Published as: *Nature*. 2014 October 9; 514(7521): 252–256.

Mechanism of Dis3L2 substrate recognition in the Lin28/let-7 pathway

Christopher R. Faehnle[#1,4], **Jack Walleshauser**[#1,2,4], and **Leemor Joshua-Tor**[1,2,3,4]

[1]W. M. Keck Structural Biology Laboratory 1 Bungtown Road, Cold Spring Harbor, NY 11724, USA

[2]Watson School of Biological Science 1 Bungtown Road, Cold Spring Harbor, NY 11724, USA

[3]Howard Hughes Medical Institute 1 Bungtown Road, Cold Spring Harbor, NY 11724, USA

[4]Cold Spring Harbor Laboratory 1 Bungtown Road, Cold Spring Harbor, NY 11724, USA

[#] These authors contributed equally to this work.

Summary paragraph

The pluripotency factor Lin28 inhibits the biogenesis of the let-7 family of mammalian microRNAs[1–4]. Lin28 is highly expressed in embryonic stem cells and has a fundamental role in regulation of development[5], glucose metabolism[6] and tissue regeneration[7]. Alternatively, Lin28 overexpression is correlated with the onset of numerous cancers[8], while let-7, a tumor suppressor, silences several human oncogenes[5]. Lin28 binds to precursor let-7 (pre-let-7) hairpins[9], triggering the 3' oligo-uridylation activity of TUT4/7[10–12]. The oligoU tail added to pre-let-7 serves as a decay signal, as it is rapidly degraded by Dis3L2[13,14], a homolog of the catalytic subunit of the RNA exosome. The molecular basis of Lin28 mediated recruitment of TUT4/7 to pre-let-7 and its subsequent degradation by Dis3L2 is largely unknown. To examine the mechanism of Dis3L2 substrate recognition we determined the structure of mouse Dis3L2 in complex with an oligoU RNA to mimic the uridylated tail of pre-let-7. Three RNA binding domains form an open funnel on one face of the catalytic domain that allows RNA to navigate a path to the active site different from its exosome counterpart. The resulting path reveals an extensive network of uracil-specific interactions spanning the first twelve nucleotides of an oligoU-tailed RNA. We identify three U-specificity zones that explain how Dis3L2 recognizes, binds and processes uridylated pre-let-7 in the final step of the Lin28/let-7 pathway.

In embryonic stem cells, Dis3L2 plays the role of the effector nuclease responsible for degrading uridylated pre-let-7 miRNAs in the Lin28/let-7 pathway[13,14]. Dis3L2 belongs to the RNase II/R 3'-5' exonuclease superfamily, which includes the catalytic subunit of the

Correspondence and requests for materials should be addressed to L.J(leemor@cshl.edu)..

Author contributions C.R.F., J.W. and L.J. designed and C.R.F and J.W. conducted all experiments. All authors contributed to data analysis and wrote the paper.

Coordinates and structure factors for Dis3L2-U_{14} complex have been deposited in the Protein Data Bank with PDB code 4PMW.

RNA exosome in yeast (yRrp44)[15,16] and humans (hDis3 and hDis3L1)[17]. Inactivation of Dis3L2 is associated with aneuploidy and mitotic abnormalities, while its overexpression suppresses cancer cell growth[18]. Genetic disruption of *DIS3L2* is the primary cause of Perlman syndrome, a congenital disorder leading to fetal overgrowth and an increased susceptibility to Wilms' tumor development[18]. Subsequent studies have shown that Wilms' tumors, a common pediatric kidney cancer, overexpress Lin28[8], underscoring the role of miRNA regulation in kidney tumorigenesis. Outside of miRNA regulation, Dis3L2 mediates 3'-5' mRNA decay in an alternative mRNA decay pathway that is independent of the exosome[19,20]. This is particularly intriguing in light of studies showing widespread uridylation of mammalian mRNAs[21,22], suggesting that RNA uridylation may be a common RNA degradation signal. Here we present the structure of mouse Dis3L2 in complex with an oligoU RNA and provide an explanation for Dis3L2's high specificity toward uridylated RNA substrates.

It was shown that Dis3L2 is a processive 3'-5' hydrolytic exonuclease that preferentially degrades uridylated pre-let-7 and other uridylated miRNAs[13,14]. Related studies showed Dis3L2 processed structured RNA more efficiently than its exosome counterparts (hDis3 and yRrp44), with no reported sequence preference[20]. However, the fission yeast Dis3L2 ortholog was shown to exhibit a preference for oligoU[19]. To clarify the extent of mammalian Dis3L2's substrate preference, we quantitatively measured mouse Dis3L2's specific activity for single-stranded RNA substrates (Fig. 1a and Extended Data Fig. 1a, b). Dis3L2 processed U_{15} ~9-fold and ~40-fold more efficiently than C_{15} and A_{15}, respectively, which are both relatively poor substrates (Fig. 1a). Dis3L2 failed to process deoxyU_{15}, thus is specific for RNA (Fig. 1a and Extended Data Fig. 1a). Next, we measured binding affinity, where Dis3L2 exhibited more than 200-fold tighter binding to U_{15} compared to A_{15} and C_{15} (Fig. 1b and Extended Data Fig. 1c, d). Interestingly, the binding constant for deoxyU_{15} was comparable to U_{15}, which further illustrates its preference for uracil (Fig. 1b). We extended our analysis to the known biological substrate, pre-let-7. Dis3L2 degraded pre-let-7-U_{15} more efficiently than the unmodified pre-let-7 hairpin and had a 10-fold higher affinity for pre-let-7-U_{15} over pre-let-7 (Fig. 1a, b). Taken together, both enhanced activity and tighter binding play a role in Dis3L2's notable preference for oligoU.

To understand Dis3L2's substrate preference we determined the crystal structure of mouse Dis3L2 in complex with an oligoU RNA substrate (Fig. 1c and Extended Data Table 1). Full length Dis3L2 failed to crystallize. We therefore crystallized a truncated form of Dis3L2, encompassing residues 37–857 with two flexible loops removed (residues 148–169 and 194–221). The construct also contained an inactivating mutation (D389N) in the active site to trap oligoU (Extended Data Fig. 2). The truncated form of Dis3L2 with an intact active site had comparable substrate preference and activity as the full-length enzyme (Extended Data Fig. 3a). Unambiguous electron density accounted for the modeling of 14 nucleotides of oligoU (U_{14}) (Extended Data Fig. 3b, see full description in Methods). Dis3L2 is composed of two cold shock domains at its N-terminus (CSD1 and CSD2), which are intimately associated with each other (Extended Data Fig. 4a, b), and form a lobe on the “top” side of the catalytic RNB domain, while the C-terminal S1 domain, sits opposite the CSD lobe (Fig. 1c, d and Extended Data Fig. 4c). The CSD lobe and S1 form an open funnel

on “top” of the RNB, where RNA substrates gain entry to the enzyme (Fig. 1c–e). The entire length of U_{14} stretches from the opening of the funnel at the “top” and threads through the RNB to the active site. The 5'-half of U_{14} (U1–U7) wraps around the mouth of the funnel, primarily interacting with loops protruding from the CSD lobe and S1 (Fig. 1d and Extended Data Fig. 4d). This is followed by U8, which resides in a pocket formed by S1 and RNB at the stem of the funnel (Extended Data Fig. 4c). The final 6 nucleotides, U9–U14, are stacked within the RNB core (Fig. 1c–e). The 3'-end nucleotides, U13–U14, reside in the active site in the “middle” of the RNB domain (Fig. 1c), where there is a solvent accessible escape “hole” through the side of the enzyme for nucleotide products to exit. We measured specific activity and binding affinity for oligoU substrates of increasing length (9–15 nucleotides) and determined that the optimal oligoU substrate (U_{15}) spans the entire length of the enzyme (Extended Data Fig. 5). The Dis3L2-U_{14} structure, therefore represents the optimal length oligoU tail required to recruit Dis3L2 and elicit the enhanced degradation of pre-let-7[13,14], and is consistent with the length of oligoU tail added to pre-let-7 by TUT4/7[10–12].

Dis3L2 is structurally similar to yRrp44[23–25] except for its lack of an N-terminal PIN domain, which mediates yRrp44 association with the exosome (Fig. 2a, b). Both Dis3L2 and yRrp44 resemble *E. coli* RNase II[26] (Extended Data Fig. 6), and share processive 3' to 5' exonuclease activity, catalyzed by a universally conserved Mg^{2+}-dependent active site (Extended Data Fig. 6a, b). Individually, the CSD1 (42% sequence identity, 0.9 Å rmsd), CSD2 (18% sequence identity, 1.5 Å rmsd) and S1 (27% sequence identity, 1.9 Å rmsd) domains of Dis3L2 align well with the corresponding domains in yRrp44 and are positioned similarly on one face of the RNB domain (39% sequence identity, 1.4 Å rmsd) (Fig. 2a, b). However, yRrp44 has a narrower pore on the “top” of its RNB compared to Dis3L2, due to closure of its CSD lobe over the “top” of the RNB to the S1. Consequently, RNA enters Rrp44 through a “side” path formed between CSD1 and the RNB[23,24] as it exits the exosome core (Exo10) (Fig. 2b). For Dis3L2, the funnel formed by the CSD lobe and S1 closes the “side” path taken in Rrp44, making it inaccessible to RNA substrates. Furthermore, the mouth of the funnel is lined with a positively charged electrostatic surface that provides an appropriate binding site for RNA substrates (Fig. 2c, d). The path taken by U_{14} in Dis3L2 is more reminiscent of the RNA path in the structure of RNase II[22] (Extended Data Fig. 6c–f). However, the opening at the “top” between the CSD lobe and S1 domain is narrower in the single-strand specific RNase II enzyme[26] (Extended Data Fig. 6d, f), which precludes structured RNA substrates from entering.

The substrate preference exhibited by Dis3L2 is explained by numerous specific interactions with the uracil base of an oligoU-tailed RNA substrate (Fig. 3a, b). While the path that the RNA must take to the RNB domain is different in Dis3L2 compared to yRrp44, the paths converge from U9–U14 within the narrow pore of the RNB domain (Fig. 2a, b). In this region, Dis3L2 residues primarily contact the RNA backbone and are generally conserved with yRrp44 (Fig. 3b). On the other hand, where the RNA paths diverge (U1–U8), we identified an extensive network of interactions between Dis3L2 and the uracil bases of U_{14}, which are not observed in structures of yRrp44[23,24] and RNase II[26] (Fig. 3b). We categorized these interactions into three U-specificity zones (U-zone 1–3) that together discriminate uracil from adenine and cytosine (Fig. 3c–d and Extended Data Fig. 7). U-zone

1, comprised of residues from the CSD lobe and S1 at the mouth of the funnel, includes U1–U4 (Fig. 3b, c). U1 is stacked between the side chains of R275 and F80, and forms hydrogen bonds with the side chain of H271 and the main chain of H271 and R275. U2 is stacked with H271 and held in place by main chain hydrogen bonds with F80 and P77. U3 and U4 interact exclusively with residues from the S1 domain (Fig. 3b, c). U3 is stacked on Y794 and interacts with the side chain of N796. U4 forms hydrogen bonds with the side chain of N777 and main chain of A779 and Q778 (Fig. 3b, c).

The bases of U6-U8 reside in the narrow stem of the funnel at the interface between the CSD lobe, S1 and RNB to make up U-zone 2 (Fig. 3b, d). U6 and U7 mediate the interaction between CSD1 and the RNB (Extended Data Fig. 4a). They are stacked and sandwiched by F84 of CSD1 and M615 of the RNB. The CSD1 side chain of R74 interacts with U6 while the side chains of CSD1-D93 and RNB-Q612 contact U7. U8 resides in a pocket at the RNB-S1 interface (Extended Data Fig. 4c), formed by Q551 and N661 of the RNB and Q790 from S1. The side chains of Q551 and Q790 make direct hydrogen bonds with U8, and are stabilized by the side chain of N661 (Fig. 3b, d).

Finally, U-zone 3 includes residues from the RNB core that interact with U9-U12 (Fig. 3b, e). The base edge of U9 is engaged in Watson-Crick like pairing with the main chain of L549 and Q551. An extended network of interactions is made between U10-U12 with the main chain of L549 and side chains of D550 and K553 (Fig. 3b, e).

Altogether we identified 22 U-specific hydrogen bonds (9 from main-chain atoms and 13 from side-chain interactions) between Dis3L2 and the uracil bases of the first 12 nucleotides of U_{14}. Most of these interactions are disrupted when we modeled A_{14} and C_{14} into the Dis3L2 structure (Extended Data Fig. 7), effectively providing an explanation for why oligoA and oligoC are poor substrates compared to oligoU. To examine the role of Dis3L2 residues in each U-zone, we mutated selected residues and measured activity with U_{15} (Fig. 3f and Extended Data Fig. 8). U-zone 1 residues primarily interact with U1-U2 through main chain atoms, which explains why side chain mutations displayed moderate impairment (R275A). Mutations of the side chains in the S1 domain that interact with U3-U4 were more variable. N796A displayed wild-type levels of activity, while mutation of the same residue to its hDis3 counterpart (N796E) is impaired (Fig. 3c, f). We suspect that the longer, negatively charged residue at this position impedes RNA progression through the enzyme compared to alanine. Mutation of Q778A is the most defective in U-zone 1 and we speculate that Q778 might be involved in the translocation of U3 to the U4 position (Fig. 3 c, f). Unexpectedly, we identified an activating mutation (N777A) in U-zone 1, which accounts for an enzyme with 40% higher levels of activity than wild-type Dis3L2. U4 is held in place through main chain interactions in addition to the side chain of N777, so that mutation to a small aliphatic residue maintains part of the interaction, but may create a hollow path to allow RNA to pass unimpeded to U-zone 2 (Fig. 3c, f). Similar “super-enzyme” mutations of functional to small aliphatic residues have also been described for RNase II[27] and yRrp44[28].

Intriguingly, two of the most impaired mutants (R74A/Q612A and Q790A) in U-zone 2 are of side chains that are not only reading the RNA sequence but are engaged in oligoU

mediated domain-domain interfaces (Fig. 3d, f and Extended Data Fig. 4a). R74 and Q612 stabilize the stacked conformation of U6-U7 and sit at the junction of the RNA stabilized CSD1-RNB interface. Q790 facilitates an RNA mediated interface between the S1 and RNB domains (Fig. 3d, f). Collectively, these data strongly support the U_{14} path and the domain configuration in the Dis3L2 structure.

U-zone 3 appears to be the least U-specific zone, as we had to mutate a stretch of 4 residues (R548-K553) to the mammalian Dis3 sequence (Fig. 3e, f) to achieve a 40 % reduction in activity. U-zone 3 reads uracil through a mixture of main chain and side chains, but is mostly composed of non-sequence specific interactions with the U_{14} backbone (Fig. 3b), which may explain the relatively modest effect on activity for this mutant.

All of the U-zone mutants examined are active, which is not entirely surprising given the extensive network of interactions maintained throughout all of the U-zones, even in the context of a single point mutation. This is especially evident when we compared the specific activity of any single U-zone mutation, where 1–2 interactions are broken, with A_{15} or C_{15}, where the bulk of the U-zone network is disrupted (Fig. 3f and Extended Data Fig. 7). To clarify the impact of U-zone mutations on Dis3L2's processing mechanism; we measured the binding affinity for U_{15} (Fig. 3g and Extended Data Fig. 9). There is little correlation between binding and activity here, given that U-zone mutations had little impact on the binding affinity for U_{15} compared to wild-type Dis3L2 (Fig. 3g). Since single U-zone mutations are not sufficient to abolish oligoU binding, again because of an overwhelming network of interactions, it may rather be the rate of oligoU translocation through the mutated U-zones to the active site that is impeded.

The structure of Dis3L2 presented here answers several important mechanistic questions about the biochemical function of this essential enzyme. Most importantly, our study identified a vast network of oligoU specific interactions, even when compared to a typical transcription factor binding to its recognition sequence. This accounts for Dis3L2's enhanced activity and higher affinity for oligoU-tailed substrates (Fig. 1a, b). The shape of the binding funnel—wide at the top and narrow at the bottom (Fig. 1c–e and Fig. 2a, c)—explains its ability to process structured RNA substrates. Indeed, Dis3L2 is more adept at processing structured RNA, even those with short 3' end overhangs, compared to hDis3[20] and yRrp44[20]. We propose that the CSD-S1 funnel is wide enough to bind structured RNA. In turn, the CSDs and S1, which are known to function as RNA chaperones in other proteins[29], may promote the remodeling of structured RNA. Our structure supports this model quite well, as we observe a far-reaching set of base interactions, including base-stacking and hydrophilic interactions throughout the CSD-S1 funnel that could play this role (Fig. 3b–e). Once the U-zones are primed with an oligoU tail, Dis3L2 will degrade it up through the RNB, where non-sequence specific interactions (Fig. 3b) and hydrolysis in the active site fuel translocation of RNA substrates, like pre-let-7, through the enzyme. In conclusion, our data provides the structural mechanism of substrate recognition that underlies Dis3L2s role as the effector in maintaining pluripotency via the Lin28/let-7 pathway.

Methods

Protein Preparation

Mouse Dis3L2 was expressed in Sf9 cells as an N-terminal Strep-sumo-TEV fusion protein from the pFL vector of the MultiBac baculovirus expression system[30]. Sf9 cells were infected with baculovirus in Hyclone CCM3 media (Thermo Scientific) at 27°C. Following 60 hours of expression the cells were centrifuged at 1200 rpm and resuspended in Wash buffer (50 mM Tris pH 8, 100 mM NaCl, and 5 mM DTT), flash frozen in liquid nitrogen and stored at −80°C. Frozen cells were thawed, NaCl concentration increased to 500 mM, and then lysed by sonication. The lysate was treated with 0.2% poly-ethylene imine (PEI) to precipitate nucleic acids prior to ultracentrifugation at 35,000 RPM at 4°C for 1 hour. The soluble fraction was incubated with 1 mL of Strep-Tactin superflow resin (IBA bioTAGnology) per 10 mL of lysate for 1 hour on a rolling shaker. The resin was applied to a gravity flow column and washed extensively with Wash buffer. The protein was eluted with Wash buffer containing 2 mM desthiobiotin. The eluted fraction was treated with TEV protease overnight at 4°C. The cleavage efficiency and purity was verified by SDS-PAGE. The cleaved protein was diluted with an equal volume of Heparin buffer A (25 mM Hepes pH 7.5, 5 mM DTT) to a final NaCl concentration of 50 mM. Dis3L2 was loaded onto a HiTrap Heparin HP column (GE Life Science) equilibrated with 25 mM Hepes pH 7.5, 50 mM NaCl and 5 mM DTT. A linear gradient between 0.05 M and 1 M NaCl was used to elute Dis3l2 at 0.25 M NaCl. Fractions that contained Dis3L2 were analyzed by SDS-PAGE, pooled and concentrated to 2 mL and loaded onto a HiLoad 16/60 Superdex 200 gel filtration column equilibrated with 10 mM Tris pH 8, 100 mM NaCl, 2 mM $MgCl_2$ and 5 mM DTT. Fractions containing Dis3L2 were pooled and concentrated to 25 mg/mL, flash frozen in liquid nitrogen and stored at −80°C. Mutants of Dis3L2 were constructed by sequence and ligation-independent cloning[31] (SLIC) using mutant primers. All mutant proteins were expressed and purified as described for wild-type Dis3L2.

Crystallization

Full length Dis3L2 was recalcitrant to crystallization despite extensive screening. We identified a protease sensitive loop within CSD1 by limited proteolysis with thermolysin. We expressed and purified a D389N active site mutant lacking the protease-sensitive loop within CSD1 (residues 148–169 and 194–221), and both the N (1–36) and C (857–870) termini (Extended Data Figure 2). The truncated Dis3L2 was incubated with a 1.2 molar excess of U_{13} RNA for 30 minutes at 20°C. Crystallization was carried out by the hanging drop vapor diffusion method by mixing the Dis3L2-U_{13} complex at 15 mg/mL with an equal volume of 100 mM tri-ammonium citrate pH 5.5, 100–200 mM ammonium chloride and 20–22% PEG 3350. Crystals appeared in 1–2 days at 18°C and were placed directly in a freshly prepared solution composed of the crystallization condition supplemented with 20% glycerol and flash frozen in liquid N_2.

Structure determination

X-ray diffraction data were collected to 2.95-Å resolution at the X25 beamline at the National Synchrotron Light Source (NSLS) at Brookhaven National Laboratory (BNL) (Extended Data Table 1). Data were processed with XDS[32] as implemented in the

autoPROC software[33]. We generated a molecular replacement (MR) search model by pruning yeast Rrp44[23] (PDB 2VNU) with the Sculptor[34] utility. The pruned model was separated into its individual CSD1, CSD2, RNB and S1 domains. Phasing was determined by MR with Phaser[35], using each of the domains of yeast Rrp44 as individual search models. Phaser found solutions for CSD2, RNB and S1 domains, but failed to find a unique solution for CSD1. The initial MR phases were input into the AutoBuild routine in Phenix[36], where the Resolve[37] and Buccaneer[38] option for density modification and model building, respectively, was critical for phase improvement and interpretation of initial electron density maps. The CSD1 of yRrp44 was placed manually into the initial map and the AutoBuild routine repeated. The resulting phases were of sufficient quality to manually build protein and RNA in Coot[39]. Although a U_{13} homo-oligonucleotide was used in crystallization, interpretable electron density accounted for 14 nucleotides. This is readily explained by slipping of the homo-oligoU into two conformations along the RNA binding cleft. One conformation is fully threaded into the active site (at the 0 position), and a second conformation occupies the −1 position. Crystallographically, these two states are indistinguishable, and thus is best accounted for by modeling 14 nucleotides. The final model includes two copies of truncated Dis3L2 with the exception of disordered residues (37–48, 119–228 and 252–258), two U_{14} molecules, two Mg^{2+} ions and 6 water molecules. The structure was refined using automatically determined NCS restraints, TLS and isotropic B-factor refinement as implemented in Phenix[40] to an R/R_{free} of 21/25 %. The final model was validated with MolProbity[41], where 96.8 % of residues reside in the favored and 3.13 % in the allowed regions of the Ramachandran plot (Extended Data Table 1). The copy represented by chain A (Dis3L2) and chain C (U_{14}) had appreciably better electron density and was used for all structural analysis described in this manuscript. Weak electron density, presumably due to disorder in the region surrounding U5, precluded accurate modeling of the position of this nucleotide, so we carefully avoided over interpretation of the interactions observed in this region, though we are confident of its presence and general location. All structural figures were generated with the PyMOL Molecular Graphics System, Version 1.6 Schrödinger, LLC.

Specific Activity Determination

All exonuclease assays were conducted under multiple turnover conditions, where we measured the linear initial rate of RNA degradation and calculated the specific activity (nt degraded $minute^{-1}$ molecule of $enzyme^{-1}$). The indicated amount of Dis3L2 was incubated at 30°C with 100 nM 5' radiolabeled RNA substrate in reaction buffer (20 mM Hepes pH 7.5, 50 mM NaCl, 1 mM DTT and 100 uM $MgCl_2$). A 50 µl reaction was initiated and 5 µl was removed and quenched at 0, 0.5, 1, 2, 4, 8 and 16 minute time points in formamide loading buffer (0.025% SDS, 95% formamide, 0.025% bromophenol blue, 0.025% xylene cyanol and 18 mM EDTA). Quenched samples were heated to 95°C for 2 minutes and resolved by denaturing urea PAGE. Products were analyzed by phosphor imaging and quantified with ImageJ software. The amount of substrate degraded (fmol) was plotted vs. time and linear regression was used to determine the initial rate (slope) of the linear portion of the curve with GraphPad Prism 6. The initial rate was converted specific activity (nt/min/molecule of enzyme) by dividing the initial rate (fmol substrate processed/min) by the fmol of enzyme used in each assay and multiplying by N-3 (where N is the number of nucleotides

of the substrate and 3 is the length of the end-product). Synthetic oligoU, oligoA and oligoC substrates were purchased from Trilink BioTechnologies. Pre-let-7 and pre-let-7-U_{15} were *in vitro* transcribed. Briefly, the RNA coding sequence was flanked by a 5' hammerhead ribozyme and 3' hepatitis delta virus ribozyme (HDV) to ensure homogeneous ends[42,43]. Constructs were cloned into a pRSF vector containing a T7 promoter and the RNA produced by run-off transcription. The pre-let-7 sequence used in this study is pre-let7a1: UGAGGUAGUAGGUUGUAUAGUUUUAGGGUCACACCCACCACUGGGAGAUAACUAUACAAUCUACUGUCUUUC Transcribed RNA was gel purified with denaturing PAGE and resuspended in DEPC-treated water.

Equilibrium Binding Assays

A range of Dis3L2 concentrations; 0–5 nM for oligoU, pre-let-7-U_{15}, deoxyU_{15} and pre-let-7 substrates and 0–150 nM for A_{15} and C_{15} were incubated with 1 pM 5' radiolabeled RNA for 1 hour in binding buffer (20 mM Hepes pH 7.5, 50 mM NaCl, 1 mM DTT, 50 μg/mL BSA and 100 μM EDTA). We determined that wild-type Dis3L2 is inactive in the presence of EDTA and no magnesium, but binds to RNA with the same affinity as the inactive D389N mutant. A slot blot filtration system was used (BioRad) to capture Dis3L2/RNA complexes on the top nitrocellulose membrane and unbound RNA on the bottom nylon membrane. The membranes were washed with 100 μl of binding buffer, prior to applying 100 μL of the binding reaction, followed by 100 μL of binding buffer. The nitrocellulose and nylon membranes were dried and analyzed by phosphor imaging. All binding assays were conducted in triplicate, were quantified and fraction bound plotted vs. free protein concentration. K_d values were determined by using nonlinear regression analysis with Graphpad Prism6 software.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank Christopher M. Hammell for comments on this manuscript and members of the Joshua-Tor laboratory for helpful comments and suggestions. We thank Annie Héroux for help at the National Synchrotron Light Source, which is supported by Department of Energy, Office of Basic Energy Sciences. We also thank the Protein Core Facility at Columbia University. This work was supported by the Watson School of Biological Sciences (to J.W. and L.J.), the Louis Morin Charitable Trust and the Robertson Research Fund of Cold Spring Harbor Laboratory (to L.J.). L.J. is an investigator of the Howard Hughes Medical Institute.

References

1. Viswanathan SR, Daley GQ, Gregory RI. Selective blockade of microRNA processing by Lin28. Science (New York, N.Y. 2008; 320:97–100. doi:10.1126/science.1154040.
2. Heo I, et al. Lin28 mediates the terminal uridylation of let-7 precursor MicroRNA. Molecular cell. 2008; 32:276–284. doi:10.1016/j.molcel.2008.09.014. [PubMed: 18951094]
3. Rybak A, et al. A feedback loop comprising lin-28 and let-7 controls pre-let-7 maturation during neural stem-cell commitment. Nature cell biology. 2008; 10:987–993. doi:10.1038/ncb1759. [PubMed: 18604195]
4. Newman MA, Thomson JM, Hammond SM. Lin-28 interaction with the Let-7 precursor loop mediates regulated microRNA processing. RNA (New York, N.Y. 2008; 14:1539–1549. doi: 10.1261/rna.1155108.

5. Thornton JE, Gregory RI. How does Lin28 let-7 control development and disease? Trends in cell biology. 2012; 22:474–482. doi:10.1016/j.tcb.2012.06.001. [PubMed: 22784697]
6. Zhu H, et al. The Lin28/let-7 axis regulates glucose metabolism. Cell. 2011; 147:81–94. doi: 10.1016/j.cell.2011.08.033. [PubMed: 21962509]
7. Shyh-Chang N, et al. Lin28 enhances tissue repair by reprogramming cellular metabolism. Cell. 2013; 155:778–792. doi:10.1016/j.cell.2013.09.059. [PubMed: 24209617]
8. Urbach A, et al. Lin28 sustains early renal progenitors and induces Wilms tumor. Genes Dev. 2014 doi:10.1101/gad.237149.113.
9. Nam Y, Chen C, Gregory RI, Chou JJ, Sliz P. Molecular basis for interaction of let-7 microRNAs with Lin28. Cell. 2011; 147:1080–1091. doi:10.1016/j.cell.2011.10.020. [PubMed: 22078496]
10. Hagan JP, Piskounova E, Gregory RI. Lin28 recruits the TUTase Zcchc11 to inhibit let-7 maturation in mouse embryonic stem cells. Nature structural & molecular biology. 2009; 16:1021–1025. doi:10.1038/nsmb.1676.
11. Heo I, et al. TUT4 in concert with Lin28 suppresses microRNA biogenesis through pre-microRNA uridylation. Cell. 2009; 138:696–708. doi:10.1016/j.cell.2009.08.002. [PubMed: 19703396]
12. Thornton JE, Chang HM, Piskounova E, Gregory RI. Lin28-mediated control of let-7 microRNA expression by alternative TUTases Zcchc11 (TUT4) and Zcchc6 (TUT7). RNA (New York, N.Y. 2012; 18:1875–1885. doi:10.1261/rna.034538.112.
13. Chang HM, Triboulet R, Thornton JE, Gregory RI. A role for the Perlman syndrome exonuclease Dis3l2 in the Lin28-let-7 pathway. Nature. 2013; 497:244–248. doi:10.1038/nature12119. [PubMed: 23594738]
14. Ustianenko D, et al. Mammalian DIS3L2 exoribonuclease targets the uridylated precursors of let-7 miRNAs. RNA (New York, N.Y. 2013; 19:1632–1638. doi:10.1261/rna.040055.113.
15. Liu Q, Greimann JC, Lima CD. Reconstitution, activities, and structure of the eukaryotic RNA exosome. Cell. 2006; 127:1223–1237. doi:10.1016/j.cell.2006.10.037. [PubMed: 17174896]
16. Dziembowski A, Lorentzen E, Conti E, Seraphin B. A single subunit, Dis3, is essentially responsible for yeast exosome core activity. Nature structural & molecular biology. 2007; 14:15–22. doi:10.1038/nsmb1184.
17. Tomecki R, et al. The human core exosome interacts with differentially localized processive RNases: hDIS3 and hDIS3L. The EMBO journal. 2010; 29:2342–2357. doi:10.1038/emboj.2010.121. [PubMed: 20531386]
18. Astuti D, et al. Germline mutations in DIS3L2 cause the Perlman syndrome of overgrowth and Wilms tumor susceptibility. Nature genetics. 2012; 44:277–284. doi:10.1038/ng.1071. [PubMed: 22306653]
19. Malecki M, et al. The exoribonuclease Dis3L2 defines a novel eukaryotic RNA degradation pathway. The EMBO journal. 2013; 32:1842–1854. doi:10.1038/emboj.2013.63. [PubMed: 23503588]
20. Lubas M, et al. Exonuclease hDIS3L2 specifies an exosome-independent 3'–5' degradation pathway of human cytoplasmic mRNA. The EMBO journal. 2013; 32:1855–1868. doi:10.1038/emboj.2013.135. [PubMed: 23756462]
21. Chang H, Lim J, Ha M, Kim VN. TAIL-seq: Genome-wide Determination of Poly(A) Tail Length and 3' End Modifications. Molecular cell. 2014 doi:10.1016/j.molcel.2014.02.007.
22. Choi YS, Patena W, Leavitt AD, McManus MT. Widespread RNA 3'-end oligouridylation in mammals. RNA (New York, N.Y. 2012; 18:394–401. doi:10.1261/rna.029306.111.
23. Lorentzen E, Basquin J, Tomecki R, Dziembowski A, Conti E. Structure of the active subunit of the yeast exosome core, Rrp44: diverse modes of substrate recruitment in the RNase II nuclease family. Molecular cell. 2008; 29:717–728. doi:10.1016/j.molcel.2008.02.018. [PubMed: 18374646]
24. Makino DL, Baumgartner M, Conti E. Crystal structure of an RNA-bound 11-subunit eukaryotic exosome complex. Nature. 2013; 495:70–75. doi:10.1038/nature11870. [PubMed: 23376952]
25. Bonneau F, Basquin J, Ebert J, Lorentzen E, Conti E. The yeast exosome functions as a macromolecular cage to channel RNA substrates for degradation. Cell. 2009; 139:547–559. doi: 10.1016/j.cell.2009.08.042. [PubMed: 19879841]

26. Frazao C, et al. Unravelling the dynamics of RNA degradation by ribonuclease II and its RNA-bound complex. Nature. 2006; 443:110–114. doi:10.1038/nature05080. [PubMed: 16957732]
27. Barbas A, et al. Determination of key residues for catalysis and RNA cleavage specificity: one mutation turns RNase II into a “SUPER-ENZYME”. The Journal of biological chemistry. 2009; 284:20486–20498. doi:10.1074/jbc.M109.020693. [PubMed: 19458082]
28. Reis FP, et al. Modulating the RNA processing and decay by the exosome: altering Rrp44/Dis3 activity and end-product. PLoS ONE. 2013; 8:e76504. doi:10.1371/journal.pone.0076504. [PubMed: 24265673]
29. Rajkowitsch L, et al. RNA chaperones, RNA annealers and RNA helicases. RNA biology. 2007; 4:118–130. [PubMed: 18347437]
30. Bieniossek C, Richmond TJ, Berger I. MultiBac: multigene baculovirus-based eukaryotic protein complex production. Curr Protoc Protein Sci. 2008; 5(5 20) doi:10.1002/0471140864.ps0520s51.
31. Li MZ, Elledge SJ. Harnessing homologous recombination in vitro to generate recombinant DNA via SLIC. Nat Methods. 2007; 4:251–256. doi:10.1038/nmeth1010. [PubMed: 17293868]
32. Kabsch W. Xds. Acta Crystallogr D Biol Crystallogr. 2010; 66:125–132. doi:10.1107/S0907444909047337. [PubMed: 20124692]
33. Vonrhein C, et al. Data processing and analysis with the autoPROC toolbox. Acta Crystallogr D Biol Crystallogr. 2011; 67:293–302. doi:10.1107/S0907444911007773. [PubMed: 21460447]
34. Bunkoczi G, Read RJ. Improvement of molecular-replacement models with Sculptor. Acta Crystallogr D Biol Crystallogr. 2011; 67:303–312. doi:10.1107/S0907444910051218. [PubMed: 21460448]
35. McCoy AJ, et al. Phaser crystallographic software. J Appl Crystallogr. 2007; 40:658–674. doi:10.1107/S0021889807021206. [PubMed: 19461840]
36. Terwilliger TC, et al. Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard. Acta Crystallogr D Biol Crystallogr. 2008; 64:61–69. doi:10.1107/S090744490705024X. [PubMed: 18094468]
37. Terwilliger TC. Maximum-likelihood density modification. Acta Crystallogr D Biol Crystallogr. 2000; 56:965–972. [PubMed: 10944333]
38. Cowtan K. The Buccaneer software for automated model building. 1. Tracing protein chains. Acta Crystallogr D Biol Crystallogr. 2006; 62:1002–1011. doi:10.1107/S0907444906022116. [PubMed: 16929101]
39. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr. 2004; 60:2126–2132. doi:10.1107/S0907444904019158. [PubMed: 15572765]
40. Adams PD, et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr D Biol Crystallogr. 2010; 66:213–221. doi:10.1107/S0907444909052925. [PubMed: 20124702]
41. Davis IW, et al. MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. Nucleic Acids Res. 2007; 35:W375–383. doi:10.1093/nar/gkm216. [PubMed: 17452350]
42. Price SR, Ito N, Oubridge C, Avis JM, Nagai K. Crystallization of RNA-protein complexes. I. Methods for the large-scale preparation of RNA suitable for crystallographic studies. Journal of molecular biology. 1995; 249:398–408. [PubMed: 7540213]
43. Shechner DM, Bartel DP. The structural basis of RNA-catalyzed RNA polymerization. Nature structural & molecular biology. 2011; 18:1036–1042. doi:10.1038/nsmb.2107.

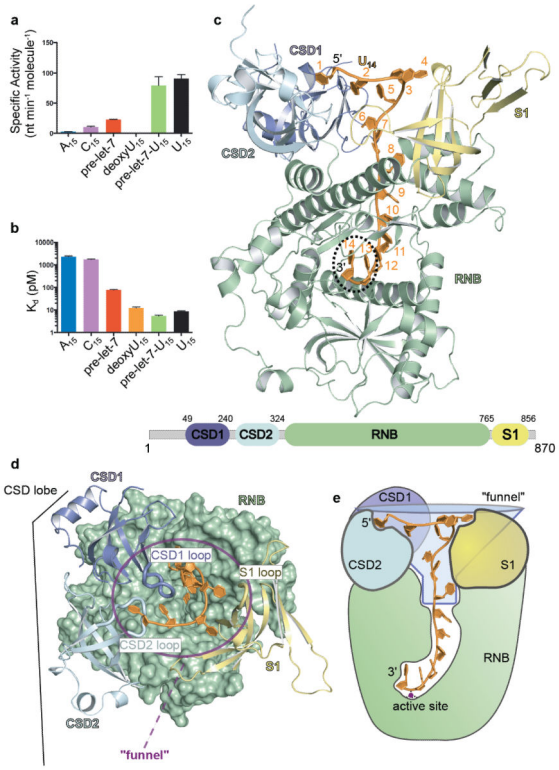

Figure 1. Structure of Dis3L2 in complex with U_{14} RNA

a, Specific activity (nt min^{-1} $molecule^{-1}$) of Dis3L2 with different RNA substrates. Specific activity was determined from the initial rate (fmol substrate degraded/minute) divided by the amount of enzyme (fmol) multiplied by N-3 (where N is the number of nucleotides of the substrate and 3 is the length of the end-product). Mean ± s.d. (n = 3) are shown. **b**, The equilibrium dissociation constant, K_d, determined by slot blot filter binding assay with different substrates conducted in triplicate. The K_d ± s.d. for each substrate is plotted. **c**, Overall structure of Dis3L2 in complex with U_{14} and a domain schematic. Every domain of Dis3L2 makes contact to the RNA. CSD1, CSD2 (CSD lobe) and S1 are oriented in an open arrangement to form a funnel, where the 5' nucleotides U1–U8 bind, funneling U9–U14 within the core of the catalytic RNB domain. U13 and U14 are positioned in the active site at the bottom of the RNB (black circle). **d**, Overall view of the “top” of the Dis3L2-U_{14} structure, rotated 90 degrees around the horizontal axis compared to panel c. The RNB is shown in surface view. CSD1 and CSD2 form a lobe (CSD lobe) opposite the S1 and are rendered as cartoons. All domains are color-coded as in panel c. A purple circle represents the mouth of the CSD-S1 funnel. CSD1, CSD2 and S1 make contact to U_{14} with residues primarily located on loops that protrude into the mouth of the funnel. **e**, Cartoon model of the funnel shaped Dis3L2-U_{14} complex.

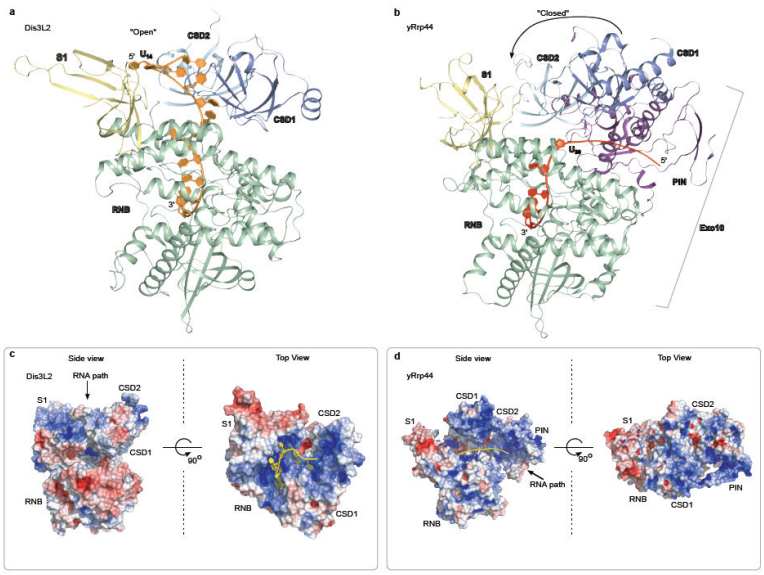

Figure 2. Comparison of Dis3L2 to yRrp44 of the RNA exosome

a, Dis3L2-U_{14} complex rotated 180° around the vertical axis compared to Fig.1c. The open funnel created by the CSD lobe and S1 allows RNA to access the "top" of the RNB. **b**, The structure of yeast Rrp44 extracted from the 11 subunit RNA exosome[20] (PDB 4IFD). The CSD lobe closes on the S1 domain leading to a closed conformation on top of the RNB, as a result a "side" RNA binding path is created between the RNB and CSD1. **c**, Perpendicular "side" and "top" views of the electrostatic surface potential of Dis3L2 (contoured at ±5 kT/e, white=neutral, blue=positive and red=negative). A positively charged electrostatic surface lines the wide portion of the funnel on the "top" of the RNB that can accommodate structured RNA substrates. **d**, Electrostatic surface potential of the exosome-associated yRrp44, in the same configuration as panel c, supports an RNA binding path along the "side" route.

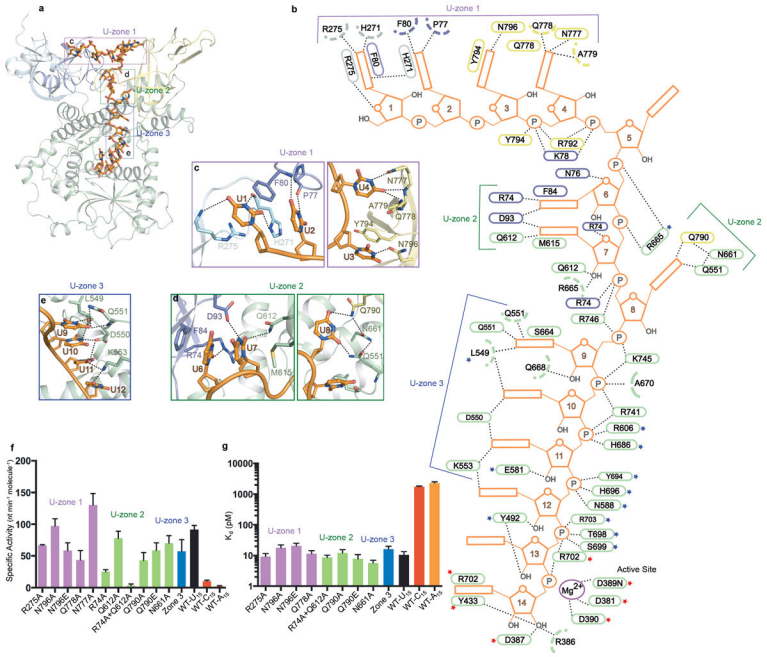

Figure 3. Mechanism of Dis3L2's oligoU specificity

a, Overall structure of Dis3L2-U_{14} with the location of three uracil specificity zones (U-zone 1–3) labeled and colored purple (U-zone 1), green (U-zone 2) and blue (U-zone 3). **b**, A schematic view of the interactions between U_{14} and Dis3L2. The interactions between U_{14} (orange) with side chain (ovals) and main chain (dashed semi-circles) residues of Dis3L2 are labeled and color-coded by domain. Dashed lines represent hydrogen bonds and/or ion pairs. The active site of Dis3L2 is labeled. Active site residues conserved with yRrp44 are highlighted with red stars. Conserved residues that interact with the RNA backbone are labeled with blue stars. U-zones are labeled as in panel a. **c–e**, Detailed U-specific interactions in U-zones 1–3 between Dis3L2 and the uracil bases of U1–U12. **c**, The bases of U1–U4 in U-zone 1. The uracil bases are shown as orange sticks and Dis3L2 residues are sticks, color-coded by domain as in Fig. 1c. Hydrogen bonds are shown as dashed lines. **d**, U-zone 2 encompassing U6–U8. **e**, U9–U12 in U-zone 3. **f**, Specific activity (nt min^{-1} $molecule^{-1}$) of selected U-zone mutants with U_{15} substrate compared to wildtype with U_{15}, C_{15} and A_{15}. Specific activity was measured with 0.5 nM of indicated mutant and 100 nM U_{15}. Mean ± s.d. (n = 3) are shown. **g**, Equilibrium dissociation constants, K_d, of U-zone mutants with U_{15}, compared to wildtype with U_{15}, C_{15} and A_{15}. All assays were conducted in triplicate. The K_d ± s.d. for each mutant is plotted.