain
AF2	AlphaFold2
Amp	Ampicillin
ANK	Ankyrin repeats
APS	Ammonium persulfate
BD	Binding domain
bE/bW	Transcription factor bEast (bE) and bWest (bW; in analogy with the re-union of Berlin)
BLAST	Basic Local Alignment Search Tool
bp	Base pairs
CBB	Coomassie Brilliant Blue
CM	Complete medium
CORVET	class C core vacuole/endosome tethering
CTD	C-terminal Annexin-repeat domain of ANXA11
C-terminal	Carboxyl-terminal
ddH2O	De-mineralized water
DIC	Differential interference contrast
DNA	Desoxy ribonucleic acid
EDD	E3 ubiquitin ligase identified by differential display
EDTA	Ethylene diamine tetra acetic acid
EE	Early endosome
EfTu	Elongation Factor-Tu
ER	Endoplasmic reticulum
et al.	And others
FERRY	Five-subunit Endosomal Rab5 and RNA/ribosome intermediarY
FYVE	Zinc finger domain found in Fab 1, YOTB, Vac 1, and EEA1
gDNA	Genomic DNA
GDP	Guanosine diphosphate
Gfp	Green fluorescence protein
Glc	Glucose
GST	Glutathione S Transferase
GTP	Guanosine triphosphate
GTPase	Enzymes that hydrolyze the nucleotide GTP to GDP
GWW	Conserved GWW domain at the C-terminus of Upa2
h	Hours
h.p.i	Hours post induction
HA	Hemagglutinin
HECT	Homologous to the E6-AP Carboxyl Terminus, an Ubiquitin-protein ligases
HF	High Fidelity

HOPS	Homotypic fusion and vacuole protein sorting, a tethering complex on late endosome via Rab7 GTPase
HRP	Horseradish peroxidase
HS	His-Sumo tag
Hyg	Hygromycin
iCLIP	Individual-nucleotide resolution UV crosslinking and immunoprecipitation
IDR	Intrinsically disordered region
IMAC	Immobilized metal ion chromatography
IMM	Inner mitochondrial membrane
in silico	By means of computer modelling or computer simulation
in situ	Examination of a phenomenon exactly in the place where it occurs
in vitro	Independent of a living organism
in vivo	In a living organism
IPTG	Isopropyl β-D-1-thiogalactopyranoside
ITC	Isothermal titration calorimetry
Kat	mKate2 - a monomeric, basic red fluorescent protein
kb	Kilo base
K_D	Equilibrium dissociation constant
kDa	Kilo Dalton
L	Liter
LC-MS/MS	Liquid Chromatography with tandem mass spectrometry
LCR	Low complexity region
LE	Late endosome
M	mol/L
MALS	Multi angle light scattering
MLLE	MademoiseLLE
$MLLE^{PAB1}$	MademoiseLLE domain of Pab1
$MLLE^{PABPC1}$	MademoiseLLE domain of human PABPC1
$MLLE1^{RRM4}$	First MademoiseLLE domain of Rrm4
$MLLE2^{RRM4}$	Second MademoiseLLE domain of Rrm4
$MLLE3^{RRM4}$	Third MademoiseLLE domain of Rrm4
$MLLE^{UBR5}$	MademoiseLLE domain of human Ubr5
mRNA	Messenger ribonucleic acid
mRNP	Messenger ribonucleotide particle
MS	Mass spectrometry
MSA	Multiple Sequence Alignment
MT	Microtubules
MTS	Mitochondrial Targeting Sequence
NCBI	National Centre for Biotechnological Information
ng	Nano gram
Ni-NTA	Nickel-Nitrilo Triacetic Acid
NLS	Nuclear localization signal
NM	Nitrate medium
nm	Nanometer
nt	Nucleotides
NT	N-terminal truncation
N-terminal	Amino terminal
OD	Optical density
OMM	Outer mitochondrial membrane
ORF	Open reading frame
Pab1	Poly(A)-binding protein in Fungi
PABPC1	Poly(A)-binding protein in human

PAM2	PABC interacting motif
$PAM2^{Upa1}$	PABC interacting motif of Upa1
PAM2L	PABC interacting motif-like region
$PAM2L^{EfTu}$	PABC interacting motif-like region of Elongation Factor Tu
$PAM2L^{Taf7}$	PABC interacting motif-like region of Taf7
$PAM2L1^{Upa1}$	First PABC interacting motif-like region of Upa1
$PAM2L2^{Upa2}$	Second PABC interacting motif-like region of Upa1
$PAM2L1^{Vps8}$	First PABC interacting motif-like region of Vps8
$PAM2L2^{Vps8}$	Second PABC interacting motif-like region of Vps8
PCR	Polymerase chain reaction
PDB	Protein Data Bank
PEG	Polyethylene glycol
PI3P	Phosphatidylinositol (3,4,5)-trisphosphate
PPI	Protein-protein interaction
PTM	Post translational modification
RBP	RNA-binding protein
Rfp	Red fluorescence protein
RING	Really interesting new gene Zn finger domain
RMSD	Root-mean-square deviation of atomic positions
RNA	Ribonucleic acid
RNA-FISH	RNA fluorescence in situ hybridization
RNA-seq	RNA sequencing
rpm	Rounds per minute
RRM	RNA recognition motif
RT	Room temperature
s	Seconds
SAXS	Small angle X-ray scattering
SD	Synthetic Dropout plates
SDS - PAGE	Sodium Dodecyl Sulfate - Polyacrylamide Gel Electrophoresis
SEC	Size Exclusion chromatography
SS	Strep-Sumo tag
SUMO	Small ubiquitin modifying protein
TAF	TBP associated factor
TBP	TATA-binding protein
TBS	Tris-Buffered Saline
TE	Tris-EDTA
TEMED	Tetramethylethylenediamin
T_{nos}	Terminator of the nopaline synthase gene of *Agrobacterium tumefaciens*
tRNA	Transfer RNA
Ub	Ubiquitin
UMAG	*Ustilago maydis* gene
UPS	Ubiquitin-proteasome system
UTR	Untranslated region
UV	Ultraviolet
ZBP	Zipcode binding protein

Table of contents

1 Introduction

1.1 mRNA transport and local translation

In the central dogma of life, messenger RNA (mRNA) conduits the flow of genetic information from DNA to protein. In eukaryotic cells, precursor mRNAs are synthesized in the nucleus, followed by 5' capping, splicing, and 3'end cleavage and polyadenylation. Matured mRNAs are exported to the cytoplasm, where ribosomes translate them and, finally, decayed. However, this textbook description is too simplistic to understand their intricate journey into the cytoplasm.

Emerging evidence from multiple model organisms has established that most mRNAs and the associated ribosomes are transported to a subcellular location for Spatiotemporal gene expression. Localized mRNAs can be translated into thousands of copies upon a local stimulus, which makes the protein synthesis process more cost-effective than transporting individual proteins. Localized protein production may avoid the expression of proteins in the undesired cellular compartments. Protein synthesis at specific compartments may support their proper folding due to their local pH and concentration, controls their function, protein-protein interaction (PPI), and multiprotein complex formation (GLOCK *et al.* 2017; ENGEL *et al.* 2020; DAS *et al.* 2021).

mRNA transport is a widespread phenomenon, which occurs in single cell microorganisms such as bacteria, fungi and also in even more complex systems like plants, insects, and animals (JUNG *et al.* 2014; DAS *et al.* 2021; VARGAS *et al.* 2022). Subcellular mRNA transport was first observed in chicken fibroblasts using *in situ* hybridization (LAWRENCE AND SINGER 1986). This observation, along with studies in *Xenopus laevis eggs* (*X. laevis*)*, Drosophila melanogaster* (*D. melanogaster*) *embryo,* and *Saccharomyces cerevisiae* (*S. cerevisiae*), laid the foundation for the concept of mRNA transport, and local translation (REBAGLIATI *et al.* 1985; BERLETH *et al.* 1988; LONG *et al.* 1997). Originally it was thought that this process affects only a tiny species of mRNA population or is limited only to specific cell types. However, recent studies have shown that a large population of mRNAs and the translation machinery in multiple cell types are transported for local translation in specific compartments. In addition, the localized pool of translating mRNAs maintains the local protein homeostasis.

Especially highly polarized cells such as neurons, *Drosophila* embryos and fungal hyphae depend on long-distance mRNA transport as their local translation regions are

drastically separated from the nucleus and cytoplasm. Defects in this process affect fundamental processes like cell polarity, migration, and differentiation (GLOCK *et al.* 2017; NIESSING *et al.* 2018; ENGEL *et al.* 2020; TURNER-BRIDGER *et al.* 2020).

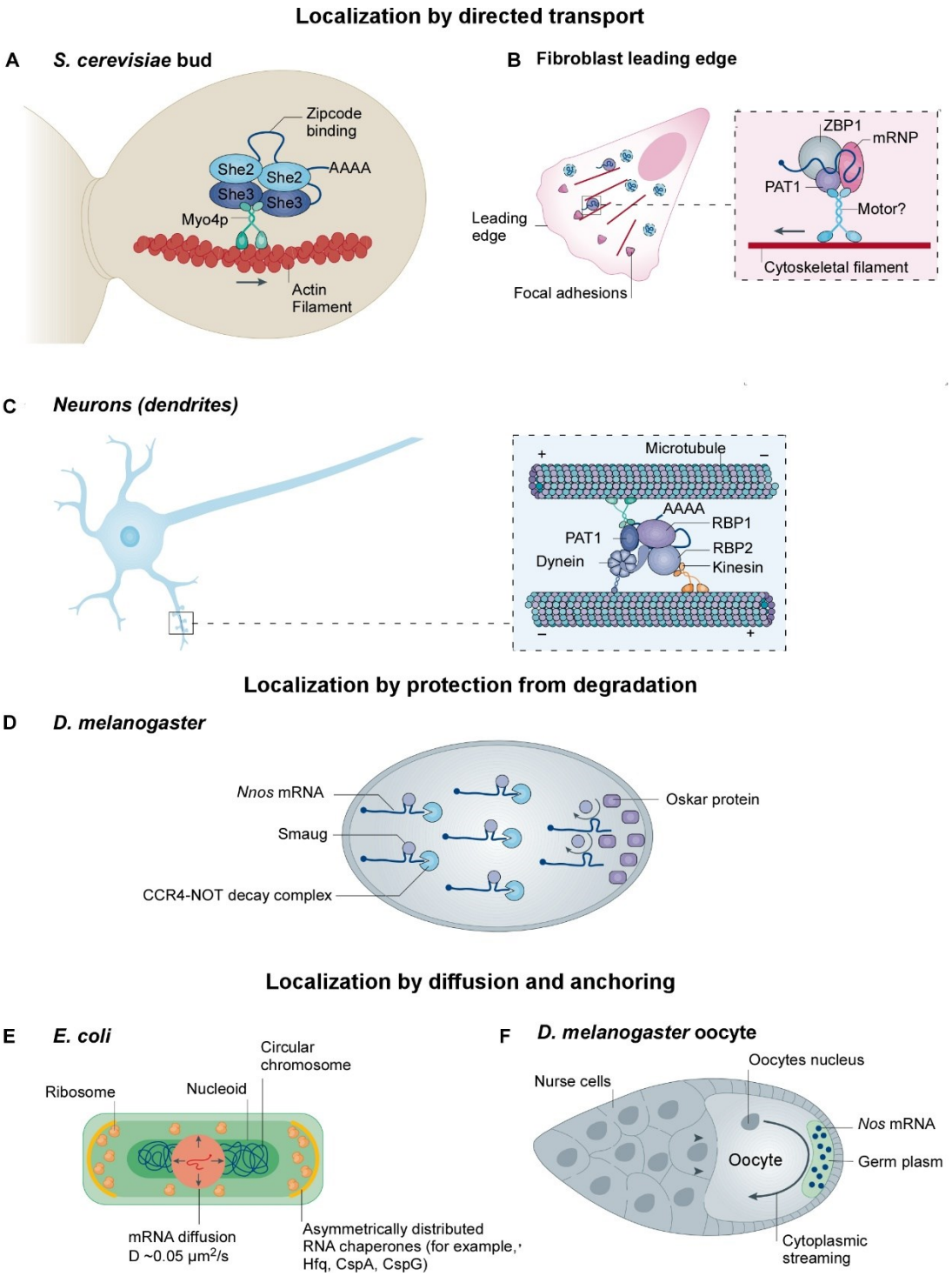
Although many studies have reported that mRNAs are translationally repressed during transport, emerging pieces of evidence have demonstrated the co-translation of mRNAs during transit (SHIBER *et al.* 2018; POPOVIC *et al.* 2020). Studying mRNA transport in multicellular organisms has established that local translation is linked to physiological changes.

1.2 General mode of mRNA transport and mechanism

mRNA localization has been well studied in many model systems, and three central mechanisms governing the localization are classified as follows. 1. Direct mRNA transport, 2. Protection from mRNA decay, 3. Passive diffusion and anchoring.

1.2.1 Direct mRNA transport

Direct mRNA transport or active transport is the most common mode of mRNA localization reported in eukaryotic cells. Motor-driven active mRNA transport is dependent on actin filaments or microtubule cytoskeleton. While actin mediates the short-distance mRNA transport, microtubules mediate the long-distance transport. A common feature of active transport is that the mRNAs transport starts with specific RNA-binding proteins (RBPs) recognizing their target mRNAs by *cis-localization*/regulatory elements, also knowns as Zipcodes. These Zipcodes are often found in 3' UTR but also on the 5'UTRs and coding regions (ANDREASSI AND RICCIO 2009; OLGEISER *et al.* 2019; ENGEL *et al.* 2020; DAS *et al.* 2021). These RNA-binding proteins and other accessory RBPs constitute the large messenger ribonucleoprotein complexes (mRNP). Transporting mRNP units are connected directly to motor proteins or indirectly via an adaptor protein. They hitchhike on organelles such as endosomes or lysosomes and are transported on the cytoskeleton to the subcellular location for local translation (MUNTJES *et al.* 2021; VARGAS *et al.* 2022). Actin-dependent mRNA transport has been well-studied in the baker's yeast *S. cerevisiae* which reproduces by budding, resulting in a large mother and a small daughter cell (Figure 1A). Actin-dependent transport of *ASH1* mRNAs to the daughter cell is vital for switching mating types. This is ensured by the inhibition of HO endonuclease by the Ash1p (Asymmetric synthesis of HO) in daughter cells (NIESSING *et al.* 2018). *ASH1* mRNA has four Zipcodes, present on the coding region and 3'UTR; these



1.2.1.1 Figure 1. Models depicting general modes of mRNA transport and localization in cells and organisms.

(A) Several mRNAs are localized to the bud of *S. cerevisiae*. She2 dimerizes and binds these mRNAs via their Zipcodes before binding She3, which bridges the interaction of the complex with the type V myosin motor Myo4. The ribonucleoparticles are actively transported along actin filaments. (B) In mammalian fibroblasts, mRNAs encoding *β-actin* are localized to the leading edge by RNA binding proteins (RBPs) such as ZBP1, which binds to the Zipcode on the 3′ untranslated regions (3′UTR) of the mRNAs to form messenger ribonucleoproteins (mRNPs) that associate with unidentified motors. PAT1 acts as a direct adapter between ZBP1 and the motor. This represents a small percentage of mRNA movement as most mRNAs undergo corralled cytoplasmic diffusion (indicated by the dashed boundaries). (C) Localization to distal

spines is achieved by packaging mRNAs involved in synaptic remodeling into transport granules composed of RBPs, the minus-end-directed motor dynein, and the plus-end-directed motor kinesin. Due to the mixed polarity of microtubules in dendrites and the presence of both motors, these granules move bi-directionally (in anterograde and retrograde motion). The net movement is proposed to occur by a 'tug-of-war' between the motors determined by their stoichiometry. (D) In *D. melanogaster* embryos, *Nos* mRNAs are bound by the RBP Smaug, which recruits the CCR4–NOT complex to initiate mRNA decay. At the posterior pole, however, *Nos* mRNAs are protected from degradation by Oskar proteins, which displace Smaug to increase local concentrations of *Nos* mRNAs. (E) In *E. coli*, mRNAs localize to ribosome-rich poles or the membrane by random diffusion at speeds of 0.05μm2/s, aided by the chaperone proteins that anchor the mRNAs. (F) During *D. melanogaster* oogenesis, several hundreds of mRNAs are deposited to the oocyte by nurse cells (dashed arrows). mRNAs such as *Nos* are localized to the posterior pole of the oocyte by cytoplasmic streaming and entrapped in the germplasm in an actin-dependent manner. (Figure and legend adapted from (DAS *et al.* 2021b), under the terms of Springer Nature license No. 5412120334416)

are recognized by RBP She2p and its binding partner Loc1 inside the nucleus.

Synergetic binding of She2p dimers to the mRNA makes a conformational switch that promotes high-affinity mRNP complex formation. She2p interacts with the adaptor protein She3p dimer, constitutively bound to the type V myosin motor protein Myo4p, which mediates the mRNP transport on the actin filaments towards the daughter cells (EDELMANN *et al.* 2017). Besides ASH1 mRNA, several other mRNAs, including *CLB2*, *TCB2*, *TCB3* and *IST2,* are actively transported on the She2p-She3p complex. She3, the adaptor protein, also has been shown to interact with mRNA. In alignment with this observation, in the opportunistic human pathogen, *C. albicans* She3p-Myo4p complex mediates the mRNA transport, yet no homologue for She2p is identified (ELSON *et al.* 2009; MCBRIDE 2017).

In mammalian fibroblast cells, mRNAs are transported to the focal adhesion points by active transport on both actin and microtubule cytoskeleton and corralled diffusion (Figure 1B). Translationally repressed *Actb* mRNAs are actively transported by the ZBP1 (Zipcode binding protein 1). However, this mode of transport represents only a tiny fraction. The majority of the mRNAs are transported by corralled cytoplasmic diffusion (Figure 1B Dashed circles) (BISWAS *et al.* 2019; MICHAEL AND PARSONS 2020; DAS *et al.* 2021).

In mammalian cells, motor protein-dependent active mRNA transport is well studied in the neuronal dendrites. Neuronal cells are highly polarized cells containing three distinct parts: a cell body and two extensions known as axons and dendrites. The cell body contains the nucleus and cytoplasm. Axons look like a stem that can extend up to hundreds of millimetres, whereas dendrites look like tree branches. To meet the rapidly changing metabolic needs in the distal reaches of the neurons, they rely on mRNA transport and local translation of proteins on demand (GLOCK *et al.* 2017; TURNER-

BRIDGER *et al.* 2020; FERNANDOPULLE *et al.* 2021b; LI *et al.* 2021). mRNA transport in neurons is orchestrated by coordinating mRNA binding proteins (mRBPs) and motor proteins on the cytoskeleton network. Adapter protein PAT1 (APP tail 1) directly tether the ZBP1 to the kinesin 1 motor protein complex, which mediates the transport of ActB mRNAs on microtubules towards the anterograde direction in the dendrites (Figure 1C). RNA granule-associated proteins such as FMRP, SFPQ, and ZBP1/IGFBP1 are shown to interact with the kinesin motor protein via the adaptor protein KLC (kinesin light chain). FMRP-bound bicaudal D (BIDC2) mRNPs are transported on the dynein towards the retrograde direction (WU *et al.* 2020; FERNANDOPULLE *et al.* 2021; FUKUDA *et al.* 2021). *In vitro* reconstitution studies have shown that adenomatous polyposis coli (APC) protein-bound mRNAs that are tethered to heterotrimeric kinesin-2 (KIF3A/B/KAP3) are sufficient for the microtubule-dependent movement (BAUMANN *et al.* 2020).

1.2.2 Protection from degradation

In *D. melanogaster* embryo posterior end, mRNAs are localized by protection from degradation. Local translation of *Nos* (Nanos) mRNA at the pole plasm (cytoplasm in the posterior pole) is essential for the abdominal development head and thorax segmentation. (DAS *et al.* 2021). *Nos* mRNAs are uniformly distributed in the bulk cytoplasm, which is selectively degraded by RBP Smaug and deadenylation complex CCR4-NOT (Figure 1D). A fraction of *Nos* mRNAs is localized at the poleplasm, which is protected by Oskar by preventing Smaug-mediated degradation. Thus, *Nos* translation is also dependent on the Oskar protein at the poleplasm (Figure 1D) (SEMOTOK *et al.* 2005; ZAESSINGER *et al.* 2006; CHEN *et al.* 2014).

1.2.3 Passive diffusion and anchoring

Prokaryotic cells were long thought to lack the mechanism of mRNA localization because of their smaller size and the non-existence of membrane-bound organelles. However, diverse mRNA localization patterns are observed in bacteria. In addition, mRNAs and the encoded proteins are localized in the same region, indicating the presence of local translation. Emerging pieces of evidence suggest that Spatio-temporal localization of mRNAs can also control the post-transcriptional regulation and translation in prokaryotes. There is no evidence for the active mRNA transport mechanism in bacterial systems reported to date (BENHALEVY *et al.* 2017; FEI AND SHARMA 2018). Recent studies have predicted bacterial mRNAs migrating towards the pole through diffusion (Figure 1E) (CASTELLANA

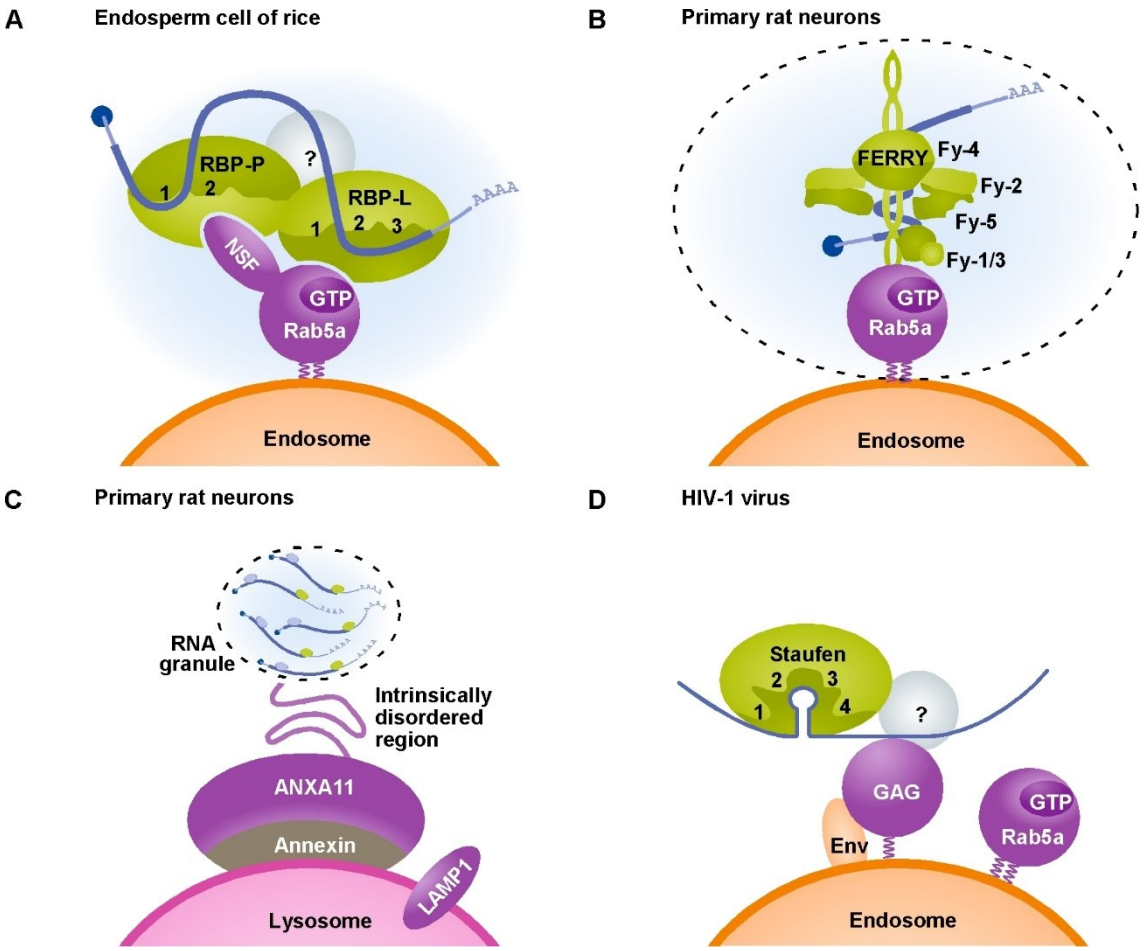
et al. 2016). Asymmetric distribution of RNA chaperones such as Hfq, CspA, and CspG are thought to be involved in the mRNA anchoring and localization (BUSKILA *et al.* 2014; DAS *et al.* 2021).

During mid-oogenesis in *D. melanogaster,* hundreds of mRNAs are selectively enriched in the germplasm by a passive diffusion-anchoring mechanism (Figure 1F). Transcriptionally active nurse cells contract and squeeze their mRNAs, including *Nos* mRNA, into the transcriptionally inactive growing oocyte. Microtubule bundles in the oocyte cause cytoplasmic streaming and swirl the content throughout the oocyte. During this process, mRNAs are entrapped into the germ granules by RNA-RNA interactions (Figure 1F). mRNA localization at the posterior pole within germ granules is essential for germ cell development and function (LU *et al.* 2018; TRCEK AND LEHMANN 2019; DAS *et al.* 2021).

1.3 Endosomal mRNA transport models

mRNAs hitchhike on organelles such as endosomes, lysosomes and mitochondria to overcome the energy cost of motor-dependent long-distance transport (VARGAS *et al.* 2022). Especially the endosomal transport of mRNAs seems to be a conserved phenomenon since several examples arise from fungi and animal model organisms. Furthermore, these examples indicate a close link between endosomal mRNA transport and mitochondrial protein import (CIONI *et al.* 2019; MUNTJES *et al.* 2021).

In developing rice (*Oryza sativa*) endosperm cells, mRNAs of Glutelin and Prolamine (major storage proteins) are transported on early endosomes to the cortical endoplasmic reticulum (ER) (Figure 2A) (TIAN *et al.* 2018). *Cis-localization* elements in these mRNAs are recognized by RRM domain-containing Zipcode binding proteins RBP-P and RBP-L. Together these mRNPs form a quaternary complex with the membrane fusion factors N-ethylmaleimide-sensitive factor (NSF), the endosomal marker Rab5a small guanosine triphosphatase (GTPase) and are transported on the cytoplasmic surface of early endosomes to two distinct subdomains in the cortical endoplasmic reticulum (Figure 2A) (TIAN *et al.* 2020a). Mutations in Rab5a lead to mislocalization of the cargo mRNAs, and mutations in the RBP-P lead to severe growth phenotypes, ranging from dwarfism, chlorophyll deficiency, sterility, late flowering and low spikelet fertility, emphasizing the importance of endosomal mRNP transport in plants (TIAN *et al.* 2020b).



1.3.1.1 Figure 2. Models depicting endosomal mRNA transport in fungi, plants and animals.

(A–E) On the cytoplasmic surface of transport endosomes or lysosomes, mRNPs are attached by different factors (purple) to endosomes. Key RNA-binding proteins (green) interact with cargo RNA (blue). mRNAs are symbolized by CAP (blue circle) and a poly(A) tail. (A). In rice endosperm cells, RBP-P and RBP-L connected with cargo mRNAs are tethered to the cytoplasmic surface of endosomes via the membrane protein N-ethylmaleimide-sensitive factor (NSF), and the endosomal marker protein small-GTPase Rab5a. (B) In primary rat neurons, the FERRY complex (Five-subunit Endosomal Rab5 and RNA/ribosome intermediary) connects the cargo mRNAs to the early endosomes via Rab5a. (C) In primary rat neurons, mRNA granules are bound to the N-terminal intrinsically disordered region of the Annexin11 protein, which is bound to the Lamp1 positive endolysosomes. (D) HIV virus hijacks the host endosomal vesicles to transport their genomic RNA. Staufen-associated mRNAs might be connected to the Rab5a-positive endosomes via the GAG protein. (Figure and legend adapted from (Muntjes et al. 2021), under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License).

Endosomal mRNA transport and endosome-coupled translation are also reported in multiple neuronal model systems in animals. The first report came from the studies using retinal ganglion neuronal cells of *X. laevis,* where several mRNAs encoding mitochondrial proteins were transported on Rab7a-positive late endosomes, which are essential for mitochondrial maintenance. Disruption of Rab7a function by expressing the Charcot Marie-Tooth disease-associated mutation in Rab7a leads to mitochondrial dysfunction, loss of axon integrity and local protein synthesis (CIONI *et al.* 2019). Another report using

image-based transcriptomics and endosome-specific RNAseq approaches has denoted that mRNAs are transported on endosomes in a translation-dependent manner. The protein coding sequence region of EEA1 (Early endosomal Antigen 1) mRNAs is sufficient for its localization on endosomes where the encoded protein resides. This study established that mRNA transport on endosomes is not limited to polarized cells (POPOVIC *et al.* 2020).

Recently, a critical study from the primary rat neurons has identified a novel Rab5 effector protein complex named the FERRY complex (Five-subunit Endosomal Rab5 and RNA/ribosome intermediary). Ferry complex mediates the endosomal mRNA transport and links the translation machinery with a subset of predominantly nuclear-encoded mitochondrial mRNAs (SCHUHMACHER *et al.* 2021b). The FERRY complex consists of five subunits (Fy1-5). It is tethered to the early endosome by the interaction between Fy2 and Rab5a GTPase (Figure 2B). Mutations in different subunits of the FERRY complex are linked to malfunctioning of the brain and neurodegenerative diseases (QUENTIN *et al.* 2021).

Another critical study using primary rat neurons and *in vivo* zebrafish model unravelled that RNA granules are hitchhiking on LAMP1-positive lysosomal vesicles. Using LAMP1-APEX labelling, Annexin A11 (ANXA11) was identified as a molecular tether connecting the β-actin-containing RNP granules to the endolysosomal compartments (LIAO *et al.* 2019). Intrinsically disordered region (IDR) in the N-terminal of ANXA11 facilitates RNA granule formation. C-terminal repeats of ANXA11 mediate calcium-dependent lysosomal binding (Figure 2C). Mutations in ANXA11 are associated with Amyotrophic Lateral Sclerosis (ALS), a progressive nervous system disease. Biochemical studies have shown that the ANXA11 mutation impairs RNA transport, indicating its importance and connection to neurodegenerative diseases (LIAO *et al.* 2019).

RNA viruses such as HIV (Human immunodeficiency virus) and MLV (Murine Leukaemia virus) hijack the host endocytosis machinery to transport their genomic RNA for maturation and fusion to the cell membrane before being secreted out (BASYUK *et al.* 2003). Staufen 1-associated mRNP complexes might be tethered to the Rab5a-positive endosomes. HIV GAG protein could be a potential tethering factor connecting the RNPs to the endosomal vesicles (Figure 2D) (LEHMANN *et al.* 2009; BASYUK *et al.* 2021).

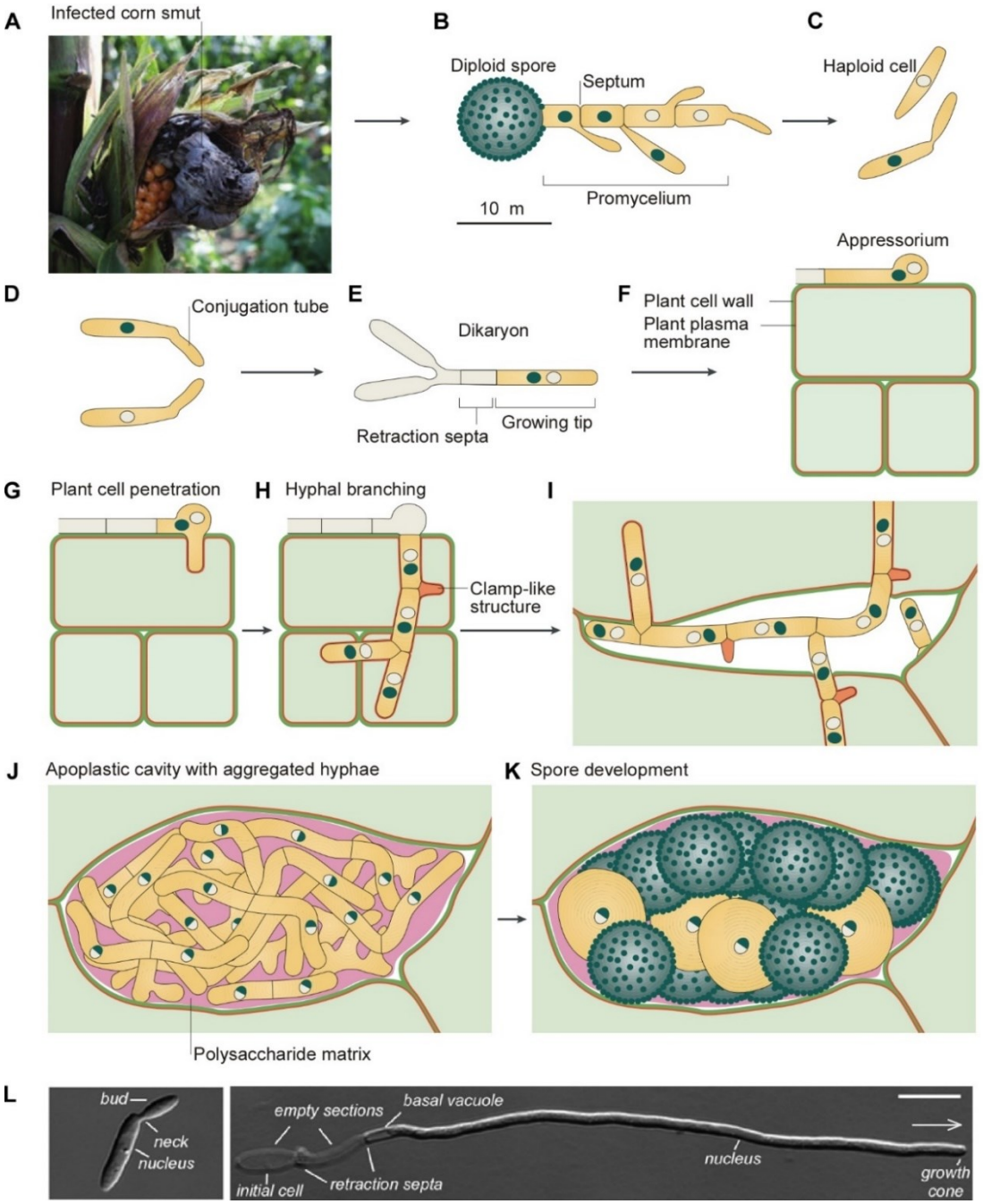
Apart from the endosome-mediated mRNA transport, mRNA hitchhiking on motor-driven membrane-bound organelles such as mitochondria is reported as an alternative transport mechanism. A recent study reported that the nuclear-encoded Pink1 mRNA co-

transport on mitochondria by the RBP Synaptojanin 2 (SYNJ2), which is bound to the outer mitochondrial membrane (OMM) via the adapter protein Synaptojanin 2 binding protein (SYNJ2BP) and supports axonal mitophagy (HARBAUER *et al.* 2022; HARBAUER AND SCHWARZ 2022). A similar study has shown that a higher fraction of nuclear-encoded mitochondrial Cox7c mRNA is hitchhiking on the mitochondria (COHEN *et al.* 2022).

Hitchhiking of mRNAs on endosomes was first reported in *Ustilago maydis (U. maydis),* where cargo mRNAs are transported on Rab5a-positive early endosomes on microtubules by motor proteins bi-directionally. It is one of the well-studied model systems for microtubule-dependent active mRNA transport (HAAG *et al.* 2015; MUNTJES *et al.* 2021).

1.4 The eukaryotic single cell model organism *U. maydis*

U. maydis is a basidiomycetes plant pathogen. It has been established as a model organism to study several cell biological processes such as DNA recombination, microtubule-dependent membrane dynamics, plant-pathogen interaction, endosome-coupled mRNA transport, unconventional secretion, etc., (HAAG *et al.* 2015a; MATEI AND DOEHLEMANN 2016; REINDL *et al.* 2019; MUNTJES *et al.* 2021). *U. maydis* causes smut disease in *Zea mays* and has a dimorphic life cycle. During its life cycle, upon pheromone sensing, yeast-like sporadic haploid cells mate with a compatible partner by conjugation on the leaf surface, which results in the formation of dikaryotic, unipolar filamentous hyphae (Figure 3A-K) (BREFORT *et al.* 2009; DJAMEI AND KAHMANN 2012). The hyphal growth is controlled by the heterodimeric transcription factor bE/bW consisting of two subunits. These subunits are encoded by two different alleles from each mating partner. An active heterodimeric transcription factor is constituted in the dikaryon, which triggers filamentous growth. Therefore, an active heterodimeric transcription factor formation is a prerequisite for hyphal growth and plant infection in wild-type cells (KAMPER *et al.* 1995; VOLLMEISTER *et al.* 2012). In the next steps of the filamentous life cycle, the fungal hyphae directly penetrate the plant cells with the formation of a specialized appressorium- like structure. During this phase, fungal cells establish an intimate and compatible relationship with the host, evade the host defence system and modulate the host metabolism by secreting a battery of effector proteins (BREFORT *et al.* 2009; DJAMEI AND KAHMANN 2012; LANVER *et al.* 2017). Fungal hyphae grow in the inter and intracellular



1.4.1.1 Figure 3. The life cycle of the model organism *U. maydis*.

(A) Symptoms of a cob infection by *U. maydis* in a maize field. The infection is locally confined and characterized by the induction of anthocyanin biosynthesis and the formation of large tumors in which fungal spores develop. (B) Diploid spores are released when tumors break open. They are dark-colored owing to their high melanin content and have a characteristic round shape and surface ornamentation. Meiosis occurs in germinating spores; the four resulting haploid nuclei migrate into a promycelium, which is delineated by septa. (C) Following mitotic divisions, haploid cells bud off from these compartments. (D) After the detection of a compatible mate, the budding program ceases, and cells develop conjugation tubes that are directed toward each other. (E) After cell fusion, a filamentous cell cycle-arrested dikaryon is produced. Only the growing tip of this filament is filled with the cytoplasm (yellow), whereas older parts are vacuolated (grey) and become sealed off by regularly spaced septa. These retraction septa enable filament elongation and the formation of an infective structure (appressorium) in extended infectious hyphae. (F, G) Hyphal tip cells develop appressoria in specific locations on the leaf surface and penetrate plant cells. (H) During the early

stages of infection, the cell cycle arrest is released, hyphae begin to branch, and clamp-like structures (orange) ensure the correct segregation of the two nuclei. The host plasma membrane completely encases hyphae during these intracellular stages (red). (I) With the onset of plant tumor formation, fungal hyphae are mainly detected intracellularly. (J) Subsequently, the two nuclei of the dikaryon fuse, followed by the substantial proliferation of diploid cells that form huge aggregates in apoplastic cavities. Aggregated hyphae become embedded in a gelatinous polysaccharide matrix (pink). (K) Hyphae then fragment and undergo spore development. In all panels, white and dark grey nuclei indicate that they are haploid and have different mating-type genes. Nuclei that are half white and half dark grey indicate diploid nuclei generated through the fusion of white and dark grey nuclei. (L) Morphology of yeast-like sporidial and hyphal cells of the laboratory strain AB33 (scale bar 10 µM; for hyphal cells: 6 h.p.i). (Figure and legend adapted from (Lanver et al. 2017), Springer Nature license No. 5376580465976)

space of the enlarged plant cells, branches at regular intervals and fragments. These fragmented hyphae are rounded off and coated with pigments. The mature teliospores with thick pigmented cell walls enter a dormant state, are dispersed from the corn smut, and continue their life cycle (B. J. SAVILLE 2012; LANVER *et al.* 2017).

Although *U. maydis* is a biotrophic fungus for which mating and dikaryon formation is a prerequisite for unipolar hyphal growth, it could be grown in laboratory conditions as a monokaryon. For example, the AB33 strain (Figure 3L) has been established in which hyphal growth can be induced by simply switching the media as the expression of an active bE/bW heterodimer is under the control of regulatable nitrate promoters (KAMPER *et al.* 1995; BRACHMANN *et al.* 2001).

Besides, *U. maydis* infected corn smut has been a delicacy in Mexico for centuries, making them a safe host organism for biotechnological and therapeutic protein production. Successful production of high-value secondary metabolites such as sesquiterpenoids by metabolic engineering, antibody fragments production via unconventional secretion, and conversion of biomass into valuable products have been demonstrated (VOLLMEISTER *et al.* 2012; FELDBRUGGE *et al.* 2013; SARKARI *et al.* 2014; REINDL *et al.* 2019; LEE *et al.* 2020).

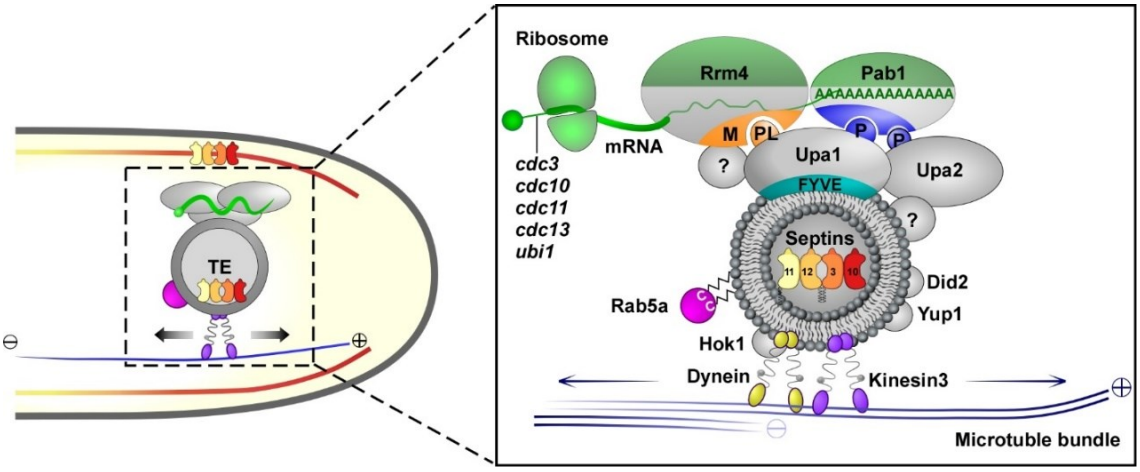
U. maydis genome is well annotated; it encodes 6902 predicted protein-coding genes. Sophisticated molecular biology tools for genetic manipulation with a wide range of vectors for homologous recombination are available. Strains can be efficiently generated by the reverse genetics approach, and selection markers can be recycled after stable insertion. State-of-the-art live cell imaging techniques and quantitative imaging tools have been successfully applied for studying the endosomal dynamics and mRNA transport machinery (KAMPER *et al.* 2006; STEINBERG AND PEREZ-MARTIN 2008; GOHRE *et al.* 2012; TERFRUCHTE *et al.* 2014; BAUMANN *et al.* 2015; MATEI AND DOEHLEMANN 2016).

Thereby. *U. maydis* is a well-characterized and sophisticated model system for basic research and biotechnological applications, especially for studying the endosome-coupled mRNA transport machinery.

1.5 Endosomal mRNA transport in *U. maydis*

Rrm4 is the essential RNA binding protein which mediates the mRNA hitchhiking on endosomes in *U. maydis* (Figure 4). Rrm4 co-localizes to the Rab5a and Yup1 positive early endosomes. The core endosomal mRNA transport machinery consists of the key RNA binding protein (Rrm4), the poly-A tail binding protein (Pab1), the FYVE and RING (Really Interesting New Gene) domain-containing endosomal adapter protein (Upa1), and the multi PAM2 containing scaffold protein (Upa2) that helps in circularizing the mRNA poly(A)-tail. Along with Rrm4, Pab1, Upa1, and Upa2 also co-localize on the cytoplasmic surface of the Rab5a-positive early endosomes (EE) (Figure 4). Loss of Rrm4 causes the formation of aberrant bipolar hyphae, which is comparable to the disturbance of microtubules or endosomal movement (BAUMANN *et al.* 2012). In line with this, the loss of Upa1 or Upa2 leads to severe bipolar growth phenotype (POHLMANN *et al.* 2015). Thus, endosome-mediated mRNA transport and local translation are essential for the efficient unipolar growth of fungal hyphae (BECHT *et al.* 2006a; HAAG *et al.* 2015; NIESSING *et al.* 2018). A nonessential component, the glycine-rich RNA binding protein 1 (Grp1) also shuttles along with the Rrm4-associated mRNPs, which serve as an RNA chaperon. Upa2 and Grp1 have a long IDR that could promote biomolecular condensation. Movement of endosomes is powered by the motor proteins Kinesin 3 (Kin3) towards the anterograde direction and cytoplasmic split Dynein 1/2 (Dyn 1/2) towards the retrograde direction along the microtubule cytoskeleton (Figure 4).

Rrm4 is a post-transcriptional regulator which transports thousands of translationally active mRNAs for local translation (KONIG *et al.* 2009; BAUMANN *et al.* 2012). In addition, it is also reported to co-transport polysomes throughout the fungal hyphae, which suggests an alternative hypothesis of the distributive function of endosomes along with hyphae (HIGUCHI *et al.* 2014). Rrm4 has three ELAV-like (Embryonic lethal abnormal vision) RRM domains in the N-terminus for RNA-binding and two MLLE domains in the C-terminus for Protein-protein interactions (BECHT *et al.* 2006; POHLMANN 2013). *In vivo* individual-nucleotide resolution UV crosslinking and immunoprecipitation (iCLIP) studies have shown that RRMs of Rrm4 binds simultaneously in multiple regions of the target mRNA, but predominantly in the 3' untranslated region (UTR). Furthermore, Rrm4



1.5.1.1 Figure 4. Model depicting mechanism of mRNA transport in *U. maydis.*

Overview of an infectious hypha (*left*). The dashed box indicates that the mRNA and RNA binding proteins are bound to the cytoplasmic surface of transporting endosome (TE), which shuttles along antiparallel microtubule bundles (dark blue) toward the growing tip of the hypha. The magnified image (*right*) depicts the mechanistic details of endosomal mRNA transport components. Cargo mRNAs, including *cdc3, cdc10, cdc11, cdc13,* and *ub3* (green), are bound by the N-terminal RRM domains of Rrm4 (green). Ribosome-bound mRNA indicates the co-translation of mRNAs during transport. The C-terminal MLLE domains of Rrm4 (orange) interact with PAM2L1,2 motifs of Upa1 (orange). The four RRMs of Pab1 (green) interact with the poly(A) tail, and the MLLE domain of Pab1 (dark blue) interacts with PAM2 motifs (dark blue) of Upa1, an adaptor protein, and Upa2, a dimerizing scaffold protein. Upa1 is attached to endosomes via its FYVE domain(cyan), and the C-terminal GWW motif of Upa2 is crucial for its endosomal binding. Rab5a (pink), Did2, and Yup1 are markers for early endosomes. Endosomes are transported on microtubule cytoskeletons by the motor proteins Kinesin3 (purple) and Dynein (yellow), anterograde and retrograde motions, respectively. Hok1 coordinates the motor attached to the endosome. (Figure is adapted from (POHLMANN *et al.* 2015a), under the Creative Commons Attribution 4.0 International Public License)

bound to UAUG binding motif by its third RRM. UAUG is enriched within the ORF of target mRNAs which could indicate that Rrm4 (especially RRM3) also plays a role in translational regulation of its targets (OLGEISER *et al.* 2019). Rrm4 mediates the long-distance transport of *ubi1 and rho3* mRNAs (KONIG *et al.* 2009). iCLIP studies combined with RNA live imaging has shown that all the four septin mRNAs *cdc3, cdc10, cdc11, and cdc12,* along with their encoded protein, were transported on the endosomes (KONIG *et al.* 2009; BAUMANN *et al.* 2014; ZANDER *et al.* 2016). In the targets of Rrm4, mitochondrial protein-encoding mRNAs were significantly enriched, with binding sites at translational landmarks such as stop codons (OLGEISER *et al.* 2019).

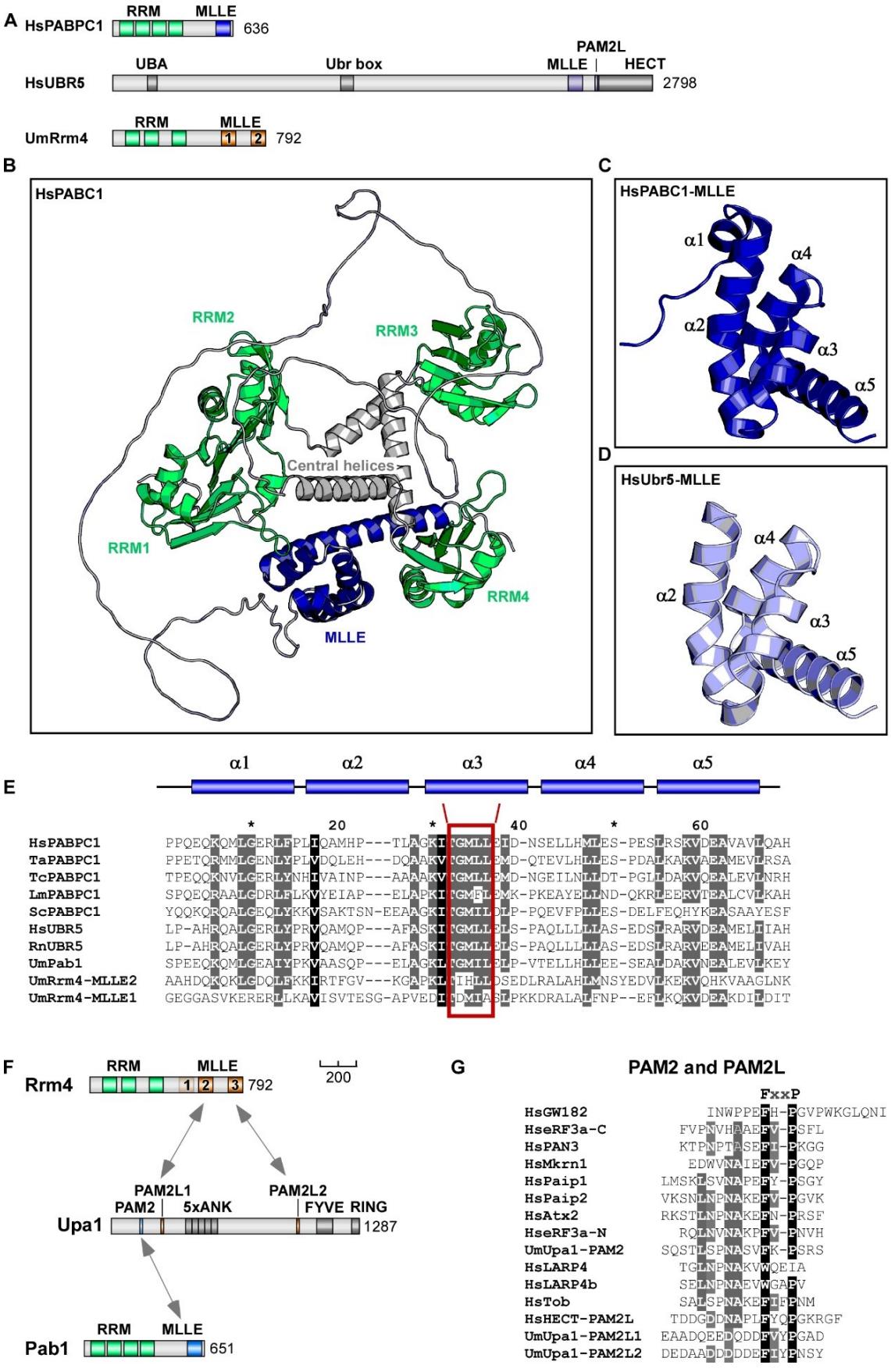
N-terminal RRM domains and C-terminal MLLE domains of Rrm4 are connected by a flexible linker region, similar to the human cytoplasmic poly(A) tail binding protein 1 (PABPC1) (DE MELO NETO *et al.* 2018). Pab1, the *U. maydis* orthologue of PABPC1 and a core component of endosomal mRNA transport machinery have comparable domain architecture. Although both RBPs Rrm4 and Pab1 have RRM domains in the N-terminal

and MLLE domain in the C-terminal, they have distinct overall domain architecture, sequence recognition and function (BECHT *et al.* 2006; POHLMANN *et al.* 2015). In the N-terminus, Rrm4 has three RRMs, whereas Pab1 has four RRMs. Similarly, in the C-terminus of Rrm4, two MLLE domains were identified based on the sequence homology. In contrast, only a single MLLE domain is present in Pab1, similar to human PABPC1. Disturbance in either the N-terminal RRM1-2 or the C-terminal MLLE domain of Rrm4 results in the loss of function phenotype and abolish the endosomal mRNA transport (BECHT *et al.* 2006).

The core components of endosomal mRNA transport (Rrm4, Upa1 and Upa2) are highly conserved in basidiomycetes filamentous smut fungus. Orthologues of Rrm4 and Upa1 are also found in the distantly related fungal phyla Mucoromycota and Zoopagomycota (MULLER *et al.* 2019). Though numerous factors involved in the endosomal mRNA transport are known, the precise mechanism of their association with the endosome still needs to be better understood. Rrm4-associated mRNPs are mainly connected to the endosome via the MLLE domains of Rrm4 and Pab1. Even though both Rrm4 and Pab1 have MLLE domains, they specifically interact with distinct targets. The loss of these interactions disturbs their endosomal association and results in bipolar phenotype.

1.6 MademoiseLLE (MLLE) domains and their binding partner PAM2 motif

The MademoiseLLE (MLLE) domain was initially identified as a protein-protein interaction (PPI) domain from the C-terminal of cytoplasmic poly(A)-binding protein (PABPC1/Pab1) therefore initially named as PABC domain (DEO *et al.* 2001a; KOZLOV *et al.* 2001). Later it was renamed as the MLLE domain due to the presence of conserved MLLE residues in the core of the PPI interaction pocket (KOZLOV AND GEHRING 2010). MLLE domains are highly conserved in eukaryotes and well characterized from humans, plants, yeast and protozoa (DEO *et al.* 2001; KOZLOV *et al.* 2001; KOZLOV *et al.* 2002; SIDDIQUI *et al.* 2003; SIDDIQUI *et al.* 2007). It is an approximately 80 amino acid (AA) long polypeptide chain consisting of right-handed supercoiled five helices. The central alpha-helix has conserved KITGMLLE residues from which the domain gets the name and abbreviation of MademoiseLLE in French (Figure 5A-C). It enables PPI by forming a peptide binding pocket for binding partners containing the poly(A)-binding protein associated motif (PAM2). PAM2 motifs consisted of conserved (L/P/F)X(P/V)XAXX(F/W)XP residues in the eukaryotes (KOZLOV *et al.* 2001a; XIE *et al.* 2014). Leucine and Phenylalanine in the PAM2 peptide are the most critical residues


A
HsPABPC1
RRM
MLLE
636
HsUBR5
UBA
Ubr box
PAM2L
MLLE
HECT
2798
UmRrm4
RRM
MLLE
1
2
792
B
HsPABC1
RRM2
RRM3
Central helices
RRM1
RRM4
MLLE
C
HsPABC1-MLLE
α1
α4
α2
α3
α5
D
HsUbr5-MLLE
α4
α2
α3
α5
E
α1
α2
α3
α4
α5
HsPABPC1
TaPABPC1
TcPABPC1
LmPABPC1
ScPABPC1
HsUBR5
RnUBR5
UmPab1
UmRrm4-MLLE2
UmRrm4-MLLE1
F
Rrm4
RRM
MLLE
1
2
3
792
200
PAM2L1
PAM2L2
PAM2
5xANK
FYVE RING
Upa1
1287
Pab1
RRM
MLLE
651
G
PAM2 and PAM2L
FxxP
HsGW182
HseRF3a-C
HsPAN3
HsMkrn1
HsPaip1
HsPaip2
HsAtx2
HseRF3a-N
UmUpa1-PAM2
HsLARP4
HsLARP4b
HsTob
HsHECT-PAM2L
UmUpa1-PAM2L1
UmUpa1-PAM2L2

1.6.1.1 Figure 5. MLLE domains and its interaction partners PAM2, PAM2L motifs.

(A) Schematic representation of MLLE domain-containing proteins drawn to scale (bar, 200 amino acids, number of amino acids indicated next to protein bars) using the following coloring: lime green, RNA recognition motif (RRM); dark blue, $MLLE^{PABPC1}$; light purple, $MLLE^{Ubr5}$; orange, $MLLE^{Rrm4}$; bright blue, $MLLE^{Pab1}$ domains; light purple, $PAM2L^{Ubr5}$; UBA, UBR box and HECT domain of Ubr5 are shaded in dark grey (B) Three dimensional (3D) structural model of human PABPC1 (left) predicted by AlphaFold2. Globular domains are represented as cartoons using the coloring: lime green, RNA recognition motif (RRM); grey, central helices; bright blue, $MLLE^{PABPC1}$. (C-D) Detailed view of MLLE domain from human PABPC1 and Ubr5, exhibiting MLLE domains are made of right-handed super helices. $MLLE^{PABPC1}$ consists of five helices, and $MLLE^{Ubr5}$ consists of four helices. (E) Sequence alignment of previously determined MLLE domains showing the degree of similarity to the two Rrm4-MLLE domains and the positions (*Hs - Homo sapiens, Ta - Triticum aestivum, La - Leishmania major, Sc - Saccharomyces cerevisiae, Tc - Trypanosoma cruzi, Rn - Rattus norvegicus, Um - Ustilago maydis, PABPC1, Pab1 – poly[A]-binding protein, UBR5 - E3 ubiquitin-protein ligase*). Conserved MLLE residues are present in the third and central helix. Accession number and sequence coverage are listed in S1 Table. Multiple sequence alignments were performed using ClustalW. (F) Schematic representation (left) of molecular interactions between the Pab1, Rrm4, and Upa1 via $MLLE^{Pab1}$-$PAM2^{Upa1}$, $MLLE^{Rrm4}$-$PAM2L1,2^{Upa1}$ interactions. Arrows indicate the interacting regions. Rrm4 and Pab1 domains are colored as described above. Upa1 domains are colored as follows: light orange, PAM2L sequence (PL1 – 2); Ankyrin repeats (5xANK), FYVE domain, and RING domain of Upa1 are given in dark grey. (G) Comparison of PAM2 and PAM2L sequences (right) found in Upa1 (A0A0D1E015) with those of human proteins, such as GW182 Q9HCJ0), eRF3 (P15170), PAN3 (Q58A45), Mkrn1 (Q9UHC7), Paip1 (Q9H074), Paip2 (Q9BPZ3), Atx2 (-Q99700), NFX (Q12986), LARP4 (Q71RC2), LARP4b (Q92615), Tob (P50616), HECT (O95071).

for the interaction as they are inserted into the hydrophobic peptide binding pocket of MLLE domains (KOZLOV *et al.* 2004; XIE *et al.* 2014).

PAM2 motifs are present only in the Low complexity region (LCR), often IDR, and strictly outside the globular domains (ALBRECHT AND LENGAUER 2004). Their phosphorylation controls PAM2 interaction at the IDR's adjacent Ser/Thr site (HUANG *et al.* 2013). In humans, the MLLE domain of PABPC1 ($MLLE^{PABPC1}$) interacted with several mRNA-associated proteins via its PAM2 (poly[A]-a binding protein associated motif). Although an MLLE domain was present in the yeast ortholog of PABPC1, Pab1p, no PAM2-containing proteins have been identified in *S. cerevisiae* (MANGUS *et al.* 1998b; KOZLOV *et al.* 2002).

1.6.2 MLLE-PAM2 interactions are mainly hydrophobic

MLLE domain is present only in two proteins other than Rrm4, which are PABPC1/Pab1 and Ubr5 (DEO *et al.* 2001; KOZLOV *et al.* 2001; BECHT *et al.* 2005). While PABPC1 and Rrm4 are mRNPs, Ubr5 is an EDD E3 ubiquitin ligase. MLLE domain of PABPC1, Ubr5 and their binding partner PAM2 are well characterized with several 3D structures available in the Protein Data Bank (PDB), whereas the tandem MLLE domains from the third protein Rrm4 remain principally uncharacterized.

1.6.3 MLLE-PAM2 connection in *U. maydis*

Bioinformatics survey of PAM2 motifs containing proteins in *the U. maydis* genome identified 14 candidates (POHLMANN *et al.* 2015), including the adapter protein Upa1, which contains a C-terminal FYVE domain, known for interacting with the Phosphatidylinositol 3-phosphates lipids, a predominant species constituting the endosomal membrane (MARAT AND HAUCKE 2016). However, in the absence of Upa1, residual endosomal shuttling of Rrm4 is still observed, suggesting the presence of additional endosomal adapter proteins. Upa1 has one PAM2 motif with which it interacts with the MLLE domain of the Pab1 ($MLLE^{Pab1}$). Besides, Upa1 has two PAM2L motifs, which resemble PAM2 and are therefore termed PAM2-like motifs, which have been demonstrated to interact with the MLLE domain of Rrm4 (POHLMANN *et al.* 2015). MLLE domains of Rrm4 and Pab1 share 42% sequence similarity and 27% identity. The dimerizing scaffold protein Upa2 contains four PAM2 motifs in the IDR, which interacts with the MLLE domain of the Pab1. Upa2 has a conserved GWW domain in the C-terminus, which is essential for endosomal attachment; regardless, the interaction partner of GWW is currently unknown (JANKOWSKI *et al.* 2019). Thus, Upa1 serves as the early endosomal adaptor protein that connects the Rrm4, Pab1 and associated mRNPs via the MLLE-PAM2/PAM2L interactions (POHLMANN *et al.* 2015).

1.7 Aim of the thesis

Rrm4-associated mRNPs in *U. maydis* is one of the well-studied models for long-distance mRNA transport. Previous studies have established that Rrm4, the key RNA binding protein, hitchhikes on Rab5a-positive early endosomes for a membrane-coupled mRNA transport. Rrm4 does not have a membrane-binding domain. It is recruited to the organelle surface by the endosomal adaptor protein Upa1, which has a PAM2 motif and two PAM2L motifs. The former interacts with the MLLE domain of Pab1, and the latter interacts with the MLLE domain(s) of Rrm4. Although MLLE domains are conserved in sequence and structure, they have distinct targets and are very specific in target recognition. The MLLE domain of Rrm4 does not interact with the PAM2 motif of Upa1; similarly, the MLLE domain of Pab1 does not recognize the PAM2L motif of Upa1. The N-terminal RRM domains of Rrm4 are well characterized with extensive *in vitro* and *in vivo* studies. In the C-terminus of Rrm4, one MLLE domain was initially identified, the disturbance of which results in loss of function phenotype. Based on the sequence similari-

ties and domain predictions second MLLE domain was identified in the C-terminal. However, its role in target recognition and function needs to be characterized. Rrm4 is the only protein containing tandem MLLE domains. How Rrm4 is mechanistically linked to dynamically shuttling endosomes remains an important question. A key aspect is how Rrm4 uses the C-terminal MLLE domains for specific endosomal localization. This dissertation is the first study to characterize the structure and function of C-terminal tandem MLLE domains of Rrm4.

The first objective of this study was to characterize the structure and function of the tandem MLLE domain-containing C-terminal of Rrm4. The second objective was to understand the molecular mechanism underlying the target recognition and specificity of the MLLE domains of Rrm4 and Pab1. The third objective was to decipher the MLLE code by identifying the crucial amino acids required for the $MLLE3^{Rrm4}$-$PAM2L^{Upa1}$ and $MLLE^{Pab1}$-$PAM2^{Upa1}$ interaction. In addition to understanding the mechanical details of endosomal Rrm4-associated mRNPs, studying the MLLE-PAM2/PAM2L connections in *U. maydis* could improve the knowledge of Rrm4-associated mRNPs networks and regulation in the endosome coupled long-distance mRNA transport, as well as it could provide critical insights in the MLLE-PAM2 connections in higher order eukaryotes.

1.8 Results and key structure of the thesis

The following chapters describe the study towards characterizing the MLLE domains of the key RBP, Rrm4, and Pab1 in *U. maydis*. The complete study is divided into three parts.

- The first part (Chapter 2) describes establishing the C-terminus of Rrm4 consisting of a PPI platform consisting of three MLLE domains. This chapter is already published in Plosgenetics 2022 as follows. “A MademoiseLLE domain binding platform links the key RNA transporter to endosomes.” Hence, it is adapted with minor changes without affecting the scientific content.
- The second part (Chapter 3) describes “Deciphering the structure and molecular basis of peptide recognition by the key MLLE domains in *U. maydis.*
- The third part (Chapter 4) describes “Identification of novel interaction partners of MLLE domains in *U. maydis*.
- The fourth part (Chapter 5) discusses the results and interpretations.
- The final and fifth part (Chapter 6) provides the Materials and method, Supplementary information for the Chapters 3-5.

2 A MademoiseLLE domain binding platform links the key RNA transporter to endosomes

Senthil-Kumar Devan[1, $], Stephan Schott-Verdugo[2, 3, $], Kira Müntjes[1], Lilli Bismar[1], Jens Reiners[4], Eymen Hachani[5], Lutz Schmitt[5], Astrid Höppner[4], Sander HJ Smits[4, 5, *], Holger Gohlke[2, 3, *], Michael Feldbrügge[1, *]

[1] Institute of Microbiology, Heinrich Heine University Düsseldorf, Cluster of Excellence on Plant Sciences, Düsseldorf, Germany
[2] John von Neumann Institute for Computing (NIC), Jülich Supercomputing Centre (JSC), Institute of Biological Information Processing (IBI-7: Structural Bioinformatics), and Institute of Bio- and Geosciences (IBG-4: Bioinformatics), Forschungszentrum Jülich GmbH, Jülich, Germany
[3] Institute for Pharmaceutical and Medicinal Chemistry, Heinrich Heine University Düsseldorf, Düsseldorf, Germany
[4] Center for Structural Studies, Heinrich Heine University Düsseldorf, Düsseldorf, Germany
[5] Institute of Biochemistry I, Heinrich Heine University Düsseldorf, Düsseldorf, Germany

$ shared first authorship, * shared corresponding authorship
This research article was published in PlosGenetics 2022 and available at https://doi.org/10.1371/journal.pgen.1010269

2.1 Abstract

Spatiotemporal expression can be achieved by transport and translation of mRNAs at defined subcellular sites. An emerging mechanism mediating mRNA trafficking is microtubule-dependent co-transport on shuttling endosomes. Although progress has been made in identifying various components of the endosomal mRNA transport machinery, a mechanistic understanding of how these RNA-binding proteins are connected to endosomes is still lacking. Here, we demonstrate that a flexible MademoiseLLE (MLLE) domain platform within RNA-binding protein Rrm4 of *Ustilago maydis* is crucial for endosomal attachment. Our structure/function analysis uncovered three MLLE domains at the C-terminus of Rrm4 with a functionally defined hierarchy. MLLE3 recognizes two PAM2-like sequences of the adaptor protein Upa1 and is essential for endosomal shuttling of Rrm4. MLLE1 and MLLE2 are most likely accessory domains exhibiting a variable binding mode for interaction with currently unknown partners. Thus, endosomal attachment of the mRNA transporter is orchestrated by a sophisticated MLLE domain binding platform.

2.2 Author summary

Eukaryotic cells rely on sophisticated intracellular logistics. Macromolecules like mRNA must be transported to defined subcellular destinations for local translation. This is mediated by active transport along the cytoskeleton. Endosomes are carrier vehicles that shuttle along microtubules by the action of molecular motors. It is currently unclear how

mRNAs are attached mechanistically to these membranous units during transport. We study the model microorganism *Ustilago maydis* where numerous components of endosomal mRNA transport have already been identified. Previously, we found that the key RNA-binding protein Rrm4 interacts with the endosomal adaptor protein Upa1. Here, we perform a structure-function analysis and discovered that Rrm4 contains not one but three different versions of a protein-protein interaction domain, called the MademoiseLLE domain, to facilitate the attachment with transport endosomes. Importantly, they function with a strict hierarchy with one essential domain and the others play accessory roles. This is currently, the most detailed mechanistic description of how an RNA-binding protein and its bound cargo mRNAs are attached to endosomes. The usage of three similar protein-protein interaction domains forming a complex binding platform with a defined hierarchy might be operational also in other unknown protein-protein interactions.

2.3 Introduction

mRNA localization and local translation are essential for spatiotemporal control of protein expression. An important mechanism to achieve localized translation is the active transport of mRNAs along the cytoskeleton (MOFATTEH AND BULLOCK 2017; DAS *et al.* 2021a; FERNANDOPULLE *et al.* 2021). Mainly, long-distance transport of mRNA is mediated by motor-dependent movement along microtubules. Transport endosomes are important carriers that move messenger ribonucleoprotein complexes (mRNPs), consisting of RNA-binding proteins and cargo mRNAs on their cytoplasmic surface (MOFATTEH AND BULLOCK 2017; NIESSING *et al.* 2018b; MÜNTJES *et al.* 2021). This process is evolutionarily conserved in fungi, plants, and animals (BAUMANN *et al.* 2012; CIONI *et al.* 2019; LIAO *et al.* 2019; TIAN *et al.* 2020; MÜNTJES *et al.* 2021; QUENTIN *et al.* 2021; SCHUHMACHER *et al.* 2021). In endosperm cells of developing rice seeds, cargo mRNAs are transported to the cortical endoplasmic reticulum ER by the action of the two RNA recognition motif (RRM)-containing proteins RBP-P and RBP-L. These form a quaternary complex with membrane trafficking factor NSF (*N*-ethylmaleimide-sensitive factor) and small GTPase Rab5a on the endosomal surface (TIAN *et al.* 2020b). In neurons, mRNA transport has been linked to early and late endosomes as well as lysosomal vesicles. Especially, local translation of mRNAs encoding mitochondrial proteins on the surface of late endosomes is needed for mitochondrial function. Importantly, this trafficking process has been associated with the neuronal Charcot Marie-Tooth disease (CIONI *et al.* 2019). Annexin 11, a factor implicated in amyotrophic lateral sclerosis (ALS), was found

as an mRNP linker on motile lysosomal vesicles (LIAO *et al.* 2019). Also, the five-membered FERRY complex was recently identified connecting mRNAs encoding mitochondrial proteins to neuronal endosomes by interaction with the active form of Rab5 (QUENTIN *et al.* 2021b; SCHUHMACHER *et al.* 2021a).

Among the best-studied examples of membrane-coupled mRNA transport is the endosomal mRNA transport in the corn pathogen *Ustilago maydis* (HAAG *et al.* 2015b; BÉTHUNE *et al.* 2019; MÜNTJES *et al.* 2021). Extensive peripheral movement of mRNAs is needed for efficient unipolar growth of infectious hyphae. These hyphae grow highly polarized by expanding at the growing tip and inserting regularly spaced septa at the basal pole. Loss of mRNA distribution causes aberrant bipolar growth (BECHT *et al.* 2006; BAUMANN *et al.* 2012; POHLMANN *et al.* 2015). Key vehicles of cargo mRNAs are Rab5a-positive endosomes that shuttle along microtubules by the concerted action of plus-end directed kinesin-3 and minus-end directed cytoplasmic dynein (BAUMANN *et al.* 2012). Important cargo mRNAs are, for example, all four septin mRNAs. Their local translation during transport is essential to form heteromeric septin complexes on the surface of transport endosomes. Endosomes deliver these complexes to the hyphal tip, forming a defined gradient of septin filaments at the growing pole (BAUMANN *et al.* 2014; ZANDER *et al.* 2016; OLGEISER *et al.* 2019).

Rrm4 is the key RNA-binding protein of the transport process that recognizes defined sets of cargo mRNAs via its three N-terminal RRMs (Fig 1A) (OLGEISER *et al.* 2019). Rrm4 and bound cargo mRNAs are linked to endosomes by Upa1, containing a FYVE zinc finger for interaction with PI_3P lipids (phosphatidylinositol 3-phosphate; Fig 1A) (STENMARK *et al.* 2002; POHLMANN *et al.* 2015). The adaptor protein Upa1 contains a PAM2 motif (poly[A]-binding protein interacting motif 2) (ALBRECHT AND LENGAUER 2004; KOZLOV *et al.* 2004; JINEK *et al.* 2010) and two PAM2-like (PAM2L) sequences. These motifs are crucial for interaction with MademoiseLLE (MLLE) domains of the poly(A)-binding protein Pab1 and Rrm4, respectively (Fig 1A) (POHLMANN *et al.* 2015).

The MLLE domain was first identified as a conserved domain at the C-terminus of the human cytoplasmic poly(A)-binding protein 1 (PABPC1) (MANGUS *et al.* 1998; MANGUS *et al.* 2003). Solution and crystal structures of PABC domains from PABPC1 and ubiquitin ligase UBR5 showed that they are structurally conserved (DEO *et al.* 2001; KOZLOV *et al.* 2001). The domain is about 70 amino acids in length and consists of five bundled α-helices. Interaction with the PAM2-binding motif (consensus sequence xxLNxxAxEFxP) is characterized by the central α-helix 3 with the sequence KITGMLLE

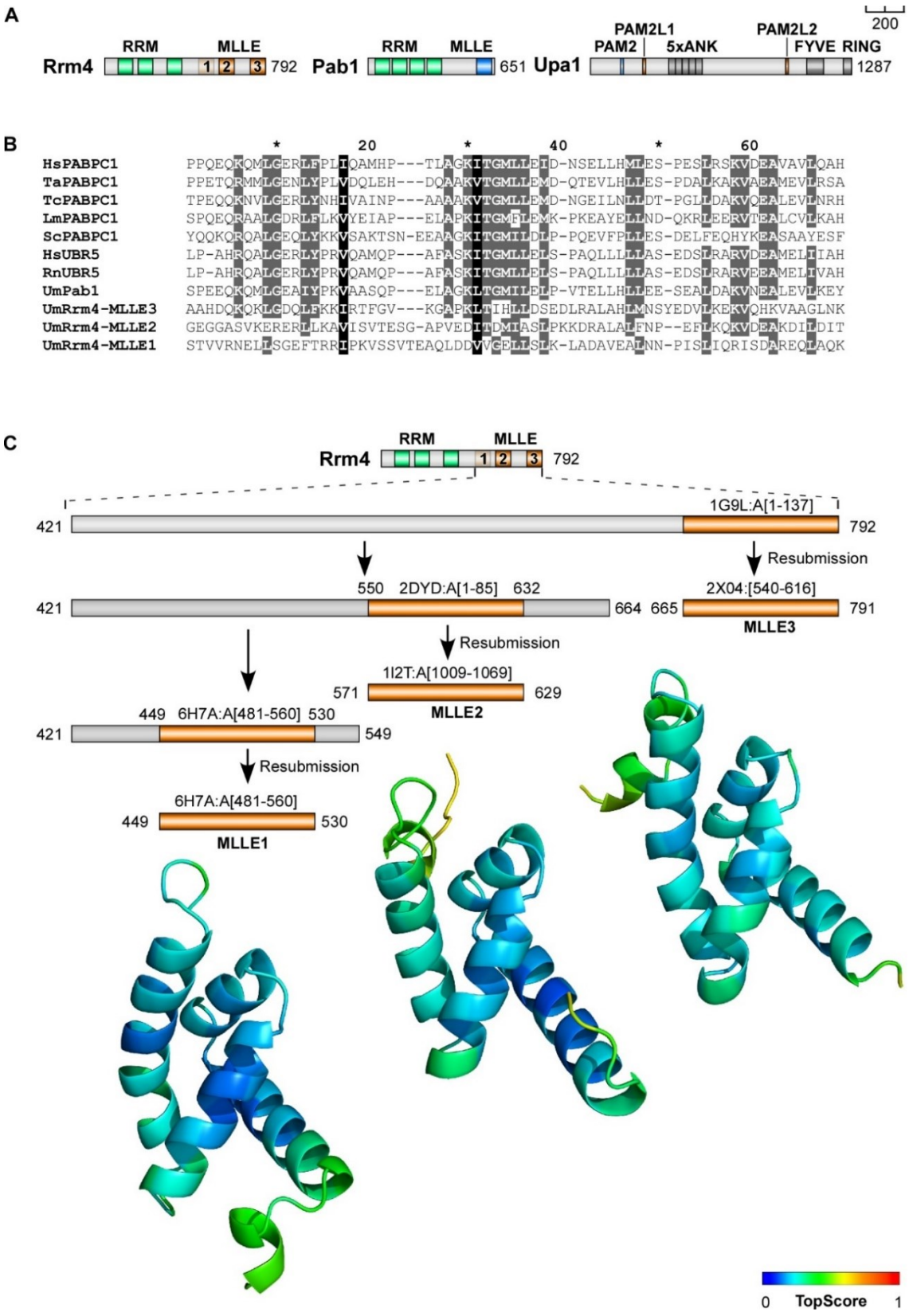
and mediated by two adjacent hydrophobic pockets (XIE *et al.* 2014), with the binding of the Phe residue of the PAM2 motif being the major determinant for this interaction (KOZLOV AND GEHRING 2010). Besides human PABPC1, there are currently only two additional proteins with MLLE domains described: the ubiquitin ligase UBR5 functioning, for example, during microRNA-mediated gene silencing (SU *et al.* 2011) and Rrm4-type RNA-binding proteins from fungi (Figure 6B) (MÜLLER *et al.* 2019).

Mutations in the C-terminal MLLE domain of Rrm4 result in the loss of Rrm4 motility, suggesting that the link to endosomes is disrupted (BECHT *et al.* 2006). Consistently, the C-terminus of Rrm4 recognizes the PAM2L sequence of the adaptor protein Upa1 (POHLMANN *et al.* 2015), suggesting that the interaction of MLLE domains with $PAM2L^{Upa1}$ sequences is responsible for its endosome association. This study combines structural biology with fungal genetics to demonstrate that the C-terminal half of Rrm4 has three divergent MLLE domains with a flexible arrangement and each domain contributes differentially to the endosomal attachment.

2.4 Results

2.4.1 Iterative structural modelling predicts three MLLE domains at the C-terminus of Rrm4

To generate structural models of the MLLE domains, present in Rrm4, we focused on the C-terminal part of the protein (residues 421 to 792). This excluded the three N-terminal RRMs but included the previously predicted two C-terminal MLLE^{Rrm4} domains (Figure 6A-B) (MÜLLER *et al.* 2019). Subjecting this region to iterative comparative modelling with TopModel (Figure 6C) (MULNAES *et al.* 2020) revealed, as expected, the previously identified two regions with homology for MLLE domains located at residues 571-629 and 712-791 (denoted MLLE2^{Rrm4} and MLLE3^{Rrm4}; Figure 6C) (MÜLLER *et al.* 2019). Unexpectedly, using the TopModel workflow with its efficient template selection capabilities (MULNAES *et al.* 2020), we identified an additional *de novo* predicted MLLE^{Rrm4} domain located at residues 451-529 (denoted MLLE1; Figures 6B-C; S1A). Although the sequence identity between templates and their respective Rrm4 sequence stretches was only 17 to 32% (Figure 6B, S1A), the generated MLLE^{Rrm4} domain models had a high predicted local structural quality, as assessed by TopScore (Figure 6C) (MULNAES AND GOHLKE 2018). The generated models were also verified by the current deep neural network modelling approaches AlphaFold2 and RoseTTAFold (Figure S1B) (BAEK *et al.* 2021; JUMPER *et al.* 2021), further indicating that the C-terminal half of Rrm4 has three MLLE domains instead of the previously identified two. All of these MLLE^{Rrm4} domains might be relevant for the interaction with Upa1.



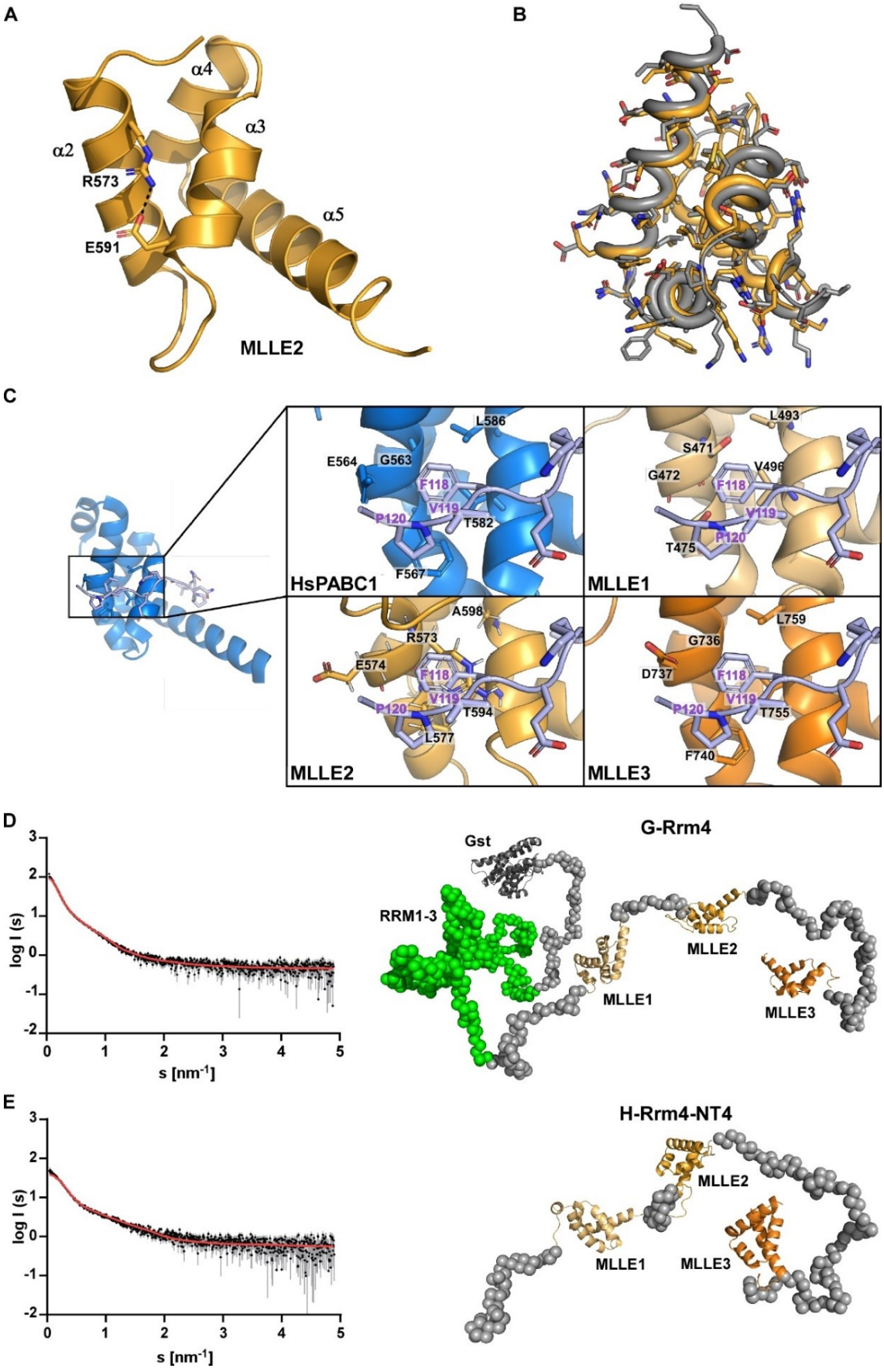
2.4.1.1 Figure 6. The C-terminal half of Rrm4 contains three MLLE domains.

(A) Schematic representation of protein variants drawn to scale (bar, 200 amino acids, number of amino acids indicated next to protein bars) using the following coloring: lime green, RNA recognition motif (RRM); orange, MLLE^{Rrm4} domains; dark blue, MLLE^{Pab1}; light blue PAM2; light orange PAM2L sequence (PL1 – 2) Ankyrin repeats (5xANK), FYVE domain, and RING domain of Upa1 are given in dark grey.

(B) Sequence alignment of previously determined MLLE domains showing the degree of similarity to the three Rrm4-MLLE domains and the positions (Hs - *Homo sapiens*, Ta - *Triticum aestivum*, La - *Leishmania major*, Sc - *Saccharomyces cerevisiae*, Tc - *Trypanosoma cruzi*, Rn - *Rattus norvegicus*, Um - *Ustilago maydis*, PABPC1, Pab1 – poly [A]-binding protein, UBR5 - E3 ubiquitin-protein ligase). Accession number and sequence coverage are listed in S1 Table. Multiple sequence alignment was performed by ClustalW. (C) Identification and modeling of C-terminal MLLE domains of Rrm4. The iterative process is depicted graphically. The best-identified template for each run, and the region of that template that aligns with Rrm4, are displayed (see also S1A Figure for the templates used for the final models). The structural models obtained are shown for the span of the first identified template and are colored according to their per-residue TopScore, where the scale from 0 to 1 indicates a low to high local structural error.

2.4.2 X-ray analysis of the second MLLE domain confirms the predicted structural models

To verify the structural models further, we expressed and purified an N-terminally truncated version of the Rrm4 carrying the three MLLE^{Rrm4} domains in *Escherichia coli* (Figure S2A-B; version H-Rrm4-NT4 carrying an N-terminal hexa-histidine-tag; Materials and methods) (POHLMANN *et al.* 2015). Size exclusion chromatography combined with Multi-angle light scattering (MALS) indicated that the protein was homogenous and did not form aggregates (Figure S2C). We thus set out to crystallize the protein for X-ray diffraction analysis (see Material and methods). Testing 2016 different conditions, crystals were only obtained in individual cases after at least 7 days of incubation. A complete dataset was collected from a single crystal diffracting to 2.6 Å resolution and a $P4_32_12$ symmetry. Data and refinement statistics are given in Table S2. Surprisingly, the unit cell dimensions were small and, with a Matthews coefficient assuming 50% solvent content, only 128 amino acids would fit into the asymmetric unit of the crystal. Hence, the unit cell had an insufficient size to cover H-Rrm4-NT4, which contains 380 amino acids. Using the predicted models of MLLE1-3^{Rrm4} as templates for molecular replacement, only MLLE2^{Rrm4} gave a clear solution, showing after refinement that two copies of MLLE2^{Rrm4} (residues 567-630) were present in the asymmetric unit. For comparison, previously, two copies of the MLLE domain in the asymmetric unit were reported in crystals of MLLE of UBR5 (MUNOZ-ESCOBAR *et al.* 2015). The structural data indicated that the protein was truncated from both termini during crystallization, resulting in a shortened version of the H-Rrm4-NT4 protein that formed stable crystals (see Material and methods). Both MLLE2^{Rrm4} copies adopted the same overall fold as seen by the RMSD of 0.29 Å over 59 C-alpha atoms. The MLLE2^{Rrm4} crystal structure displayed high similarity with the MLLE domain of the ubiquitin ligase UBR5 (MLLE^{UBR5}; PDB ID: 3NTW, RMSD of 0.97 Å over 56 amino acids) (MUNOZ-ESCOBAR *et al.* 2015) and the MLLE domain of


A
α4
α3
α2
R573
α5
E591
MLLE2
B
C
L586
E564
G563
F118
V119
P120
T582
F567
HsPABC1
L493
S471
V496
G472
F118
V119
T475
P120
MLLE1
A598
R573
E574
F118
V119
P120
T594
L577
MLLE2
L759
G736
D737
F118
V119
P120
T755
F740
MLLE3
D
log I (s)
s [nm-1]
G-Rrm4
Gst
RRM1-3
MLLE1
MLLE2
MLLE3
E
log I (s)
s [nm-1]
H-Rrm4-NT4
MLLE2
MLLE1
MLLE3

2.4.2.1 Figure 7. Rrm4 contains a C-terminal tripartite MLLE binding platform.

(A) Crystal structure of the MLLE2^{Rrm4} domain is highlighted in orange (PDB ID: 7PZE). The four helices are indicated by α2-5 according to the 5 helix nomenclature found in MLLE domains (XIE *et al.* 2014a). Note that the first short helix a1 is missing. Arg573 and Glu591 are highlighted in the sticks. These side chains would interfere with binding the canonical Phe of PAM2-type motifs. (B) Structural alignment of the MLLE2^{Rrm4} model generated by TopModel and the X-ray crystal structure of this domain (grey or orange, respectively). The all-atom RMSD is 0.69 Å, resulting mostly from different rotamers of solvent-exposed sidechains. (C) Comparison of peptide-binding sites after structural alignment of the models of Rrm4 MLLE domains (orange shades) and the canonical MLLE domain of HsPABPC1 (blue; PDB ID: 3KUS) and manually placing the PAM2 motif of PAIP2 (lilac). In the interaction of MLLE^{PABPC1} with PAM2 of PAIP2, Phe118 of PAM2 is the major determinant for binding and is present in all the PAM2 motifs except LARP4a and b ((KOZLOV AND GEHRING 2010; XIE *et al.* 2014); S3A Figure). Of the identified Rrm4 MLLE domains, only MLLE3^{Rrm4} retains all sidechains that favor the binding of this characteristic Phe; particularly, Gly736 should allow the Phe to bind into a pocket. MLLE1^{Rrm4} and MLLE2^{Rrm4} have Ser471 and Arg573 instead of G in this position, suggesting that Phe binding would be sterically hindered in these interfaces. (D) *The left panel* Experimental data curve for GST-Rrm4 is shown in black dots with grey error bars, and the EOM fits as a red line ($\chi 2$ =1.289). The intensity is displayed as a function of momentum transfers. *Right panel* Selected model of the EOM analysis from GST-Rrm4 with an R_g of 8.75 nm, a D_{max} of 23.99 nm with a volume fraction of~0.25 (SASBDB ID: SASDMS5). (E) *The left panel* Experimental data curve for H-Rrm4-NT4 is shown in black dots with grey error bars; the EOM fit as the red line ($\chi 2$ =1.262). The intensity is displayed as a function of momentum transfers. *Right panel* Selected model of the EOM analysis from H-Rrm4-NT4 with an R_g of 5.10 nm, a D_{max} of 16.43 nm, and a volume fraction of~0.75 (SASBDB ID: SASDMT5). The MLLE^{Rrm4} subdomains are shown in cartoon representation (MLLE1^{Rrm4} in light orange, MLLE2^{Rrm4} in orange, MLLE3^{Rrm4} in dark orange, and the GST in dark grey), and the missing amino acids as grey spheres (all other models and the SAXS data are available in S2E Figure).

PABPC1 (PDB ID: 3KUS, RMSD of 1.34 Å over 61 amino acids) (KOZLOV *et al.* 2010). The MLLE2^{Rrm4} domain consisted of four helices (designated α2 - 5; Figure 7A, (PDB ID: 7PZE) arranged as a right-handed superhelix similar to MLLE^{UBR5}. In comparison to the MLLE domain of PABPC1, the first short helix was absent in both MLLE2^{Rrm4} and MLLE^{UBR5} structures.

When comparing the obtained crystal structure with the MLLE2 ^{Rrm4} model generated by TopModel, the average RMSD was 0.69 Å over the backbone atoms, close to the uncertainty of the atomic coordinates of the experimental structure (Figure 7B). Importantly, this confirmed our structural model of MLLE2^{Rrm4} and strongly suggested that the modelled MLLE1^{Rrm4} and MLLE3^{Rrm4} domains should be of equally high quality.

We compared the predicted models of MLLE1-3^{Rrm4} with the known structure of the human PABPC1 focusing on the well-described PAM2 peptide-binding pocket. This revealed that MLLE3^{Rrm4} maintained a characteristic Gly residue at position 736 that binds the conserved Phe residue of the PAM2 motifs, a major binding determinant in PABPC1 and UBR5 (Figure 7C) (KOZLOV AND GEHRING 2010; MUNOZ-ESCOBAR *et al.* 2015). However, the binding interfaces of MLLE1^{Rrm4} and MLLE2^{Rrm4} were altered compared to the 'canonical' binding site in PABPC1 and UBR5 (Figure 7C). Instead of Gly,

MLLE1^{Rrm4} and MLLE2^{Rrm4} had a Ser and Arg in the corresponding positions 471 and 573. The notion that MLLE1^{Rrm4} and MLLE2^{Rrm4} may differ from canonical MLLE domains was also supported by the lower sequence identity of MLLE1^{Rrm4} and MLLE2^{Rrm4} when compared to previously characterized MLLE domains (Figure 6B; Figure S1A). In summary, structural modelling revealed the presence of three MLLE domains at the C-terminus of Rrm4. Furthermore, the structure of the MLLE2^{Rrm4} domain was successfully verified by X-ray crystallographic analysis. MLLE1^{Rrm4} and MLLE2^{Rrm4} are divergent in the key region of PAM2 binding, suggesting that these domains might employ a different binding mode or show a different binding specificity.

2.4.3 The MLLE domains of Rrm4 form a binding platform with flexible arrangement of the individual domains

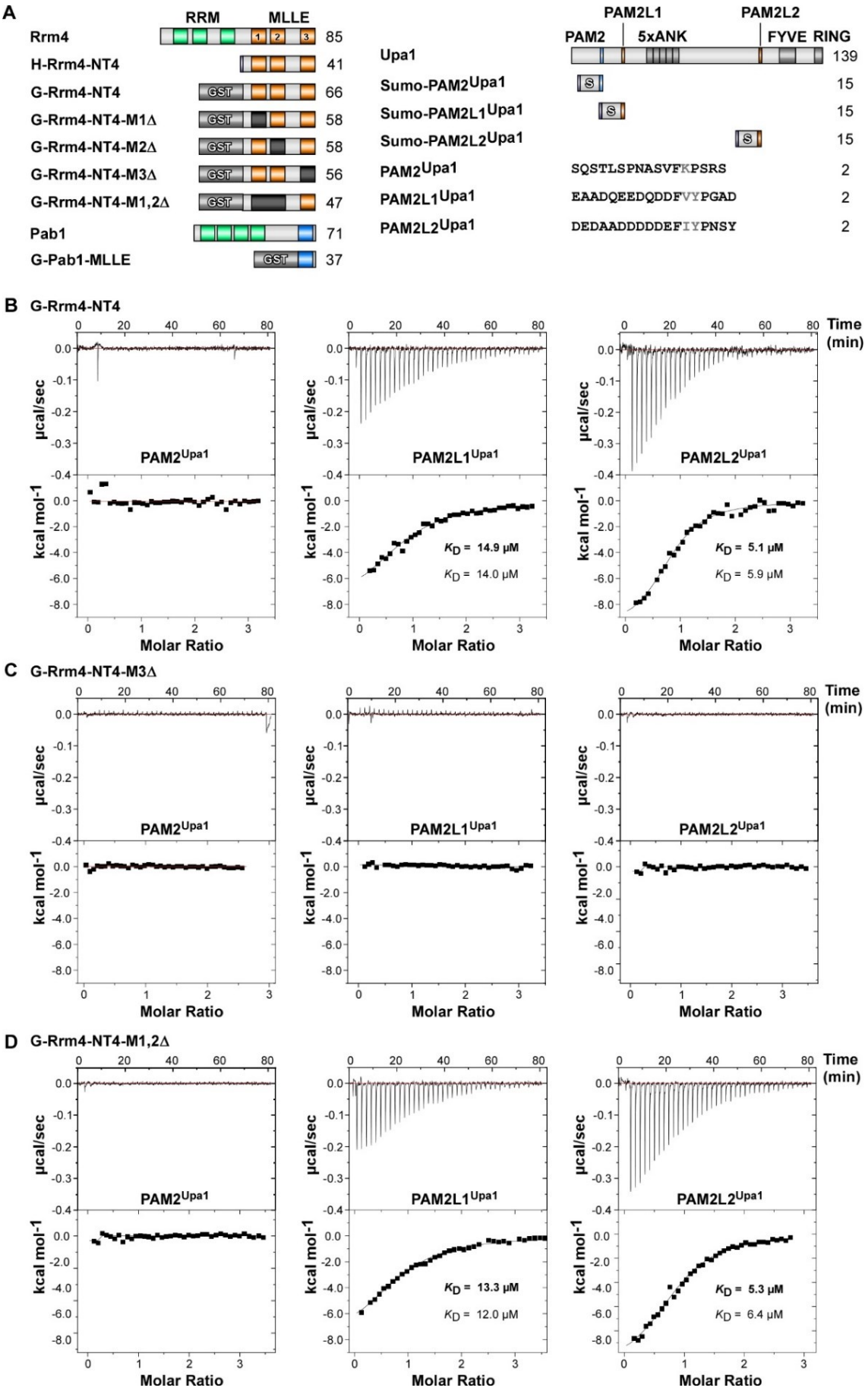
To study the relative arrangement of all three MLLE^{Rrm4} domains and the orientation to the N-terminal RRMs, we performed Small-Angle X-ray Scattering (SAXS) experiments. We expressed and purified H-Rrm4-NT4 as well as the full-length protein with N-terminal GST fusion (glutathione S-transferase; G-Rrm4) from *E. coli* (see Materials and methods). Primary data analysis of the scattering curves (KONAREV *et al.* 2003; TRIA *et al.* 2015) revealed that both proteins were monomeric and highly flexible in solution (Table S3; Figures 7D-E; S2D-E). To visualize the different protein conformations, we performed an Ensemble Optimization Method (EOM) analysis for both the G-Rrm4 and H-Rrm4-NT4 proteins (Figure 7D-E, SASBDB ID: SASDMS5, SASDMT5). We used our MLLE models and a GST model (PDB ID: 1UA5) together with the protein sequence for G-Rrm4 as input, yielding a distribution of different conformations of the protein in solution (representative models in Figure S2E). One model of G-Rrm4, representing 25% of the population, revealed that the C-terminal part containing MLLE1-3^{Rrm4} adopted an elongated and mainly unfolded but open conformation (Figures 7D; S2D-E). The N-terminal part, containing RRM domains of the GST fusion protein, adopted a more globular structure, indicating less flexibility within this region (Figures 7D; S2D-E). Studying only the C-terminal part of Rrm4 revealed that the most prominent model of this analysis (75 % of the population) had a nearly identical conformation as the one selected for G-Rrm4 (Figures 7D-E, S2E). This suggests that the C- terminal part of Rrm4-NT4 adopts a very similar orientation when expressed by itself. This analysis deduced that the MLLE1-3^{Rrm4} domains form a C-terminal binding platform with a flexible arrangement for multiple

contact sites for binding partners. Thus, the RRM domains for RNA interaction are spatially separated from the protein interaction platform.

2.4.4 The third MLLE is essential for interaction with PAM2-like sequences of Upa1

To evaluate the interaction capacity of MLLE1-3^{Rrm4}, we performed *in vitro* binding studies. We expressed different deletion versions of Rrm4 as N-terminal GST fusions in *E. coli*. As a control, we expressed an N-terminal GST fusion of the MLLE domain of Pab1 (Figure 8A; Materials and methods). To check the physical interaction with PAM2 and PAM2L sequences of Upa1, we expressed 18 amino acid fragments (Figure 8A) as N-terminal hexa-histidine-SUMO (HS) fusion proteins (see Materials and methods). In GST pull-down experiments using GST fusion proteins as bait, G-Pab1-MLLE interacted with HS-PAM2 but not with the HS-PAM2L motifs of Upa1 (Figure S3B, lane 2). Conversely, G-Rrm4-NT4 recognized the two HS-PAM2L motifs of Upa1 but not the HS-PAM2 motif (Figure S3 B, lane 3) (POHLMANN *et al.* 2015). Interestingly, the interaction with both PAM2L motifs was lost when MLLE3^{Rrm4} was deleted (G-Rrm4-NT4-M3Δ; Figure S3B, lane 6), while constructs with deletion of MLLE1^{Rrm4} or MLLE2^{Rrm4}, or both MLLE1^{Rrm4} and MLLE2^{Rrm4}, still interacted with the HS-PAM2L motifs of Upa1 (Figure S3B, lane 4,5 and 7).

To validate qualitatively whether these results also hold true for full-length proteins, we performed yeast two-hybrid experiments comparable to previous studies (POHLMANN *et al.* 2015). To this end, Upa1 or Rrm4 versions were fused at the N-terminus with the DNA-binding domain (BD) and activation domain (AD) of Gal4p, respectively (see Materials and methods; the C-termini were fused with the enhanced version of the green fluorescent protein [Gfp], Clontech; or the monomeric version of red fluorescent protein mKate2 [Kat], respectively) (POHLMANN *et al.* 2015; MÜNTJES *et al.* 2020). Rrm4-Kat interacted with full-length Upa1-Gfp (Figure S4A) (POHLMANN *et al.* 2015) and interaction was lost when MLLE3^{Rrm4} was deleted. Mutations in MLLE1^{Rrm4}, MLLE2^{Rrm4} or MLLE1,2^{Rrm4} did not alter the interaction with Upa1-Gfp (Figure S4B-D). To further investigate the presence of unknown interaction motifs in Upa1-Gfp, variants carrying block mutations in either or both PAM2L1 and PAM2L2 motifs were tested against the Rrm4-Kat versions (Figure S4B-C). Upa1-Gfp versions with block mutations in either PAM2L1 or PAM2L2 still interacted with the Rrm4-Kat versions (Figure S4D).


A
RRM
MLLE
Rrm4
85
H-Rrm4-NT4
41
G-Rrm4-NT4
66
G-Rrm4-NT4-M1Δ
58
G-Rrm4-NT4-M2Δ
58
G-Rrm4-NT4-M3Δ
56
G-Rrm4-NT4-M1,2Δ
47
Pab1
71
G-Pab1-MLLE
37
PAM2L1
PAM2L2
PAM2
5xANK
FYVE
RING
Upa1
139
Sumo-PAM2Upa1
15
Sumo-PAM2L1Upa1
15
Sumo-PAM2L2Upa1
15
PAM2Upa1
SQSTLSPNASVFKPSRS
2
PAM2L1Upa1
EAADQEEDQDDFVYPGAD
2
PAM2L2Upa1
DEDAADDDDDEFIYPNSY
2
B G-Rrm4-NT4
C G-Rrm4-NT4-M3Δ
D G-Rrm4-NT4-M1,2Δ
µcal/sec
kcal mol-1
Molar Ratio
Time (min)
PAM2Upa1
PAM2L1Upa1
PAM2L2Upa1
KD = 14.9 µM
KD = 14.0 µM
KD = 5.1 µM
KD = 5.9 µM
KD = 13.3 µM
KD = 12.0 µM
KD = 5.3 µM
KD = 6.4 µM

2.4.4.1 Figure 8. MLLE3^{Rrm4} is crucial for PAM2L1^{Upa1} and PAM2L2^{Upa1} binding.

(A) Schematic representation of protein variants (Molecular weight in kilo Dalton indicated) using the following coloring: lime green, RNA recognition motif (RRM); orange, MLLE^{Rrm4} domains; dark blue, MLLE^{Pab1}; light blue PAM2^{Upa1}; light orange PAM2L^{Upa1} sequence (PL1 – 2). Ankyrin repeats (5xANK), FYVE domain, and RING domain of Upa1 are in dark grey. GST and SUMO tags are labeled. Variant amino acids of the FxP and FxxP of PAM2 and PAM2L sequences are printed in grey font. (B-D) Representative isothermal titration calorimetry (ITC) binding curves of MLLE domains. Experiments were performed using GST- or Histidine-tagged MLLE variants and synthetic PAM2^{Upa1} peptide variants. K_D values of two independent measurements are given (values corresponding to the indicated data are in bold).

However, when both PAM2L1,2 motifs were mutated, the interaction between the Upa1 and Rrm4 was lost, comparable to earlier observation (Figure S4D) (POHLMANN *et al.* 2015). Invariably, MLLE3^{Rrm4} deletion caused loss of interaction with all Upa1-Gfp versions (Figure S4A-D). These results confirm that the MLLE3^{Rrm4} domain is essential for the interaction with Upa1. MLLE1^{Rrm4} and MLLE2^{Rrm4} appear to be dispensable for the interaction with Upa1 suggesting the presence of additional interaction partners (see below).

To obtain quantitative data on the protein/peptide interactions, we performed isothermal titration calorimetry (ITC) experiments with purified proteins (Figure S5A) and synthetic peptides with a length of 18 amino acids (PAM2^{Upa1}, PAM2L1^{Upa1}, and PAM2L2^{Upa1}; Figure 8A). The binding constant K_D and the binding stoichiometry were calculated from the curves, which in all cases indicated a 1:1 ratio between the G-Rrm4-NT4 protein and the binding partner.

Testing G-Pab1-MLLE with the peptides revealed a K_D of 14.6 µM for PAM2^{Upa1} (Figure S5B), which is within the range of observed K_D of 0.2 to 40 µM for known MLLE/PAM2 interactions like the MLLE domain of PABPC1 with various PAM2 sequences (MATTIJSSEN *et al.* 2021). Testing G-Pab1-MLLE with PAM2L1^{Upa1} and PAM2L2^{Upa1} peptides, no indication for binding was observed. PAM2L^{Upa1} sequences are rich in acidic residues and exhibit a different FxxP spacing than the canonical FxP sequence of PAM2 sequences in the core region (Figure 8A, S3A; see Discussion). The observed binding behavior indicated a clear binding specificity differentiating PAM2 and PAM2L peptides. This was in line with the previously published GST pull-down experiments (POHLMANN *et al.* 2015).

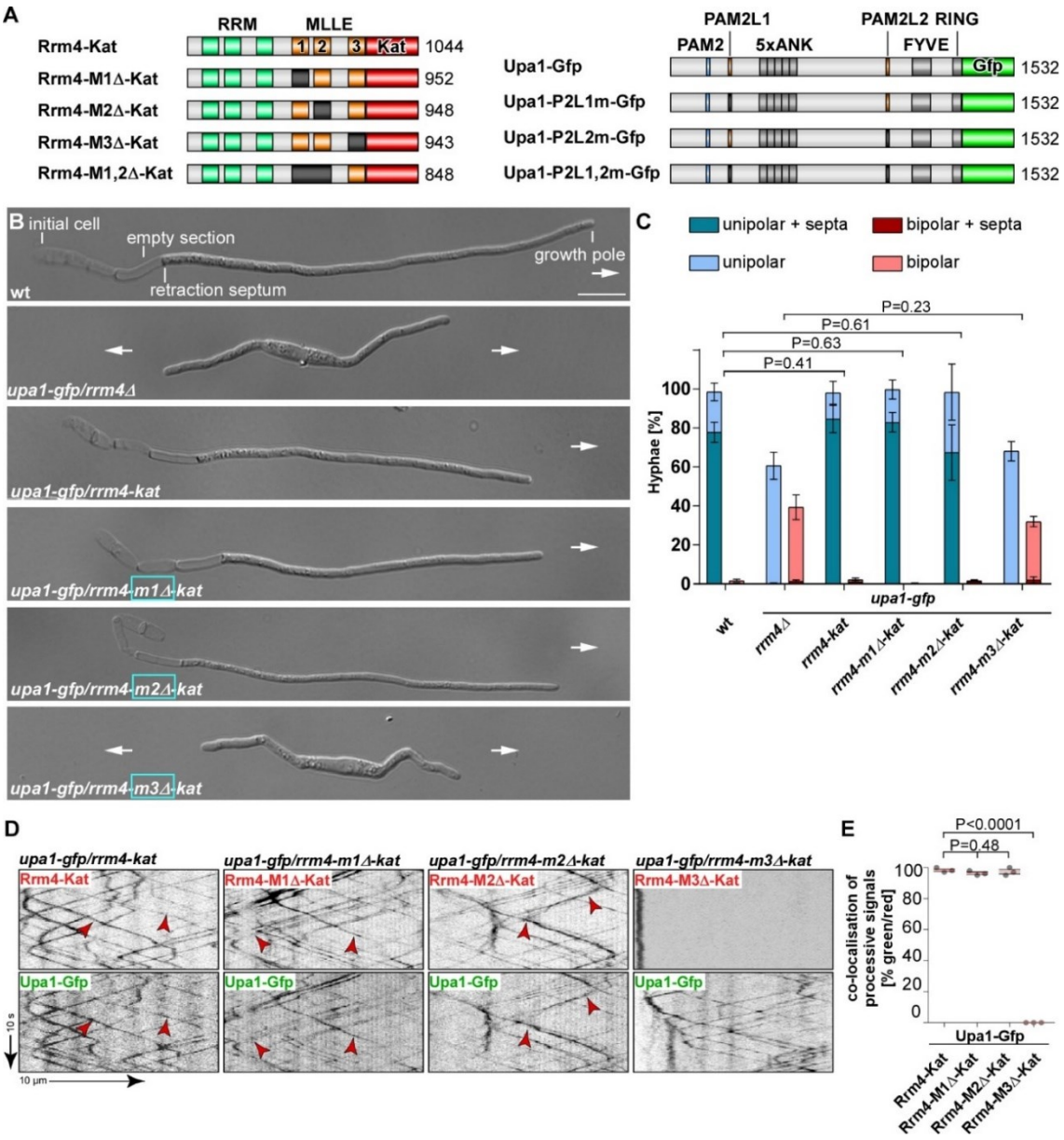
In comparison, testing G-Rrm4-NT4 with the peptides revealed a K_D of 14.9 µM for PAM2L1^{Upa1} and 5.1 µM for PAM2L2^{Upa1} and no binding to PAM2^{Upa1} (Figure 8B). This suggested a similar affinity when compared to the interactions of MLLE^{Pab1} with PAM2 and demonstrated the high sequence specificity of the MLLE domains to their respective PAM2L sequences (see Discussion).

Analyzing G-Rrm4-NT4-M3Δ with a deletion of MLLE3^{Rrm4} revealed that binding to PAM2L1^{Upa1} and PAM2L2^{Upa1} was no longer detectable (Figure 8C). This was in line with observations from the GST pull-down experiments (Figure S3B). This suggests that MLLE3^{Rrm4} is essential for binding. Testing G-Rrm4-NT4 versions carrying deletions in either MLLE1^{Rrm4} or MLLE2^{Rrm4} showed no difference in binding affinity (Figure S6A, B). Even testing G-Rrm4-NT4 with a deletion in both MLLE1,2^{Rrm4} domains exhibited a binding affinity in the same range as the wild type version containing all three MLLEs (Figure 8D). We conclude that (i) MLLE3^{Rrm4} is vital for recognizing PAM2L sequences with a higher affinity to PAM2L2 and (ii) neither MLLE1^{Rrm4} nor MLLE2^{Rrm4} contributed to the binding of PAM2L or PAM2 motifs (Figure. 8, S3-6; results summarized in Figure S6D, see Materials and methods). This is consistent with our structural analysis revealing the differences in the binding site for these MLLE domains (see Discussion).

2.4.5 The third MLLE domain of Rrm4 is essential for its function

To address how the different MLLE domains, contribute to the biological function of Rrm4, we generated *U. maydis* strains carrying deletions in the respective domains of Rrm4 (Figure 9A). As genetic background, we used laboratory strain AB33, expressing the heteromeric master transcription factor of hyphal growth (bE/bW) under control of the nitrate inducible promoter P$_{nar1}$. Thereby, polar hyphal growth can be elicited efficiently and in a highly reproducible fashion by changing the nitrogen source (Figure 9B, top) (BRACHMANN *et al.* 2001). To investigate dynamic endosomal transport, we used strains expressing functional C-terminal fusion Upa1-Gfp and Rrm4-Kat (see Materials and methods).

The resulting hyphae grew with a defined axis of polarity, i.e., they expanded at the hyphal tip and inserted basal septa leading to the formation of empty sections (Figure 9B-C). Loss of Rrm4 (*rrm4Δ* strain) caused the formation of hyphae growing at both ends, characteristic of aberrant bipolar growth (Figure 9B-C) (POHLMANN *et al.* 2015). Rrm4-Kat versions carrying deletions of MLLE1^{Rrm4} or MLLE2^{Rrm4} did not cause alterations in unipolar growth (Figure 9B-C). Furthermore, endosomal shuttling and co-localization were indistinguishable from wild type (Figure 9D-E). Also, the number of endosomes (number of signals / 10 µM, Figure S9A), velocity, and processivity (Figure S9B--C) were comparable to wild type. Hence, the first two MLLE^{Rrm4} domains were dispensable for polar growth and endosomal shuttling under optimal growth conditions. Since the deletion of the first two MLLEs^{Rrm4} did not substantially alter the function of Rrm4, we



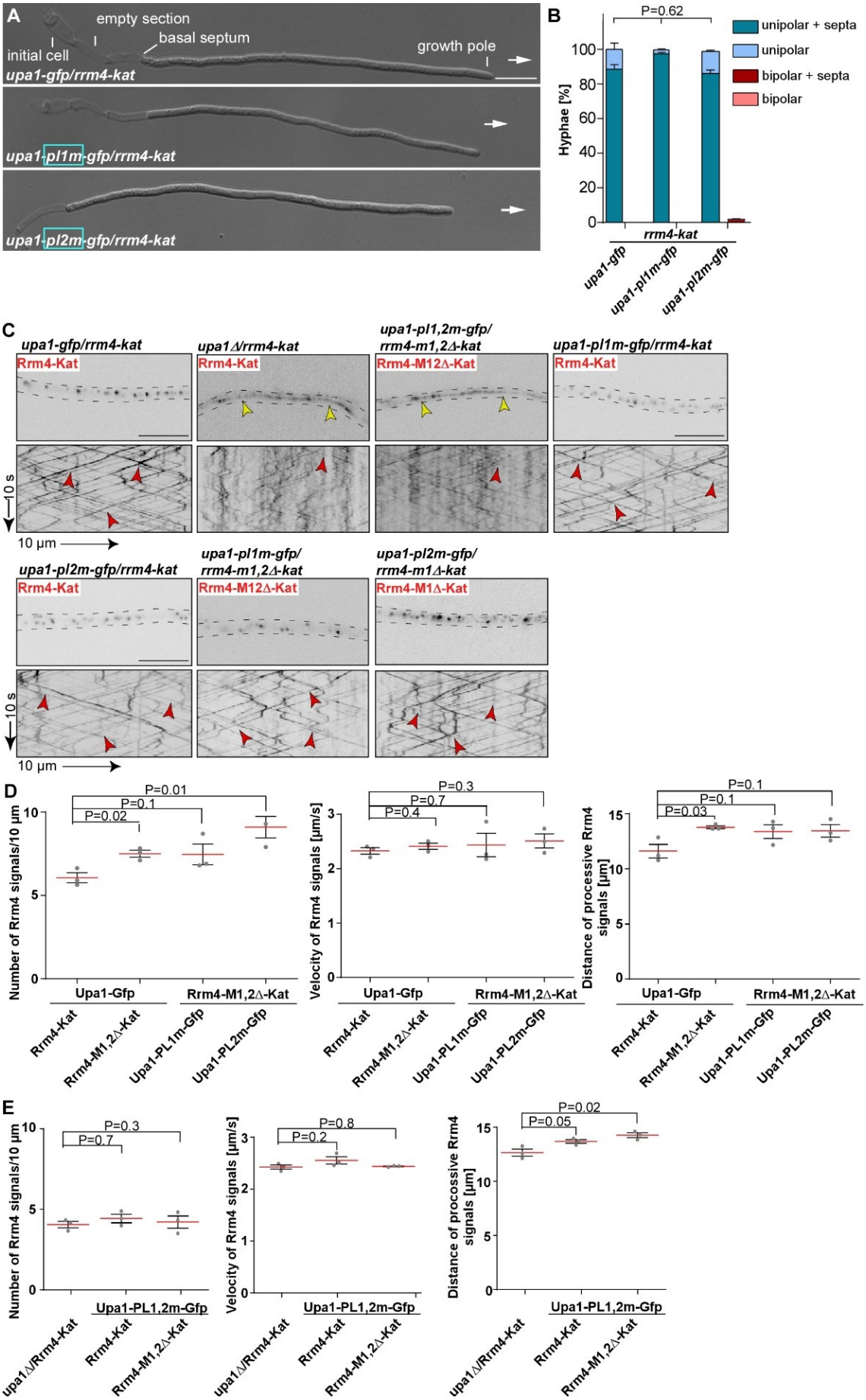
2.4.5.1 Figure 9. MLLE3 is key for endosomal mRNA transport.

(A) Schematic representation of Rrm4 and Upa1 variants drawn not to scale (number of amino acids indicated next to protein bars) using the following coloring: lime green, RNA recognition motif (RRM); orange, $MLLE^{Rrm4}$ domains; red, mKate2, blue, $PAM2^{Upa1}$, light orange $PAM2^{Upa1}$-like sequence (PL1 – 2) and light green, Gfp. Ankyrin repeats (5xANK), FYVE domain, and RING domain of Upa1 are given in dark grey. (B) Growth of AB33 derivatives in their hyphal form (6 h.p.i.; size bar 10 µM). Arrows mark the growth direction. (C) Quantification of hyphal growth of AB33 derivatives shown in B (6 h.p.i.): unipolarity, bipolarity, and basal septum formation were quantified (error bars, SEM.; n = 3 independent experiments, > 100 hyphae were counted per strain; for statistical evaluation, the percentage of uni- and bipolarity was investigated by using unpaired two-tailed Student's t-test ($a<0.05$). (D) Kymographs of AB33 hyphae derivatives (6 h.p.i.; inverted fluorescence images) expressing pairs of red and green fluorescent proteins as indicated. Fluorescence signals were detected simultaneously using dual-view technology (arrow length on the left and bottom indicate time and distance, respectively). Red arrowheads mark processive co-localizing signals. (E) Percentage of processive signals exhibiting co-localization for strains shown in D (data points represent means from n = 3 independent experiments, with mean of means, red line, and SEM; unpaired two-tailed Student's t-test ($a<0.05$); for each experiment, 10 hyphae per strains were analyzed).

infer that the deletion neither affected the overall structure of the protein nor interfered with other domains like the RNA-binding domain of the protein. This supports the conclusions of our biochemical experiments (see above). Importantly, testing strains expressing Rrm4-Kat with deletion of the third MLLE^{Rrm4} domain revealed a loss-of-function phenotype similar to *rrm4Δ* strains. The number of bipolar hyphae was comparable to *rrm4Δ* strains (Figure 9B-C; mutation identical to allele *rrm4G$^{P\Delta}$*) (BECHT *et al.* 2006). We observed drastic alteration in shuttling, and Rrm4 aggregates did not co-localize with motile Upa1-positive signals (Figure 9D-E). While the Rrm4 signals were static (Figures 9D-E; S7A-C), the number of motile Upa1-Gfp positive endosomes, their velocity, and their processivity were not affected (Figure S7A-C, summarized in S7D). This is consistent with previous results showing that the third MLLE^{Rrm4} domain is important for the movement of Rrm4 and that endosomal shuttling of Upa1 is not affected if Rrm4 is missing (BECHT *et al.* 2006; BAUMANN *et al.* 2012; POHLMANN *et al.* 2015). To conclude, MLLE3^{Rrm4} is an essential domain for Rrm4 attachment to endosomes in contrast to MLLE1^{Rrm4} and MLLE2^{Rrm4}.

2.4.6 The second MLLE domain plays accessory roles in endosomal Rrm4 attachment

To investigate the biological role of MLLE1^{Rrm4} and MLLE2^{Rrm4} in more detail, we generated strains expressing Rrm4-M1,2Δ-Kat, lacking both MLLE1^{Rrm4} and MLLE2^{Rrm4} domains, and tested the influence on hyphal growth. Unipolar growth was not disturbed (Figure 10A-B). To challenge the endosomal attachment of Rrm4, we expressed Upa1-Gfp versions carrying mutations in PAM2L motif 1 or 2 as well as in both motifs; these motifs are important for Rrm4 interaction (Figure 9A) (POHLMANN *et al.* 2015). Strains expressing Rrm4-M1,2Δ-Kat in combination with mutated PAM2L1 or PAM2L2 of Upa1 showed unipolar growth comparable to wild type (Figures 10A-B; S8A-B), indicating that MLLE1^{Rrm4} and MLLE2^{Rrm4} were dispensable for unipolar growth even when the endosomal attachment was weakened by expressing Upa1 versions with mutated PAM2L motifs (Figure S8D). When studying Upa1 mutated in both PAM2L motifs, we observed an aberrant bipolar growth phenotype comparable to the *upa1Δ* strain (Figure 10A-B). This was expected, since the interaction of Rrm4 to endosomes is mediated by both PAM2L motifs (POHLMANN *et al.* 2015). Analysing Rrm4-M1,2Δ-Kat in this genetic


A
empty section
basal septum
initial cell
growth pole
upa1-gfp/rrm4-kat
upa1-pl1m-gfp/rrm4-kat
upa1-pl2m-gfp/rrm4-kat
B
P=0.62
unipolar + septa
unipolar
bipolar + septa
bipolar
Hyphae [%]
rrm4-kat
upa1-gfp
upa1-pl1m-gfp
upa1-pl2m-gfp
C
upa1-gfp/rrm4-kat
Rrm4-Kat
upa1Δ/rrm4-kat
upa1-pl1,2m-gfp/ rrm4-m1,2Δ-kat
Rrm4-M12Δ-Kat
upa1-pl1m-gfp/rrm4-kat
upa1-pl2m-gfp/rrm4-kat
upa1-pl1m-gfp/ rrm4-m1,2Δ-kat
upa1-pl2m-gfp/ rrm4-m1Δ-kat
Rrm4-M1Δ-Kat
10 s
10 µm
D
Number of Rrm4 signals/10 µm
P=0.01
P=0.1
P=0.02
Velocity of Rrm4 signals [µm/s]
P=0.3
P=0.7
P=0.4
Distance of processive Rrm4 signals [µm]
P=0.1
P=0.1
P=0.03
Upa1-Gfp
Rrm4-M1,2Δ-Kat
Rrm4-Kat
Rrm4-M1,2Δ-Kat
Upa1-PL1m-Gfp
Upa1-PL2m-Gfp
E
Number of Rrm4 signals/10 µm
P=0.3
P=0.7
Velocity of Rrm4 signals [µm/s]
P=0.8
P=0.2
Distance of processive Rrm4 signals [µm]
P=0.02
P=0.05
Upa1-PL1,2m-Gfp
upa1Δ/Rrm4-Kat
Rrm4-Kat
Rrm4-M1,2Δ-Kat

2.4.6.1 Figure 10. MLLE2 plays an accessory role in endosomal attachment of Rrm4.

(A) Growth of AB33 derivatives in their hyphal form (6 h.p.i.; size bar 10 µM). Arrows mark the growth direction. (B). Quantification of hyphal growth of AB33 derivatives shown in A (6 h.p.i.): unipolarity, bipolarity, and basal septum formation were quantified (error bars, SEM; n = 3 independent experiments, > 100 hyphae were analyzed per strain; For statistical analysis, the percentage of uni- and bipolarity was investigated by using unpaired two-tailed Student's t-test ($a<0.05$). (C) Micrographs (inverted fluorescence image; size bar, 10 µM) and corresponding kymographs of AB33 hyphae derivatives (6 h.p.i.) co-expressing various Upa1-Gfp and Rrm4-Kat versions as indicated. Movement of Rrm4-Kat versions is shown (arrow length on the left and bottom indicates time and distance, respectively). Bidirectional movement is visible as diagonal lines (red arrowheads). A yellow arrowhead indicates aberrant microtubule staining. (D) Percentage of hyphae (6 h.p.i.) exhibiting aberrant microtubule association as indicated in panel C and Figure S8C. The set of strains that were analyzed simultaneously is shown in the same yellow shading (error bars, SEM; for statistical evaluation, the percentage of hyphae with abnormal microtubule staining was compared by using unpaired two-tailed Student's t-test ($a<0.05$); n = 3 independent experiments, > 25 hyphae were analyzed per strain). (E) Normalized minimum and maximum grey level intensities of shuttling signals measured in Rrm4 kymographs shown in Figures 5C and S5C (error bars, SEM; n = 3 independent experiments, 100 shuttling signals kymographs were analyzed per strain, two-tailed Student's t-test ($a > 0.05$).

background revealed no additive phenotype (Figure 10A-B). This reinforces that the interactions of PAM2L motifs of Upa1 are the major determinants for endosomal attachment of Rrm4.

Next, we investigated endosomal shuttling. In strains expressing Rrm4-M1,2Δ-Kat missing MLLE1^{Rrm4} and MLLE2^{Rrm4} endosomal shuttling was not disturbed (Figure 10C).

The number of motile Rrm4-M1,2Δ-Kat positive signals, their velocity, and their processivity were not affected (Figure S8C-D). Like above, we challenged the endosomal attachment of Rrm4 by expressing Upa1 versions with mutations in the PAM2L motifs. As expected, simultaneous mutation of both PAM2L motifs of Upa1 resulted in a reduction in the number of Rrm4-Kat positive shuttling endosomes ((Figure 10C, S8E; POHLMANN *et al.* 2015). When both PAM2L motifs were mutated, the Rrm4-Kat version was mislocalized and exhibited aberrant staining of filamentous structures in about 80% of hyphae (Figure 10C-D). This staining pattern was reminiscent of the microtubule association of Rrm4 previously observed during altered accumulation of static Rrm4-Kat in upa1Δ strains ((Figure S8C; (JANKOWSKI *et al.* 2019)). Quantifying Rrm4-Kat signals exhibiting processive movement in kymographs revealed that strains exhibiting aberrant staining of filamentous structures resulted in reduced fluorescence (Figure 10E) indicating fewer Rrm4-Kat versions on shuttling endosomes. As an important control, we treated the strains with the microtubule inhibitor benomyl, demonstrating that aberrant staining was microtubule-dependent (Figure S9A). Furthermore, Western blot analysis demonstrated that mutations in Rrm4 do not alter the protein amount (Figure S9B). Comparable to previous reports, we observed residual motility of Rrm4-Kat on shuttling endosomes if both

PAM2L motifs were mutated or if upa1 was deleted (Figure 10C). This indicates additional proteins besides Upa1 are involved in the endosomal attachment of Rrm4 (POHLMANN *et al.* 2015).

To analyze the influence of individual PAM2L motifs, we determined the number of hyphae with aberrant microtubule staining in strains co-expressing Rrm4-Kat versions and an Upa1-Gfp version with mutations of PAM2L sequence 1 or 2. Mutations in PAM2L1 and PAM2L2 caused 8% and 19% of hyphae with aberrant MT staining, respectively (Figure 10D). Hence, the interaction of PAM2L2 is more important for correct endosomal attachment of Rrm4. This is consistent with our biochemical results demonstrating that MLLE3^{Rrm4} binds stronger to PAM2L2 of Upa1 than to PAM2L1 (Figure 8B).

Next, we investigated the association of Rrm4-M1,2Δ-Kat in strains expressing Upa1 with mutated PAM2L1. In this strain, the endosomal attachment was solely dependent on the interaction of MLLE3^{Rrm4} with the PAM2L2 sequence of Upa1. We did observe 6% of hyphae with aberrant MT staining (Figures 10D; S7D). This was comparable to strains expressing Rrm4-Kat, suggesting no clear difference (Figure 10D). However, testing Upa1 with its PAM2L2 mutated, leaving only PAM2L1 for interaction with Rrm4, we observed a clear increase in hyphae with aberrant MT staining when comparing strains co-expressing Rrm4-M1,2Δ-Kat versus Rrm4-Kat (52% versus 19%; Figure 10D-E). Hence, the region covering MLLE1^{Rrm4} and MLLE2^{Rrm4} was important for Rrm4 attachment. Finally, we tested individual deletions in MLLE1^{Rrm4} and MLLE2^{Rrm4} in combination with mutated PAM2L2 in Upa1 to dissect the role of the different MLLE^{Rrm4} domains. In strains expressing Rrm4-Kat or Rrm4-M1Δ-Kat with this type of Upa1 mutation, the number of hyphae with aberrant MT staining was comparable (18% versus 11%, respectively; Figure 10D). However, strains expressing Rrm4-M2Δ-Kat exhibited an increased number of hyphae with aberrant MT staining that was comparable to Rrm4-M1,2Δ-kat (51% versus 52% respectively; Figure 10D; S7B). As mentioned above, aberrant MT localization of mutated Rrm4-M2Δ-Kat and Rrm4-M1,2Δ-Kat also exhibited reduced intensity of processive signals in Rrm4 kymographs (Figure 10E) suggesting that the endosomal association was altered. To conclude, for MLLE1^{Rrm4}, we were unable to assign a clear function yet. However, MLLE2^{Rrm4} plays an accessory role in the endosomal attachment of Rrm4. In essence, the C-terminus of Rrm4 contains three MLLE domains, with MLLE2^{Rrm4} fulfilling an accessory role and MLLE3^{Rrm4} having an essential function during the attachment of mRNPs to endosomes.

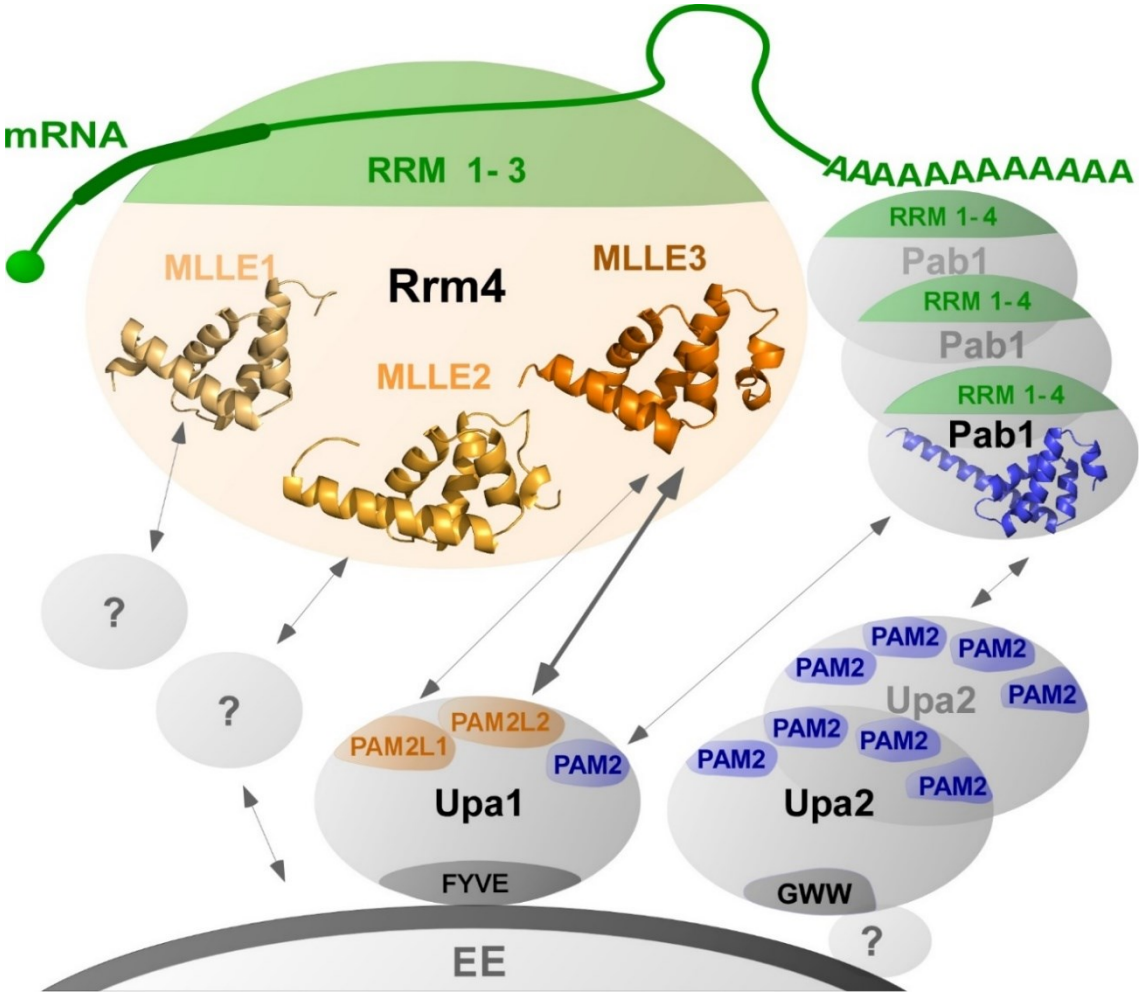
2.5 Discussion

Combining structural biology and biophysical techniques with fungal genetics and cell biology, we addressed how mRNPs can be mechanistically linked to endosomes in the model fungus *U. maydis*. Previously, it was found that the C-terminal MLLE domain of Rrm4 is needed for shuttling (BECHT *et al.* 2006) and that the C-terminus of Rrm4 interacts with two PAM2L motifs of Upa1 (POHLMANN *et al.* 2015). Now, we demonstrate that this region of Rrm4 contains not only two MLLE^{Rrm4} domains, but a sophisticated binding platform consisting of three MLLE^{Rrm4} domains with MLLE2^{Rrm4} and MLLE3^{Rrm4} functioning in linking the key RNA transporter to endosomes. We disclose a strict hierarchy with main and accessory domains. The accessory MLLE2^{Rrm4} domain shows variations in the critical region of the predicted PAM2 binding pocket, suggesting a novel mode of interaction with currently unknown interaction partners. Rrm4 represents the first protein containing multiple MLLE domains to form a binding platform to the best of our knowledge. This interaction unit is essential for the correct endosomal attachment and, hence, mRNP trafficking.

2.5.1 The MLLE/PAM2 connection

The founding member of the MLLE domain family is present at the C-terminus of the poly(A)-binding protein PABPC1. This domain interacts with PAM2 motifs of numerous interaction partners such as GW182, eRF3, and the RNA-binding protein LARP4 functioning in microRNA biology, translational termination, and posttranscriptional control, respectively (JINEK *et al.* 2010; KOZLOV AND GEHRING 2010; YANG *et al.* 2011). Structural analysis revealed a common mode of binding, where the Leu and particularly the Phe of the PAM2 consensus motif xxLNxxAxEFxP (Figure S3A) are interacting with helix 2 and 3 as well as helix 3 and 5 of MLLE domain, respectively (KOZLOV AND GEHRING 2010; XIE *et al.* 2014). Indeed, the interaction of MLLE with a hydrophobic amino acid is highly conserved, which in most cases is Phe with a known exception in the variant PAM2w motif of LARP4 and LARP4A, where Trp is found (Figure S3A) (YANG *et al.* 2011; XIE *et al.* 2014; GRIMM *et al.* 2020).

Studying the MLLE domain-containing protein Rrm4, we discover that it has three MLLE domains in its C-terminal half. MLLE3^{Rrm4} binds PAM2L motifs of Upa1 with a K_D of 5 and 15 µM for PAM2L2^{Upa1} and PAM2L1^{Upa1}, respectively. The binding affinities



2.5.1.1 Figure 11. Schematic model of endosomal attachment of mRNPs via MLLE domains.

Cargo mRNAs (green) are bound by the N-terminal RRM (1-3) domains of Rrm4 (green). The C-terminal MLLE domains (orange) form a binding platform: MLLE3^{Rrm4} interacts with PAM2L1^{Upa1} and PAM2L2^{Upa1}(orange), MLLE1 and -2^{Rrm4} might interact with currently unknown factors to support the endosomal binding. In particular, MLLE2^{Rrm4} has an accessory role during endosomal interaction. The four RRMs of Pab1 (green) interact with the poly(A) tail, and the MLLE^{Pab1} (blue) interacts with PAM2 of Upa1 and with the four PAM2 motifs of Upa2 (dark blue), a dimerizing scaffold protein. Upa1 is attached to endosomes via its FYVE domain, and the C-terminal GWW motif of Upa2 is crucial for its endosomal binding.

are in the same range as described for other MLLE/PAM2 interactions: for example, the binding of MLLE^{PABPC1} with PAM2^{LARP1}, PAM2$^{Tob2-125}$, PAM2^{LARP4} exhibit a K_D of 3.8, 16 and 22 µM, respectively (MATTIJSSEN *et al.* 2021). Importantly, our biophysical assessment confirms the exquisite binding specificity of MLLE^{Rrm4} that recognizes PAM2L1^{Upa1} and PAM2L2^{Upa1} but not the PAM2^{Upa1} version. PAM2L sequences contain a stretch of acidic amino acids in the N-terminal half, and the spacing of FxxP in the core sequence is altered (Figure S3A). These variations might account for the differential binding mode. Visual inspection of the potential PAM2L binding region in the predicted model revealed that MLLE3^{Rrm4} contains a Gly at position 736 to sustain the binding of

an aromatic residue of PAM2L as described for other MLLE domains (see above). However, we were unable to uncover the structural basis for the sequence specificity. Towards this end, future structural studies are required to provide detailed information on the interaction of MLLE3^{Rrm4} with PAM2L sequences.

Differential PAM2 binding has also been described for the MLLE^{UBR5}. This MLLE domain interacts with PAM2^{PAIP} with an affinity of 3.4 µM (LIM *et al.* 2006), whereas it binds a PAM2L sequence (Figure S3A) in its own HECT domain with lower affinity (K_D of 50 µM). The latter interaction has been implicated in regulating the HECT ligase activity (MUNOZ-ESCOBAR *et al.* 2015). Interestingly, the PAM2L sequence within the HECT domain of UBR5 is highly similar to the PAM2L1 and -2 of Rrm4: (i) the sequences contain an acidic stretch N-terminal to the conserved Phe (Figure S3A), (ii) the distance between Phe and Pro is two instead of one amino acid, and (iii) the PAM2L sequence contains an additional bulky Tyr in close vicinity to the Phe residue. Remarkably, MLLE^{PABC1} does not recognise the PAM2L sequence of UBR5 (MUNOZ-ESCOBAR *et al.* 2015). In essence, although the strong sequence specificity of MLLE^{Rrm4} and MLLE^{Pab1} from *U. maydis* is, to the best of our knowledge, so far unique, we hypothesize that differential PAM2 and PAM2L interactions are evolutionarily conserved and might be more widespread than currently anticipated.

We also observed a clear binding specificity for MLLE^{Pab1} from *U. maydis* that interacts with PAM2^{Upa1} but not the PAM2L sequences from Upa1 (Figure 11). MLLE^{Pab1} binds with comparable affinity to the PAM2^{Upa1} (K_D of about 14 µM, Figure S5B). Previously, we showed that mutations in PAM2^{Upa1} strongly decreased MLLE^{Pab1} binding but did not interfere with the endosomal shuttling of Pab1 (POHLMANN *et al.* 2015). Thus, there might be other members of the endosomal mRNPs interacting with Pab1 and stabilizing its endosomal association. In fact, the dimerizing scaffold protein Upa2 of endosomal mRNA transport contains four PAM2 motifs offering eight potential PAM2 motifs for interaction with Pab1 (Figure 11). However, mutating all four PAM2 motifs did not interfere with endosomal mRNA transport, although interaction with MLLE^{Pab1} was lost (JANKOWSKI *et al.* 2019), confirming a potential redundancy. Consistently, mutations in PAM2 of human LARP4B did not interfere with the function of stress granule recruitment, suggesting additional factors in this case (GRIMM *et al.* 2020)].

Studying the other two MLLE domains of Rrm4 revealed that both lack the canonical Gly for interactions with PAM2 or PAM2L sequences. MLLE1^{Rrm4} and MLLE2^{Rrm4} have Ser471 and Arg573 instead, respectively. Consistently, MLLE1^{Rrm4} and MLLE2^{Rrm4} do

not bind PAM2 or PAM2L sequences. Thus, although the general fold of the MLLE domain is probably conserved in MLLE1^{Rrm4} and MLLE2^{Rrm4}, these domains most likely exhibit a different binding specificity to their potential protein partner. Our detailed *in vivo* analysis revealed that MLLE2^{Rrm4} carries out an accessory function for the correct attachment of Rrm4 during endosomal shuttling. In the case of MLLE1^{Rrm4}, we did not identify a clear function so far. However, we believe that all three MLLE^{Rrm4} domains are functionally important. This is supported by the fact that the presence of an MLLE^{Rrm4} binding platform with three MLLE domains is evolutionarily conserved. Even Rrm4 versions of the distantly related fungus *Rhizophagus irregularis* contains three MLLE domains (Mucoromycota, determined by AlphaFold2) (MÜLLER *et al.* 2019).

Studying the spatial arrangement of the three MLLE^{Rrm4} domains revealed that they form a highly flexible binding platform pertinent for the regulation of Rrm4 mRNP transport. This would allow for the simultaneous interaction of several binding partners and potential rearrangements like an induced fit after binding. Such a scenario might be crucial during the loading and unloading of mRNPs to endosomes. Noteworthy, the N-terminal RNA-binding domain consisting of three RRMs is clearly separated from the MLLE^{Rrm4} domains for endosomal attachment. This is comparable with the arrangement of RRM and MLLE domains in human PABPC1: the four N-terminal RRM domains interact with the poly(A) tail of mRNAs, and a flexible spacer region exposes the MLLE^{PABPC1} domain for protein/protein interactions (SCHÄFER *et al.* 2019). Within the spacer region, additional interactions with the RRM2 of PABPC1 were found, suggesting a function in multimerization of the protein on the poly(A) tail of mRNAs (SAWAZAKI *et al.* 2018).

2.5.2 Conclusion

Endosomal mRNA transport is evolutionarily highly conserved. Besides hyphal growth in fungi, it is important for endosperm development in plants as well as neuronal functions in animals and humans (BÉTHUNE *et al.* 2019; TIAN *et al.* 2020; FERNANDOPULLE *et al.* 2021; MÜNTJES *et al.* 2021). Malfunctioning of this process causes defects in polar growth in fungi and has been implicated in neuronal diseases such as Charcot-Marie-Tooth type 2B neuropathy or amyotrophic lateral sclerosis in humans (CIONI *et al.* 2019; LIAO *et al.* 2019).

A key question is how mRNPs are linked to endosomes. In plants, two RRM-type RNA-binding proteins form a complex with cargo mRNAs and the endosomal component *N*-

ethylmaleimide-sensitive factor NSF as well as Rab5a (TIAN *et al.* 2018; TIAN *et al.* 2020a; TIAN *et al.* 2020b). Comparably, the FERRY complex (Five-subunit Endosomal Rab5 and RNA/ribosome intermediarY) interacts with the activated guanosine triphosphate (GTP) bound form of Rab5 during endosomal mRNA transport in neurons (SCHUHMACHER *et al.* 2021; QUENTIN *et al.* 2021). Further examples are the membrane-associated protein ANXA11 that links large RNA granules to lysosomal vesicles during mRNA transport in neuronal axons and dendrites (LIAO *et al.* 2019). Thus, a number of components and interactions are known, however detailed structural insights are scarce. Here, we have demonstrated that in hyphae, endosomal attachment of Rrm4 is mediated by an $MLLE^{Rrm4}$ binding platform with a non-canonical accessory domain joining an essential $MLLE^{Rrm4}$ domain for perfect interaction with Upa1 on the endosomal surface (Figure 11). This binary interaction in the core of the transport mRNPs is supported by numerous interactions of additional protein partners such as Upa2 and Pab1 that assist in attaching components to the endosomal surfaces (Figure 11). In closing, studying endosomal mRNP transport in fungal model systems might guide future research endeavors in plant and neuronal systems.

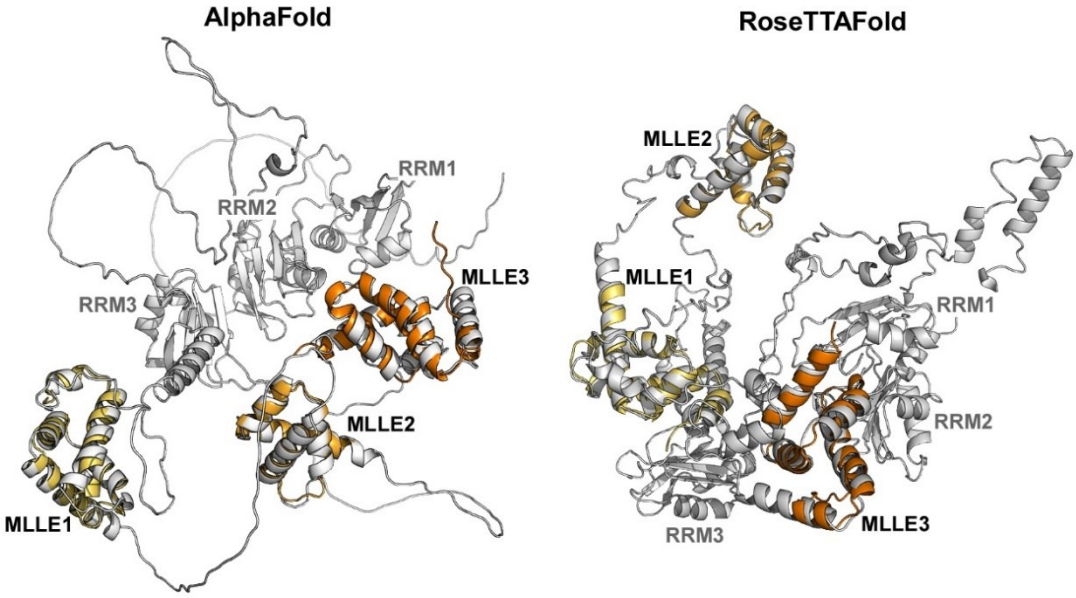
2.5.3 Supporting information

A

Domain	Templates[a]	Rrm4 span	Template span	Identity (Similarity)[b]
MLLE1	6H7A:A	451-529	481-560	22.1 (38.4)
	3PTH:A	451-528	544-618	18.4 (34.5)
	1I2T:A	464-528	1009-1069	25.8 (40.9)
MLLE2	1I2T:A	571-629	1009-1069	23.8 (44.4)
	3KTP:A	550-631	554-620	17.6 (35.3)
	3KUR:A	550-627	554-616	18.5 (37.0)
MLLE3	2X04:A	712-791	540-616	30.6 (44.7)
	3KUR:A	716-791	544-616	32.9 (45.6)
	3KTP:A	716-791	544-616	32.9 (45.6)

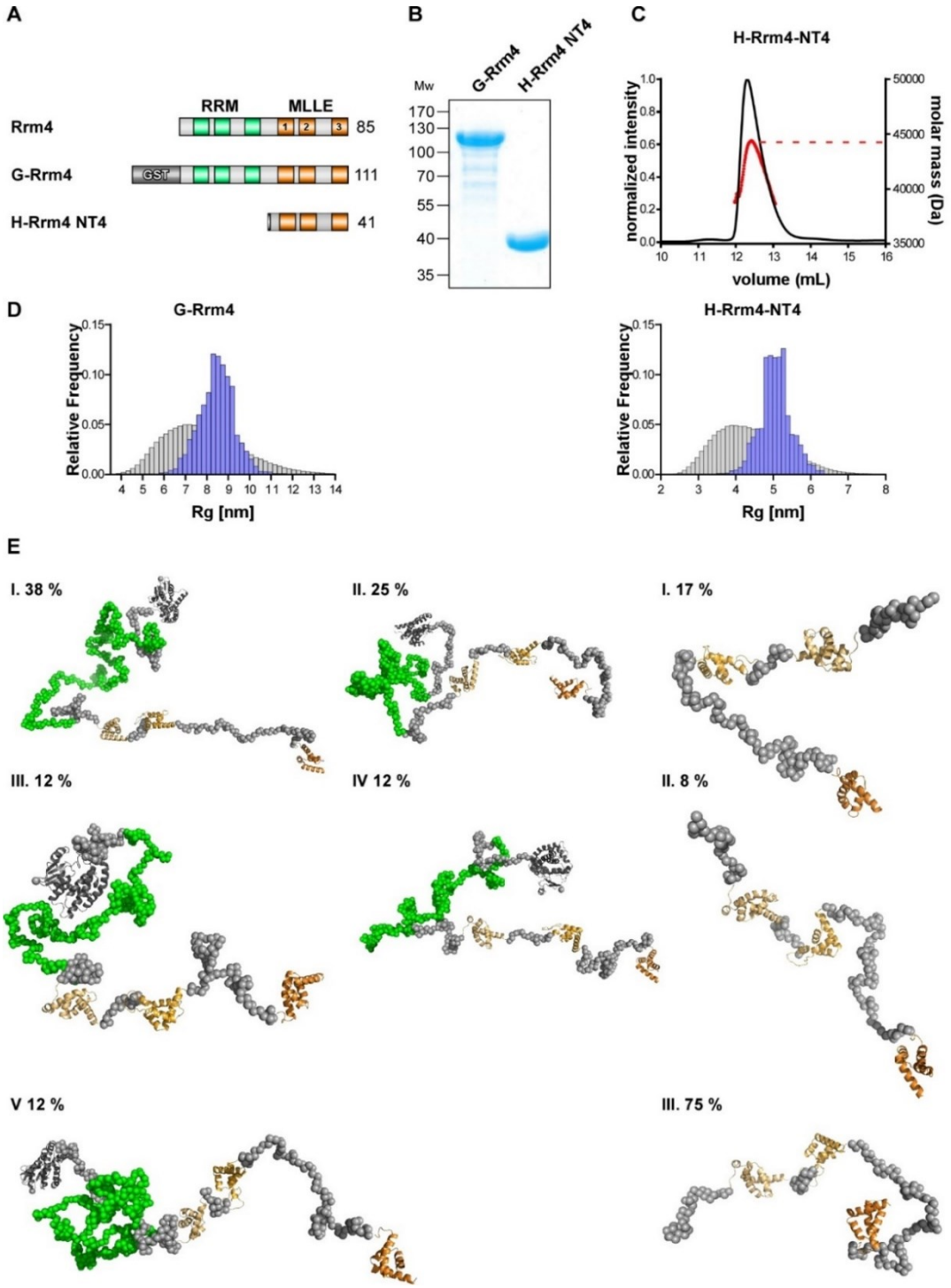
a. PDB ID and chain identifier. b. in percentage

B



2.5.3.1 Figure S1. The presence of three MLLEs is verified by additional modelling predictions.

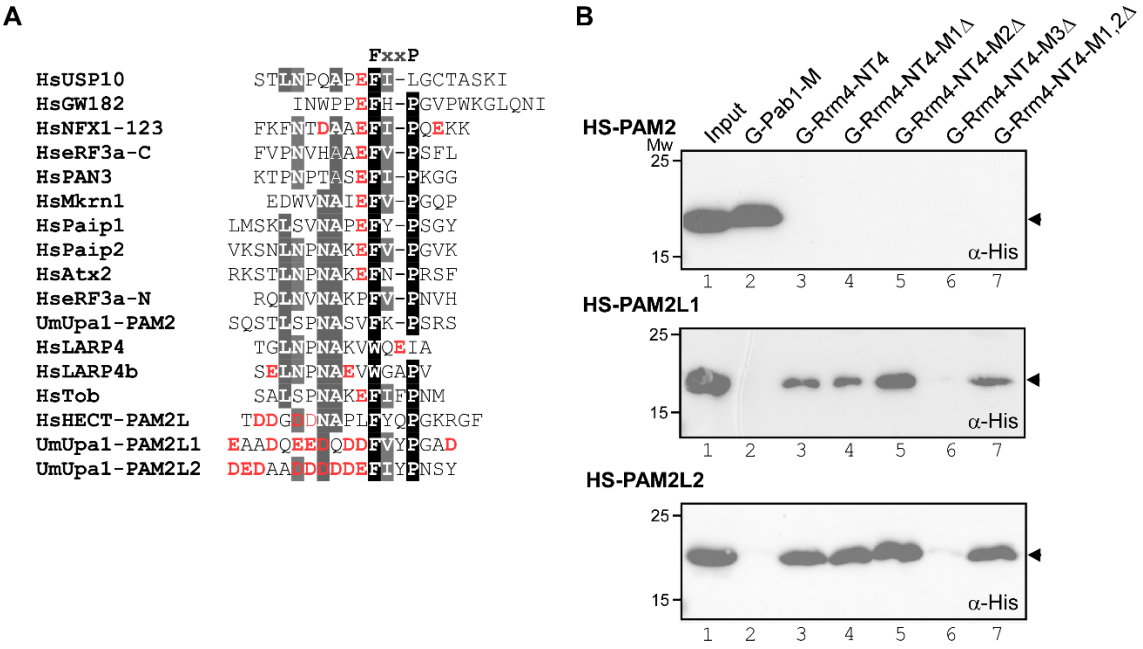
(A) Compilation of MLLE sequences used for modeling with the highest similarity of MLLE1-3^{Rrm4}. (B) Structural models obtained with TopModel overlaid to Rrm4 full-length models obtained with the recently available tools as indicated. Natural alignments between corresponding MLLE^{Rrm4} domains have an RMSD < 2Å, mutually confirming the quality of the independently modeled structures. The differences in the relative domain arrangements in both full-length models and the disordered regions between the domains suggest high mobility within Rrm4.



2.5.3.2 Figure S2. The three MLLEs of Rrm4 are located in a flexible C-terminal region.

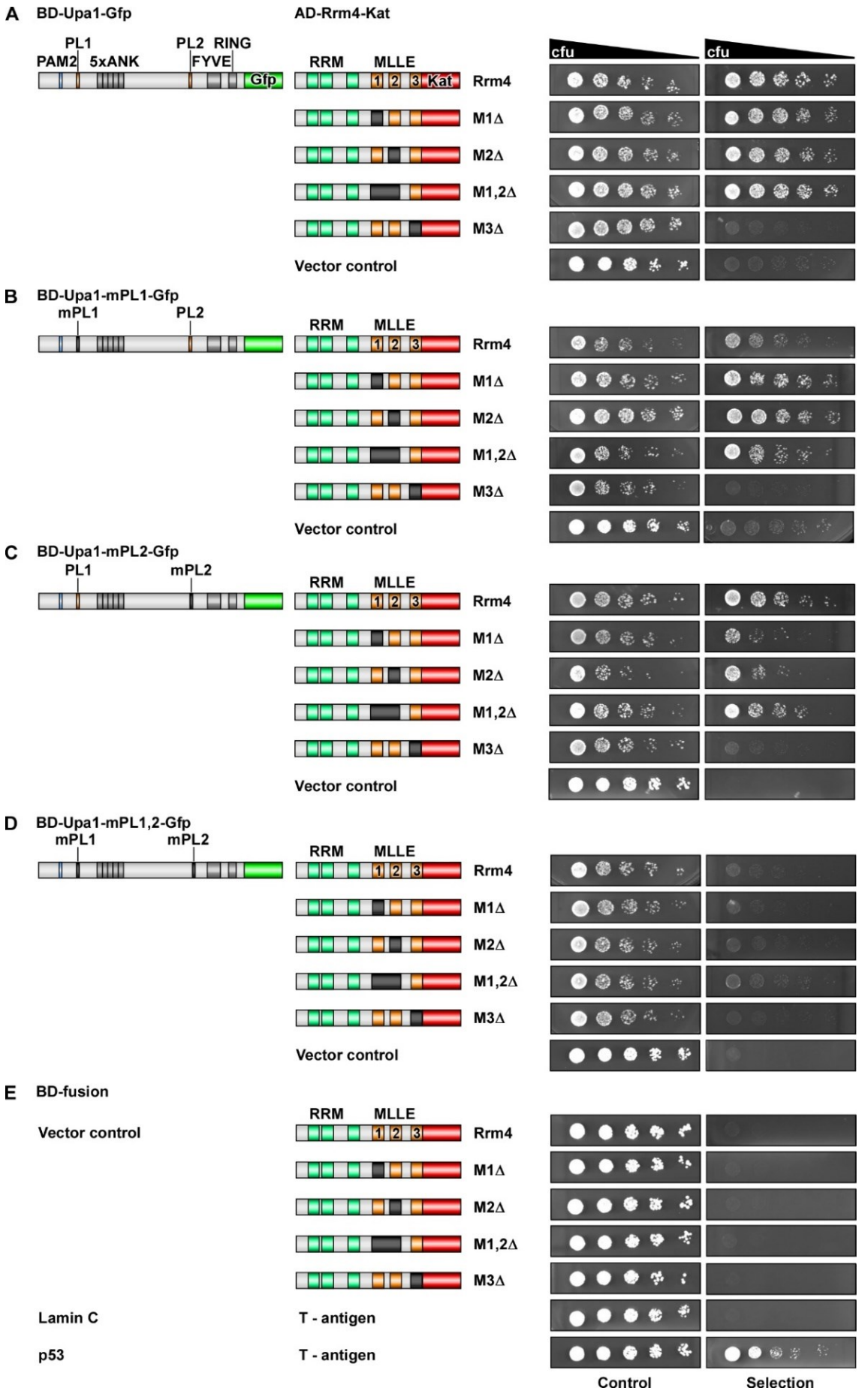
(A) Schematic representation of protein variants drawn to scale (molecular weight in kilo Dalton indicated next to protein bar) using the following coloring: lime green, RNA recognition motif (RRM); orange, $MLLE^{Rrm4}$ domains; (B) SDS-PAGE analysis of purified G-Rrm4, H-Rrm4-NT4 used in crystallography and SAXS measurement. (C) MALS-SEC analysis of H-Rrm4-NT4. Graph shows the elution profile. Dotted line in red indicate the apparent molecular weight as observed in the light scattering. (D) R_g distribution calculated by EOM pool is shown in grey bars and the selected models in blue bars left GST_Rrm4 right H-Rrm4-NT4 (E) Left Selected models of the EOM analysis for GST-Rrm4. The $MLLE^{Rrm4}$ subdomains and the GST are shown in cartoon representation (MLLE1 in light orange, $MLLE2^{Rrm4}$ in orange, MLLE 3^{Rrm4} in dark orange) and the missing amino acids as grey spheres. I: The model has a R_g of 8.94 nm, a

D_{max} of 29.22 nm with a volume fraction of~0.38. II: The model has a R_g of 8.75 nm, a D_{max} of 23.99 with a volume fraction of~0.25. III: The model has a R_g of 7.74 nm, a D_{max} of 25.90 with a volume fraction of~0.12. IV: The model has a R_g of 8.33 nm, a D_{max} of 28.79 with a volume fraction of~0.12. V: The model has a R_g of 9.14 nm, a D_{max} of 33.73 with a volume fraction of~0.12. Right Selected models of the EOM analysis for H-Rrm4-NT4. The MLLE subdomains are shown in cartoon representation (MLLE1 in light orange, MLLE2 in orange, MLLE3 in dark orange) and the missing amino acids as grey spheres. I: The model has a R_g of 5.12 nm, a D_{max} of 15.56 with a volume fraction of~0.17. II: The model has a R_g of 5.90 nm, a D_{max} of 18.73 nm with a volume fraction of~0.08. III: The model has a R_g of 5.10 nm, a D_{max} of 16.43 nm with a volume fraction of~0.75.



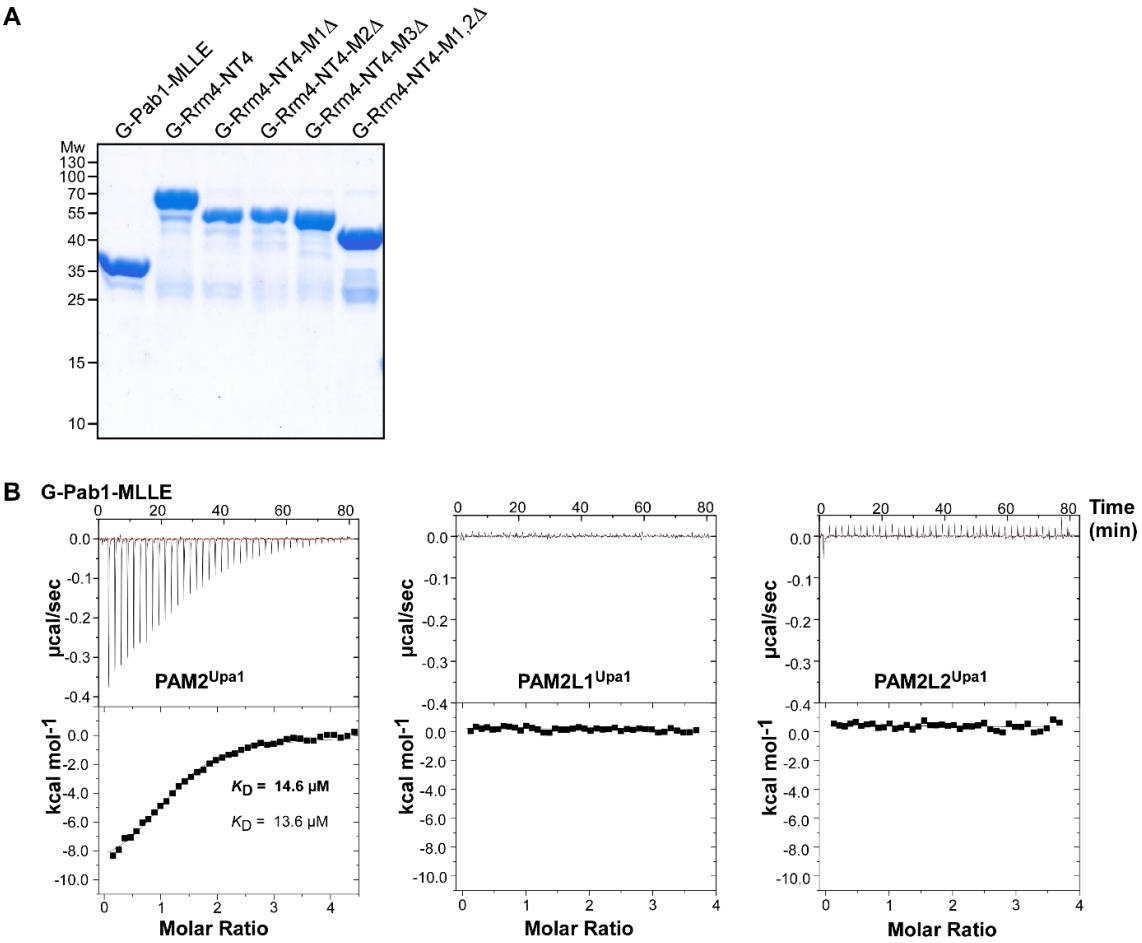
2.5.3.3 Figure S3. MLLE1^{Rrm4}, -2^{Rrm4} are not essential for PAM2L1^{Upa1} and -L2^{Upa1} binding in GST pull-down assay.

(A) Comparison of PAM2 sequences found in Upa1 (UniprotKB ID A0A0D1E015) with those of human proteins, such as Usp10 (Q14694), GW182 Q9HCJ0), Mkrn1 (Q9UHC7), Paip1 (Q9H074), Paip2 (Q9BPZ3), Atx2 (-Q99700), NFX (Q12986), eRF3 (P15170), PAN3 (Q58A45), LARP4 (Q71RC2), LARP4b (Q92615), Tob (P50616), HECT (O95071), Asp and Glu are indicated in red stressing the highly negative charges in PAM2L sequences. (B) Western blot analysis of GST co-purification experiments with components expressed in *E. coli*: N-terminal Hexa-Histidine-SUMO-tagged PAM2 variants were pulled down by N-terminal GST fused MLLE variants of Rrm4 and Pab1. Experiment was performed with the soluble fraction of *E. coli* cell lysate to demonstrate specific binding. Results were analyzed with α-His immunoblotting.


A
BD-Upa1-Gfp
AD-Rrm4-Kat
PL1
PAM2
5xANK
PL2
RING
FYVE
Gfp
RRM
MLLE
1
2
3
Kat
Rrm4
M1Δ
M2Δ
M1,2Δ
M3Δ
Vector control
cfu
cfu
B
BD-Upa1-mPL1-Gfp
mPL1
PL2
C
BD-Upa1-mPL2-Gfp
mPL2
D
BD-Upa1-mPL1,2-Gfp
E
BD-fusion
Lamin C
T - antigen
p53
T - antigen
Control
Selection

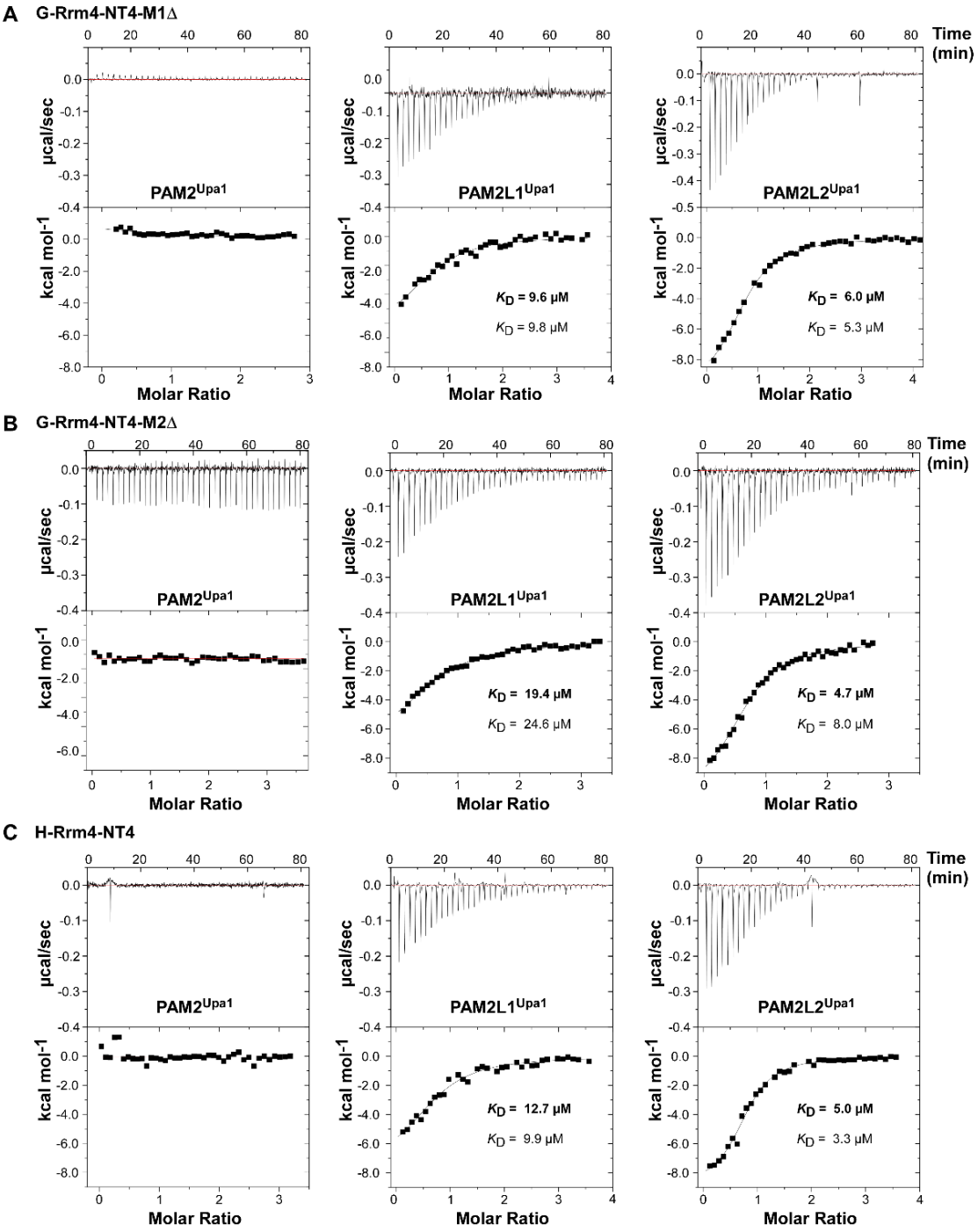
2.5.3.4 Figure S4. MLLE1^{Rrm4}, -2^{Rrm4} are not essential for PAM2L1^{Upa1} and -L2^{Upa1} binding in yeast two-hybrid experiments.

(A-E) Yeast two-hybrid analyses with schematic representation of protein variants tested on the left. Cultures were serially diluted 1:5 (decreasing colony-forming units, cfu) and spotted on respective selection plates controlling transformation and assaying reporter gene expression (see Materials and methods).



2.5.3.5 Figure S5. MLLE^{Pab1} does not bind PAM2L1^{Upa1} and -L2 ^{Upa1}.

(A) SDS-PAGE analysis of purified GST-MLLE variants used in ITC experiments (see also S6 Figure). (B) Representative isothermal titration calorimetry (ITC) binding curves of MLLE^{Pab1} domain. Experiments were performed using GST or hexa-histidine-tagged MLLE variants and synthetic PAM2^{Upa1} and PAM2L^{Upa1} peptide variants. K_D values of two independent measurements are given (values corresponding to the indicated data are given in bold).

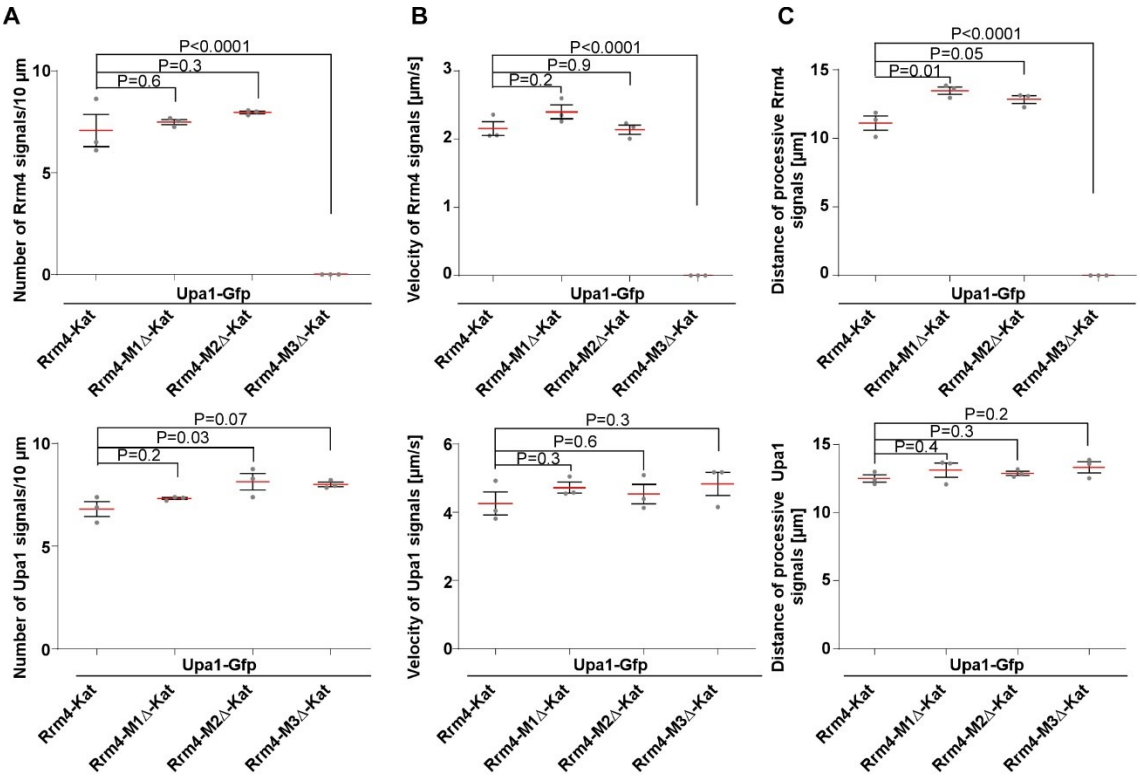


D

	PAM2^{Upa1}	PAM2L1^{Upa1}	PAM2L2^{Upa1}
G-Pab1-MLLE	14.6 / 13.6	No binding	No binding
G-Rrm4-NT4	No binding	14.9 / 14.0	5.1 / 5.9
G-Rrm4-NT4-M1Δ	No binding	9.6 / 9.8	6.0 / 5.3
G-Rrm4-NT4-M2Δ	No binding	19.4 / 24.6	4.7 / 8.0
G-Rrm4-NT4-M3Δ	No binding	No binding	No binding
G-Rrm4-NT4-M1,2Δ	No binding	13.3 / 12.0	5.3 / 6.4
H-Rrm4-NT4	No binding	12.7 / 9.9	5.0 / 3.3

2.5.3.6 Figure S6. MLLE1^{Rrm4}, -2^{Rrm4} do not contribute to the binding of PAM2L1^{Upa1} and -L2^{Upa1}.

(A-C) Representative isothermal titration calorimetry (ITC) binding curves of MLLE domains. Experiments were performed using GST or hexa-histidine-tagged MLLE variants and synthetic PAM2^{Upa1} and PAM2L^{Upa1} peptide variants. K_D values of two independent measurements are given (values corresponding to the indicated data are given in bold). (D) Summary of ITC results shown in Figures 3 and S3. K_D values are given in µM.



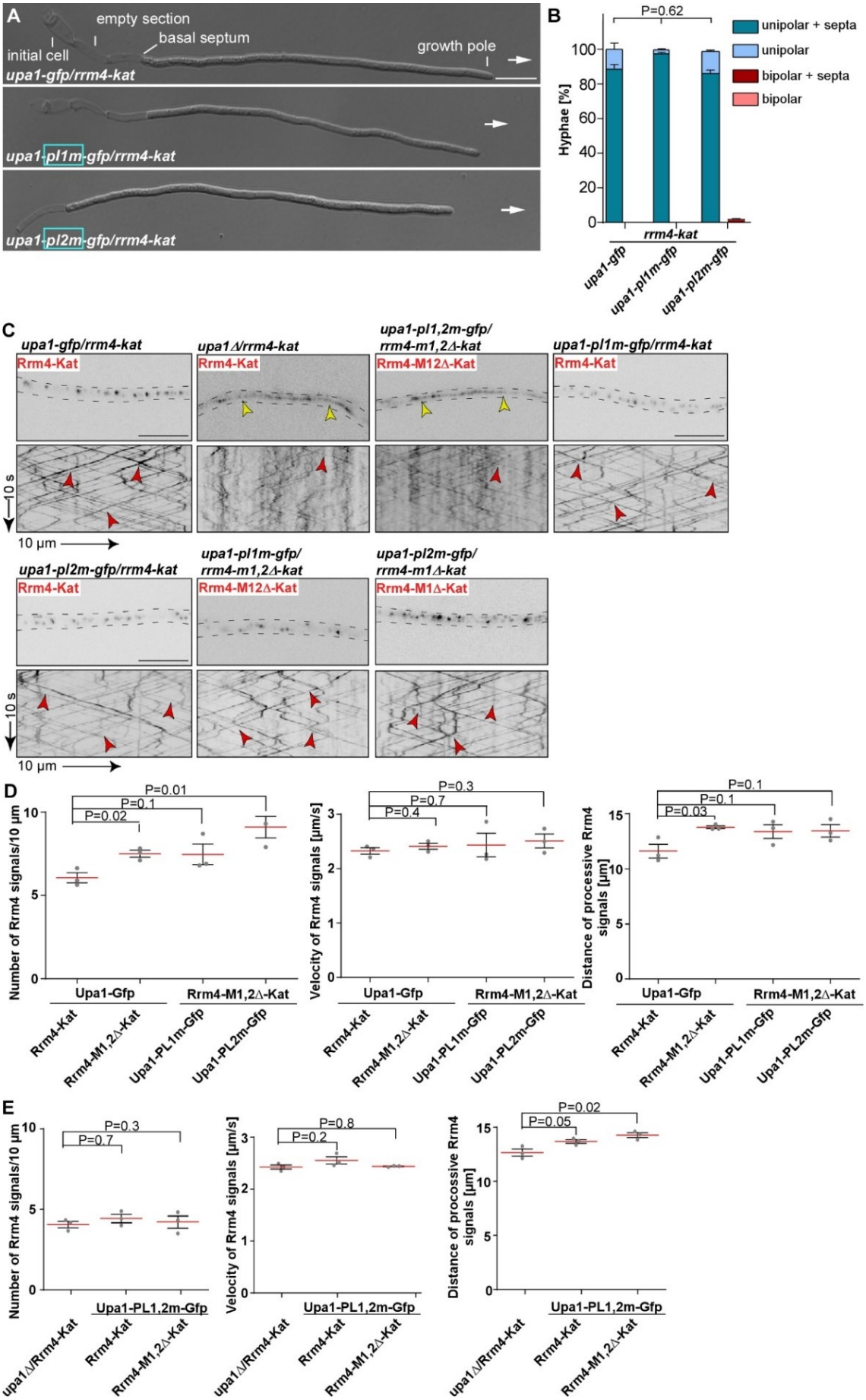
D

upa1	*rrm4*	Phenotype	Endosomal shuttling of Rrm4	Microtuble binding of Rrm4
wt	*wt*	*wt*	*wt*	*wt*
wt	*rrm4Δ*	bipolar	not applicable	not applicable
wt	*m1Δ*	*wt*	*wt*	*wt*
wt	*m2Δ*	*wt*	*wt*	*wt*
wt	*m3Δ*	bipolar	abolished	not applicable
wt	*m1,2Δ*	*wt*	*wt*	*wt*
pl1,2m	*wt*	bipolar	aberrant	strong
upa1Δ	*wt*	bipolar	aberrant	strong
pl1,2m	*m1,2Δ*	bipolar	aberrant	strong
pl1m	*wt*	*wt*	*wt*	weak
pl2m	*wt*	*wt*	*wt*	medium
pl1m	*m1,2Δ*	*wt*	*wt*	weak
pl2m	*m1,2Δ*	*wt*	*wt*	strong
pl2m	*m1Δ*	*wt*	*wt*	medium
pl2m	*m2Δ*	*wt*	*wt*	strong

2.5.3.7 Figure S7. Deletion of MLLE3^{Rrm4} abolishes endosomal movement of Rrm4.

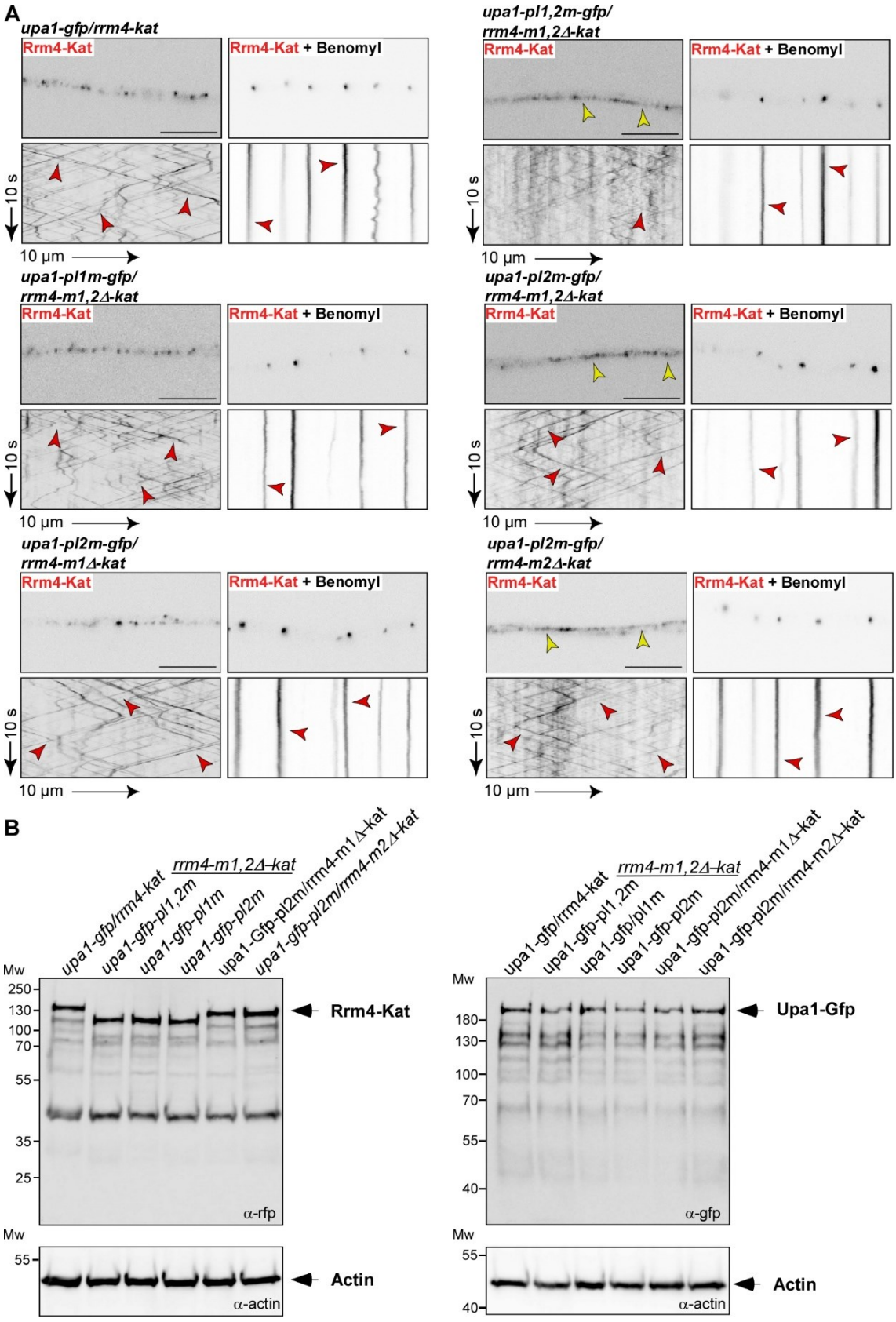
(A-C) Quantification of processive Rrm4-Kat (top) and Upa1-Gfp signals (bottom; (A)), velocity of fluorescent Rrm4-Kat (top) and Upa1-Gfp signals (bottom; (B)) and the travelled distance of processive Rrm4-Kat (top) and Upa1-Gfp signals (bottom; (C); per 10 µM of hyphal length; only particles with a processive

movement of > 5 µM were conducted; data points representing mean from n =3 independent experiments, with mean of means, red line and SEM; unpaired two-tailed Student's t-test (α <0.05), for each experiment at least 25 hyphae were analyzed per strain). (D) Summary of the *in vivo* analysis is shown in Figures 4,5 and S7-10.


A
empty section
basal septum
initial cell
growth pole
upa1-gfp/rrm4-kat
upa1-pl1m-gfp/rrm4-kat
upa1-pl2m-gfp/rrm4-kat
B
P=0.62
unipolar + septa
unipolar
bipolar + septa
bipolar
Hyphae [%]
rrm4-kat
upa1-gfp
upa1-pl1m-gfp
upa1-pl2m-gfp
C
upa1-gfp/rrm4-kat
Rrm4-Kat
upa1Δ/rrm4-kat
upa1-pl1,2m-gfp/ rrm4-m1,2Δ-kat
Rrm4-M12Δ-Kat
upa1-pl1m-gfp/rrm4-kat
upa1-pl2m-gfp/rrm4-kat
upa1-pl1m-gfp/ rrm4-m1,2Δ-kat
upa1-pl2m-gfp/ rrm4-m1Δ-kat
Rrm4-M1Δ-Kat
10 s
10 µm
D
Number of Rrm4 signals/10 µm
P=0.01
P=0.1
P=0.02
Velocity of Rrm4 signals [µm/s]
P=0.3
P=0.7
P=0.4
Distance of processive Rrm4 signals [µm]
P=0.1
P=0.1
P=0.03
Upa1-Gfp
Rrm4-M1,2Δ-Kat
Rrm4-Kat
Rrm4-M1,2Δ-Kat
Upa1-PL1m-Gfp
Upa1-PL2m-Gfp
E
Number of Rrm4 signals/10 µm
P=0.3
P=0.7
Velocity of Rrm4 signals [µm/s]
P=0.8
P=0.2
Distance of processive Rrm4 signals [µm]
P=0.02
P=0.05
Upa1-PL1,2m-Gfp
upa1Δ/Rrm4-Kat
Rrm4-Kat
Rrm4-M1,2Δ-Kat

2.5.3.8 Figure S8. Deletion of MLLE1^{Rrm4} and -2 cause aberrant staining of microtubules.

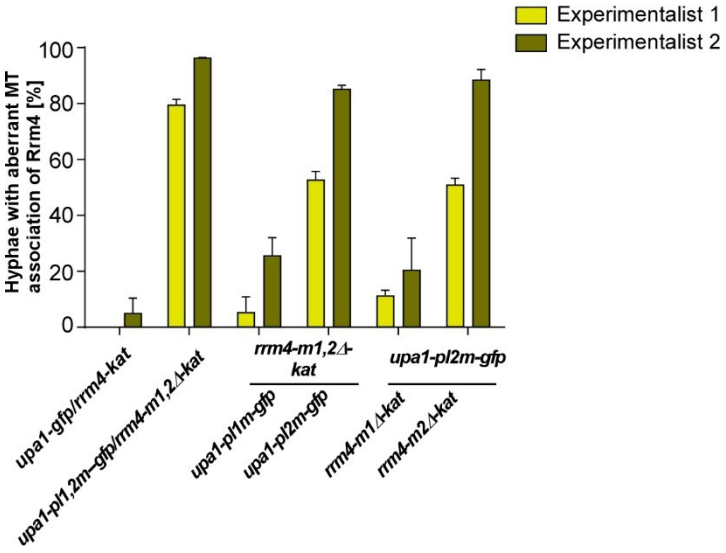
(A) Growth of AB33 derivatives in their hyphal form (6 h.p.i.; size bar 10 µM). Growth direction is marked by arrows. (B) Quantification of hyphal growth of AB33 derivatives shown in panel A (6 h.p.i.): unipolarity, bipolarity and basal septum formation were quantified (error bars, SEM.; n = 3 independent experiments, > 100 hyphae were analyzed per strain; For statistical evaluation, the percentage of uni- and bipolarity was investigated by using unpaired two-tailed Student's t-test ($\alpha<0.05$). (C) Micrograph and Kymograph of AB33 hyphae derivatives (6 h.p.i.) expressing red and green fluorescent proteins as indicated. Fluorescence signals were detected simultaneously using dual-view technology (arrow length on the left and bottom indicates time and distance, respectively). Processive co-localizing signals are marked by red arrowheads. Aberrant microtubule staining is indicated by a yellow arrowhead. (D-E) Quantification of processive Rrm4-Kat signals (left), velocity of fluorescent Rrm4-Kat (middle) and the travelled distance of processive Rrm4-Kat signals (right) related to Figure 5C and EV5C, respectively (per 10 µM of hyphal length; only particles with a processive movement of > 5 µM were conducted; data points representing mean from n =3 independent experiments, with mean of means, red line and SEM; unpaired two-tailed Student's t-test ($\alpha<0.05$), for each experiment at least 25 hyphae were analyzed per strain).



2.5.3.9 Figure S9. Mislocalization of Rrm4 is microtubule-dependent.

(A) Benomyl treatment is shown in micrograph and kymograph of AB33 hyphae derivatives (6 h.p.i.) expressing red and green fluorescent proteins. Processive signals, as well as static signals, post benomyl treatment are marked by red arrowheads. Aberrant microtubule staining is indicated by a yellow arrowhead. (B)

Western blot analysis of the expression levels of Rrm4 and Upa1 variants 6 h.p.i. of hyphal growth. Rrm4 and Upa1 variants were detected via mKate2 and Gfp, respectively. Actin was detected as loading control. Bands representing full-length proteins are marked with arrows.



2.5.3.10 Figure S10. Independent testing of aberrant microtubule staining.

Evaluation of the most important strains showing aberrant microtubule staining analyzed by two experimentalists (we used the data obtained by the more experienced microscopist in Figure 5D; see Materials and methods).

2.5.3.11 Spreadsheet S11. Microscopic data used for quantifications in Figures 4, 5, S7, S8 and S10.

Available online at https://doi.org/10.1371/journal.pgen.1010269.s020

2.6 Materials and methods

2.6.1 Structural modelling of C-terminal MLLE domains of Rrm4

To obtain structural models of the C-terminal region of Rrm4, an iterative homology modelling approach was used with the TopModel workflow (MULNAES *et al.* 2020). Initially, the entire C-terminal region (421 to 792) was submitted as input in TopModel and identified templates for MLLE3^{Rrm4} (665 – 791 AA; Figures 6C, S1A). Then, the rest of the C-terminal part comprising amino acids 421 to 664 was resubmitted as input identifying other templates as a new starting point for the MLLE2^{Rrm4} (571-629). Likewise, the remaining C-terminal sequence comprising amino acids 421 to 549 was resubmitted as input, identifying other templates as a new starting point for the MLLE1^{Rrm4} (446-530). In total, this led to the identification of three MLLE^{Rrm4} domains, for which structural models were generated using default TopModel parameters. The quality of the structural models was assessed with TopScore (MULNAES AND GOHLKE 2018).

2.6.2 Plasmids, strains, and growth conditions

For molecular cloning of plasmids, *Escherichia coli* Top10 cells (Thermofisher C404010) and for recombinant protein expression *E .coli Lobstr* cells (Kerafast EC1002) were used respectively. Sequence encoding H-Rrm4-NT4 was inserted into the pET22 vector (Merck 69744) with an N-terminal hexa-histidine tag for crystallization studies. Sequence encoding MLLE variants were inserted into the pGEX-2T vector (Merck GE28-9546-53) containing GST sequence in N-terminus for pulldown and ITC experiments. Sequence encoding PAM2 variants were inserted into the Champion pET-Sumo vector (Thermofisher K30001). pRarepLys plasmid was co-transformed in *E. coli Lobstr strain* to supplement the rare codons for efficient recombinant protein production. *E. coli* transformation, cultivation, and plasmid isolation were conducted using standard techniques. For yeast two-hybrid analyses *S. cerevisiae* strain AH109 (Matchmaker 3 system, Clontech) was used. Yeast cells were transformed and cultivated using standard techniques. All *U. maydis* strains are derivatives of AB33, in which hyphal growth can be induced by switching the nitrogen source in the medium (BRACHMANN *et al.* 2001). *U. maydis* yeast cells were incubated in complete medium (CM) supplemented with 1% glucose, whereas hyphal growth was induced by changing to nitrate minimal medium (NM) supplemented with 1% glucose, both at 28 °C (BRACHMANN *et al.* 2001). Detailed growth conditions and general cloning strategies for *U. maydis* are described elsewhere (BRACHMANN *et al.*

2004; BAUMANN *et al.* 2012; TERFRÜCHTE *et al.* 2014). All plasmids were verified by sequencing. Strains were generated by transforming progenitor strains with linearized plasmids. Successful integration of constructs was verified by diagnostic PCR and by Southern blot analysis (BRACHMANN *et al.* 2004). For ectopic integration, plasmids were linearized with *Ssp*I and targeted to the ip^S locus (LOUBRADOU *et al.* 2001). A detailed description of all plasmids, strains, and oligonucleotides is given in S3–S9 Tables. Sequences are available upon request.

2.6.3 Recombinant protein expression and purification

E. coli cells from freshly transformed plates were inoculated in 20 ml expression media. To produce high-density expression cultures with tight regulation of induction and expression in shake flasks we designed a complex media inspired by the principle of Studier's autoinduction media (STUDIER 2005). In essence, we used adequate amount of glucose to prevent the unintended induction and leaky expression of target protein as well as phosphate buffer to prevent acidity as a result of glucose metabolism from the excessive glucose in the media. In addition, the medium contained glycerol, nitrogen, sulphur, and magnesium for promoting high-density growth. Unlike the Studier's autoinduction media our media lack lactose therefore expression can be induced with IPTG and expressed at required temperature (1.6 % Trypton, 1% Yeast extract, 50 mM Na_2HPO_4, 50 mM KH_2PO_4, 25 mM $[NH_4]_2SO_4$, 0.5% Glycerol, 0.5% Glucose, 2 mM $MgSO_4$) (STUDIER 2005) with ampicillin (100 mg/ml) and chloramphenicol (34 mg/ml) or kanamycin (200 mg/ml) and chloramphenicol (34 mg/ml) and grown overnight (16 hours) at 37 °C, 200 rpm. Note that the high concentration of kanamycin was used to prevent the unintended resistance promoted by high phosphate concentration (STUDIER 2005). Supernatant from the overnight culture was removed by centrifugation at 4 °C, 5000 × g for 2 minutes. Cells were resuspended in fresh media with a starting OD_{600} of 0.1 and grown at 37 °C, 200 rpm for about 2 hours 30 minutes until the OD_{600} = 1. Protein expression was induced at 28 °C, 200 rpm, for 4 hours by addition of 1 mM IPTG, and harvested by centrifugation at 4°C, 6,000 × g for 5 minutes. Protein purification was performed as per the previous report (ABTS *et al.* 2013). Hexa-histidine tagged protein was purified using Nickel-based affinity chromatography (HisTrap HP, GE Healthcare) on Akta prime FPLC system. Cells were thawed on ice and resuspended in buffer A (20 mM HEPES pH 8.0, 200 mM NaCl, 1 mM EDTA, 10 mM Imidazole pH 8.0; 1 mM PMSF, 0.5 mg/ml Lysozyme, 0.5 mg/ml DNase, 1mM β-mercaptoethanol [β-ME]). Subsequently, cells were

lysed by sonication on ice and centrifuged at 4 °C 18,000 × g for 30 minutes. Resulting supernatant was loaded onto a pre-equilibrated column with buffer B (20 mM HEPES pH 8.0, 200 mM NaCl,10 mM Imidazole), washed with buffer C (20 mM HEPES pH 8.0, 200 mM NaCl, 50 mM Imidazole, 1 mM β-ME), eluted with buffer D (20 mM HEPES pH 8.0, 200 mM NaCl, 300 mM Imidazole, 1mM β-ME) and further purified by size exclusion chromatography (HiLoad 26/600 Superdex 200, GE Healthcare), pre-equilibrated with buffer E (20 mM HEPES pH 8.0, 200 mM NaCl, 1 mM β-ME). For crystallization studies, H-Rrm4-NT4 was purified as above except that the buffers were prepared with high salt (500 mM NaCl) and without β-ME.

GST-tagged protein was purified using Glutathione-based affinity chromatography (GSTrap FF GE Healthcare). Cells were thawed on ice and resuspended in Buffer F (20 mM HEPES pH 8.0, 200 mM NaCl, 1 mM EDTA, pH 8.0; 1 mM PMSF, 0.1 mg/ml Lysozyme, 1 mM β-ME). Subsequently, cells were lysed by sonication on ice and centrifuged at 4 °C, 18,000 g for 30 minutes. The resulting supernatant was loaded onto a pre-equilibrated column with buffer E (20 mM HEPES pH 8.0, 200 mM NaCl, 1 mM β-ME) and washed with the same buffer, eluted with buffer H (20 mM HEPES pH 8.0, 200 mM NaCl, 10 mM reduced glutathione, 1 mM β-ME), and further purified by size exclusion chromatography (HiLoad 16/600 Superdex 200 GE Healthcare), pre-equilibrated with buffer E. Protein purity was analyzed on SDS-PAGE. All the purified proteins were concentrated, centrifuged at 4°C, 100,000 × g for 30 minutes, quantified by Nanodrop A280, aliquoted, and stored at -80 °C. Peptides for ITC experiments were custom-synthesized and purchased from Genscript, USA (see Figure 3A for peptide sequence).

2.6.4 GST pull-down experiments

Pull-down assays were performed as per the previous report (JANKOWSKI *et al.* 2019). In short, GST-MLLE variants and HS-PAM2^{Upa1} variants were expressed in *E. coli*. Cell pellets from 50 ml *E. coli* expression culture were resuspended in 10 ml buffer F (20 mM HEPES pH 8.0, 200 mM NaCl, 1 mM EDTA; 0.5% Nonidet P-40, 1 mM PMSF, 0.1 mg/ml Lysozyme). Cells were lysed by sonication on ice and centrifuged at 4 °C, 16,000 × g for 10 minutes. 1 mL of the resulting supernatant was incubated for 1 hour at 4 °C on constant agitation of 1,000 rpm with 100 µL glutathione sepharose (GS) resin (GE Healthcare), pre-equilibrated in buffer F. The GS resin was washed three times with 1 ml of buffer G (20 mM HEPES pH 8.0, 200 mM NaCl, 1 mM EDTA, 0.5 % Nonidet P-40). Subsequently, supernatant of HS-PAM2 variants was added to the GST-MLLE variant

bound resins and incubated for 1 hour at 4 °C on agitation. The resins were washed as aforementioned, resuspended in 100 µL of 2x Laemmli loading buffer, boiled for 10 minutes at 95 °C and analyzed by western blotting.

2.6.5 *U. maydis* cell disruption and sample preparation for immunoblotting

U. maydis hyphae were induced as described earlier (see Plasmids, strains, and growth conditions). 50 ml of hyphal cells (6 h.p.i) were harvested in 50 ml conical centrifuge tubes by centrifugation at 7,150 × g, for 5 minutes. Cell pellets were resuspended in 2 ml phosphate-buffered saline pH 7.0 (PBS; 137 mM NaCl, 2.7mM KCl, 8 mM Na_2HPO_4 and 2 mM KH_2PO_4) and transferred to a 2 ml centrifuge tubes. Cells were harvested at 7,150 × g for 5 minutes and supernatant was removed completely. The resulting cell pellets were flash-frozen in liquid nitrogen and stored at -80 °C until use. Sample tubes were placed on 24 well TissueLyser adapter (Qiagen 69982) and soaked in liquid nitrogen for 1 minute, 5 mm stainless steel bead was added to each sample tube and the cells were disrupted at 30 Hz for 3 times 1 minute in Mixer Mill MM400 (Retsch, Germany), with intermittent cooling between shaking. At the end of the cell disruption dry homogenized powder of cells was resuspended in 1 ml urea buffer (8 M urea, 50 mM Tris/HCl pH 8.0 containing one tablet of 'cOmplete' protease inhibitor per 25 ml, Roche, Germany; 1 mM DTT; 0.1 M PMSF) and centrifuged at 16,000 × g for 10 minutes at 4 °C. The supernatant was used for subsequent analysis. Samples were diluted ten times and protein concentrations were measured by BCA assay (Thermofisher 23225). Samples were diluted to 1 mg/ml final concentration in Laemmli buffer and boiled at 95 °C for 10 minutes. 40 µg of each sample was loaded in 1.5 mm thickness gels for SDS-PAGE, subsequently analyzed by Western blotting.

2.6.6 Yeast two-hybrid analysis

Yeast two-hybrid analyses were performed as per the previous report (POHLMANN *et al.* 2015). The two-hybrid system Matchmaker 3 from Clontech was used as per manufacturer's instructions. Yeast strain AH109 was co-transformed with derivatives of pGBKT7-DS and pGADT7-Sfi (S8 Table, S4 Figure) and were grown on synthetic drop-out plates (SD) without leucine and tryptophan at 28 ˚C for 2 days. Transformants were patched on SD plates without leucine and tryptophan (control) or on SD plates without leucine, tryptophan, histidine, and adenine (selection). Plates were incubated at 28 ˚C for 2 days to test for growth under selection conditions. For qualitative plate assays, cells (SD

-leu, -trp, OD600 of 0.5) were serially diluted 1:5 with sterile water, spotted 4 µl each on control and selection plates and incubated at 28 ˚C for 2 days. Colony growth was documented with a LAS 4000 imaging system (GE Healthcare).

2.6.7 SDS-PAGE and immunoblotting

All SDS–PAGE and Western blotting experiments were performed as reported previously (JANKOWSKI *et al.* 2019). Western blotting samples were resolved by 8 or 10 or 12 % SDS-PAGE and transferred and immobilized on nitrocellulose membrane (Amersham Protran) by semi-dry blotting using Towbin buffer (25 mM Tris pH 8.6, 192 mM Glycine, 15% Methanol). Proteins were detected using α -His from mouse (Sigma H1029), α-Gfp from mouse, (Roche, Germany), α-tRfp from rabbit (AB233-EV, Evrogen) and α-Actin from mouse (MP Biomedicals, Germany) as primary antibodies. As secondary antibodies α-mouse IgG HRP conjugate (Promega W4021) or α-rabbit IgG HRP conjugate (Cell Signaling #7074) were used. Antibodies bound to nitrocellulose membranes were removed by incubating in TBS buffer pH 3.0 (50 Tris pH 3.0, 150mM NaCl) at room temperature, before detecting with the constitutively expressed control (α-Actin). Detection was carried out by using ECL^{TM} Prime (Cytiva RPN2236). Images were taken by luminescence image analyzer, LAS4000 (GE Healthcare) as per the manufacturer's instructions.

2.6.8 Multiangle light scattering (MALS)

MALS was performed as per the previous report (Weiler et al. 2021). Superdex 200 Increase 10/300 GL column (GE Healthcare) was pre-equilibrated overnight at 0.1 ml/minute flow rate with buffer E (20 mM HEPES pH 8.0, 200 mM NaCl, 1 mM β-ME). For each analysis, 200 µL of a protein sample at 2.0 mg/ml concentration was loaded onto the column at 0.6 ml/minute flow rate using a 1260 binary pump (Agilent Technologies). The scattered light was measured with a miniDAWN TREOS II light scatterer, (Wyatt Technologies), and the refractive index was measured with an Optilab T-rEX refractometer, (Wyatt Technologies). Data analysis was performed with ASTRA 7.3.2.21 (Wyatt Technologies) (Slotboom et al. 2008).

2.6.9 Crystallization of H-Rrm4 NT4

Initial crystallization conditions were searched using MRC 3 96-well sitting drop plates and various commercially available crystallization screens at 12 °C. 0.1 µL homogeneous protein solution (10 mg/ml in 20 mM Hepes pH 8.0, 500 mM NaCl) was mixed with 0.1

µL reservoir solution and equilibrated against 40 µL of the reservoir. After one week, initial rod-shaped crystals were found, which were then further optimized by slightly varying the precipitant concentrations. Optimization was also performed in sitting drop plates (24-well) at 12 °C but by mixing 1 µL protein solution with 1 µL of the reservoir solution, equilibrated against 300 µL reservoir solution. Best diffracting crystals were grown within 7 days in 0.1 M Hepes pH 7.5, 20% (w/v) PEG 10000 (Qiagen PEG I, D5). Before harvesting the crystal, crystal-containing drops were overlaid with 2 µL mineral oil and immediately flash-frozen in liquid nitrogen.

2.6.10 Data collection, processing, and structure refinement

A complete data set of the H-Rrm4-NT4 were collected at beamline ID23EH1 (ESRF, France) at 100 K and wavelength 0.98 Å up to 2.6 Å resolution. All data were processed using the automated pipeline at the EMBL HAMBURG and reprocessed afterwards using XDS (KABSCH 2014). Above obtained model for MLLE2^{Rrm4} by TopModel was successfully used to phase the 2.6 Å data set of Rrm4 MLLE using the program Phaser from the program suite Phenix (AFONINE *et al.* 2012). The structure was then refined in iterative cycles of manual building and refinement in Coot (EMSLEY AND COWTAN 2004), followed by software-based refinements using the program suite Phenix (AFONINE *et al.* 2012). All residues were in the preferred and additionally allowed regions of the Ramachandran plot (S2 Table). The data collection and refinement statistics are listed in S2 Table. The structure and models were compared using the superpose tool of PHENIX to calculate the corresponding RMSD. The images of the models were prepared using PyMOL. The structure was deposited at the worldwide protein data bank under the accession code 7PZE.

2.6.11 Small-angle X-ray scattering

We collected all SAXS data on beamline BM29 at the ESRF Grenoble (PERNOT *et al.* 2013). The beamline was equipped with a PILATUS 2M detector (Dectris) with a fixed sample to a distance of 2.827 m. To prevent concentration-dependent oligomerization, we performed the measurements with 0.6 mg/ml protein concentrations at 10 °C in buffer E. We collected one frame each second and scaled the data to absolute intensity against water. All used programs for data processing were part of the ATSAS Software package (Version 3.0.3) (MANALASTAS-CANTOS *et al.* 2021). The primary data reduction was performed with the program Primus (KONAREV *et al.* 2003). With Primus and the included

Guinier approximation (FRANKE AND SVERGUN 2009), we determined the forward scattering *I(0)* and the radius of gyration (R_g). The pair-distribution function *p(r)* was calculated with Gnom (SVERGUN 1992) and was used to estimate the maximum particle dimension (D_{max}). Due to the high flexibility of the proteins we performed an Ensemble Optimization Method (EOM) (TRIA *et al.* 2015); default parameters, 10,000 models in the initial ensemble, native-like models, constant subtraction allowed) with the predicted MLLE domains from TopModel (MULNAES *et al.* 2020; MULNAES *et al.* 2021) for H-Rrm4-NT4 and G-Rrm4 with an additional GST (PDB ID: 1UA5). We uploaded the data to the Small Angle Scattering Biological Data Bank (SASBDB) (VALENTINI *et al.* 2015; KIKHNEY *et al.* 2020) with the accession codes SASDMS5 (G-Rrm4) and SASDMT5 (H-Rrm4-NT4).

2.6.12 Isothermal titration calorimetry

All ITC experiments were performed as per the previous report (ABTS *et al.* 2013). All the protein samples used in ITC were centrifuged at 451,000 × g for 30 minutes and quantified by Nanodrop (A280) before use. The concentration of GST or His-tagged MLLE variants was adjusted to 30 µM and PAM2 peptide variants was adjusted to 300 µM using buffer G (20 mM HEPES pH 8.0, 200 mM NaCl, 1 mM 2 ME). Using an MicroCal iTC200 titration calorimeter (Malvern Panalytical technologies), a PAM2 peptide variant with a volume of 40 µL was titrated to the different GST-MLLE variants. All experiments were repeated at least twice. ITC measurements were performed at 25 °C with 40 injections (1 µL each). Only the first injection had a volume of 0.5 µL and was discarded from the isotherm. The other technical parameters were reference power, 5 µcal s^{-1}; stirring speed, 1000 rpm; spacing time, 120 s, and a filter period, 5 s. The resulting isotherm was fitted with a one-site binding model using MicroCal Origin for ITC software (MicroCal LLC). Note, that the binding of GST-Rrm4-NT4 and H-Rrm4-NT4 were comparable indicating that tagging of the Rrm4 versions did not influence the binding affinity (Figures 3B; S5D).

2.6.13 Microscopy, image processing and image analysis

Laser-based epifluorescence-microscopy was performed on a Zeiss Axio Observer.Z1 as previously described (JANKOWSKI *et al.* 2019). Co-localization studies of dynamic processes were carried out with a two-channel imager (DV2, Photometrics, Tucson, AZ, USA) (BAUMANN *et al.* 2016). To quantify uni- and bipolar hyphal growth, cells were

grown in 30 ml cultures to an OD_{600} of 0.5, and hyphal growth was induced. After 6 hours, more than 100 hyphae were analyzed per strain towards their growth behaviour (n = 3). Cells were assessed for unipolar and bipolar growth as well as the formation of a basal septum. To analyze the signal number, velocity, and travelled distance of fluorescently labelled proteins, movies with an exposure time of 150 ms and 150 frames were recorded. More than 25 hyphae were analyzed per strain (n = 3). To inhibit microtubule polymerization, hyphal cultures were incubated with 50 µM of benomyl (Sigma Aldrich) for 2 h at 28 °C and 200 rpm (BECHT *et al.* 2006). All movies and images were processed and analyzed using the Metamorph software (Version 7.7.0.0, Molecular Devices, Seattle, IL, USA). For the generation of kymographs, 20 µM of hyphal cell starting at the hyphal tip were used. To determine the minimum and maximum grey level intensities of shuttling endosomes, 100 signals were analyzed per strain (the ten most prominent signals per kymograph that showed processive movement of > 20 µM without changes in directions were chosen per strain). The minimum and maximum grey level intensities were measured using the region measurement tool of the Metamorph software. All pixel intensities were measured, and minimum as well as maximum intensities for each region were listed (16-bit images). The grey level intensities were normalized to the wild-type intensity, which was set to 100%. For statistical analysis of the signal number, velocity, and travelled distance, processive signals with a travelled distance of more than 5 µM were conducted and counted manually. For determination of aberrant microtubule staining, hyphae were counted manually as well. Data points represent means from three independent experiments ($n = 3$) with mean of means (red line) and SEM. For all statistical evaluations, two-tailed Student´s *t*-tests were used. Determination of strains exhibiting aberrant staining of microtubules was scored manually. For verification, key comparisons were evaluated independently by two experimentalists. Importantly, the key findings were confirmed (S5H Figure). We used the data obtained by the more experienced microscopist in the main figure (Figure 5D). All evaluated data are compiled in Spreadsheat S1.

2.7 Appendix I

2.7.1 Table S1. Accession numbers for protein sequences used in multiple sequence alignment of MLLE domains

Organism Name	Protein Name	Domain Name	Uniprot KB Number	Sequence coverage
Homo sapiens	Poly[A] binding protein, PABP	$MLLE^{PABP}$	P11940	554 - 617
Triticum aestivum	Poly[A] binding protein, PABP	$MLLE^{PABP}$	P93616	564 - 627
Trypanosoma cruzi	Poly[A] binding protein, PABP	$MLLE^{PABP}$	Q27335	484 - 547
Leishmania major	Poly[A] binding protein, PABP	$MLLE^{PABP}$	E9AFX7	494 - 557
Saccharomyces cerevisiae	Poly[A] binding protein, PABP	$MLLE^{PABP}$	P04147	501 - 566
Homo sapiens	E3 ubiquitin-protein ligase UBR5, EDD	$MLLE^{Ubr5}$	O95071	2390 - 2452
Rattus norvegicus	E3 ubiquitin-protein ligase UBR5	$MLLE^{Ubr5}$	Q62671	2380-2442
Ustilago maydis	Poly[A] binding protein, Pab1	$MLLE^{Pab1}$	Q4P8R9	567 - 630
Ustilago maydis	Rrm4	$MLLE3^{Rrm4}$	A0A0D1DWZ5	727 - 792
Ustilago maydis	Rrm4	$MLLE2^{Rrm4}$	A0A0D1DWZ5	564 - 629
Ustilago maydis	Rrm4	$MLLE1^{Rrm4}$	A0A0D1DWZ5	462 – 528

2.7.2 Table S2. Data collection and refinement statistics

Parameters	$MLLE2^{Rrm4}$
Wavelength	0.979340
Resolution range	33.51 - 2.6 (2.693 - 2.6)
Space group	P 43 21 2
Unit cell	53.455 53.455 144.873 90 90 90
Total reflections	50419 (4828)
Unique reflections	6898 (665)
Multiplicity	7.3 (7.3)
Completeness (%)	98.34 (98.08)
Mean I/sigma(I)	17.90 (2.79)
Wilson B-factor	70.38
R-merge	0.08252 (0.9043)
R-meas	0.08902 (0.9692)
R-pim	0.03265 (0.3441)
CC1/2	0.997 (0.859)
CC*	0.999 (0.961)
Reflections used in refinement	6879 (665)
Reflections used for R-free	688 (65)
R-work	0.2189 (0.3054)
R-free	0.2646 (0.3718)
CC(work)	0.969 (0.844)
CC(free)	0.970 (0.761)
Number of non-hydrogen atoms	1003
Macromolecules	1002
Solvent	1
Protein residues	131
RMS(bonds)	0.010
RMS(angles)	1.24
Ramachandran favored (%)	96.85
Ramachandran allowed (%)	2.36
Ramachandran outliers (%)	0.79
Rotamer outliers (%)	1.79
Clashscore	5.87
Average B-factor	76.62
Macromolecules	76.62
Solvent	78.27

Statistics for the highest-resolution shell are shown in parentheses.

2.7.3 Table S3. Overall SAXS Data

SAXS Device	**BM29, ESRF Grenoble [1, 2]**	
Data collection parameters		
Detector	PILATUS 2 M	
Detector distance (m)	2.827	
Beam size	200 µM x 200 µM	
Wavelength (nm)	0.099	
Sample environment	Quartz capillary,1 mm ø	
s range (nm-1)‡	0.025–6.0	
Exposure time per frame (s)	1 (10 frames each concentration)	
Sample	**H-Rrm4-NT4**	**G-Rrm4**
Organism	*Ustilago maydis*	*Ustilago maydis*
UniProt ID and range	A0A0D1DWZ5	A0A0D1DWZ5
Mode of measurement	Batch	Batch
Temperature (°C)	10	10
Protein buffer	20mM Hepes pH 8.0, 200mM NaCl, 1mM βME	20mM Hepes pH 8.0, 200mM NaCl, 1mM βME
Protein concentration (mg/ml)	0.6	0.6
Structural parameters		
I(0) from P(r)	41.90	103.50
R_g (real-space from P(r)) (nm)	5.55	8.99
I(0) from Guinier fit	43.34	104.01
s-range for Guinier fit (nm-1)	0.060 – 0.230	0.054 – 0.147
R_g (from Guinier fit) (nm)	5.60	8.78
Points from Guinier fit	4 - 37	3 - 21
Dmax (nm)	18.49	30.74
POROD volume estimate (nm3)	122.73	586.72
Molecular mass (kDa)		
From I(0)	43.34	104.01
From MoW2 [3]	22.14	74.22
From Vc [4]	36.57	147.88
From POROD	61.37 – 76.71	293.36 – 366.70
From sequence	40.33	110.95
Structure Evaluation		
EOM fit $\chi 2$	1.262	1.289
Ambimeter score	2.307	2.530
Software		
ATSAS Software Version [5]	3.0.3	
Primary data reduction	PRIMUS [6]	
Data processing	GNOM [7]	
Ensemble modelling	EOM [8]	
Structure evaluation	AMBIMETER [9]	
Model visualization	PyMOL [10]	

‡$s = 4\pi\sin(\theta)/\lambda$, 2θ – scattering angle, λ – X ray-wavelength

Reference

1. Pernot, P., et al., *New beamline dedicated to solution scattering from biological macromolecules at the ESRF.* Journal of Physics: Conference Series, 2010. **247**(1): p. 012009.
2. Pernot, P., et al., *Upgraded ESRF BM29 beamline for SAXS on macromolecules in solution.* J Synchrotron Radiat, 2013. **20**(Pt 4):p. 660-4.
3. Fischer, H., et al., *Determination of the molecular weight of proteins in solution from a single small-angle X-ray scattering measurement on a relative scale.* Journal of Applied Crystallography, 2010. **43**: p. 101-109.
4. Rambo, R.P. and J.A. Tainer, *Accurate assessment of mass, models and resolution by small-angle scattering.* Nature, 2013. **496**(7446): p. 477-81.
5. Manalastas-Cantos, K., et al., *ATSAS 3.0: expanded functionality and new tools for small-angle scattering data analysis.* Journal of Applied Crystallography, 2021. **54**(1).
6. Konarev, P.V., et al., *PRIMUS: a Windows PC-based system for small-angle scattering data analysis.* Journal of Applied Crystallography, 2003. **36**: p. 1277-1282.
7. Svergun, D.I., *Determination of the Regularization Parameter in Indirect-Transform Methods Using Perceptual Criteria.* Journal of Applied Crystallography, 1992. **25**: p. 495-503.
8. Tria, G., et al., *Advanced ensemble modelling of flexible macromolecules using X-ray solution scattering.* IUCrJ, 2015. **2**(Pt 2): p.207-17.
9. Petoukhov, M.V. and D.I. Svergun, *Ambiguity assessment of small-angle scattering curves from monodisperse systems.* Acta Crystallogr D Biol Crystallogr, 2015. **71**(Pt 5): p. 1051-8.
10. PyMOL, *The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC.* 2015

2.7.4 Table S4. Description of *U. maydis* strains used in this study

Strain name with code	Locus	Progenitor strain	Short description
AB33 (UMa133)	*b*	FB2	*Pnar:bW2bE1*, expression of active b heterodimer under control of the *nar1* promoter, strain grows filamentous upon changing the nitrogen source.
AB33rrm4Δ/upa1-gfp (UMa2769)	*rrm4* *upa1*	AB33rrm4-Cherry/ upa1-gfp	carrying a deletion of *rrm4* and expressing Upa1 C-terminally fused to eGfp
AB33upa1-gfp/rrm4-kat (UMa2976)	*rrm4* *upa1*	AB33rrm4Δ/upa1-gfp	expressing Upa1 C-terminally fused to eGfp and Rrm4 C-terminally fused to mKate2
AB33upa1-gfp/rrm4-m1Δ-kat (UMa2977)	*rrm4* *upa1*	AB33rrm4Δ/upa1-gfp	expressing Upa1 C-terminally fused to eGfp and Rrm4-M1Δ-C-terminally fused to mKate2. Like rm4-kat but carrying the deletion of 1st MLLE domain. Residues of Rrm4 from 447 to 540 were replaced with a HAtag-HRV3C protease recognition site.
AB33upa1-gfp/rrm4-m2Δ-kat (UMa2978)	*rrm4* *upa1*	AB33rrm4Δ/upa1-gfp	expressing Upa1 C-terminally fused to eGfp and Rrm4-M2Δ-C-terminally fused to mKate2. Like rrm4-kat but carrying the deletion of 2nd MLLE domain. Residues of Rrm4 from 547 to 644 were replaced with a HAtag-HRV3C protease recognition site.
AB33upa1-gfp/rrm4-m3Δ-kat (UMa2979)	*rrm4* *upa1*	AB33rrm4Δ/upa1-gfp	expressing Upa1 C-terminally fused to eGfp and Rrm4-M3Δ-C-terminally fused to mKate2. Like rrm4-kat but carrying the deletion of 3rd MLLE domain. Residues of Rrm4 from 689 to 792 were replaced with a HAtag-HRV3C protease recognition site.
AB33upa1-gfp/rrm4-m1,2Δ-kat (Uma2981)	*rrm4* *upa1*	AB33rrm4Δ/upa1-gfp	expressing Upa1 C-terminally fused to eGfp and Rrm4-M1,2Δ-C-terminally fused to mKate2. Like rrm4-kat but carrying the deletion of 1st and 2nd MLLE domains. Residues of Rrm4 from 447 to 644 were replaced with a HAtag-HRV3C protease recognition site.
AB33upa1-pl1m-gfp/rrm4-m1,2Δ-kat (Uma2982)	*rrm4* *upa1*	AB33rrm4Δ/upa1-pl1m-gfp	expressing Upa1-PL1m- C-terminally fused to eGfp, carries block mutations leading to the amino acid substitutions AA-SAAATAAS from residues 242-251 in the N-terminal PAM2L-motif (PAM2L-1). Rrm4-M1,2Δ C-terminally fused to mKate2 and carrying the deletion of 1st and 2nd MLLE domains. Residues of Rrm4 from 447 to 644 were replaced with a HAtag-HRV3C protease recognition site.
AB33upa1-pl2m-gfp/rrm4-m1,2Δ-kat (Uma2983)	*rrm4* *upa1*	AB33rrm4Δ/upa1-pl2m-gfp	expressing Upa1-PL2m- C-terminally fused to eGfp, carries block mutations leading to the amino acid substitutions AA-SAAATAAS from residues 949-958 in the C-terminal PAM2L-motif (PAM2L-2). Rrm4-M1,2Δ C-terminally fused to mKate2 and carrying the deletion of 1st and 2nd MLLE domains. Residues of Rrm4 from 447 to 644 were replaced with a HAtag-HRV3C protease recognition site.
AB33upa1-pl1,2m-gfp/rrm4-m1,2Δ-kat (Uma3177)	*rrm4* *upa1*	AB33rrm4Δ/upa1-pl1,2m-gfp	expressing Upa1-PL1,2m C-terminally fused to eGfp, carries block mutations leading to the amino acid substitutions AA-SAAATAAS from residues 242-251 in the N-terminal PAM2L-motif (PAM2L-1) and from residues 949-958 in the C-terminal PAM2L-motif (PAM2L-2). Rrm4-M1,2Δ C-terminally fused to mKate2 and carrying the deletion of 1st and 2nd MLLE domains. Residues of Rrm4 from 447 to 644 were replaced with a HAtag-HRV3C protease recognition site.
AB33upa1Δ/rrm4-kat (Uma3179)	*rrm4* *upa1*	AB33upa1-gfp/rrm4-kat	Carrying a deletion of *upa1* and Rrm4 C-terminally fused to mKate2
AB33upa1-pl1,2m-gfp/rrm4-kat (Uma3355)	*rrm4* *upa1*	AB33rrm4Δ/upa1-pl1,2m-gfp	expressing Upa1-PL1,2m C-terminally fused to eGfp, carries block mutations leading to the amino acid substitutions AA-SAAATAAS from residues 242-251 in the N-terminal PAM2L-motif (PAM2L-1) and from residues 949-958 in the

			C-terminal PAM2L-motif (PAM2L-2). Rrm4 C-terminally fused to mKate2.
AB33upa1-pl1m-gfp/rrm4-kat (UL46)	*rrm4* *upa1*	AB33rrm4Δ/upa1-pl1m-gfp	expressing Upa1-PL1m C-terminally fused to eGfp, carries block mutations leading to the amino acid substitutions AA-SAAATAAS from residues 242-251 in the N-terminal PAM2L-motif (PAM2L-1) and Rrm4 C-terminally fused to mKate2.
AB33upa1-pl2m-gfp/rrm4-kat (UL47)	*rrm4* *upa1*	AB33rrm4Δ/upa1-pl2m-gfp	expressing Upa1-PL2m C-terminally fused to eGfp, carries block mutations leading to the amino acid substitutions AA-SAAATAAS from residues 949-958 in the C-terminal PAM2L-motif (PAM2L-2) and Rrm4 C-terminally fused to mKate2.
AB33upa1-pl2m-gfp/rrm4-m1Δ-kat (UL48)	*rrm4* *upa1*	AB33rrm4Δ/upa1-pl1,2m-gfp	expressing Upa1-PL2m C-terminally fused to eGfp, carries block mutations leading to the amino acid substitutions AA-SAAATAAS from residues 949-958 in the C-terminal PAM2L-motif (PAM2L-2). Rrm4-M1Δ C-terminally fused to mKate2. Like rrm4-kat but carrying the deletion of 1st MLLE domain. Residues of Rrm4 from 447 to 540 were replaced with a HAtag-HRV3C protease recognition site.
AB33upa1-pl2m-gfp/rrm4-m1,2Δ-kat (UL49)	*rrm4* *upa1*	AB33rrm4Δ/upa1-pl2m-gfp	expressing Upa1-PL2m C-terminally fused to eGfp, carries block mutations leading to the amino acid substitutions AA-SAAATAAS from residues 949-958 in the C-terminal PAM2L-motif (PAM2L-2). Rrm4-M1,2Δ C-terminally fused to mKate2 and carrying the deletion of 1st and 2nd MLLE domains. Residues of Rrm4 from 447 to 644 were replaced with a HAtag-HRV3C protease recognition site.

2.7.5 Table S5. Generation of *U. maydis* strains used in this study

Strains	Relevant genotype	Strain code	Refer-ence	Transformed plasmid	Locus	Progenitor
AB33	*a2 P_{nar}:bW2 bE1*	UMa 133	[1]	pAB33	*b*	FB2
AB33rrm4Δ/upa1-gfp	*rrm4 upa1-gfp*	UMa 2769	this study	pRrm4Δ_genitR (pUMa1755)	*rrm4*	AB33rrm4-mCherry/upa1-gfp (UMa1594)
AB33upa1-gfp/rrm4-kat	*upa1-gfp rrm4-kat*	Uma 2976	this study	pRrm4-kat-hygR (pUMa3908)	*rrm4*	AB33rrm4Δ/upa1-gfp (UMa2769)
AB33upa1-gfp/rrm4-m1Δ-kat	*upa1-gfp rrm4*-m1Δ-kat	UMa 2977	this study	pRrm4-m1Δ-kat-hygR (pUMa4433)	*rrm4*	AB33rrm4Δ/upa1-gfp (UMa2769)
AB33upa1-gfp/rrm4-m2Δ-kat	*upa1-gfp rrm4-m2Δ-kat*	UMa 2978	this study	pRrm4- m2Δ-kat-hygR (pUMa4434)	*rrm4*	AB33rrm4Δ/upa1-gfp (UMa2769)
AB33upa1-gfp/rrm4-m3Δ-kat	*upa1-gfp rrm4-m3Δ-kat*	UMa 2979	this study	pRrm4-m3Δ-kat-hygR (pUMa4435)	*rrm4*	AB33rrm4Δ/upa1-gfp (UMa2769)
AB33upa1-gfp/rrm4-m1,2Δ-kat	*upa1-gfp rrm4-m1,2Δ-kat*	UMa 2981	this study	pRrm4-m1,2Δ-kat-hygR (pUMa4578)	*rrm4*	AB33rrm4Δ/upa1-gfp (UMa2769)
AB33upa1-pl1m-gfp/ rrm4- m1,2Δ-kat	*upa1-pl1m-gfp rrm4-m1,2Δ-kat*	UMa 2982	this study	pRrm4-m1,2Δ-kat-hygR (pUMa4578)	*rrm4*	AB33rrm4Δ/upa1-pl1m-gfp (UMa2766)
AB33upa1-pl2m-gfp/ rrm4-m1,2Δ-kat	*upa1-pl2m-gfp rrm4-m1,2Δ-kat*	UMa 2983	this study	pRrm4-m1,2Δ-kat-hygR (pUMa4578)	*rrm4*	AB33rrm4Δ/upa1-pl2m-gfp (UMa2767)
AB33upa1-pl1,2m-gfp/ rrm4-m1,2Δ-kat	*upa1-pl1,2m-gfprrm4-m1,2Δ-kat*	UMa 3177	this study	pRrm4-m1,2Δ-kat-hygR (pUMa4578)	*rrm4*	AB33rrm4Δ/upa1-pl1,2m-gfp (UMa2768)
AB33upa1Δ/rrm4-kat	*upa1Δ rrm4-kat*	UMa 3179	this study	pUpa1Δ-genitR (pUMa1915)	*upa1*	AB33upa1-gfp/rrm4-kat (UMa2976)
AB33upa1-pl1,2m-gfp/ rrm4-kat	*upa1-pl1,2m-gfp rrm4-kat*	UMa 3355	this study	pRrm4-kat-hygR (pUMa3908)	*rrm4*	AB33rrm4Δ/upa1-pl1,2m-gfp (UMa2768)
AB33upa1-pl1m- gfp/ rrm4-kat	*upa1-pl1m-gfp rrm4-kat*	UL46	this study	pRrm4-kat-hygR (pUMa3908)	*rrm4*	AB33rrm4Δ/upa1-pl1m-gfp (UMa2766)
AB33upa1-pl2m-gfp/ rrm4-kat	*upa1-pl2m-gfprrm4-kat*	UL47	this study	pRrm4-kat-hygR (pUMa3908)	*rrm4*	AB33rrm4Δ/upa1-pl2m-gfp (UMa2767)
AB33upa1-pl2m-gfp/ rrm4-m1Δ-kat	*upa1-pl2m-gfp rrm4-m1Δ-kat*	UL48	this study	pRrm4-m1Δ-kat-hygR (pUMa4433)	*rrm4*	AB33rrm4Δ/upa1-pl2m-gfp (UMa2767)

AB33upa1-pl2m-gfp/ rrm4-m2Δ-kat	*upa1-pl2m-gfp rrm4-m2Δ-kat*	UL49	this study	pRrm4-m2Δ-kat-hygR (pUMa4434)	*rrm4*	AB33rrm4Δ/upa1-pl2m-gfp (UMa2767)

Reference

Brachmann A, Weinzierl G, Kämper J, Kahmann R. Identification of genes in the bW/bE regulatory cascade in *Ustilago maydis*. Mol Microbiol. 2001 42:1047-63. https://doi.org/10.1046/j.1365-2958.2001.02699.x. PMID: 11737646.

2.7.6 Table S6. Description of plasmids used for *U. maydis* strain generation

Plasmid	pUMa	Resistance cassette	Short description
pRrm4Δ	1755	genitR (G418 resistance - SfiI insert of pMF1g) [1]	Plasmid vector for generating deletion mutants of *rrm4*.
pUpa1Δ_genitR	1915	genitR [2]	Plasmid vector for generating deletion mutants of *upa1*.
pRrm4-kat-hygR	3908	hygR (Hygromycin resistance - SfiI insert of pMF1h) [3]	Plasmid vector for the expression of Rrm4 C-terminally fused to mKate2. The mKate2 cassette contains the Tnos terminator and the Hyg resistance. The entire coding sequence for the fusion protein is flanked by a 1025 bp upstream region and a 1396 bp downstream region for homologous recombination.
pRrm4-m1Δ-kat-hygR	4433	hygR	Plasmid vector for the expression of Rrm4-M1Δ C-terminally fused to mKate2. Like pRrm4-mK-HygR, but carrying the deletion of 1st MLLE domain. Residues of Rrm4 from 447 to 540 were replaced with a HAtag-HRV3C protease recognition site.
pRrm4-m2Δ-kat-hygR	4434	hygR	Plasmid vector for the expression of Rrm4-M2Δ C-terminally fused to mKate2. Like pRrm4-mK-HygR, but carrying the deletion of the 2nd MLLE domain. Residues of Rrm4 from 547 to 644 were replaced with a HAtag-HRV3C protease recognition site.
pRrm4-m3Δ-kat-hygR	4435	hygR	Plasmid vector for the expression of Rrm4-M3Δ C-terminally fused to mKate2. Like pRrm4-mK-HygR, but carrying the deletion of the 3rd MLLE domain. Residues of Rrm4 from 689-792 were replaced with a HAtag-HRV3C protease recognition site.
pRrm4-m1,2Δ-kat-hygR	4578	hygR	Plasmid vector for the expression of Rrm4-M1,2Δ C-terminally fused to mKate2. Like pRrm4-mK-HygR, but carrying the deletion of 1st and 2nd MLLE domains. Residues of Rrm4 from 447 to 644 were replaced with a HAtag-HRV3C protease recognition site.

Reference

1. Baumann S, Pohlmann T, Jungbluth M, Brachmann A, Feldbrügge M. Kinesin-3 and dynein mediate microtubule-dependent co-transport of mRNPs and endosomes. J Cell Sci. 2012; 125:2740-52. https://doi.org/10.1242/jcs.101212 PMID: 22357951
2. Pohlmann T, Baumann S, Haag C, Albrecht M, Feldbrügge M. A FYVE zinc finger domain protein specifically links mRNA transport to endosome trafficking. Elife. 2015 4:e06041. https://doi.org/10.7554/eLife.06041. PMID: 25985087
3. Brachmann A, König J, Julius C, Feldbrügge M. A reverse genetic approach for generating gene replacement mutants in *Ustilago maydis*. Mol Genet Genomics. 2004 272:216-26. https://doi.org/10.1007/s00438-004-1047-z. PMID: 15316769

2.7.7 Table S7. Description of plasmids used for recombinant expression in *E. coli*

Plasmid	pUMa	Short description
pGEX-G-Pab1-MLLE	2187	Plasmid for the expression of the G-Pab1-MLLE. C-terminal region of Pab1 comprising amino acid residues 566-651 were N-terminally fused to a GST-tag. [1]
pGEX-G-Rrm4-NT4	3920	Plasmid for the expression of the G-Rrm4-NT4. C-terminal region of Rrm4 comprising amino acid residues 421 to 792 was N-terminally fused to a GST-tag.
pGEX-G-Rrm4-NT4-M1Δ	4616	Plasmid for the expression of the G-Rrm4-NT4-M1Δ. Same as pUMa3920 but carrying the deletion of 1st MLLE domain. Residues of Rrm4 from 447 to 540 were replaced with a HAtag-HRV3C protease recognition site.
pGEX-G-Rrm4-NT4-M2Δ	4617	Plasmid for the expression of the G-Rrm4-NT4-M2Δ. Same as pUMa3920 but carrying the deletion of the 2nd MLLE domain. Residues of Rrm4 from 547 to 644 were replaced with a HAtag-HRV3C protease recognition site.
pGEX-G-Rrm4-NT4-M3Δ	4618	Plasmid for the expression of the G-Rrm4-NT4-M3Δ. Same as pUMa3920 but carrying the deletion of the 3rd MLLE domain. Residues of Rrm4 from 689 to 792 were replaced with a HAtag-HRV3C protease recognition site.
pGEX-G-Rrm4-NT4-M1,2Δ	4619	Plasmid for the expression of the G-Rrm4-NT4-M1,2Δ. Same as pUMa3920 but carrying the deletion of 1st to 2nd MLLE domains. Residues of Rrm4 from 447 to 644 were replaced with a HAtag-HRV3C protease recognition site.
pET28-HS-PAM2^{Upa1}	4296	Plasmid for the expression of the PAM2 motif of Upa1 (SQSTLSPNASVFKPSRS) as a fusion protein with an N-terminal 6xHis-Sumo-tag.
pET28-HS_PAM2L1^{Upa1}	4297	Plasmid for the expression of PAM2L1 motif of Upa1 (EAADQEEDQDDFVYPGAD) as a fusion protein with an N-terminal 6xHis-Sumo-tag.
pET28-HS-PAM2L2^{Upa1}	4298	Plasmid for the expression of PAM2L2 motif of Upa1 (DEDAADDDDDEFIYPNSY) as a fusion protein with an N-terminal 6xHis-Sumo-tag.
pET22-H-Rrm4-NT4	3552	Plasmid for the expression of H-Rrm4-NT4. C-terminal region of Rrm4 comprising amino acid 421 to 792 were N-terminally fused to 6xHis-tag.
pGX-G-Rrm4	429	Plasmid for the expression of G-Rrm4. Rrm4 full-length protein was N-terminally fused to GST.

Reference

Pohlmann T, Baumann S, Haag C, Albrecht M, Feldbrügge M. A FYVE zinc finger domain protein specifically links mRNA transport to endosome trafficking. Elife. 2015 4:e06041. https//:10.7554/eLife.06041. PMID: 25985087

2.7.8 Table S8. Description of plasmids used for yeast two-hybrid analyses

Plasmid	Plasmid code	Gene	Short description
pGADT7-DS	pUMa1624		Plasmid for the expression of hybrid proteins, N-terminally fused to a nuclear localization signal (NLS) of the simian virus 40 (SV40), followed by the Gal4 activation domain (aa 768-881) and an HA-epitope for Western Blot detection. The resulting hybrid proteins are termed AD-"X". For the positive selection of transformants on minimal medium, this plasmid carries a *LEU2* auxotrophy marker. This plasmid contains two diverse SfiI-restriction sites for cloning purposes (Dualsystems Biotech, Schlieren, Switzerland).
pGBKT7-SfiI MCS	pUMa1625		Plasmid for the expression of hybrid proteins, N-terminally fused to the Gal4 DNA-binding domain (aa 1-147), followed by a c-Myc-epitope for Western Blot detection. The resulting hybrid proteins are termed BD-"X". For the positive selection of transformants on minimal medium, this plasmid carries a *TRP1* auxotrophy marker. This plasmid contains two diverse SfiI-restriction sites for cloning purposes (Clontech Laboratories, Inc., Mountain View, CA, USA).
pGADT7-T	pUMa1636		Plasmid for the expression of an N-terminal AD-fusion of the large T-antigen of SV40. It interacts with BD-p53 as a positive control (Clontech).
pGBKT7-p53	pUMa1638		Plasmid for the expression of an N-terminal BD-fusion of the murine p53. It interacts with AD-T as a positive control (Clontech).
pGBKT7-Lam	pUMa1637		Plasmid for the expression of an N-terminal BD-fusion with the human nuclear protein Lamin C, which shows no interaction with most proteins and serves as negative control (Clontech).
pGBKT7-Upa1-Gfp	pUL0128	*upa1*	Plasmid for the expression of BD-Upa1-Gfp, where eGfp is fused C-terminally to the BD-Upa1-hybrid.
pGBKT7-Upa1-pl1m-Gfp	pUL0120	*upa1*	Like GBKT7-Upa1-Gfp, expressing BD-Upa1-pl1-Gfp, where eGfp is fused C-terminally to the BD-Upa1-pl1m hybrid but carries block mutations leading to the amino acid substitutions AASAAATAAS from residues 242-251 in the N-terminal PAM2L-motif (PAM2L-1) of Upa1.
pGBKT7-Upa1-pl2m-Gfp	pUL0121	*upa1*	Like pGBKT7-Upa1-Gfp, expressing BD-Upa1-pl2-Gfp, where eGfp is fused C-terminally to the BD-Upa1-pl2m hybrid. but carries block mutations leading to the amino acid substitutions AASAAATAAS from residues 949-958 in the C-terminal PAM2L-motif (PAM2L-2) of Upa1.
pGBKT7-Upa1-pl1,2m-Gfp	pUL0122	*upa1*	Like GBKT7-Upa1-Gfp, expressing BD-Upa1-pl1,2-Gfp, where eGfp is fused C-terminally to the BD-Upa1-pl1,2m hybrid but carries block mutations leading to the amino acid substitutions AASAAATAAS from residues 242-251 in the N-terminal PAM2L-motif (PAM2L-1) and from residues 949-958 in the C-terminal PAM2L-motif (PAM2L-2).
pGADT7-Rrm4-kat	pUL0112	*rrm4*	Plasmid for the expression of AD-Rrm4-kat, where mKate2 is fused C-terminally to the AD-Rrm4-hybrid.
pGADT7-Rrm4-m1Δ-kat	pUL0116	*rrm4*	Plasmid for the expression of AD-Rrm4-M1Δ-kat where mKate2 is fused C-terminally to the AD-Rrm4-M1Δ hybrid. Like pGADT7-Rrm4-kat, but carrying the deletion of 1st MLLE domain. Residues of Rrm4

			from 447 to 540 were replaced with a HAtag-HRV3 C protease recognition site.
pGADT7-Rrm4-m2Δ-kat	pUL0117	*rrm4*	Plasmid for the expression of AD-Rrm4-M2Δ-kat, where mKate2 is fused C-terminally to the AD-Rrm4-M2Δ hybrid. Like pGADT7-Rrm4-kat, but carrying the deletion of 2nd MLLE domain. Residues of Rrm4 from 547 to 644 were replaced with a HAtag-HRV3C protease recognition site.
pGADT7-Rrm4-m1,2Δ-kat	pUL0118	*rrm4*	Plasmid for the expression of AD-Rrm4-M1,2Δ-kat, where mKate2 is fused C-terminally to the AD-Rrm4-M1,2Δ hybrid. Like pGADT7-Rrm4-kat, but carrying the deletion of 1st and 2nd MLLE domains. Residues of Rrm4 from 447 to 644 were replaced with a HAtag-HRV3C protease recognition site.
pGADT7-Rrm4-m3Δ-k	pUL0119	*rrm4*	Plasmid for the expression of AD-Rrm4-M3Δ-kat, where mKate2 is fused C-terminally to the AD-Rrm4-M3Δ hybrid. Like pGADT7-Rrm4-kat, but carrying the deletion of 3rd MLLE domain. Residues of Rrm4 from 689-792 were replaced with a HAtag-HRV3C protease recognition site.

2.7.9 Table S9. DNA oligonucleotides used in this study

Designation	Nucleotide sequence (5' --> 3')	Remarks
oUM727	GTATTCGAGCCAAGCATCTACGTATGTCGACCCTTGCAACC	Rrm4-internal-Gibson cloning-fwd
oAB354	GGGCCCCTGGAACAGTACTTCCAGGGCGTAGTCGGGCACGTCGTAAGGG-TAAGGCACACCTGCTTTGAAG	Rrm4-M1Δ_Gibson cloning-rev
oAB355	TACCCTTACGACGTGCCCGACTACGCCCTG-GAAGTACTGTTCCAGGGGCCCCTGTCTGCTGAACACCCAGC	Rrm4-M1Δ_Gibson cloning-fwd
oAB359	CGATCGCCGGGCGGCCGGCGCGCCACCGGTTTAGCGGTGACCGAG-TTTCGAGG	mKate2-rev
oAB345	GGGCCCCTGGAACAGTACTTCCAGGGCGTAGTCGGGCACGTCGTAAGGG-TATGCAGGAAGCGCAGCAAGCG	Rrm4-M3Δ_Gibson cloning-rev
oAB346	TACCCTTACGACGTGCCCGACTACGCCCTG-GAAGTACTGTTCCAGGGGCCCGCGGCCAACGCGGCCACCATGGTG	Rrm4-M3Δ_Gibson cloning-fwd
oAB356	GGGCCCCTGGAACAGTACTTCCAGGGCGTAGTCGGGCACGTCGTAAGGG-TATGGGTGTTCAGCAGACAGTG	Rrm4-M2Δ_Gibson cloning-rev
oAB357	TACCCTTACGACGTGCCCGACTACGCCCTG-GAAGTACTGTTCCAGGGGCCCAGCGCTCCGGTGCCATTGTC	Rrm4-M2Δ-Gibson cloning-fwd
oAB312	CATGCCATGGCCAGCAGCAACAGTCCGCCCAC	NcoI_Rrm4-NT4-fwd
oAB45	CGGCCATATGGGCAGCAGCCATCATC	pET28 vector-ORF-fwd
oAB46	CTCACTCGAGTTAGGATCGGGACGGCTTGAAGACGGAGGCGTTGGGAGA-CAAGGTGCTTTGCGAACCACCAATCTGTTCTCTGTGAGC	Sumo-PAM2-XhoI-rev

oAB47	CTCACTCGAGTTAGTCGGCTCCTGGGTAGACAAAGTCATCTT-GATCTTCCTCTTGGTCTGCAGCCTCACCACCAATCTGTTCTCTGTGAG	Sumo-PAM2L1-XhoI-rev

2.8 Acknowledgements

We thank Drs. Florian Altegoer and Dierk Niessing as well as laboratory members for critical reading of the manuscript. We acknowledge Drs. Georg Groth, Alexander Minges, and Mohanraj Gopalswamy for support with ITC experiments. We gratefully acknowledge the computational support provided by the "Center for Information and Media Technology" (ZIM) at the Heinrich Heine University Düsseldorf and the computing time provided by the John von Neumann Institute for Computing (NIC) on the supercomputer JUWELS at Jülich Supercomputing Centre (JSC) (user IDs: HKF7, VSK33). The synchrotron MX data were collected at ESRF ID23-EH1. We thank Sylvain Engilberge for the assistance in using the beamline. The final SAXS experiments were performed on beamline BM29 at the European Synchrotron Radiation Facility (ESRF), Grenoble, France. We are grateful to Dihia Moussaoui at the ESRF for assisting with data collection.

2.9 Author contributions

Conceptualization: Senthil-Kumar Devan, Stephan Schott-Verdugo, Lilli Bismar, Lutz Schmitt, Sander HJ Smits, Holger Gohlke, Michael Feldbrügge.
Data curation: Senthil-Kumar Devan, Stephan Schott-Verdugo, Kira Müntjes, Lilli Bismar, Jens Reiners, Astrid Höppner, Sander HJ Smits, Michael Feldbrügge.
Funding acquisition: Sander HJ Smits, Holger Gohlke, Michael Feldbrügge.
Investigation: Senthil-Kumar Devan, Stephan Schott-Verdugo, Kira Müntjes, Eymen Hachani.
Methodology: Kira Müntjes, Lilli Bismar, Jens Reiners, Eymen Hachani, Astrid Höppner.
Project Administration: Sander HJ Smits, Holger Gohlke, Michael Feldbrügge.
Software: Jens Reiners.
Supervision: Kira Müntjes, Lilli Bismar, Lutz Schmitt, Sander HJ Smits, Holger Gohlke, Michael Feldbrügge.
Validation: Lutz Schmitt, Astrid Höppner, Sander HJ Smits, Holger Gohlke, Michael Feldbrügge.
Writing – original draft: Senthil-Kumar Devan, Stephan Schott-Verdugo, Sander HJ Smits, Holger Gohlke, Michael Feldbrügge.
Writing – review & editing: Sander HJ Smits, Holger Gohlke, Michael Feldbrügge.

3 Deciphering the structure and molecular basis of peptide recognition by the key MLLE domains in *U. maydis*

3.1 Introduction

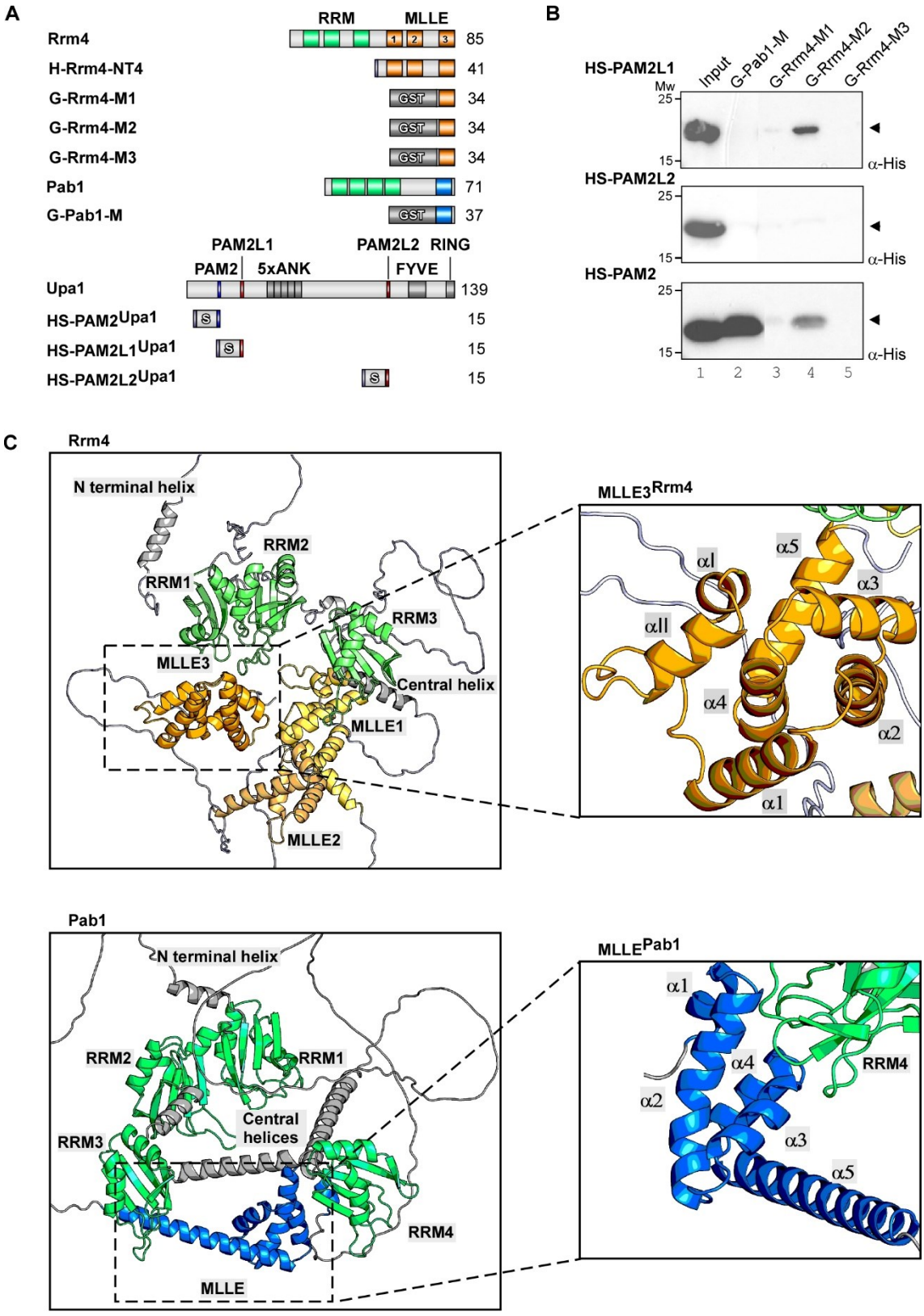
Structure functional studies described in the previous chapter established that the key RBP in endosome-mediated mRNA transport, Rrm4, is consisted of three tandem MLLE domains that form a sophisticated PPI platform. MLLE3^{Rrm4} is the main domain for Upa1 interaction *in vitro* and *in vivo*. MLLE3^{Rrm4} interacted with the PAM2L1,2 motifs of Upa1 specifically, but it did not recognize the PAM2^{Upa1} motifs of Upa1. On the other hand, MLLE^{Pab1} interacted with the PAM2^{Upa1} motif of Upa1 specifically but did not recognize the PAM2L1,2^{Upa1} motifs of Upa1. Regardless, the molecular basis of the peptide recognition by the MLLE^{Pab1} and MLLE3^{Rrm4} needed to be understood. Therefore, structural studies were carried out to obtain the co-crystallized structures of the MLLE3^{Rrm4} and MLLE^{Pab1} with PAM2L1,2^{Upa1} and PAM2^{Upa1}, respectively. This study revealed that MLLE3^{Rrm4} possessed a novel domain architecture that consisted of seven helices and exhibited a non-canonical peptide recognition, whereas MLLE^{Pab1} possessed a typical five-helix structure and exhibited a canonical peptide recognition.

3.2 Structural characterization of MLLE3^{Rrm4} - PAM2L^{Upa1} complex

3.2.1 The third MLLE domain of Rrm4 with five helices is not sufficient for interaction with PAM2-like motifs of Upa1

To evaluate the interaction capacity of each MLLE domain of Rrm4, *in vitro* binding studies were carried out with GST pull-down assay. To this end, all three MLLE domains (MLLE1^{Rrm4}, MLLE2^{Rrm4}, MLLE3^{Rrm4}) of Rrm4 were individually expressed as a fusion protein with an N-terminal GST tag in *E. coli* (Figure 12A; G-Rrm4-M1, G-Rrm4-M2, G-Rrm4-M3; Materials and methods). An N-terminal GST fused MLLE domain of Pab1 was expressed as a control (Figure 12A; G-Pab1-M; Materials and methods). To check the physical interaction with PAM2 and PAM2L sequences of Upa1, 18 amino acid fragments were expressed (Figure 12A) as N-terminal hexa-histidine-SUMO (HS) fusion proteins (Figure 12A; HS-PAM2^{Upa1}, HS-PAM2L1^{Upa1}, HS-PAM2L2^{Upa1}, Materials, and methods).

In GST pull-down experiments using GST fusion proteins as bait, G-Pab1-MLLE interacted with PAM2 but not with the PAM2L motifs of Upa1 (G-Pab1-MLLE; Figure 12B, lane 2) which is in line with previous results (FigS3B and ;(POHLMANN *et al.* 2015; DEVAN *et al.* 2022)). Earlier *in vitro* and *in vivo* studies have established that MLLE3 of Rrm4 is the main interaction partner for both PAM2L1,2^{Upa1} (Figure 8, Figure 9, Figure S3-S7, (DEVAN *et al.* 2022)) and MLLE1,2^{Rrm4} are not interacting with both these peptides. In line with the earlier observation, G-Rrm4-MLLE1 did not interact with any of these peptides (G-Rrm4-M1, Figure 12B, lane 3). In contrast to the earlier observations, G-Rrm4-MLLE2 interacted with both HS-PAM2L1 and 2^{Upa1} in the *in vitro* binding studies (G-Rrm4-M2, Figure 12B, lane 4). This suggested that the presence of MLLE1,3^{Rrm4} might regulate or hinder MLLE2^{Rrm4} interaction. In their absence, MLLE2^{Rrm4} might be freely accessible for interacting with PAM2L1,2^{Upa1} peptides. In addition, G-Rrm4-MLLE2 showed a weak interaction with HS-PAM2^{Upa1} (G-Rrm4-M2 Figure 12B, lane 4), which suggests that MLLE2^{Rrm4} has flexible target specificity. Surprisingly, G-Rrm4-MLLE3 did not interact with both PAM2L1,2^{Upa1} (G-Rrm4-M3, Figure 12B, lane 5), whereas in the earlier experiments, constructs lacking both MLLE1,2^{Rrm4} and having only MLLE3^{Rrm4} interacted with both PAM2L1,2^{Upa1} (Figure 8D, G-Rrm4-NT4-M1,2D, FigS3B, lane 5,(DEVAN *et al.* 2022)). This observation suggested that the G-Rrm4-M3 construct might lack some crucial amino acids necessary for interaction.



3.2.1.1 Figure 12. Five helix structure of MLLE3^{Rrm4} is not sufficient for PAM2L1,2^{Upa1} interaction.

(A) Schematic representation of protein variants (Molecular weight in kilo Dalton indicated) using the following coloring: lime green, RNA recognition motif (RRM); orange, MLLE^{Rrm4} domains; bright blue, MLLE^{Pab1}; dark blue PAM2 motif; light red PAM2L1 motif; dark red PAM2L2 motif; Ankyrin repeats (5xANK), FYVE domain, and RING domain of Upa1 are given in dark grey. GST and SUMO tags are

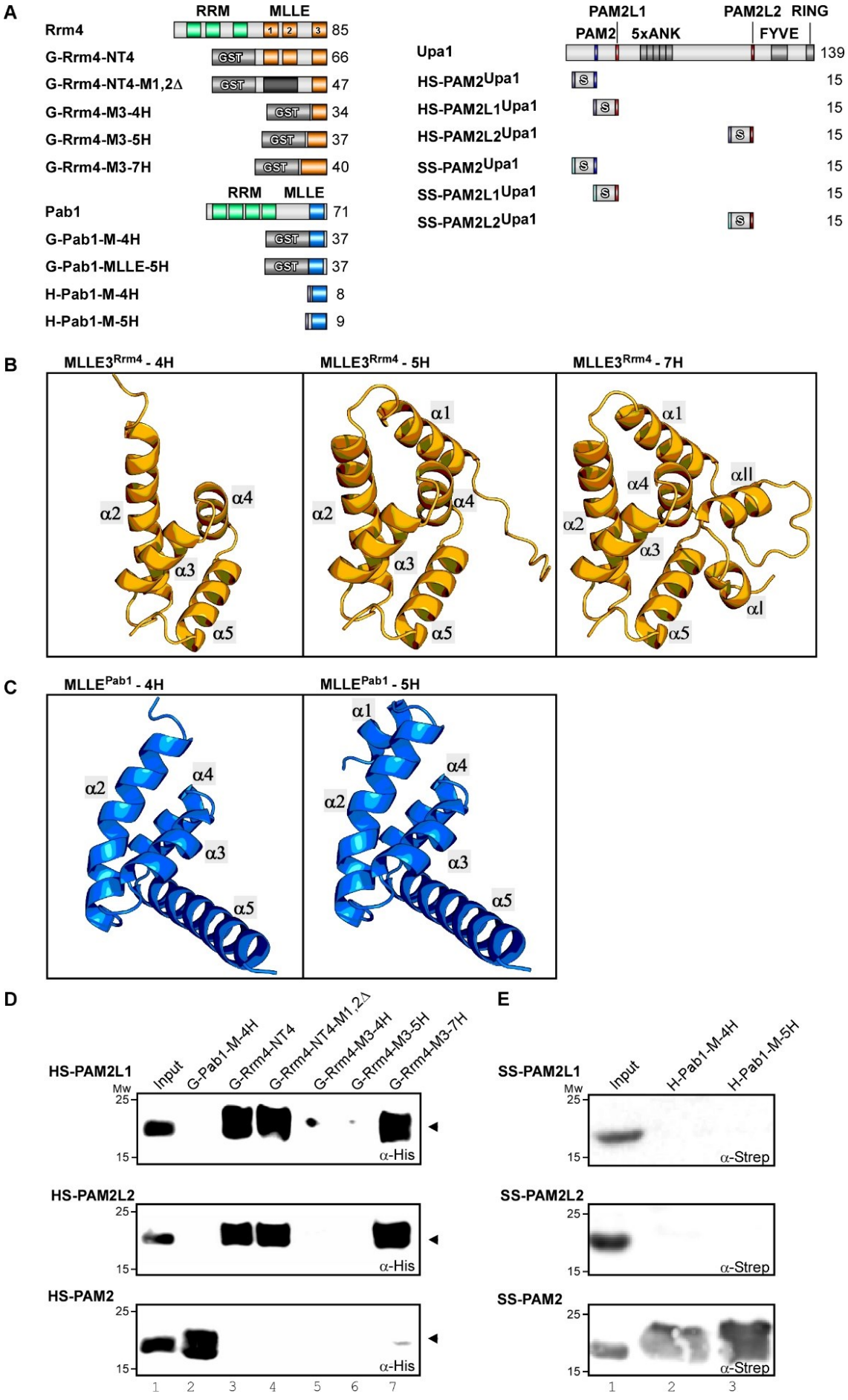
labeled. **(B)** Western blot analysis of GST pull-down experiments with components expressed in *E. coli*: N-terminal hexa-histidine-SUMO-tagged PAM2 variants were pulled-down by N-terminal GST fused MLLE variants of Rrm4 and Pab1. The experiment was performed with the soluble fraction of *E. coli* cell lysate to demonstrate specific binding. Results were analyzed with α-His immunoblotting. **(C)** Three-dimensional (3D) structural models of Rrm4 *(top panel)* and Pab1 *(lower panel)* predicted by AlphaFold2 algorithm. Globular domains are represented as cartoons using the following coloring: lime green, RNA recognition motif (RRM); golden yellow, MLLE1^{Rrm4}; light orange, MLLE2^{Rrm4}; dark orange MLLE3^{Rrm4}; bright blue, MLLE^{Pab1}. The dotted box indicates the MLLE3^{Rrm4} *(top panel)* and MLLE^{Pab1} *(lower panel)*. Magnified images on the right box display the detailed view of MLLE3^{Rrm4} with seven helices *(top panel)* and MLLE^{Pab1} with five helices *(lower panel)*.

To investigate this further, AlphaFold2 predicted Rrm4 structure was analyzed in the structure visualization tool Pymol. Careful inspection revealed that MLLE3^{Rrm4} has two additional helices (αI, II) in the N-terminus of the Topmodel predicted MLLE3^{Rrm4} (α1-5) structure (Figure 12C, Rrm4, MLLE3^{Rrm4}). However, the MLLE domain of Pab1 from the AlphaFold2 predicted structure exhibited the typical a 5-helices (α1-5) without any additional helices (Figure 12C, Pab1, MLLE^{Pab1}). These observations indicated that 7-helix MLLE domain structure might be specific for Rrm4. In summary, MLLE3^{Rrm4} with 5-helices is insufficient for interacting with neither PAM2L1^{Upa1} nor PAM2L2^{Upa1} *in vitro,* and AlphaFold2 structure identified two additional helices (αI, II) on the N-terminus of MLLE3^{Rrm4}.

3.2.2 The N-terminal helices (αI, II) of third MLLE domain of Rrm4 are essential for interaction with PAM2L1,2 motifs of Upa1

In vitro binding studies were performed using GST pull-down assay with three different G-MLLE3^{Rrm4} versions to verify the role of αI, II from MLLE3 domain of Rrm4, identified from the AlphaFold2 prediction (Figure 13A, G-Rrm4-M3-4H, G-Rrm4-M3-5H, G-Rrm4-M3-7H, Materials, and Methods) at different length of MLLE3^{Rrm4} comprising the 4, 5 or 7 helices. MLLE3^{Rrm4} versions were generated as a fusion protein with an N-terminal GST tag. G-Pab1-MLLE, G-Rrm4-NT4, and G-Rrm4-NT4-M1,2D were used as control (Figure 13A; Materials and methods). AlphaFold2 predicted models of the newly designed constructs depicted that deletion of αI, II did not affect the overall protein folding of MLLE3^{Rrm4} (Figure 13B, MLLE3^{Rrm4}-4H, MLLE3^{Rrm4}-5H, MLLE3^{Rrm4}-7H). To check the physical interaction with PAM2 and PAM2L sequences of Upa1, HS-PAM2^{Upa1}, and HS-PAM2L^{Upa1} constructs were used as described earlier.

In GST pull-down experiments using GST fusion proteins as bait, as expected, G-Pab1-MLLE interacted with HS-PAM2^{Upa1} but not with the HS-PAM2L^{Upa1} (Figure 13D, G-Pab1-MLLE, lane 2). G-Rrm4-NT4 version carrying all the three MLLE domains and


A
RRM
MLLE
Rrm4
1
2
3
85
G-Rrm4-NT4
GST
66
G-Rrm4-NT4-M1,2Δ
47
G-Rrm4-M3-4H
34
G-Rrm4-M3-5H
37
G-Rrm4-M3-7H
40
Pab1
71
G-Pab1-M-4H
37
G-Pab1-MLLE-5H
37
H-Pab1-M-4H
8
H-Pab1-M-5H
9
PAM2L1
PAM2
5xANK
PAM2L2
RING
FYVE
Upa1
139
HS-PAM2Upa1
S
15
HS-PAM2L1Upa1
15
HS-PAM2L2Upa1
15
SS-PAM2Upa1
15
SS-PAM2L1Upa1
15
SS-PAM2L2Upa1
15
B
MLLE3Rrm4 - 4H
MLLE3Rrm4 - 5H
MLLE3Rrm4 - 7H
α1
α2
α3
α4
α5
αII
αI
C
MLLEPab1 - 4H
MLLEPab1 - 5H
D
HS-PAM2L1
Input
G-Pab1-M-4H
G-Rrm4-NT4
G-Rrm4-NT4-M1,2Δ
G-Rrm4-M3-4H
G-Rrm4-M3-5H
G-Rrm4-M3-7H
Mw
25
15
α-His
HS-PAM2L2
HS-PAM2
1
2
3
4
5
6
7
E
SS-PAM2L1
H-Pab1-M-4H
H-Pab1-M-5H
α-Strep
SS-PAM2L2
SS-PAM2

3.2.2.1 Figure 13. The N-terminal αI, II of MLLE3^{Rrm4} are important for recognition of PAM2L1,2^{Upa1} *in vitro.*

(A) Schematic representation of protein variants (Molecular weight in kilo Dalton indicated) using the following coloring: lime green, RNA recognition motif (RRM); orange, MLLE^{Rrm4} domains; bright blue, MLLE^{Pab1}; dark blue PAM2^{Upa1} motif; light red PAM2L1^{Upa1} motif; dark red PAM2L2^{Upa1} motif; Ankyrin repeats (5xANK), FYVE domain, and RING domain of Upa1 are given in dark grey. GST and SUMO tags are labeled. (B-C) Three-dimensional (3D) AlphaFold2 predicted three-dimensional (3D) structural models of MLLE^{Rrm4}, MLLE^{Pab1} variants in the above mentioned color code (D) Western blot analysis of GST pull-down experiments with components expressed in *E. coli*: N-terminal hexa-histidine-SUMO-tagged PAM2 variants were pulled-down by N-terminal GST fused MLLE variants of Rrm4. The experiment was performed with the soluble fraction of *E. coli* cell lysate to demonstrate specific binding. Results were analyzed with α-His immunoblotting. (E) Western blot analysis of GST pull-down experiments with components expressed in *E. coli*: N-terminal Streptavidin-SUMO-tagged PAM2 variants were pulled-down by N-terminal hexa-histidine fused MLLE variants of Pab1. The experiment was performed with the soluble fraction of *E. coli* cell lysate to demonstrate specific binding. Results were analyzed with α-His immunoblotting.

the G-Rrm4-NT4 version lacking both MLLE1,2 interacted with both HS-PAM2L1 and 2^{Upa1} (Figure 13D, G-Rrm4-NT4, G-Rrm4-NT4-M1,2D, lane 3,4). G-MLLE3^{Rrm4} version lacking helices I, II, and 1 (G-Rrm4-M3-4H), and G-MLLE3^{Rrm4} version lacking helices I and II (G-Rrm4-M3- 5H) did not interact with both HS-PAM2L1 and 2^{Upa1} (Figure 13D, G-Rrm4-M3-4H, G-Rrm4-5H, lane 5,6), which confirmed that MLLE3^{Rrm4} domain with 4 or a 5-helices is not sufficient for the interaction. In comparison, human MLLE^{PABPC1} with five helices or human MLLE^{Ubr5} with four helices is sufficient for interaction (XIE *et al.* 2014). G-MLLE3^{Rrm4} version with all of the 7 predicted helices showed the interaction with both HS-PAM2L1 and 2^{Upa1} (Figure 13D, G-Rrm4-M3-7H, lane 7), which emphasizes that the two newly identified N-terminal helices are essential for the interaction and MLLE3^{Rrm4} could be comprised of 7 helices as observed in AlphaFold2 prediction. In summary, *in vitro* pull-down assays demonstrated that two additional helices (αI, II) of MLLE3^{Rrm4} predicted from the AlphaFold2 are crucial for PAM2L1,2^{Upa1} peptide binding.

3.2.3 The N-terminal alpha helix (α1) of MLLE domain of Pab1 is not essential for interaction with PAM2 motif of Upa1

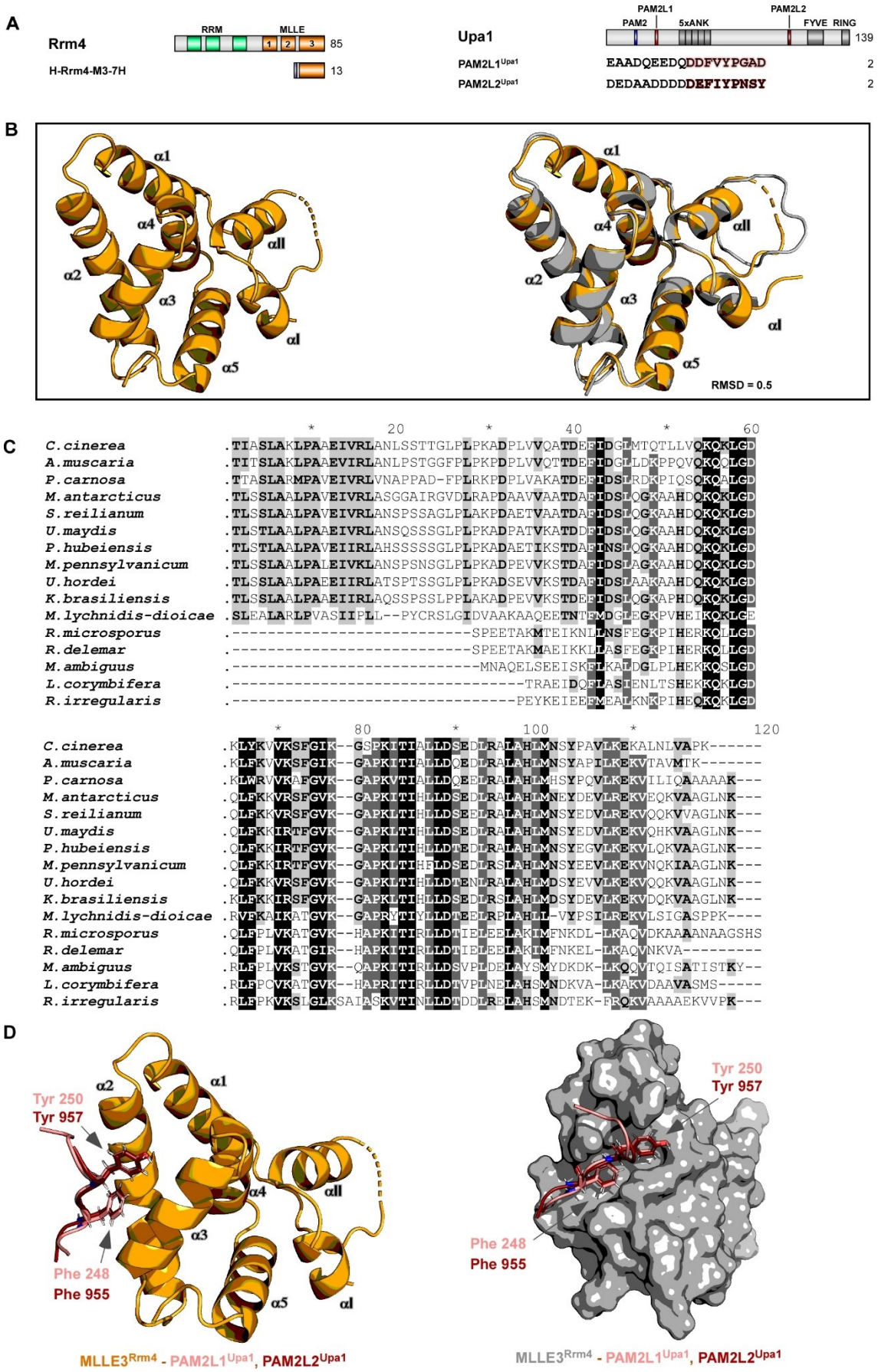
In humans, the N-terminal helix (α1) of MLLE^{PABPC1} does not participate in the PAM2^{Upa1} binding (KOZLOV *et al.* 2010). AlphaFold2 prediction revealed that MLLE^{Pab1} consisted of five alpha helices. *In vitro* binding studies using Ni-NTA beads were performed to verify if the N-terminal helix of MLLE^{Pab1} of *U. maydis* is important for its ligand recognition. To this end, two Pab1-MLLE versions were generated as N-terminal His-tag fusion proteins (Figure 13A, H-Pab1-M-4H, H-Pab1-M-5H) and (Figure 13A, Materials and Methods). AlphaFold2 predicted models of the newly designed constructs depicted

that deletion of α1 did not affect the overall protein folding of MLLE^{Pab1} structure (Figure 13C, MLLE^{Pab1}-4H, MLLE^{Pab1}-5H). 17 amino acid fragments of PAM2 peptide were expressed (Figure 13A, Materials and methods) as N-terminal Strep-SUMO (SS) fusion proteins to check the physical interaction with PAM2^{Upa1}. In pull-down experiments using His tag fusion proteins as bait, H-MLLE^{Pab1} versions carrying either 4 helices or 5 helices showed interaction with SS-PAM2^{Upa1} (Figure 13E, Materials and Methods), which confirmed that the first helix (α1) of the MLLE^{Pab1} is dispensable for interaction *in vitro*. In summary, Pull-down assays have confirmed that the first helix (α1) of MLLE^{Pab} is not essential for PAM2^{Upa1} peptide binding *in vitro,* and this is in alignment with the observation in human MLLE^{PABPC1} (Figure S13,KOZLOV *et al.* 2010).

3.2.4 Crystal structure of MLLE3^{Rrm4}-PAM2L1,2^{Upa1} complexes establish that third MLLE domain of Rrm4 consists of 7 α-helices and non-canonically interacts with its ligands

To obtain the structural insights into the Rrm4 and Upa1 interaction, the MLLE3^{Rrm4} was co-crystallized with its binding partner PAM2L1 or 2 peptides from Upa1 (Figure 14A, H-Rrm4-M3-7H, PAM2L1, PAM2L2 synthetic peptides). To this end, an MLLE3^{Rrm4} version comprising Rrm4 residues 679-791AA was cloned, expressed, and purified to homogeneity (Figure 14A, S12A, B; H-Rrm4-M3-7H carrying an N-terminal hexa-histidine-tag; Materials and methods). Co-crystallization trials were carried out with both MLLE3^{Rrm4}-PAM2L1^{Upa1} and MLLE3^{Rrm4}-PAM2L2^{Upa1} complexes and apoprotein. However, only protein-peptide complexes formed crystals (Figure S12C). Both crystals displayed very low solvent content resulting from tight packing. Diffraction datasets for both complexes were solved by molecular replacement and refined to 2.4 Å for MLLE3^{Rrm4}-PAM2L1^{Upa1} and 1.7 Å for MLLE3^{Rrm4}-PAM2L2^{Upa1} complexes. Data and refinement statistics are given in Table S16. X-ray structure revealed that MLLE3^{Rrm4} polypeptide folds into a stable, compact domain with a 7-helices wound into a right-handed, supercoiled structure (Figure 14B) and confirmed that the AlphaFold2 prediction is accurate with the RMSD of 0.5 Å (Figure 14B).

In general, the MLLE^{PABPC1} domain consists of five helices, the MLLE^{Ubr5} domain consists of 4 helices, whereas the MLLE3^{Rrm4} domain consists of 7 helices. To check if such 7-helix structure is limited to *U. maydis* or is it conserved in other fungi, multiple


A
Rrm4
RRM
MLLE
1
2
3
85
H-Rrm4-M3-7H
13
Upa1
PAM2L1
PAM2L2
PAM2
5xANK
FYVE
RING
139
PAM2L1Upa1
EAADQEEDQDDFVYPGAD
2
PAM2L2Upa1
DEDAADDDDDEFIYPNSY
2
B
α1
α4
αII
α2
α3
αI
α5
RMSD = 0.5
C
C.cinerea
A.muscaria
P.carnosa
M.antarcticus
S.reilianum
U.maydis
P.hubeiensis
M.pennsylvanicum
U.hordei
K.brasiliensis
M.lychnidis-dioicae
R.microsporus
R.delemar
M.ambiguus
L.corymbifera
R.irregularis
D
Tyr 250
Tyr 957
Phe 248
Phe 955
MLLE3Rrm4 - PAM2L1Upa1, PAM2L2Upa1
MLLE3Rrm4 - PAM2L1Upa1, PAM2L2Upa1

3.2.4.1 Figure 14. X-ray structure of MLLE3^{Rrm4}-PAM2L1,2^{Upa1}.

(A) Schematic representation of protein variants (Molecular weight in kilo Dalton indicated) using the following coloring: lime green, RNA recognition motif (RRM); orange, MLLE^{Rrm4} domains; dark blue PAM2^{Upa1} motif; light red PAM2L1^{Upa1} motif; dark red PAM2L2^{Upa1} motif; Ankyrin repeats (5xANK), FYVE domain, and RING domain of Upa1 are given in dark grey. Amino acids bound in the crystal structures of the PAM2L1^{Upa1} sequence are highlighted in light red, PAM2L2^{Upa1} sequence is in dark red. (B) Crystal structure of MLLE3^{Rrm4} (left), comparison of crystal structure and AlphaFold2 predicted structures of MLLE3^{Rrm4} (Right) represented as cartoon models using the following coloring: orange, crystal structure; grey AlphaFold2 predicted. The all-atom RMSD is 0.5 Å, resulting mostly from different rotamers of solvent-exposed sidechains (C) Sequence alignment of MLLE3-like domains from the Rrm4 orthologues in representative organisms of different fungal phyla (Basidiomycetes: *Amanita muscaria, Phanerochaete carnosa, Moesziomyces antarcticus, Sporisorium reilianum, Ustilago maydis, Pseudozyma hubeiensis, Melanopsichium pennsylvanicum, Ustilago hordei, Kalmanozyma brasiliensis, Microbotryum lychnidis-dioicae*, Mucoromycota: *Rhizopus microspores, Rhizopus delemar, Mucor ambiguous, Lichtheimia corymbifera, Rhizophagus irregularis*) showing that the third MLLE domain of Rrm4 is conserved only in the Basidiomycetes. Accession number and sequence coverage are listed in S1 Table. Multiple sequence alignment was performed using ClustalW. (D) Crystal structure of PAM2L1,2^{Upa1}-MLLE3^{Rrm4}. The PAM2L1,2^{Upa1} peptides are inserted into the hydrophobic pocket formed by the helices α2,3 of MLLE3^{Rrm4} models represented as a cartoon in blue (left) and surface in grey (right), Arrows point toward the key residues Phe and Tyr in the PAM2L1,2^{Upa1} peptides.

sequence alignment (MSA) of the MLLE3^{Rrm4} domain from the Rrm4 type proteins were analyzed from the representative organisms in basidiomycetes and other fungi. MSA revealed that the amino acid stretch, corresponding to the helices I, II (α I, II) of MLLE3^{Rrm4}, are well conserved only in the Rrm4 orthologues of basidiomycetes but not found in Rrm4 orthologues of other fungi (Figure 14C). PAM2L^{Upa1} peptides bound to the same face of the MLLE3^{Rrm4} domain comparable to the human PAM2 peptides of GW182 and Paip2 binding to the PABPC1 and Ubr5 MLLE domains (JINEK *et al.* 2010b; XIE *et al.* 2014b).

Overlay of the MLLE3^{Rrm4}-PAM2L1^{Upa1} and MLLE3^{Rrm4}-PAM2L2^{Upa1} complex structures revealed striking similarity of the bound peptide conformations (Figure 14D). In both structures, PAM2L^{Upa1} peptides were bound non-canonically to the hydrophobic pocket formed between the α2 and α3 helices of MLLE3^{Rrm4} by inserting the bulky sidechains of Phe248 and Tyr250 of PAM2L1^{Upa1}, Phe955, and Tyr957 of PAM2L2^{Upa1} (Figure 14D). In both complexes, only a short region consisting of the last 9 residues of the PAM2L1^{Upa1} or 8 residues of PAM2L2^{Upa1} peptide (Figure 14A, PAM2L1^{Upa1}, PAM2L2^{Upa1} highlighted in light red color) was bound in the crystal structure which indicates that only these residues are essential for the interaction and stabilizing the protein complex during crystallization process (Figure 14D).

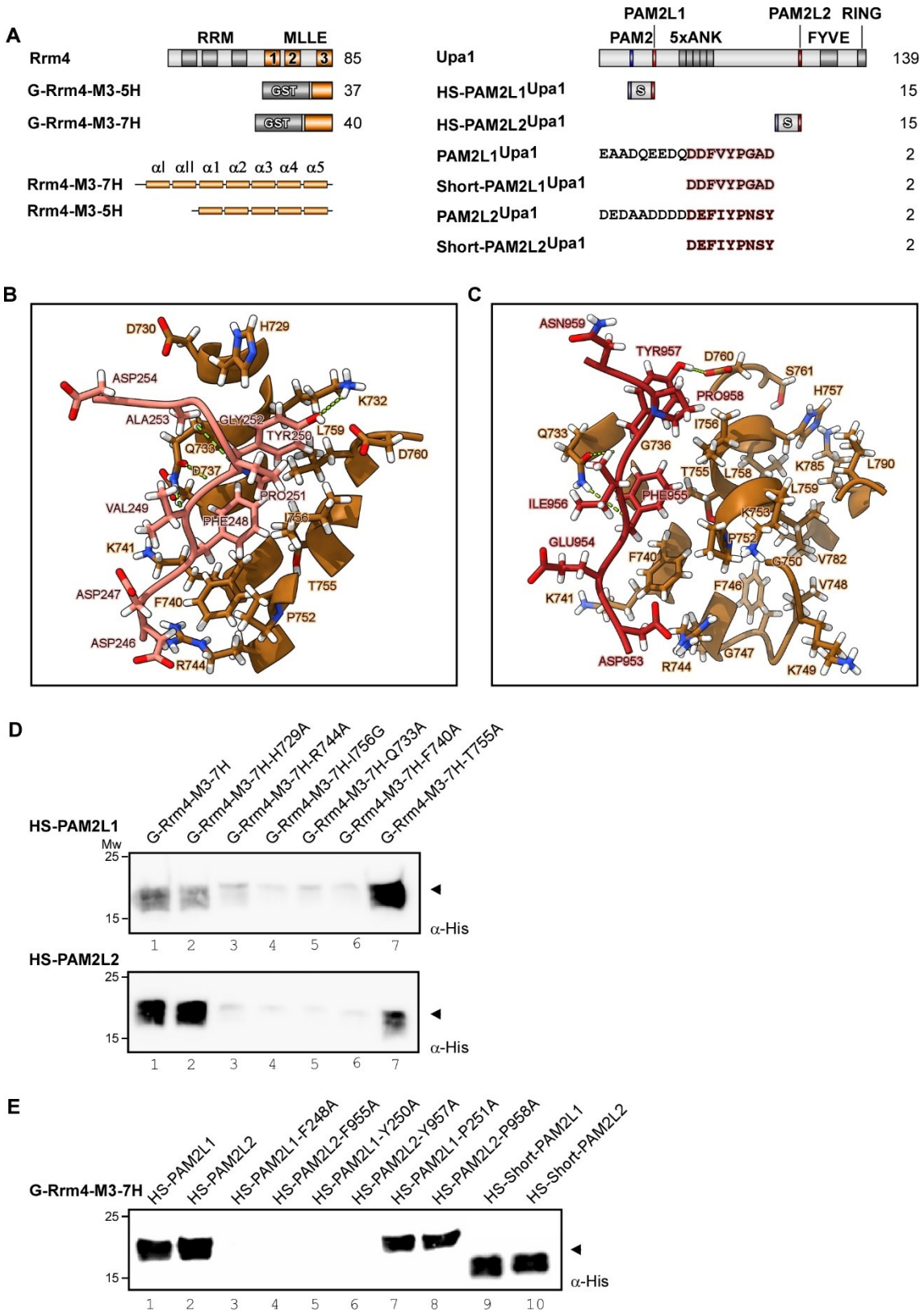
In conclusion, co-crystallization structures confirmed that MLLE3^{Rrm4} consisted of a 7-alpha-helix structure as predicted by AlphaFold2. MSA revealed a 7-helix containing MLLE3^{Rrm4} is conserved in the basidiomycetes, both PAM2L1 and 2^{Upa1} peptides interact

with the MLLE3^{Rrm4} via alpha-helix 2,3 non-canonically, only the last 8 to 9 residues in the C-terminal of PAM2L1 and 2 peptides respectively are bound in the crystal structure.

3.2.5 Point mutation studies identify the key residues in the third MLLE domain of Rrm4 for the interaction with PAM2L1,2 motifs of Upa1

The crystal structures were analyzed in PDBePisa server to understand the key residues in the MLLE3^{Rrm4}-PAM2L1,2^{Upa1} interaction (KRISSINEL AND HENRICK 2007) which listed out the interface residues (Figure 15B, C). Hydrophobic residues (Gly736, Phe740, Pro752, Ile756, and Leu759**)** in the helix α2-α3 make the core of the peptide binding pocket in MLLE3^{Rrm4} (Figure 15B, C). Gln732 of MLLE3^{Rrm4} is the coordinator residue stabilizing the PAM2L^{Upa1}-MLLE3^{Rrm4} interaction by forming hydrogen bonds with the peptide backbones of the two key bulky aromatic residues Phe248 and Tyr250 in the case of PAM2L1^{Upa1} (Figure 15B), Phe955 and Tyr957 in case of PAM2L2^{Upa1} (Figure 15C). In MLLE3^{Rrm4}-PAM2L1^{Upa1} interaction, the positively charged side chain of Lys732 in MLLE3^{Rrm4} makes a polar contact with the hydroxyl group of the Tyr250 of PAM2L1^{Upa1} (Figure 15B), whereas in MLLE3^{Rrm4}-PAM2L2 interaction negatively charged side chain of Asp760 in MLLE3^{Rrm4} makes a polar contact with the hydroxyl group of the Tyr957 of PAM2L2^{Upa1} (Figure 15C).

To further characterize the MLLE3^{Rrm4} and PAM2L^{Upa1} interactions, site-directed point mutations were generated on key residues of MLLE3^{Rrm4}, and *in vitro* binding studies were performed using GST pull-down assays (Figure 15A, Materials and method). In pull-down experiments performed using the G-MLLE3^{Rrm4} version as a bait carrying a point mutation in either of these residues, Gln733, Phe740, Arg744, Ile756 interaction was completely abolished with both HS-PAM2L1,2^{Upa1} versions (Figure 15D lane 3-6) confirming that these are crucial residues for MLLE3^{Rrm4}-PAM2L1,2^{Upa1} interactions. Mutation of His729 in the G-MLLE3^{Rrm4} version slightly inhibited the binding with the HS-PAM2L1 version but not with HS-PAM2L2^{Upa1} (Figure 15D lane2). This observation could be because His729 in MLLE3^{Rrm4} makes a polar contact with Tyr250, Pro251, Gly252, and Ala253 residues of PAM2L1^{Upa1}, which strengthens the interaction; however, such His729 mediated polar contacts are not found in MLLE3^{Rrm4}-PAM2L2^{Upa1}. Therefore, His729 mutation does not affect the MLLE3^{Rrm4}- PAM2L2^{Upa1} interaction (Figure 15 B, C, D lane 2). Mutation in Thr755 of MLLE3^{Rrm4} interacts with HS-PAM2L1,2^{Upa1} similar to the wild-type protein, indicating that it is not participating in the



3.2.5.1 Figure 15. Key residues in the MLLE3^{Rrm4}-PAM2L1 and 2^{Upa1} interaction.

(A) Schematic representation of protein variants (Molecular weight in kilo Dalton indicated) using the following coloring; orange, MLLE^{Rrm4} domains; light red PAM2L1^{Upa1} motif; dark red PAM2L2^{Upa1} motif; RRM domain of Rrm4, Ankyrin repeats (5xANK), FYVE domain, and RING domain of Upa1 are given in dark grey. GST and SUMO tags are labeled. Amino acids bound in the crystal structures of the PAM2L1^{Upa1} sequence are highlighted in light red, PAM2L2^{Upa1} sequence is in dark red. (B) A view of the interface

between the PAM2L1^{Upa1} peptide (light red; interacting side chains are shown in stick format) and MLLE3^{Rrm4} domain (Brown; interacting residues in stick format). Dashed lines indicate hydrogen bonding interactions (C) A view of the interface between the PAM2L2^{Upa1} peptide (dark red; interacting side chains are shown in stick format) and MLLE3^{Rrm4} domain (Brown; interacting residues in stick format). Dashed lines indicate hydrogen bonding interactions (D) Western blot analysis of GST pull-down experiments with components expressed in *E. coli*: N-terminal Hexa-Histidine-SUMO(HS)-tagged PAM2L1,2^{Upa1} versions were pulled-down by N-terminal GST fused MLLE3-7H point mutation variants of Rrm4. The experiment was performed with the soluble fraction of *E. coli* cell lysate to demonstrate specific binding. Results were analyzed with α-His immunoblotting. (E) Western blot analysis of GST pull-down experiments performed as above; HS-PAM2L1,2^{Upa1}-point mutation variants were pulled-down by N-terminal GST fused MLLE3-7H of Rrm4. Results were analyzed with α-His immunoblotting.

peptide binding, as the sidechain of Thr755 is facing outwards from the peptide binding pocket.

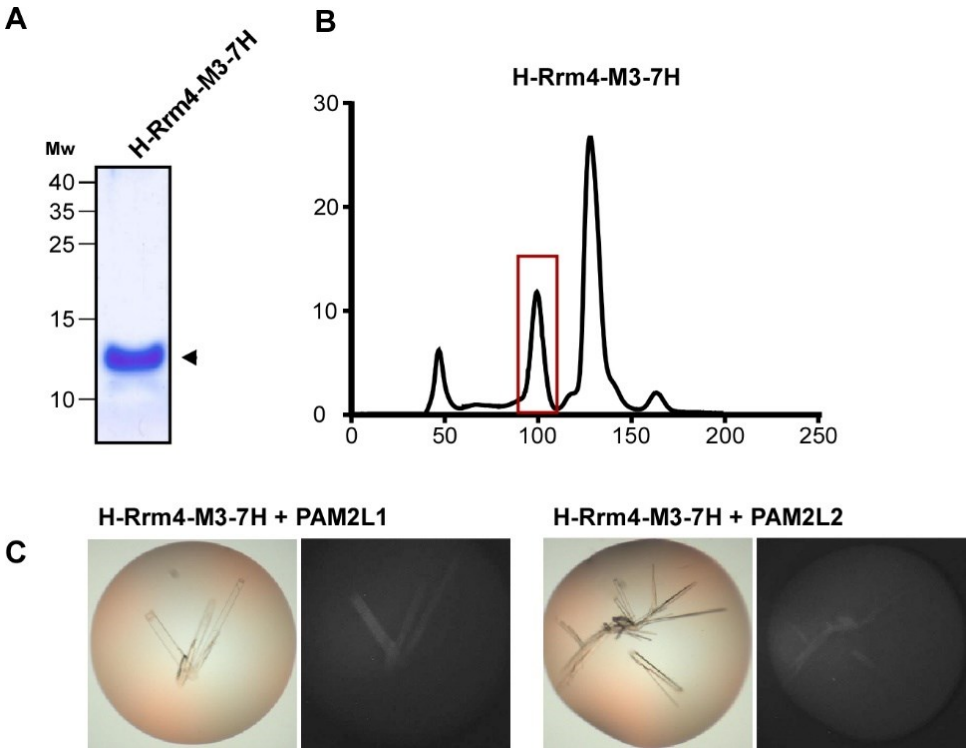
Site-directed point mutations were generated on the key residues of both HS-PAM2L1,2^{Upa1} versions, and *in vitro* binding studies were performed using GST pull-down assays to identify the contribution of the key residues in PAM2L1,2^{Upa1} peptides (Figure 15A, Materials and method). In pull-down experiments performed using G-MLLE3^{Rrm4} as a bait, HS-PAM2L1^{Upa1} versions carrying a point mutation in either one of the bulky residues Phe248, Tyr250 interaction was completely lost (Figure 15E lane 3,5). Similarly, HS-PAM2L2^{Upa1} versions carrying a point mutation in either one of the bulky residues Phe955 or Tyr957 completely prevented the binding with MLLE3^{Rrm4} (Figure 15E lane 4,6), indicating that these residues are crucial for PAM2L^{Upa1}-MLLE3^{Rrm4} interaction. This observation is in alignment with the interactions observed in the crystal structure. Mutation in Pro251^{PAM2L1} or Pro958^{PAM2L2} did not inhibit the interaction. This could be because, as observed in the crystal structure, Proline residues of both PAM2L1,2^{Upa1} are exposed outside the peptide binding pocket and do not make a major contribution to PAM2L-MLLE3^{Rrm4} interaction.

To test whether the short version of the PAM2L1,2^{Upa1} peptides, as seen in the crystal structures (Figure 14D, 15 B, C, highlighted in light red), is sufficient for binding, N-terminally truncated PAM2L^{Upa1} versions comprising only 9 amino acids in the C-termini of the PAM2L1 or 2^{Upa1} were generated and *in vitro* binding studies were performed using GST pull-down assays (Figure 15A, B, C). In pull-down experiments performed using G-MLLE3^{Rrm4} as a bait, shorter versions of both HS-PAM2L1 and 2^{Upa1} showed interaction bands similar to longer versions of the HS-PAM2L1,2^{Upa1} versions, confirming that these regions are sufficient for binding (Figure 15E lane 9,10).

In summary, pull-down results using point mutations in G-MLLE3^{Rrm4} versions confirmed that Arg744, Ile756, Q733, and Phe740 are key residues in MLLE3^{Rrm4} for peptide

recognition. Using a similar strategy and point mutations in HS-PAM2L1,2^{Upa1} versions confirmed that Phe248 and Tyr250 are the key residues in PAM2L1^{Upa1}, as well as Phe955 and Tyr957 are the key residues in PAM2L2^{Upa1}. In addition, pull-down results using a truncated version of the PAM2L1 or 2^{Upa1} versions confirmed that the shorter version of the peptides found in the co-crystallized structure (Figure 15A) is sufficient for the MLLE3^{Rrm4}-PAM2L1,2^{Upa1} interactions.

3.2.6 Supporting information



3.2.6.1 Figure S12. Purified MLLE3 domain of Rrm4 co-crystallized with PAM2L1 and 2 peptides of Upa1.

(A) SDS-PAGE analysis of purified H-Rrm4-M3-7H, used in X-ray crystallography. (B) Size exclusion chromatography analysis of H-Rrm4- M3-7H. The graph shows the elution profile. The red box indicates the monomeric fraction of the purified protein. (C) Light microscopy images depicting co-crystalized MLLE3^{Rrm4}- PAM2L1^{Upa1} crystals left, MLLE3^{Rrm4}- PAM2L2^{Upa1} crystals right. Light microscopy images taken under UV light (dark images) indicate that these are protein crystals and not salt crystals.

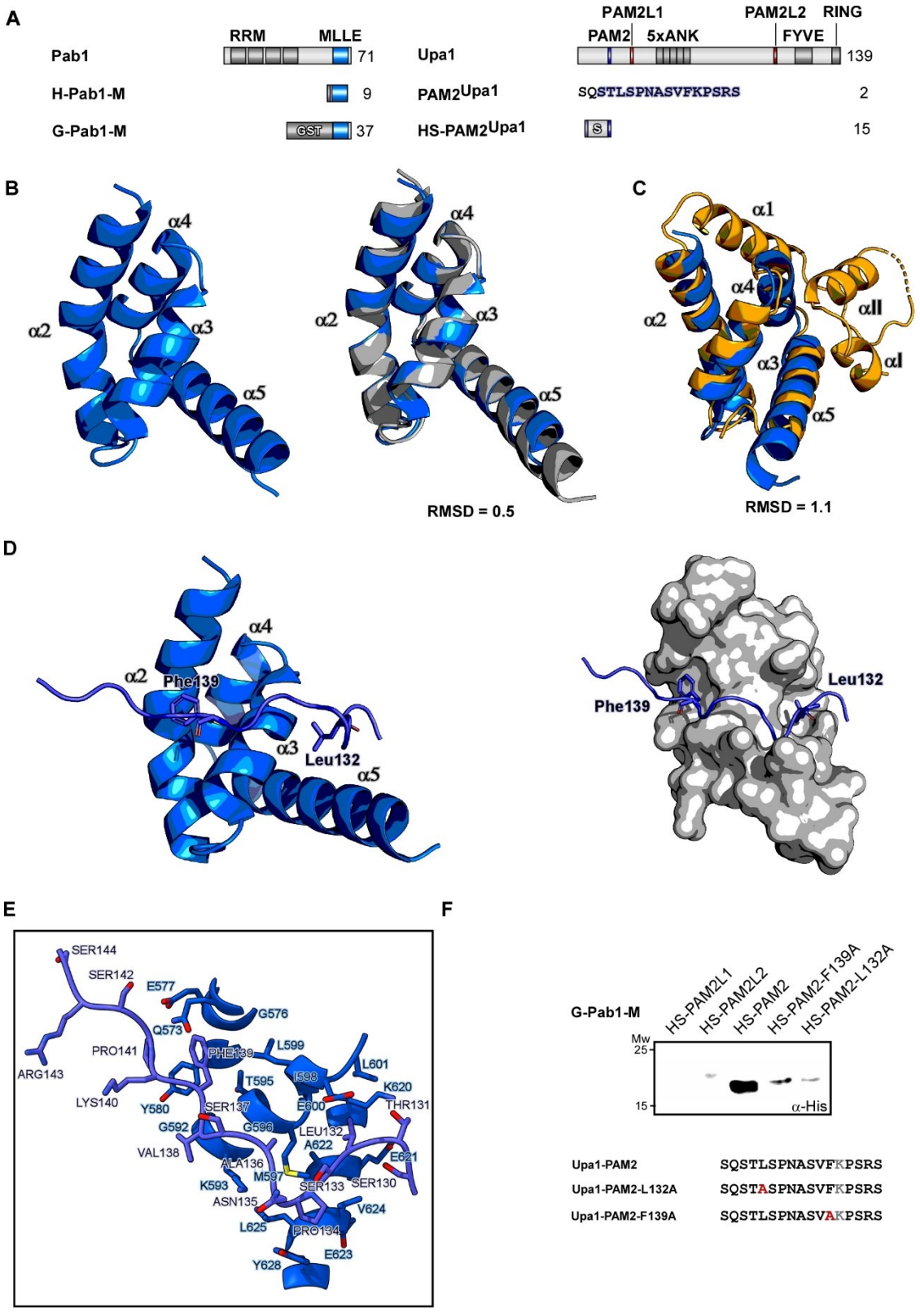
3.3 Structural characterization of MLLE^{Pab1} - PAM2^{Upa1} complex

3.3.1 Crystal structure of MLLE^{Pab1}-PAM2^{Upa1} complex established that MLLE domain of Pab1 consists of 5 α-helices and canonically interacts with its ligand

To obtain the structural insights into the Upa1 and Pab1 interaction, the MLLE^{Pab1} with PAM2^{Upa1} peptide was co-crystallized. To this end, the MLLE^{Pab1} version comprising Pab1 residues 567-636 AA was cloned, expressed, and purified to homogeneity (Figure 16A, S13A, B; H-Pab1-M-4H, carrying an N-terminal hexa-histidine-tag; Materials and methods). Co-crystallization trials were carried out with H-MLLE^{Pab1}-PAM2^{Upa1} complex as well as apoprotein. However, only H-MLLE^{Pab1}-PAM2^{Upa1} complexes formed crystals, highlighting that the PAM2^{Upa1} peptide stabilizes the MLLE^{Pab1} protein and forms a stable complex that, in turn, favors the crystallization condition (Figure S13C). Obtained crystals displayed very low solvent content resulting from tight packing. Diffraction datasets for MLLE^{Pab1}-PAM2^{Upa1} complex were solved by molecular replacement and refined to 1.8 Å. Data and refinement statistics are given in Table S16.

X-ray structure revealed that MLLE^{Pab1} polypeptide folds into a stable, compact domain with 4 helices wound into a right-handed supercoil structure (Figure 16B) and confirmed that the AlphaFold2 prediction is accurate with the RMSD of 0.5 Å (Figure 16B). PAM2^{Upa1} peptide bound to the same face of the MLLE^{Pab1} domain comparable to the human PABPC1 and Ubr5, MLLE domains that bind to the PAM2 peptides of GW182 and Paip2 respectively (JINEK *et al.* 2010; XIE *et al.* 2014). The peptide binds by wrapping around the highly conserved KLTGMILE signature motif of MLLE and interacting with the hydrophobic pockets between helices α2 and α3 and between helices α3 and α5 of MLLE^{Pab1}.

In alignment with the conventional PAM2-MLLE interactions, PAM2^{Upa1} peptide canonically interacted with the α2 and α5 helices of MLLE^{Pab1}, similar to the MLLE^{PABPC}-PAM2^{PAIP2} complex (Figure S13D). In MLLE^{Pab1}-PAM2^{Upa1} complexes, only a short region consisting of the last 15 residues in the C-terminal of the PAM2^{Upa1} (Figure16A, D - PAM2 highlighted in light blue color) is bound in the crystal structure, which indicates that only these residues are essential for the interaction and stabilizing the protein complex during the crystallization process.



3.3.1.1 Figure 16. X-ray structure of MLLE^{Pab1}-PAM2^{Upa1} complex.

(A) Schematic representation of protein variants (Molecular weight in kilo Dalton indicated) using the following coloring: bright blue, MLLE^{Pab1}; dark blue PAM2^{Upa1} motif; light red PAM2L1^{Upa1} motif; dark red PAM2L2^{Upa1} motif; RRM domain of Rrm4, Ankyrin repeats (5xANK), FYVE domain, and RING domain of Upa1 are shaded in dark grey. GST and SUMO tags are labeled. Amino acids bound in the crystal struc-

tures of the PAM2 sequence are highlighted in light blue. (B) Crystal structure of MLLE^{Pab1} (left), comparison of crystal structure and AlphaFold2 predicted structures of MLLE^{Pab1} (Right) represented as cartoon models using the following coloring: blue, crystal structure; grey AlphaFold2 predicted. The all-atom RMSD is 0.5 Å, resulting mostly from different rotamers of solvent-exposed sidechains. (C) Comparison of MLLE^{Pab1} and MLLE3^{Rrm4} crystal structures, models depicted in the following coloring: blue, MLLE^{Pab1}; orange, MLLE3^{Rrm4}. The all-atom RMSD is 1.1Å, revealing the differences between the two MLLE variants. (D) Crystal structure of PAM2^{Upa1}-MLLE^{Pab1}. The PAM2^{Upa1} peptide wraps around the MLLE^{Pab1}, and the key residues Leucine (Leu132) and Phenylalanine (Phe139) are inserted into the hydrophobic pocket formed in between the helices α3,5 and in between the helices α2,3 of MLLE^{Pab1}. Models are represented as a blue cartoon (left) and a grey surface (right). (E) A view of the interface between the PAM2^{Upa1} peptide (dark blue; interacting side chains in stick format) and MLLE^{Pab1} domain (blue cartoon; interacting residues in stick format). Dashed lines indicate hydrogen bonding interactions (F) Western blot analysis of GST pull-down experiments with components expressed in *E. coli*: N-terminal Hexa-Histidine-SUMO-tagged PAM2 mutant versions were pulled-down by N-terminal GST fused MLLE variants of Pab1. The experiment was performed with the soluble fraction of *E. coli* cell lysate to demonstrate specific binding. Results were analyzed with α-His immunoblotting.

In conclusion, co-crystallization structures confirmed that the H-MLLE^{Pab1} version used for crystallization consisted of 4 alpha-helix structures as predicted by AlphaFold2, PAM2^{Upa1} peptide interacts with the MLLE^{Pab1} via alpha-helix 2,5 canonically, and only the last 15 residues in the C-terminal of PAM2^{Upa1} peptide is bound in the crystal structure.

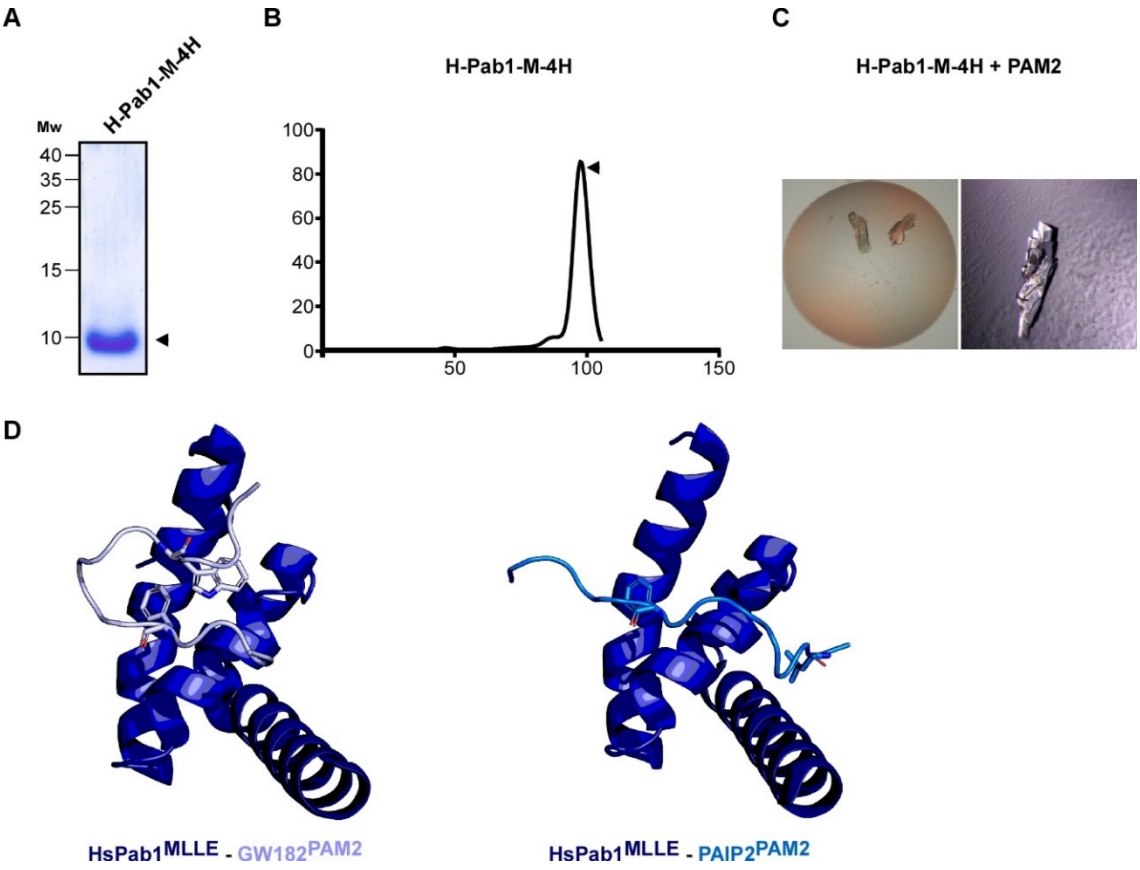
3.3.2 Point mutation studies identify the key residues in the MLLE domain of Pab1 for interaction with PAM2 motif of Upa1

To further understand the key residues in the MLLE^{Pab1}-PAM2^{Upa1} interaction, the crystal structures were analyzed in the PDBePisa server (Krissinel and Henrick 2007), which listed out the interface residues (Figure 16E). Typically PAM2-MLLE interactions are mediated by hydrophobic interactions (XIE *et al.* 2014). In alignment with this, hydrophobic residues (Gly576, Tyr580, Gly592, Gly596, Met597, Ile598, Leu599, Leu600, Glu601 and Leu602) in the α2, α3, α5 of MLLE^{Pab1} make the core of the peptide binding pocket in MLLE^{Pab1} (Figure 16E). The two hydrophobic pockets between the helices α2-α3 and between helices α3-α5 of MLLE^{Pab1} are occupied by the Leucine (Leu132) and Phenylalanine (Phe139) residues of the PAM2^{Upa1} peptide (Figure 16A, D, E).

To further characterize the MLLE^{Pab1} and PAM2^{Upa1} interaction, site-directed point mutations were inserted into key residues in HS-PAM2^{Upa1} versions, and *in vitro* binding studies were performed using GST pull-down assays (Figure 16A, Materials and method). In pull-down experiments performed using the G-Pab1-M version as bait, the interaction was completely abolished with HS-PAM2^{Upa1} carrying a point mutation in either Leucine (Leu132^{Upa1}) or Phenylalanine (Phe139^{Upa1}) (Figure 16F lane 2,3), which confirmed that

these two are crucial residues in PAM2^{Upa1} for MLLE^{Pab1} binding. Mutation in Pro 141^{Upa1} slightly reduced the interaction but did not completely abolish the interaction, which suggested that Proline is not essential for peptide recognition (Figure 16F lane 2,3). In summary, the Pull-down results confirmed that Leu132 and Phe139 are the crucial residues in PAM2^{Upa1} for interaction with MLLE^{Pab1}.

3.3.3 Supporting information



3.3.3.1 Figure S13. Purified MLLE domain of Pab1, co-crystallized with PAM2 peptide of Upa1.

(A) SDS-PAGE analysis of purified H-Pab1-M-4H, used in X-ray crystallography. (B) Size exclusion chromatography analysis of H-Pab1-M-4H. The graph shows the elution profile. The arrowhead indicates the monomeric fraction of the purified protein. (C) Light microscopy images depicting co-crystalized MLLE^{Pab1}- PAM2^{Upa1} crystals *left*, magnified image *right*. (D) Crystal structure of human MLLE^{PABPC1}-PAM2^{GW182} *(left)* (PDB ID: 3KTP). Key residues (Phe1389, Trp1395) of the PAM2^{GW182} peptide are inserted into the hydrophobic pocket formed in between the helices α2,3 of MLLE^{PABPC1} and an example of non-canonical MLLE-PAM2 interaction. Crystal structure of human MLLE^{PABPC1}-PAM2 ^{PAIP2} *(right)* (PDB ID: 3KUS). Key residues (Leu111, Phe118) of the PAM2^{PAIP2} peptide are inserted into the hydrophobic pocket formed between the helices α3,5 of MLLE^{PABPC1} and an example for canonical MLLE-PAM2 interaction.

4 Identification of novel interaction partners of MLLE domains in *U. maydis*

4.1 Introduction

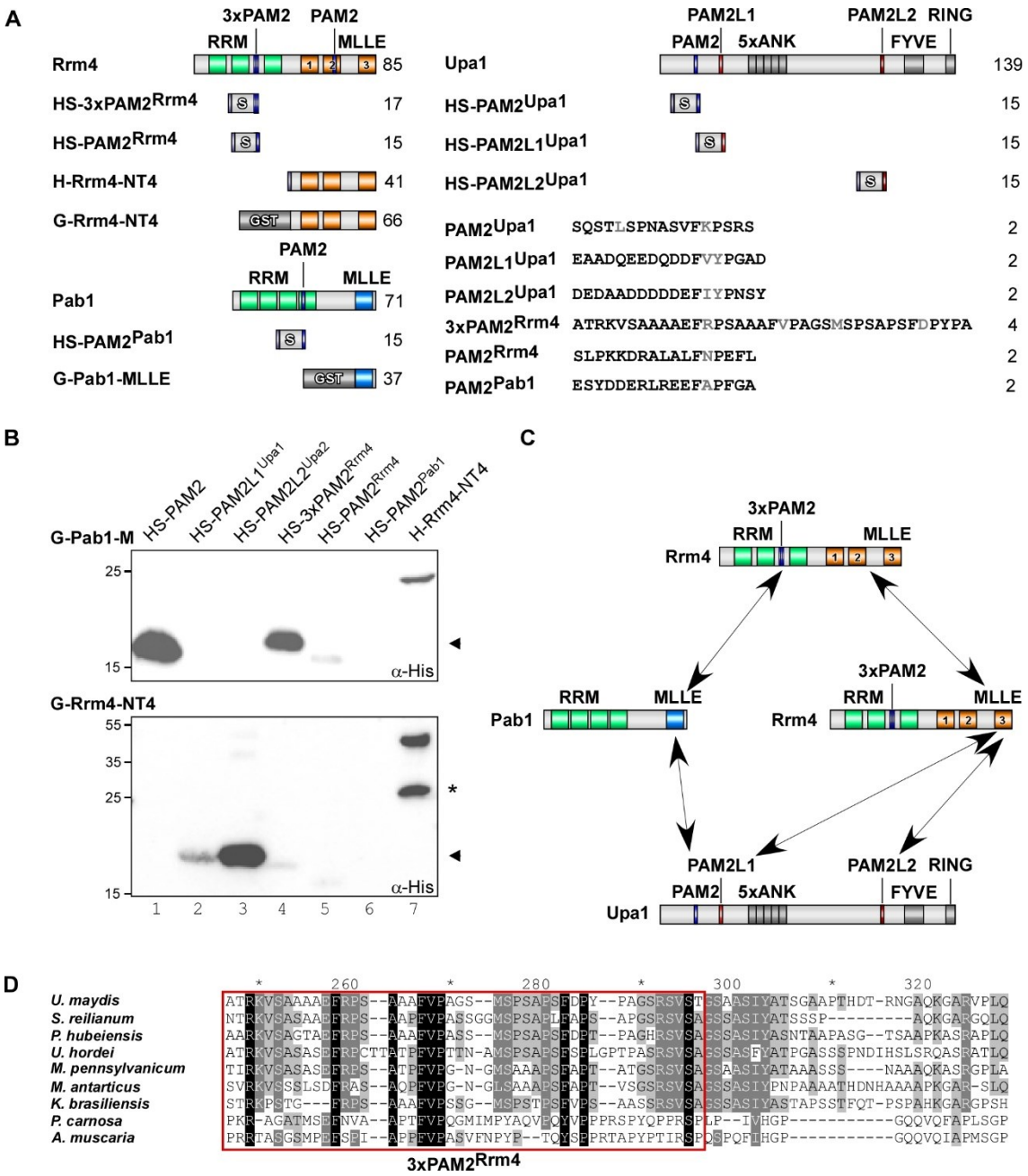
Structural studies described in the previous chapter deciphered the molecular basis of the peptide recognition by the MLLE3^{Rrm4} and MLLE^{Pab1} domains. Typically, MLLE domains in humans recognized the PAM2 peptides by hydrophobic interactions with the conserved Leu (n3) and Phe (n10) residues. In agreement with this, MLLE^{Pab1} recognized the Leu (n3) and Phe (n10) residues of the PAM2Upa1 peptide by the canonical hydrophobic residues. Nevertheless, MLLE3^{Rrm4} recognized the PAM2L1,2^{Upa1} peptides by the non-canonical interactions with the FxY residues in the C-terminus of the peptide. Truncated versions of the peptides containing only the C-terminus of PAM2L were sufficient for interaction (short PAM2L1,2^{Upa1}). A preliminary MLLE3^{Rrm4} recognition sequence was formulated based on this observation. Novel PAM2L and PAM2-containing proteins were identified from the *U. maydis* genome using reiterative BLAST search.

In addition, a few PAM2 and PAM2L sequences were identified by visual inspection of the Rrm4-associated protein sequences. For example, Rrm4 has an internal, overlapping PAM2 (3x PAM2Rrm4) sequence between the RRM2 and RRM3 domains. 3x PAM2^{Rrm4} interacted with MLLE^{Pab1} *in vitro*. EfTu was discovered as a novel PAM2L-containing interaction partner of MLLE3^{Rrm4} by Mass spectrometry analysis of the co-purified proteins in the *in vitro* pull-down assays. Thus, a set of novel interaction partners of MLLE domains of Rrm4 and Pab1 were identified.

4.2 Inter and intra molecular interactions of Rrm4 and Pab1 via MLLE domains

To get insights into the molecular interactions of the Rrm4 and Pab1 through their MLLE domains, Rrm4 and Pab1 protein sequences were manually analyzed for internal PAM2 and PAM2L sequences. Visual inspection and multiple sequence analysis revealed the presence of a potential tandem 3x $PAM2^{Rrm4}$ sequences in the linker region between the $RRM2^{Rrm4}$ and $RRM3^{Rrm4}$ domains, a potential PAM2 motif within the $MLLE2^{Rrm4}$, and a potential PAM2 motif in the $RRM4^{Pab1}$ (Figure 17A) To evaluate the interaction capacity of the *de novo* predicted internal PAM2 sequences in Rrm4 and Pab1, *in vitro* binding studies, were performed using GST pull-down assay as described earlier. To this end, *de novo* predicted internal PAM2 sequences in Rrm4 and Pab1 were expressed as N-terminal hexa-histidine-SUMO (HS) fusion proteins (Figure. 17A, Materials and methods). As bait proteins in the GST pull-down assay, N-terminal truncated version of Rrm4 and Pab1 were expressed as a fusion protein with an N-terminal GST tag (Figure. 17 A, G-Rrm4-NT4, G-Pab1-M). $MLLE3^{Rrm4}$ specifically interacted with PAM2L1 motif of Upa1 but not with PAM2 motif of Upa1 whereas $MLLE^{Pab1}$ specifically interacts with PAM2 motif of Upa1 but not with PAM2L1,2 motifs of Upa1, (POHLMANN *et al.* 2015; DEVAN *et al.* 2022). Therefore, HS-$PAM2^{Upa1}$ and $PAM2L1,2^{Upa1}$ were used as positive and negative controls for interaction studies with G-Pab1-M, whereas $PAM2L1,2^{Upa1}$ and HS-$PAM2^{Upa1}$ were used as positive and negative controls respectively for interaction studies with G- Rrm4-NT4.

In GST pull-down experiments using GST fusion proteins as bait, G-Pab1-M interacted with the PAM2 motif of Upa1 (positive control) but not with PAM2L1 and 2 motifs of Upa1 (negative control) (Figure. 17B, G-Pab1-M, lane 1-3). Conversely, G-Rrm4-NT4 interacted with PAM2L1 and 2 motifs of Upa1 (positive control) but not with PAM2 motif Upa1 (negative control) (Figure. 17B, G-Rrm4-NT4, lane 1-3). Interestingly, G-Pab1-M showed a strong binding signal for the internal tandem 3x PAM2 motif of Rrm4, whereas G-Rrm4-NT4 showed a very weak binding signal for the same, indicating that $MLLE^{Pab1}$ interacted with Rrm4 via 3x PAM2 motif of Rrm4 *in vitro* (Figure 17B, Lane 4, HS-3x PAM2Rm4). G-Pab1 and G-Rrm4-NT4 showed a weak binding with the PAM2 motif of Rrm4 (Figure 17B, Lane 5, HS-PAM2Rm4). Likewise, both G-Pab1-M and G-Rrm4-NT4 did not bind with the PAM2 motif of Pab1, indicating that both the PAM2



4.2.1.1 Figure 17. MLLE mediated inter and intra molecular interactions of Rrm4 and Pab1.

(A) Schematic representation of protein variants (Molecular weight in kilo Dalton indicated) using the following coloring: lime green, RNA recognition motif (RRM); orange, $MLLE^{Rrm4}$ domains; dark blue, $MLLE^{Pab1}$; light blue, $PAM2^{Upa1}$; light orange, $PAM2L^{Upa1}$ sequence (PL1 – 2); Ankyrin repeats (5xANK), FYVE domain, and RING domain of Upa1 are shaded in dark grey. GST and SUMO tags are labeled. Variant amino acids of the FxP and FxxP of $PAM2^{Upa1}$ and $PAM2L^{Upa1}$ sequences are printed in grey font. (B) Western blot analysis of GST pull-down experiments with components expressed in *E. coli*: N-terminal Hexa-Histidine-SUMO-tagged PAM2 variants were pulled-down by N-terminal GST fused MLLE variants of Rrm4 and Pab1. The experiment was performed with the soluble fraction of *E. coli* cell lysate to demonstrate specific binding. Results were analyzed with α-His immunoblotting. (C) Schematic representation of molecular interactions between the Pab1, Rrm4, and Upa1, identified from the *in vitro* binding studies. Arrows indicate the interacting regions. Domains are colored as described above. (D) Sequence alignment of potential tandem 3x PAM2 motif of Rrm4 orthologues in representative organisms (*Ustilago*

maydis, Sporisorium reilianum, Pseudozyma hubeiensis, Ustilago hordei, Melanopsichium pennsylvanicum, Moesziomyces antarcticus, Kalmanozyma brasiliensis, Phanerochaete carnosa, and *Amanita muscaria*)

motif of Rrm4, PAM2 motif of Pab1 might not be a real PAM2 motif (Figure 7B, Lane 6, HS-PAM2^{Pab1}). Surprisingly, both G-Pab1-MLLE and G-Rrm4-NT4 interacted with the C-terminal half of Rrm4, containing three tandem MLLE^{Rrm4} domains (Figure 18B, Lane 7, H-Rrm4-NT4), revealing that the MLLE domain of Pab1 and MLLE1-3 domains of Rrm4 are capable of interacting with each other, MLLE1-3^{Rrm4} is capable of interacting with its kind to form higher order oligomers.

In summary, *in vitro* binding studies have confirmed that the MLLE domain of Pab1 specifically recognizes the *de novo* predicted tandem 3x PAM2 motif present in between the RRM2-3 of Rrm4 and C-terminal region of Pab1 as well as the C-terminal region of Rrm4 interacts with the Rrm4-NT4 (C-terminal region of Rrm4) (Figure. 17C). Collectively these *in vitro* results pointing out that mRNA binding proteins Rrm4 and Pab1 are capable of physically interacting with each other through their MLLE domains. The C-terminal half of Rrm4 might be involved in homo oligomerization.

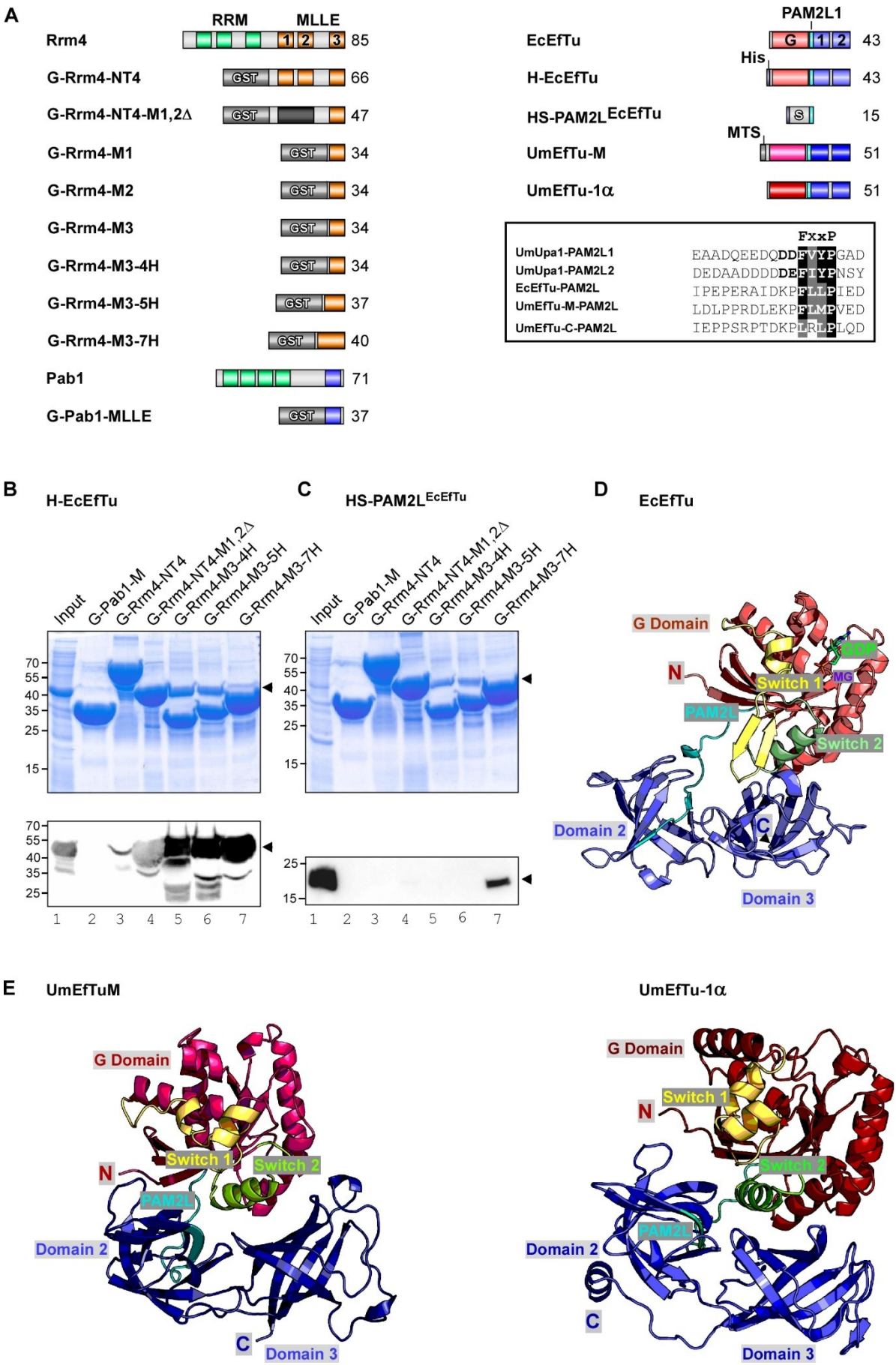
4.3 *E. coli*-EfTu interacts with Rrm4-MLLE3

To assess the ligand recognition capacity of each MLLE domain of Rrm4, *in vitro* binding studies were carried out. To this end, all three MLLE domains (MLLE1^{Rrm4}, MLLE2^{Rrm4}, MLLE3^{Rrm4}) of Rrm4 were individually expressed as a fusion protein with an N-terminal GST tag in *E. coli* (Figure 18A; G-Rrm4-M1, G-Rrm4-M2, G-Rrm4-M3; Materials and methods). As a control, N-terminal GST fused, MLLE domain of Pab1 was expressed (Figure 18A; G-Pab1-M; Materials and methods). 18 amino acid fragments were expressed (Figure 18A) as N-terminal hexa-histidine-SUMO (HS) fusion proteins (Figure 18A; HS-PAM2^{Upa1}, HS-PAM2L1^{Upa1}, HS-PAM2L2^{Upa1}, Materials, and methods) to check the physical interaction of PAM2 and PAM2L sequences of Upa1 with MLLE domain variants of Rrm4 and Pab1.

In GST pull-down experiments using GST fusion proteins as bait, G-Pab1-MLLE interacted with PAM2 but not with the PAM2L motifs of Upa1 (G-Pab1-MLLE; Figure S14A lane 1-3). Conversely, G-Rrm4-NT4 recognized the two HS-PAM2L motifs of Upa1 but not the HS-PAM2 motif (G-Rrm4-NT4, Figure S14A lane 1-3), which is in alignment with the earlier observations (Figure.S3B). Interestingly, in experiments where the GST-Rrm4-M3 variant was used as a bait, an unknown protein band was enriched

corresponding to the molecular weight of 40,000 Daltons (Figure S14A, HS-PAM2L1^{Upa1}, HS-PAM2L2^{Upa1}, HS-PAM2^{Upa1}, lane 4). This protein size was too big compared to the prey protein HS-PAM2^{Upa1} variants and was found in all three pull-down experiments irrespective of the prey protein variant used. Surprisingly this band was not detected in western blotting with anti-his antibodies. This result indicated that an unknown *E. coli* protein could interact with the MLLE3^{Rrm4} and co-eluted in pull-down experiments. The gel fragment containing the enriched protein band was excised and analyzed by liquid chromatography coupled with mass spectrometric analysis (LC-MS/MS) to identify the unknown *E. coli* protein. The elongation factor EfTu (*tufB*) from *E. coli* with a molecular weight of 43,000 Daltons was the most abundant candidate among the 50 enriched candidates identified by LC-MS/MS (Figure S14B, Table S17). Visual inspection of the amino acid sequence of the *E. coli* EfTu (*Ec*EfTu) revealed a potential PAM2L motif (IPEPERAIDKP**F**L**LP**IED), consisting of a conserved hydrophobic FxxP pocket which could serve as a potential interaction motif of MLLE3^{Rrm4}.

To verify the molecular interactions of *E. coli* EfTu with the MLLE3^{Rrm4} versions, *in vitro* binding studies were performed using GST pull-down assay with three different G-MLLE3^{Rrm4} versions (Figure 18A, G-Rrm4-M3-4H, G-Rrm4-M3-5H, G-Rrm4-M3-7H, Materials, and Methods) at different length of MLLE3^{Rrm4} comprising the 4, 5 or 7 helices. G-MLLE3^{Rrm4} versions were generated as a fusion protein with an N-terminal GST tag. G-Pab1-MLLE, G-Rrm4-NT4, and G-Rrm4-NT4-M1,2D were used as control (Figure 18A; Materials and methods). To verify the physical interaction, *E. coli* EfTu version with an N-terminal Histidine tag (H-EcEfTu), the potential PAM2L^{EfTu} motif with an N-terminal Sumo protein tag (HS-PAM2L^{EcEfTu}) were cloned and expressed as fusion proteins (Fig18A, Materials, and Method). In GST pull-down experiments using GST fusion proteins as bait, G-Pab1-M did not recognize both the H-EcEfTu and HS-PAM2L motif of *E. coli* EfTu (Figure 18B, lane 2). However, all the versions containing Rrm4-MLLE3 domain showed strong interaction with H-EcEfTu (Figure 18B, lane 3-7). This observation confirmed that *E. coli* EfTu specifically interacted with MLLE3^{Rrm4} *in vitro* but did not interact with the MLLE^{Pab1}. In contrast, the HS-PAM2L^{EfTu} was only binding to G-Rrm4-MLLE3-7H (Figure 18C, lane 7) but not to any other G-Rrm4-MLLE3 variants, albeit they have intact MLLE3^{Rrm4} (Figure 18C, lane 2-6), which raises more questions and is a subject for further investigation.


A
RRM
MLLE
Rrm4
1 2 3
85
G-Rrm4-NT4
GST
66
G-Rrm4-NT4-M1,2Δ
47
G-Rrm4-M1
34
G-Rrm4-M2
34
G-Rrm4-M3
34
G-Rrm4-M3-4H
34
G-Rrm4-M3-5H
37
G-Rrm4-M3-7H
40
Pab1
71
G-Pab1-MLLE
37
PAM2L1
EcEfTu
G 1 2
43
His
H-EcEfTu
43
HS-PAM2L EcEfTu
S
15
MTS
UmEfTu-M
51
UmEfTu-1α
51
FxxP
UmUpa1-PAM2L1 EAADQEEDQDDFVYPGAD
UmUpa1-PAM2L2 DEDAADDDDDEFTYPNSY
EcEfTu-PAM2L IPEPERAIDKPFLLPIED
UmEfTu-M-PAM2L LDLPPRDLEKPFLMPVED
UmEfTu-C-PAM2L IEPPSRPTDKPLRLPLQD
B
H-EcEfTu
Input
G-Pab1-M
G-Rrm4-NT4
G-Rrm4-NT4-M1,2Δ
G-Rrm4-M3-4H
G-Rrm4-M3-5H
G-Rrm4-M3-7H
70
55
40
35
25
15
1 2 3 4 5 6 7
C
HS-PAM2L EcEfTu
D
EcEfTu
G Domain
GDP
N
Switch 1
MG
PAM2L
Switch 2
Domain 2
C
Domain 3
E
UmEfTuM
G Domain
Switch 1
Switch 2
N
PAM2L
Domain 2
C
Domain 3
UmEfTu-1α
G Domain
Switch 1
N
Switch 2
PAM2L
Domain 2
C
Domain 3

4.3.1.1 Figure 18. *E. coli*-EfTu interacts with MLLE3^{Rrm4}.

(A) Schematic representation of protein variants (Molecular weight in kilo Dalton indicated) using the following coloring: lime green, RNA recognition motif (RRM); orange, MLLE^{Rrm4} domains; dark blue, MLLE^{Pab1}; light blue PAM2^{Upa1}; light orange PAM2L^{Upa1} sequence (PL1 – 2). Ankyrin repeats (5xANK), FYVE domain, and RING domain of Upa1 are given in dark grey. GST and SUMO tags are labeled. GDP binding domain (G) domain of EfTu (Red variants), Domain II and III of EfTu (Blue variants), and Variant amino acids of the FxP and FxxP of PAM2^{Upa1} and PAM2L^{Upa1} sequences are shaded in grey. (B) SDS-PAGE and Western blot analysis of GST pull-down experiments with components expressed in *E. coli*: N-terminal GST fused MLLE variants of Rrm4 and Pab1 were used as a bait to pull down the N-terminal hexa-histidine-tagged EfTu variants, (C) SDS-PAGE and Western blot analysis of GST pull-down experiments with components expressed in *E. coli*: N-terminal GST fused MLLE variants of Rrm4 and Pab1 were used as a bait to pull down the N-terminal Hexa-Histidine-SUMO-tagged PAM2L^{EfTu} variants. The experiment was performed with the soluble fraction of *E. coli* cell lysate to demonstrate specific binding. Results were analyzed with α-His immunoblotting. (D) 3D structure of *E. coli* EfTu (PDB ID: 1EFC) depicted in similar coloring given in the schematic protein bars (E) AlphaFold2 predicted mitochondrial *(left)* and cytoplasmic *(right)* EfTu variant structures from *U. maydis*.

Inspired by these results, the *U. maydis* proteome was analyzed to look for an orthologue of *E. coli*^{EfTu} using NCBI-BLAST (National Center for Biotechnology Information-Basic Local Alignment Search Tool). BLAST search identified two proteins with high sequence identity and similarity. One of them was the mitochondrial EfTu of *U. maydis* (*Um*EfTuM, UMAG_00138) which have 98% coverage and 62% sequence identity. The second one was the cytoplasmic Elongation factor 1 alpha subunit (*Um*EfTuM-1α, UMAG_00924) which has 96% coverage and 30% sequence identity.

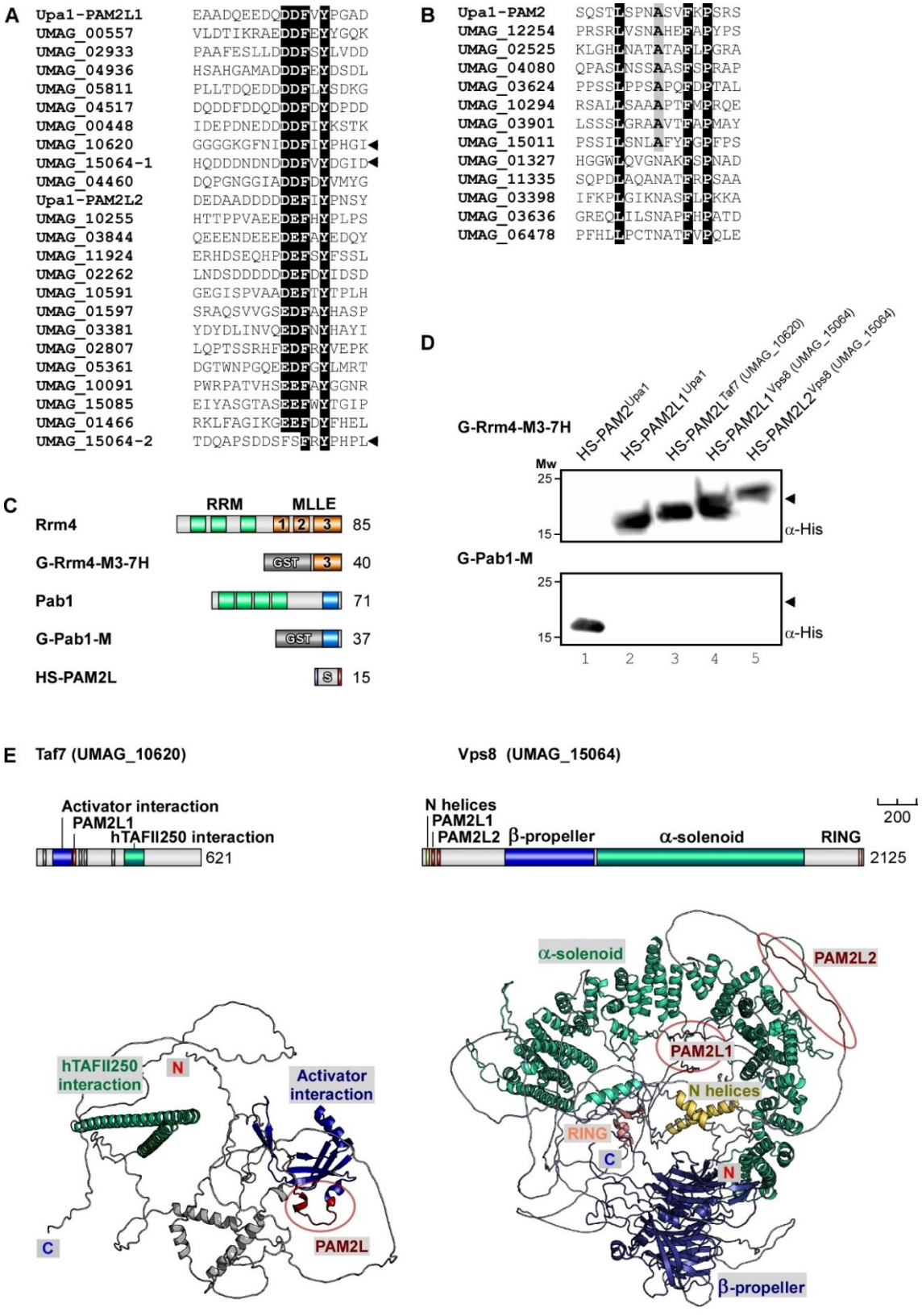
The structures of *Ec*EfTu (Figure 18 D, PDB ID: 1EFC) were compared with AlphaFold2 predicted structures of *Um*EfTuM and *Um*EfTuM-1α to obtain structural insights. All EfTu tertiary structures are well conserved, with all of them having an N-terminal guanosine diphosphate (GDP) binding domain (Figure. 18A, D, E, shown in red) comprising switch 1 (yellow) and switch 2 (green) helices, followed by domain II and III (blue) (SONG *et al.* 1999). Although the domain architecture is similar, the switch 1 region in *E. coli* and *U. maydis* have a different secondary and tertiary structure. In *E. coli* this switch consists of a helix and β-hairpin, whereas in *U. maydis*, both EfTu versions contain only two helices (Figure. 18 D, E, shown in yellow) . PAM2 motifs are found only in the low complexity region and not in the globular domains (XIE *et al.* 2014). The potential PAM2L motif in all three EfTu is present in a short, low complexity region (LCR) that connects the N-terminal GDP domain with the C-terminal domains II and III (Figure. 18 D, E, shown in cyan). Multiple sequence analysis (MSA) showed that this region containing the PAM2L motif is well conserved from *E. coli* to higher eukaryotes (Figure S14C, (SONG *et al.* 1999; ANDERSEN *et al.* 2000)).

It is noteworthy that many more *E. coli* proteins co-purified in GST-Rrm4-M3 pull-down experiments were identified along with EfTu by Mass spectrometry analysis. However, they showed less enrichment, suggesting these were not strong interaction partners. However, some of them have orthologues in *U. maydis*, for example (UMAG_10397, UMAG_10213, UMAG_11194, UMAG_04871, UMAG_10528, UMAG_00115, UMAG_10836, UMAG_05019, UMAG_06461, UMAG_05776, UMAG_11855, UMAG_04472, UMAG_10659). This observation indicates that these proteins are conserved in evolution. The complete list is provided in Supplementary Table S17. Based on the above-mentioned strategy, it is an interesting list for finding more MLLE3^{Rrm4} interaction partners in *U. maydis* in the future.

In summary, LC-MS analysis identified *E. coli* EfTu as a highly enriched candidate co-eluted with the G-MLLE3^{Rrm4} in the GST pull-down assays. *In vitro* binding studies confirmed that *E. coli* EfTu specifically interacts with MLLE^{Rrm4} *in vitro* but not with MLLE^{Pab1}. BLAST search identified mitochondrial and cytoplasmic EfTu of *U. maydis* are the orthologues of *E. coli* EfTu with high sequence coverage and identity. MSA and structural comparison suggest that the potential PAM2L motif in *E. coli* EfTu is highly conserved in mitochondrial as well as cytoplasmic EfTu of *U. maydis*.

4.4 Identification of novel PAM2L containing proteins as interaction partners of Rrm4 in *U. maydis*

Structural studies identified the critical residues of PAM2L1,2^{Upa1} motifs required for the MLLE3^{Rrm4} binding. Essentially, these are the conserved FxY hydrophobic residues in the C-terminus of the PAM2L1,2^{Upa1} peptides. Three levels of screening were performed to find more potential PAM2L-containing proteins as novel interaction partners of the MLLE3^{Rrm4} in *the U maydis* genome. As a first step, a reiterative BLAST search was performed by submitting PAM2L1^{Upa1} and PAM2L2^{Upa1} sequences (Figure 19A, Upa1-PAM2L1, Upa1-PAM2L2) as inputs in NCBI-BLAST which identified Upa1 as the top candidate containing PAM2L sequences, with high sequence similarity and identity as expected. In addition, it fetched several new PAM2L-containing candidates with the crucial FxY pocket. In the next step, accessibility of the PAM2L motif for MLLE interaction was analyzed by looking for the presence of PAM2L sequences within the Low Complexity Region (LCR) using AlphaFold2 predicted 3D structure. In the third step, MSA of the chosen candidate was performed with its orthologue proteins from the representative basidiomycetes smut fungi (*U. maydis, Pseudozyma hubeiensis, Kalmanozyma*



4.4.1.1 Figure 19. *De novo* predicted PAM2L peptides interact with MLLE3^{Rrm4}.

(A) *De novo* predicted PAM2L sequences, conserved key residues are in the black shade (B) *De novo* predicted PAM2^{Upa1} sequences, conserved key residues are in the black shade (C) Schematic representation of protein variants used in the *in vitro* binding studies (Molecular weight in kilo Dalton indicated) using the following coloring: lime green, RNA recognition motif (RRM); orange, MLLE^{Rrm4} domains; dark blue,

MLLE^{Pab1}; GST and SUMO tags are labeled, Red, PAM2L motif. Variant amino acids of the FxP and FxxP of PAM2 and PAM2L sequences are printed in grey font. (D) Western blot analysis of GST pull-down experiments with components expressed in *E. coli*: N-terminal Hexa-Histidine-SUMO-tagged PAM2L variants were pulled-down by N-terminal GST fused MLLE variants of Rrm4 and Pab1. The experiment was performed with the soluble fraction of *E. coli* cell lysate to demonstrate specific binding. Results were analyzed with α-His immunoblotting. (E) Schematic representation of proteins identified as potential interaction partners of MLLE3^{Rrm4} from the *in vitro* binding studies. (*Top panel*) (bar, 200 amino acids, number of amino acids indicated next to protein bars), 3D structure predicted from AlphaFold2 *(lower panel)*. TAF7 (UMAG_10620) domains are depicted in the coloring dark blue, activator interaction domain; dark green, dark red, PAM2L; homolog of human TAFII250 interaction domain; grey, N-terminal helices. Vps8 (UMAG_15064) domains are depicted in the following coloring golden yellow, N-terminal helices; dark red, PAM2L1,2; dark blue, b-propeller domain; dark green, α-solenoid domain; light pink, RING domain. AlphaFold2 predicted 3D structures of TAF7 (*left*) and Vps8 (*right*) using the similar coloring used in the schematic protein bars. PAM2L motifs in TAF7 and Vps8 are present in the low-complexity regions (Circled in red).

brasiliensis, Sporisorium reilianum, Melanopsichium pennsylvanicum, Ustilago hordei, Moesziomyces antarcticus, Testicularia cyperi) which contains an Rrm4 type protein with a 7-helix-MLLE3 (MULLER et al. 2019). Since the known PAM2L1,2^{Upa1} sequences are ligands for the MLLE3^{Rrm4}, the newly identified PAM2L sequences were expected to be conserved in basidiomycetes fungi with a conserved MLLE3^{Rrm4}. Noteworthy to mention, sequences of the chosen PAM2L containing candidate from *Malassezia globosa* (*M. globose)* and *Cryptococcus neoformans var grubii* (*C. neoforman)* were included as a negative control in MSA because these organisms do not have an MLLE3 containing Rrm4 orthologue (MULLER et al. 2019). Therefore, it is expected that PAM2L sequences are not conserved as well in the respective protein orthologue due to evolutionary selection pressure. With the three-level, as mentioned earlier, stringent selection criteria, about two dozen novel PAM2L-containing candidates were identified with high sequence identity in the crucial FxY binding residues (Figure 19A, Table S18). Manual annotation of the highly conserved PAM2L-containing candidates revealed that several RNA-binding proteins, Ubiquitin ligases, Proteases, Carbohydrate enzymes, and an endosomal tethering protein were present in the list (Figure 19A, Table 19).

Three PAM2L candidates from the list were further analyzed (Figure 19A, 1. UMAG_10620-PAM2LTaf7, 2. UMAG_15064-PAM2L1Vps8, 3. UMAG_15064-PAM2L2Vps8). The first PAM2L candidate from UMAG_10620 is Taf7, the orthologue of yeast *Schizosaccharomyces pombe (S. pombe)* (Transcription initiation factor TFIID subunit 7 - *Taf7*, also known as *Ptr6*) (SHIBUYA *et al.* 1999). The second and third PAM2L candidates are from UMAG_15064, Vps8 - Early endosome specific subunit of CORVET complex (SCHNEIDER *et al.* 2022).

AlphaFold2 predicted structures of Taf7 and Vps8 exhibited that the PAM2L motif of Taf7 and PAM2L1,2 motifs of Vps8 are present in the low complexity region (LCR) (Figure 19E), which is comparable to the presence of PAM2L^{Upa1} motifs in LCR in Upa1 (Figure S15 C). Multiple sequence analysis revealed that the PAM2L motif of Taf7 and both PAM2L1,2 motifs of Vps8 are conserved in the representative basidiomycetes smut fungi but not conserved in the *M. globose* and *C. neoformans* (Figure S15A, B).

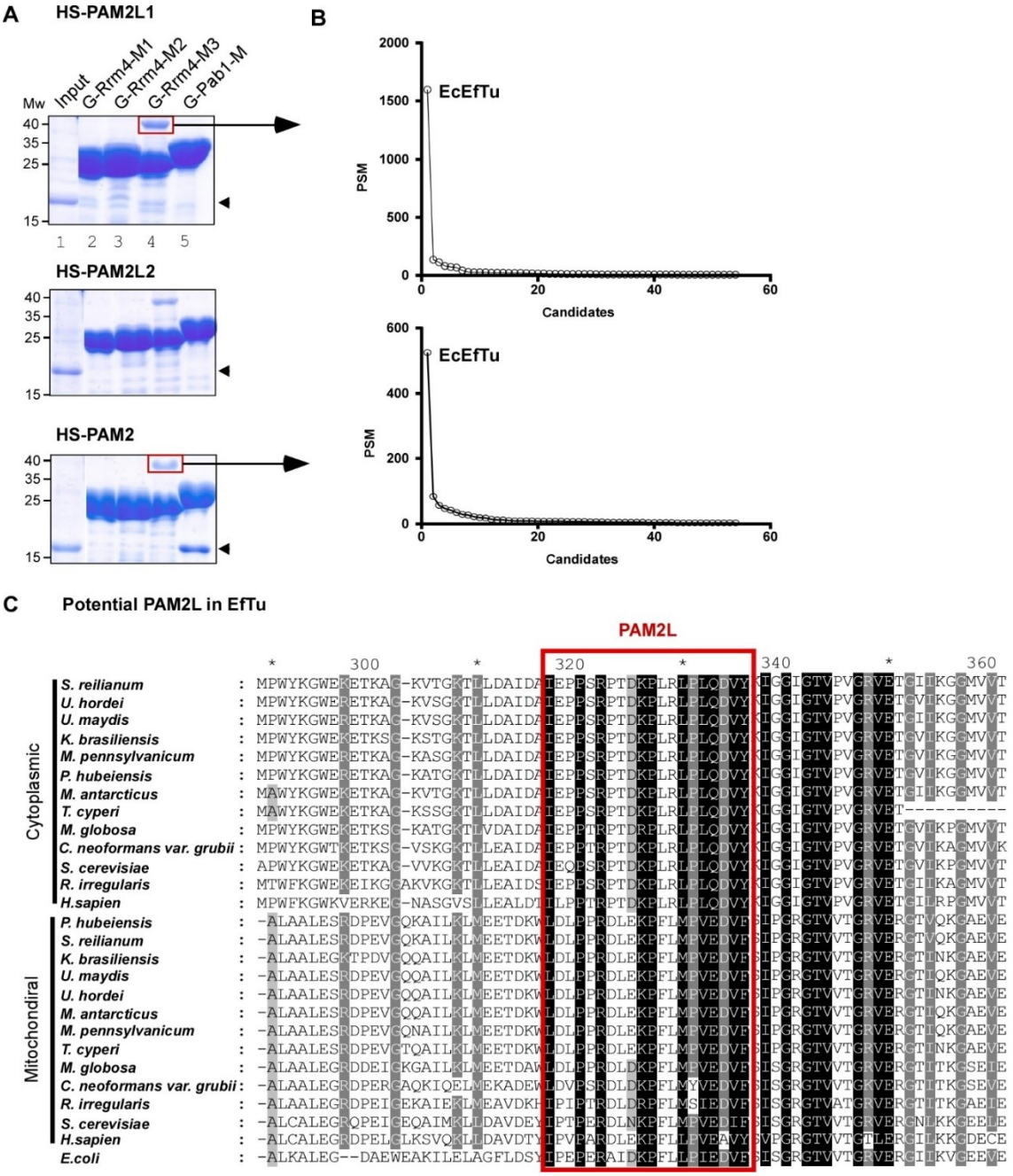
To evaluate the interaction capacity of the *de novo* predicted PAM2L sequences, *in vitro* binding studies were performed using the GST pull-down assay as described earlier. To this end, PAM2L sequences comprising 18 amino acid fragments of computationally predicted candidates (Figure 19A, C, E, HS-PAM2L^{Taf7}, HS-PAM2L1^{Vps8}, HS-PAM2L2^{Vps8}) were expressed as N-terminal hexa-histidine-SUMO (HS) fusion proteins (see Materials and methods). MLLE3^{Rrm4} version specifically interacted with PAM2L1 motif of Upa1 but not with PAM2 motif of Upa1, whereas MLLE^{Pab1} specifically interacted with PAM2 motif of Upa1 but not with PAM2L1,2 motif of Upa1 (POHLMANN *et al.* 2015; DEVAN *et al.* 2022). Therefore, HS-PAM2L1^{Upa1} and HS-PAM2^{Upa1} variants were used as positive and negative controls, respectively. As bait proteins in the GST pull-down assay, MLLE3^{Rrm4} and MLLE^{Pab1} were expressed as a fusion protein with an N-terminal GST tag (Figure 19C, G-Rrm4-M3-7H, G-Pab1-M).

In GST pull-down experiments using GST fusion proteins as bait, G-Pab1-MLLE interacted with the positive control HS-PAM2 motif of Upa1 but did not recognize the negative control HS-PAM2L1,2 motifs of the Upa1 or the *de novo* predicted HS-PAM2L motifs (Figure 19D, G-Pab1-M, lane 1-5). In contrast, G-Rrm4-M3-7H did not interact with the negative control HS-PAM2 motif of Upa1 (Figure 19D, G-Rrm4-M3, lane 1) but interacted with the positive control HS-PAM2L1 motif of Upa1 (Figure 19D, G-Rrm4-M3-7H, lane 2) and recognized all the three *de novo* predicted HS-PAM2L motifs (Figure 19D, G-Rrm4-M3, lane 3-5, PAM2L^{Taf7}, HS-PAM2L1^{Vps8}, HS-PAM2L2^{Vps8}). G-Rrm4-M3 showed moderate binding with HS- PAM2L motif of Taf7 and the HS-PAM2L2 motif of Vps8 (Figure 19D; lane 3, 5) where, as it showed a strong binding with HS-PAM2L1 motif of Vps8 (Figure 19D lane 4). These results demonstrated that the *de novo* predicted PAM2L motifs are recognized by the MLLE3^{Rrm4} *in vitro.* Therefore, Taf7 and Vps8 could be the potential interaction partners of MLLE3^{Rrm4}.

A similar approach using reiterative BLAST and LCR search in AlphaFold2 predicted structures identified, 12 novel potential PAM2 containing proteins in *U. maydis* (Figure 19B, Table S19) which has to be verified experimentally.

To sum up, bioinformatics analysis with stringent screening conditions identified several novel PAM2L-containing candidates. PAM2L sequences from Taf7 and Vps8 were taken for further analysis. AlphaFold2 structures revealed that PAM2L sequences of both Taf7 and Vps8 are present in the LCR. MSA showed that even though PAM2L sequences of both Taf7 and Vps8 are conserved in the basidiomycetes fungi, they are not conserved in the organisms that do not have MLLE3^{Rrm4} (*M. globose,* and *C. neoformans). In vitro* pull-down assays confirmed that PAM2L sequences of Taf7 and Vps8 are recognized by MLLE3^{Rrm4}. Therefore, they could be potential interaction partners of MLLE3^{Rrm4}.

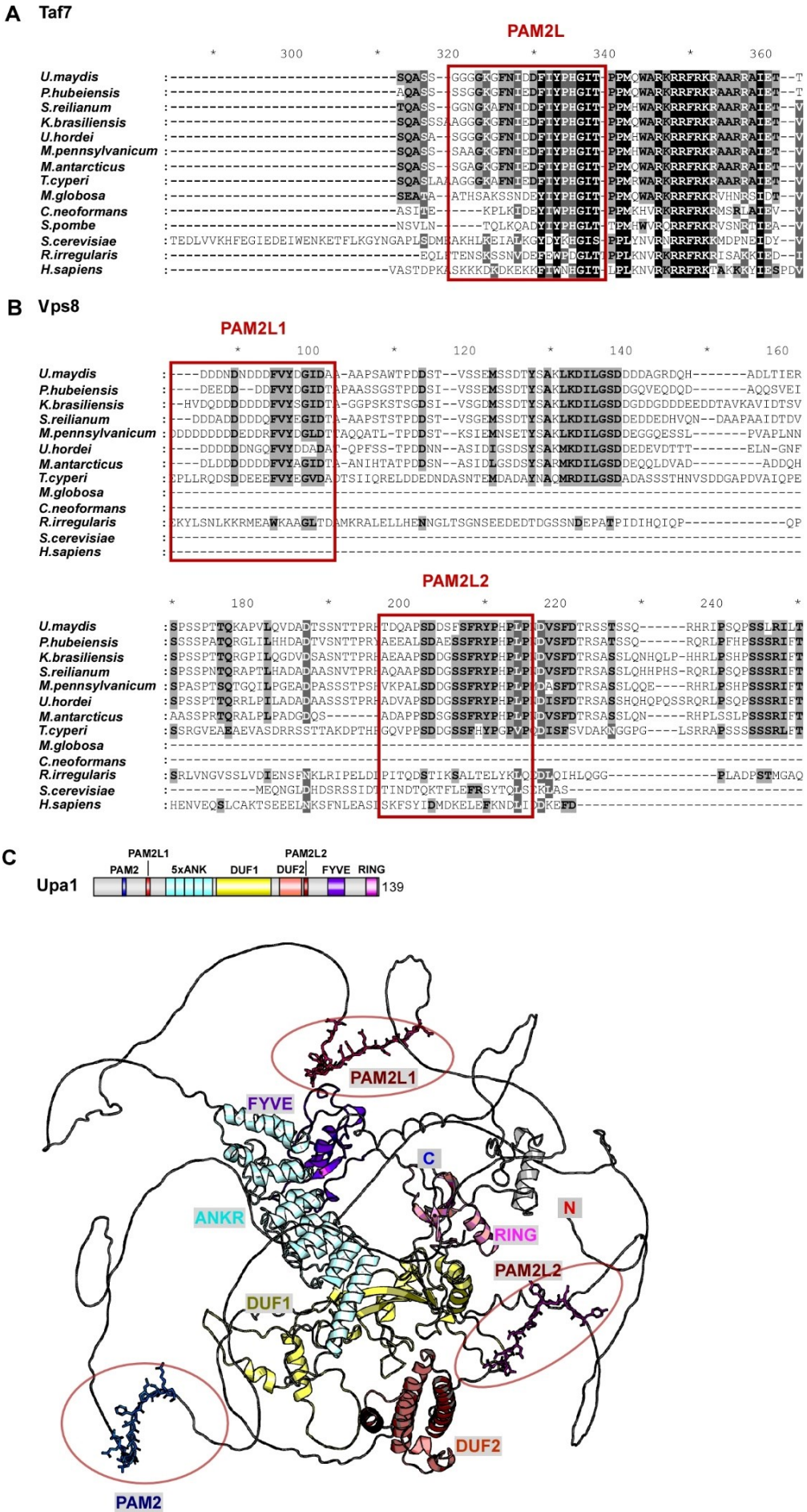
4.5 Supporting information



4.5.1.1 Figure S14. *E. coli*-EfTu interacts with MLLE3^{Rrm4}.

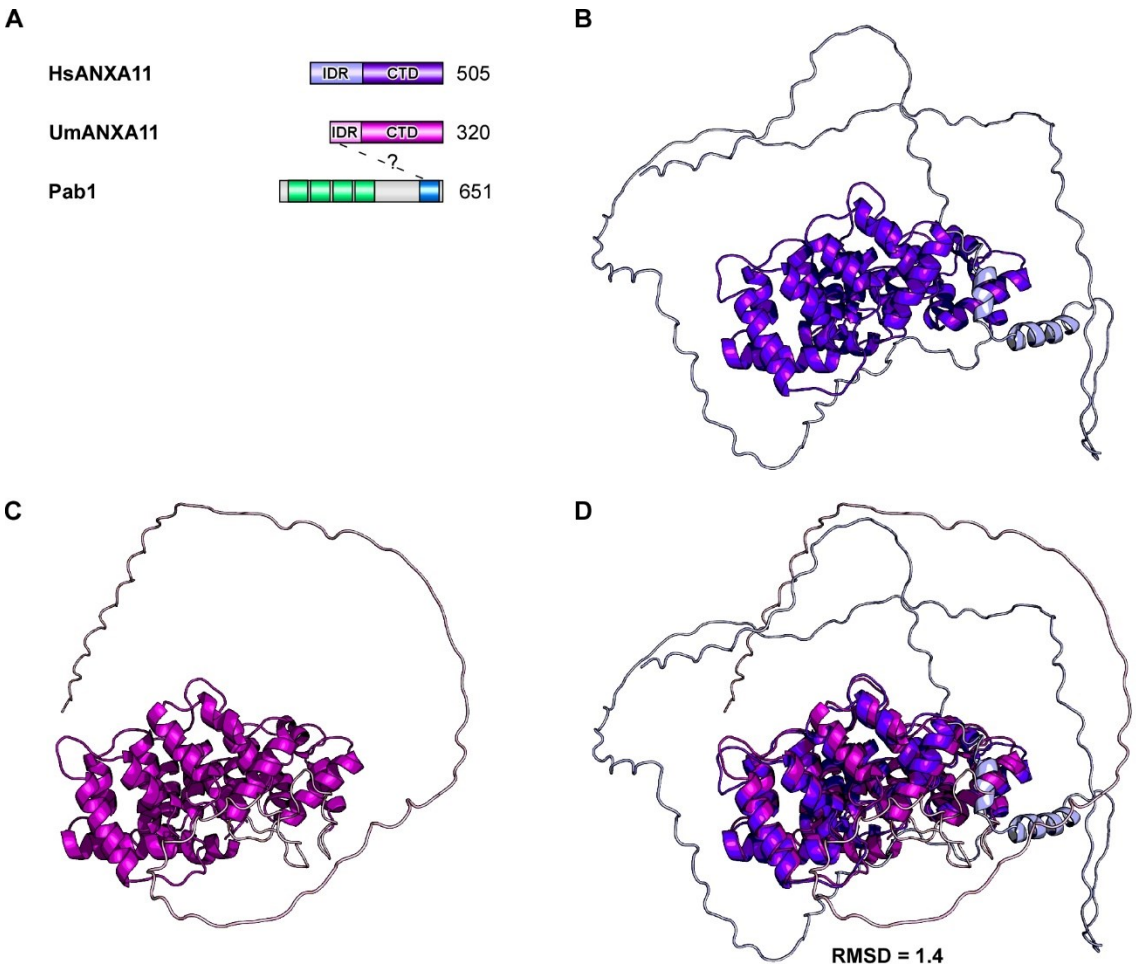
(A) SDS-PAGE analysis of GST pull-down experiments with components expressed in *E. coli*: N-terminal Hexa-Histidine-SUMO-tagged PAM2 variants were pulled-down by N-terminal GST fused MLLE variants of Rrm4 and Pab1. The experiment was performed with the soluble fraction of *E. coli* cell lysate to demonstrate specific binding. (B) LC-MS/MS analysis of the unknown protein band found in the GST pull-down experiments using G-Rrm4-M3 as bait and HS-PAM2L1 as prey (Top), HS-PAM2 as prey (bottom) identified that *E. coli*-EfTu is the most enriched candidate in both the samples. (C) Sequence alignment of potential PAM2L containing region of EfTu orthologues in representative organisms (Basidiomycetes: *S. reilianum, U. hordei, U. maydis, K. brasiliensis, M. pennsylvanicum, P. hubeiensis, M. antarcticus, T. cyperi, M. globosa and C. neoformans var grubii*; Ascomycota: *S. cerevisiae*; Mucoromycota: *R. irregularis*; Metazoa: *H. sapiens*; and *E. coli*) showing that this region is conserved well across all the organism

in both cytoplasmic and mitochondrial Eftu. Especially PAM2L sequence in mitochondrial EfTu contains a conserved FxxP pocket for MLLE3^{Rrm4} interaction. Accession number are provided in the S13 Table. Multiple sequence alignment was performed using ClustalW.


A Taf7
PAM2L
U.maydis
P.hubeiensis
S.reilianum
K.brasiliensis
U.hordei
M.pennsylvanicum
M.antarcticus
T.cyperi
M.globosa
C.neoformans
S.pombe
S.cerevisiae
R.irregularis
H.sapiens
B Vps8
PAM2L1
PAM2L2
C
Upa1
PAM2
5xANK
DUF1
DUF2
FYVE
RING
139
PAM2L1
FYVE
C
ANKR
N
RING
PAM2L2
DUF1
DUF2
PAM2

4.5.1.2 Figure S15. *De novo* predicted PAM2L sequences are conserved only in basidiomycetes filamentous fungi.

(A) Sequence alignment of *de novo* predicted PAM2L containing region of Taf7 orthologues in representative organisms (Basidiomycetes: *U. maydis, P. hubeiensis, K. brasiliensis, S. reilianum, M. pennsylvanicum, U. hordei, M. antarcticus, T. cyperi, M. globose, and C. neoformans var grubii.* Mucoromycota: *R. irregularis*; Ascomycota: *S. cerevisiae*, Metazoa: *H. sapiens*) (B) Sequence alignment of *de novo* predicted PAM2L containing region of Vps8 orthologues in representative organisms as described above. MSA analysis exhibited that *de novo* predicted PAM2L sequences are conserved in the basidiomycetes unipolar filamentous fungi having a conserved MLLE3^{Rrm4} type domain which excludes *M. globose*, and *C. neoformans var grubii*. Accession number and sequence coverage are listed in Tables S12 and S13. Multiple sequence alignment was performed using ClustalW. (C) 3D structure of Upa1 predicted from AlphaFold2. Upa1 (UMAG_12183) domains are depicted in the following coloring dark blue, PAM2^{Upa1}; dark red, PAM2L^{Upa1}; yellow, DUF domain1; Cyan, Ankyrin repeat domain, Salmon red, DUF domain2, Pink, RING; PAM2 and PAM2L motifs of Upa1 are present in the low complexity regions (Amino acid resides are shown as sticks and circled in red).



4.5.1.3 Figure S16. ANXA11 in human and *U. maydis* are conserved with a long IDR.

(A) Schematic representation of protein variants drawn to scale (bar, 200 amino acids, number of amino acids indicated next to protein bars) using the following coloring: light blue, intrinsically disordered region (IDR) of *Hs*ANXA11; Purple, C-terminal domain (CTD) of *Hs*ANXA11; light pink, IDR of *Um*ANXA11; Magenta, CTD of *Um*ANXA11. (B) AlphaFold2 predicted 3D structure of human ANXA11, (C) AlphaFold2 predicted 3D structure of *U. maydis* ANXA11 (D) Structural alignment of the CTD domains from human and *U. maydis* ANXA11. The all-atom RMSD is 1.4 Å showing a highly similar domain structure.

5 Discussion

Long-distance mRNA transport is crucial for unipolar growth in *Ustilago maydis*. Rrm4 is the key mRNA binding protein that mediates the mRNA transport on early endosomes in a microtubule-dependent manner (MUNTJES *et al.* 2021). Rrm4 has three RRM domains in the N-terminal for mRNA binding and interacts with thousands of mRNAs (OLGEISER *et al.* 2019). The C-terminal of Rrm4 is essential for endosomal attachment. Loss of the C-terminus of Rrm4 results in bipolar filaments, similar to the deletion phenotype, suggesting loss of Rrm4 function (BECHT *et al.* 2006; POHLMANN *et al.* 2015). Rrm4 contained two MLLE domains in the C-terminus, which interacted with the PAM2L1, 2 motifs of the endosomal adaptor protein Upa1. Both PAM2L^{Upa1} motifs are important for interaction with the C-terminus of Rrm4. Loss of Upa1 also results in the bipolar phenotype (POHLMANN *et al.* 2015).

In the first part, using a combination of structural biology, biochemistry, biophysics, molecular genetics, and cell biology methods, this study characterized the C-terminal region of the Rrm4 in detail. It demonstrated that, 1. It consisted of three tandem MLLE^{Rrm4} domains, contrary to two MLLE^{Rrm4} as previously thought. 2. These tandem MLLE^{Rrm4} domains formed a sophisticated protein-protein interaction platform. 3. Elucidated their binding behavior and revealed that a strict hierarchy governs them.

MLLE2^{Rrm4} and MLLE3^{Rrm4} domains are involved in endosomal attachment. While MLLE2^{Rrm4} is an accessory domain, MLLE3^{Rrm4} is an essential domain and the main interaction partner for the PAM2L1,2 motifs of Upa1. Of the three MLLE domains described here, the first, MLLE1^{Rrm4}, was *de novo* predicted in this work by homology modeling without a known function. The interaction partners of MLLE2^{Rrm4} and MLLE1^{Rrm4} are yet to be discovered.

In the second part, using X-ray crystallization and *in vitro* biochemical assays, this study demonstrated that MLLE3^{Rrm4} consisted of 7-alpha helices, the newly found helices (α I, II) are important for PAM2L1,2^{Upa1} interaction. Co-crystallized structures of the MLLE3 domain of Rrm4 with PAM2L1,2 peptides of Upa1 have revealed that MLLE3^{Rrm4} interacted with PAM2L1,2^{Upa1} peptides non-canonically and identified the crucial residues mediating the protein-peptide interactions.

In the third part, by applying the knowledge obtained from the MLLE3^{Rrm3}, MLLE^{Pab1} biochemical and structural studies, this study identified several novel interaction partners of Rrm4 and Pab1 and verified some of these predicted candidates by *in*

vitro biochemical assays. MLLE domain of Pab1 interacted with Rrm4 via the internal PAM2 motif present in the Rrm4. Mitochondrial EfTu was identified as a potential interaction partner of Rrm4 based on the interaction of MLLE3^{Rrm4} with the structurally similar *E. coli* EfTu. Endosomal covert complex protein Vps8 and Nucleocytoplasmic transport protein Taf7 were identified as a potential interaction partners of Rrm4 by *in vitro* studies.

5.1 Tripartite MLLE platform in *U. maydis*

MademoiseLLE (MLLE) domains were first identified in the C-terminal region of the Poly(A)-binding protein (PABPC1) as a conserved protein-protein interaction (PPI) domain (KUHN AND PIELER 1996). MLLE interacts with the PAM2 motif-containing proteins specifically. Only two proteins contain MLLE domain in higher eukaryotes, for example, human PABPC1 and Ubr5 (XIE *et al.* 2014). Rrm4 is the third protein known to contain the MLLE domain. However, it specifically interacts with the PAM2L motifs but not the PAM2 motif of the endosomal adaptor protein Upa1 in *U. maydis* (POHLMANN *et al.* 2015).

Neural network-based computational modeling with Topmodel, AlphaFold2, and RosettaFold2 predicted that Rrm4 has three MLLE domains in its C-terminal portions (Figure 6C). SAXS experiments performed using purified recombinant proteins confirmed the volumetric shape and relative spatial arrangements of the three MLLE domains in the C-terminus of Rrm4 and the three RRM domains in the N-terminus (Figure 7). Such a domain organization is conserved in fungi; even Rrm4 versions of the distantly related fungus *Rhizophagus irregularis* (Mucoromycota) contain three MLLE domains in its C-terminus (Figure 20C, determined by AlphaFold2). Tandem MLLE^{Rrm4} domains are not found elsewhere in plants or animals (XIE *et al.* 2014). Rrm4, an mRNA transport protein on a dynamic endosomal membrane, might be involved in numerous PPI interactions. Structural studies have shown that the stoichiometry of MLLE-PAM2 interaction is 1:1 (XIE *et al.* 2014) which indicates that each MLLE domain of Rrm4 should interact with one PAM2 containing protein at a time. Therefore, having tandem MLLE domains is likely advantageous to Rrm4 for multiple MLLE^{Rrm4}-mediated protein-protein interactions simultaneously. For example, these three MLLE domains enable the Rrm4 to be involved in protein interactions necessary for endosomal tethering and association with accessory RBPs and cargo proteins during endosomal hitchhiking.

5.2 Spatial arrangements of the RRM and MLLE domains

Structural studies using SAXS and AlphaFold2 predictions have shown that N-terminal RRM and C-terminal MLLE domains are separated from each other by intrinsically disordered regions (IDR) (Figure 7D, S1B). This arrangement keeps the RRM domains required for RNA interactions distinct from the MLLE domains required for protein-protein interactions farther from the mRNA strand, which may relax steric or dynamic constraints in assembling the mRNPs. Human PABPC1 also has a similar spatial arrangement (MELO *et al.* 2003; SCHAFER *et al.* 2019).

All three MLLE^{Rrm4} domains are individually separated from each other by flexible linker regions and have differential peptide specificity, pointing out that their spatial arrangement should be dynamic and regulated by their binding partners. In comparison, RRM domains of Pab1 are individually separated from each other by the flexible linker region and have differential binding preferences. In human PABPC1, RRMs 1 and 2 bind eIF4G and PAIP1 with high affinity to the poly(A) binding site, whereas RRMs 3 and 4 bind poly(A) with reduced affinity but, in addition, binds AU-rich RNA and mediate protein-protein interaction with eEF1α (GOSS AND KLEIMAN 2013). Interestingly, MLLE1^{Rrm4}, MLLE2^{Rrm4}, and MLLE3^{Rrm4} are more similar to their counterpart in other *Basidiomycota* species than within the same species. For example, in all the representative organisms, MLLE1^{Rrm4} consisted of a five-helical, MLLE2^{Rrm4} consisted of a four-helical, and MLLE3^{Rrm4} consisted of a seven helical structures (AlphaFold2 predictions; (MULLER *et al.* 2019)). A similar observation is reported for RRM domains of PABPC1. For example, each of the four RRM domains of the PABPC1 is conserved and similar to different organisms than the RRM domains within the same organism (GOSS AND KLEIMAN 2013). Tandem MLLE domains of Rrm4 may interact with multi-PAM2L variants in yet unknown proteins non-canonically.

5.3 IDR mediated regulation of MLLE-PAM2/PAM2L interactions

Structural analysis using SAXS and AlphaFold2 predictions showed that each MLLE domain of Rrm4, PAM2, and PAM2L motifs of Upa1 was surrounded by IDRs (Figure S15 C), which provides flexibility and mobility for protein-protein interactions. In humans, most RNA-binding proteins are highly enriched with IDRs (CASTELLO *et al.* 2012)**.** PAM2 motifs are generally located in the IDR with adjacent clusters of phosphorylation sites. MLLE-PAM2 interaction is proposed to be a two-step process. Initial interaction is

made by the direct contact between MLLE and PAM2, which induces the folding of IDR and promotes further contacts to strengthen the interaction. During this process, reversible phosphorylation at the clusters of serine or threonine residues adjacent to the PAM2 motifs within the IDR serves to modulate the interaction (HUANG *et al.* 2013). In humans, MLLE domain of Ubr5 interacts with its own PAM2L motif present adjacent to the HECT. Ubr5 is heavily phosphorylated, which may lead to a conformational change in the protein that regulates the ubiquitin activity (BETHARD *et al.* 2011; MUNOZ-ESCOBAR *et al.* 2015). Ser2484 is one of the phosphorylation sites in Ubr5 located between the $MLLE^{Ubr5}$ domain and the $PAM2L^{Ubr5}$ motif. This residue could be a potential regulator of the MLLE-PAM2L interaction in Ubr5.

In alignment with this, clusters of serine and threonine are found adjacent to $PAM2^{Upa1}$ and $PAM2L^{Upa1}$. This observation indicates potential phosphorylation-mediated modulation of MLLE-PAM2/PAM2L interactions in endosomal mRNPs in *U. maydis*.

5.4 $MLLE3^{Rrm4}$ is the essential domain

In vitro and *in vivo* studies have demonstrated that $MLLE3^{Rrm4}$ is the only domain interacting with the PAM2L1,2 motifs of Upa1 (Figure 8, 9). Quantitative ITC measurements revealed that the $MLLE3^{Rrm4}$ domain binds to the $PAM2L2^{Upa1}$ motif with higher affinity than the $PAM2L1^{Upa1}$ motif but not the PAM2 motif, with the binding affinities (Figure 8B, K_D = 5 and 15 µM respectively) comparable to the human $MLLE^{PABPC1}$ PAM2 motifs interactions (MATTIJSSEN *et al.* 2021). Similarly, the MLLE domain of Pab1 showed binding to PAM2 of Upa1 (Figure S5B, K_D = 5 µM) but no binding with the $PAM2L1,2^{Upa1}$ peptides. In addition, ITC results also reconfirmed the binding specificities of the $MLLE^{Pab1}$ and $MLLE^{Rrm4}$ motifs reported earlier (POHLMANN *et al.* 2015). In comparison, the human MLLE domain of PABPC1 did not interact with the PAM2L motif of Ubr5, emphasizing the binding specificity of $MLLE^{Pab1}$ to the PAM2. The MLLE domain of Ubr5 also exhibited differential binding, which binds to the PAM2 of PAIP1 with a K_D of 3.4 µM and PAM2 of $HECT^{Ubr5}$ with a K_D of 50 µM (MUNOZ-ESCOBAR *et al.* 2015).

5.5 $MLLE1,2^{Rrm4}$ are accessory domains

Generally, $MLLE^{PABPC1}$ consists of five helices, but $MLLE^{Ubr5}$ and yeast $MLLE^{Pab1}$ have four helices (DEO *et al.* 2001; KOZLOV *et al.* 2002). The crystal structure revealed that

MLLE2^{Rrm4} has only four helices (α2-5), and resembles the MLLE domain of human Ubr5 and yeast MLLE^{Pab1} (DEO *et al.* 2001; KOZLOV *et al.* 2002). Computational models showed MLLE1^{Rrm4} has a typical five-helix structure. Although MLLE1^{Rrm4} and MLLE2^{Rrm4} have a conserved overall structure similar to MLLE3^{Rrm4}, the absence of these domains did not affect the PAM2L1,2^{Upa1} peptide binding with the MLLE3^{Rrm4} *in vitro*. Remarkably, careful analysis of MLLE domains *in vivo* revealed that MLLE2^{Rrm4} plays an accessory role in the proper attachment of Rrm4 during endosomal shuttling. Nevertheless, this study could not identify a clear interaction partner for MLLE2^{Rrm4}. Notably, MLLE2^{Rrm4}, when expressed alone, interacted with PAM2L1^{Upa1} motifs in the *in vitro* pull-down experiments. Surprisingly it also recognized the PAM2 peptide (Figure 12B, lane 4). This observation indicated that MLLE2^{Rrm4} could recognize both PAM2^{Upa1} and PAM2L^{Upa1} sequences, has a flexible target specificity, and the presence of the other two MLLE^{Rrm4} domains might be regulating MLLE2^{Rrm4} interactions. As an alternate hypothesis, MLLE2^{Rrm4} might interact with a PAM2 variant, an intermediate version of PAM2 and PAM2L sequences. Hence it could recognize both these kinds. Therefore, it requires further investigation with a chimeric peptide.

In the case of MLLE1^{Rrm4}, we have yet to identify a clear function. Nonetheless, the observation that the tripartite MLLE^{Rrm4} domain platform is conserved across basidiomycetes fungi and even in the distantly related Mucoromycota fungus *R. irregularis* indicates that all the three MLLE domains could be functionally important (Figure 20B, (MULLER *et al.* 2019; DEVAN *et al.* 2022)).

5.6 MLLE3^{Rrm4} - PAM2L1,2^{Upa1} interaction is non-canonical

Biochemical and biophysical studies confirmed that MLLE3^{Rrm4} is the main domain interacting with the PAM2L1,2 motifs of Upa1 (Figure 8B-D). Initial crystallization attempts with the C-terminal domains of the Rrm4 (Rrm4-NT4) version carrying all three MLLE domains only produced crystals containing a truncated version of the MLLE2^{Rrm4} domain (Figure 7A). AlphaFold2 predictions identified that MLLE3^{Rrm4} has two additional helices on the N-terminal (α I, II), making it an exceptional MLLE domain consisting of seven helices (Figure 12C). Biochemical studies conducted with MLLE3^{Rrm4} variants carrying all seven helices or the last five helices demonstrated that helices α I and II were essential for interaction with PAM2L1,2^{Upa1} and a five-helices containing MLLE3^{Rrm4} domain was insufficient for interaction. In comparison, human MLLE^{PABPC1}

with five helices, or yeast MLLE^{Pab1} and human MLLE^{Ubr5} with four helices, are sufficient for interaction (DEO *et al.* 2001; KOZLOV *et al.* 2001; KOZLOV *et al.* 2002; XIE *et al.* 2014; MUNOZ-ESCOBAR *et al.* 2015).

Co-crystallization studies were carried out with PAM2L1,2^{Upa1} peptides, and the newly designed MLLE3^{Rrm4} constructs with seven helices based on the AlphaFold2 predicted structure. Unlike the earlier attempts, protein-peptide complexes yielded crystals rapidly, whereas MLLE3^{Rrm4} failed to form crystals independently. Similarly, MLLE^{Pab1} also crystallized with its ligand PAM2, whereas it did not crystallize independently. Likewise, in humans, MLLE^{PABPC1} domain was unable to form crystals but rapidly crystallized upon adding 'PAM2w' peptide variants (GRIMM *et al.* 2020). One explanation for this observation is that during the crystallization process, the peptides stabilize flexible regions of the MLLE domain, which results in stable protein-peptide complex formation that, in turn, favors the crystallization condition (Figure S12C, S13C). It is supported by the fact that both PAM2L^{Upa1} and PAM2^{Upa1} peptides bind to their respective MLLE domains with high affinity in the ITC experiments (Figure 8B).

X-ray structures of complexes confirmed that MLLE3^{Rrm4} and MLLE^{Pab1} fold into a stable, compact domain with seven and four helices, respectively. Both the MLLE polypeptides were wound into a right-handed, superhelix structure as predicted by AlphaFold2 (Figure 14B, 16B). Structural comparison of MLLE3^{Rrm4} and MLLE^{Pab1} exhibited that both MLLE domains are structurally similar, having characteristic folds of the MLLE domains from helices 2-5, except that MLLE3^{Rrm4} has a longer first helix and two additional helices (α I, II) compared to the MLLE^{Pab1} (Figure 16C). In the case of MLLE^{Pab1}, the protein variant with five helices did not form the crystals, whereas the MLLE^{Pab1} variant with four helices formed the crystals rapidly. In humans MLLE^{PABPC1} having four helices, has been crystallized with PAM2^{Paip2} (Figure S13D, PDB ID: 3KUS). Unlike MLLE3^{Rrm4}, the MLLE^{Pab1} variant without its N-terminal α1-helix could interact with its ligand PAM2 in GST pull-down experiments (Figure 13E, lane 2). In humans, α1-helix did not directly make contact with the PAM2 peptide, but it may affect the interaction indirectly through its effect on the α2-helix (KOZLOV *et al.* 2004).

Generally, MLLE domains have 100% sequence conservation in many regions. The longest stretch is KIT**G**MLLE residues at the core of the hydrophobic peptide binding pocket (KOZLOV *et al.* 2001). In *U. maydis* MLLE3^{Rrm4}, K**L**T**I**H**L**L**D** has replaced this stretch (Figure 6B, 14C). Notably, Ile756 at this stretch is a critical difference compared to the small glycine residue in human MLLE^{PABPC1} or MLLE^{Ubr5}. Ile756 makes crucial

hydrophobic interaction with Phe248, Tyr250, Pro251 of PAM2L1^{Upa1} and Phe955, Tyr957, Pro958 of PAM2L2^{Upa1} at 3.0 Å distance. When this isoleucine is replaced with a glycine to mimic the human MLLE domain, the lack of sidechain should have increased the distance between these residues to approximately 6 Å, which is too far to make sidechain contacts and failed to show binding as expected (Figure 15B, C, D Lane 4).

In humans, PAM2 motifs consist of 18 AA residues (Fig.S3A; XIE *et al.* 2014), and Leucine (n3) and Phenylalanine (n10) are the two key residues from PAM2 that mediates the hydrophobic interactions with MLLE domain. Existing MLLE-PAM2 structures in PDB contain 15-18 residues bound to the MLLE domains. In contrast, only the last eight residues in the C-terminal of PAM2L1,2^{Upa1} motifs were bound in the crystal structures of MLLE3^{Rrm4}. Biochemical analysis demonstrated that these eight residues are sufficient for the interaction. Similar to PAM2 motifs, both PAM2L1,2^{Upa1} motifs also have a conserved bulky Phenylanine residue (Phe248 in PAM2L1^{Upa1}, Phe955 in PAM2L2^{Upa1}) crucial for MLLE3^{Rrm4} interaction. However, Tyrosine is the second crucial residue in PAM2L1,2^{Upa1} for hydrophobic interactions (Tyr250 in PAM2L1^{Upa1}, Tyr957 in PAM2L2^{Upa1}) that is inserted into the hydrophobic pocket formed by the helices α2-3. A similar mode of PAM2 interaction is observed in MLLE^{PABPC1}-PAM2w^{GW182} in which the second key residue Tryptophan is inserted into the hydrophobic pocket formed by the helices α2-3 of MLLE^{PABPC1} (Figure S13D, PDB ID: 3KTP, XIE *et al.* 2014) and reported as non-canonical MLLE-PAM2 interaction.

Both PAM2L1,2^{Upa1} motifs have identical key residues (Phenylalanine and Tyrosine) inserted into the MLLE3^{Rrm4} hydrophobic pocket, preceded with similarly charged residues (Aspartate or Glutamate) is a very important outcome from this structural study. In essence, this structural study identified that MLLE3^{Rrm4}-PAM2L1,2^{Upa1} interaction is non-canonical, and crucial residues in PAM2L^{Upa1} peptide is (D/E)(D/E)**F**x**Y** in contrast to the (**L**/P/F)x(P/V)xAxx(**F**/W)xP in PAM2 (XIE *et al.* 2014).

5.7 MLLE3^{Rrm4} seven helix structure is conserved in basidiomycetes

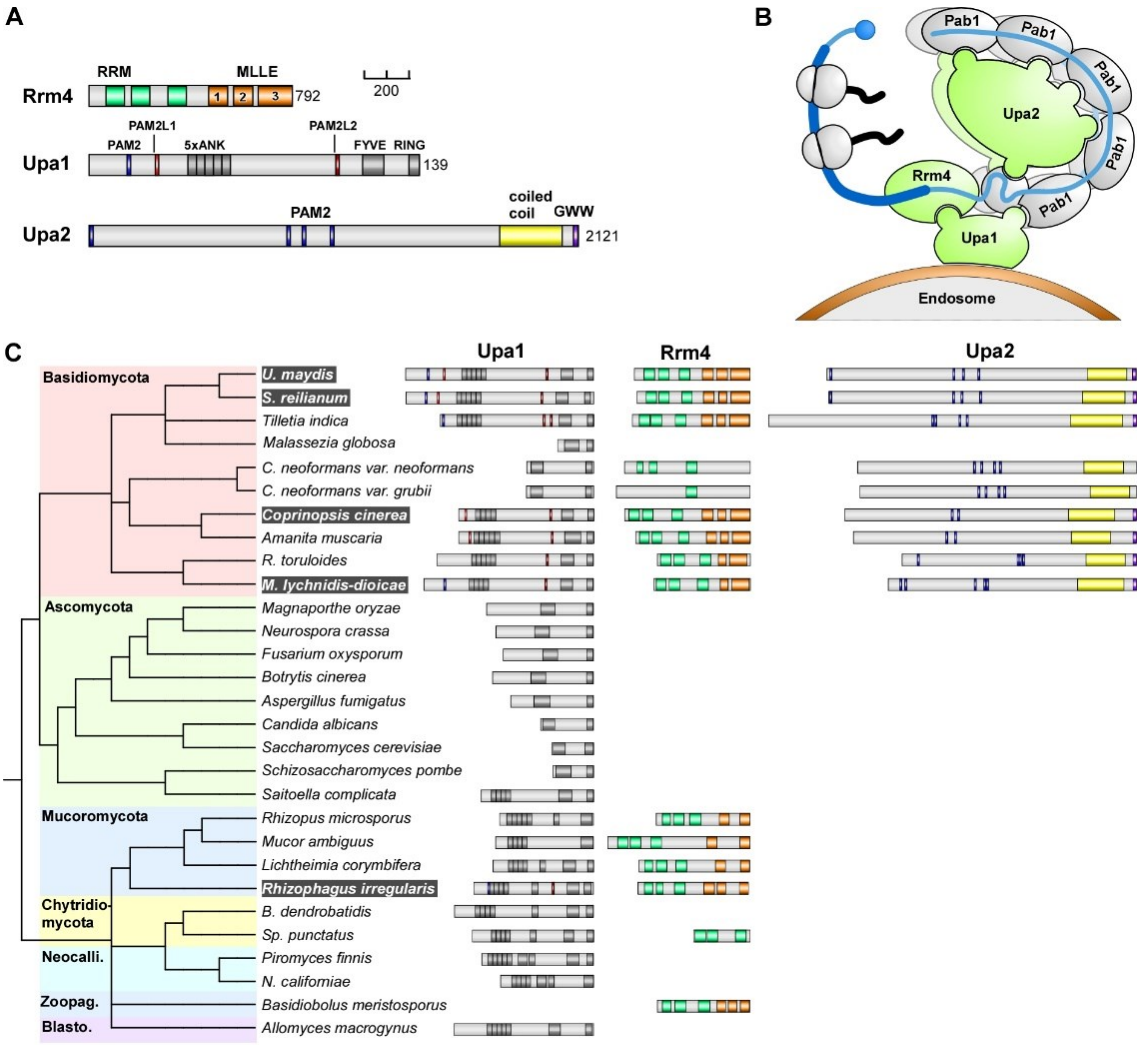
This study established that the MLLE3 domain of Rrm4 consisted of seven helices (MLLE3-7H^{Rrm4}) and not a five alpha-helix structure (MLLE-5H^{Rrm4}) using X-ray crystallography, contrary to previous knowledge.

Previously, 5-helix version of the MLLE3 domain of Rrm4 (MLLE3-5H^{Rrm4}) was deleted, and this truncated version of Rrm4 did not shuttle on endosomes, and exhibited a loss-of-function phenotype similar to rrm4Δ strains (Figure 9B-D; BECHT *et al.* 2006).

However, in the previous *in vivo* studies, MLLE3-5H^{Rrm4}-Gfp containing only five helices did not shuttle on endosomes (data not shown). The current study demonstrated that MLLE3-5H^{Rrm4} is insufficient for PAM2L1,2^{Upa1} interactions, whereas MLLE3-7H^{Rrm4} is sufficient for the PAM2L1,2^{Upa1} interaction *in vitro.* In addition, an N-terminally truncated version of Rrm4 consisted of only MLLE3-7H^{Rrm4}-kat fusion protein moved bi-directionally on endosomes *in vivo* in the recent follow-up study (data not shown).
Multiple sequence alignment (MSA) and structure comparison of AlphaFold2 predicted structures of Rrm4 orthologues showed that MLLE3-7H^{Rrm4} is conserved in Basidiomycetes but not in Rrm4 orthologues of other fungi (Figure 20C, MULLER *et al.* 2019). Interestingly, Rrm4 orthologue from the distantly related Mucoromycota *R. irregularis* shuttled bi-directionally throughout *U. maydis* hyphae with velocities resembling the microtubule-dependent movement of transport endosomes. Core components of the endosomal mRNA transport machinery, including Rrm4, Upa1, and Upa2, are conserved in fungi (MULLER *et al.* 2019). Therefore, hypothetically *R. irregularis* Rrm4 should be interacting with the Upa1 via its PAM2L motif in *U. maydis*. However, Rrm4 of *R. irregularis* has only five helices in its MLLE3^{Rrm4}-type domain and does not possess the N-terminal α I, II helices. Despite this observation, how it interacts with the PAM2L sequences of Upa1 in *U. maydis* for endosomal shuttling is an intriguing question.

The fact that MLLE3^{Rrm4} with even 4 or 5 helices interacted with full-length *E. coli* EfTu *in vitro* (Figure 18B Lane 5, 6) indicates that the MLLE3^{Rrm4} domain is capable of interacting with other proteins even in the absence of the two short helices (α I, II). Therefore, the requirement of the N-terminal helices might be varying case to case. Another point to note is that although Upa1 of *Tilletia indica* (Basidiomycota) has the non-conserved key hydrophobic residues in its PAM2L motif (DIDDGP**LRTP**TRL) it has been considered as a valid PAM2L motif and annotated (MULLER *et al.* 2019). Similarly, Upa1 of *R. irregularis* (Mucoromycota) has a potential PAM2L motif with non-conserved key hydrophobic residues (TAPTSPSFKDV**PITP**TIS; Figure 20E, TiPAM2L1, RiPAM2L), which could be an interaction partner of MLLE3 of *Ri*Rrm4. However, this may have been overlooked and not annotated as a PAM2L motif in the past (MULLER *et al.* 2019).

Interestingly the conserved key Tyrosine residue in the PAM2L^{Upa1} of *U. maydis* is replaced with Threonine in the PAM2L^{Upa1} in *T. indica* and *R. irregularis* (Figure 20E). Both these residues contain hydroxyl group in their sidechain which is a hotspot for


A
RRM
MLLE
Rrm4
1
2
3
792
200
PAM2L1
PAM2L2
PAM2
5xANK
FYVE
RING
Upa1
139
coiled coil
GWW
PAM2
Upa2
2121
B
Pab1
Pab1
Pab1
Pab1
Upa2
Rrm4
Upa1
Endosome
C
Upa1
Rrm4
Upa2
Basidiomycota
U. maydis
S. reilianum
Tilletia indica
Malassezia globosa
C. neoformans var. neoformans
C. neoformans var. grubii
Coprinopsis cinerea
Amanita muscaria
R. toruloides
M. lychnidis-dioicae
Ascomycota
Magnaporthe oryzae
Neurospora crassa
Fusarium oxysporum
Botrytis cinerea
Aspergillus fumigatus
Candida albicans
Saccharomyces cerevisiae
Schizosaccharomyces pombe
Saitoella complicata
Mucoromycota
Rhizopus microsporus
Mucor ambiguus
Lichtheimia corymbifera
Rhizophagus irregularis
Chytridio-mycota
B. dendrobatidis
Sp. punctatus
Neocalli.
Piromyces finnis
N. californiae
Zoopag.
Basidiobolus meristosporus
Blasto.
Allomyces macrogynus
D
α1
α4
αII
α2
α3
αI
α5
Basidiomycota
α1
α2
α4
α3
α5
Other fungi
E
PAM2L
FxY
AmPAM2L1 DEEEEEFIYPSVR
AmPAM2L2 DSSDEEFHYPGAS
CcPAM2L1 EDDEDVFVYPGNT
CcPAM2L2 DDDDDEFVYPGAE
MlPAM2L1 DPTTSSFVLPILL
MlPAM2L2 EESDDEFVYDGQD
RtPAM2L2 DDSDDEFVYSGRD
SrPAM2L1 SDEEDNFVYPATT
SrPAM2L2 DDDDDDFIYPNSY
TiPAM2L1 DIDDGPLRTPTRL
TiPAM2L2 ADDDEEFIYPNSY
UmPAM2L1 EEDQDDFVYPGAD
UmPAM2L2 DDDDDEFIYPNSY
RiPAM2L PSFKDVPITPTIS
F
UmMLLE3Rrm4 - UmPAM2L2Upa1
RiMLLE3Rrm4 - UmPAM2L2Upa1
G
α1
α2
α4
αII
α3
αI
α5

5.7.1.1 Figure 20: Components of endosomal transport are conserved in distinct fungal phyla.

(A) Schematic representation of core components of endosomal mRNA transport: the MLLE domain-containing protein Rrm4 as well as the PAM2 motif-containing proteins Upa1 and Upa2 drawn to scale (bar, 200 amino acids; green, RRM domain; blue, MLLE domain; orange, PAM2 motif; dark blue, PAM2-like motif; dark grey, Ankyrin repeats; light blue, FYVE domain; blue, RING domain; yellow, coiled-coil region; red, GWW). Some orthologues of Upa1 contain a DysFN and DysFC domain indicated in rose and lilac, respectively. (B) Schematic representation of the core components (green) of endosomal mRNA transport. (C) Schematic representation of orthologues of Upa1, Rrm4, and Upa2 in representative organisms of different fungal phyla (Neocalli., Neocallimastigomycota; Zoopag, Zoopagomycota; Blasto, Blastocladiomycota. The phylogenetic tree was generated by applying the phyloT tool using the NCBI taxonomy (no evolutionary distances). If no protein is depicted, a clear orthologue cannot be identified. Accession numbers are listed in (MULLER *et al.* 2019), Supplemental Tables S5. *S., Sporisorium; C., Cryptococcus; R., Rhodotorula; M., Microbotryum; B., Basidiobolus; Sp., Spizellomyces; N., Neocallimastix*. Fungal species studied in this publication are highlighted. (D) Comparison of AlphaFold2 predicted structures of MLLE3^{Rrm4}-type domain consisted of 7 helices from Rrm4 homologs in Basidiomycetes fungi (left), consisted of 5 helices in other fungi (right), represented as cartoon models. AlphaFold2 structures were obtained from Uniprot using the respective protein's accession numbers from the list as stated above, and alignment was done in Pymol. (E) Comparison of PAM2L sequences found in Upa1-like proteins in fungi. Conserved key residues for MLLE3^{Rrm4} interactions are shaded in the black (F) Crystal structure of PAM2L2^{Upa1}-MLLE3^{Rrm4}. The PAM2L2^{Upa1} peptides are inserted into the hydrophobic pocket formed by the helices α2,3 of MLLE3^{Rrm4} in *U. maydis* (*left*). Overlaid structure of MLLE3^{Rrm4} in *R. irregularis* with the PAM2L2^{Upa1}-MLLE3^{Rrm4} structure showed that the hydrophobic pocket of MLLE3^{Rrm4} domain of *R. irrugularis* has sufficient space for PAM2L2 peptide recognition (*right*). Models of PAM2L2^{Upa1} are represented as cartoons and MLLE3^{Rrm4} as the surface in the following colors. PAM2L2 peptide ruby red, MLLE3^{Rrm4} domain dark grey. (G) Alignment of MLLE3^{Rrm4} domain from *R. irregularis* and *U. maydis* (Figure 20A-C adapted from (MULLER *et al.* 2019), under the terms of CCC Order No. 5430730645089).

phosphorylation by kinases. Thereby, PTM of these residues might be regulating the interaction of PAM2L1,2^{Upa1} with MLL3^{Rrm4} *in vivo*.

Structural comparison of the AlphaFold2 predicted RiMLLE3^{Rrm4} domain with the X-ray structure of *Um*MLLE3^{Rrm4} revealed some clues on how *Ri*MLLE3^{Rrm4} with five helices might interact with PAM2L motifs of Upa1 whereas *Um*MLLE3^{Rrm4} is not able to do so. First, the peptide binding pocket in *Ri*MLLE3^{Rrm4} looks perfect fit for recognizing the PAM2L^{Upa1}-containing proteins (Figure 20F, *Ri*MLLE3^{Rrm4}).

Although i*n vitro* binding studies showed that *Um*MLLE3^{Rrm4} seven helix structure is essential for interacting with PAM2L1,2^{Upa1} peptides, the structural analysis showed that the αI, II helices did not participate in the peptide binding. However, they are making intramolecular interaction with helix-5 and pushing it towards helices 3 and 2 to keep the MLLE3^{Rrm4} domain tightly packed to bind the PAM2L1,2^{Upa1} (Figure 20G). Compared with *Um*MLLE3^{Rrm4}, the fifth helix of the RiMLLE3^{Rrm4} is slightly deviated by pushing the helix3 towards helix 2 (Figure 20G).

Therefore, it is possible that even in the absence of the α I, II helices, RiM-LLE3^{Rrm4} might be forming a tightly packed peptide binding pocket that could recognize the PAM2L1,2^{Upa1} in *U. maydis* hyphae which results in endosomal shuttling.

5.8 Rrm4 and Pab1 interacts with each other directly

Previously, biochemical studies with *in vitro* pull-down experiments had been used to demonstrate the MLLE-PAM2 interaction (KOZLOV *et al.* 2004; JINEK *et al.* 2010; MUNOZ-ESCOBAR *et al.* 2015). In this study, *in silico* sequence analysis and biochemical analysis demonstrated that overlapping triplex PAM2 sequences (3x PAM2^{Rrm4}) are present in the linker region connecting the second and third RRM domains of Rrm4. This triplex 3x PAM2^{Rrm4} interacted with MLLE^{Pab1}. It is a preliminary evidence for the direct interaction between the Pab1 and Rrm4. Previous studies have shown that the MLLE domain of Pab1 interacted with the PAM2 motif in Upa1 and Upa2. However, PAM2 motif(s) in the Upa1 and Upa2 were dispensable for the protein function as the mutations in the respective PAM2 motifs did not affect the endosomal shuttling of Pab1 which could be due to the redundancy (POHLMANN *et al.* 2015; JANKOWSKI *et al.* 2019). MLLE^{Pab1} is connected to the endosome via the PAM2 motifs of Upa1 and Upa2. When PAM2^{Upa2} motifs are mutated, MLLE^{Pab1} should be connected to the PAM2^{Upa1}. When the PAM2^{Upa1} motif is mutated, MLLE^{Pab1} should be connected via PAM2^{Upa2}. *In vitro* 3x PAM2^{Rrm4} results suggest that even when both PAM2^{Upa1} and PAM2^{Upa2} motifs are mutated, MLLE^{Pab1} could still be connected to the endosomes via an indirect route by interacting with the triplex PAM2 motif of Rrm4, independent of the pol(A)-binding.

In humans, overlapping PAM2 of eRF3 showed very high affinity to the MLLE^{PABPC1}, preventing the interaction of PABPC1 with deadenylase complexes Caf1–Ccr4 and Pan2–Pan3 by competing for MLLE^{PABPC1} (OSAWA *et al.* 2012; MATTIJSSEN *et al.* 2021). Overlapping PAM2 in eRF3a-C did not have a conserved Leucine residue for hydrophobic interaction; instead, its second Phenylalanine residue (Phe76) bends back to partly occupy the space vacated by the missing leucine side chain for MLLE^{PABPC1} interaction (KOZLOV AND GEHRING 2010). Similarly, Rrm4 triplex PAM2 sequences, should possess a very high affinity towards MLLE^{Pab1}, and the second or third Phenylalanine residues of 3xPAM2 should be complementing the missing Leucine residue for MLLE^{Pab1} interaction (Figure 17A). Due to its expected high affinity, it should compete with the PAM2 motifs of Upa1 and Upa2 or other factors to prevent their interaction, suggesting

a potential mechanism of 1) Pab1 association with Rrm4 independent of endosomal hitchhiking, 2) Rrm4 controlling Pab1's interaction with Upa1 and Upa2 for loading and unloading on endosomes.

In addition, $MLLE^{Pab1}$ directly interacted with the C-terminal of Rrm4, containing MLLE domains (Figure 17B, lane 7, G-Pab1-M), indicating heteromeric MLLE interactions mediating a direct interaction between Rrm4 and Pab1. Rrm4 variants expressed with C-terminal regions containing three $MLLE^{Rrm4}$ domains interacted with its kind (Figure 17B, lane 7, G-Rrm4-NT4). MLLE domains are involved in homo oligomerization and form regularly spaced, multimeric complexes of PABPC1 on the poly(A) tail of mRNA (KUHN AND PIELER 1996). It is conceivable that Rrm4 and Pab1 are tightly linked to each other despite their association with the Upa1 and Upa2. This observation points out that Rrm4 and Pab1 interaction is not necessarily limited to endosomes but may exist in another part of the cell or organelles. A tight interaction between the two mRNA-binding proteins Rrm4 and Pab1 might be necessary to stabilize the whole mRNP complex on a dynamic endosomal compartment.

5.9 $MLLE3^{Rrm4}$ interacts with EfTu

Studying an unknown co-eluted protein band in the *in vitro* pull-down assay identified that *E. coli,* EfTu (*tufB*), interacted with $MLLE3^{Rrm4}$. EfTu is an ancient molecule, universally conserved GTPase, and probably the most abundant protein in many bacterial species , (CALDON AND MARCH 2003; HUGHES 2017). Higher eukaryotes sense microbes through the perception of pathogen-associated molecular patterns (PAMPs). EfTu is used as a PAMP by the host to recognize the pathogens due to its abundance (ZIPFEL *et al.* 2006). EfTu can repetitively engage aa-tRNA within the ribosome during the proofreading stage of tRNA selection (MORSE *et al.* 2020). During translation elongation, in its active form, GTP-bound EfTu transports aminoacyl tRNA (aa-tRNA) to the ribosome A-site by forming a ternary complex EfTu plays a critical role in mRNA decoding by increasing the rate and fidelity of aa-tRNA selection at each mRNA codon.

Eukaryotic cells have two translation systems, one in the cytoplasm and the other in the mitochondria. The eukaryotic equivalent of EfTu is elongation factor 1 alpha (EF-1α). Bacterial and eukaryotic EfTu differ in how they recharge the EfTu-GDP complex. This recharging function is performed by the Elongation Factor Thermo stable (EfTs) in prokaryotes and mitochondria, whereas by eukaryotic Elongation Factor 1B (eEF1B) in eukaryotes (HARVEY *et al.* 2019). The mitochondrial equivalent of EfTu is mt-EfTu (*TufM*).

Mitochondrial elongation factors are encoded in the nuclear genome with a mitochondrial targeting signal (MTS) for mitochondrial import. Mitochondrial elongation factors are distinct from their cytosolic counterparts and are more similar to those of bacteria (HUGHES 2017). In alignment with this, mitochondrial EfTu of *U. maydis* has a higher sequence identity (*Um*EfTuM, UMAG_00138, 62%) with *E. coli* EfTu than the cytoplasmic counterpart (*Um*EfTu-1α, UMAG_00924, 30%).

Generally, EfTu structures consist of three major domains (Figure 18A, D,E GDP-binding domain, II, and III Domain), which have evolved a high degree of molecular flexibility. Domain I forms a helix structure with α/β Rossmann fold topology, while domains II and III are largely comprised of β-barrel structures (SONG *et al.* 1999; ANDERSEN *et al.* 2000). Structural comparison of *E. coli* EfTu with *Um*EfTuM and *Um*EfTu-1α revealed that all the three EfTu structures are conserved and have the said domains. However, the switch 1 region in *E. coli* consists of an α-helix followed by a β-hairpin (Figure. 18D). In contrast, in *U. maydis*, both EfTu versions contain only two helices in the Switch 1 region (Figure. 18E), a similar switch 1 consisting of only helices is observed in human mt-EfTu AlphaFold2 predicted structure.

A recent study exhibited that during the conformational change from GTP bound to GDP bound state, Switch 1 of EfTu rapidly converted from an α-helix into a β-hairpin and moved to interact with the acceptor stem of the aa-tRNA. Thereby, switch 1 acted as a gate to control the movement of the aa-tRNA during accommodation through steric interactions with the acceptor stem (GIRODAT *et al.* 2020). The EfTu structure of *E. coli* is experimentally determined with a bound GDP (Figure 18D). In contrast, the UmEfTuM and UmEfTu-1α structures of *U. maydis* or human EfTuM structures are predicted by AlphaFold2, which lacks the GDP (Figure 18E, PDB ID: 1EFC). This observation explains the variations in the switch 1 region of *E. coli*. In addition, it serves as a proof of principle example that the switch region present in the low complexity structure of a protein can undergo conformational changes to modify the protein function or interaction.

EfTu Domain I is connected to domain II by a 16 Å-long disordered-like linker region, with a single helical turn in the middle (SONG *et al.* 1999). In order to perform its canonical function, the linker region has to undergo a conformational change and turn the GDP domain 90° to align with domains II and III (HARVEY *et al.* 2019). The potential PAM2L of the EfTu is present in this linker region connecting the GDP domain and domain II (Figure 18D, E); therefore, it is likely that the PAM2L region should have a dynamic structure and be involved in the modification of the function.

PAM2L motifs of EfTu are conserved from *E. coli* to higher eukaryotes (Figure S14C). PAM2L region in *E. coli* EfTu has KP**F**L**L**P, *U. maydis* PAM2L^{EfTuM} has KP**F**L**M**P, and *U. maydis* PAM2L$^{EfTu\text{-}1\alpha}$ has KP**L**R**L**P as their respective hydrophobic key residues required for the MLLE3^{Rrm4} interaction (Figure 18A). In Upa1-PAM2L motifs 1,2, these are DD**F**V**Y**P, and DE**F**I**Y**P, respectively. The Phenylalanine in the PAM2L^{EfTuM} is replaced with the Leucine in PAM2L$^{EfTu\text{-}1\alpha}$. Since both Phenylalanine and Leucine are hydrophobic residues, they should be able to interact with the hydrophobic binding pocket of MLLE3^{Rrm4}. It is supported by the fact that the canonical Leucine required for MLLE^{PABPC1} interaction is substituted by Phenylalanine in the overlapping duplex PAM2 motif of eRF3-C and is capable of making hydrophobic interactions, (KOZLOV AND GEHRING 2010; OSAWA *et al.* 2012; MATTIJSSEN *et al.* 2021). On the other hand, the essential key residue Tyrosine in the FxY pocket has been replaced by Leucine in *E. coli* PAM2L^{EfTu} and *U. maydis* PAM2L$^{EfTu\text{-}1\alpha}$, whereas in the *U. maydis* PAM2L^{EfTuM} it is replaced by the Methionine. Since all these are hydrophobic residues, they should be capable of making hydrophobic contacts with MLLE3^{Rrm4}. PAM2L1^{Upa1} and PAM2L2^{Upa1} have minor sequence variations in their key residues, which results in differential binding affinities (Figure 8A, B) and regulating their function (Figure 10, S8, S9). Therefore, although the overall PAM2L pattern is conserved in elongation factors of *U. maydis* (PAM2L^{EfTuM}, PAM2L$^{EfTu\text{-}1\alpha}$), the differences in the key residues should be dictating their binding affinity and function.

A recombinant variant of *E. coli* PAM2L^{EfTu} interacted only with MLLE3^{Rrm4} but not with any other variants, but the full-length EfTu interacted with all the Rrm4 variants. With the current results from this study, this observation cannot be completely explained. However, a partial explanation is that the full-length EfTu has more contact sites to the Rrm4 C-terminal region than the short PAM2L variant. Therefore, it could interact with shorter and longer versions of the MLLE3^{Rrm4} variants. The short PAM2L variant lacking the extended contact sites can only interact with the defined seven helix MLLE3^{Rrm4}. However, further investigation is required to decipher the mechanism completely.

Interestingly, earlier studies have demonstrated that *Um*EfTuM was enriched in GFPTrap, MS/LC-MS experiments performed using the Rrm4-Gfp version in *U. maydis* (TULINSKI 2021 unpublished data). EfTuM is a target of Rrm4 in iCLIP experiments, and Rrm4 is bound to *EfTuM* mRNA at the stop codon (STOFFEL 2022 unpublished data). The

higher sequence identity, and conserved structures, combined with *in vitro* pull-down experiments using *E. coli* EfTu and *U. maydis* GFPTrap experiments, support that *Um*EfTuM is a potential, direct binding partner of Rrm4.

Mitochondrial EfTu transports aa-tRNA to the A site of mitochondrial ribosomes by forming a ternary complex. In addition, mt-EfTu has been reported to have several regulatory functions, including aa-tRNA surveillance in mammalian mitochondria, chaperone activity, preventing thermal aggregation of proteins, and enhancing protein refolding *in vitro*. In humans, mutations in EfTuM are associated with brain and neuronal disorders. For example, an infant with a mutation in EfTuM (Gly322Arg) showed persistent axial hypotonia and limb spasticity at 6 months and did not survive beyond 10 months. Glycine 322 residue is highly conserved from yeast to humans. The equivalent of a human EfTuM G322R mutation in yeast EfTuM is G311R; it impaired mitochondrial respiration (DI NOTTIA *et al.* 2017). This glycine residue is highly conserved from *E. coli* to fungi and humans. In *U. maydis*, this is Gly343; mutation of this residue might impair mitochondrial respiration, similar to yeast.

Nuclear encoded mitochondrial mRNAs are targets of Rrm4, and expression of mitochondrial proteins is altered in the absence of Rrm4 in *U. maydis* hyphae (KOEPKE *et al.* 2011; OLGEISER *et al.* 2019). Translationally active ribosomes are co-localized with Rrm4 on endosomes in an mRNA-dependent manner (BAUMANN *et al.* 2014). A close link between Rrm4-mediated mRNA transport and local mitochondrial translation has been proposed in *U. maydis* (MUNTJES *et al.* 2021). 99% of the mitochondrial proteins are encoded by the nuclear genome, synthesized as precursor proteins in the cytosol, and translocated in the unfolded form to the mitochondria by mitochondria protein import machinery (WIEDEMANN AND PFANNER 2017). EfTu is synthesized in the cytosol and imported as unfolded linear protein through the OMM and IMM translocases TOM40 complex and TIM23 complex, respectively. MTS is cleaved by the matrix's mitochondria processing peptidase (MPP) (CHOI *et al.* 2022). Therefore, two possible scenarios exist for the Rrm4 and EfTuM interaction. In the first case, RRM domains of Rrm4 interact with the EfTu mRNA during the translation-coupled mRNA transport on the endosomal surface where newly synthesized precursor EfTuM, in its unfolded state, might be interacting with the MLLE3 domain of Rrm4.

EfTuM is a target of Ubiquitin specific peptidase 5 (USP5), a ubiquitously expressed deubiquitinating enzyme, which regulated and stabilized EfTuM level through deubiquitination in colorectal cancer (XU *et al.* 2019). In addition to translocation via the import

pathway, a fraction of the pre-matured protein with MTS is localized to the cytoplasmic surface of OMM by an unidentified process. OMM-localized EfTuM is subject to ubiquitination and proteasomal degradation (CHOI *et al.* 2022).

Structural studies have established that the MLLE2^{Rrm4} domain resembles the human MLLE^{Ubr5}, a HECT domain containing E3 ubiquitinase (DEVAN *et al.* 2022). iCLIP experiments combined with RNA live imaging has demonstrated that Rrm4 transports *ubi1* mRNA on endosomes encoding a fusion protein of ubiquitin and the ribosomal protein Rpl40 (OLGEISER *et al.* 2019). Rrm4's role in *ubi1* mRNA transport is yet unknown. Rrm4 is tethered to the endosome via the adaptor protein Upa1, which has the RING domain in its C-terminal (POHLMANN *et al.* 2015). RING and HECT are E3 ubiquitin ligases specifically targeting their substrate proteins for ubiquitin-mediated proteolysis by adding the 76-residue ubiquitin moiety (VARSHAVSKY 2012; SCHEFFNER AND KUMAR 2014). The above facts indicate a potential link between Rrm4 and Ubiquitination. Therefore, as a second scenario, Rrm4 may interact with the EfTu on the OMM surface and mediate the ubiquitination and proteasome degradation.

5.10 Novel interaction partners of MLLE3 domain of Rrm4

By using the preliminary consensus sequence `(D/E)(D/E)FxY` of PAM2L^{Upa1} required for the MLLE3^{Rrm4} binding, identified from the crystallization studies as a key, reiterative BLAST search combined with stringent selection criteria identified about twenty new potential PAM2L containing proteins in *U. maydis* from the NCBI non-redundant sequence database. A similar approach has identified several PAM2-containing interaction partners of MLLE^{PABPC1} in humans, plants, and fungi. The potential hits were further screened based on two criteria, first the sequence conservation, second the accessibility of the PAM2L motif as judged from the presence of surrounding low complexity sequence (KOZLOV *et al.* 2001a; ALBRECHT AND LENGAUER 2004; BRAVO *et al.* 2005; HUANG *et al.* 2013; JIMENEZ-LOPEZ *et al.* 2015; POHLMANN *et al.* 2015). In addition, the current study used the absence of PAM2L hits or sequence conservation in the fungal species that do not have MLLE domains in basidiomycetes as the third criterion to shortlist the potential PAM2L containing Rrm4 interaction partners. Because the previous report has revealed that *M. globosa* does not have Rrm4, and *C. neoformans* do not have MLLE domain in the Rrm4 orthologue (MULLER *et al.* 2019).

In addition to the known interaction partner Upa1, this strategy identified several putative RBPs, for example, Taf7, a nucleocytoplasmic mRNA transporter

(UMAG_10620), RNA-directed RNA polymerase (UMAG_02933), DNA-directed RNA polymerase (UMAG_04460), Brix domain-containing RBP (UMAG_03844), CWF19-RNA lariat debranching enzyme (UMAG_11924), Transcriptional activator HAP2 (UMAG_01597), Transcription factor CBF (UMAG_02262), Transcription corepressor (UMAG_10255), FCP1 - RNA polymerase II phosphatase (UMAG_04936). Other candidates include Vps8 – an early endosome-specific subunit of CORVET complex (UMAG_15064), a ubiquitination-related RING domain and Ankyrin repeats containing protein (UMAG_10591), Serine carboxypeptidase (UMAG_10091), Endo/exonuclease-phosphatase (UMAG_03381), GH16 - glycosyl hydrolases (UMAG_05811), Methyl-itaconate delta2-delta3-isomerase (UMAG_02807), DUF3835- Calsequestrin and nucleolin related protein (UMAG_04517), mitochondrial 5-demethoxyubiquinone hydroxylase (UMAG_15085), Methyl-accepting transducer (UMAG_00448), alpha-mannosidase (UMAG_00557), Laccase I (UMAG_05361), Cytochrome b5 heme-binding protein (UMAG_01466). The complete list with more information is provided in Table S18.

Most studies have shown that PAM2 motif-containing proteins are associated with different mRNA-related processes by interacting with MLLE^{PABPC1} (XIE *et al.* 2014). As Rrm4 is an mRNA binding protein, similar to PABPC1, it is natural that it interacts with more RBPs. Therefore, most of the *de novo* predicted candidates are RBP, which is not just a coincidence but a strong indication of the interrelation. The interaction of PAM2-containing proteins with MLLE^{Ubr5} is less understood and an area for further investigation (XIE *et al.* 2014). As stated in the previous section, the MLLE2^{Rrm4} domain of Rrm4 resembles the MLLE^{Ubr5} and transports *Ubi1* mRNA on endosome, suggesting a role of Rrm4 in the ubiquitination (Figure 7A, KONIG *et al.* 2009; DEVAN *et al.* 2022). Rrm4 interacted with the endosomal adaptor protein Upa1 via the PAM2L1,2 motifs, which, in addition, has Ankyrin repeats and RING domain-containing protein (POHLMANN *et al.* 2015). A similar candidate in the predicted list is UMAG_10591, which has a PAM2L motif, ankyrin repeats, and a RING domain. Considering its well-conserved PAM2L motif, it could be a potential interaction partner of Rrm4 and an additional factor that points out the potential link between Rrm4 and ubiquitination.

All the newly identified PAM2L candidates have the consensus `(D/E)(D/E)FxY` sequences, two negatively charged residues followed by the hydrophobic key residues for MLLE3^{Rrm4} interaction. *In vitro*, pull-down assays have demonstrated that the representative PAM2L motifs from Vps8 and Taf7 interacted with MLLE3^{Rrm4}. PAM2L1 motif of

Vps8 (PAM2L1^{Vps8}- HQDDDNDNDDD**F**V**Y**DGID) and Taf7 (PAM2L^{Taf7} - GGGGKGF-NIDD**F**I**Y**PHGI) showed strong binding to MLLE3^{Rrm4} whereas PAM2L2 motif of Vps8 (PAM2L2^{Vps8}- TDQAPSDDSFS**F**R**Y**PHPL) showed week binding to MLLE3^{Rrm4} and one explanation for this observation is that the later lacks the negatively charged residues immediately preceding the key hydrophobic residues FxY. Co-crytallized MLLE3^{Rrm4} - PAM2L1,2^{Upa1} structures have shown that these two negatively charged residues adjacent to the FxY residues are making important contributions and stabilizing the complex (Figure 15A-E). Albeit a lower affinity, PAM2L2^{Vps8} is a proof of principle example that a peptide motif containing just an FxY motif can interact with the MLLE3^{Rrm4}. More details on Vps8 and Taf7 are given in the following sections.

5.10.1 Taf7

Eukaryotic transcription of protein-coding genes involves the formation of a multiprotein complex containing the RNA polymerase II core enzyme and a set of accessory factors (TFIIA, TFIIB, TFIID, TFIIE, TFIIF, and TFIIH). The transcription factor TFIID by itself is a multiprotein complex comprising the TATA-binding protein (TBP) and TBP-associated factors (TAFs) (LAVIGNE *et al.* 1996). Taf 7, encoded by *the Ptr6* gene, is a TAF and subunit 7 of the TFIID complex in *S. pombe*, an essential gene previously known as Ptr6. Taf7 orthologues are found in *S. cerevisiae* (yTAFII67) and humans (hTAFII55) with 54% and 36% sequence similarity, respectively. In human Taf7, the N-terminal transcription activator interaction domain and central TAFII250 interaction domain are reported as functional PPI domains and are highly conserved in all three organisms Human Taf7 (hTAFII55) interacts with TAFII250, TAFII100, TAFII28, TAFII20, and TAFII18, but not with TAFII30 or TBP. Although human Taf7 did not interact by itself with TBP, stable ternary complexes containing Taf7 and TBP can be formed in the presence of TAFII250, TAFII100, or TAFII28 (LAVIGNE *et al.* 1996).

Taf7 in *S. pombe* and *S. cerevisiae* contains a stretch of acidic amino acids (aspartic acid and glutamic acid) in the carboxy-terminal region, which is lacking in human Taf7. The acidic stretches of Taf7 in *S. cerevisiae* are longer than those of *S. pombe*. In addition, the former contains a stretch of basic amino acids (lysine and arginine) in the amino-terminal region, however in *S. pombe* and human this stretch is missing (SHIBUYA *et al.* 1999).

Temperature sensitive mutation in Taf7 (G182E) a conserved residue within the conserved central TAFII250 interaction domain caused accumulation of mRNA in the nucleus and inhibition of growth at the nonpermissive temperature. Taf7 (*Ptr6*) play a role in the regulation of transcription of many genes at the nonpermissive temperature. Taf7-Gfp fusion protein was localized in the nucleus, suggesting that it is mainly a nuclear protein and functions there. Taf7 is proposed to be nucleocytoplasmic transport protein involved in transcription and nuclear mRNA export (SHIBUYA *et al.* 1999).

Taf7 was identified as a PAM2L-containing interaction partner of the MLLE3^{Rrm4} domain of Rrm4 by *de novo* prediction and confirmed by in vitro pull-down experiments. *U. maydis* Taf7 has the conserved activator interaction domain and TAFII250 interaction domain similar to *S. pombe*, *S. cerevisiae*, and humans. In *U. maydis* PAM2L^{Taf7} motif is located adjacent to the C-terminal of the activator interaction domain in a low complexity switch region (Figure 19E, AlphaFold2 prediction)., indicating that it may undergo a conformational change during PPIs. Due to sequence and structural conservation, Taf7 may play a similar role in *U. maydis* as a TAF and a nucleocytoplasmic transport protein. In addition, it may be involved in the endosomal loading of mRNA by interacting with the Rrm4 in the cytoplasm.

MSA alignment showed that the PAM2L region of the Taf7 is highly conserved in basidiomycetes and partially conserved in other fungi (Figure S15A). In *S. pombe* and *S. cerevisiae*, this region contains the residues `ADYIYPHGLT` and `KGYDYKHGIS`. In *U. maydis* PAM2L^{Taf7} motif has the important key residues required for MLLE3^{Rrm4} interaction as `DD`**`F`**`I`**`Y`**`PHGIT`, in which F(Phe) and Y(Tyr) are critical for hydrophobic interactions. Because Tyrosin is a hydrophobic residue with a bulky aromatic side chain, comparable size, and properties of Phenylalanine, it can complement the Phenylalanine and interact with the yeast MLLE domains.

Interestingly unlike other eukaryotes, *S. cerevisiae* has only one MLLE domain-containing protein, that is Pab1p. Although yeast MLLE^{Pab1p} domain structure is solved (PDB ID: 1IFW), and detailed biochemical and genetics studies are available, it is reported as an orphan domain without any known binding partner (KOZLOV *et al.* 2002). A C-terminal portion of yeast Pab1p was used as bait in a yeast two-hybrid (Y2H) screen to identify yeast proteins interacting with MLLE^{Pab1p}. However, none of the proteins identified from this screen contained a consensus PAM2 sequence (MANGUS *et al.* 1998b; KOZLOV *et al.* 2002). *S. cerevisiae* genome was searched for new PAM2 motifs using the previously published consensus sequence of human PAM2 motifs. This search identified several

PAM2 motifs in the known PAM2-containing proteins based on the human MLLE-PAM2 interaction network, for example, PAN1, Pbp1p, and RF3. None of these predicted PAM2 peptides bound to MLLE^{Pab1p} in NMR (KOZLOV *et al.* 2002).

The Yeast MLLE^{Pab1p} domain is more closely related to the human MLLE^{Ubr5} than human MLLE^{PABPC1}, with 57 and 40% identity, respectively (KOZLOV *et al.* 2002). MSA showed that the PAM2L motif in Taf7 is conserved in *S. cerevisiae* (Figure S15A). Therefore, the interaction partners of yeast MLLE^{Pab1p} might be a PAM2L motif rather than PAM2, which may be the reason why the predicted PAM2 sequences in yeast did not interact with the MLLE domain of yeast Pab1p. Although Pab1p is an essential gene, the MLLE^{Pab1p} domain is dispensable for Pab1p functions such as translation initiation and poly(A)-shortening (MANGUS *et al.* 1998). It is possible that the MLLE domain of yeast Pab1p is not involved in the typical Pab1p functions but is involved in ubiquitination similar to human Ubr5 by interacting with the PAM2L-containing proteins in yeast.

5.10.2 Vps8

CORVET and HOPS multi-protein complexes mediate the maturation of early endosome (EE) to late endosome (LEs)/vacuoles. These heterohexameric complexes share four 'class C core' components, Vps11, Vps16, Vps18, and Vps33. Besides the core components, Vps8 and Vps3 are CORVET-specific subunits that bind to the Rab5 GTPase, whereas Vps39 and Vps41 are HOPS-specific subunits that bind to the Rab7 GTPase (BALDERHAAR AND UNGERMANN 2013; LOPEZ-BERGES *et al.* 2017). The CORVET and HOPS complexes are mediating membrane fusion events at the EEs and LEs/vacuoles, respectively, by acting as molecular tethers bringing together acceptor and donor membranes and by regulating the actual fusion step through their regulation of the SNARE (soluble N-ethylmaleimide-sensitive fusion protein attachment protein receptors) machinery (BALDERHAAR AND UNGERMANN 2013; LOPEZ-BERGES *et al.* 2017).

Although they interact with different GTPases or interaction partners, HOPS and CORVET complexes are conserved in eukaryotes from yeast to mammals (PERINI *et al.* 2014). Most HOPS and CORVET subunits are structurally very similar, with a predicted N-terminal β propeller and a C-terminal a-solenoid. Furthermore, yeast Vps8, Vps11, and Vps18, as well as mammalian Vps41, have C-terminal RING domains associated with ubiquitination (BALDERHAAR AND UNGERMANN 2013). CORVET-specific subunits Vps3 and Vps8 interact with Rab5 and require their N-terminal domains for localization and

function. CORVET subunits Vps3 and Vps8 may lack either of their two N-terminal domains, but not both, to promote protein sorting via the endosome. Vps8 N-terminal domain contributes to the interaction with Vps21 (Rab5 GTPase homolog in yeast), which could be important for efficient tethering at endosomal membranes. C-terminal α-solenoid domains of both proteins are necessary for CORVET assembly and function, as mutants lacking these parts behave like deletions of the entire protein (EPP AND UNGERMANN 2013).

In *S. cerevisiae*, loss of Vps8 results in clustered, abnormally large vacuoles, whereas in *A. nidulans*, Vps8 deletion causes severe growth defect phenotype and contains numerous small vacuoles (LOPEZ-BERGES *et al.* 2017). In *Arabidopsis*, the knockdown of Vps8 resulted in abnormal root morphology (TAKEMOTO *et al.* 2018). Genetic analysis of *Drosophila* revealed that depletion of Vps8 leads to fragmentation of endosomes. Vps8 mutants are semi-lethal and develop melanotic tumors (LORINCZ *et al.* 2016). In human cell culture studies, Vps3 and Vps8 interact directly with each other, control integrin trafficking, and regulate cell adhesion and migration (JONKER *et al.* 2018). These observations could be because the EE failed to mature into LE in the absence of Vps8 and emphasized Vps8's role in EE functions.

The tethering- and core-subunits Vps3 and Vps8 of CORVET are essential, buttressing a central role for EE transport in *U. maydis*. Vps3 and Vps8 fused C-terminally to the fluorescent proteins co-localized with Rab5a positive motile EE. Vps3 deletion resulted in cytokinesis defect in yeast-like sporadic cells leading to chains of non-separated cells. Attempts to produce Vps8 deletion did not produce transformants, indicating it is an essential gene (SCHNEIDER *et al.* 2022).

Vps8 was identified as a PAM2L-containing interaction partner of the MLLE3^{Rrm4} domain of Rrm4 by *de novo* prediction and confirmed by *in vitro* pull-down experiments. *U. maydis* Vps8 has the conserved N-terminal β propeller, C-terminal a-solenoid, and a RING domain structure similar to yeast and human versions. Interestingly, *U. maydis* Vps8 has a long IDR at the very N-terminal where the two PAM2L motifs are located (Figure 19E, AlphaFold2 prediction). As Rrm4 shuttles on endosomes in the absence of Upa1, and the MLLE3 domain of Rrm4 interacts with PAM2L1,2 motifs of Vps8, it suggests that Vps8 could be a secondary binding partner for Rrm4 on endosomes.

As PAM2L1,2^{Vps8} motifs interacted with MLLE3^{Rrm4} *in vitro,* they might also interact with the Rrm4 *in vivo*. Therefore, hypothetically Vps8 could be a secondary tethering

molecule of Rrm4 on EE. It is supported by the fact that residual movement of Rrm4 is observed in Upa1 deletion strains (Figure 10C, POHLMANN *et al.* 2015). However, the Vps8 protein was not found in the Rrm4-GFP trap combined with MS analysis (TULINSKI 2021) nor in Ferry-GST Trap (SCHUHMACHER *et al.* 2021). One explanation for this observation is that Rrm4 interaction with Vps8 could be transient.

vps8 mRNA was a target of Rrm4 in iCLIP experiments (STOFFEL 2022 unpublished data), and human Ferry complex (SCHUHMACHER *et al.* 2021) suggesting that *vps8* mRNA may be transported to EE for local translation. Therefore, by local translation, Rrm4 might recruit the *vps8* mRNA to the EE, where the newly synthesized protein interacts with MLLE3^{Rrm4}. In essence, Vps8 is a potential interaction partner of Rrm4 and may act as a secondary tethering molecule on EE, therefore an interesting candidate for further investigation.

Endosomes are multipurpose platforms on which unique sets of molecular machines can assemble and perform different cellular roles (GOULD AND LIPPINCOTT-SCHWARTZ 2009). During endocytosis, cell surface receptors are degraded or recycled at the endosomes. Proteins are marked by ubiquitylation and sorted into intraluminal vesicles (ILVs) for degradation. Ubiquitylated receptors are cleared off from the endosomal surface by ESCRT machinery, thus converting or maturing the early endosome into the multivesicular late endosome or multivesicular body (MVB) (BALDERHAAR AND UNGERMANN 2013). The C-terminal domains of Upa1 are an orthologue of yeast Pib1, which mediates the Ubiquitination of a subset of cellular proteins localized to endosomes (SHIN *et al.* 2001). Pib1 and Vps8 are RING domains possessing E3 ubiquitin ligases on EE in *S. cerevisiae* (MACDONALD *et al.* 2017). Rrm4 transports *ubi1* mRNA on endosomes encoding a fusion protein of ubiquitin and the ribosomal protein Rpl40 (OLGEISER *et al.* 2019). The Association of Rrm4 with Upa1 and Vps8 on EE and its role in *ubi1* mRNA transport in *U. maydis* stipulate further investigations on Rrm4's association in Ubiquitination mediated protein degradation on the endosomal surface in addition to the mRNA transport.

5.11 Novel interaction partners of Pab1

In the past, 14 PAM2-containing proteins were identified in *the U. maydis* genome using the human PAM2 consensus sequences as an input in a reiterative BLAST search, which includes the Upa1 and Upa2 (POHLMANN *et al.* 2015). This study identified twelve new

PAM2-containing proteins using the consensus sequences from the known PAM2 motifs from Upa1 and Upa2. Although these PAM2 motifs are not verified experimentally, these proteins are potential interaction partners for the MLLE domain of Pab1 as they have the conserved Leu (n3) and Phe (n10) residues for hydrophobic interactions.

These new PAM2 proteins lists include RNA binding proteins DEAD-box ATP-dependent RNA helicase DED1 (UMAG_04080), MMS19 nucleotide excision repair protein (UMAG_03624), Serine/threonine-protein kinase MRCK-β (UMAG_03901), Serine/threonine-protein kinase ATM (UMAG_15011), E3 ubiquitin ligase (UMAG_02525), sepA-related formin (UMAG_12254), Diphthamide biosynthesis protein 4 (UMAG_01327), Beta-lactamase domain-containing protein (UMAG_03398), Mediator of RNA polymerase II transcription subunit 19 (UMAG_11335), SWIRM domain-containing protein (UMAG_03636), BTB/POZ domain-containing protein (UMAG_10294), Alpha/beta-hydrolase (UMAG_06478). A complete list with more information is given in Table S19.

As described earlier, PAM2 motifs are embedded adjacent to the IDR's Ser/Thr clusters for phosphorylation. PAM2-MLLE mediated PPI is modulated by phosphorylation at Ser/Thr residues (Huang et al. 2013). Ser-Pro, and Thr-Pro Phosphorylation sites are found in the linker region close to the N-terminal of the MLLE domain of PABPC1 in the protozoan Leishmania, which controls the PPI and function during translation (de Melo Neto et al. 2018). Phosphorylation is one of the most prevalent and important post-translational modifications to proteins (PTMs). It is a reversible process mediated by the protein kinases by adding a phosphate group (PO4) (Ardito et al. 2017). Currently, no MLLE interacting PAM2 containing kinases are reported to the best of my knowledge. Interestingly, this study identified two kinases (UMAG_03901, UMAG_15011) with PAM2 motifs, which suggests a potential link between the MLLE domain and the kinases.

The most poorly understood aspect of MLLE domain function is its role in protein ubiquitination. Besides PABC1, the only other protein that contains MLLE domain in human is the E3 ubiquitin ligase UBR5. USP10, a PAM2 containing interaction partner of $MLLE^{PABPC1}$ is known to stabilize the mRNA and is associated with the stress granules (KOZLOV *et al.* 2010c; TAKAHASHI *et al.* 2013; XIE *et al.* 2014b). Intriguingly this study identified a PAM2 containing E3 ubiquitin ligase (UMAG_02525) and opens up the possibility to investigate the role of the *U. maydis* $MLLE^{Pab1}$ in ubiquitination.

Apart from the sequence based PAM2-motif predictions described above, Annexin A11 (ANXA11), could be an $MLLE^{Pab1}$ interaction partner based on the literature.

ANXA11 is an RNA granule-associated phosphoinositide-binding protein, and a molecular tether between RNA granules and lysosomes mediate the RNA granule transport in human neurons (Figure 2C). ANXA11 has a long N-terminal IDR, facilitating its phase separation into membraneless RNA granules, and a C-terminal phosphoinositide membrane- binding domain, enabling interactions with lysosomes. Amyotrophic lateral sclerosis (ALS)-related mutations in ANXA11 disrupted the interaction with lysosomes and impair RNA granule transport (LIAO *et al.* 2019). ANXA11 orthologue in *U. maydis* has (UMAG_03580) 42.5% sequence similarity and possess a 120 amino acid long, N-terminal IDR and a C-terminal domain comparable to Annexin 11. Structural comparison with AlphaFold2 predicted structures of human and *U. maydis* ANXA11 revealed that both these proteins are having conserved domain architecture (Figure S16).

mRNA interacting proteins are highly enriched in IDRs in the human proteome (Castello et al. 2012). PAM2 motifs are embedded in the IDR and interact with MLLE^{PABPC1} (Huang et al. 2013). Almost all the eukaryotic mRNAs have poly(A)-tail in the 3' end. PABPC1 binds to the poly(A)-tail in the cytoplasm. This interaction is essential for the poly(A)-tail-stabilization by protecting it from the nucleases, important for translation initiation, 60s ribosomal subunit binding and mRNA decay (Mangus et al. 2003b; Kuhn and Wahle 2004). PABPC1 is a component of RNA granules, and proteins containing PAM2 motifs are over-represented in P-bodies and RNA stress granules (KOZLOV *et al.* 2010c). Therefore, hypothetically ANXA11 orthologue in *U. maydis* may form RNA granules on lysosomes similar to humans. Endosome-mediated mRNA transport is a conserved phenomenon in eukaryotes (Muntjes et al. 2021), similarly, membraneless RNA granules transport on lysosomes could be a conserved phenomenon too. Furthermore, ANXA11 IDRs might possess a PAM2 motif that interacts with the MLLE^{Pab1}. Visual inspection did not reveal any canonical PAM2 motifs; however, potential non-canonical PAM2 sequences could be identified by *in silico* analysis the N-terminal IDR (data not shown). This observation indicates that ANXA11 could be hypothetically an interaction partner of Pab1 in eukaryotes.

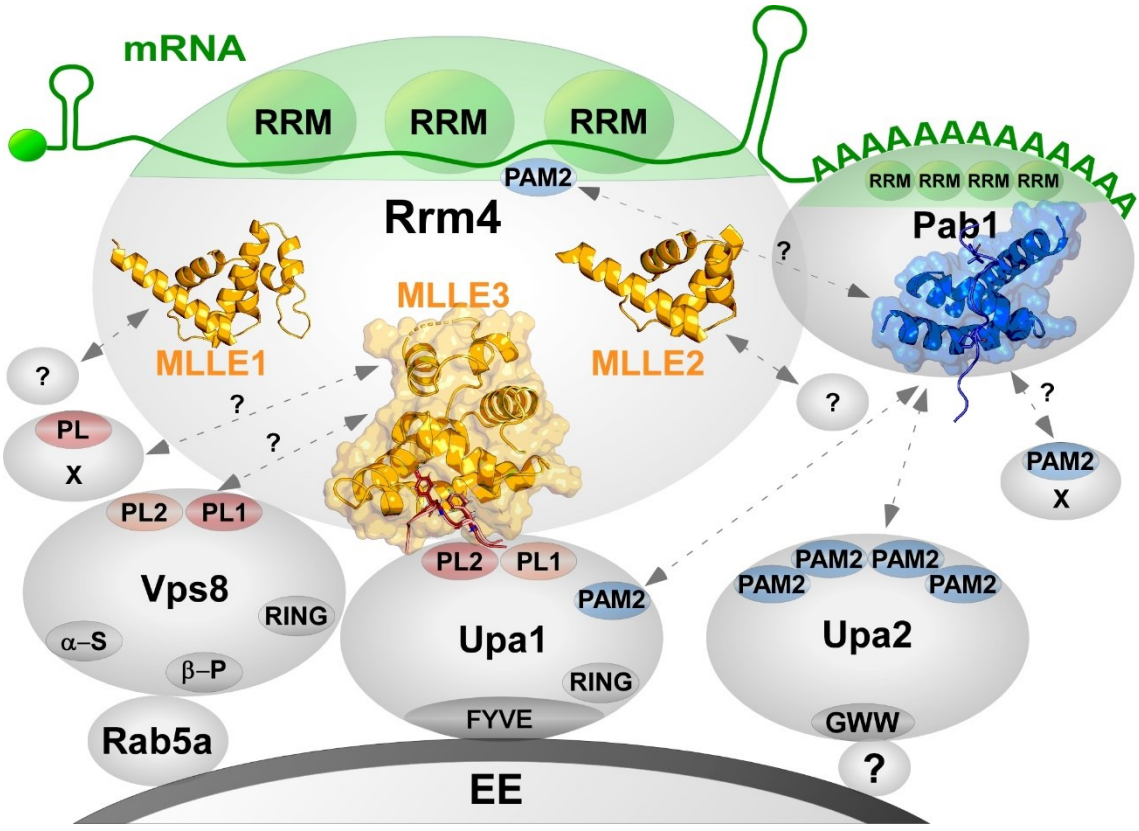
These findings collectively indicate that MLLE domains of Rrm4 and Pab1 might be interacting with multiple proteins in space and time in *U. maydis* for not only endosomal attachment but for several other biological functions, similar to the human MLLE domains from PABPC1 and Ubr5, which interact with dozens of proteins containing PAM2 motif (XIE *et al.* 2014). Since many newly identified proteins with potential PAM2 or

PAM2L candidates are conserved in eukaryotes, their interaction with the MLLE domains is likely to be conserved too.

5.12 New aspects in the model of endosomal attachment of mRNPs via MLLE domains

mRNA transport and local translation are conserved biological processes from prokaryotes to eukaryotes. Active transport of mRNA by the associated RBPs on the cytoplasmic surface of endosomes is widespread, from fungal hyphae to plant endosperms and neuronal cells. Defect in this process results in aberrant growth phenotype in fungi and plants and are associated with neuronal disorders in human (MUNTJES *et al.* 2021).

A key question is how the mRNPs are attached to the endosomes. *U. maydis* is the best-studied model organism for endosome-mediated mRNA transport. Although various components of the endosomal mRNA transport machinery had been identified, a mechanistic understanding of how these RBPs and endosomes are connected was unclear. The



5.12.1.1 Figure 21: Schematic model of endosomal attachment of mRNPs via MLLE domains.

Cargo mRNAs (green) are bound by the N-terminal RRM (1-3) domains of Rrm4 (green). The C-terminal MLLE^{Rrm4} domains (orange) form a binding platform: MLLE3^{Rrm4} interacts with PAM2L1,2 motifs of Upa1. PAM2L1^{Upa1} (Salmon red) and PAM2L2^{Upa1} (dark red) interact with the hydrophobic binding pocket formed by the α2-3 helices of the MLLE3^{Rrm4} domain, MLLE1, 2^{Rrm4} might interact with currently unknown

factors denoted with a question mark, to support the endosomal binding. In particular, MLLE2^{Rrm4} has an accessory role during endosomal interaction. The four RRMs of Pab1 (green) interact with the poly(A) tail. The MLLE^{Pab1} (bright blue) interacts with PAM2 of Upa1 and with the four PAM2 motifs of Upa2 (light blue), a dimerizing scaffold protein, particularly, PAM2^{Upa1} is interacting with the hydrophobic binding pocket formed by the α2,3 and α3,5 helices of the MLLE^{Pab1} domain. Upa1 is attached to endosomes via its FYVE domain. The C-terminal GWW motif of Upa2 is crucial for its endosomal binding through a yet unknown interaction partner, denoted with a question mark. Vps8 is a newly identified interaction partner of MLLE3^{Rrm4} preliminarily confirmed by in vitro assays, which could be a secondary tethering molecule of Rrm4 on endosomes. Vps8 is connected to the endosomes by direct interaction with the Rab5 GTPase. X denotes the *de novo* predicted uncharacterized PAM2 and PAM2L-containing proteins.

structural and functional studies described in this dissertation have provided the mechanistic details of the endosomal attachment of the key RBP Rrm4 and the accessory protein Pab1 via the adapter protein Upa1 in *U. maydis*.

These findings improved the endosomal mRNA transport unit model, and new questions were postulated (Figure 21). Rrm4 is tethered to the endosomes via its C-terminus, which is consisted of three tandem MLLE^{Rrm4} domains that form a sophisticated PPI platform. Of which MLLE1^{Rrm4} is dispensable for function, MLLE2^{Rrm4} is an accessory domain involved in ensuring the correct attachment of Rrm4 to the endosome, and MLLE3^{Rrm4} is the main domain for Upa1 interaction, MLLE3^{Rrm4} consists of a conserved seven alpha helices, non-canonically interacts with the PAM2L1, 2 motifs of Upa1 via the α2-3. By deciphering the key residues in the PAM2L1,2^{Upa1} and PAM2^{Upa1} peptides for MLLE3^{Rrm4} and MLLE^{Pab1} recognition, novel interaction partners of MLLE3^{Rrm4} and MLLE^{Pab1} with PAM2L and PAM2 motifs were identified, respectively.

Endosomal CORVET complex subunit Vps8 is one of the newly found proteins with PAM2L motifs. In the past, the loss of Upa1 did not abolish the endosomal movement of Rrm4. This observation could be due to accessory tethering factors (POHLMANN *et al.* 2015). An additional adaptor protein designated as 'protein X' was postulated to mediate the Rrm4 attachment to the endosomes in cooperation with Upa1 or independently (POHLMANN 2013). The preliminary results obtained from this study have shown that Vps8 has two PAM2L motifs, and both interacted with MLLE3^{Rrm4} *in vitro*. Besides, Vps8 is a specific early endosomal CORVET complex subunit, localized to early endosomes and interacting directly with Rab5a in *U. maydis* (SCHNEIDER *et al.* 2022). Therefore, Vps8 could be the most likely secondary tethering molecule connecting Rrm4 to the endosomes via PAM2L^{Vps8}-MLLE3^{Rrm4} interactions. Since Rrm4-mediated endosomal mRNA transport is essential, it is very important to have an alternative tethering point for the Rrm4-associated mRNPs as a backup.

In vitro binding studies demonstrated that the MLLE domain of Pab1 interacts directly with Rrm4 via a tandem PAM2 motif. This interaction might be important for a potential mRNP stabilization. The localization of the newly identified interaction partners of Rrm4 and Pab1 currently needs to be discovered. For example, Taf7, a TAF, and the nucleocytoplasmic mRNA transporter characterized in fission yeast may play a role in mRNA loading on endosomes by interacting with the MLLE3^{Rrm4} in the cytoplasm. Mitochondrial elongation factor EfTuM may be co-translated at the EE and interacts with the MLLE3^{Rrm4} on the cytoplasmic surface of endosomes before being delivered at their destination or interacting at the surface of the OMM. Although several PAM2 and PAM2L motifs have been reported already, to the best of my knowledge, this is the first study that provided the mechanistic differences between the PAM2^{Upa1} and PAM2L1,2^{Upa1} recognition by MLLE domains.

To conclude, this study has enhanced our understanding of the MLLE domains, their target recognition, and their role in endosomal mRNP attachment in the fungal model system. Future studies will define more clearly the function and specificity of different MLLE domains of Rrm4, characterize the newly identified MLLE3^{Rrm4} and MLLE^{Pab1} interaction partners, and provide a detailed picture of the MLLE-mediated PPI network in *U. maydis*. This might serve as a reference for future studies in higher eukaryotes.

5.13 Outlook

The structural and functional studies described in this thesis have uncovered the mechanistic details of MLLE-mediated mRNP attachment to endosomes in detail and postulated insights into how new interaction partners might be recruited to endosomes. As a result, the model of the endosome-coupled mRNA transport unit has been improved with fine details (Figure 21). Furthermore, the new findings now allow to address new questions. For example, what is the role of the MLLE1^{Rrm4} and MLLE2^{Rrm4} domains in endosomal attachment? Deletion of MLLE2^{Rrm4} in Rrm4-Gfp variants leads to an aberrant accumulation of MLLE2^{Rrm4} on microtubules. What is the mechanism behind this observation? What are the interaction partners of MLLE2^{Rrm4} on endosomes and microtubules? Since the MLLE3^{Rrm4} domain interacts with the PAM2L motif-containing proteins, MLLE1,2^{Rrm4} are likely interacting with PAM2L-containing proteins well. In *in vitro* pull-down experiments, G-MLLE2^{Rrm4} interacted with HS-PAM2L1^{Upa1} and HS-PAM2^{Upa1} peptides. However, such an interaction was not observed in the longer version of the Rrm4 variants carrying MLLE3^{Rrm4} mutations.

Does it mean MLLE2^{Rrm4} has a flexible target selection as it can recognize both PAM2 and PAM2L1 sequences? Does it mean that MLLE1^{Rrm4} and MLLE3^{Rrm4} in full-length protein regulate the target recognition of MLLE2^{Rrm4}, or is it due to steric hindrance?

This study has shown that 7 helices of the MLLE3^{Rrm4} are important for the PAM2L recognition *in vitro*. MLLE3^{Rrm4} without the helices α I, II failed to recognize the PAM2L peptides of Upa1 *in vitro*. However, this still needs to be tested *in vivo*. As the first step, this has to be studied in the fungal hyphae with Rrm4-Gfp, Rrm4-M3 (α I, IIΔ) , and MLLE3^{Rrm4}-Gfp variants. Furthermore, Rrm4 from *R. irregularis* (RiRrm4) was shuttling on the endosomes in *U. maydis* (Muller 2019). However, AlphaFold2 structures showed that they did not have a seven-helix MLLE3^{Rrm4} domain. The interesting question is, how does the 5-helical structure from the RiRrm4 recognize the PAM2L motifs of Upa1 in *U. maydis* for endosomal attachment? Did it interact with the PAM2L motifs of Upa1 or Vps8 or something else? To address this, initially, this can be tested simply by *in vitro* pull-down experiments using GST-MLLE3^{Rrm4} from the *R. irregularis* as bait and HS-PAM2L motifs from Upa1 or Vps8 from *U. maydis*.

Another important question is how, when, and where the newly identified PAM2L and PAM2 candidates interact with the Rrm4 and Pab1 correspondingly. For example, MLLE^{Pab1} recognized the overlapping PAM2 motif of Rrm4. This PAM2 is comparable to the overlapping PAM2 in human eRF3, which prevented the interaction of PABPC1 with deadenylase complexes Caf1–Ccr4 and Pan2–Pan3 by competing with its high affinity for MLLE^{PABPC1} (OSAWA *et al.* 2012; MATTIJSSEN *et al.* 2021). It is interesting to understand if the overlapping PAM2^{Rrm4} regulates MLLE^{Pab1} by competing with other PAM2 proteins. A Rrm4-Kat variant carrying a deletion of overlapping PAM2^{Rrm4} in the Upa1-Gfp background has been generated to address this question (UMAG_3356). It could be initially tested in ITC using PAM2 peptide variants from Upa1, Upa2, and Rrm4 in ITC and further characterized *in vivo*.

MLLE3^{Rrm4} recognized the newly identified PAM2L motifs of the endosomal protein Vps8 *in vitro*. In humans and yeast, Vps8 ensures the maturation of early endosomes to late endosomes (BALDERHAAR AND UNGERMANN 2013). These interactions must be verified *in vivo* using Vps8-Gfp variants carrying N-terminus truncations or mutations in the PAM2L motifs in the Rrm4-mKate background. Vps8-Gfp is functional and co-localized with Rab5-positive EE in *U. maydis* (SCHNEIDER *et al.* 2022). Further, the Vps8-Rrm4

interaction has to be tested in the Upa1D background. Another interesting question is, does Vps8-Rrm4 interaction permanent or transient?

Taf7 is reported as a nucleocytoplasmic mRNA transporter and is associated with transcription. MLLE3^{Rrm4} interacted with PAM2L^{Taf7} *in vitro*. Does Rrm4-Taf7 interact *in vivo* as well? Where do they interact? What is Taf7's role in *U. maydis*? Does Taf7 mediate the mRNA cargo loading on the Rrm4? Taf7-Gfp fusion is functional in fission yeast (SHIBUYA *et al.* 1999). Taf7-Gfp variants with PAM2L mutations can be used to verify the Taf7-Rrm4 interaction and their localization *in vivo* in *U. maydis*. Taf7 does not seem to have an RNA binding domain. However, it has two PPI domains for interacting with the transcription activators and regulators. Studies on these domains may answer the Taf7's role in nucleocytoplasmic mRNA transport, endosomal loading and regulation in *U. maydis*.

PAM2L^{EfTu} from *E. coli* interacted with the MLLE3^{Rrm4} *in vitro.* Both mitochondrial and cytoplasmic EfTu from *U. maydis* have a conserved PAM2L motif which indicates that they are potential interaction partners of MLLE3^{Rrm4}. As a first step this interaction has to be verified with the full-legth or HS-PAM2L variants from the cytoplasmic and mitochondrial EfTu variants from *U. maydis in vitro*. Suppose this interaction is verified in the next step. The interaction and co-localization of Rrm4 and EfTu can be further studied *in vivo* with the EfTu-Gfp variants carrying mutations in the PAM2L motifs. The EfTu-Yfp fusion protein is functional in plants (SHARMA *et al.* 2018). EfTu is a highly conserved globular protein with very less IDR (Figure 18D, E). Intact EfTu variants from Bovine and *E. coli* have been crystallized (SONG *et al.* 1999; ANDERSEN *et al.* 2000). Therefore, it is an interesting target for co-crystallization studies with the MLLE3^{Rrm4}.

Twenty-one new PAM2L motifs and 12 PAM2 motifs were identified from *de novo* prediction in this study, yet only a few were tested. The remaining candidates can be shortlisted based on their functional relevance to the endosomal mRNP association or function and tested by the GST pull-down assays. Since the predicted PAM2L and PAM2 motifs have the conserved key residues (Figure 19A), most of them are likely to interact with the Rrm4 and Pab1, respectively. Based on the *in vitro* results, interesting candidates from the above list *of U. maydis* deletion strains could be generated to understand their functional importance and phenotype quickly. Many of these candidates are essential genes in other organisms. Therefore, it might be very difficult or impossible to generate a deletion strain. In this case, strains with PAM2L mutation in the respective protein could be generated.

RING domains are known to be involved in the Ubiquitination of proteins. Vps8, Upa1, and the other predicted interaction partner UMAG_10591 are PAM2L-containing proteins with a RING domain. Besides, MLLE domains of Rrm4 are more related to the Ubr5 than the PABPC1. Therefore, it is interesting and important to study the role of Rrm4 in Ubiquitination in the future.

Phosphorylation modulates MLLE-PAM2 interactions at the Ser/Thr residues near the PAM2 motifs (HUANG *et al.* 2013). This aspect has not been studied in the $MLLE^{Pab1}$-$PAM2^{Upa1}$ or $MLLE3^{Rrm4}$-$PAM2L1,2^{Upa1}$ interactions in *U. maydis*. Studying the phosphorylation's role in MLLE interaction in *U. maydis* can unravel the questions on loading and unloading of the Rrm4 and Pab1 on endosomes. In addition, a PAM2 containing Serine/Threonine kinase (UMAG_15011) is predicted as an interaction partner of Pab1. If this interaction is true, this may be the first Serine / Threonine kinase protein interacting and may regulate the phosphorylation of MLLE-PAM2 interactions.

Annexin11 mediated RNA granules transport in human lysosomes and it has a large N-terminal IDR. ANXA11 is structurally conserved in *U. maydis*. Interesting question is, does the ANXA11 mediated RNA granules transport is also conserved in *U. maydis*? Does, ANXA11 has an important function in *U. maydis*? Does Annexin11 IDR interact with $MLLE^{Pab1}$?

The *de novo* prediction of PAM2 and PAM2L motifs were based on the conserved key residues identified from the PAM2 and PAM2L1,2 motifs of Upa1. However, PAM2L variants from the EfTu have non-canonical hydrophobic residues. Initially, the interaction of such non-canonical PAM2L residues with $MLLE3^{Rrm4}$ must be tested *in vitro* and *in vivo*. Upon successful confirmation, these PAM2L codes could be used for identifying more non-canonical PAM2L motifs from the *U. maydis* genome.

Reiterative blast search was used as a primary tool to identify the new PAM2L candidates, followed by AlphaFold2 structure prediction and sequence conservation using MSA in the Basidiomycota fungi. Although BLAST search is a simple tool, not all the proteins in *U. maydis* have the AlphaFold2 structures at the time of this study. AlphaFold2 structure prediction of each protein and MSA analysis with its orthologues in Basidiomycota is a lingering and lengthy task. Developing an automated tool that could perform all three functions to predict the PAM2 or PAM2L and verify their location in IDR and conservation in Basidiomycota would make this task easier and more efficient for future work.

Several PAM2 motif-containing proteins are identified and characterized in humans (KOZLOV *et al.* 2001; ALBRECHT AND LENGAUER 2004). However, the known PAM2L-containing proteins are limited. This study identified PAM2L motifs that are conserved from fungal hyphae to human proteins (Figure S14C, S15A-B) for example, EfTu, Taf7 and Vps8 and they specifically interact with the MLLE3^{Rrm4} *in vitro* (Figure 18,19). If these interactions are true, they could be conserved in human MLLE^{Ubr5} too, which will be interesting to study in the future. On the other hand, Yeast MLLE^{Pab1} has been reported as an orphan domain without any known interaction partners from the yeast genome. Yeast has only one MLLE^{Pab1p} domain in its genome, which resembles the MLLE domain of Ubr5 (KOZLOV *et al.* 2002). Ubr5 and Rrm4 are known to interact with PAM2L motifs. Therefore, it is likely that yeast MLLE recognizes the PAM2L peptides and not the PAM2 peptides. Human MLLE^{Ubr5} and yeast MLLE^{Pab1} might recognize the PAM2L peptides from the EfTu and Taf7 from the respective orthologue. This can be tested using the *in vitro* pull-down experiments with GST fused MLLE variants from yeast Pab1 or human Ubr5. Novel PAM2L candidates from the human yeast genome could be predicted by using the conserved residues of these PAM2L motifs in reiterative BLAST.

In addition to the sequence-based PAM2L predictions, biochemical methods can be employed to find the new identification partners of MLLE domains. Especially in the case of MLLE1,2^{Rrm4}, where the interaction partners are unknown. For example, 'Far-western blot and mass spectrometry experiments can be carried out with the purified MLLE(1-3)^{Rrm4} domains with a C-terminal epitope tag from *E. coli* to identify the interaction partner of the respective MLLE domain specifically. A similar approach can be used for identifying the interacting partners of the Upa2-GWW domain. Recently human EfTuM was identified as an interaction partner of viral IFN regulatory factor 1 (vIRF-1) using the above strategy (CHOI *et al.* 2022).

This study explained the mechanism of PAM2L1,2^{Upa1} peptide interaction with MLLE3^{Rrm4} using X-ray crystallization studies. However, non-canonical PAM2L sequences are identified with various hydrophobic (Leu, Met, Thr, Pro) key residues instead of the canonical Phe and Tyr residues in the EfTu variants (Figure 19A) and Upa1 orthologues (Figure 20E). The important question is, what are the residues accepted or tolerated at these positions? Is its sufficient to have any hydrophobic residues instead of the FxY residues? In order to understand the PAM2L peptide recognition by MLLE3^{Rrm4} entirely, FxY residues of the PAM2L1 or 2^{Upa1} could be substituted with each hydrophobic residue and tested systematically using *in vitro* pull-down assays. Above and beyond,

it will be interesting to understand the newly identified non-canonical PAM2L peptide variant recognition by MLLE3^{Rrm4} using crystallization studies in the future. In addition, each newly identified PAM2 and PAM2L motifs could be used in synthetic biology approaches to generate artificial protein-protein interactions at various binding affinities.

Alphafold2 structure prediction revealed the presence of two DUF domains (Domain of unknown function) in Upa1 (Figure S15C) which escaped the analysis in the past due to the limitation of the earlier structure prediction tools (POHLMANN *et al.* 2015). Function of these domains in Upa1 is completely unknown. It is important to be characterized in future.

Furthermore, it is important to understand the mechanism of mRNP interactions using full-length proteins of core endosomal mRNA transport components with their interaction partners. Full-length fungal protein production and purification is a complicated process in *E. coli* due to the large size and unstable IDR. Nonetheless, simple eukaryotic protein production hosts, for example, *Pichia Pastoris* (extra cellular expression) and *S. cerevisiae* (intra cellular expression) could be used to overcome this limitation. Previously all the core components of endosomal mRNPs, Rrm4, Pab1, Upa1, and Upa2 have been successfully expressed in *S. cerevisiae* for the Y2H analysis (POHLMANN *et al.* 2015; JANKOWSKI *et al.* 2019; DEVAN *et al.* 2022). Endosomal mRNP complex could be reconstituted using the purified full-length proteins for structural studies using cryo-electron microscopy and membrane binding assays.

6 Materials and methods

6.1 AlphaFold2 structure prediction and structure comparison.

The AlphaFold2 algorithm was used to obtain computationally predicted structural models of the Rrm4, Pab1 full-length, and truncated versions. AlphaFold2 is a novel machine-learning approach that incorporates physical and biological knowledge about protein structure, leveraging multi-sequence alignments, into the design of the deep learning algorithm (JUMPER *et al.* 2021). The desired protein sequence was submitted to the colab notebook encoding the AlphaFold2 algorithm at the following link. https://colab.research.google.com/github/sokrypton/ColabFold/blob/main/beta/AlphaFold2_advanced.ipynb#scrollTo=pc5-mbsX9PZC (MIRDITA *et al.* 2022). Five models for each protein sequence were generated using default parameters. The model with the highest score was used in the structural analysis. The structures were compared using the superpose tool of PyMOL to calculate the corresponding RMSD. The images of the AlphaFold2 predicted models were prepared using PyMOL (version 2.5.2).

6.2 Multiple sequence alignment

Multiple sequence alignment (MSA) were performed using the clustalX2 (LARKIN *et al.* 2007) by submitting the protein sequences obtained from the NCBI protein database. MSA images were prepared using the Genedoc sequence editor tool (version 2.6) (CAFFREY *et al.* 2007)

6.3 Plasmids, strains, and growth conditions

For molecular cloning of plasmids, *E. coli* Top10 cells (Thermofisher C404010) and recombinant protein expression *E .coli Lobstr* cells (Kerafast EC1002) were used, respectively. Sequence encoding H-Rrm4-M3-7H, H-Pab1-M-4H was cloned into the pET22 vector (Merck 69744) with an N-terminal hexa-histidine tag for crystallization studies. Sequence encoding MLLE variants were cloned into the pGEX-2T vector (Merck GE28-9546-53) containing GST sequence in N-terminus for pulldown experiments. Sequence encoding PAM2 variants were cloned into the Champion pET-Sumo vector (Thermofisher K30001). pRarepLys plasmid was co-transformed in *the E. coli Lobstr strain* to supplement the rare codons for efficient recombinant protein production. *E. coli* transfor-

mation, cultivation, and plasmid isolation were conducted using standard molecular biology techniques. A detailed description of all plasmids, strains, and oligonucleotides is given in the Supplementary tables S17–S18. Sequences are available upon request.

6.4 Recombinant protein expression and purification

Recombinant protein expression in *E. coli* was performed as per the previous report (DEVAN *et al.* 2022). In a nutshell, *E. coli* Lobstr cells from freshly transformed plates were inoculated in 20 ml expression media (1.6 % Trypton, 1% Yeast extract, 50 mM Na2HPO4, 50 mM KH2PO4, 25 mM [NH4]2SO4, 0.5% Glycerol, 0.5% Glucose, 2 mM MgSO4) with ampicillin (100 mg/ml) and chloramphenicol (34 mg/ml) or kanamycin (200 mg/ml) and chloramphenicol (34 mg/ml) and grown overnight (16 hours) at 37 °C, 200 rpm. Note that the high concentration of kanamycin was used to prevent the unintended resistance promoted by high phosphate concentration (STUDIER 2005). Supernatant from the overnight culture was removed by centrifugation at 4 °C, 5000 × g for 2 minutes. Cells were resuspended in fresh media with a starting OD600 of 0.1 and grown at 37 °C, 200 rpm for about 2 hours 30 minutes until the OD600 = 1. Protein expression was induced at 28 °C, 200 rpm, for 4 hours by adding 1 mM IPTG and harvested by centrifugation at 4 °C, 6,000 × g for 5 minutes.

6.5 Recombinant protein purification

H-Rrm4-M3-7H versions were purified for crystallization studies as per the previous report (DEVAN *et al.* 2022). In essence, Hexa-histidine tagged protein was purified using Nickel-based affinity chromatography (HisTrap HP, GE Healthcare) on Akta prime FPLC system. Cells were thawed on ice and resuspended in buffer A (20 mM HEPES pH 8.0, 200 mM NaCl, 1 mM EDTA, 10 mM Imidazole pH 8.0; 1 mM PMSF, 0.5 mg/ml Lysozyme, 0.5 mg/ml DNase). Subsequently, cells were lysed by sonication on ice and centrifuged at 4 °C 18,000 × g for 30 minutes. The resulting supernatant was loaded onto a pre-equilibrated column with buffer B (20 mM HEPES pH 8.0, 200 mM NaCl,10 mM Imidazole), washed with buffer C (20 mM HEPES pH 8.0, 200 mM NaCl, 50 mM Imidazole), eluted with buffer D (20 mM HEPES pH 8.0, 200 mM NaCl, 300 mM Imidazole.) and further purified by size exclusion chromatography (HiLoad 26/600 Superdex 200, GE Healthcare), pre-equilibrated with storage buffer E (20 mM HEPES pH 8.0, 200 mM NaCl). H-Pab1-M-4H versions were purified as above, except that the wash buffer C was

prepared with 20 mM Imidazole (20 mM HEPES pH 8.0, 200 mM NaCl, 20 mM Imidazole).

6.6 GST pull-down experiments

Pull-down assays were performed as per the previous report (DEVAN *et al.* 2022). In short, GST-MLLE variants and HS-PAM2^{Upa1} variants were expressed in *E. coli*. Cell pellets from 50 ml *E. coli* expression culture were resuspended in 10 ml buffer F (20 mM HEPES pH 8.0, 200 mM NaCl, 1 mM EDTA; 0.5% Nonidet P-40, 1 mM PMSF, 0.1 mg/ml Lysozyme). Cells were lysed by sonication on ice and centrifuged at 4 °C, 16,000 × g for 10 minutes. 1 mL of the resulting supernatant was incubated for 1 hour at 4 °C on constant agitation of 1,000 rpm with 100 µL glutathione sepharose (GS) resin (GE Healthcare), pre-equilibrated in buffer F. The GS resin was washed three times with 1 ml of buffer G (20 mM HEPES pH 8.0, 200 mM NaCl, 1 mM EDTA, 0.5 % Nonidet P-40). Subsequently, supernatant of HS-PAM2^{Upa1} variants was added to the GST-MLLE variant bound resins and incubated for 1 hour at 4 °C on agitation. The resins were washed as aforementioned, resuspended in 100 µL of 2x Laemmli loading buffer, boiled for 10 minutes at 95 °C, and analyzed by western blotting.

6.7 SDS-PAGE and immunoblotting

All SDS–PAGE and Western blotting experiments were performed as reported previously (DEVAN *et al.* 2022). Briefly, Western blotting samples were resolved by 12 % SDS-PAGE and transferred and immobilized on nitrocellulose membrane (Amersham Protran) by semi-dry blotting using Towbin buffer (25 mM Tris pH 8.6, 192 mM Glycine, 15% Methanol). Proteins were detected using α-Histidine from mouse (Sigma H1029) as the primary antibody. As secondary antibodies α-mouse IgG HRP conjugate (Promega W4021) was used. Detection was carried out by using ECLTM Prime (Cytiva RPN2236). Images were taken with a luminescence image analyzer, LAS4000 (GE Healthcare), according to the manufacturer's instructions.

6.8 Protein crystallization

For co-crystallization studies, H-Rrm4-M3-7H and H-Pa1-M-4H versions of purified recombinant proteins were used. PAM2^{Upa1} and PAM2L1,2^{Upa1} peptides were custom-synthesized and purchased from Genescript, USA. The peptide was dissolved in storage buffer E (20 mM Hepes pH 8.0, 200 mM NaCl). All the protein samples were centrifuged

at 100,000 × g for 30 minutes and quantified by Nanodrop (A280) before use and mixed with the PAM2^{Upa1} or PAM2L1,2^{Upa1} peptide variant in a 1:1.5 molar ratio to obtain the final protein concentration of 12 mg/ml. Initial crystallization conditions were searched using MRC-3, 96-well sitting drop plates, and various commercially available crystallization screens at 12 °C. 0.1 µL homogeneous protein-peptide solution was mixed with 0.1 µL reservoir solution and equilibrated against 40 µL of the reservoir. After one week, initial rod-shaped crystals were found, further optimized by slightly varying the precipitant concentrations. Optimization was also performed in sitting drop plates (24-well) at 12 °C but by mixing 1 µL protein solution with 1 µL of the reservoir solution, equilibrated against 300 µL reservoir solution. Best diffracting crystals of H-Rrm4-M3-7H with PAM2L1^{Upa1} and H-Rrm4-M3-7H with PAM2L2^{Upa1} complexes were grown within 7 days in 0.1 M Sodium HEPES pH 7.5, 25% PEG 3000. Best diffracting crystals of H-Pa1-M-4H with PAM2 complexes were grown within 7 days in 3.2 M AmSO4 + 0.1M MES pH 6. Before harvesting the crystal, crystal-containing drops were overlaid with 2 µL mineral oil and immediately flash-frozen in liquid nitrogen.

6.9 Data collection, processing, and structure refinement

A complete data set of the MLLE^{Rrm4}-PAM2L1,2^{Upa1}, MLLE^{Pab1}-PAM2^{Upa1} complexes was collected at beamline ID23EH1 (ESRF, France) at 100 K and wavelength 0.98 Å up to 2.6 Å resolution. All data were processed using the automated pipeline at the EMBL HAMBURG and reprocessed afterward using XDS (KABSCH 2014). AlphaFold2 predicted models for MLLE3^{Rrm4} and MLLE^{Pab1} successfully phase the 1.7 Å data set of MLLE^{Rrm4}-PAM2L1^{Upa1}, 2.4 Å data set of MLLE^{Rrm4}-PAM2L2^{Upa1}, 1.7 Å data set of MLLE^{Pab1}-PAM2^{Upa1}, using the program Phaser from the program suite Phenix (AFONINE *et al.* 2012). The structure was then refined in iterative cycles of manual building and refinement in Coot (EMSLEY AND COWTAN 2004), followed by software-based refinements using the program suite Phenix (AFONINE *et al.* 2012). All residues were in the preferred and additionally allowed regions of the Ramachandran plot (S16 Table). The data collection and refinement statistics are listed in the S2 Table. The structure and models were compared using the superpose tool of PyMOL to calculate the corresponding RMSD. The protein-Peptide interface resides identified using the Ligplot software tool (LASKOWSKI AND SWINDELLS 2011). The images of the co-crystallized X-ray structure models were prepared using PyMOL (version 2.5.2) and UCSF ChimeraX (version 1.4rc202205290614 (2022-05-29)).

6.10 Appendix II

6.10.1 Table S10. Accession numbers for protein sequences used in multiple sequence alignment of MLLE3^{Rrm4} type domain from Rrm4 orthologues

Organism Name	Protein Name	Uniprot/*NCBI* ID	Sequence coverage
Coprinopsis cinerea	Rrm4 orthologue	A8NCM2	751-862
Amanita muscaria	Rrm4 orthologue	A0A0C2SHU5|	675-786
Phanerochaete carnosa	Rrm4 orthologue	*XP_007393387.1*	672-785
Moesziomyces antarcticus	Rrm4 orthologue	*XP_014657015.1*	684-798
Sporisorium reilianum	Rrm4 orthologue	*CBQ73718.1*	660-785
Ustilago maydis	Rrm4	A0A0D1DWZ5	752 - 792
Pseudozyma hubeiensis	Rrm4 orthologue	*XP_012192836.1*	537-651
Melanopsichium pennsylvanicum	Rrm4 orthologue	*CDI54139.1*	687-800
Ustilago hordei	Rrm4 orthologue	I2FYN4	685-798
Kalmanozyma brasiliensis	Rrm4 orthologue	*XP_016289934.1*	573-686
Microbotryum lychnidis-dioicae	Rrm4 orthologue	*KDE02990.1*	551-661
Rhizopus microspores	Rrm4 orthologue	*ORE17079.1*	553-642
Rhizopus delemar	Rrm4 orthologue	I1CQR1	487-567
Mucor ambiguous	Rrm4 orthologue	GAN10032.1	888-974
Lichtheimia corymbifera	Rrm4 orthologue	CDH52259.1	682-760
Rhizophagus irregularis	Rrm4 orthologue	GBC41783.1	572-655

6.10.2 Table S11. Accession numbers for protein sequences used in multiple sequence alignment of *EfTu* orthologues

Organism Name	Protein Name	Target sequence	Uniprot/*NCBI* ID
Moesziomyces antarcticus	EfTu-Mitochondrial	PAM2L	*XP_014659838.1*
Sporisorium reilianum	EfTu-Mitochondrial	PAM2L	*CBQ67532.1*
Ustilago maydis	EfTu-Mitochondrial	PAM2L	A0A0D1CZK5
Pseudozyma hubeiensis	EfTu-Mitochondrial	PAM2L	*XP_012190298.1*
Melanopsichium pennsylvanicum	EfTu-Mitochondrial	PAM2L	*CDI51084.1*
Ustilago hordei	EfTu-Mitochondrial	PAM2L	XP_041409347.1
Kalmanozyma brasiliensis	EfTu-Mitochondrial	PAM2L	*XP_016294755.1*
Testicularia cyperi	EfTu-Mitochondrial	PAM2L	*PWZ02900.1*
Malassezia globosa	EfTu-Mitochondrial	PAM2L	*XP_001729762.1*
Cryptococcus neoformans var grubii	EfTu-Mitochondrial	PAM2L	*XP_012051135.1*
Saccharomyces cerevisiae	EfTu-Mitochondrial	PAM2L	*P02992*
Rhizophagus irregularis	EfTu-Mitochondrial	PAM2L	*PKY46354.1*
Homo sapiens	EfTu-Mitochondrial	PAM2L	*P49411*
E. coli	EfTu	PAM2L	*A0A140NCI6*
Moesziomyces antarcticus	Elongation factor EF-1 alpha	PAM2L	XP_014658646.1
Sporisorium reilianum	Elongation factor EF-1 alpha	PAM2L	SJX60979.1
Ustilago maydis	Elongation factor EF-1 alpha	PAM2L	XP_011386931.1
Pseudozyma hubeiensis	Elongation factor EF-1 alpha	PAM2L	XP_012192602.1
Melanopsichium pennsylvanicum	Elongation factor EF-1 alpha	PAM2L	CDI51837.1
Ustilago hordei	Elongation factor EF-1 alpha	PAM2L	XP_041415607.1
Kalmanozyma brasiliensis	Elongation factor EF-1 alpha	PAM2L	XP_016295136.1
Testicularia cyperi	Elongation factor EF-1 alpha	PAM2L	PWZ00941.1
Malassezia globosa	Elongation factor EF-1 alpha	PAM2L	XP_001732312.1
Cryptococcus neoformans var grubii	Elongation factor EF-1 alpha	PAM2L	OWZ73299.1
Saccharomyces cerevisiae	Elongation factor EF-1 alpha	PAM2L	NP_009676.1
Rhizophagus irregularis	Elongation factor EF-1 alpha	PAM2L	PKY45134.1
Homo sapiens	Elongation factor EF-1 alpha	PAM2L	NP_001949.1

6.10.3 Table S12. Accession numbers for protein sequences used in multiple sequence alignment of Taf7 orthologues

Organism Name	Protein Name	Target sequence	Uniprot/*NCBI* ID
Ustilago maydis	Taf7 orthologue	PAM2L	*XP_011387353.1*
Pseudozyma hubeiensis	Taf7 orthologue	PAM2L	*XP_012192603.1*
Sporisorium reilianum	Taf7 orthologue	PAM2L	*SJX60978.1*
Kalmanozyma brasiliensis	Taf7 orthologue	PAM2L	*EST10146.2*
Ustilago hordei	Taf7 orthologue	PAM2L	*XP_041415606.1*
Melanopsichium pennsylvanicum	Taf7 orthologue	PAM2L	*CDI51836.1*
Moesziomyces antarcticus	Taf7 orthologue	PAM2L	*XP_014658645.1*
Testicularia cyperi	Taf7 orthologue	PAM2L	*PWZ00942.1*
Malassezia globosa	Taf7 orthologue	PAM2L	*XP_001732308.1*
Cryptococcus neoformans var grubii	Taf7 orthologue	PAM2L	*OXB39574.1*
Saccharomyces pombe	Taf7	PAM2L	O13701
Saccharomyces cerevisiae	Taf7 orthologue	PAM2L	Q05021
Rhizophagus irregularis	Taf7 orthologue	PAM2L	*PKC16038.1*
Homo sapiens	Taf7 orthologue	PAM2L	Q15545

6.10.4 Table S13. Accession numbers for protein sequences used in multiple sequence alignment of Vps8 orthologues

Organism Name	Protein Name	Target sequence	Uniprot/*NCBI* ID
Ustilago maydis	Vps8	PAM2L	A0A0D1DXQ1
Pseudozyma hubeiensis	Vps8 orthoogue	PAM2L	R9P7V6
Kalmanozyma brasiliensis	Vps8 orthoogue	PAM2L	*EST05200.2*
Sporisorium reilianum	Vps8 orthoogue	PAM2L	A0A2N8UEF7
Melanopsichium pennsylvanicum	Vps8 orthoogue	PAM2L	*CDI53747.1*
Ustilago hordei	Vps8 orthoogue	PAM2L	*XP_041410257.1*
Moesziomyces antarcticus	Vps8 orthoogue	PAM2L	*XP_014656004.1*
Testicularia cyperi	Vps8 orthoogue	PAM2L	*PWY99436.1*
Malassezia globosa	Vps8 orthoogue	PAM2L	A8QA38
Cryptococcus neoformans var grubii	Vps8 orthoogue	PAM2L	J9VWP5
Rhizophagus irregularis	Vps8 orthoogue	PAM2L	A0A2I1GF06
Saccharomyces cerevisiae	Vps8 orthoogue	PAM2L	P39702
Homo sapiens	Vps8 orthoogue	PAM2L	Q8N3P4

6.10.5 Table S14. Data collection and refinement statistics for X-ray structures

Parameters	MLLE3^{Rrm4}-PAM2L1^{Upa1}	MLLE3^{Rrm4}-PAM2L2^{Upa1}	MLLE^{Pab1}-PAM2^{Upa1}
Wavelength	0.9253	1	0.8856
Resolution range	46.61 - 1.743 (1.805 - 1.743)	56.17 - 2.4 (2.486 - 2.4)	40.27 - 1.67 (1.73 - 1.67)
Space group	C 2 2 21	C 1 2 1	C 1 2 1
Unit cell	74.715 83.486 170.398 90 90 90	82.881 74.633 168.511 90 90.081 90	88.416 47.851 45.524 90 117.806 90
Total reflections	224109 (17004)	120083 (11963)	66371 (6926)
Unique reflections	54171 (4939)	39386 (3912)	19112 (1893)
Multiplicity	4.1 (3.4)	3.0 (3.1)	3.5 (3.7)
Completeness (%)	98.91 (91.55)	97.17 (98.27)	97.18 (97.03)
Mean I/sigma(I)	9.06 (0.86)	8.20 (1.40)	16.07 (2.45)
Wilson B-factor	26.97	49.66	26.59
R-merge	0.08617 (1.193)	0.09393 (1.046)	0.03473 (0.532)
R-meas	0.09871 (1.409)	0.1134 (1.261)	0.04111 (0.6231)
R-pim	0.04717 (0.73)	0.06273 (0.6948)	0.0217 (0.3213)
CC1/2	0.998 (0.361)	0.997 (0.581)	0.999 (0.851)
CC*	0.999 (0.728)	0.999 (0.857)	1 (0.959)
Reflections used in refinement	54087 (4938)	39281 (3912)	19103 (1893)
Reflections used for R-free	2667 (232)	897 (93)	1015 (108)
R-work	0.2200 (0.3471)	0.2399 (0.3498)	0.1959 (0.2816)
R-free	0.2498 (0.3993)	0.2760 (0.4326)	0.2201 (0.3148)
CC(work)	0.944 (0.628)	0.923 (0.644)	0.947 (0.754)
CC(free)	0.921 (0.592)	0.900 (0.261)	0.951 (0.682)
Number of non-hydrogen atoms	4087	7457	1363
Macromolecules	3816	7424	1254
Solvent			10
Protein residues	271	33	99
RMS(bonds)	504	968	163
RMS(angles)	0.008	0.025	0.015
Ramachandran favored (%)	1.30	2.34	1.78
Ramachandran allowed (%)	96.52	96.69	99.35
Ramachandran outliers (%)	2.87	2.46	0.65
Rotamer outliers (%)	0.61	0.85	0.00
Clashscore	0.00	0.12	5.67
Average B-factor	6.89	18.74	2.76
Macromolecules	33.35	65.37	32.48
Solvent	32.93	65.42	31.77

Statistics for the highest-resolution shell are shown in parentheses.

6.10.6 Table S15. Description of plasmids used for recombinant expression in *E. coli*

Plasmid	Plasmid ID	Short description
pGEX-G-Pab1-M	pUMa2187	Plasmid for the expression of the G-Pab1-M. C-terminal region of Pab1 comprising amino acid residues from 566 to 651 was fused to an N-terminal GST- tag (POHLMANN *et al.* 2015).
pGEX-G-Rrm4-NT4	pUMa3920	Plasmid for the expression of the G-Rrm4-NT4. C-terminal region of Rrm4 comprising amino acid residues from 421 to 792 was fused to an N-terminal GST- tag (POHLMANN *et al.* 2015).
pGEX-G-Rrm4-NT4-M1,2Δ	pUMa4619	Plasmid for the expression of the G-Rrm4-NT4-M1,2Δ. Same as pUMa3920 but carrying the deletion of 1st to 2nd MLLE domains. Amino acid residues of Rrm4 from 447 to 644 were replaced with a HA tag-HRV3C protease recognition site (DEVAN *et al.* 2022).
pGX-G-Rrm4-M1	pUMa4620	Plasmid for the expression of G-Rrm4-M1. MLLE1 domain of Rrm4 comprising amino acid residues from 462-630 was fused to an N-terminal GST- tag.
pGX-G- Rrm4-M2	pUMa4621	Plasmid for the expression of G-Rrm4-M2. MLLE2 domain of Rrm4 comprising amino acid residues from 568 to 626 was N-terminally fused to a GST-tag.
pGX-G-Rrm4-M3 or pGX-G-Rrm4-M3-4H	pUMa4639	Plasmid for the expression of G- Rrm4-M2. MLLE3 domain of Rrm4 comprising amino acid residues from 723 to 792 was fused to an N-terminal GST- tag.
pGX-G- Rrm4-M3-5H	pUMa4701	Plasmid for the expression of G-Rrm4-M3-5H. MLLE2 domain of Rrm4 comprising amino acid residues from 700 to 792 was fused to an N-terminal GST- tag.
pGX-G-Rrm4-M3-7H	pUMa4702	Plasmid for the expression of G-Rrm4-M3-7H. MLLE3 domain of Rrm4 comprising amino acid residues from 679 to 792 was fused to an N-terminal GST- tag.
pGX-G-Pab1-M-4H	pUMa4790	Plasmid for the expression of G-Pab1-M-4H. MLLE domain of Pab1 comprising amino acid residues from 558 to 651 was fused to an N-terminal GST- tag.
pGX-G-Pab1-M-5H	pUMa4791	Plasmid for the expression of G-Pab1-M-5H. MLLE domain of Pab1 comprising amino acid residues from 567 to 636 was fused to an N-terminal GST- tag.
pGX-G- Rrm4-M3-7H-H729A	pUL0198	Plasmid for the expression of G-Rrm4-M3-7H. MLLE3 domain of Rrm4 comprising amino acid residues from 679 to 792 with a point mutation (Histidine 729 to Alanine) was fused to an N-terminal GST- tag.
pGX-G-Rrm4-M3-7H-R744A	pUL200	Plasmid for the expression of G-Rrm4-M3-7H. MLLE3 domain of Rrm4 comprising amino acid residues from 679 to 792 with a point mutation (Arginine 744 to Alanine) were fused to an N-terminal GST- tag.
pGX-G-Rrm4-M3-7H-I756G	pUL0201	Plasmid for the expression of G-Rrm4-M3-7H. MLLE3 domain of Rrm4 comprising amino acid residues from 679 to 792 with a point mutation (Isoleucine 756 to Glycine) was fused to an N-terminal GST- tag.
pGX-G- Rrm4-M3-7H-Q733A	pUL0206	Plasmid for the expression of G-Rrm4-M3-7H. MLLE3 domain of Rrm4 comprising amino acid residues from 679 to 792 with a point mutation (Glutamine 733 to Alanine) was fused to an N-terminal GST- tag.
pGX-G-Rrm4-M3-7H-F740A	pUL0207	Plasmid for the expression of G-Rrm4-M3-7H. MLLE3 domain of Rrm4 comprising amino acid residues from 679 to 792 with a point mutation (Phenylalanine 740 to Alanine) was fused to an N-terminal GST- tag.

pGX-G-Rrm4-M3-7H-T755A	pUL0208	Plasmid for the expression of G-Rrm4-M3-7H. MLLE3 domain of Rrm4 comprising amino acid residues from 679 to 792 with a point mutation (Threonine 755 to Alanine) was fused to an N-terminal GST- tag.
H-Rrm4-M3-7H	pUMa4704	Plasmid for the expression of Rrm4-M3-7H. MLLE3 domain of Rrm4 comprising amino acid residues from 679 to 792 was fused to an N-terminal 6xHis-Sumo tag.
H-Pab1-M-4H	pUMa4794	Plasmid for the expression of Pab1-M-4H. MLLE domain of Pab1 comprising amino acid residues from 567 to 636AA was fused to an N-terminal 6xHis-Sumo tag.
pET28-HS-PAM2^{Upa1}	pUMa4296	Plasmid for the expression of the PAM2 motif of Upa1 (SQSTLSPNASVFKPSRS) was fused to an N-terminal 6xHis-Sumo tag (DEVAN *et al.* 2022).
pET28-HS_PAM2L1^{Upa1}	pUMa4297	Plasmid for the expression of PAM2L1 motif of Upa1 (EAADQEEDQDDFVYP-GAD) was fused to an N-terminal 6xHis-Sumo tag (DEVAN *et al.* 2022).
pET28-HS-PAM2L2^{Upa1}	pUMa4298	Plasmid for the expression of PAM2L2 motif of Upa1 (DEDAADDDDDEFIYPNSY) was fused to an N-terminal 6xHis-Sumo tag (DEVAN *et al.* 2022).
pET28-SS-PAM2^{Upa1}	pUMa4687	Plasmid for the expression of the PAM2 motif of Upa1 (SQSTLSPNASVFKPSRS) was fused to an N-terminal Strep-Sumo tag.
pET28-SS_PAM2L1^{Upa1}	pUMa4688	Plasmid for the expression of PAM2L1 motif of Upa1 (EAADQEEDQDDFVYP-GAD) was fused to an N-terminal Strep-Sumo tag.
pET28-SS-PAM2L2^{Upa1}	pUMa4689	Plasmid for the expression of PAM2L2 motif of Upa1 (DEDAADDDDDEFIYPNSY) was fused to an N-terminal Strep-Sumo tag.
pET28-HS-short-PAM2L1^{Upa1}	pUL0138	Plasmid for the expression of PAM2L1 motif of Upa1 (DDFVYPGAD) was fused to an N-terminal 6xHis-Sumo tag.
pET28-HS-short-PAM2L2^{Upa1}	pUL0139	Plasmid for the expression of PAM2L2 motif of Upa1 (DDEFIYPNSY) was fused to an N-terminal 6xHis-Sumo tag.
pET28-HS-PAM2L1-F248A^{Upa1}	pUL0169	Plasmid for the expression of PAM2L1 motif of Upa1 (EAADQEEDQDD**A**VYP-GAD) carrying a point mutation (Phenylalanine 248 to Alanine) was fused to an N-terminal 6xHis-Sumo tag.
pET28-HS-PAM2L1-Y250A^{Upa1}	pUL0166	Plasmid for the expression of PAM2L1 motif of Upa1 (EAADQEEDQDDFV**A**P-GAD) carrying a point mutation (Tyrosine 250 to Alanine) was fused to an N-terminal 6xHis-Sumo tag.
pET28-HS-PAM2L1-P251A^{Upa1}	pUL0171	Plasmid for the expression of PAM2L1 motif of Upa1 (EAADQEEDQDD-FVY**A**GAD) carrying a point mutation (Proline 251 to Alanine) was fused to an N-terminal 6xHis-Sumo tag.
pET28-HS-PAM2L2-F955A^{Upa1}	pUL0170	Plasmid for the expression of PAM2L2 motif of Upa1 (DEDAADDDDDE**A**IYPNSY) carrying a point mutation (Phenylalanine 955 to Alanine) was fused to an N-terminal 6xHis-Sumo tag.
pET28-HS-PAM2L2-Y957A^{Upa1}	pUL0167	Plasmid for the expression of PAM2L2 motif of Upa1 (DEDAADDDDDEFI**A**PNSY) carrying a point mutation (Tyrosine 957 to Alanine) was fused to an N-terminal 6xHis-Sumo tag.
pET28-HS-PAM2L2-P958A^{Upa1}	pUL0172	Plasmid for the expression of PAM2L2 motif of Upa1 (DEDAADDDDDEFIY**A**NSY) carrying a point mutation (Proline 958 to Alanine) was fused to an N-terminal 6xHis-Sumo tag.
pET28-HS-PAM2-L132A^{Upa1}	pUL202	Plasmid for the expression of the PAM2 motif of Upa1 (SQST**A**SPNASVFKPSRS) carrying a point mutation (Leucine 132 to Alanine) was fused to an N-terminal 6xHis-Sumo tag.

pET28-HS-PAM2-F139A^{Upa1}	pUL0173	Plasmid for the expression of the PAM2 motif of Upa1 (SQSTLSPNASV**A**KPSRS) carrying a point mutation (Phenylalanine 139 to Alanine) was fused to an N-terminal 6xHis-Sumo tag.
pET28-HS-3x PAM2^{Rrm4}	pUMa4676	Plasmid for the expression of the HS-3x PAM2^{Rrm4} in which amino acid residues from 235 to 271 of Rrm4 (ATRKVSAAAAEFRPSAAAFVPAGSMSPSAPSFDPYPA) were fused to an N-terminal 6xHis-Sumo tag.
pET28-HS-PAM2^{Rrm4}	pUMa4721	Plasmid for the expression of the HS-PAM2^{Rrm4} in which amino acid residues from 599 to 615 of Rrm4 (SLPKKDRALALFNPEFL) were were fused to an N-terminal 6xHis-Sumo tag.
HS-PAM2^{Pab1}	pUMa4722	Plasmid for the expression of the HS-PAM2^{Pab1} in which amino acid residues from 339 to 355 of Pab1 (ESYDDERLREEFAPFGA) were fused to an N-terminal 6xHis-Sumo tag.
pET28-H-EcEfTu	pUMa4700	Plasmid for the expression of the His-EfTu from *E. coli* in which amino acid residues from 1 to 394 of *E. coli* EfTu (ORF) were fused to an N-terminal 6xHis-tag and a Thrombin cleavage site.
pET28-HS-PAM2L^{EcEfTu}	pUMa4720	Plasmid for the expression of the HS- PAM2L^{EcEfTu} in which amino acid residues from 200 to 217 of *E. coli*-EfTu (IPEPERAIDKPFLLPIED) were fused to an N-terminal 6xHis-Sumo tag.
pET28-HS-PAM2L^{Taf7}	pUL0224	Plasmid for the expression of the HS- PAM2L^{Taf7} in which amino acid residues from 212 to 229 of Taf7 (GGGGKGFNIDDFIYPHGI) were fused to an N-terminal 6xHis-Sumo tag.
pET28-HS-PAM2L1^{Vps8}	pUL0225	Plasmid for the expression of the HS- PAM2L1^{Vps8} in which amino acid residues from 70 to 87 of Vps8 (HQDDDNDNDDDFVYDGID) were fused to an N-terminal 6xHis-Sumo tag.
pET28-HS-PAM2L2^{Vps8}	pUL0226	Plasmid for the expression of the HS- PAM2L2^{Vps8} in which amino acid residues from 173 to 190 of Vps8 (TDQAPSDDSFSFRYPHPL) were fused to an N-terminal 6xHis-Sumo tag.

6.10.7 Table S16. DNA oligonucleotides used in this study

Designation	Nucleotide sequence (5' --> 3')
AB45	CGGCCATATGGGCAGCAGCCATCATC
AB611	ATGCGAATTCGGTACCTTACGCAGAGTCGTTCTGTTG
AB967	ACAGAGAACAGATTGGTGGTGCTACGCGCAAGGTCAGCGC
AB968	GCCGGATCTCACTCGAGTTAAGCAGGGTAAGGGTCAAAGG
AB969	TGAGATCCGGCTGCTAACAAAGCCCGAAAGG
AB970	ACCACCAATCTGTTCTCTGTGAGCCTC
AB972	CTCACTCGAGTTAGAGAAACTCTGGATTGAAGAGCGCAAGGGCCCGATCCTTCTTTGG-CAGGCTACCACCAATCTGTTCTCTG
AB973	CTCACTCGAGTTAGGCACCAAAGGGAGCAAACTCCTCGCGAAGACGCTCGTCGTCGTAC-GACTCACCACCAATCTGTTCTCTG
CD214	GGTCTCGCCTGCATATGAGTCCCGAGGAGCAGAAGC
CD503	CATGCCATGGGAGCTAGCGCGGCCGCCAGCTCCGGTCTGCCTCTCC
CD504	CATGCCATGGGAGCTAGCGCGGCCGCCAGCTCCATCTCGCCTGGCGCTG
CD532	CATGCCATGGGGCATATGATGTCTAAAGAAAAGTTTGA
CD533	ATGCCTCGAGGAATTCTTAGCTCAGAACTTTTGCTAC
CD534	CTCACTCGAGTTAGTCTTCGATCGGCAGCAGGAACGGCTTGTCAATCGCAC-GCTCTGGTTCCGGAATACCACCAATCTGTTCTCTGTGAG
CD605	ATGCATCATATGCATCATCATCATCATCACACATTATCCACGCTTGCTGC
CD658	CTTTAAGAAGGAGATATACATATGGGCAGCAGCCATCATCATC
CD694	ACTGCACATATGCATCATCATCATCATCACAGTCCCGAGGAGCAGAAGC
CD695	GTTAGCAGCCGGATCTCACTCGAGTTAGTCGGCTCCTGGGTAGACAAAGTCATCACCAC-CAATCTGTTCTCTGTG
CD696	GTTAGCAGCCGGATCTCACTCGAGTTAGTACGAGTTCGGGTAGATGAATTCATCACCAC-CAATCTGTTCTCTGTG
CD808	GTGGTGGTGGTGGTGCTCGAG-TTAATCCGCGCCCGGCGCCACAAAATCATCCTGATCTTCTTCCTGATCCGCCGCTTCACCAC-CAATCTGTTCTCTG
CD809	GTGGTGGTGGTGGTGCTCGAGTTAATAGCTGTTCGGCGCAA-TAAATTCATCATCATCATCATCCGCCGCATCTTCATCACCACCAATCTGTTCTCTG
CD811	GTGGTGGTGGTGGTGCTCGAGTTAATCCGCGCCCGGATACACCG-CATCATCCTGATCTTCTTCCTGATCCGCCGCTTCACCACCAATCTGTTCTCTG
CD812	GTGGTGGTGGTGGTGCTCGAGTTAATAGCTGTTCGGA-TAAATCGCTTCATCATCATCATCATCCGCCGCATCTTCATCACCACCAATCTGTTCTCTG
CD813	GTGGTGGTGGTGGTGCTCGAGTTAATCCGCGCCCGCATA-CACAAAATCATCCTGATCTTCTTCCTGATCCGCCGCTTCACCACCAATCTGTTCTCTG
CD814	GTGGTGGTGGTGGTGCTCGAGTTAATAGCTGTTCGCATAAA-TAAATTCATCATCATCATCATCCGCCGCATCTTCATCACCACCAATCTGTTCTCTG
CD815	GTGGTGGTGGTGGTGCTCGAGTTAGCTGCGGCTCGGTTTCGCCAC-GCTCGCGTTCGGGCTCAGGGTGCTCTGGCTACCACCAATCTGTTCTCTG
EF001	GCAAAGCAGCGGCGGATCAGAAGCAGAAGCTGGGTGATCAGCTC
EF008	CTGCTTCTGATCCGCCGCTGCTTTGCCCTGAAGCGAGTCG
EF009	CTAGGCCGCGTTGGCCTCGAGTCACTTGTTCAGACC
EF010	GGTTCCGCGTGGATCCATGGGAGCTAGCGCGG

EF013	CTTCAAAAAGATCGCTACGTTCGGCGTCAAAGGCGCACCCAAGC
EF014	CGAACGTAGCGATCTTTTTGAAGAGCTGATCACCCAGC
EF015	GCACCCAAGCTCACCGGCCACTTGTTGGATTCCGAAGACTTGCGGGCG
EF016	CCAACAAGTGGCCGGTGAGCTTGGGTGCGCCTTTGACGCC
EF033	GCAGCGCACGATCAGAAGGCGAAGCTGGGTGATCAGCTCTTCAAAAAGATCC
EF034	ACCCAGCTTCGCCTTCTGATCGTGCGCTGCTTTGCC
EF035	CAGAAGCTGGGTGATCAGCTCGCCAAAAAGATCCGTACGTTCGGCGTCAAAGGC
EF036	ATCTTTTTGGCGAGCTGATCACCCAGCTTCTGCTTCTGATCG
EF037	GGCGTCAAAGGCGCACCCAAGCTCGCAATCCACTTGTTGGATTCCGAAGACTTGCG
EF038	CCAACAAGTGGATTGCGAGCTTGGGTGCGCCTTTGACGCC
EF107	GTGGTGGTGGTGGTGCTCGAGTTACAGCGGGTGCGGGTAACGGAAAGAGAAAGAG-TCGTCAGACGGCGCCTGGTCGGTACCACCAATCTGTTCTCTGTG
EF17	GTTAGCAGCCGGATCTCACTCGAGTTAGGATCGGGACGGCTTGAAGACGGAGGCGTTGGGA-GACGCGGTGCTTTGCGAACCACCAATCTGTTCTCTGTG
EF87	GTGGTGGTGGTGGTGCTCGAGTTAGATACCGTGCGGGTAGATGAAGTCGTCGATGTTGAAAC-CTTTACCACCACCACCACCACCAATCTGTTCTCTGTG
EF92	GTGGTGGTGGTGGTGCTCGAGTTAGTCGATACCGTCGTAAACGAAGTCGTCGTCGTT-GTCGTTGTCGTCGTCCTGGTGACCACCAATCTGTTCTCTGTG
MB696	TCGACTCGAGTCACTTGTTCAGACCTGCAGC

6.10.8 Table S17. Highly enriched interaction partners of G-MLLE3-5H^{Rrm4} in *E. coli* identified by the LC-MS/MS

S.No	UniprotKB Number	Gene Symbol	# PSMs	Orthologues in *U. maydis*
1	A0A140NCI6	*tufB*	524	UMAG_00138, UMAG_00924
2	A0A140NFM6	*ftsZ*	83	UMAG_01221, UMAG_05540
3	A0A140NF01	*rho*	56	UMAG_10397, UMAG_10213
4	A0A140NEQ0	*clpX*	46	UMAG_11194
5	A0A140N8E1	*pgk*	41	UMAG_04871
6	A0A140N655	*proV*	34	UMAG_10528
7	A0A140NAR5	*ychF*	29	UMAG_06117
8	A0A140N4P3	*ispG*	28	UMAG_00115
10	A0A140NEK1	*purA*	25	UMAG_03851
11	A0A140N627	*metK*	21	UMAG_05019
12	A0A140N6G0	*eno*	20	UMAG_03356
13	A0A140NHG8	*malK*	18	UMAG_06461
14	A0A140N9D5	*iscS*	17	UMAG_05776
15	A0A140N7J1	*rplB; CSK29544_RS06430; AL524_RS00875*	15	UMAG_04511, UMAG_11233
16	A0A140NB96		13	UMAG_11855
17	A0A140N6P6	*serA*	12	UMAG_01233
18	A0A140ND70	*acrA*	11	UMAG_04472
19	A0A140N5W3	*yfbQ*	11	UMAG_10659
20	A0A140NAX3	*ndh*	11	UMAG_02164
21	A0A140N7U4	*gnd*	10	UMAG_02577

6.10.9 Table S18. *De novo* predicted PAM2L candidates in *U. maydis*

UMAG	Protein Name
UMAG_15064	Vps8, early endosome specific subunit of CORVET complex
UMAG_10255	Orthologue of transcription corepressor in Yeast, transcription elongation regulator 1 in humans
UMAG_02933	RNA-directed RNA polymerase in fungi
UMAG_03844	Brix domain, rRNA binding, SSF1-Nucleolar protein related
UMAG_11924	CWF19-RNA lariat debranching enzyme activator activity
UMAG_04936	FCP1 domain-containing protein, that dephosphorylates the C-terminal domain (CTD) of RNA polymerase II.
UMAG_01597	Transcriptional activator HAP2
UMAG_02262	Transcription factor CBF
UMAG_10591	Ankyrin repeats, RING domain containing protein
UMAG_10091	Serine carboxypeptidase
UMAG_05811	GH16 domain-containing protein, glycosyl hydrolases family 16
UMAG_02807	Methyl-itaconate delta2-delta3-isomerase
UMAG_04517	DUF3835 domain protein in fungi, Human Calsequestrin, Nucleolin related
UMAG_15085	Peptide-N4 asparagine amidase A. mitochondrial 5-demethoxyubiquinone hydroxylase,
UMAG_00448	Methyl-accepting transducer domain-containing protein in fungi
UMAG_03381	Endo/exonuclease/phosphatase domain-containing protein
UMAG_00557	alpha-mannosidase
UMAG_04460	DNA-directed RNA polymerases I, II, and III subunit RPABC3
UMAG_05361	Laccase I (multicopper oxidases)in fungi
UMAG_01466	Cytochrome b5 heme-binding domain-containing protein
UMAG_10620	Taf7, Ptr6, TAFII55_N domain, RNA polymerase II preinitiation complex assembly

6.10.10 Table S19. *De novo* predicted PAM2 candidates in *U. maydis*

UMAG	Protein Name
UMAG_12254	sepA-related Formin
UMAG_02525	E3 ubiquitin ligase
UMAG_04080	DEAD-box ATP-dependent RNA helicase DED1
UMAG_03624	MMS19 nucleotide excision repair protein
UMAG_10294	BTB/POZ domain-containing protein
UMAG_03901	Protein kinase
UMAG_15011	Serine / Threonine-protein kinase
UMAG_01327	Diphthamide biosynthesis protein 4
UMAG_11335	Mediator of RNA polymerase II transcription subunit 19
UMAG_03398	Beta-lactamase domain-containing protein
UMAG_03636	SWIRM domain-containing protein
UMAG_06478	Alpha/beta-hydrolase

7 Author contributions

Part of this dissertation (Chapter II: text and figures) was already published (DEVAN *et al.* 2022) and reproduced under the terms of the Creative Commons Attribution 4.0 International License (CC BY).
No changes were made to the published materials besides the position of figures or formatting of text and tables. Results or interpretation of data were not changed in this dissertation concerning the published materials. The source and co-authors of the figures used/reused in this dissertation are listed below.

Figure No.	Source / Co-author	License
1	(DAS *et al.* 2021)	Springer Nature: 5412120334416
2	(MUNTJES *et al.* 2021)	CC BY
3	(LANVER *et al.* 2017)	Springer Nature: 5376580465976
4	(POHLMANN *et al.* 2015)	CC BY
6-11, S1-10	(DEVAN *et al.* 2022)	CC BY
6C, 7B	Stephan Schott-Verdugo	
7D-E, S2D-E	Jens Reiners	
7C	Sander Smiths	
S2C	Eymen Hachani	
9B-E, 10, 11, S7A-C, S8, S9A, S10	Kira Muntjes	
S12C, S13C	Violetta Applegate, Stefanie Galle (CSS, HHU)	
15D,E, 16F, S13A,B, 19D	Sainath Shanmugasundaram	
20A-C	(MULLER *et al.* 2019)	RightsLink/Elsevier 5430730645089
All the figures except those listed above.	Senthil Kumar Devan	

Further contributions in publications that are not part of this thesis but related to the PhD topic:

Müntjes, K., **S. K. Devan**, A. S. Reichert and M. Feldbrügge, 2021 Linking transport and translation of mRNAs with endosomes and mitochondria. EMBO Rep accepted for publication.

Jankowski, S., T. Pohlmann, S. Baumann, K. M. Müntjes, **S. K. Devan** et al., 2019 The multi PAM2 protein Upa2 functions as novel core component of endosomal mRNA transport. EMBO Rep. 24: e47381.

8 References

Abts, A., M. Montalban-Lopez, O. P. Kuipers, S. H. Smits and L. Schmitt, 2013 NisC binds the FxLx motif of the nisin leader peptide. Biochemistry 52**:** 5387-5395.

Afonine, P. V., R. W. Grosse-Kunstleve, N. Echols, J. J. Headd, N. W. Moriarty *et al.*, 2012 Towards automated crystallographic structure refinement with phenix.refine. Acta Crystallogr. D. Biol. Crystallogr. 68**:** 352-367.

Albrecht, M., and T. Lengauer, 2004a Survey on the PABC recognition motif PAM2. Biochem Biophys Res Commun 316**:** 129-138.

Andersen, G. R., S. Thirup, L. L. Spremulli and J. Nyborg, 2000 High resolution crystal structure of bovine mitochondrial EF-Tu in complex with GDP. J Mol Biol 297**:** 421-436.

Andreassi, C., and A. Riccio, 2009 To localize or not to localize: mRNA fate is in 3'UTR ends. Trends Cell Biol 19**:** 465-474.

Ardito, F., M. Giuliani, D. Perrone, G. Troiano and L. Lo Muzio, 2017 The crucial role of protein phosphorylation in cell signaling and its use as targeted therapy (Review). Int J Mol Med 40**:** 271-280.

B. J. Saville, M. E. D. a. C. E. D., 2012 Investigating Host Induced Meiosis in a Fungal Plant Pathogen, pp. in *Meiosis - Molecular Mechanisms and Cytogenetic Diversity*, edited by D. A. Swan. InTech.

Baek, M., F. DiMaio, I. Anishchenko, J. Dauparas, S. Ovchinnikov *et al.*, 2021 Accurate prediction of protein structures and interactions using a three-track neural network. Science 373**:** 871-876.

Balderhaar, H. J., and C. Ungermann, 2013 CORVET and HOPS tethering complexes - coordinators of endosome and lysosome fusion. J Cell Sci 126**:** 1307-1316.

Basyuk, E., T. Galli, M. Mougel, J. M. Blanchard, M. Sitbon *et al.*, 2003 Retroviral genomic RNAs are transported to the plasma membrane by endosomal vesicles. Dev Cell 5**:** 161-174.

Basyuk, E., F. Rage and E. Bertrand, 2021 RNA transport from transcription to localized translation: a single molecule perspective. RNA Biol 18**:** 1221-1237.

Baumann, S., A. Komissarov, M. Gili, V. Ruprecht, S. Wieser *et al.*, 2020 A reconstituted mammalian APC-kinesin complex selectively transports defined packages of axonal mRNAs. Sci Adv 6**:** eaaz1588.

Baumann, S., J. Konig, J. Koepke and M. Feldbrugge, 2014a Endosomal transport of septin mRNA and protein indicates local translation on endosomes and is required for correct septin filamentation. EMBO Rep 15**:** 94-102.

Baumann, S., T. Pohlmann, M. Jungbluth, A. Brachmann and M. Feldbrugge, 2012a Kinesin-3 and dynein mediate microtubule-dependent co-transport of mRNPs and endosomes. J Cell Sci 125**:** 2740-2752.

Baumann, S., N. Takeshita, N. Grun, R. Fischer and M. Feldbrugge, 2015 Live cell imaging of endosomal trafficking in fungi. Methods Mol Biol 1270**:** 347-363.

Baumann, S., S. Zander, S. Weidtkamp-Peters and M. Feldbrügge, 2016 Live cell imaging of septin dynamics in *Ustilago maydis*, pp. 143-159 in *Methods in Cell Biol: Septins*, edited by A. S. Gladfelter. Elsevier Inc.

Becht, P., J. Konig and M. Feldbrugge, 2006 The RNA-binding protein Rrm4 is essential for polarity in Ustilago maydis and shuttles along microtubules. J Cell Sci 119**:** 4964-4973.

Becht, P., E. Vollmeister and M. Feldbrugge, 2005 Role for RNA-binding proteins implicated in pathogenic development of Ustilago maydis. Eukaryot Cell 4**:** 121-133.

Benhalevy, D., I. Biran, E. S. Bochkareva, R. Sorek and E. Bibi, 2017 Evidence for a cytoplasmic pool of ribosome-free mRNAs encoding inner membrane proteins in Escherichia coli. PLoS One 12**:** e0183862.

Berleth, T., M. Burri, G. Thoma, D. Bopp, S. Richstein *et al.*, 1988 The role of localization of bicoid RNA in organizing the anterior pattern of the Drosophila embryo. EMBO J 7**:** 1749-1756.

Bethard, J. R., H. Zheng, L. Roberts and S. T. Eblen, 2011 Identification of phosphorylation sites on the E3 ubiquitin ligase UBR5/EDD. J Proteomics 75**:** 603-609.

Béthune, J., R. P. Jansen, M. Feldbrügge and K. Zarnack, 2019 Membrane-associated RNA-binding proteins orchestrate organelle-coupled translation Trends Cell Biol. 29**:** 178-188.

Biswas, J., L. Nunez, S. Das, Y. J. Yoon, C. Eliscovich *et al.*, 2019 Zipcode Binding Protein 1 (ZBP1; IGF2BP1): A Model for Sequence-Specific RNA Regulation. Cold Spring Harb Symp Quant Biol 84**:** 1-10.

Brachmann, A., J. König, C. Julius and M. Feldbrügge, 2004 A reverse genetic approach for generating gene replacement mutants in *Ustilago maydis*. Mol. Gen. Genom. 272**:** 216-226.

Brachmann, A., G. Weinzierl, J. Kamper and R. Kahmann, 2001 Identification of genes in the bW/bE regulatory cascade in Ustilago maydis. Mol Microbiol 42**:** 1047-1063.

Bravo, J., L. Aguilar-Henonin, G. Olmedo and P. Guzman, 2005 Four distinct classes of proteins as interaction partners of the PABC domain of Arabidopsis thaliana Poly(A)-binding proteins. Mol Genet Genomics 272**:** 651-665.

Brefort, T., G. Doehlemann, A. Mendoza-Mendoza, S. Reissmann, A. Djamei *et al.*, 2009 Ustilago maydis as a Pathogen. Annu Rev Phytopathol 47**:** 423-445.

Buskila, A. A., S. Kannaiah and O. Amster-Choder, 2014 RNA localization in bacteria. RNA Biol 11**:** 1051-1060.

Caffrey, D. R., P. H. Dana, V. Mathur, M. Ocano, E. J. Hong *et al.*, 2007 PFAAT version 2.0: a tool for editing, annotating, and analyzing multiple sequence alignments. BMC Bioinformatics 8**:** 381.

Caldon, C. E., and P. E. March, 2003 Function of the universally conserved bacterial GTPases. Curr Opin Microbiol 6**:** 135-139.

Castellana, M., S. Hsin-Jung Li and N. S. Wingreen, 2016 Spatial organization of bacterial transcription and translation. Proc Natl Acad Sci U S A 113**:** 9286-9291.

Castello, A., B. Fischer, K. Eichelbaum, R. Horos, B. M. Beckmann *et al.*, 2012 Insights into RNA biology from an atlas of mammalian mRNA-binding proteins. Cell 149**:** 1393-1406.

Chen, L., J. G. Dumelie, X. Li, M. H. Cheng, Z. Yang *et al.*, 2014 Global regulation of mRNA translation and stability in the early Drosophila embryo by the Smaug RNA-binding protein. Genome Biol 15**:** R4.

Choi, C. Y., M. T. Vo, J. Nicholas and Y. B. Choi, 2022 Autophagy-competent mitochondrial translation elongation factor TUFM inhibits caspase-8-mediated apoptosis. Cell Death Differ 29**:** 451-464.

Cioni, J. M., J. Q. Lin, A. V. Holtermann, M. Koppers, M. A. H. Jakobs *et al.*, 2019 Late Endosomes Act as mRNA Translation Platforms and Sustain Mitochondria in Axons. Cell 176**:** 56-72 e15.

Cohen, B., T. Altman, A. Golani-Armon, A. F. Savulescu, A. Ibraheem *et al.*, 2022 Co-transport of the nuclear-encoded Cox7c mRNA with mitochondria along axons occurs through a coding-region-dependent mechanism. J Cell Sci 135.

Das, S., M. Vera, V. Gandin, R. H. Singer and E. Tutucci, 2021a Intracellular mRNA transport and localized translation. Nat. Rev. Mol. Cell. Biol. 22**:** 483-504.

de Melo Neto, O. P., T. D. C. da Costa Lima, K. C. Merlo, T. P. Romao, P. O. Rocha *et al.*, 2018 Phosphorylation and interactions associated with the control of the Leishmania Poly-A Binding Protein 1 (PABP1) function during translation initiation. RNA Biol 15**:** 739-755.

Deo, R. C., N. Sonenberg and S. K. Burley, 2001a X-ray structure of the human hyperplastic discs protein: an ortholog of the C-terminal domain of poly(A)-binding protein. Proc Natl Acad Sci U S A 98**:** 4414-4419.

Devan, S. K., S. Schott-Verdugo, K. Muntjes, L. Bismar, J. Reiners *et al.*, 2022 A MademoiseLLE domain binding platform links the key RNA transporter to endosomes. PLoS Genet 18**:** e1010269.

Di Nottia, M., A. Montanari, D. Verrigni, R. Oliva, A. Torraco *et al.*, 2017 Novel mutation in mitochondrial Elongation Factor EF-Tu associated to dysplastic leukoencephalopathy and defective mitochondrial DNA translation. Biochim Biophys Acta Mol Basis Dis 1863**:** 961-967.

Djamei, A., and R. Kahmann, 2012 Ustilago maydis: dissecting the molecular interface between pathogen and plant. PLoS Pathog 8**:** e1002955.

Edelmann, F. T., A. Schlundt, R. G. Heym, A. Jenner, A. Niedner-Boblenz *et al.*, 2017 Molecular architecture and dynamics of ASH1 mRNA recognition by its mRNA-transport complex. Nat Struct Mol Biol 24**:** 152-161.

Elson, S. L., S. M. Noble, N. V. Solis, S. G. Filler and A. D. Johnson, 2009 An RNA transport system in Candida albicans regulates hyphal morphology and invasive growth. PLoS Genet 5**:** e1000664.

Emsley, P., and K. Cowtan, 2004 Coot: model-building tools for molecular graphics. Acta Crystallogr. D. Biol. Crystallogr. 60**:** 2126-2132.

Engel, K. L., A. Arora, R. Goering, H. G. Lo and J. M. Taliaferro, 2020 Mechanisms and consequences of subcellular RNA localization across diverse cell types. Traffic 21**:** 404-418.

Epp, N., and C. Ungermann, 2013 The N-terminal domains of Vps3 and Vps8 are critical for localization and function of the CORVET tethering complex on endosomes. PLoS One 8**:** e67307.

Fei, J., and C. M. Sharma, 2018 RNA Localization in Bacteria. Microbiol Spectr 6.

Feldbrugge, M., R. Kellner and K. Schipper, 2013 The biotechnological use and potential of plant pathogenic smut fungi. Appl Microbiol Biotechnol 97**:** 3253-3265.

Fernandopulle, M. S., J. Lippincott-Schwartz and M. E. Ward, 2021 RNA transport and local translation in neurodevelopmental and neurodegenerative disease. Nat Neurosci 24**:** 622-632.

Franke, D., and D. I. Svergun, 2009 DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering. J. Appl. Crystallogr. 42**:** 342-346.

Fukuda, Y., M. F. Pazyra-Murphy, E. S. Silagi, O. E. Tasdemir-Yilmaz, Y. Li *et al.*, 2021 Binding and transport of SFPQ-RNA granules by KIF5A/KLC1 motors promotes axon survival. J Cell Biol 220.

Girodat, D., S. C. Blanchard, H. J. Wieden and K. Y. Sanbonmatsu, 2020 Elongation Factor Tu Switch I Element is a Gate for Aminoacyl-tRNA Selection. J Mol Biol 432**:** 3064-3077.

Glock, C., M. Heumuller and E. M. Schuman, 2017 mRNA transport & local translation in neurons. Curr Opin Neurobiol 45**:** 169-177.

Gohre, V., E. Vollmeister, M. Bolker and M. Feldbrugge, 2012 Microtubule-dependent membrane dynamics in Ustilago maydis: Trafficking and function of Rab5a-positive endosomes. Commun Integr Biol 5**:** 485-490.

Goss, D. J., and F. E. Kleiman, 2013 Poly(A) binding proteins: are they all created equal? Wiley Interdiscip Rev RNA 4**:** 167-179.

Gould, G. W., and J. Lippincott-Schwartz, 2009 New roles for endosomes: from vesicular carriers to multi-purpose platforms. Nat Rev Mol Cell Biol 10**:** 287-292.

Grimm, C., J. P. Pelz, C. Schneider, K. Schaffler and U. Fischer, 2020a Crystal Structure of a Variant PAM2 Motif of LARP4B Bound to the MLLE Domain of PABPC1. Biomolecules 10.

Haag, C., B. Steuten and M. Feldbrugge, 2015a Membrane-Coupled mRNA Trafficking in Fungi. Annu Rev Microbiol 69**:** 265-281.

Harbauer, A. B., J. T. Hees, S. Wanderoy, I. Segura, W. Gibbs *et al.*, 2022 Neuronal mitochondria transport Pink1 mRNA via synaptojanin 2 to support local mitophagy. Neuron 110**:** 1516-1531 e1519.

Harbauer, A. B., and T. L. Schwarz, 2022 Mitochondrial hitch-hiking of Pink1 mRNA supports axonal mitophagy. Autophagy**:** 1-2.

Harvey, K. L., V. M. Jarocki, I. G. Charles and S. P. Djordjevic, 2019 The Diverse Functional Roles of Elongation Factor Tu (EF-Tu) in Microbial Pathogenesis. Front Microbiol 10**:** 2351.

Higuchi, Y., P. Ashwin, Y. Roger and G. Steinberg, 2014 Early endosome motility spatially organizes polysome distribution. J Cell Biol 204**:** 343-357.

Huang, K. L., A. B. Chadee, C. Y. Chen, Y. Zhang and A. B. Shyu, 2013 Phosphorylation at intrinsically disordered regions of PAM2 motif-containing proteins modulates their interactions with PABPC1 and influences mRNA fate. RNA 19**:** 295-305.

Hughes, D., 2017 Elongation Factors: Translation in *Reference Module in Life Sciences*. Elsevier.

Jankowski, S., T. Pohlmann, S. Baumann, K. Muntjes, S. K. Devan *et al.*, 2019a The multi PAM2 protein Upa2 functions as novel core component of endosomal mRNA transport. EMBO Rep 20**:** e47381.

Jimenez-Lopez, D., J. Bravo and P. Guzman, 2015 Evolutionary history exposes radical diversification among classes of interaction partners of the MLLE domain of plant poly(A)-binding proteins. BMC Evol Biol 15**:** 195.

Jinek, M., M. R. Fabian, S. M. Coyle, N. Sonenberg and J. A. Doudna, 2010a Structural insights into the human GW182-PABC interaction in microRNA-mediated deadenylation. Nat. Struct. Mol. Biol. 17**:** 238-240.

Jonker, C. T. H., R. Galmes, T. Veenendaal, C. Ten Brink, R. E. N. van der Welle *et al.*, 2018 Vps3 and Vps8 control integrin trafficking from early to recycling endosomes and regulate integrin-dependent functions. Nat Commun 9**:** 792.

Jumper, J., R. Evans, A. Pritzel, T. Green, M. Figurnov *et al.*, 2021 Highly accurate protein structure prediction with AlphaFold. Nature 596**:** 583-589.

Jung, H., C. G. Gkogkas, N. Sonenberg and C. E. Holt, 2014 Remote control of gene function by local translation. Cell 157**:** 26-40.

Kabsch, W., 2014 Processing of X-ray snapshots from crystals in random orientations. Acta Crystallogr. D. Biol. Crystallogr. 70**:** 2204-2216.

Kamper, J., R. Kahmann, M. Bolker, L. J. Ma, T. Brefort *et al.*, 2006 Insights from the genome of the biotrophic fungal plant pathogen Ustilago maydis. Nature 444**:** 97-101.

Kamper, J., M. Reichmann, T. Romeis, M. Bolker and R. Kahmann, 1995 Multiallelic recognition: nonself-dependent dimerization of the bE and bW homeodomain proteins in Ustilago maydis. Cell 81**:** 73-83.

Kikhney, A. G., C. R. Borges, D. S. Molodenskiy, C. M. Jeffries and D. I. Svergun, 2020 SASBDB: Towards an automatically curated and validated repository for biological scattering data. Protein Sci. 29**:** 66-75.

Koepke, J., F. Kaffarnik, C. Haag, K. Zarnack, N. M. Luscombe *et al.*, 2011 The RNA-binding protein Rrm4 is essential for efficient secretion of endochitinase Cts1. Mol Cell Proteomics 10**:** M111 011213.

Konarev, P. V., V. V. Volkov, A. V. Sokolova, M. H. J. Koch and D. I. Svergun, 2003 PRIMUS: a Windows PC-based system for small-angle scattering data analysis. J. Appl. Crystallogr. 36**:** 1277-1282.

Konig, J., S. Baumann, J. Koepke, T. Pohlmann, K. Zarnack *et al.*, 2009 The fungal RNA-binding protein Rrm4 mediates long-distance transport of ubi1 and rho3 mRNAs. EMBO J 28**:** 1855-1866.

Kozlov, G., G. De Crescenzo, N. S. Lim, N. Siddiqui, D. Fantus *et al.*, 2004 Structural basis of ligand recognition by PABC, a highly specific peptide-binding domain found in poly(A)-binding protein and a HECT ubiquitin ligase. EMBO J 23**:** 272-281.

Kozlov, G., and K. Gehring, 2010 Molecular basis of eRF3 recognition by the MLLE domain of poly(A)-binding protein. PLoS One 5**:** e10169.

Kozlov, G., M. Menade, A. Rosenauer, L. Nguyen and K. Gehring, 2010 Molecular determinants of PAM2 recognition by the MLLE domain of poly(A)-binding protein. J Mol Biol 397**:** 397-407.

Kozlov, G., N. Safaee, A. Rosenauer and K. Gehring, 2010 Structural basis of binding of P-body-associated proteins GW182 and ataxin-2 by the Mlle domain of poly(A)-binding protein. J Biol Chem 285**:** 13599-13606.

Kozlov, G., N. Siddiqui, S. Coillet-Matillon, J. F. Trempe, I. Ekiel *et al.*, 2002 Solution structure of the orphan PABC domain from Saccharomyces cerevisiae poly(A)-binding protein. J Biol Chem 277**:** 22822-22828.

Kozlov, G., J. F. Trempe, K. Khaleghpour, A. Kahvejian, I. Ekiel *et al.*, 2001 Structure and function of the C-terminal PABC domain of human poly(A)-binding protein. Proc Natl Acad Sci U S A 98**:** 4409-4413.

Krissinel, E., and K. Henrick, 2007 Inference of macromolecular assemblies from crystalline state. J Mol Biol 372**:** 774-797.

Kuhn, U., and T. Pieler, 1996 Xenopus poly(A) binding protein: functional domains in RNA binding and protein-protein interaction. J Mol Biol 256**:** 20-30.

Kuhn, U., and E. Wahle, 2004 Structure and function of poly(A) binding proteins. Biochim Biophys Acta 1678**:** 67-84.

Lanver, D., M. Tollot, G. Schweizer, L. Lo Presti, S. Reissmann *et al.*, 2017 Ustilago maydis effectors and their impact on virulence. Nat Rev Microbiol 15**:** 409-421.

Larkin, M. A., G. Blackshields, N. P. Brown, R. Chenna, P. A. McGettigan *et al.*, 2007 Clustal W and Clustal X version 2.0. Bioinformatics 23**:** 2947-2948.

Laskowski, R. A., and M. B. Swindells, 2011 LigPlot+: multiple ligand-protein interaction diagrams for drug discovery. J Chem Inf Model 51**:** 2778-2786.

Lavigne, A. C., G. Mengus, M. May, V. Dubrovskaya, L. Tora *et al.*, 1996 Multiple interactions between hTAFII55 and other TFIID subunits. Requirements for the formation of stable ternary complexes between hTAFII55 and the TATA-binding protein. J Biol Chem 271**:** 19774-19780.

Lawrence, J. B., and R. H. Singer, 1986 Intracellular localization of messenger RNAs for cytoskeletal proteins. Cell 45**:** 407-415.

Lee, J., F. Hilgers, A. Loeschke, K. E. Jaeger and M. Feldbrugge, 2020 Ustilago maydis Serves as a Novel Production Host for the Synthesis of Plant and Fungal Sesquiterpenoids. Front Microbiol 11**:** 1655.

Lehmann, M., M. P. Milev, L. Abrahamyan, X. J. Yao, N. Pante *et al.*, 2009 Intracellular transport of human immunodeficiency virus type 1 genomic RNA and viral production are dependent on dynein motor function and late endosome positioning. J Biol Chem 284**:** 14572-14585.

Li, L., J. Yu and S. J. Ji, 2021 Axonal mRNA localization and translation: local events with broad roles. Cell Mol Life Sci 78**:** 7379-7395.

Liao, Y. C., M. S. Fernandopulle, G. Wang, H. Choi, L. Hao *et al.*, 2019 RNA Granules Hitchhike on Lysosomes for Long-Distance Transport, Using Annexin A11 as a Molecular Tether. Cell 179**:** 147-164 e120.

Lim, N. S., G. Kozlov, T. C. Chang, O. Groover, N. Siddiqui *et al.*, 2006 Comparative peptide binding studies of the PABC domains from the ubiquitin-protein isopeptide ligase HYD and poly(A)-binding protein. Implications for HYD function. J. Biol. Chem. 281**:** 14376-14382.

Long, R. M., R. H. Singer, X. Meng, I. Gonzalez, K. Nasmyth *et al.*, 1997 Mating type switching in yeast controlled by asymmetric localization of ASH1 mRNA. Science 277**:** 383-387.

Lopez-Berges, M. S., H. N. Arst, M. Pinar and M. A. Penalva, 2017 Genetic studies on the physiological role of CORVET in Aspergillus nidulans. FEMS Microbiol Lett 364.

Lorincz, P., Z. Lakatos, A. Varga, T. Maruzs, Z. Simon-Vecsei *et al.*, 2016 MiniCORVET is a Vps8-containing early endosomal tether in Drosophila. Elife 5.

Loubradou, G., A. Brachmann, M. Feldbrügge and R. Kahmann, 2001 A homologue of the transcriptional repressor Ssn6p antagonizes cAMP signalling in *Ustilago maydis.* Mol. Microbiol. 40**:** 719-730.

Lu, W., M. Lakonishok, A. S. Serpinskaya, D. Kirchenbuechler, S. C. Ling *et al.*, 2018 Ooplasmic flow cooperates with transport and anchorage in Drosophila oocyte posterior determination. J Cell Biol 217**:** 3497-3511.

MacDonald, C., S. Winistorfer, R. M. Pope, M. E. Wright and R. C. Piper, 2017 Enzyme reversal to explore the function of yeast E3 ubiquitin-ligases. Traffic 18**:** 465-484.

Manalastas-Cantos, K., P. V. Konarev, N. R. Hajizadeh, A. G. Kikhney, M. V. Petoukhov *et al.*, 2021 ATSAS 3.0: expanded functionality and new tools for small-angle scattering data analysis. J. Appl. Crystallogr. 54**:** 343-355.

Mangus, D. A., N. Amrani and A. Jacobson, 1998 Pbp1p, a factor interacting with *Saccharomyces cerevisiae* poly(A)-binding protein, regulates polyadenylation. Mol. Cell. Biol. 18**:** 7383-7396.

Mangus, D. A., M. C. Evans and A. Jacobson, 2003 Poly(A)-binding proteins: multifunctional scaffolds for the post-transcriptional control of gene expression. Genome Biol. 4**:** 223.

Marat, A. L., and V. Haucke, 2016 Phosphatidylinositol 3-phosphates-at the interface between cell signalling and membrane traffic. EMBO J 35**:** 561-579.

Matei, A., and G. Doehlemann, 2016 Cell biology of corn smut disease-Ustilago maydis as a model for biotrophic interactions. Curr Opin Microbiol 34**:** 60-66.

Mattijssen, S., G. Kozlov, B. D. Fonseca, K. Gehring and R. J. Maraia, 2021 LARP1 and LARP4: up close with PABP for mRNA 3' poly(A) protection and stabilization. RNA Biol. 18**:** 259-274.

McBride, A. E., 2017 Messenger RNA transport in the opportunistic fungal pathogen Candida albicans. Curr Genet 63**:** 989-995.

Melo, E. O., R. Dhalia, C. Martins de Sa, N. Standart and O. P. de Melo Neto, 2003 Identification of a C-terminal poly(A)-binding protein (PABP)-PABP interaction domain: role in cooperative binding to poly (A) and efficient cap distal translational repression. J Biol Chem 278**:** 46357-46368.

Michael, M., and M. Parsons, 2020 New perspectives on integrin-dependent adhesions. Curr Opin Cell Biol 63**:** 31-37.

Mirdita, M., K. Schutze, Y. Moriwaki, L. Heo, S. Ovchinnikov *et al.*, 2022 ColabFold: making protein folding accessible to all. Nat Methods 19**:** 679-682.

Mofatteh, M., and S. L. Bullock, 2017 SnapShot: Subcellular mRNA localization. Cell 169**:** 178.e171.

Morse, J. C., D. Girodat, B. J. Burnett, M. Holm, R. B. Altman *et al.*, 2020 Elongation factor-Tu can repetitively engage aminoacyl-tRNA within the ribosome during the proofreading stage of tRNA selection. Proc Natl Acad Sci U S A 117**:** 3610-3620.

Muller, J., T. Pohlmann and M. Feldbrugge, 2019 Core components of endosomal mRNA transport are evolutionarily conserved in fungi. Fungal Genet Biol 126**:** 12-16.

Mulnaes, D., and H. Gohlke, 2018 TopScore: using deep neural networks and large diverse data sets for accurate protein model quality assessment. J. Chem. Theory Comput. 14**:** 6117-6126.

Mulnaes, D., F. Koenig and H. Gohlke, 2021 TopSuite Web Server: A Meta-Suite for Deep-Learning-Based Protein Structure and Quality Prediction. J Chem Inf Model 61**:** 548-553.

Mulnaes, D., N. Porta, R. Clemens, I. Apanasenko, J. Reiners *et al.*, 2020 TopModel: template-based protein structure prediction at low sequence identity using top-down consensus and deep neural networks. J. Chem. Theory Comput. 16**:** 1953-1967.

Munoz-Escobar, J., E. Matta-Camacho, G. Kozlov and K. Gehring, 2015a The MLLE domain of the ubiquitin ligase UBR5 binds to its catalytic domain to regulate substrate binding. J Biol Chem 290**:** 22841-22850.

Muntjes, K., S. K. Devan, A. S. Reichert and M. Feldbrugge, 2021 Linking transport and translation of mRNAs with endosomes and mitochondria. EMBO Rep 22**:** e52445.

Müntjes, K., M. Philipp, L. Husemann, N. Heucken, S. Weidtkamp-Peters *et al.*, 2020 Establishing polycistronic expression in the model microorganism *Ustilago maydis*. Front. Microbiol. 11**:** 1384.

Niessing, D., R. P. Jansen, T. Pohlmann and M. Feldbrugge, 2018a mRNA transport in fungal top models. Wiley Interdiscip Rev RNA 9.

Olgeiser, L., C. Haag, S. Boerner, J. Ule, A. Busch *et al.*, 2019a The key protein of endosomal mRNP transport Rrm4 binds translational landmark sites of cargo mRNAs. EMBO Rep 20.

Osawa, M., N. Hosoda, T. Nakanishi, N. Uchida, T. Kimura *et al.*, 2012 Biological role of the two overlapping poly(A)-binding protein interacting motifs 2 (PAM2) of eukaryotic releasing factor eRF3 in mRNA decay. RNA 18**:** 1957-1967.

Perini, E. D., R. Schaefer, M. Stoter, Y. Kalaidzidis and M. Zerial, 2014 Mammalian CORVET is required for fusion and conversion of distinct early endosome subpopulations. Traffic 15**:** 1366-1389.

Pernot, P., A. Round, R. Barrett, A. De Maria Antolinos, A. Gobbo *et al.*, 2013 Upgraded ESRF BM29 beamline for SAXS on macromolecules in solution. J. Synchrotron. Radiat. 20**:** 660-664.

Pohlmann, T., 2013 Die Rolle von PAM2-Proteinen während des endosomalen mRNA-Transports in *Ustilago maydis*, pp. in *Institut für Mikrobiologie*. Heinrich-Heine-Universität Düsseldorf.

Pohlmann, T., S. Baumann, C. Haag, M. Albrecht and M. Feldbrugge, 2015a A FYVE zinc finger domain protein specifically links mRNA transport to endosome trafficking. Elife 4.

Popovic, D., W. Nijenhuis, L. C. Kapitein and L. Pelkmans, 2020 Co-translational targeting of transcripts to endosomes. bioRxiv**:** 2020.2007.2017.208652.

Quentin, D., J. S. Schuhmacher, B. U. Klink, J. Lauer, T. R. Shaikh *et al.*, 2021a Structure of the human FERRY Rab5 effector complex. bioRxiv**:** 2021.2006.2021.449265.

Rebagliati, M. R., D. L. Weeks, R. P. Harvey and D. A. Melton, 1985 Identification and cloning of localized maternal RNAs from Xenopus eggs. Cell 42**:** 769-777.

Reindl, M., S. Hansch, S. Weidtkamp-Peters and K. Schipper, 2019 A Potential Lock-Type Mechanism for Unconventional Secretion in Fungi. Int J Mol Sci 20.

Sarkari, P., M. Reindl, J. Stock, O. Muller, R. Kahmann *et al.*, 2014 Improved expression of single-chain antibodies in Ustilago maydis. J Biotechnol 191**:** 165-175.

Sawazaki, R., S. Imai, M. Yokogawa, N. Hosoda, S. I. Hoshino *et al.*, 2018 Characterization of the multimeric structure of poly(A)-binding protein on a poly(A) tail. Sci. Rep. 8**:** 1455.

Schafer, I. B., M. Yamashita, J. M. Schuller, S. Schussler, P. Reichelt *et al.*, 2019 Molecular Basis for poly(A) RNP Architecture and Recognition by the Pan2-Pan3 Deadenylase. Cell 177**:** 1619-1631 e1621.

Scheffner, M., and S. Kumar, 2014 Mammalian HECT ubiquitin-protein ligases: biological and pathophysiological aspects. Biochim Biophys Acta 1843**:** 61-74.

Schneider, K., T. Farr, N. Pinter, K. Schmitt, O. Valerius *et al.*, 2022 The Nma1 protein promotes long distance transport mediated by early endosomes in Ustilago maydis. Mol Microbiol 117**:** 334-352.

Schuhmacher, J. S., S. T. Dieck, S. Christoforidis, C. Landerer, L. Hersemann *et al.*, 2021a The novel Rab5 effector FERRY links early endosomes with the translation machinery. BioRiv 10.1101/2021.06.20.449167.

Semotok, J. L., R. L. Cooperstock, B. D. Pinder, H. K. Vari, H. D. Lipshitz *et al.*, 2005 Smaug recruits the CCR4/POP2/NOT deadenylase complex to trigger maternal transcript localization in the early Drosophila embryo. Curr Biol 15**:** 284-294.

Sharma, M., B. Bennewitz and R. B. Klosgen, 2018 Dual or Not Dual?-Comparative Analysis of Fluorescence Microscopy-Based Approaches to Study Organelle Targeting Specificity of Nuclear-Encoded Plant Proteins. Front Plant Sci 9**:** 1350.

Shiber, A., K. Doring, U. Friedrich, K. Klann, D. Merker *et al.*, 2018 Cotranslational assembly of protein complexes in eukaryotes revealed by ribosome profiling. Nature 561**:** 268-272.

Shibuya, T., S. Tsuneyoshi, A. K. Azad, S. Urushiyama, Y. Ohshima *et al.*, 1999 Characterization of the ptr6(+) gene in fission yeast: a possible involvement of a transcriptional coactivator TAF in nucleocytoplasmic transport of mRNA. Genetics 152**:** 869-880.

Shin, M. E., K. D. Ogburn, O. A. Varban, P. M. Gilbert and C. G. Burd, 2001 FYVE domain targets Pib1p ubiquitin ligase to endosome and vacuolar membranes. J Biol Chem 276**:** 41388-41393.

Siddiqui, N., G. Kozlov, I. D'Orso, J. F. Trempe and K. Gehring, 2003 Solution structure of the C-terminal domain from poly(A)-binding protein in Trypanosoma cruzi: a vegetal PABC domain. Protein Sci 12**:** 1925-1933.

Siddiqui, N., M. J. Osborne, D. R. Gallie and K. Gehring, 2007 Solution structure of the PABC domain from wheat poly (A)-binding protein: an insight into RNA metabolic and translational control in plants. Biochemistry 46**:** 4221-4231.

Slotboom, D. J., R. H. Duurkens, K. Olieman and G. B. Erkens, 2008 Static light scattering to characterize membrane proteins in detergent solution. Methods 46**:** 73-82.

Song, H., M. R. Parsons, S. Rowsell, G. Leonard and S. E. Phillips, 1999 Crystal structure of intact elongation factor EF-Tu from Escherichia coli in GDP conformation at 2.05 A resolution. J Mol Biol 285**:** 1245-1256.

Steinberg, G., and J. Perez-Martin, 2008 Ustilago maydis, a new fungal model system for cell biology. Trends Cell Biol 18**:** 61-67.

Stenmark, H., R. Aasland and P. C. Driscoll, 2002 The phosphatidylinositol 3-phosphate-binding FYVE finger. FEBS Lett. 513**:** 77-84.

Stoffel, N. K., 2022 Deciphering RRM code of Rrm4 using iCLIP experiments in *Ustilago maydis*, pp. Heinrich-Heine-Universität Düsseldorf.

Studier, F. W., 2005 Protein production by auto-induction in high density shaking cultures. Protein Expr. Purif. 41**:** 207-234.

Su, H., S. Meng, Y. Lu, M. I. Trombly, J. Chen *et al.*, 2011 Mammalian hyperplastic discs homolog EDD regulates miRNA-mediated gene silencing. Mol. Cell 43**:** 97-109.

Svergun, D. I., 1992 Determination of the regularization parameter in indirect-transform methods using perceptual criteria. . J. Appl. Crystallogr. 25**:** 495-503.

Takahashi, M., M. Higuchi, H. Matsuki, M. Yoshita, T. Ohsawa *et al.*, 2013 Stress granules inhibit apoptosis by reducing reactive oxygen species production. Mol Cell Biol 33**:** 815-829.

Takemoto, K., K. Ebine, J. C. Askani, F. Kruger, Z. A. Gonzalez *et al.*, 2018 Distinct sets of tethering complexes, SNARE complexes, and Rab GTPases mediate membrane fusion at the vacuole in Arabidopsis. Proc Natl Acad Sci U S A 115**:** E2457-E2466.

Terfrüchte, M., B. Joehnk, R. Fajardo-Somera, G. Braus, M. Riquelme *et al.*, 2014 Establishing a versatile Golden Gate cloning system for genetic engineering in fungi. Fungal Genet. Biol. 62**:** 1-10.

Tian, L., H. L. Chou, M. Fukuda, T. Kumamaru and T. W. Okita, 2020a mRNA localization in plant cells. Plant Physiol. 182**:** 97-109.

Tian, L., H. L. Chou, L. Zhang, S. K. Hwang, S. R. Starkenburg *et al.*, 2018 RNA-Binding Protein RBP-P Is Required for Glutelin and Prolamine mRNA Localization in Rice Endosperm Cells. Plant Cell 30**:** 2529-2552.

Tian, L., K. A. Doroshenk, L. Zhang, M. Fukuda, H. Washida *et al.*, 2020b Zipcode RNA-Binding Proteins and Membrane Trafficking Proteins Cooperate to Transport Glutelin mRNAs in Rice Endosperm. Plant Cell 32**:** 2566-2581.

Trcek, T., and R. Lehmann, 2019 Germ granules in Drosophila. Traffic 20**:** 650-660.

Tria, G., H. D. Mertens, M. Kachala and D. I. Svergun, 2015 Advanced ensemble modelling of flexible macromolecules using X-ray solution scattering. IUCrJ 2**:** 207-217.

Tulinski, M., 2021 Identifizierung einer neuen Rrm4-abhängigen Komponente des endosomalen mRNA-Transports in *Ustilago maydis*, pp. Heinrich-Heine-Universität Düsseldorf.

Turner-Bridger, B., C. Caterino and J. M. Cioni, 2020 Molecular mechanisms behind mRNA localization in axons. Open Biol 10**:** 200177.

Valentini, E., A. G. Kikhney, G. Previtali, C. M. Jeffries and D. I. Svergun, 2015 SASBDB, a repository for biological small-angle scattering data. Nucleic Acids Res. 43**:** D357-363.

Vargas, J. N. S., J. N. Sleigh and G. Schiavo, 2022 Coupling axonal mRNA transport and local translation to organelle maintenance and function. Curr Opin Cell Biol 74**:** 97-103.

Varshavsky, A., 2012 The ubiquitin system, an immense realm. Annu Rev Biochem 81**:** 167-176.

Vollmeister, E., K. Schipper, S. Baumann, C. Haag, T. Pohlmann *et al.*, 2012 Fungal development of the plant pathogen Ustilago maydis. FEMS Microbiol Rev 36**:** 59-77.

Weiler, A. J., O. Spitz, M. Gudzuhn, S. N. Schott-Verdugo, M. Kamel *et al.*, 2021 Novel intracellular phospholipase B from *Pseudomonas aeruginosa* with activity towards endogenous phospholipids affects biofilm assembly. bioRiv doi.org/10.1101/2021.06.15.448513.

Wiedemann, N., and N. Pfanner, 2017 Mitochondrial Machineries for Protein Import and Assembly. Annu Rev Biochem 86**:** 685-714.

Wu, H., J. Zhou, T. Zhu, I. Cohen and J. Dictenberg, 2020 A kinesin adapter directly mediates dendritic mRNA localization during neural development in mice. J Biol Chem 295**:** 6605-6628.

Xie, J., G. Kozlov and K. Gehring, 2014a The "tale" of poly(A) binding protein: the MLLE domain and PAM2-containing proteins. Biochim. Biophys. Acta 1839**:** 1062-1068.

Xu, X., A. Huang, X. Cui, K. Han, X. Hou *et al.*, 2019 Ubiquitin specific peptidase 5 regulates colorectal cancer cell growth by stabilizing Tu translation elongation factor. Theranostics 9**:** 4208-4220.

Yang, R., S. A. Gaidamakov, J. Xie, J. Lee, L. Martino *et al.*, 2011 La-related protein 4 binds poly(A), interacts with the poly(A)-binding protein MLLE domain via a variant PAM2w motif, and can promote mRNA stability. Mol. Cell Biol. 31**:** 542-556.

Zaessinger, S., I. Busseau and M. Simonelig, 2006 Oskar allows nanos mRNA translation in Drosophila embryos by preventing its deadenylation by Smaug/CCR4. Development 133**:** 4573-4583.

Zander, S., S. Baumann, S. Weidtkamp-Peters and M. Feldbrugge, 2016a Endosomal assembly and transport of heteromeric septin complexes promote septin cytoskeleton formation. J Cell Sci 129**:** 2778-2792.

Zander, S., S. Baumann, S. Weidtkamp-Peters and M. Feldbrügge, 2016b Endosomal assembly and transport of heteromeric septin complexes promote septin cytoskeleton formation. J. Cell Sci. 129**:** 2778-2792.

Zipfel, C., G. Kunze, D. Chinchilla, A. Caniard, J. D. Jones *et al.*, 2006 Perception of the bacterial PAMP EF-Tu by the receptor EFR restricts Agrobacterium-mediated transformation. Cell 125**:** 749-760.

9 Acknowledgment

Foremost, I owe my highest gratitude to my Ph.D. supervisor Prof. Dr. Michael Feldbrügge, for allowing me to pursue a Ph.D. in this intriguing topic. All the more for providing a friendly, motivating work environment, excellent guidance, constructive criticisms, support, and setting up high standards in research.

I would like to thank Prof. Holger Gohlke for accepting to be my second supervisor and for his support and feedback. I would like to thank Prof. Dr. Lutz Schmitt for his generous support and collaboration and SFB1208 for the funding.

I would like to thank Dr. Sander HJ Smits and his CSS team, Astrid, Jens, Stefanie, and Violetta, for supporting the structural studies. I thank Dr. Stephan Schott-Verdugo for his valuable input with Topmodel computational studies.

I would like to thank Kira Müntjes for proofreading and correcting the complete thesis, for critical feedback, and for supporting the in vivo studies. I thank Sainath for helping with the structural analyses and in vitro experiments.

I would like to thank my co-workers Lilli, Florian Altegoer, Seomun (Woogie), Srimeenakshi, Nina, Sainath, and Taniya for proofreading parts of this thesis and for critical feedback.

I thank Ute, Simone, and Jessica for their valuable technical support. I thank Carola Leitsch and Elisabeth Stratmann, who made the complicated official formalities simple. I thank Uli, who ensured everything worked smoothly in the institute.

I thank Dr. Sigrun Feldbrügge and Han dang Klein at JUNO, HHU, who made my life easier in Düsseldorf with their kind help and assistance with visa and official formalities at the city office.

I thank all the current and past members of the Institute of Microbiology, HHU, for their unlimited support and for making the work environment friendly with the shared social activities, especially the Carnival, Christmas, graduation parties, and hat crafting. I would like to thank the Rab-lab community Thomas, Carl, Ben, Ute, Sabrina, Silke, Lilli, Kira, Jungho, Markus, Minerva, Woogie, Sri, Nina, Lea, Sachiko, Jessica, Lisa, Pascal, Simon and my intern student Albert for their support and feedback in my projects at various time points. In addition, I would like to thank all the present and past co-workers from the Rab-express and Rab-attack members Vera, Kerstin, Bettina, Theresa, Jörn, Marius, Michèle, Peter, Kai, Magnus, Kaitlyn, Kristin, Lesly, Summia, Natascha, Ashiq, Sanchi, Trusha, and Joana for their generous help and support at various time points of my Ph.D.

A big thanks to Thomas and Lilli for their kind help and support in settling down at Düsseldorf in my initial days and for introducing me to the project, life on the campus, and Duesseldorf.

Special thanks to Lilli, Woogie, Sri, Nina, and Taniya who motivated and backed me up during the stressful span and kept my spirits high.

I thank my Ph.D.-batchmates and friends at Düsseldorf Sathish, Arun, Ruchika, Bala, Sushma, Prashanth Desai, Prashanth Shinde, Anand, Priya, Mohan, Sriram, Rizwan, Sri, Karthika, Sarnava, Kumari, Kumaresan, Kannan, Sakshi, Sabah, Aiswaria, Ritu, Marie, Miriam, Rajesh, Taniya, Dibin, Feby, Sai, Pavithran, Thirulok Chandar, Vedanta, Harsha, Deepitha at Heidelberg, Vivek at Barcelona, Imran at Hamburg, Arun at Groningen, and Prem at Amsterdam for the priceless memories during my Ph.D. with their joyful company.

This journey wouldn't have been possible without the support of many people who helped, mentored, and guided me at various career stages. I would like to thank Prof. Dr. P. Loganathan, who mentored me during my bachelor's degree, and my friends Praveen, Meenakshi Kumar, Charles, Jeeva, Shins, Renjan, Sekar, Nini, Emy and Venita for their help and backing.

I would like to thank Prof. Dr. S. Karutha Pandian at Alagappa University, who inspired me to pursue a career in science. I thank my friends, seniors, and batchmates Ganesh Babu, Saravana Kumar, Nisha, Sushma, Gowri, SK, Muthu Kumar, Karthik, Boopathy, M. Karuppia, and Baskar for their help and encouragement.

I would like to thank Prof. Dr. Gaiti Hasan, who provided the opportunity to carry out my master's thesis in her laboratory and nurtured me as a molecular biologist at National Center for biological sciences (NCBS), Bangalore. I would like to thank Prof. Dr. K. Vijayaragavan, Former director of NCBS, and my colleagues and friends at NCBS, Senthil Ramnad, Velumani, Jeevan, Vaigundan, Vivek, Manivannan, Gayathri, Shalima, Ramesh Kumar and Nagaraj for their support and memorable times.

I would like to thank Prof. Dr. K. Suguna, who introduced me to the world of structural biologists at the Molecular Biophysics Unit (MBU), Indian Institute of Science (IISc), Bangalore. I want to thank my labmates and friends at IISc, Koustav, Anita, Karthika, Barathy, Swetha, Nandini, Dinesh, Mohan, and Yamuna, for their support and good old memories.

I would like to thank Prof. Dr. N. Manoj for providing an independent structural biology project and letting me mature as a senior researcher in his laboratory at the Department of Biotechnology, Indian Institute of Technology Madras (IITM), Chennai. I thank my co-workers and friends during my stay at IITM Gayathri, Thirumalai, Puneet, Samar, Mrityunjey, Navin, Deepak, Infant, Viveka, Tanushree, Swetha, Roshini, Sneha, Barath Pikaji, Ravendiran and Varatharajan for their help, and unforgettable fun junctures.

I would like to thank industrialist Dr. Ganesan at Microtherpeutics Research Laboratory (MTR), Chennai, for introducing me to the wondrous world of the biopharma industry and for his support. Special thanks to my colleagues at MTR, Selvaraj, Murthy, Suthakaran, Bilgi, Shanthi, Kirubakaran, Saranya, Krishna, Siva, Arun, Raja, Raj Kumar, Panneer, Madhurima, Jamima, and Iniyan, for their kind support and memorable times.

I thank all my friends and peers at the Vishnupuram Literature circle who inspired me with their intellectual pursuits and for their support and guidance. A big thanks to writer Jeyamohan who continues to inspire me day-to-day toward excellence through his writings and activities. Special thanks to Saran and Venu for their guidance, love, and support.

I thank my friends Balaji Pandian, Raghu Kumar, M Karthik, Ajatha Sathru, and Saminathan, who supported me during difficult times. I thank my cousin and nice friend Suresh who always helped and encouraged me to pursue my goals and aspirations, my uncle Venkatesh who guided me in the right direction in my college days and my aunt Meenambaal for her eternal love.

Words are not enough to express gratitude to my dear parents and heroes, N.S. Devan and M. Dhanalakshmi, without their love, dedication, and support, I wouldn't have made it here. I thank my dear brothers Mohan and Das, who are my strength and always stood by me. I would like to thank my father in-law Mr. Renganathan and mother in law Ms. Latha Maheswari for their love, support, trust and patience. Finally, I thank my lovely wife, an accomplished engineer and a beautiful person, Nirosha VR (Keerthi), for her love, sacrifice, and commitment to making my life easier, especially while writing this thesis.