Molecular insights into RNA recognition and gene regulation by the TRIM-NHL protein Mei-P26

Anna Salerno-Kochan[1,2], Andreas Horn[3], Pritha Ghosh[4], Chandran Nithin[4,*], Anna Kościelniak[1], Daniela Strauss[3], Oliver Rossbach[5], Janusz M. Bujnicki[4,6], Monika Gaik[1,#], Jan Medenbach[3,#] and Sebastian Glatt[1,#]

[1] Malopolska Centre of Biotechnology, Jagiellonian University Krakow, Poland

[2] Postgraduate School of Molecular Medicine, Warsaw, Poland

[3] Biochemistry I, University of Regensburg, Germany

[4] Laboratory of Bioinformatics and Protein Engineering, International Institute of Molecular and Cell Biology in Warsaw, Poland

[5] Institute of Biochemistry, University of Giessen, Germany

[6] Bioinformatics Laboratory, Institute of Molecular Biology and Biotechnology, Faculty of Biology, Adam Mickiewicz University, Poznan, Poland

[*] Present address: Laboratory of Computational Biology, Faculty of Chemistry, Biological and Chemical Research Centre, University of Warsaw, Warsaw 02-089, Poland

Correspondence to MG (monika.gaik@uj.edu.pl), JM (Jan.Medenbach@vkl.uni-regensburg.de) and SG (sebastian.glatt@uj.edu.pl)

Keywords: Mei-P26, NHL, RNA recognition, TRIM-NHL proteins, stem cell differentiation, translational regulation

Running title: Structural and functional characterization of the Mei-P26 NHL domain

Abstract

The TRIM-NHL protein Meiotic P26 (Mei-P26) acts as a regulator of cell fate in *Drosophila*. Its activity is critical for ovarian germline stem cell maintenance, differentiation of oocytes and spermatogenesis. Mei-P26 functions as a post-transcriptional regulator of gene expression, however, the molecular details of how its NHL domain selectively recognizes and regulates its mRNA targets have remained elusive. Here, we present the crystal structure of the Mei-P26 NHL domain at 1.6 Å resolution and identify key amino acids that confer substrate specificity and distinguish Mei-P26 from closely related TRIM-NHL proteins. Furthermore, we identify mRNA targets of Mei-P26 in cultured *Drosophila* cells and show that Mei-P26 can act as either a repressor or activator of gene expression on different RNA targets. Our work reveals the molecular basis of RNA recognition by Mei-P26 and the fundamental functional differences between otherwise very similar TRIM-NHL proteins.

Introduction

RNA binding proteins (RBPs) play key roles in the post-transcriptional regulation of gene expression. They comprise a large and functionally diverse group of proteins involved in all aspects of RNA biology from RNA synthesis to its degradation. RBPs typically bind RNAs through dedicated RNA-binding domains (RBDs)(1). Several members of the evolutionary conserved TRIM-NHL family employ their NHL domains to interact with RNA (2–6). The TRIM-NHL protein family shares a common architecture comprised of an N-terminal tripartite motif (TRIM, consisting of a RING domain, one or two B-Box type zinc fingers and a coiled-coil domain) followed by a C-terminal NCL-1, HT2A, LIN-41 (NHL) domain (7–11). The NHL domain folds into a β-propeller that typically acts as a scaffold to mediate interactions with other biomolecules such as proteins, DNA or RNA (6, 12, 13).

TRIM-NHL proteins play important roles in development where they control cell fate decisions to regulate differentiation and cell growth (11, 14). The *Drosophila melanogaster* genome encodes several proteins with a TRIM-NHL-like architecture, among them Brain tumor (Brat) and Meiotic-P26 (Mei-P26). Mei-P26 was identified as a regulator of differentiation in the male and female germline and its loss results in over-proliferation of germline cells, tumor formation and sterility (15–19). Germline homeostasis depends on the maintenance of germline stem cells (GSCs) in the stem cell niche and on the proper differentiation of their progeny into gametes (20). In the female germline, Mei-P26 supports both cellular programs. It ensures maintenance of GSCs through control of BMP signaling (17), but also promotes differentiation of daughter cells upon exit from the stem cell niche (18). In the male germline, Mei-P26 limits mitotic divisions during the differentiation process of precursor germ cells preventing over-proliferation (16). Moreover, ovarian cells lacking Mei-P26 activity grow abnormally large and exhibit increased nucleolar size (21).

The function of the closely-related Brat protein critically depends on its NHL domain as its deletion or other sequence alterations result in strong phenotypes that can be partially rescued by expression of the NHL domain alone (22, 23). The Brat NHL domain participates in multiple protein-protein interactions involving factors such as Pum, eIF4EHP and Miranda as binding partners (23–28). Similarly, Mei-P26 function requires several additional proteins, among them, Sxl (Sex-lethal), Bam (Bag of marbles), Bgcn (Benign gonial cell neoplasm), Wuho and Ago1 (Argonaute-1) (17, 21, 29, 30). The Mei-P26 NHL domain is required for the interaction with at least some of these proteins, as mutations leading to substitutions in its sequence abolish the interaction with Ago1 and impair ovarian stem cell maintenance (21). Finally, the NHL domains of both Brat and Mei-P26 interact with RNA in a sequence-specific manner (3, 6, 31).

Considering the functional importance of the Mei-P26 NHL domain, we conducted comprehensive structural and functional analyses. We confirm the high structural similarity between the NHL domains of Mei-P26 and Brat, but also detect subtle differences that influence their function and affect target specificity. Computational modeling of the Mei-P26 NHL interaction with its RNA substrate allowed us to predict and experimentally validate key residues critical for sequence-specific RNA binding. Finally, we employed individual-nucleotide resolution cross-linking and immunoprecipitation (iCLIP) to identify cellular mRNA targets of full-length Mei-P26 and its NHL domain. We confirm the recognition of several identified target sequences by Mei-P26 using individual EMSA analyses. Furthermore, we show that Mei-P26 directly affects the expression of reporters that carry these newly identified Mei-P26 binding sites and dissect the domain requirements for regulation. In summary, our results deepen the molecular understanding of NHL-RNA interactions that play a key function in governing cell fate decisions.

Materials and Methods

DNA constructs

First, the coding sequence of Mei-P26 NHL (isoform C/E) was amplified by RT-PCR from total RNA (prepared from *Drosophila* embryos, strain Oregon-R) and cloned into a modified pFastBac HTA vector carrying an additional N-terminal GST tag sequence followed by a TEV cleavage site. Mutated variants of Mei-P26 NHL were obtained using the QuickChange cloning protocol. For Brat-NHL, the His6-ubiquitin fusion protein produced from the pHUE vector was used (a gift from I. Loedige) (6). For the luciferase reporter plasmids, 3′ UTR fragments of select genes were RT-PCR amplified from total RNA prepared from *Drosophila* S2R+ cells (primers and RNA regions are provided in supplementary table 1) and ligated into a modified pCasPeR-Heatshock vector containing a Firefly luciferase open reading frame (pHS-FLuc (32)) using the *Hpa*I and *Bgl*II restriction sites. A vector encoding Renilla luciferase (pHS-B_m-RLuc-(EF)m-SV40 aka pHS-RL) (32) was used for co-transfection and served as a reference for normalization. Plasmids encoding HA-tagged or lambda N-HA-tagged, full-length Mei-P26 or Brat were kindly provided by I. Loedige (MDC Berlin) and G. Meister (University of Regensburg). For iCLIP experiments, the coding sequences of Mei-P26 and its NHL domain were subcloned into a modified pAc5.1 vector using the *Sbf*I and *Not*I restriction sites to generate expression plasmids encoding N-terminally 2x FLAG-tagged proteins. Point mutations in the NHL domain were subsequently introduced using site-directed mutagenesis. The sequences of all oligonucleotides used for cloning are provided in the table (see **Supplementary Table 1**). All vectors were validated by sequencing.

Protein expression and purification

Mei-P26 NHL and its mutated variants were obtained using Bac-to-Bac baculovirus expression system according to standard protocols. For protein expression, Hi5 insect cells

were infected at a MOI of 0.5 and grown for 3 days at 27°C. After harvesting (4°C, 15′, 7 000 rcf), the cell pellets were resuspended in lysis buffer (50 mM HEPES pH 7.5, 600 mM NaCl, 10% Glycerol, 5 mM $MgCl_2$, 1 mM DTT, DNase), snap frozen in liquid nitrogen and stored in -80°C. For purification, cells were lysed by 3 cycles of freezing and thawing, followed by mild sonication and centrifugation (4°C, 1h, 80 000 rcf). The supernatant was applied and circulated on a Glutathione Sepharose 4 Fast Flow 16/10 column (GE Healthcare) for 16 h, then washed with buffer A (50 mM HEPES pH 7.5, 600 mM NaCl, 10% Glycerol, 1 mM DTT) and high salt buffer A (50 mM HEPES pH 7.5, 1 M NaCl, 10% Glycerol, 1 mM DTT). Elution occurred in a buffer containing 50 mM HEPES pH 8, 600 mM NaCl,1 mM DTT, 10 mM Glutathione. The elution fractions were pooled, supplemented with TEV protease and dialysed overnight against the dialysis buffer (20 mM HEPES, 300 mM NaCl, 1 mM DTT). Subsequently, the samples were re-adsorbed onto a Glutathione Sepharose 4 Fast Flow 16/10 column (GE Healthcare) and the flow-through was purified by size exclusion chromatography using a Superdex 200 10/300 GL column (GE Healthcare). Purified samples were stored in gel filtration buffer (20 mM HEPES, 300 mM NaCl, 5 mM DTT). Brat-NHL was expressed and purified as previously described (31).

Crystallization and structure determination

Mei-P26 NHL crystals were grown at 20°C using the sitting drop vapor diffusion method. The purified protein was concentrated to 11 mg/ml in gel filtration buffer and combined with an equal volume of reservoir solution (0.2 Potassium thiocyanate and 20% PEG3350). After 2 weeks of incubation, crystals were fished, cryo-protected with 30% glycerol and subsequently frozen in liquid nitrogen. Datasets were collected at BESSY II Helmholtz-Zentrum Berlin beamlines 14.1 and 14.2. Data processing was performed using XDS (33) and initial phases were obtained by molecular replacement using a search poly-alanine model based of Brat-NHL (PDB code 1Q7F) (28) in Phaser (34). A structure model

was built in COOT (35) and subsequently refined in Phenix (36). All dataset and refinement statistics are given in table **(Table 1)**. Model figures were generated using Pymol (37).

RNAcompete analysis

RNAcompete data for the Brat and Mei-P26 NHL domains were obtained from Prof. Timothy Hughes' laboratory at the University of Toronto (6). The raw data was computationally split into two halves (Set A and B) in the RNAcompete pipeline (38) to facilitate internal data comparisons. Next, the Z-scores for 7-mers for Set A, Set B and the average of Set A and Set B (Set A+B) were calculated, ranked and presented on scatter plots. Consensus motifs recognized by Mei-P26 and Brat were calculated from average of two halves of the RNA pool.

Modelling of Mei-P26-NHL:RNA complexes

The initial structure of the Mei-P26-NHL:RNA complex was generated by superimposing the Mei-P26 protein with the Brat-RNA complex (PDB ID: 5EX7) and copying the coordinates of the Brat RNA ligand. The superposition was performed for the 257 Cα atoms (RMSD = 1.343 Å), which matches between the Mei-26-NHL and Brat proteins. The sequence of that RNA was modified to UUUUUUU, UUUUACA, or UUUGUUGU using UCSF Chimera (39) to prepare three starting structures for Mei-P26-NHL:RNA complexes. Molecular dynamics simulations for Mei-P26-NHL:RNA complexes for the three cases were performed using the Amber18 package (40). Molecular dynamics simulations were run for Brat in complex with UUGUUAA, UUUUACA and UUGUUGU RNAs as controls. The input structure for the simulation was prepared using tleap in a truncated octahedral box of 10 Å allowance using the TIP3P water model (41). Simulations were performed using the combination of the Amber ff14sb force field for proteins (42) and the χOL3 force field for RNA (43, 44). The structure was energy-minimized for 10,000 cycles with restraints, followed by 10,000 cycles without restraints. The minimized structures were

subjected to heating, density equilibration and short runs of equilibration. The heating was done from 100 K to 300 K for 500 ps with restraints on the entire structure and the density equilibration was performed for 500 ps, also with restraints on the entire structure. The equilibration of the structures was run for four short rounds. The first three rounds of equilibration were run for 200 ps each with the main chain atoms constrained. The final round of equilibration was performed for 2 ns to ensure full convergence and reliability of the models. The production run was run for 1 µs. We have used constant pressure periodic boundary conditions (ntb = 2) with isotropic position scaling (ntp = 1) with a pressure relaxation time taup = 2.0 ps for the production run. The particle-mesh Ewald (PME) procedure (45) was used for computing the electrostatic interactions. The cut-off values used for electrostatics and LJ interactions were set as 12 Å. The equilibration steps were run with the NVT ensemble (ntb = 1) while the production run was performed with the NPT ensemble. The minimization was performed using Sander and the subsequent steps were performed using the CUDA version of PMEMD available in the Amber package (46–48). The simulation trajectories were clustered using the reimplementation of NMRCLUST algorithm (49) available in UCSF Chimera and the representative frames (**Supplementary Fig. 4, 5**) are provided as PDB files in the supplementary materials. All simulated models and the used restraints were deposited to a publicly available data repository (Mendeley Data; doi:10.17632/jvkcfwyz47.1). The Theseus analysis was performed to evaluate the MD simulation by simultaneous superposition of the Brat NHL - UUGUUGU crystal structure (PDB ID: 5EX7) with the Brat NHL - UUGUUGU model and the top five clusters from representative models (52).

Microscale thermophoresis

Experiments were conducted using 20 nM of Cy5-labeled oligonucleotides in the buffer containing 20 mM HEPES, 100 mM NaCl, 5 mM $MgCl_2$, 5 mM DTT, 0.0125% Tween

20 for Mei-P26 NHL and 20 mM Tris 8.0, 150 mM NaCl, 1 mM $MgCl_2$, 5 mM DTT, 0.0125% Tween 20 for Brat NHL. 3:1 serial dilutions of unlabelled protein (starting from the highest concentration of 5 µM) was mixed with labelled oligonucleotide, incubated for 15 minutes and applied for measurements. Measurements were conducted at 40% MST power and light-emitting diode (LED)/excitation power setting 20% in Premium Coated capillaries on the Monolith NT.115 at 25°C (Nanotemper Technologies). Each experiment was performed in at least three replicates. The data were analysed using the MO.Affinity software (Nanotemper Technologies) at the standard MST on time off 5 s. To calculate dissociation constants (K_d), Hill models were fitted to each dataset. The graphs were prepared in GraphPad Prism software. Oligonucleotides used in the experiments are listed in **Supplementary Table 1**.

Electromobility shift assays (EMSA)

15 pmol of RNA/DNA were radioactively labeled for 1 h at 37 °C using 5 U T4 Polynucleotide Kinase (Thermo Fisher) and 10 µCi γ-^{32}P ATP. The reaction was inactivated at 75 °C for 10 min and the labeled RNA/DNA was purified by gel filtration (Illustra MicroSpin G-25 columns; GE Healthcare). Recombinant, wild-type Mei-P26 NHL was prepared in 20 mM HEPES/KOH pH 7.5, 100 mM NaCl, 5 mM DTT, 5% glycerol, variants were prepared in the same buffer but containing 300 mM NaCl. 10 fmol of RNA/DNA were diluted in 10 mM Tris/Cl pH 7.4, 50 mM KCl, 1mM EDTA, 1 mM DTT, 0.4 mg/ml yeast tRNA (Invitrogen) and mixed with the indicated amounts of recombinant protein in a total reaction volume of 10 µl. After incubation for 30 min at 4°C, samples were separated by electrophoresis (6% polyacrylamide [37.5:1 acrylamide:bisacrylamide], 5% glycerol, 44.5 mM Tris, 44.5 mM boric acid) at 230 V for 60 min at 4 °C. Gels were dried for 2 h at 80 °C and analysed on a Personal Molecular Imager (Bio-Rad). For RNA stability assays, 10 µl reactions were set up as described above, containing radioactively labeled RNA and 1280 nM

of the purified NHL domain. Control reactions were supplemented with buffer instead of the protein preparation. After 30 minutes of incubation at 4°C, RNAs were purified by organic extraction, separated by 15% denaturing PAGE and visualized by autoradiography.

Thermal shift assays

Recombinant Mei-P26 NHL domain and variants thereof (each at a concentration of 1g/l) were incubated with SYPRO Orange and 20 mM HEPES pH 7.5, 300 mM NaCl, 5 mM DTT buffer in 96-well plates followed by centrifugation (5', 1000 rpm). Subsequently, the samples were gradually heated from 4 to 98°C with a rate of 0.2°C/10 s in the CFX96 Real-Time System C1000 Touch Thermal Cycler (Biorad). The fluorescence intensity was measured using an excitation wavelength of 470 nm while monitoring emission at 570 nm.

Tissue culture

Drosophila S2R+ cells were propagated at 25°C at 80% confluency in Express Five SFM supplemented with 10X Glutamax (Thermo Fisher).

Western blotting

Cultured cells were harvested and resuspended in lysis buffer (20 mM Tris/Cl pH 8.0, 150 mM NaCl, 5 mM EDTA, 1% NP-40, 2% SDS). Protein concentration of cleared lysates was determined using the BioRad protein assay reagent. 25 µg of total protein were separated by denaturing PAGE and subjected to Western blotting using mouse monoclonal anti-FLAG antibody (M2, Sigma Aldrich, 1:1000) followed by probing with an HRP-coupled anti-mouse light chain-specific secondary antibody (1:10000, Jackson Immuno Research). Detection occurred by using Clarity Western ECL substrate and a ChemiDoc Touch Imaging System (BioRad). After stripping, the membrane was re-probed using mouse anti-alpha-tubulin antibody (DM1A, Sigma Aldrich, 1:1000).

Individual-nucleotide cross-linking and immunoprecipitation (iCLIP)

Cells were transfected for 48 h with 15 µg DNA per 15 cm dish using Fugene HD (Promega) following the manufacturer's instructions using plasmids encoding full-length FLAG-tagged MeiP26 protein, its NHL domain only, or their respective mutant versions. A novel, improved version of the iCLIP method was utilized, termed "iCLIP2" (53). Briefly, a 15-cm dish of cells was washed with PBS and UV-irradiated (120mJ/cm² at 254 nm) using a UV Stratalinker 2400 (Stratagene). Next, cell extract was prepared and subjected to RNase treatment using 10 U of RNase I (Ambion). Immunoprecipitation was performed with anti-FLAG antibody (M2, Sigma Aldrich) or control serum on Dynabeads Protein A (Life Technologies) for 2 h at 4°C. After washing four times with washing buffer (50 mM Tris/Cl pH 7.4, 1 M NaCl, 0.05% Tween 20), the co-immunoprecipitated RNA was dephosphorylated, ligated to a 3'-RNA linker and 5'-radiolabeled with T4 PNK and [γ-^{32}P]-ATP. Samples were subjected to neutral SDS-PAGE (NuPAGE, Invitrogen) and transferred to a nitrocellulose membrane. Protein/RNA-complexes were visualized by autoradiography. Mei-P26:RNA-complexes were cut from the membrane, proteins were digested with Proteinase K and RNA was subjected to iCLIP2 library preparation as previously described (53). Sequencing occurred on a HiSeq 2500 (Illumina). Three independent biological replicates were performed for each protein construct; as a control, non-transfected cells were processed in parallel.

iCLIP data analysis

The iCLIP data were processed using the iCount software suite and analysis pipeline (54). The sequencing reads were demultiplexed based on barcodes for individual replicates (allowing one mismatch), PCR duplicates were removed and adapters were trimmed. The reads for each of the replicates were aligned to the *Drosophila melanogaster* genome (ENSEMBL release 98) and processed separately. Cross-linked nucleotides (peaks) were identified and then clustered. Gene loci that produced iCLIP peaks in the experiments

conducted with both, the full-length protein and the NHL domain were manually curated. When crosslinking occurred to mitochondrially-encoded RNAs (4 loci in total) or known contaminants such as snoRNA/scaRNA/snRNA sequences (or similar) present in the host genes (17 loci), the gene loci were excluded from further analyses. Similarly, crosslinking to low complexity regions (mostly A-stretches that resemble polyA-tails) or sequences derived from the transfected plasmids (originating from the Actin5C promoter or the Mei-P26 coding region) were not considered. In the remaining 214 genes, crosslinking positions were considered equivalent between the individual experiments when the distance between the crosslinking peaks (full-length versus NHL domain) was <50 nt. In these loci, the presence of Mei-P26 RNA target motifs was scored up to 30 nt upstream and 20 nt downstream of the crosslink positions, considering full matches (UUUUACN, UUUUANA, UUUUNCA, or UUUUUUU, 65 loci) or U-rich sequences with four consecutive U residues followed by 3 nt containing at least one additional U residue (63 loci).

Tethering- and reporter assays

For tethering assays, per well of a 96-well plate $4x10^4$ *Drosophila* SR2+ cells were seeded and transfected using Fugene (Promega) according to the manufacturer's instructions with the following plasmids: (1) 10 ng of a Firefly luciferase reporter with five BoxB elements in the 3′ UTR (pAC-FL-5boxB, kindly provided by I. Loedige), (2) 30 ng of a control plasmid encoding Renilla luciferase and (3) 60 ng of plasmids encoding either HA- or λNHA-tagged proteins of interest. 48 h after transfection, the cells were lysed with 1x Passive Lysis Buffer (Promega) and luciferase activities were determined using the Dual-Luciferase Reporter Assay System (Promega) and a Centro LB 960 luminometer (Berthold). Relative luciferase units (RLUs) were calculated for each sample by dividing Firefly luciferase activity by Renilla luciferase activity. Activities are expressed as the ratio of RLUs obtained for the tethered, λNHA-tagged proteins relative to the untethered control proteins. Depicted are mean

values ± standard deviation of at least three independent biological experiments performed in technical triplicates. 3′ UTR reporter assays were performed analogously. The transfection mixture contained 100 ng of plasmid DNA consisting of 23.3 ng Firefly luciferase reporter plasmid (pHS-FL bearing different 3' UTR sequences), 1.7 ng Renilla luciferase encoding plasmid (pHS-RL, normalization control) and 75 ng of a pAc5.1 plasmid encoding either 2xFLAG-tagged Mei-P26, or a mutant version thereof. An empty pAc5.1 plasmid served as a control. 48 h after transfection, RLUs were determined as described above and normalized to the empty pAc5.1 vector control. Depicted are mean values ± standard deviation of at least three independent biological experiments performed in technical triplicates.

Results

The crystal structure of the *Drosophila* Mei-P26 NHL domain

To gain molecular insight into RNA binding of *Drosophila melanogaster* Mei-P26, we aimed to obtain structural information of its C-terminal NHL domain (Mei-P26 NHL) at high resolution. Therefore, we predicted the beginning of the NHL domain and expressed different constructs using a baculovirus expression system. We obtained large quantities of Mei-P26 $NHL_{aa908\text{-}1206}$ producing a homogenous monomer of approximately 30 kDa, which was completely free of any proteinaceous contaminants and nucleic acids (**Fig. 1A**). Mei-P26 NHL crystallized in several tested conditions and we collected numerous complete datasets at various synchrotron sources. We solved the structure of the Mei-P26 NHL domain at 1.6 Å resolution by molecular replacement using the backbone of the previously determined Brat NHL domain (28) as a reference model. The N-terminus of the NHL domain remains invisible due to averaging of its conformations throughout the crystal but is present as there was no indication of proteolytic degradation during purification and crystallization (**Supplementary Fig. 1A, B)**. After refinement, an atomic model could be obtained with R/R_{free} values of

20.6%/21.9% obeying all basic rules of proftein stereochemistry (**Table 1**). Mei-P26 NHL folds into a six-bladed β-propeller with a donut-like shape and a diameter of ~45 Å and a height of ~25 Å (**Fig. 1A**). The two molecules located in the asymmetric unit are connected by a di-sulfide bond between Cys1030, but the functional relevance of this dimerization under reducing conditions *in vivo* remains questionable. The six blades are asymmetrically distributed in a radial fashion around a central axis. Each unit is composed of four antiparallel β-strands connected by loops of various length and degrees of flexibility. In the center of the molecule the six β-sheets form a solvent channel with a diameter of approximately 12 Å. The overall structure is stabilized by a β-sheet complementation of the first N-terminal β-strand (aa 933-940) with the sixth sheet (aa 1178-1202) in a molecular velcro-like fashion.

Mei-P26 NHL binds to single-stranded RNA

We compared the structure of Mei-P26 NHL to the NHL domains of Brat, Lin41 and Thin/Abba to gain further insights into their unique and commonly shared features (3, 31, 55). All four proteins act as post-transcriptional repressors of gene expression and recognize their target RNAs via their NHL domains. Recently, the RNA binding modes for Brat and Lin41 have been reported, highlighting the existence of two distinct recognition mechanisms. In detail, the NHL domain of Brat binds to linear single-stranded RNA motifs (6), while the NHL domain of Lin41 prefers RNA hairpins (3). The NHL domains of the *D. melanogaster* Mei-P26 and Brat are closely related to each other, whereas the NHL domain of Lin41 from *D. rerio* noticeably resembles the NHL domains of the fly Wech and human TRIM71 proteins (3, 6). Thin/Abba is a member of a distinct branch of the TRIM-NHL phylogeny, which also includes *H. sapiens* TRIM32 and *C. elegans* NHL-1 (14). These evolutionary relationships are also observed on the structural level, although all four NHL domains in principle share a similar overall architecture and differ only in the length of individual loops and β-strands (**Supplementary Fig. 1C)**. As with Brat and Lin41, the NHL domain of Mei-P26 and

Thin/Abba also exhibit a highly positively charged patch on their top surface (**Supplementary Fig. 1D)** (3, 31, 55, 56). Because it has been experimentally demonstrated that RNA binding in NHL domains occurs via this charged surface area, it was previously assumed that Mei-P26 NHL can also bind to RNA (3, 6, 31). However, 11 out of 15 of the residues important for RNA binding in the closely related Brat NHL domain are not conserved in the Mei-P26 NHL domain (**Supplementary Fig. 1E, Supplementary Fig. 2**). There is, however, one notable exception: three amino acids important for the recognition of the first position of the RNA target by Brat (a uridine base interacting with Asn800, Tyr829 and Arg847) are conserved in Mei-P26 (Asn970, Tyr 999 and Arg1017). This suggests that Mei-P26 also binds a uridine base in the same region but utilizes varying, non-conserved amino acids to recognize different sequence motifs in its target RNAs.

To assess the binding specificity of the Mei-P26 NHL domain experimentally, we performed complementary Microscale Thermophoresis (MST) and Electrophoretic Mobility Shift Assay (EMSA) experiments. MST experiments showed purified Mei-P26 NHL binds a single stranded polyU RNA sequence (U_7, K_d = 1.8 ± 0.2 µM), while no interaction could be detected with a folded RNA hairpin (recognized by Lin41) or a single stranded 7 nt DNA oligonucleotide (TTTTACA; **Fig. 1B**). EMSA experiments confirmed the complex formation between Mei-P26 NHL and a polyU RNA oligonucleotide without affecting RNA integrity (**Supplementary Fig. 3C**) and verified its inability to bind other tested nucleic acids (**Fig. 1C**). The similarity between Mei-P26 NHL and Brat as well as the interaction data indicate that both proteins preferentially recognize linear motifs in single-stranded RNA (ssRNA) molecules.

Mei-P26 NHL recognizes U-rich RNA sequences

To gain further insight into the sequence specificity of Mei-P26 NHL, we used recently developed algorithms (38) to re-analyze the data from RNAcompete experiments previously carried out by the Hughes and Morris groups (6). We sought to obtain a complete list of oligonucleotide sequences that may interact with the NHL domain of Mei-P26, as an initial step of defining Mei-P26 NHL targets. In our analyses we utilized a previously generated library of almost all possible 7-mer oligonucleotide fragments that have been tested for binding to purified GST-tagged Mei-P26 NHL (6). The relative binding of each of the sequences is presented as a Z-score calculated from a dataset computationally separated into two subsets termed Set A and Set B and for a combined subsets A+B. Consensus motifs recognized by Mei-P26 NHL (**Fig. 2A**) were calculated from the ten top-scored sequences of 7-mer motifs. Of note, we performed the same control analysis for the Brat NHL domain and subsequently confirmed the binding of its RNAcompete-driven consensus motif (BRAT1: UUGUUAA) using MST (**Supplementary Fig. 3A, B**). Our data are consistent with previously published results (6) and show that Mei-P26 NHL preferentially binds oligonucleotides including three or more consecutive uridines commonly flanked by additional adenine and/or cytosine residues.

Modeling of the NHL:RNA complexes

To better understand the mechanistic nature of RNA binding, we performed molecular dynamics (MD) simulations for the NHL domains of Mei-P26 and Brat with the U-rich RNA sequences derived from RNAcompete (**Fig. 2B, Supplementary Fig. 4, Supplementary Fig. 5)**. Initial models of Mei-P26 in complex with RNA (SEQ1: UUUUUUU, SEQ3: UUUUACA, or BRAT1: UUGUUAA) were generated by modeling the RNA to reflect the conformation of the UUGUUGU nucleotide bound to the Brat NHL domain (PDB ID: 5EX7). According to our analysis, SEQ3 binds most stably to the Mei-P26 NHL domain and converges into a well-defined conformation (**Fig. 2B, Supplementary Fig. 4**). Simulations

with the top-scored polyU oligonucleotide (SEQ1) and the inverted SEQ3 sequence (ACAUUUU) resulted in much higher RMSD values, indicating a weaker fit in comparison to the SEQ3 RNA (**Supplementary Fig. 4A**). The top five clusters obtained by the MD simulations with SEQ3 are much more similar to each other than any of them compared to the clusters obtained with SEQ1. Moreover, we observed lower variance in the 3' region and more structural heterogeneity of the 5'-RNA docking site between the individual clusters of the Mei-P26:UUUUACA models (**Supplementary Fig. 4B**). Together, our simulations emphasize the importance of the ACA trinucleotide anchor adjacent to poly-uridine stretches for Mei-P26 RNA recognition (**Fig. 2B, Supplementary Fig. 4**). Although similar affinities were determined for both oligonucleotides (**Fig. 2C, Supplementary Fig. 6**), an elevated value for the polyU could have been indirectly caused by the so called avidity effect (57) (**Fig. 2D)**. This effect describes a scenario where an individual binding event increases the probability of additional interactions occurring in its close proximity. Our *in silico* analysis, together with RNAcompete results, suggests that Mei-P26 NHL may use various RNA recognition modes to interact with different RNA sequences despite similar binding constants.

To obtain a relative comparison between models and crystal structures using the same modeling restraints, we performed three control MD simulations of the Brat NHL protein with both the Brat NHL and the Mei-P26 NHL top-scoring sequences from the RNAcompete experiment (BRAT1: UUGUUAA, SEQ3: UUUUACA) and the co-crystallized UUGUUGU nucleotide (PDB ID: 5EX7) (**Supplementary Fig. 5A)**. Interestingly, the Brat NHL top-scoring sequence (BRAT1) from the RNAcompete experiment showed a relatively similar RMSD profile to the Mei-P26 NHL top-scoring sequence (SEQ3). Even if the model of crystallized Brat: UUGUUGU complex shows the lowest RMSD in comparison to the starting model, we still observed a slight structural divergence from the obtained crystal structure (**Supplementary Fig. 5B)**. Strikingly, the MD simulations for both Mei-P26 NHL:SEQ3 and

Brat NHL:BRAT1 complexes are of a similar range (**Supplementary Fig. 4, Supplementary Fig. 5)**, but to show the atomic importance of individual residues the obtained models cannot substitute the crystal structures. Notably, the RMSD profile of Brat NHL:SEQ3 increases later during simulations, while the Mei-P26 NHL:SEQ3 complex remains stable, indicating that the ACA trinucleotide does not adapt well to the Brat NHL surface. During the simulation of the Brat-BRAT1 complex, the guanosine nucleotide remains in a stable interaction with the binding pocket further supporting its essential role in sequence-specific RNA recognition by Brat. Hence, our computational analyses support the premise that Mei-P26 and Brat have very distinct RNA recognition modes.

Experimental characterization of the RNA target specificity of Mei-P26 NHL

Next, we set out to experimentally validate the target predictions and performed *in vitro* binding experiments using the six top ranked RNA sequences. Our results show that untagged Mei-P26 NHL interacts with four of the identified 7-nt RNA oligonucleotides at affinities that would be expected for specific but transiently associated RBPs (**Fig. 2C, Supplementary Fig. 6**). Among the identified sequences, Mei-P26 NHL showed the highest affinities towards SEQ3 (K_d 1.7 µM) and polyU (K_d 1.8 µM). Mei-P26 NHL also exhibited measurable affinities to the SEQ2 and SEQ5 RNA oligonucleotides but binding to SEQ4 and SEQ6 was reduced. Limited solubility of purified Mei-P26 NHL protein at higher concentrations did not allow determination of the dissociation constants for these oligonucleotides. These data demonstrate sequence-specific recognition and discrimination by the Mei-P26 NHL domain.

To gain further insights into the RNA sequence preferences of Mei-P26 NHL, we systematically altered individual positions and features of oligonucleotides identified as Mei-P26 NHL ligands. First, we added uridine nucleotides at the 5′ end of SEQ3, which resulted in

an increased binding affinity (**Fig. 2D**). This observation agrees well with the finding that Mei-P26 NHL also displays an increased affinity towards extended polyU sequences (e.g. U_9; **Fig. 2D**). Of note, NHL domains usually bind to relatively short RNA motifs and specifically recognize only a few nucleobases. For instance, the NHL domain of Brat establishes contacts with only six consecutive bases (6), while Lin41 only recognizes two nucleobases and makes most of its contacts via the ribose-phosphate backbone (3). Hence, we assume that the NHL domain of Mei-P26 also recognizes a short stretch of bases (**Fig. 2B, Supplementary Fig. 4**). Furthermore, we created an inverted SEQ3 oligonucleotide (ACAUUUU; **Fig. 2D**), which exhibited decreased affinity and indicated the necessary 5′ to 3′ directionality of the motif. Analysis performed with a UUUUCAC sequence also displayed significantly lower affinity further indicating the importance of the ACA anchor (**Fig. 2D).** We also managed to convert a poor substrate (UUUACAA) into a substrate able to be bound by the addition of a single U nucleotide (**U**UUUACAA), highlighting the necessity of at least four consecutive uridines. Conversely, insertion of a guanosine nucleotide at asymmetric and symmetric positions into the polyU sequence (SEQ1) reduced the affinity to Mei-P26 (**Fig. 2D**). Extending SEQ1 by two additional uridines, lead to the observation of increased affinity, illustrating the aforementioned avidity effect for polyU sequences (57) (**Fig. 2D**). In summary, our results show sequence-specific RNA recognition by the Mei-P26 NHL domain and confirm computational predictions based on the RNAcompete data.

Identification of key residues involved in RNA recognition

The comparison of the Mei-P26 and Brat NHL domain structures prompted us to analyze the key features of these domains involved in sequence-specific RNA recognition. The models of the Mei-P26 NHL domain in complex with different RNA ligands (**Supplementary Fig. 4**) allowed high confidence prediction of surface areas and individual amino acids important for RNA binding. We predicted Arg1017 to be in close proximity to

the first base (U1) that could stack with Y999 (**Fig. 3A**). Lys1172 and Arg1175 are predicted to be in hydrogen bonding distance to the fourth and sixth base (U4, C6) of the RNA substrate, respectively. Moreover, in our MD experiment for Mei-P26 with SEQ3, Arg1175 was predicted to interact with the second, third, fourth and fifth base (U2, U3, U4, A5), while Arg1150 is likely to interact with A5, C6 and A7. Importantly, for Mei-P26 simulations with polyU, we did not observe an interaction of K1172 and R1150 with the RNA, suggesting that they might be involved in the recognition of the ACA anchor of the target RNAs. The docking analyses both emphasize the importance of the positively charged surface patch for RNA binding and highlight the potential contributions of three flexible loop regions that contain several positively charged residues. To confirm this hypothesis, we generated variants of Mei-P26 NHL carrying alanine substitutions of the individual residues (Mei-P26 NHL_{Y999A}, Mei-P26 NHL_{R1017A}, Mei-P26 NHL_{K1172A} and Mei-P26 NHL_{R1175A}), combinations thereof (Mei-P26 $NHL_{K1172A/R1175A}$, Mei-P26 $NHL_{R1150A/K1172A/R1175A}$ (Mei-$P26^{RKR}$)) and deletions of the longest identified loop region (Mei-P26 $NHL_{\Delta P1169\text{-}R1175}$). Finally, we also substituted core residues in the center of the positively charged cavity (Mei-P26 $NHL_{R1001A/K1002A/K1192A}$). All variants were individually expressed in insect cells, purified and subjected to thermal shift assays to confirm proper folding and stability. With the exceptions of substitutions of Tyr999 (Mei-P26 NHL_{Y999A}) and residues in its proximity Arg1001 and Lys1002 (Mei-P26 $NHL_{R1001A/K1002A/K1192A}$), which strongly destabilized the domain or led to protein degradation (**Supplementary Fig. 7A,B**), all variants showed purity and stability parameters comparable to the wild type Mei-P26 NHL protein (**Fig. 3B, Supplementary Fig. 7A,B**). We tested all stable Mei-P26 NHL variants in RNA interaction assays and observed severely compromised or even complete loss of complex formation with SEQ3 (**Fig. 3C**). To understand whether some of the variants exhibit different motif and sequence specificity, we also tested their binding to a polyU oligonucleotide (U_9 and U_{16}). Most of the variants still associate with the

458 polyU oligonucleotides, albeit with reduced affinity (**Fig. 3D**). Of note, our MD simulation
459 indicated that Arg1175 participates exclusively in SEQ3 binding, but appears dispensable for
460 polyU binding. The substitution of Arg1175 indeed strongly reduces Mei-P26 NHL ability to
461 recognize the SEQ3 sequence, but shows only minor influence on its interaction with U_9. In
462 contrast, residue Arg1017 that seems to stabilize U1 in both simulations likewise strongly
463 contributes to the binding of both, SEQ3 and U_9. This data experimentally indicates that Mei-
464 P26 might employ different modes of RNA recognition for the association with different
465 RNA motifs. Furthermore, we show that simultaneous substitution of residues R1150A,
466 K1172A and R1175A fully abolishes the binding to any of the RNA sequences tested.
467 Therefore, our findings are consistent with the MD simulations and conclude that R1150 and
468 K1172 residues are responsible for anchoring the ACA trinucleotide, while R1175 stabilizes
469 the uridine tract. In summary, our results identify particular amino acid residues critical for
470 RNA recognition.

471 We also tested binding of Mei-P26 NHL to the consensus Brat recognition sequence
472 (UUGUUAA, BRAT1), which we obtained from the RNAcompete data (**Supplementary Fig.**
473 **2A**) and which we used in our MD simulations (**Supplementary Fig. 4)**. While this RNA was
474 only weakly bound by Mei-P26, a single substitution from G to U ($BRAT_{mut}$UU**U**UUAA;
475 **Fig. 2D**) generated a Mei-P26-like sequence motif and allowed binding. Similarly, in our
476 experiments, Brat NHL weakly interacts with Mei-P26 SEQ3, demonstrating the specificity of
477 its NHL domain **(Supplementary Fig. 3A).** In the NHL domain of Brat, a single point
478 substitution (R875A) located in the center of the positively charged surface area impairs
479 binding to the target sequence in the 3′ UTR of the *hunchback* (*hb)* mRNA (31). In Mei-P26,
480 the corresponding residue at the same position is already an alanine (Ala1046)
481 (**Supplementary Fig. 8A, 8B**). In an attempt to mimic Brat, we replaced the alanine with
482 arginine (A1046R) in the Mei-P26 NHL domain. The substitution resulted in an overall

483 destabilization of the domain and its failure to bind RNA (**Supplementary Fig. 7B,**
484 **Supplementary Fig. 8C**).

485 ## Identification of Mei-P26 target mRNAs by iCLIP

486 To identify cellular RNA targets of Mei-P26, we performed individual-nucleotide
487 resolution UV crosslinking and immunoprecipitation (iCLIP2 (53)) experiments. As
488 *Drosophila* tissues that express Mei-P26 are not readily accessible for biochemical
489 experimentation such as iCLIP, we turned to cultured *Drosophila* Schneider 2 (S2) cells as a
490 model system. S2 cells have been frequently and successfully employed to study gene
491 regulation processes that are operating in select, specialized tissues or cells (such as the
492 germline or neurons), or that occur during specific developmental stages (6, 32, 58).
493 Moreover, these cells have served to identify the mRNA targets of numerous RNA-binding
494 proteins employing CLIP or related methodology (59–61).

495 S2 cells express only low levels of mRNAs encoding Mei-P26 (62), suggesting that
496 endogenous Mei-P26 protein is not abundant. To overcome this limitation, we transfected
497 constructs encoding FLAG-tagged proteins as bait for the iCLIP experiments. We used full-
498 length Mei-P26 protein, its NHL domain, or the respective derivatives thereof that carry
499 substitutions that impair RNA binding (R1150A, K1172A and R1175A, **Supplementary Fig.**
500 **9A**). As expected, compared to the wild-type counterparts, variants with substitutions
501 immunoprecipitate strongly reduced amounts of RNA (**Supplementary Fig. 9B**).

502 Analyses of the iCLIP data from the wild-type proteins identify a local enrichment of
503 cross-link positions in 751 protein-coding genes for the full-length protein and 623 for the
504 NHL domain that do not exhibit a bias regarding their chromosomal origin (**Fig. 4A,B,**
505 **Supplementary Fig. 9C, Supplementary Data 1**). For both proteins, cross-linking mostly
506 occurs at sites located in the 3′ UTRs of the target genes (**Fig. 4C, Supplementary Fig. 9C**).

When comparing the iCLIP datasets obtained for full-length Mei-P26 and the NHL domain, only a moderate overlap is observed (22.1%, 249 loci, **Fig. 4A, Supplementary Data 2**). In 71.9% of the shared target genes, crosslinking is observed in comparable positions in the gene body (at a distance of 50 nt or less, **Fig. 4D**). After removal of contaminating sequences (e.g. mitochondrial transcripts, sno-RNA-derived reads, vector-derived sequences), the remaining 214 mRNA targets that are bound by the full length protein and the NHL domain were analyzed for the occurrence of Mei-P26 binding motifs in a region encompassing 30 nts upstream and 20 nts downstream of the crosslink positions. This revealed the presence of U-rich motifs in proximity to the respective crosslink peaks in 128 mRNAs (59.8%), 65 (30%) of which represent *bona fide* Mei-P26 target sequences **(Fig. 4D, 4F, Supplementary Data 2**). Hence, in proximity to the iCLIP peaks, *bona fide* Mei-P26 motifs occur at a ~9 fold higher frequency than expected by chance (f=3.4% expected in random 50mer RNA fragments), validating the computational and experimental *in vitro* binding studies. Further analyses of the mRNAs that are bound by both, full-length Mei-P26 and its NHL domain, reveal an enrichment of genes encoding ribosomal proteins and translation factors (**Fig. 4E**), hinting at a potential function of Mei-P26 in the control of the translation machinery which was previously proposed (21). Several of the newly identified, potential RNA targets from the iCLIP analyses were chosen for validation experiments employing the recombinant Mei-P26 NHL domain. In all cases, EMSA experiments confirmed the interaction and again the Mei-P26 NHL^{RKR} mutant did not associate with the RNAs (**Fig 4F, Supplementary Fig. 10**).

The NHL domain is important for Mei-P26 gene-regulatory activity

To better understand the impact of Mei-P26 on gene expression, we conducted a series of reporter assays in cultured *Drosophila* Schneider 2 cells. We first used the phage-derived lambda-boxB system to artificially recruit full-length Mei-P26 to an RNA. For this, the lambda bacteriophage antiterminator protein N (lambdaN) was fused to Mei-P26 and two

control proteins (Brat and GW182) to tether them to a firefly reporter RNA that contains within its 3' UTR several copies of the lambdaN binding site (boxB). The same proteins without the lambda peptide served as controls. In addition, a co-transfected plasmid encoding a renilla luciferase mRNA that lacks the boxB elements was used for normalization. In this experimental setup, Brat and the positive control GW182, which is involved in miRNA-mediated gene silencing, convey robust repression of the firefly luciferase reporter mRNA (31). In contrast, Mei-P26 exhibits only a weak gene regulatory activity (**Fig. 5A**).

Next, we analyzed expression of a reporter RNA that contains a fragment of the *nanos* 3′UTR, a genetically identified target of Mei-P26 (18). Co-expression of full-length Mei-P26 resulted in silencing of the reporter relative to a control mRNA that bears a 3' UTR from an unrelated RNA (*msl-2*). Repression is dependent on the ability of Mei-P26 to bind to RNA through its NHL domain since the RKR substitution abolishes regulation (**Fig. 5B**). A previously identified U-rich RNA sequence element in the 3′UTR of *nanos* mRNA (29) is critical for Mei-P26-mediated repression as its deletion completely abrogates regulation of the reporter (**Fig. 5B**). Using a similar experimental setup, we tested a series of selected mRNA targets that we identified in our iCLIP analyses (**Fig. 5C**). We included all RNA candidates for which we confirmed Mei-P26 interaction and targets that contain (e.g. *Hrb27c*) or lack (e.g. *spz*) the recognition motif (**Fig. 4F, Supplementary Fig. 10, Supplementary data 2**). The effect of Mei-P26 on the expression of the reporters was diverse. For instance, we could not detect a significant change to the expression of reporters that bear the 3′UTRs of *bic*, *chic* and *RpS23*. In contrast, *LanA*-, *Mlc-c*-, *lost*- and *Hrb27C*-derived reporters exhibited significant repression. Unexpectedly, *RpS20*-, *sqd*-, *sta*-, *Swip*-1, *spz*-, *eIF4A*-, *Col4A1*- and *Hsp83*-derived reporters showed Mei-P26-dependent activation (**Fig. 5C**).

To further dissect the functional domain requirements of Mei-P26, we analyzed the regulatory activity of mutant derivatives. Substitutions in the NHL domain of Mei-P26 that

abrogate RNA binding activity severely blunted regulation of all reporter RNAs, underlining the functional importance of the domain for activity (**Fig. 5C**). Despite being expressed at a comparable level (**Supplementary Fig. 11**), a construct encompassing only the NHL domain was not sufficient for regulation (**Fig. 5D**), demonstrating that additional sequences outside the NHL domain are required. Previously, it has been debated whether the N-terminal RING domain and its ubiquitin ligase activity contribute to the gene regulatory activity of Mei-P26 (21). However, deletion of the N-terminal RING domain neither abolished activation of *spz*, or repression of *nos* or *Hrb27c* reporters, demonstrating that ubiquitin ligase activity is dispensable in this experimental setup (**Fig. 5D**).

Discussion

TRIM-NHL proteins are required for proper development in metazoans and their NHL domains are crucial for function (25, 31, 63, 64). Despite their evolutionarily conserved architecture, NHL domains exhibit clear differences in RNA binding, recognizing diverse RNA sequences or RNA hairpin structures (3, 6). A structural comparison of the NHL domains of Brat, Lin41, Thin/Abba and Mei-P26 reveals similarities between Brat and Mei-P26 regarding their interactions with ssRNA, while Lin41 employs a different mode of interaction. Although none of our numerous attempts yielded co-crystals of the Mei-P26 NHL domain in complex with its RNA substrate, computational modelling allowed us to identify and experimentally validate key amino acid residues involved in RNA recognition. For instance, Mei-P26 and Brat employ an evolutionary conserved interaction site to specifically recognize a uridine base in the first position of their RNA target. In contrast, recognition of other bases differs between the two proteins and they utilize different amino acid residues for substrate binding, resulting in different RNA specificity. While Brat preferentially associates with a UUGUUGU RNA sequence, Mei-P26 recognizes a linear UUUUACA core motif indicating at least two distinct interaction modes between these two similar NHL domains.

582 Typically, individual RNA-binding domains recognize short RNA motifs of 3-5
583 nucleotides in length with rather moderate affinities (65). To increase affinity and specificity,
584 often multiple binding domains are combined, either in tandem in the same polypeptide, or *in*
585 *trans* through formation of protein complexes (66). Our measured affinities are in the
586 micromolar range and are lower for Brat NHL than previously described for longer
587 oligonucleotide sequences (6), but comparable to those reported for CeLin41 NHL for shorter
588 hairpins (3). We provide experimental evidence (**Supplementary Fig. 9B**) that in Mei-P26
589 protein regions outside the NHL domain also contribute to RNA target recognition/binding. *In*
590 *vivo*, it is most likely that additional protein partners contribute to stable complex formation
591 between Mei-P26 and its RNA targets. Genetic experiments demonstrated that besides Mei-
592 P26, the proteins Sxl, Bam, Bgcn and Wuho are crucial for the repression of Nanos protein
593 production and differentiation of ovarian stem cells in the female germline (19, 29, 30, 67,
594 68). For regulation, these proteins likely form a large repressor complex on the 3′UTR of *nos*
595 mRNA (18, 67). Genetic ablation of any one of these proteins impaired regulation and
596 resulted in strong phenotypes (19, 30, 69–73), demonstrating that these factors need to act
597 jointly to achieve their function.

598 Moreover, we find surprising evidence that Mei-P26 not only functions in the
599 repression of selected target mRNAs but acts as an activator on other mRNAs (**Fig. 5C**). How
600 can these two seemingly different activities be explained? In tethering experiments, where the
601 closely related protein Brat acts as a strong silencer of gene expression, Mei-P26 exhibits only
602 weak activity. This suggests that either tethering disturbs its function or Mei-P26 itself is not a
603 strong regulator of gene expression. In the latter scenario, Mei-P26 might function in
604 promoting complex assembly, recruiting other factors that act in gene regulation. Previously,
605 Ago1, Sxl, Bam, Bgcn and Wuho have been identified as co-repressors that act in concert
606 with Mei-P26 in post-transcriptional regulation of gene expression (**Fig. 5E**) (18, 19, 21, 30,

74). Potential activators that can be recruited by Mei-P26 for gene regulation remain to be identified.

Previously, it has been speculated that the N-terminal RING finger ubiquitin ligase domain is important for Mei-P26-mediated regulation of gene expression by promoting the turnover of RNA regulatory factors (21). Our data demonstrate that the Mei-P26 RING domain is dispensable for regulatory activity. Similarly, the closely related protein Brat, a potent post-transcriptional regulator of gene expression, has a truncated TRIM domain that lacks the RING motif. However, in contrast to Brat, where the NHL domain alone was able to provide most of the activity of the full-length protein (22, 23), the isolated NHL domain of Mei-P26 does not exhibit gene regulation in functional assays (**Fig. 5D**).

iCLIP experiments allowed us to identify and validate Mei-P26 binding sites in numerous transcripts. However, the S2 cells that we employed for our experimentation express only a limited set of mRNAs, lacking most germline- and neuron-specific transcripts. Our analysis is thus limited to the identification of Mei-P26 binding sites present in the repertoire of expressed RNAs. Moreover, interacting partners likely shape the interaction profile of Mei-P26 *in vivo*. In *Drosophila* this has been well documented and analyzed in molecular detail for the protein Upstream of N-ras (Unr). In female flies, Unr is recruited to the *msl-2* mRNA by the protein Sxl through highly synergistic interactions (75–78). We have provided the first experimental evidence that the NHL domain of Mei-P26 binds to the 3' UTR of *nanos* mRNA (which is not expressed in S2 cells (62)) in a mutually exclusive fashion, suggesting that repressive complex formation on this transcript requires additional factors such as Sxl, Bam and Bgcn. As judged by the expression level of the corresponding mRNAs (Cherbas et al., Genome Res, 2011), these proteins presumably have extremely low abundance in S2 cells and thus, most likely, cannot significantly impact on the RNA interaction profile of Mei-P26.

Despite these limitations, our experimentation allowed us to identify numerous transcripts that encode translation factors and ribosomal proteins among the Mei-P26 target mRNAs. Previously, Mei-P26, like Brat, has been broadly implicated in the regulation of ribosome biogenesis by controlling the expression of *Myc* (21, 79, 80), which stimulates the expression of the Pol I transcriptional machinery thus promoting ribosome biogenesis (21, 81, 82). The coordinate regulation of ribosomal proteins and ribosomal RNA (rRNA) synthesis by Mei-P26 might allow it to efficiently tune ribosome biosynthesis which is linked to cell growth and the switch between proliferation and differentiation.

In summary, our computational, *in vitro* and *in vivo* studies provide a comprehensive insight into RNA recognition by Mei-P26 and reveal differences to the closely related protein Brat. We further identify and validate numerous novel mRNA targets of Mei-P26 and provide unexpected evidence that it can function in both repression and activation of gene expression.

Data availability

The atomic coordinates and respective structure factors for Mei-P26 NHL (PDB ID: 7NYQ) have been validated and deposited at the European Protein Data Bank. Computational models and description of the restraints have been deposited with Mendeley Data (doi:10.17632/jvkcfwyz47.1). All sequencing data has been deposited at GEO under the following accession number: GSE152013 **(**token for reviewers: kfwtcywofxaxlkv).

Funding

National Science Centre, Poland [UMO-2015/19/P/NZ1/02514 to M.G., UMO-2019/32/T/NZ1/00420 to A.S.K, MAESTRO 2017/26/A/NZ1/01083 to J.M.B.]. German Research Foundation [SFB960 TP B11, ME4238/1-1 to J.M., RTG2355 325443116 to O.R.]; German Federal Ministry of Education and Research [01ZX1401D to J.M.]; Foundation for

655 Polish Science (FNP)[TECH CORE FACILITY/2017-4/6 to S.G., TEAM/2016-3/18 to
656 J.M.B.]; IIMCB statutory funds to J.M.B.

657 Acknowledgements

658 We thank the staff at the beamlines 14.1 (BESSY) and P11 (PETRAIII) for support
659 during data collection and the MCB Structural Biology Core Facility (supported by the
660 TEAM TECH CORE FACILITY/2017-4/6 grant from Foundation for Polish Science) for
661 providing instruments and support. We thank Kate Nie, Debashish Ray, Timothy R. Hughes
662 and Quaid D. Morris at the University of Toronto for providing the RNAcompete data and for
663 valuable discussions on the data analysis. We thank Gunter Meister at the University of
664 Regensburg, Inga Loedige at the Max Delbrück Center for Molecular Medicine, Berlin, and
665 Julian König at the Central European Institute of Molecular Biology, Mainz, for sharing
666 reagents. We thank Panagiotis Alexiou at the Central European Institute of Technology, Brno,
667 for discussions on CLIP data analysis. We thank Elizabeth Michalczyk for useful comments
668 on the revised manuscript. This project has received funding from the European Union's
669 Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant
670 agreement No. 665778 (M.G. and A.S.K.). Computational analyses were performed using the
671 resources of IIMCB, the Poznań Supercomputing and Networking Center at the Institute of
672 Bioorganic Chemistry, Polish Academy of Sciences (grant: 312), the Polish Grid
673 Infrastructure (grants: rnpmd, rnpmc and simcryox) and the Interdisciplinary Centre for
674 Mathematical and Computational Modelling at the University of Warsaw (grants: G73-4 and
675 GB76-30). The open-access publication of this article was funded by the BioS Priority
676 Research Area under the program "Excellence Initiative – Research University" at the
677 Jagiellonian University in Krakow

Author contributions

ASK performed biochemical, biophysical and crystallographic analyses; ASK cloned and purified proteins and characterized variants using biophysical analyses with the help of AK and MG; ASK and SG collected crystallographic data, refined structures and analyzed the crystallographic results. ASK and AH analyzed RNA-protein interaction *in vitro*. AH and OR performed *in vivo* analysis with the help of DS and JM. PG, CN and JMB conducted computational analyses and structure modeling. JM, MG and SG envisioned the project and designed the experimental concepts. ASK, JM and SG prepared figures. ASK, MG, JM and SG wrote the manuscript, with the input from all other authors.

Conflict of interest

Janusz M. Bujnicki is Executive Editor of Nucleic Acids Research.

References

1. Corley,M., Burns,M.C. and Yeo,G.W. (2020) How RNA-Binding Proteins Interact with RNA: Molecules and Mechanisms. *Mol. Cell*, **78**, 9–29.
2. Castello,A., Fischer,B., Eichelbaum,K., Horos,R., Beckmann,B.M., Strein,C., Davey,N.E., Humphreys,D.T., Preiss,T., Steinmetz,L.M., *et al.* (2012) Insights into RNA Biology from an Atlas of Mammalian mRNA-Binding Proteins. *Cell*, **149**, 1393–1406.
3. Kumari,P., Aeschimann,F., Gaidatzis,D., Keusch,J.J., Ghosh,P., Neagu,A., Pachulska-Wieczorek,K., Bujnicki,J.M., Gut,H., Großhans,H., *et al.* (2018) Evolutionary plasticity of the NHL domain underlies distinct solutions to RNA recognition. *Nat. Commun.*, **9**.
4. Kwon,S.C., Yi,H., Eichelbaum,K., Föhr,S., Fischer,B., You,K.T., Castello,A., Krijgsveld,J., Hentze,M.W. and Kim,V.N. (2013) The RNA-binding protein repertoire of embryonic stem cells. *Nat. Struct. Mol. Biol.*, **20**, 1122–1130.
5. Loedige,I., Gaidatzis,D., Sack,R., Meister,G. and Filipowicz,W. (2013) The mammalian TRIM-NHL protein TRIM71/LIN-41 is a repressor of mRNA function. *Nucleic Acids Res.*, **41**, 518–532.
6. Loedige,I., Jakob,L., Treiber,T., Ray,D., Stotz,M., Treiber,N., Hennig,J., Cook,K.B., Morris,Q., Hughes,T.R., *et al.* (2015) The Crystal Structure of the NHL Domain in Complex with RNA Reveals the Molecular Basis of Drosophila Brain-Tumor-Mediated Gene Regulation. *Cell Rep.*, **13**, 1206–1220.
7. Borden,K.L.B. (1998) RING fingers and B-boxes: Zinc-binding protein-protein interaction domains. *Biochem. Cell Biol.*, **76**, 351–358.
8. Reddy Bramham A., Etkin Laurence D.,F.P.S. (1992) A novel zinc finger coiled-coil domain in a family of nuclear proteins.
9. Reymond,A., Meroni,G., Fantozzi,A., Merla,G., Cairo,S., Luzi,L., Riganelli,D., Zanaria,E., Messali,S., Cainarca,S., *et al.* (2001) The tripartite motif family identifies cell compartments. **20**.
10. Slack,F.J. and Ruvkun,G. (1998) A novel repeat domain that is often associated with RING finger and B-box motifs. *Trends Biochem. Sci.*, **23**, 474–475.
11. Williams,F.P., Haubrich,K., Perez-Borrajero,C. and Hennig,J. (2019) Emerging RNA-binding roles in the TRIM family of ubiquitin ligases. *Biol. Chem.*, **400**, 1443–1464.
12. Jin,W., Wang,Y., Chao-Pei,L., Yang,N., Jin,M., Cong,Y., Wang,M. and Rui-Ming,X. (2016) Structural basis for snRNA recognition by the double-WD40 repeat domain of Gemin5. *Genes Dev.*, **30**, 2391–2403.
13. Stirnimann,C.U., Petsalaki,E., Russell,R.B. and Muller,C.W. (2010) WD40 proteins propel cellular networks. *Trends Biochem Sci*, **35**, 565–574.
14. Tocchini,C. and Ciosk,R. (2015) TRIM-NHL proteins in development and disease. *Semin. Cell Dev. Biol.*, **47**–**48**, 52–59.
15. Chen,D., Wu,C., Zhao,S., Geng,Q., Gao,Y., Li,X., Zhang,Y. and Wang,Z. (2014) Three RNA Binding Proteins Form a Complex to Promote Differentiation of Germline Stem Cell Lineage in Drosophila. *PLoS Genet.*, **10**.
16. Insco,M.L., Bailey,A.S., Kim,J., Olivares,G.H., Wapinski,O.L., Tam,C.H. and Fuller,M.T. (2012) A self-limiting switch based on translational control regulates the transition from proliferation to differentiation in an adult stem cell lineage. *Cell Stem Cell*, **11**, 689–700.
17. Li,Y., Maines,J.Z., Tastan,O.Y., McKearin,D.M. and Buszczak,M. (2012) Mei-P26 regulates the maintenance of ovarian germline stem cells by promoting BMP signaling. *Development*, **139**, 1547–1556.
18. Li,Y., Zhang,Q., Carreira-Rosario,A., Maines,J.Z., McKearin,D.M. and Buszczak,M. (2013) Mei-P26 Cooperates with Bam, Bgcn and Sxl to Promote Early Germline Development in the Drosophila Ovary. *PLoS One*, **8**, 1–7.
19. Page,S.L., McKim,K.S., Deneen,B., Van Hook,T.L. and Hawley,R.S. (2000) Genetic studies of mei-P26 reveal a link between the processes that control germ cell proliferation in both sexes and those that control meiotic exchange in Drosophila. *Genetics*, **155**, 1757–1772.
20. Slaidina,M. and Lehmann,R. (2014) Translational control in germline stem cell development. *J. Cell Biol.*, **207**, 13–21.
21. Neumüller,R.A., Betschinger,J., Fischer,A., Bushati,N., Poernbacher,I., Mechtler,K., Cohen,S.M. and Knoblich,J.A. (2008) Mei-P26 regulates microRNAs and cell growth in the Drosophila ovarian stem cell lineage. *Nature*, **454**, 241–245.
22. Arama,E., Dickman,D., Kimchie,Z., Shearn,A. and Lev,Z. (2000) Mutations in the β-propeller domain of the Drosophila brain tumor (brat) protein induce neoplasm in the larval brain. *Oncogene*, **19**, 3706–3716.
23. Sonoda,J. and Wharton,R.P. (2001) Drosophila brain tumor is a translational repressor. *Genes Dev.*, **15**, 762–773.
24. Cho,P.F., Gamberi,C., Cho-Park,Y.A., Cho-Park,I.B., Lasko,P. and Sonenberg,N. (2006) Cap-Dependent

Translational Inhibition Establishes Two Opposing Morphogen Gradients in Drosophila Embryos. *Curr. Biol.*, **16**, 2035–2041.

25. Harris,R.E., Pargett,M., Sutcliffe,C., Umulis,D. and Ashe,H.L. (2011) Brat Promotes Stem Cell Differentiation via Control of a Bistable Switch that Restricts BMP Signaling. *Dev. Cell*, **20**, 72–83.
26. Lee,C.Y., Wilkinson,B.D., Siegrist,S.E., Wharton,R.P. and Doe,C.Q. (2006) Brat is a Miranda cargo protein that promotes neuronal differentiation and inhibits neuroblast self-renewal. *Dev. Cell*, **10**, 441–449.
27. Muraro,N.I., Weston,A.J., Gerber,A.P., Luschnig,S., Moffat,K.G. and Baines,R.A. (2008) Pumilio Binds para mRNA and Requires Nanos and Brat to Regulate Sodium Current in Drosophila Motoneurons. *J. Neurosci.*, **28**, 2099–2109.
28. Edwards,T.A., Wilkinson,B.D., Wharton,R.P. and Aggarwal,A.K. (2003) Model of the Brain Tumor-Pumilio translation repressor complex. *Genes Dev.*, **17**, 2508–2513.
29. Chau,J., Kulnane,L.S. and Salz,H.K. (2012) Sex-lethal enables germline stem cell differentiation by down-regulating Nanos protein levels during Drosophila oogenesis. *Proc. Natl. Acad. Sci. U. S. A.*, **109**, 9465–9470.
30. Rastegari,E., Kajal,K., Tan,B.S., Huang,F., Chen,R.H., Hsieh,T.S. and Hsu,H.J. (2020) WD40 protein Wuho controls germline homeostasis via TRIM-NHL tumor suppressor Mei-p26 in Drosophila. *Dev.*, **147**.
31. Loedige,I., Stotz,M., Qamar,S., Kramer,K., Hennig,J., Schubert,T., Löffler,P., Längst,G., Merkl,R., Urlaub,H., *et al.* (2014) The NHL domain of BRAT is an RNA-binding domain that directly contacts the hunchback mRNA for regulation. *Genes Dev.*, **28**, 749–764.
32. Medenbach,J., Seiler,M. and Hentze,M.W. (2011) Translational control via protein-regulated upstream open reading frames. *Cell*, **145**, 902–913.
33. Kabsch,W. (2010) *XDS*. *Acta Crystallogr. Sect. D Biol. Crystallogr.*, **66**, 125–132.
34. McCoy,A.J., Grosse-Kunstleve,R.W., Adams,P.D., Winn,M.D., Storoni,L.C. and Read,R.J. (2007) Phaser crystallographic software. *J Appl Crystallogr*, **40**, 658–674.
35. Emsley,P. and Cowtan,K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr*, **60**, 2126–2132.
36. Afonine,P. V., Grosse-Kunstleve,R.W., Echols,N., Headd,J.J., Moriarty,N.W., Mustyakimov,M., Terwilliger,T.C., Urzhumtsev,A., Zwart,P.H. and Adams,P.D. (2012) Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. Sect. D Biol. Crystallogr.*, **68**, 352–367.
37. DeLano,W.L. (2002) Pymol: An open-source molecular graphics tool. *CCP4 Newsl. Protein Crystallogr.*, **40**, 82–92.
38. Ray,D., Ha,K.C.H., Nie,K., Zheng,H., Hughes,T.R. and Morris,Q.D. (2018) RNAcompete methodology and application to determine sequence preferences of unconventional RNA-binding proteins. *Methods*, **8260**, 3–15.
39. Pettersen,E.F., Goddard,T.D., Huang,C.C., Couch,G.S., Greenblatt,D.M., Meng,E.C. and Ferrin,T.E. (2004) UCSF Chimera - A visualization system for exploratory research and analysis. *J. Comput. Chem.*, **25**, 1605–1612.
40. Salomon-Ferrer,R., Case,D.A. and Walker,R.C. (2013) An overview of the Amber biomolecular simulation package. *Wiley Interdiscip. Rev. Comput. Mol. Sci.*, **3**, 198–210.
41. Jorgensen,W.L., Chandrasekhar,J., Madura,J.D., Impey,R.W. and Klein,M.L. (1983) Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.*, **79**, 926–935.
42. Maier JA, Martinez C, Kasavajhala K, Wickstrom L, Hauser KE,S.C. (2017) ff14SB: Improving the accuracy of protein side chain and backbone parameters from ff99SB. *Physiol. Behav.*, **176**, 139–148.
43. Banáš,P., Hollas,D., Zgarbová,M., Jurečka,P., Orozco,M., Cheatham,T.E., Šponer,J. and Otyepka,M. (2010) Performance of molecular mechanics force fields for RNA simulations: Stability of UUCG and GNRA hairpins. *J. Chem. Theory Comput.*, **6**, 3836–3849.
44. Zgarbová,M., Otyepka,M., Šponer,J., Mládek,A., Banáš,P., Cheatham,T.E. and Jurečka,P. (2011) Refinement of the Cornell et al. Nucleic acids force field based on reference quantum chemical calculations of glycosidic torsion profiles. *J. Chem. Theory Comput.*, **7**, 2886–2902.
45. Essmann,U., Perera,L., Berkowitz,M.L., Darden,T., Lee,H. and Pedersen,L.G. (1995) A smooth particle mesh Ewald method. *J. Chem. Phys.*, **103**, 8577–8593.
46. Salomon-Ferrer,R., Götz,A.W., Poole,D., Le Grand,S. and Walker,R.C. (2013) Routine microsecond molecular dynamics simulations with AMBER on GPUs. 2. Explicit solvent particle mesh ewald. *J. Chem. Theory Comput.*, **9**, 3878–3888.
47. Götz,A.W., Williamson,M.J., Xu,D., Poole,D., Le Grand,S. and Walker,R.C. (2012) Routine microsecond molecular dynamics simulations with AMBER on GPUs. 1. generalized born. *J. Chem. Theory Comput.*, **8**, 1542–1555.
48. Le Grand,S., Götz,A.W. and Walker,R.C. (2013) SPFP: Speed without compromise - A mixed precision model for GPU accelerated molecular dynamics simulations. *Comput. Phys. Commun.*, **184**, 374–380.
49. Kelley,L.A., Gardner,S.P. and Sutcliffe,M.J. (1996) An automated approach for clustering an ensemble of

NMR-derived protein structures into conformationally related subfamilies. *Protein Eng.*, **9**, 1063–1065.
50. Morris,J.H., Huang,C.C., Babbitt,P.C. and Ferrin,T.E. (2007) StructureViz: Linking Cytoscape and UCSF Chimera. *Bioinformatics*, **23**, 2345–2347.
51. Shannon,P., Markiel,A., Ozier,O., Baliga,N.S., Wang,J.T., Ramage,D., Amin,N., Schwikowski,B. and Ideker,T. (2003) Cytoscape: A software Environment for integrated models of biomolecular interaction networks. *Genome Res.*, **13**, 2498–2504.
52. Theobald,D.L. and Steindel,P.A. (2012) Optimal simultaneous superpositioning of multiple structures with missing data. *Bioinformatics*, **28**, 1972–1979.
53. Buchbender,A., Mutter,H., Sutandy,F.X.R., Körtel,N., Hänel,H., Busch,A., Ebersberger,S. and König,J. (2019) Improved library preparation with the new iCLIP2 protocol. *Methods*, **178**, 33–48.
54. Curk,T. et al. (2019) iCount: protein-RNA interaction iCLIP data analysis. *Prep.*
55. Bawa,S., Gameros,S., Baumann,K., Brooks,D.S., Kollhoff,J.A., Zolkiewski,M., David Re Cecconi,A., Panini,N., Russo,M., Piccirillo,R., *et al.* (2020) Costameric Integrin and Sarcoglycan protein levels are altered in a Drosophila model for Limb Girdle Muscular Dystrophy type 2H . *Mol. Biol. Cell*, 10.1091/mbc.e20-07-0453.
56. Bawa,S., Brooks,D.S., Neville,K.E., Tipping,M., Sagar,M.A., Kollhoff,J.A., Chawla,G., Geisbrecht,B. V., Tennessen,J.M., Eliceiri,K.W., *et al.* (2020) Drosophila TRIM32 cooperates with glycolytic enzymes to promote cell growth. *Elife*, **9**, 1–29.
57. Helder,S., Blythe,A.J., Bond,C.S. and Mackay,J.P. (2016) Determinants of affinity and specificity in RNA-binding proteins. *Curr. Opin. Struct. Biol.*, **38**, 83–91.
58. Weidmann,C.A., Qiu,C., Arvola,R.M., Lou,T.F., Killingsworth,J., Campbell,Z.T., Tanaka Hall,T.M. and Goldstrohm,A.C. (2016) Drosophila nanos acts as a molecular clamp that modulates the RNA-binding and repression activities of pumilio. *Elife*, **5**.
59. Moschall,R., Rass,M., Rossbach,O., Lehmann,G., Kullmann,L., Eichner,N., Strauss,D., Meister,G., Schneuwly,S., Krahn,M.P., *et al.* (2019) Drosophila Sister-of-Sex-lethal reinforces a male-specific gene expression pattern by controlling Sex-lethal alternative splicing. *Nucleic Acids Res.*, **47**, 2276–2288.
60. Stepien,B.K., Oppitz,C., Gerlach,D., Dag,U., Novatchkova,M., Krüttner,S., Stark,A. and Keleman,K. (2016) RNA-binding profiles of Drosophila CPEB proteins Orb and Orb2. *Proc. Natl. Acad. Sci. U. S. A.*, **113**, E7030–E7038.
61. Hansen,H.T., Rasmussen,S.H., Adolph,S.K., Plass,M., Krogh,A., Sanford,J., Nielsen,F.C. and Christiansen,J. (2015) Drosophila Imp iCLIP identifies an RNA assemblage coordinating F-actin formation. *Genome Biol.*, **16**.
62. Cherbas,L., Willingham,A., Zhang,D., Yang,L., Zou,Y., Eads,B.D., Carlson,J.W., Landolin,J.M., Kapranov,P., Dumais,J., *et al.* (2011) The transcriptional diversity of 25 Drosophila cell lines. *Genome Res.*, **21**, 301–314.
63. Nicklas,S., Otto,A., Wu,X., Miller,P., Stelzer,S., Wen,Y., Kuang,S., Wrogemann,K., Patel,K., Ding,H., *et al.* (2012) TRIM32 regulates skeletal muscle stem cell differentiation and is necessary for normal adult muscle regeneration. *PLoS One*, **7**.
64. Schwamborn,J.C., Berezikov,E. and Knoblich,J.A. (2009) The TRIM-NHL Protein TRIM32 Activates MicroRNAs and Prevents Self-Renewal in Mouse Neural Progenitors. *Cell*, **136**, 913–925.
65. Maris,C., Dominguez,C. and Allain,F.H.T. (2005) The RNA recognition motif, a plastic RNA-binding platform to regulate post-transcriptional gene expression. *FEBS J.*, **272**, 2118–2131.
66. Hennig,J. and Sattler,M. (2015) Deciphering the protein-RNA recognition code: Combining large-scale quantitative methods with structural biology. *BioEssays*, **37**, 899–908.
67. Li,Y., Park,J.K., Minor,N.T., Maines,J.Z. and McKearin,D.M. (2009) Bam and Bgcn antagonize Nanos-dependent germ-line stem cell maintenance. *Proc. Natl. Acad. Sci.*, **106**, 9304–9309.
68. Chau,J., Kulnane,L.S. and Salz,H.K. (2009) Sex-lethal facilitates the transition from germline stem cell to committed daughter cell in the Drosophila ovary. *Genetics*, **182**, 121–132.
69. Schupbach,T. (1985) Normal Female Germ Cell Differentiation Autosome Ratio and Expression O F Sex-Lethal. *Construction.*
70. Hashiyama,K., Hayashi,Y. and Kobayashi,S. (2011) Drosophila Sex Lethal Gene Initiates Female Development in Germline Progenitors. *Science (80-.).*, **885**, 885–888.
71. Salz,H.K., Cline,T.W. and Schedl,P. (1987) Functional changes associated with structural alterations induced by mobilization of a P element inserted in the Sex-lethal gene of Drosophila. *Genetics*, **117**, 221–231.
72. Ohlstein,B. and McKearin,D. (1997) Ectopic expression of the Drosophila Bam protein eliminates oogenic germline stem cells. *Development*, **124**, 3651–3662.
73. Gönczy,P., Matunis,E. and DiNardo,S. (1997) bag-of-marbles and benign genial cell neoplasm act in the germline to restrict proliferation during Drosophila spermatogenesis. *Development*, **124**, 4361–4371.
74. Fu,Z., Geng,C., Wang,H., Yang,Z., Weng,C., Li,H., Deng,L., Liu,L., Liu,N., Ni,J., *et al.* (2015) Twin Promotes the Maintenance and Differentiation of Germline Stem Cell Lineage through Modulation of

Multiple Pathways. *Cell Rep.*, **13**, 1366–1379.
75. Grskovic,M., Hentze,M.W. and Gebauer,F. (2003) A co-repressor assembly nucleated by Sex-lethal in the 3′UTR mediates translational control of Drosophila msl-2 mRNA. *EMBO J.*, **22**, 5571–5581.
76. Abaza,I., Coll,O., Patalano,S. and Gebauer,F. (2006) Drosophila UNR is required for translational repression of male-specific lethal 2 mRNA during regulation of X-chromosome dosage compensation. *Genes Dev.*, **20**, 380–389.
77. Duncan,K., Grskovic,M., Strein,C., Beckmann,K., Niggeweg,R., Abaza,I., Gebauer,F., Wilm,M. and Hentze,M.W. (2006) Sex-lethal imparts a sex-specific function to UNR by recruiting it to the msl-2 mRNA 3′ UTR: Translational repression for dosage compensation. *Genes Dev.*, **20**, 368–379.
78. Hennig,J., Militti,C., Popowicz,G.M., Wang,I., Sonntag,M., Geerlof,A., Gabel,F., Gebauer,F. and Sattler,M. (2014) Structural basis for the assembly of the Sxl–Unr translation regulatory complex. *Nature*, **515**, 287–290.
79. Betschinger,J., Mechtler,K. and Knoblich,J.A. (2006) Asymmetric Segregation of the Tumor Suppressor Brat Regulates Self-Renewal in Drosophila Neural Stem Cells. *Cell*, **124**, 1241–1253.
80. Ferreira,A., Boulan,L., Perez,L. and Milán,M. (2014) Mei-P26 mediates tissue-specific responses to the brat tumor suppressor and the dMyc Proto-Oncogene in Drosophila. *Genetics*, **198**, 249–258.
81. Frank,D.J., Edgar,B.A. and Roth,M.B. (2002) The Drosophila melanogaster gene brain tumor negatively regulates cell growth and ribosomal RNA synthesis. *Development*, **129**, 399–407.
82. Grewal,S.S., Li,L., Orian,A., Eisenman,R.N. and Edgar,B.A. (2005) Myc-dependent regulation of ribosomal RNA synthesis during Drosophila development. *Nat. Cell Biol.*, **7**, 295–302.

Figure 1. Structure of the Mei-P26 NHL domain and its specific ssRNA recognition

A. Mei-P26 domain organization with NHL domain shown in blue (top left). Abbreviations: RING domain (R), B-Boxes (B), coiled coil (CC), glutamine rich region (Q). Size exclusion chromatography profile and SDS PAGE gel for Mei-P26 NHL protein with mass markers indicated at the top. Blue line on the chromatogram corresponds to 280nm wavelength, orange line to 254 nm wavelength (top right). Cartoon and surface representation of the crystal structure of the Mei-P26 NHL domain in top surface and side orientations (encompassing amino acids 931-1203) (left bottom). **B.** MST results for the fraction of bound DNA, dsRNA, or ssRNA (U_7) ligands for different concentrations of Mei-P26 NHL. The dissociation constant (K_d) was calculated from at least three independent experiments ($n \geq 3$). **C.** EMSAs employing various concentrations of the Mei-P26 NHL domain (as indicated above each lane) on a single-stranded DNA oligonucleotide, an RNA hairpin structure (both as indicated in the MST panel), or a single-stranded U_9 RNA oligonucleotide. Depicted is a representative gel of three independent replicates.

Figure 2. The Mei-P26 NHL domain specifically recognizes single-stranded, U-rich RNA motifs

A. Analysis of the Mei-P26 NHL RNAcompete data. Scatter plot representing Z-scores for two halves of the RNA pool (set A and set B) shown at the bottom; six highest scoring 7-mers were highlighted in color in right corner of the plot. The consensus motif derived from the indicated sequence motifs is indicated at the top. **B.** Comparison between the Mei-P26 NHL in complex with UUUUACA (left) and UUUUUUU (right) (color-coded RNA sequences depicted at the top). The complexes were computationally modelled using molecular dynamics simulations by introducing oligonucleotides onto the surface of the Mei-P26 NHL domain. **C, D:** MST binding studies of the NHL domain to the 6 highest scoring 7-mer sequences (C), variants thereof, or the motif recognized by the closely related protein Brat (D). K_d values at least three independent experiments ($n \geq 3$). MST binding curves for protein-oligonucleotide complexes with calculated K_d were marked in solid lines, curves with $K_d > 3$ μM were presented as experimental points.

Figure 3. Identification of amino acids in the NHL domain important for RNA binding

A. Model of the Mei-P26 NHL domain in complex with a UUUUACA RNA sequence. Select amino acids predicted to be in close proximity to the RNA are highlighted. **B.** Size exclusion chromatography profile and SDS-PAGE gel for selected Mei-P26 mutated variants. Abbreviations for K1172A/R1175A (KR) and R1150A/K1172A/R1175A (RKR) are used in the inset. **C** and **D**. MST- (panels C and D (left)) or EMSA-based (panel D (right)) *in vitro* binding analyses of the recombinant NHL domain or variants thereof (as indicated in each graph) to RNA oligonucleotides with the sequence UUUUACA (C), U_9 (D (left)), or U_{16} (D (right)). Dissociation constants (K_d) were calculated from three independent experiments (n=3).

Figure 4. Identification of Mei-P26 target RNAs in cultured cells

A. Comparison of iCLIP data derived from either full-length Mei-P26 protein (depicted in blue) or its NHL domain (depicted in salmon). Number of genes containing a statistically significant local enrichment of cross-link positions (cross-link peaks) and percent values are given for each fraction. **B.** Chromosomal origin of the 249 shared target genes identified in both iCLIP datasets. **C.** Location of the full-length Mei-P26 (left) or its NHL domain-derived (right) iCLIP clusters within transcripts. **D.** Presence of potential Mei-P26 RNA motifs in vicinity of the cross-link peaks in the shared target genes. No overlap: cross-linking of the full-length protein and the NHL domain occurred in different regions of the gene locus. **E.** Gene Ontology (GO)-term analysis of biological processes enriched in the shared target genes. **F.** Validation of select Mei-P26 binding sites identified by iCLIP. Top left: schematic depiction of the *Hrb27c* and *Col4a1* gene loci. Introns are depicted as lines, exons as boxes; the grey shading indicates the protein coding region. Below: iCLIP read depth analyses of 3′ UTR regions (as indicated by the solid black lines) from experiments performed with either the full-length Mei-P26 protein (light blue), its NHL domain (dark blue), or from control experiments (grey). Potential RNA motifs recognized by Mei-P26 are highlighted in yellow. Right: EMSA analyses using RNA fragments derived from the iCLIP clusters (as indicated by the red arrows, sequences provided at the bottom of the gels), using different concentrations of the recombinant Mei-P26 NHL domain or its mutant derivative (Mei-P26^{RKR}: R1150A, K1172A, R1175A, as indicated above the gels). Free RNA probe and NHL:RNA complexes are indicated on the left.

Figure 5. Mei-P26 regulates gene expression via 3′ UTR binding sites

A. Tethered function assay using Brat, Gw182 (positive control) and full-length Mei-P26. Activities are calculated relative to proteins that lack the tag for tethering. **B.** Mei-P26-mediated repression of a reporter that either bears the *nanos* mRNA 3′ UTR *(nos* wt) or a version thereof in which a U-rich sequence element previously implicated in regulation was mutated (*nos* mut). Reporter activity was determined in the presence of co-expressed wild-type Mei-P26 protein (grey bars), or a Mei-P26 protein carrying substitutions that affect RNA interaction. (Mei-P26^{RKR}:R1150A, K1172A, R1175A; black bars). All activities are expressed relative to control reactions without overexpression of Mei-P26 (using an empty plasmid). **C.** Reporter assays using 3′ UTR sequences derived from various Mei-P26-bound genes (as indicated at the bottom) and employing wt Mei-P26-FL protein (grey bars) or its mutant derivative (black bars) as described for panel B. All activities are expressed relative to control reactions without overexpression of Mei-P26 (using an empty plasmid). **D.** Reporter assays using 3′ UTR sequences of *nos* (grey), *Hrb27c* (blue) and *spz* (dark blue) derived from Mei-P26-FL protein or its mutant derivatives (as indicated at the bottom) as described for panels B and C. For panels a-d mean values ±SD are depicted of at least three independent biological experiments performed in three technical replicates each. P-values were calculated with a two-sided student's t-test relative to the control reactions described for each panel; n.s. not significant, * P-value <0.05, **: P-value <0.01, *** P-value <0.001. **E.** Schematic depiction of Mei-P26-mediated post-transcriptional regulation of gene expression. Mei-P26 (NTD in grey, NHL domain in blue, mutated residues that are crucial for RNA binding are highlighted) associates with U-rich RNA motifs present in the 3′ UTRs of its RNA targets (depicted schematically) to regulate their expression. Regulation involves additional factors such as Ago1, Sxl, Bam, Bgcn, and Wuho that have been implicated in Mei-P26-dependent repression, while potential co-factors involved in Mei-P26-dependent gene activation remain to be identified.

978 # Table 1: Data collection and refinement statistics

		Mei-P26 NHL PDB ID 7NYQ BESSY II, MX14-1
Data collection		
Space group		$P2_1$
Cell dimensions		
a, *b*, *c* (Å)		34.71 116.50 64.70
α, β, *γ* (°)		90 96.362 90
Wavelength		0.9184
Resolution (Å) [†]		43.17-1.6 (1.64-1.6)
R_{meas} (%)		7.7 (160.9)
$I/\sigma(I)$		10.15 (0.85)
$CC_{1/2}$		0.99 (0.46)
Completeness (%)		98.8 (97.0)
Redundancy		3.82
Refinement		
Resolution (Å)		43.17-1.6
No. reflections		66477
R_{work} / R_{free}		0.196 / 0.215
No. atoms		
Protein		8923
Ligand/ion		no ligands
Water		320
B factors		
Protein		36.7
Ligand/ion		not applicable
Water		38.3
r.m.s deviations		
Bond lengths (Å)		0.007
Bond angles (°)		0.689
Ramachandran statistics (%)	outliers	0
	allowed regions	4.59
	favored regions	95.41
Rotamer outliers (%)		0.21
Clash score		3.02
MolProbity score		1.41

979 [a] Values in parentheses are for highest-resolution shell.

980 [†] Resolution limits according to I/σ of 2 is 1.75 Å