The pluripotency factor Lin28 inhibits the biogenesis of the let-7 family of mammalian microRNAs. Lin28 is highly expressed in embryonic stem cells and has a fundamental role in regulation of development, glucose metabolism and tissue regeneration. Over-expression of Lin28 is correlated with the onset of numerous cancers, whereas let-7, a tumour suppressor, silences several human oncogenes. Lin28 binds to precursor let-7 (pre-let-7) hairpins, triggering the 3'-oligo-uridylation activity of TUT4 and TUT7 (refs 10–12). The oligoU tail added to pre-let-7 serves as a decay signal, as it is rapidly degraded by Dis3l2 (refs 13, 14), a homologue of the catalytic subunit of the RNA exosome. The molecular basis of Lin28-mediated recruitment of TUT4 and TUT7 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 that of its exosome counterpart. The resulting path reveals an extensive network of uracil-specific interactions spanning the first 12 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 functions as the effector nuclease responsible for degrading uridylated pre-let-7 miRNAs in the Lin28–let-7 pathway. Dis3l2 belongs to the RNase II/R 3'-5' exonuclease superfamily, which includes the catalytic subunit of the RNA exosome (Rrp44) in yeast (Rrp44) and humans (Dis3 and Dis3l1). Inactivation of Dis3l2 is associated with aneuploidy and mitotic abnormalities, whereas its overexpression suppresses cancer cell growth. Genetic disruption of DIS3L2 is the primary cause of Perlman’s syndrome, a congenital disorder leading to fetal overgrowth and an increased susceptibility to Wilms’ tumour development. Subsequent studies have shown that Wilms’ tumours, a common paediatric kidney cancer, overexpress Lin28 (ref. 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. This is particularly intriguing in light of studies showing widespread uridylation of mammalian mRNAs 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 towards 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. Related studies showed that Dis3l2 processed structured RNA more efficiently than its exosome counterparts (human DIS3 and yeast Rrp44), with no reported sequence preference. However, the fission yeast Dis3l2 orthologue was shown to exhibit a preference for oligoU. To clarify the extent of mammalian Dis3l2’s substrate preference, we quantitatively measured the specific activity of mouse Dis3l2 for single-stranded RNA substrates (Fig. 1a and Extended Data Fig. 1a, b). Dis3l2 processed U15 ~9-fold and ~40-fold more efficiently than C15 and A15, respectively, which are both relatively poor substrates (Fig. 1a). Dis3l2 failed to process deoxyU15, thus is specific for RNA (Extended Data Fig. 1a). Next, we measured binding affinity, where Dis3l2 exhibited more than 200-fold tighter binding to U15 compared to A15 and C15 (Fig. 1b and Extended Data Fig. 1c, d). Notably, the binding constant for deoxyU15 was comparable to U15, 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–U15 more efficiently than the unmodified pre-let-7 hairpin and had a tenfold higher affinity for pre-let-7–U15 over pre-let-7 (Fig. 1a, b). Taken together, both enhanced activity and tighter binding have a role in Dis3l2’s notable preference for oligoU.

To understand the substrate preference of Dis3l2 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 modelling of 14 nucleotides of oligoU (U14) (Extended Data Fig. 3b, see full description in Methods). Dis3l2 is composed of two cold shock domains at its amino 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 carboxy-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 U14 stretches from the opening of the funnel at the top and threads through the RNB to the active site. The 5’-half of U14 (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. 4e). The final 6 nucleotides, U9–U14, are stacked within the RNB core (Fig. 1c–e). The 3’-end nucleotides, U13–U14, are situated 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 (U13) spans the entire length of the enzyme (Extended Data Fig. 5). The Dis3l2–U13 structure, therefore, represents the optimal length oligoU tail required to recruit Dis3l2 and elicit the enhanced degradation of pre-let-7 (refs 13, 14) and is consistent with the length of oligoU tail added to pre-let-7 by TUT4 and TUT7 (refs 10–12).
Dis3l2 is structurally similar to yeast Rrp44 (refs 23–25) except for its lack of an N-terminal PIN domain, which mediates yeast Rrp44 association with the exosome (Fig. 2a, b). Both Dis3l2 and yeast Rrp44 resemble Escherichia coli RNase II22 (Extended Data Fig. 6a), and share processive 3' to 5' exonuclease activity, catalysed by a universally conserved Mg$^{2+}$-dependent active site (Extended Data Fig. 6a, b). Individualy, the CSD1 (42% sequence identity, 0.9 Å root mean squared deviation (r.m.s.d.)), CSD2 (18% sequence identity, 1.5 Å r.m.s.d.) and S1 (27% sequence identity, 1.9 Å r.m.s.d.) domains of Dis3l2 align well with the corresponding domains in yeast Rrp44 and are positioned similarly on one face of the RNB domain (39% sequence identity, 1.4 Å r.m.s.d.) (Fig. 2a, b). However, yeast Rrp44 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 RNB23-24 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 U14 in Dis3l2 is more reminiscent of the RNA path in the structure of RNase II22 (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 enzyme26 (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 yeast Rrp44, the paths converge from U9 to 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 yeast Rrp44 (Fig. 3b). However, where the RNA paths diverge (U1–U8), we identified an extensive network of interactions between Dis3l2 and the uracil bases of U14 that are not observed in structures of yeast Rrp44 (refs 23, 24) and RNase II26 (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 H721 and R275. U2 is stacked with H271 and is 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 are located 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 is located 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 (Extended Data 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 U14. Most of these interactions are disrupted when we modelled A15 and C15 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 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, whereas mutation of the same residue to its human DI35 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. 3c, f). Unexpectedly, we identified an activating mutation (N777A) in U-zone 1 that 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 (ref. 27) and yeast Rrp44 (ref. 28).

Notably, two of the most impaired mutants (R74A/Q612A and Q790A) in U-zone 2 are of side chains that not only read the RNA sequence but also engage in oligoU-mediated domain–domain interfaces (Fig. 3d, f). 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 U14 path and the domain configuration in the Dis3l2 structure.

U-zone 3 seems to be the least U-specific zone, as we had to mutate a stretch of six residues in Dis3l2 (R548–K553) to the mammalian Dis3 sequence (Fig. 3e, f and Extended Data Fig. 2) 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 U14 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 particularly evident when we compared the specific activity of any single U-zone mutation, where 1–2 interactions are broken, with A15 or C15, 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 U15 (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 U15 compared to wild-type Dis3l2 (Fig. 3g). Because 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 human DIS3 (ref. 20) and yeast Rrp44 (ref. 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 (ref. 29), may promote the remodelling 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 have 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 provide the structural mechanism of substrate recognition that underlies Dis3l2’s role as the effector in maintaining pluripotency via the Lin28–let-7 pathway.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 9 April; accepted 2 June 2014.

1. Viswanathan, S. R., Daley, G. Q. & Gregory, R. I. Selective blockade of microRNA processing by Lin28. Science 320, 97–100 (2008).
2. Heo, I. et al. Lin28 mediates the terminal undylation of let-7 precursor MicroRNA. Mol. Cell 32, 276–284 (2008).
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 Biol. 10, 987–993 (2008).
4. Newman, M. A., Thomson, J. M. & Hammond, S. M. Lin28 interaction with the Let-7 precursor loop mediates regulated microRNA processing. RNA 14, 1539–1549 (2008).
5. Thornton, J. E. & Gregory, R. I. How does Lin28 let-7 control development and disease? Trends Cell Biol. 22, 474–482 (2012).
6. Zhu, H. et al. The Lin28/let-7 axis regulates glucose metabolism. Cell 147, 81–94 (2011).
7. Shyh-Chang, N. et al. Lin28 enhances tissue repair by reprogramming cellular metabolism. Cell 155, 778–792 (2013).
8. Urbach, A. et al. Lin28 sustains early renal progenitors and induces Wilms tumor. Genes Dev. 28, 971–982 (2014).
9. Nam, Y., Chen, C., Gregory, R. I., Chou, J. J. & Sliz, P. Molecular basis for interaction of let-7 microRNAs with Lin28. Cell 147, 1080–1091 (2011).
10. Hagan, J. P., Piskounova, E. & Gregory, R. I. Lin28 recruits the TUTase Zcchc11 to inhibit let-7 maturation in mouse embryonic stem cells. Nature Struct. Mol. Biol. 16, 1021–1025 (2009).
11. Heo, I. et al. TUT4 in concert with Lin28 suppresses microRNA biogenesis through pre-microRNA uridylation. Cell 138, 696–708 (2009).
12. Thornton, J. E., Chang, H. M., Piskounova, E. & Gregory, R. I. Lin28-mediated control of let-7 microRNA expression by alternative TUTases Zcchc11 (TUT4) and Zcchc6 (TUT7). RNA 18, 1875–1889 (2012).
13. Chang, H. M., Triboulet, R., Thornton, J. E. & Gregory, R. I. A role for the Perlman syndrome exonuclease Dis3l2 in the Lin28–let-7 pathway. Nature 497, 244–248 (2013).
14. Ustianenko, D. et al. Mammalian DIS3L2 exoribonuclease targets the uridylated precursors of let-7 microRNAs. RNA 19, 1632–1638 (2013).
15. Liu, Q., Greimann, J. C. & Lima, C. D. Reconstitution, activities, and structure of the eukaryotic RNA exosome. Cell 127, 1223–1237 (2006).
16. Dziembowski, A., Lorentzen, E., Conti, E. & Seraphin, B. A single subunit, Dis3, is essentially responsible for yeast exosome core activity. Nature Struct. Mol. Biol. 14, 22–22 (2007).
17. Tomecki, R. et al. The human core exosome interacts with differentially localized processive RNAases: hDIs3 and hDIs3L. EMBO J. 29, 2342–2357 (2010).
18. Astuti, D. et al. Germine mutations in DIS3L2 cause the Perlman syndrome of overgrowth and Wilms tumor susceptibility. Nature Genet. 44, 277–284 (2012).
19. Malecki, M. et al. The exoribonuclease Dis3L2 defines a novel eukaryotic RNA degradation pathway. EMBO J. 32, 1842–1854 (2013).
20. Lubas, M. et al. Exonuclease hDIS3L2 specifies an exosome-independent 3’–5’ degradation pathway of human cytoplasmic mRNA. EMBO J. 32, 1855–1868 (2013).
21. Chang, H., Lim, J., Ha, M. & Kim, V. N. TAIL-seq: Genome-wide determination of poly(A) tail length and 3’ end modifications. Mol. Cell 53, 1044–1052 (2014).
22. Choi, Y. S., Patena, W., Leavitt, A. D. & McManus, M. T. Widespread RNA 3′-end oligouridylation in mammals. RNA 18, 394–401 (2012).
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 nuclelease family. Mol. Cell 29, 717–728 (2008).
24. Makino, D. L., Baumgartner, M. & Conti, E. Crystal structure of an RNA-bound 11-subunit eukaryotic exosome complex. Nature 495, 70–75 (2013).
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 139, 547–559 (2009).
26. Frãzão, C. et al. Unravelling the dynamics of RNA degradation by ribonuclease II and its RNA-bound complex. Nature 443, 110–114 (2006).
27. Barbas, A. et al. Determination of key residues for catalysis and RNA cleavage specificity: one mutation turns RNase II into a “SUPER-ENZYME”. J. Biol. Chem. 284, 20486–20498 (2009).
28. Reis, F. P. et al. Modulating the RNA processing and decay by the exosome: altering Rrp44/Dis3 activity and end-product. PLoS ONE 8, e76504 (2013).
29. Rajkowitsch, L. et al. RNA chaperones, RNA annealers and RNA helicases. RNA Biol. 4, 118–130 (2007).

Acknowledgements We thank C. M. Hammell for comments on this manuscript and members of the Joshua-Tor laboratory for comments and suggestions. We thank A. 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.

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.

Author Information Coordinates and structure factors for the Dis3l2–U14 complex have been deposited in the Protein Data Bank under accession code 4PMW. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to L.J. (leemor@cshl.edu).
METHODS

Protein preparation. Mouse Dis3l2 was expressed in S9 cells as an N-terminal Strep–sumo–TEV fusion protein from the pFL vector of the MultiBac baculovirus expression system.29 S9 cells were infected with baculovirus in HyClone CCM3 media (Thermo Scientific) at 27°C. Following 60 h of expression, the cells were centrifuged at 1,200 r.p.m. and re-suspended 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 before ultracentrifugation at 35,000 r.p.m. at 4°C for 1 h. The soluble fraction was incubated with 1 ml of Strep-Tactin superflow resin (IBA bioTAgNology) per 10 ml of lysate for 1 h 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 HitTrap 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 analysed by SDS–PAGE, pooled and concentrated to 2 ml and loaded onto an HiLoad 16/60 Superdex 200 gel filtration column equilibrated with 10 mM Tris pH 8, 100 mM NaCl, 2 mM MgCl2 and 5 mM DTT. Fractions containing Dis3l2 were pooled and concentrated to 25 mg ml−1, flash frozen in liquid nitrogen and stored at −80°C. Mutants of Dis3l2 were constructed by sequence and ligation-independent cloning28 (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 Fig. 2). The truncated Dis3l2 was incubated with a 1.2 molar excess of U13 RNA for 30 min at 20°C. Crystallization was carried out by the hanging-drop vapour diffusion method by mixing the Dis3l2–U13 complex at 15 mg ml−1 with 10 mM HEPES pH 7.5, 50 mM NaCl, 1 mM DTT and 100 mM MgCl2. A 50 μl reaction was initiated and 5 μl was removed and quenched at 0, 0.5, 1, 2, 4, 8 and 16 min 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 min and resolved by denaturing urea PAGE. Products were analysed by phosphor imaging and quantified with ImageJ software. The amount of substrate degraded (fmol) was plotted against 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 to 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-7-15 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 homogenous ends34,35. Constructs were cloned into a pRSV 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: UGGAGGUGAAUGGUUGUAUAUUUGGAAACACGCCUCUCUGGGAGUAUCCUAUCUGCUU.

Transcribed RNA was gel purified with denaturing PAGE and re-suspended in DEPC-treated water.

Equilibrium binding assays. A range of Dis3l2 concentrations—0–5 nM for oligoU, pre-let-7–U13, deoxyU13 and pre-let-7 substrates and 0–150 nM for A9 and C15—were incubated with 1 pM 5’ radiolabelled RNA for 1 h in binding buffer (20 mM HEPES pH 7.5, 50 mM NaCl, 1 mM DTT, 50 μg ml−1 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, before applying 100 μl of the binding reaction, followed by 100 μl of binding buffer. The nitrocellulose and nylon membranes were dried and analysed by phosphor imaging. All binding assays were conducted in triplicate, were quantified and fraction bound plotted versus free protein concentration. Kd values were determined by nonlinear regression analysis with GraphPad Prism software.

30. Bieniossek, C., Richardson, T. J. & Berger, I. MultiBac: multigene baculovirus-based eukaryotic protein complex production. Curr. Protoc. Protein Sci. Ch. 5, Unit 5.20 (2010).
31. Li, M. Z. & Elledge, S. J. Harnessing homologous recombination for yeast and mammalian cell-based systems. Yeast 17, 645–654 (2000).
32. McCoy, A. J. et al. Phaser crystallographic software. J. Appl. Crystallogr. 40, 658–674 (2007).
33. Terwilliger, T. C. et al. Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard. Acta Crystallogr. D 64, 61–69 (2008).
34. Price, S. R., Ito, N., Oubridge, C., Avis, J. M. & Nagai, K. Crystallization of RNA-protein complexes on the top nitrocellulose membrane and unbound RNA on the bottom nylon membrane. The membranes were washed with 100 μl of binding buffer, before applying 100 μl of the binding reaction, followed by 100 μl of binding buffer. The nitrocellulose and nylon membranes were dried and analysed by phosphor imaging. All binding assays were conducted in triplicate, were quantified and fraction bound plotted versus free protein concentration. Kd values were determined by nonlinear regression analysis with GraphPad Prism software.
Extended Data Figure 1 | Substrate specificity of Dis3l2. a, Dis3l2 exonuclease assays were conducted with 5' radio-labelled RNA substrates at 30 °C over a 16 min time course and products resolved by denaturing urea PAGE. Two representative gels, one from an exonuclease assay with U15 that was used to quantify Dis3l2 enzymatic activity and another for deoxyU15, showing no detectable activity under the initial rate conditions tested. b, The initial rate plots for each substrate used to calculate specific activity as shown in Fig. 1a. Exonuclease assays were conducted in triplicate and quantified with ImageJ. The concentration of enzyme used to measure the initial rate is indicated on each plot. The amount of substrate degraded (fmol) was plotted against time and the initial rate (slope) was determined with linear regression using GraphPad Prism 6. Mean ± s.d. (n = 3) are shown. c, Representative slot blot filter-binding assay of Dis3l2 with U15. Dis3l2–RNA complexes were captured on the top nitrocellulose membrane and unbound RNA to the bottom nylon membrane. d, The equilibrium dissociation constant (Kd) for each substrate in Fig. 1b was determined by plotting the fraction RNA bound against the concentration of free Dis3l2 and fit by nonlinear regression with GraphPad Prism 6. The mean of three independently measured replicates with error bars representing ± s.d. are shown with the Kd.
Extended Data Figure 2 | Sequence alignment of mouse and human Dis3l2, yeast Rap44 and human Dis3. Dis3l2 is composed of two cold shock domains (CSD1, residues 49–240, α1–β5, and CSD2, residues 241–324, β6–β10, α2) at its N terminus followed by a catalytic RNB domain (residues 325–765, α3–α17, β11–β21) and flanked by an S1 domain (residues 766–856, β22–β29). The secondary structure elements deduced from the structure of mouse Dis3l2 are shown on top of the sequences, colour-coded by domain as in Fig. 1a. The segments of mouse Dis3l2 that were truncated to facilitate crystallization are outlined in a black box. Conserved amino acid residues are coloured blue. Blue diamonds indicate residues interacting with the backbone of U14 RNA and red diamonds denote conserved active-site residues. Residues in mouse Dis3l2 that interact with U14 bases are shaded purple.
Extended Data Figure 3 | Substrate specificity of truncated Dis3l2 and electron density map for U14 RNA. a, Comparison of specific activity (nt min⁻¹ molecule⁻¹) of truncated Dis3l2 construct resembling the one used in crystallization, but without the D389N mutation with the same set of substrates analysed for wild-type Dis3l2. Truncated Dis3l2 had comparable levels of activity and displayed the same substrate preference as wild-type Dis3l2. The specific activity was calculated from the initial rate plots, in the same way as described in Fig. 1a and Extended Data Fig. 1. The mean ± s.d. (n = 3) are shown. b, The final model of Dis3l2 is shown as a transparent white cartoon. The final model of U14 is rendered as yellow sticks. Initial molecular replacement phases were improved by density modification with Resolve and automated model building with Buccaneer, as implemented by the AutoBuild wizard in Phenix. The unbiased density modified electron density map before inclusion of RNA is shown contoured at 1σ. Despite the crystallization of Dis3l2 in complex with U13, unambiguous density allowed modelling of U14 (details are discussed in the Methods). Clear electron density accounted for all 14 nucleotides, except some disorder contributed to weak electron density surrounding U5, precluding its accurate placement in density.
Extended Data Figure 4 | Dis3l2 domain interface analysis. a, Analysis of the CSD1–RNB interface. The conformation of CSD1 is stabilized by two protein–protein interactions with the RNB (K240 with D739 and D91 with T613) and an RNA-mediated interaction with the RNB through U6–U7 and α11.
b, Analysis of the CSD2–RNB interface. CSD2 is intimately associated with CSD1, but also forms an interface with the RNB through α3 (S242 with E337 and K319 with E332). c, Analysis of the S1–RNB interface. S1 is part of a large hydrophobic interface with RNB (α11, α13 and α17, β18 and β19). The S1 domain also forms an interface with the RNB through interactions with U8 (also see Fig. 3d) d, CSD1 is further stabilized through an RNA-mediated interaction with S1. The backbone phosphate of U4 bridges K78 of CSD1 and R792 of S1.
Extended Data Figure 5 | OligoU length preference of Dis3l2.  

Extended Data Figure 5 | OligoU length preference of Dis3l2. a, Specific activity (nt min⁻¹ molecule⁻¹) of Dis3l2 with increasing length oligoU (U₉, U₁₁, U₁₃ and U₁₅). The calculation of specific activity was conducted as in Fig. 1a. Mean ± s.d. (n = 3) are shown. b, The initial rate plots for each substrate used to calculate specific activity in panel a. The amount of substrate degraded (fmol) was plotted against time and the initial rate (slope) was determined with linear regression using GraphPad Prism 6. Mean ± s.d. (n = 3) are shown. c, Equilibrium dissociation constants (Kₐ) for increasing oligoU length. The Kₐ ± s.d. determined from three independent replicates is shown. d, The equilibrium dissociation constant (Kₐ) for each substrate in panel c was determined by plotting the fraction RNA bound versus the concentration of free Dis3l2 and fit by nonlinear regression with GraphPad Prism 6. The mean of three independently measured replicates with error bars representing ± s.d. are shown with the Kₐ.
Extended Data Figure 6 | Dis3l2 has a conserved active site and an RNA path that resembles RNase II. a, The active site of the Dis3l2–U14 complex. Conserved active site residues are shown as green sticks and U13 and U14 are shown as orange sticks. A single Mg$^{2+}$ ion (purple sphere) is modelled in the active site. As proposed for RNase II, Dis3l2 may utilize a two Mg$^{2+}$ ion mechanism during catalysis. b, Superposition of Dis3l2 and yeast Rrp44 (PDB 2VNU) active sites. Dis3l2 side chains and the Mg$^{2+}$ ion are coloured the same as in panel a, and yeast Rrp44 residues are shown as grey sticks. c, Dis3l2–U14 complex in an identical layout as Fig. 2a. A wide-open funnel created by the CSD lobe and S1 allows RNA to access the top of the RNB. d, The structure of E. coli RNase II$^{22}$ (PDB 2IX1). The path of RNA in RNase II more closely resembles that in Dis3l2, compared to yeast Rrp44, although narrow along its length, underscoring its ability to accommodate only single-stranded RNA substrates. e, Perpendicular side and top views of the electrostatic surface potential of Dis3l2 (contoured at $\pm 5 \text{kT}e^{-1}$; white is neutral, blue is positive and red is negative). A much wider positively charged funnel on the top of the RNB supports the ability of Dis3l2 to degrade structured RNA substrates. f, Electrostatic surface potential of the single-strand-specific RNase II, in the same configuration as panel e. A narrow RNA binding channel can only accommodate single-stranded RNA. 

©2014 Macmillan Publishers Limited. All rights reserved
Extended Data Figure 7 | Dis3l2 discriminates U over A and C.  

**a** Dis3l2 selects oligo-U-tailed substrates by way of an extensive U-specific interaction network along most of its binding path. In U-zone 1, Dis3l2 makes both main-chain and side-chain-mediated hydrogen bonds with uracil. The hydrogen bond network with U14 is disrupted for modelled A14 and C14 RNAs. Dis3l2 residues are shown as sticks colour-coded by domain as in Fig. 1c. RNA is shown as orange sticks. Hydrogen bond pairs are shown as dashed lines and disrupted hydrogen bonds are shown as curved red lines.  

**b** U-zone 2, with the same layout as panel a.  

**c** U-zone 3.
Extended Data Figure 8 | Initial rate plots of Dis3l2 U-zone mutants.

(a) Initial rate plots for U-zone 1 used to calculate specific activity as shown in Fig. 3f. Assays were conducted under the same conditions as described for U15 in Extended Data Fig. 1, where reactions contained 0.5 nM enzyme and 100 nM 5′ radio-labelled U15. Mean ± s.d. (n = 3) are shown. (b) Initial rate plots for U-zone 2, conducted as described for panel a. (c) Initial rate plots for U-zone 3, conducted as described for panel a.
Extended Data Figure 9 | Binding affinity of Dis3l2 U-zone mutants. a, Slot blot filter-binding assay was used to measure the binding affinity of selected Dis3l2 mutants with U15. The equilibrium dissociation constant ($K_d$) for U-zone 1 mutants as shown in Fig. 3g was determined by plotting the fraction RNA bound versus the concentration of free protein and fit by nonlinear regression with GraphPad Prism 6. The mean of three independently measured replicates with error bars representing ± s.d. are shown with the $K_d$. b, The equilibrium dissociation constant ($K_d$) for U-zone 2 mutants as shown in Fig. 3g, and measured as described for panel a. c, The $K_d$ for U-zone 3 as shown in Fig. 3g, and measured as described for panel a.
Extended Data Table 1 | Data collection and refinement statistics

|                         |                  |
|-------------------------|-----------------|
| **Data collection**     |                 |
| Space group             | \( P2_1 \)      |
| Cell dimensions         |                 |
| \( a, b, c \) (Å)       | 63.9 96.1 157.3 |
| \( \alpha, \beta, \gamma \) (°) | 90 98.8 90   |
| Wavelength (Å)          | 1.1             |
| Resolution (Å)          | 63.18 - 2.95 (3.06 - 2.95) |
| \( R_{merge} \) (%)    | 0.128 (0.636)   |
| \( I/\sigma(I) \) (%)  | 8.4 (2.1)       |
| Completeness (%)        | 99.8 (99.9)     |
| Redundancy              | 3.5 (3.5)       |

| **Refinement**          |                 |
| Resolution (Å)          | 63.2 - 2.95 (3.02 - 2.95) |
| No. reflections         | 39755 (3968)    |
| \( R_{work} / R_{free} \) (%) | 0.202/0.251 |
| No. atoms (non-hydrogen)| 22942           |
| Protein                 | 22092           |
| RNA                     | 842             |
| \( \text{Mg}^{2+} \)    | 2               |
| Water                   | 6               |
| \(<B\text{-factors}\> \) (Å²) |          |
| Protein                 | 78.3            |
| RNA                     | 101.9           |
| \( \text{Mg}^{2+} \)    | 46.6            |
| Water                   | 34.2            |
| r.m.s. deviations       |                 |
| Bond lengths (Å)        | 0.012           |
| Bond angles (°)         | 0.95            |
| Ramachandran            |                 |
| Favored (%)             | 96.8            |
| Allowed (%)             | 3.13            |
| Outliers (%)            | 0.07            |

Values in parentheses are for the highest resolution shell.