Multi-domain utilization by TUT4 and TUT7 in control of let-7 biogenesis

Christopher R. Faehnle[1,4,6], **Jack Walleshauser**[1,2,4,5,6], and **Leemor Joshua-Tor**[1,3,4]

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

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

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

[4]Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 11724, USA

Abstract

The uridyl transferases TUT4 and TUT7 switch between two modes of activity that either promotes let-7 expression (monoU) or marks it for degradation (oligoU). Lin28 modulates the switch via recruitment of TUT4(7) to pre-let-7 in stem cells and human cancers. We found TUT4(7) utilize two multi-domain functional modules during the switch from mono- to oligoU. The catalytic module (CM) is essential for both activities, while the Lin28-interacting module (LIM) is indispensible for oligoU. The TUT7 CM structure trapped in the monoU state, revealed a duplex RNA binding pocket that orients group II pre-let-7 hairpins to position the 1-nt overhang favor monoU addition. Conversely, the switch to oligoU requires the ZK domain of Lin28 to drive the formation of a stable ternary complex between pre-let-7 and the inactive LIM. Finally, ZK2 of TUT4(7) aids oligoU addition by engaging the growing oligoU tail through uracil-specific interactions.

The let-7 miRNA is broadly expressed in somatic cells and regulates cellular proliferation and differentiation, as well as the repression of several oncogenes and key regulators of mitogenic pathways, including HMGA2, MYC and RAS[1]. Humans have 12 let-7 variants (let-7a-1, -2, -3; let-7b-e; let-7f-1, -2; let-7g; let-7i; miR-98), with many tumors associated with the coordinated downregulation of multiple let-7 family members[2]. Mature let-7 is produced via the canonical miRNA biogenesis pathway, however its production levels are tightly controlled post-transcriptionally through the Lin28/let-7 pathway[3]. Elevated levels of Lin28 in stem cells[4] and a subset of human cancers[5–7] triggers the destruction of let-7 precursors[8–12] (pre-let-7) via recruitment of redundant terminal uridyltransferases, TUT4

(ZCCHC11) and TUT7 (ZCCHC6), referred to collectively as TUT4(7). The oligoU tail added by TUT4(7)[13,14] is a signal for degradation of pre-let-7 by Dis3L2[15–17], which guarantees suppression of let-7 expression.

As cells differentiate, Lin28 levels decline, causing TUT4(7) to switch their catalytic mode from processive oligouridylation to distributive monouridylation of group II pre-let-7[18]. Group II pre-let-7s (and other miRNAs) acquire a 1-nt overhang from Drosha processing and must be monouridylated prior to serving as a substrate for Dicer. Monouridylation by TUT4(7) promotes let-7 biogenesis by supplying the precursor with a proper 2-nucleotide (nt) 3′-end overhang.

TUT4(7) are members of the non-canonical poly(A) polymerases of the DNA polymerase β superfamily. Unlike most uridyltransferases, TUT4(7) are modular, multi-domain enzymes composed of a N-terminal CCHH zinc finger (ZF), two nucleotidyltransferase domains (NTD1 and NTD2) connected by a flexible linker, and three CCHC zinc knuckle domains (ZK1-3) (Fig. 1a). Structurally, the TUT4(7) NTDs closely resemble TUTases from trypanosomes[19–23] and the *S. pombe* CID1 polyU polymerase[24–27]. However, the domain architecture of TUT4(7) is unique and more complex, with two tandem NTDs accompanied by ZF and ZK domains, whose function remains to be determined. Curiously, NTD1 is not an active nucleotidyl transferase, as it lacks critical catalytic aspartate residues.

Besides pre-let-7, a plethora of mammalian RNAs are targets of TUT4(7) uridylation; including mature miRNAs[28,29], other pre-miRNAs[30], mRNAs[30,31] and non-coding RNAs[32–34]. Small RNA uridylation is also widespread in several animal model organisms, with Argonaute bound siRNA in flies[35] and worms[36] as examples. In plants, the TUTase HESO1 competes with methylation in an active system of miRNA turnover[37,38].

The mechanisms driving the Lin28/let-7 pathway are of particular interest to human health. The principal players in the pathway are linked to numerous diseases. Lin28 strongly down regulates the tumor-suppressor activity of let-7. As such, Lin28 activation is associated with several human primary tumors and its ectopic expression promotes cellular transformation[39,40]. Furthermore, Lin28 expression is correlated with Wilms' tumor development[41] and the advanced stages and poor clinical outcome of ovarian carcinoma, colon adenocarcinoma, germ cell tumors, hepatocellular carcinoma (HCC) and chronic myeloid leukemia (CML)[39,40,42–44]. Because Lin28 recruits TUT4, it is not surprising that TUT4 can promote tumor growth and metastasis and is often overexpressed in human cancers[45], especially Lin28 expressing tumors[7]. Interestingly, TUT4 inhibition blocks tumorigenicity and invasiveness of breast cancer cells *in vivo* and *in vitro*[7], making the Lin28/let-7 pathway a superb therapeutic target. Finally, genetic deletion of Dis3L2, the effector of the Lin28/let-7 pathway is the primary cause of Perlman Syndrome, a congenital growth defect that often leads to Wilms' tumors[46].

In this study we set out to determine the mechanisms utilized by TUT4(7) in carrying out two alternative activities that occur in different cellular contexts with opposite outcomes. We used a structural biochemistry approach to show that TUT4(7) are multi-domain modular enzymes. We define TUT4(7) as composed of two functional modules, the catalytic module

(CM), which is essential for both modes of U addition, and the Lin28-interacting module (LIM), which is critical for interaction with Lin28 and the switch to processive oligoU addition.

RESULTS

TUT4(7) utilize multiple domains to regulate pre-let-7

To understand the TUT4(7) activity switch on pre-let-7g substrate, we examined the uridylation activities of mouse TUT4 and human TUT7 purified from insect cells. We selected pre-let-7g as a substrate for biochemical characterization, because it has been extensively studied[47] and there is a published structure of the Lin28-pre-let-7g pre-element[48]. It had previously been reported that immunopurified and/or bacterially expressed TUT4(7) catalyzed mono- and Lin28-dependent oligoU addition to pre-let-7[13,14,18,47,49]. We found that baculovirus infected insect cells produced higher yields and more stable TUT4(7) proteins for *in vitro* studies and therefore established assay conditions to monitor both activities. In accordance with previous biochemical data, we found that both TUT4(7) catalyze distributive monoU addition of group II pre-let-7g (1-nt 3′-end overhang) (Supplementary Fig. 1a,b) with a marked preference over group I pre-let-7g (existing 2-nt 3′-end overhang) (Supplementary Fig. 1a,c) substrate. Also, in the presence of Lin28, TUT4(7) switched to processive oligoU addition for both group I and group II pre-let-7 substrates (Supplementary Fig. 1a).

Our principal aim is to understand the TUT4(7) activity switch. We reasoned that a structure of TUT4(7) in complex with Lin28 and pre-let-7 would be invaluable toward this goal. Unfortunately, recombinant full-length TUT4(7) is difficult to purify and is recalcitrant to crystallization. We therefore examined several mouse TUT4 truncation mutants (mT1-mT7) expressed in insect cells and assayed for monoU and oligoU activities (Supplementary Fig. 1e–g). We expected to identify functional truncation mutants that would be more amenable to structural studies yet also provide insight to the Lin28-induced activity switch. We identified construct mT5 (composed of NTD2 and three ZK domains), as the minimal domain tested that catalyzed monoU addition (Supplementary Fig. 1f). We refer to mT5 as the catalytic module (CM), as it is also indispensible for Lin28-dependent oligoU activity (Supplementary Fig. 1g). Yet, oligoU activity also required the ZF and NTD1 domains in combination with the CM. We termed the ZF and NTD1, the Lin28-interacting module (LIM), which was inactive on its own, yet together with the CM facilitates processive oligoU addition (Supplementary Fig. 1g)[47,49]. Deletion up to and including the ZF within the LIM abrogated Lin28-dependent oligoU addition as observed previously[47]. Meanwhile, mT4 (deletion of ZK2) catalyzed processive Lin28-dependent oligoU addition, but consistently added shorter oligoU tails (oligoU$_{short}$, up to 10 uridines) to pre-let-7 than mT1 (>30 uridines), suggesting a role for ZK2 in oligoU extension (Supplementary Fig. 1g). Overall, our data is consistent with previous domain mapping studies reported for TUT4(7)[47]. However, that study reported no activity for ZK2 truncation mutants in contrast to our oligoU$_{short}$ phenotype. This discrepancy could be due to a difference in the constructs used or expression system.

Next, we incubated mT1 with pre-let-7g and analyzed the complex by gel filtration chromatography (GF). The mT1-pre-let-7g complex is transient, as they eluted separately by GF, consistent with distributive monoU addition[49] in the absence of Lin28 (Fig. 1c). We hypothesized Lin28 would stabilize the interaction between TUT4 and pre-let-7 RNA, to switch on processive oligoU activity. Indeed, Lin28 compels the stable association of TUT4 with pre-let-7g[50], since all three components eluted as a ternary complex (Fig. 1c). Given that the LIM is essential for oligoU activity, we next asked if the LIM is in communication with Lin28 during ternary complex formation. Indeed, construct mT2 (Fig. 1a), encompassing only the LIM, while not catalytically active, formed a ternary complex with Lin28-pre-let-7g (Fig. 1d).

Pre-let-7 family members are highly conserved in the double-stranded stem, but slightly more divergent in the terminal loop. However, a conserved GGAG element within the terminal loop of pre-let-7 is essential for the oligoU activity switch of TUT4(7)[13] (Supplementary Fig. 1b). Lin28 uses its two RNA binding domains; N-terminal cold shock domain (CSD) and two CCHC zinc knuckles (ZK_{Lin28}) (Fig. 1b) to bind the terminal loop of pre-let-7. Because ZK_{Lin28} bind GGAG[48], we asked whether ZK_{Lin28} is also driving TUT4 recruitment and the switch to oligoU addition. Lin28 constructs encompassing ZK_{Lin28} (mL2), but not the CSD (mL5), promoted oligoU addition (Supplementary Fig. 2). In agreement with the oligoU assay, ZK_{Lin28} also promoted ternary complex formation (Fig. 1e). Finally, we mutated the zinc coordinating cysteine residues in ZK_{Lin28} (mL3, C139A, C142A) and (mL4, C161A, C164A) to disrupt the GGAG interaction[13,48], and found the mL3 mutant to be especially defective in oligoU addition assays (Supplementary Fig. 2e). Unexpectedly, mL3 still bound pre-let-7g, but failed to retain mT1 in our GF assay (Fig. 1e), suggesting Lin28 binding, in itself, is not sufficient to recruit TUT4. Interestingly, mutants of the GGAG motif in pre-let-7 failed to undergo oligouridylation by TUT4(7)[13], further signifying the importance of the ZK_{Lin28}/GGAG interaction in TUT4(7) recruitment and the switch to oligoU. We imagine that a specific surface or conformation is adopted between the ZK_{Lin28} and GGAG motif of pre-let-7 that is recognized by the LIM of TUT4(7).

Structure of the CM in the "monoU" state

Since the CM is essential for both modes of TUT4(7) activities we set out to understand the molecular underpinnings of "monoU" and "oligoU addition". We reasoned that a structure of the CM in complex with a group II pre-let-7 double-helical stem would effectively represent "monoU addition". To this end, we identified two truncated forms of the human TUT7 CM (CM1-aa 963–1365 and CM2-aa 983–1365) suitable for crystallization. CM1 includes ZK1 (Figure 3a) while CM2 begins at NTD2 (Figure 2b). Crystals for both constructs diffracted X-rays to a comparable resolution limit. Overall, human TUT7 and mouse TUT4 share 48% sequence identity, while the CMs are very similar with 69% sequence identity. We conclude that our structural interpretations of the TUT7 CM will likely be conserved for TUT4. We determined the structure of CM1 by SAD phasing of a selenomethionine derivative to 3.0 Å resolution, followed by molecular replacement (MR) phasing of native data to 2.6 Å (CM-apo) (Table 1) (Supplementary Fig. 3a). Next, we co-crystallized CM2 (D1060A active site mutant) in complex with a 14 bp palindromic RNA duplex and UTP nucleotide (CM-dsRNA) (Fig. 2) and determined its structure by MR using

CM-apo as a search model (Table 1). Two CMs are observed at each end of the dsRNA, but we only considered one CM along with the dsRNA duplex as representing the "monoU" relevant structure (Supplementary Fig. 3b,c). The dsRNA is a mimic of the duplex stem of group II pre-let-7 (Supplementary Fig. 1d), as it contains a 1-nt 3′-end U overhang. We observed electron density accounting for most of NTD2, but curiously, no density was observed for ZK2 (residues 1337–1365). NTD2 is bilobal, the N-lobe (residues 987–1124) consists of a mixed 5-stranded ß-sheet wrapped tightly by two α-helices, while the C-lobe (residues 1125–1336) is made of 6 α-helices joined by several long loops (Fig. 2c). The overall fold of NTD2 resembles other determined single-domain TUTase (including *S. pombe* CID1[24–27] and the *Trypanosoma brucei* TbTUT1[22], TbTUT4[19], TbRET2[23] and TbMEAT1[20]) and non-canonical polyA polymerase (including GLD-2[51], Trf4p[52] and mtPAP[53,54]) structures, however these enzymes lack ZK domains (Supplementary Fig. 4). Interestingly, TbTUT1 contains a structured CCHH type zinc finger domain N-terminal to its NTD domain, reminiscent of the domain architecture of the TUT4(7) LIM[22]. A large substrate-binding cleft is formed between the N- and C-lobes of NTD2, whose conformation is similar to the closed conformation observed in *S. pombe* CID1[55]. Incoming UTP nucleotide is located at the bottom of the cleft between the lobes in the +1 position. Meanwhile dsRNA is positioned on the outer rim of the cleft with the 1-nt U overhang occupying the −1 position. The dsRNA duplex is clamped in position by the "5′-anchor" stemming from the N-lobe and the "groove loop" from the C-lobe (Fig. 2c).

Structure of the CM in the "oligoU" state

Next we set out to better understand the role of the CM during "Lin28-dependent oligoU addition". We imagined the CM in complex with nucleotide and oligoU substrate would suitably mimic the "oligoU state". We determined the structure of CM1 in complex with a 2-nt oligoU RNA substrate and a non-hydrolyzable UTP analog, UMPNPP (CM-U$_2$) (Table 1) (Fig. 3a,b and Supplementary Fig. 5a,b). We did not observe density for ZK1 (aa 963–984); however, unlike CM-dsRNA we observed clear density for the entire ZK2 domain (aa 1337–1363) (Fig. 3b). The fold of NTD2 in CM-U$_2$ was identical to that observed in CM-dsRNA. However, when we compared the CM-U$_2$ and CM-dsRNA structures to CM-apo we noticed a modest conformation change within the N-lobe (Fig. 3c). The N-lobe in CM-apo adopts a more "open" conformation, while the RNA bound CM-U$_2$ and CM-dsRNA structures adopt a "closed" state. This is best illustrated by the displacement of V1104 (from the N-lobe) by ~2.7 Å going from "open" to "closed" (Fig. 3d). Similar domain movements have been observed in *S. pombe* CID1[55]. The conformation of the "5′-anchor" and "groove loop" is indistinguishable between CM-dsRNA and CM-U$_2$.

Strikingly, ZK2 snakes between the N- and C-lobes where it engages the substrate U nucleotide in the −2 position (Fig. 3b). In contrast, the nucleotide in the −2 position (C14) in the CM-dsRNA structure is base-paired with G1, displacing ZK2 (Fig. 2c). Both the CM-dsRNA and CM-U$_2$ structures represent the pre-catalytic state prior to U addition with an incoming nucleotide occupying the +1 position (Fig. 2c and Fig. 3b).

ZK2 stabilizes oligoU in the pre- and post-catalytic state

Next, we determined the structure of CM2 in complex with a longer 5-nt oligoU RNA, (CM-U_5), for which U2-U5 were accounted for in the density (Supplementary Fig. 5c,d). The U_5 substrate adopted the post-catalytic state, imitating the product following uridyl transfer (Fig. 4b). Here, U5 sits in the +1 nucleotide-binding site with the remaining nucleotides occupying the −1, −2, and −3 positions. In both the pre- and post-catalytic states, ZK2 remains engaged with the U in the -2 position (Fig. 4a,b) (Supplementary Fig. 5e). We hypothesize ZK2 might aid translocation of the oligoU tail (Fig. 4c). Earlier biochemical studies of TUT4 have shown that ZK2 is essential for catalysis[47]. However, our activity data showed that ZK2 mutants still maintain monoU and processive oligoU addition activities (Supplementary Figs. 1 and 6), albeit with oligoU$_{short}$ as product in the presence of Lin28. These data support a role for ZK2 in oligoU extension, perhaps by stabilizing the oligoU product or by aiding translocation. We also find it plausible that ZK1 and ZK3 could act redundantly with ZK2 to aid oligoU addition to pre-let-7[13,14] and perhaps other RNA substrates[31].

ZK2 is displaced by pre-let-7 stem during monoU addition

We compared the structures representing both "monoU" and "oligoU addition". The oligoU in CM-U_5 tracks along the same path as the 3′-strand in CM-dsRNA in the −1 and −2 positions, but U2 in the −3 position takes an alternate path (Supplementary Fig. 7a). The deviation in the path taken by oligoU compared to the 3′-strand of dsRNA is due to the engagement of ZK2 (Supplementary Fig. 7a,b). When ZK2 is engaged (oligoU state) it occupies the same position as the first ~3-nt of the 5′-strand of dsRNA (monoU state), which forces oligoU to diverge at the −3 position (Supplementary Fig. 7a,b). Conversely, in CM-dsRNA, since ZK2 is displaced by the 5′-strand of the RNA duplex (Fig. 4d,e), the nucleotide in the −2 position (C14) is stabilized by a base pair with G1 (Fig. 4e) and the −3 position (G13) is base paired with C2 (Supplementary Fig. 7b). In the absence of Lin28 (monoU addition mode), the pre-let-7 hairpin stem excludes ZK2 from engaging substrate, since the 5′-stand occupies the same position (Supplementary Fig. 7b). When Lin28 is present, it first promotes the association of TUT4(7) with pre-let-7 through the LIM. Based on our structures representing the "monoU" and "oligoU" states, we hypothesize ZK2 is initially displaced by the pre-let-7 hairpin stem (Fig. 4e and Supplementary Fig. 7b), since an unpaired uridine at the −2 position would not be available yet. But following the initial rounds of U-addition, ZK2 eventually engages the growing U tail to promote oligoU addition (Fig. 4d). It remains unclear how many U additions must occur for ZK2 to engage, as well as what conformational changes must occur in the CM to accommodate this model. Interestingly, even in the absence of ZK2, TUT4 (mT4) produced Lin28-dependent oligoU$_{short}$ products (Supplementary Fig. 1g). Therefore, the primary role for ZK2 appears to be the extension of oligoU$_{short}$ (~8–10 Us) to oligoU (>30). The length of oligoU tail is almost certainly important to trigger the effector nuclease Dis3L2, which is activated by longer oligoU tails[15,16].

TUT4(7) active site and conserved UTP selectivity

Next we evaluated the substrate interactions occurring in the "oligoU" state of the CM (Fig. 5). In CM-U_2, the incoming UMPNPP base is wedged between Tyr1171, the uracil base of U2, and Val1104. These interactions were found to help determine substrate specificity in TbTUT4[19], and likely contribute to specificity for UTP as the incoming nucleotide in the CM as well (Fig. 6). UTP specificity is also facilitated through U-specific interactions with C-lobe residues Asn1130, His1286, and a water-mediated interaction with Asp1280. The role of His1286 in UTP selection is conserved in *S. pombe* CID1[24–26], but this residue is not conserved in TbTUTases (Fig. 6). Otherwise, the UTP binding site (+1 position) and the catalytic site are remarkably conserved between the TUT7 CM and other TUTases (Fig. 6), despite overall low sequence identity (~20%). Although not involved in nucleotide selection, C-lobe (Lys1152, Ser1170) and N-lobe (Ser1047, Ser1057) residues stabilize the γ phosphate of UMPNPP. The catalytic Mg^{2+} binding residues Asp1058 and Asp1060 coordinate a single Mg^{2+} which further stabilizes the β and γ phosphates of UMPNPP with the help of a water. Asp1119 completes the active site, where the α phosphate of incoming UTP is positioned for uridyltransfer (Fig. 5 and Fig. 6).

U-specificity is not unique to the +1 position, but is also evident in the substrate RNA binding pocket. The N-lobe residue Asn1124 makes a U-specific interaction with U2. This asparagine is conserved and essential for CID1 activity[27] (Fig. 6a). Furthermore, residues His1355 and Lys1352 of ZK2, which engages only in the oligoU circumstance, make U-specific interactions with U1 in the −2 position (Fig. 5). These interactions are not observed in any other TUTase structure, which all lack ZK domains (Fig. 6). Ultimately, the TUT4(7) CM is elegantly designed to catalyze oligoU addition. This is exemplified by the U-specific nature of interactions that occur with UTP, the 3′-terminal nucleotide, and the penultimate nucleotide in the −2 position by ZK2.

5′-anchor orients pre-let-7 duplex

UTP binding and selection is accomplished in an identical manner in the "monoU" structure, CM-dsRNA (Supplementary Fig. 7c,d). However, since ZK2 is displaced and unavailable to stabilize the nucleotide in the −2 position, this role is taken over by the 5′-strand of the duplex (Fig. 4e and Supplementary Fig. 7d). Here, we find the "5′-anchor" to be crucial for positioning of the dsRNA hairpin stem resulting in precise placement of U15 in the active site (Fig. 7a). The "5′-anchor" residues Leu1097 and Ile1099 create a hydrophobic platform for the first base pair of the duplex, C14:G1, while Val1104 positions the U15 base. The hydrophobic platform is conserved in mammalian TUT4(7), but quite divergent in CID1 and TbTUTs. Additionally, Thr1101 forms a H-bond with C14 and Lys1103 connects the bridging phosphate of C12 and G13 (Fig. 7a). Opposite the "5′-anchor", the "groove loop" from the C-lobe stabilizes the minor groove of the dsRNA duplex primarily through van der Waals' forces with Ala1163, Ser1164, and Arg1165 and a H-bond with Ser1164 (Supplementary Fig. 7e). The interactions with the duplex stem are non-sequence specific and are consistent with TUT4(7)'s ability to monouridylate multiple pre-let-7 family members as well as other group II pre-miRNAs[18]. Interestingly, the 5′-anchor is also involved in oligoU substrate binding, as Val1104 interacts with the −1 nucleotide base regardless of whether the substrate is dsRNA or ssRNA (Fig. 5). In addition, the

hydrophobic anchor residue Ile1099 stabilizes the base of the −2 nucleotide for both dsRNA and ssRNA substrates, while L1097 is only engaged with the 5′-strand of dsRNA duplex (Fig. 7a).

Group II pre-let-7 must be repaired by TUT4(7) monoU addition activity prior to their biogenesis. We asked how TUT4(7) might discriminate between group I (2-nt overhang) (Supplementary Fig. 1c) and group II (1-nt overhang) (Supplementary Fig. 1b) pre-let-7 substrates. First, TUT4(7) must ensure precise positioning of group II pre-let-7 in the pre-catalytic state, with UTP in the +1 position so that monoU addition can proceed (Fig. 7b). Based on the CM-dsRNA structure, we hypothesized the CM of TUT4(7) measures the end structure of group II substrates through the hydrophobic platform of the "5′-anchor". To validate the CM-dsRNA structure we mutated the hydrophobic platform of full-length hTUT7 to make it more bulky (L1097W, I1099W) and measured monoU addition. Indeed, a bulky "5′-anchor" is inactive towards monoU addition (Fig. 7c,d), likely a result of misplacement of the 1-nt overhang.

We propose that following monoU addition of group II pre-let-7, the newly formed 2-nt 3′-end overhang product will occupy the post-catalytic state and dissociate (Fig. 7b). If we assume that group I pre-let-7 (with a pre-existing 2-nt overhang) binds to the hydrophobic platform of the CM as observed in the CM-dsRNA structure, we imagine it would associate with the CM in the post-catalytic state, which is unfavorable for monoU addition (Fig. 7b).

DISCUSSION

Our study is the first to describe a structural model of the two catalytic modes of mammalian TUT4(7) (Fig. 8). We find that group II pre-miRNAs bind TUT4(7) in the pre-catalytic state, which is favored for monoU addition (Fig. 8a). We propose the transient interaction between TUT4(7) and pre-let-7 favors the addition and release of the RNA substrate before oligouridylation can occur. In contrast, group I pre-miRNAs bind in the inactive post-catalytic state, which are released prior to uridylation (Fig. 7b). All the while, the double-stranded helical stem of pre-let-7 prevents ZK2 engagement.

Lin28 controls the "oligoU switch" by recruiting the LIM of TUT4(7) to the GGAG motif within the terminal loop of pre-let-7 (Fig. 8b). The stable ternary complex supports processive oligoU addition by the CM, aided by the U-specific engagement of ZK2 (Fig. 8b). Although our CM structures lack the processivity factor Lin28, RNA, and LIM, we identify a clear and unexpected role for ZK2 in oligoU extension. This is not entirely unprecedented, as the processivity of the TRAMP complex is mediated by an analogous interaction between Trf4 (polyA polymerase) and the ZK-containing Air2 protein[52]. Our model of the Lin28-induced switch from monoU to oligoU is also supported by single molecule studies, where Lin28 dramatically increases the dwell time of TUT4 on pre-let-7 substrate[50]. It remains unclear exactly how TUT4(7) accommodate group I (2-nt overhang) pre-let-7 during Lin28-dependent oligoU addition. Since Lin28 induces a stable long-lived ternary complex, it most likely allows the CM adequate time to engage both group I and group II pre-let-7 substrates to catalyze oligoU addition (Supplementary Fig. 1).

While our manuscript was under review, a parallel study was published that described the minimal components necessary for Lin28-dependent oligoU addition[56]. The authors also identified the ZK domain of Lin28 as the sole element that recruits TUT4(7). In addition, they describe the N-terminal half of TUT4 (similar to the LIM shown here) as the interaction partner for the Lin28/pre-let-7 binary complex. Importantly, together with our study, the entities mediating the Lin28 triggered activity switch in TUT4(7) are better understood and are in agreement with our proposed model.

To date, several questions remain regarding the function of TUT4(7). First, it is unclear how the LIM of TUT4(7) recognizes the Lin28-GGAG binary complex to initiate oligoU addition. Lin28 and TUT4(7) do not interact without pre-let-7. It is tempting to speculate that the ZK_{Lin28}/GGAG complex creates a unique binding surface for the LIM, although it is unclear where and what the nature of that interaction is. Notably, Lin28 ZK mutants will bind pre-let-7 but not recruit TUT4(7). Finally, our model does not address how the LIM and CM are oriented with respect to each other during the switch from monoU to oligoU (Fig. 8b). Future structural studies aimed at the intact TUT4(7) and/or LIM in complex with Lin28/pre-let-7 would be crucial to shedding light on these details.

Taken together, this study takes the first steps in distinguishing the mechanism of the two catalytic modes of TUT4(7). We imagine this approach will serve as a blueprint for the design of therapeutic compounds to target TUT4(7) activity, in particular Lin28-dependent oligoU activity in human cancers.

ONLINE METHODS

Protein expression and purification

Plasmids containing the ORF of mouse TUT4 (mTUT4) and human TUT7 (hTUT7) were kind gifts from Richard I. Gregory and V. Narry Kim, respectively. TUT4(7) constructs were expressed in Sf9 or High Five insect cells as N-terminal Strep-sumo-TEV fusion proteins from the pFL vector of the MultiBac baculovirus expression system. Insect cells were infected with baculovirus at 27°C for 60 hours of protein expression, after which the cells were centrifuged at 1200 rpm and 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. Cells were thawed, then the NaCl concentration was increased to 500 mM, and the cells were lysed by sonication. The lysate was treated with 0.2% poly-ethylene imine (PEI) to precipitate bulk nucleic acid 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 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 (1:100 TEV:protein ratio) and incubated overnight at 4°C. The cleavage efficiency and purity was verified by SDS-PAGE. TEV protease treated protein was diluted with an equal volume of MonoS buffer A (25 mM Hepes pH 7.5, 5 mM DTT) to a final NaCl concentration of 50 mM. TUT4(7) proteins were loaded onto a Poros HS or HiTrap SP HP cation exchange column 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 TUT4(7). Fractions that contained TUT4(7) were analyzed

by SDS-PAGE, pooled and concentrated and loaded onto either a HiLoad 16/60 Superdex 200 or Superdex 200 10/300 gel filtration column equilibrated in 20 mM Hepes pH 7.5, 150 mM NaCl, and 5 mM DTT. TUT4(7) was concentrated to 15 mg/mL, flash frozen in liquid nitrogen and stored at −80°C. Selenomethionine (SeMet) substituted human TUT7 constructs were expressed in High-five cells infected with baculovirus in ESF921 media (Expression Systems) for 6 hr, followed by media exchange to a methionine-free media. After growth for an additional 4 hr, SeMet was added to a final concentration of 198 mg/L. The cells were harvested after 48 hr. SeMet-substituted TUT7 was purified in the same manner as the native protein. All mutant proteins (mouse TUT4 and human TUT7) used in this study were expressed and purified in an identical manner as the wild-type proteins, and exhibited similar size exclusion chromatography profiles, indicative of proper folding.

Mouse Lin28 (mLin28) was cloned into the pET28 vector and expressed as an N-terminal His$_6$-SUMO fusion protein. mLin28 was expressed in BL21-RIPL (DE3) cells grown in TB media. Cells were grown to an OD>1, followed by addition of 100 μM zinc chloride, then expression was induced with 0.1 mM IPTG for 3 hours at 37°C. Cell pellets were collected by centrifugation at 18,000 RPM and suspended in Wash buffer (50 mM Tris pH 8.0, 400 mM NaCl, 20 mM imidazole, and 2 mM β-ME). mLin28 was purified by Ni^{2+}-affinity chromatography, followed by elution with Wash buffer containing 250 mM imidazole. Eluted mLin28 fractions were treated with Ulp1 protease to remove the SUMO tag and diluted into ion exchange buffer A (25 mM Hepes pH 7.5 and 5 mM DTT) to a final NaCl concentration of 200 mM. mLin28 was contaminated with substantial amounts of nucleic acids, which were removed by passage through a HiTrap Q HP column. The flow through fraction (nucleic acid-free mLin28) was loaded onto a HiTrap SP HP column and eluted with a linear gradient from 0.2 M to 1M NaCl. mLin28 fractions were loaded onto a HiLoad Superdex 75 16/60 gel filtration column equilibrated in 20 mM Hepes pH 7.5, 150 mM NaCl, and 5 mM DTT, and then flash frozen in liquid nitrogen for storage at −80°C.

Structure determination of human TUT7 CM complexes

We identified two hTUT7 catalytic module constructs (CM1 and CM2) suitable for crystallization by limited proteolysis with thermolysin. CM1 is composed of zinc knuckle 1 (ZK1), nucleotidyltransferase domain 2 (NTD2), and zinc knuckle 2 (ZK2) (aa 963–1365), while CM2 lacks ZK1 (aa 983–1365). Both CM constructs were crystallized at 18°C by sitting drop vapor diffusion by mixing protein (12 mg/ml concentration) 1:1 with well solution (0.8 to 1 M lithium sulfate, 75 to 200 mM sodium citrate pH 5–6, and 50 to 200 mM potassium iodide) using a mosquito® Crystal liquid handling robot (TTP Labtech). Crystals were cryoprotected by increasing the lithium sulfate concentration to 80 to 90% saturation, and flash cooled in liquid nitrogen. A SeMet derivative of CM1 was crystallized and flash cooled in an identical manner. Crystals were also obtained in similar conditions for both wild type and active site mutant CMs (D1060A) by co-crystallization with 2 mM UTP, 2 mM UMPNPP and the dinucleotide U$_2$, and 1 mM U$_5$ RNA. Initial phases were determined from a SAD data set collected on a SeMet substituted crystal at beamline X25 at the NSLS at Brookhaven National Laboratory to 3.0 Å resolution (λ=0.979 Å). Se site determination, phasing, and automatic model building were performed with the SAD phasing module as implemented in Phenix[57]. The SAD phased map was of excellent quality,

which allowed the AutoBuild[58] utility in Phenix to build a near complete atomic model. The partial model was subsequently used for molecular replacement (MR) phasing of a native data set collected at the Zn^{2+} absorption edge (λ=1.28 Å) on apo CM crystals to 2.6 Å resolution (CM-apo). The CM-apo structure contained three copies of the CM in the asymmetric unit. We did not observe density for ZK1, due to disorder in the region, but a nearly complete model was built for the NTD2 domain (aa 986–1335) in all three copies. In addition, we modeled ZK2 into ordered density for one copy (chain A, aa 1336–1361), but observed disorder in the remaining copies, so we excluded ZK2 from chains B and C in the final refined model. To confirm the presence and location of the Zn^{2+} coordinated by the CCHC zinc-binding motif within ZK2 we calculated an anomalous difference map, which clearly identified the presence and location of Zn^{2+} (ED Fig. 3). Next, we examined data sets collected from crystals co-crystallized with nucleotides and/or substrate RNA. Upon inspection of the electron density maps it was evident that sulfate and iodide ions occupied the nucleotide and substrate binding sites in place of the desired nucleotide or RNA. We assumed that high lithium sulfate and iodide concentrations are detrimental to substrate binding. We therefore devised an alternative strategy to lower the ionic strength of the crystal mother liquor prior to soaking of substrate.

Apo hTUT7 CM1 or CM2 crystals were transferred to a stabilizing cryo-solution consisting of 20% PEG 3350, 20–30% glucose, 150 mM sodium citrate pH 6, 100 mM lithium sulfate. Next, nucleotide and/or RNA substrate was introduced (at concentrations of 5 mM UTP or non-hydrolyzable analog UMPNPP, 1 mM U_2 or U_5 RNA) to the drop for 1.5 to 24 hours prior to cryo-cooling in liquid nitrogen. Complete X-ray diffraction data sets were collected for CM-U_2 (apo CM1 crystals soaked in 2 mM $MgCl_2$, 2 mM UMPNPP, 1 mM U_2) at the Advanced Photon Source (Beamline 19ID, λ=0.979 Å) and CM-U_5 (apo CM2, D1060A active site mutant crystals soaked in 1 mM U_5) at the Advanced Light Source (Beamline 8.2.2, λ=1.0 Å). Data sets were processed with XDS[59] as implemented in the AutoPROC[60] software package. The unit cell, space group, and crystal packing arrangement were identical to that observed for CM-apo. Phases were determined by molecular replacement (MR) using the CM-apo structure as a search model in Phaser[61]. Model refinement was performed with Phenix and manual correction to the models was performed with Coot[62]. Difference electron density maps clearly indicate the unambiguous presence of substrate for both CM-U_2 and CM-U_5. Final refinement statistics are presented in Table 1.

To determine the structure of the CM-dsRNA complex we incubated CM2 (D1060A active site mutant) with a 14 bp palindromic RNA duplex with a 1 nt U overhang at a 1:1 molar ratio and 5 mM UTP. The complex crystallized at 10 mg/ml using sitting drop vapor diffusion in a well solution of 200 mM ammonium acetate, 100 mM sodium citrate pH 5.5, and 24% PEG 400. Crystals were harvested and flash-cooled directly from the mother liquor. Diffraction data were collected to 2.50 Å resolution at the Advanced Photon Source (Beamline 19ID). Data were processed with XDS as implemented in the AutoPROC software package and phases determined by MR using the CM-apo structure as a search model in Phaser. Refinement cycles were performed with Refmac[63] and Phenix and manual correction to the model was performed with Coot. The final model consists of 2 copies of the CM (clear density for residues 987–1068, 1073–1195, 1202–1336) bound to each 1 nt overhang end of the palindromic RNA duplex. Final refinement statistics are presented in

Table 1. Throughout this manuscript we describe one CM and the duplex as representative of the monoU state structure. We selected the CM copy that generally displayed more ordered density for further analysis. The geometry of the refined structures reported here were validated with MolProbity[64]. All structure figures were created with the Pymol Molecular Graphics System.

Uridylation assay

Pre-let-7g hairpins were *in vitro* transcribed and purified by denaturing urea PAGE. We used pre-let-7g in all of our biochemical assays to maintain consistency between activity and binding assays. Also, pre-let-7g was previously used to characterize the biochemistry of TUT4(7)[47], and the structure of the Lin28/pre-let-7g pre-element has been determined[48]. We have tested other pre-let-7 family members and observed similar results reported here, as have others in the literature[13]. Monouridylation time course assays were performed by incubation of purified TUT4(7) constructs (5 nM concentration) with 5′ ^{32}P radiolabeled pre-let-7 RNA (100 nM) in U-assay buffer (25 mM Hepes pH 7.5, 150 mM NaCl, 10% glycerol, 2 mM DTT, 3 mM MgCl2) in a final volume of 50 uL. Mixtures were pre-heated to 30°C followed by addition of 1 mM UTP to trigger the reaction. 5 uL aliquots were removed at the indicated time points and were quenched by dilution into 20 uL of stop buffer (95% formamide, 10 mM EDTA, 0.1% bromophenol blue, and 0.1% xylene cyanol) and heated for 10 minutes at 65°C. 1 uL of each reaction was resolved by denaturing urea PAGE using 10% sequencing gels. The RNA products were visualized with phosphor imaging.

Oligouridylation time course assays were performed in an identical manner as described for monouridylation. Except, 5′ ^{32}P radiolabeled pre-let-7 RNA (100 nM) was pre-incubated with 150 nM of the indicated Lin28 construct in U-assay buffer (25 mM Hepes pH 7.5, 150 mM NaCl, 10% glycerol, 2 mM DTT, 3 mM MgCl2) in a final volume of 50 uL. Mixtures were pre-heated to 30°C followed by addition of 5 nM of the indicated TUT4(7) construct. Following an additional 1-minute incubation, 1 mM UTP was added to trigger the reaction. Reactions were prepared and resolved in the same manner as described for monouridylation assays.

Gel filtration chromatography assay

To monitor TUT4, Lin28, and pre-let-7 ternary complex formation we resolved the complexes by gel filtration chromatography (GF) on a Superdex 200 10/300 column equilibrated in 25 mM Hepes pH 7.5, 150 nM NaCl, 5 mM DTT. Combinations of TUT4, Lin28, and pre-let-7 were mixed at concentrations of 20 uM in 100 uL total volume and incubated on ice for 1 hour, prior to loading of the entire volume on the column. Control GF experiments, using individual Lin28 and TUT4 proteins, as well as pre-let-7 were loaded alone and as binary complexes and loaded as described. All GF experiments were monitored by measuring the absorbance at 280 nm and 260 nm.

Data Availability and Accession Code Availability Statements

Coordinates and structure factors for CM-apo (PDB ID 5W0B), CM-U$_2$ (PDB ID 5W0N), CM-U$_5$ (PDB ID 5W0M), and CM-dsRNA (PDB ID 5W0O) have been deposited in the Protein Data Bank.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

1. Büssing I, Slack FJ, Grosshans H. let-7 microRNAs in development, stem cells and cancer. Trends Mol Med. 2008; 14:400–409. [PubMed: 18674967]

2. Shell S, et al. Let-7 expression defines two differentiation stages of cancer. Proc Natl Acad Sci USA. 2007; 104:11400–11405. [PubMed: 17600087]

3. Thomson JM, et al. Extensive post-transcriptional regulation of microRNAs and its implications for cancer. Genes Dev. 2006; 20:2202–2207. [PubMed: 16882971]

4. Yu J, et al. Induced pluripotent stem cell lines derived from human somatic cells. Science. 2007; 318:1917–1920. [PubMed: 18029452]

5. Johnson SM, et al. RAS is regulated by the let-7 microRNA family. Cell. 2005; 120:635–647. [PubMed: 15766527]

6. Viswanathan SR, et al. Lin28 promotes transformation and is associated with advanced human malignancies. Nat Genet. 2009; 41:843–848. [PubMed: 19483683]

7. Piskounova E, et al. Lin28A and Lin28B inhibit let-7 microRNA biogenesis by distinct mechanisms. Cell. 2011; 147:1066–1079. [PubMed: 22118463]

8. Heo I, et al. Lin28 mediates the terminal uridylation of let-7 precursor MicroRNA. Mol Cell. 2008; 32:276–284. [PubMed: 18951094]

9. Newman MA, et al. Lin-28 interaction with the Let-7 precursor loop mediates regulated microRNA processing. RNA. 2008; 14:1539–1549. [PubMed: 18566191]

10. Rybak A, et al. A feedback loop comprising lin-28 and let-7 controls pre-let-7 maturation during neural stem-cell commitment. Nat Cell Biol. 2008; 10:987–993. [PubMed: 18604195]

11. Viswanathan SR, et al. Selective blockade of microRNA processing by Lin28. Science. 2008; 320:97–100. [PubMed: 18292307]

12. Piskounova E, et al. Determinants of microRNA processing inhibition by the developmentally regulated RNA-binding protein Lin28. J Biol Chem. 2008; 283:21310–21314. [PubMed: 18550544]

13. Heo I, et al. TUT4 in concert with Lin28 suppresses microRNA biogenesis through pre-microRNA uridylation. Cell. 2009; 138:696–708. [PubMed: 19703396]

14. Hagan JP, Piskounova E, Gregory RI. Lin28 recruits the TUTase Zcchc11 to inhibit let-7 maturation in mouse embryonic stem cells. Nat Struct Mol Biol. 2009; 16:1021–1025. [PubMed: 19713958]

15. 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. [PubMed: 23594738]

16. Faehnle CR, Walleshauser J, Joshua-Tor L. Mechanism of Dis3l2 substrate recognition in the Lin28-let-7 pathway. Nature. 2014; 514:252–256. [PubMed: 25119025]

17. Ustianenko D, et al. Mammalian DIS3L2 exoribonuclease targets the uridylated precursors of let-7 miRNAs. RNA. 2013; 19:1632–1638. [PubMed: 24141620]

18. Heo I, et al. Mono-uridylation of pre-microRNA as a key step in the biogenesis of group II let-7 microRNAs. Cell. 2012; 151:521–532. [PubMed: 23063654]

19. Stagno J, Aphasizheva I, Aphasizhev R, Luecke H. Dual role of the RNA substrate in selectivity and catalysis by terminal uridylyl transferases. Proc Natl Acad Sci USA. 2007; 104:14634–14639. [PubMed: 17785418]

20. Stagno J, et al. Structure of the mitochondrial editosome-like complex associated TUTase 1 reveals divergent mechanisms of UTP selection and domain organization. J Mol Biol. 2010; 399:464–475. [PubMed: 20403364]

21. Stagno J, et al. UTP-bound and Apo structures of a minimal RNA uridylyltransferase. J Mol Biol. 2007; 366:882–899. [PubMed: 17189640]

22. Rajappa-Titu L, et al. RNA Editing TUTase 1: structural foundation of substrate recognition, complex interactions and drug targeting. Nucleic Acids Res. 2016; 44:10862–10878. [PubMed: 27744351]

23. Deng J, Ernst NL, Turley S, Stuart KD, Hol WGJ. Structural basis for UTP specificity of RNA editing TUTases from Trypanosoma brucei. EMBO J. 2005; 24:4007–4017. [PubMed: 16281058]

24. Lunde BM, et al. Crystal structures of the Cid1 poly (U) polymerase reveal the mechanism for UTP selectivity. Nucleic Acids Res. 2012; 40:9815–9824. [PubMed: 22885303]

25. Munoz-Tello P, Gabus C, Thore S. Functional implications from the Cid1 poly(U) polymerase crystal structure. Structure. 2012; 20:977–986. [PubMed: 22608966]

26. Yates LA, et al. Structural basis for the activity of a cytoplasmic RNA terminal uridylyl transferase. Nat Struct Mol Biol. 2012; 19:782–787. [PubMed: 22751018]

27. Munoz-Tello P, Gabus C, Thore S. A critical switch in the enzymatic properties of the Cid1 protein deciphered from its product-bound crystal structure. Nucleic Acids Res. 2014; 42:3372–3380. [PubMed: 24322298]

28. Thornton JE, et al. Selective microRNA uridylation by Zcchc6 (TUT7) and Zcchc11 (TUT4). Nucleic Acids Res. 2015; 42:11777–11791.

29. Jones MR, et al. Zcchc11-dependent uridylation of microRNA directs cytokine expression. Nat Cell Biol. 2009; 11:1157–1163. [PubMed: 19701194]

30. Liu X, et al. A MicroRNA precursor surveillance system in quality control of MicroRNA synthesis. Mol Cell. 2014; 55:868–879. [PubMed: 25175028]

31. Lim J, et al. Uridylation by TUT4 and TUT7 Marks mRNA for Degradation. Cell. 2014; 159:1365–1376. [PubMed: 25480299]

32. Pirouz M, Du P, Munafò M, Gregory RI. Dis3l2-Mediated Decay Is a Quality Control Pathway for Noncoding RNAs. Cell Rep. 2016; 16:1861–1873. [PubMed: 27498873]

33. Łabno A, et al. Perlman syndrome nuclease DIS3L2 controls cytoplasmic non-coding RNAs and provides surveillance pathway for maturing snRNAs. Nucleic Acids Res. 2016; 44:10437–10453. [PubMed: 27431325]

34. Ustianenko D, et al. TUT-DIS3L2 is a mammalian surveillance pathway for aberrant structured non-coding RNAs. EMBO J. 2016; 35:2179–2191. [PubMed: 27647875]

35. Ameres SL, et al. Target RNA-directed trimming and tailing of small silencing RNAs. Science. 2010; 328:1534–1539. [PubMed: 20558712]

36. van Wolfswinkel JC, et al. CDE-1 affects chromosome segregation through uridylation of CSR-1-bound siRNAs. Cell. 2009; 139:135–148. [PubMed: 19804759]

37. Ren G, Chen X, Yu B. Uridylation of miRNAs by hen1 suppressor1 in Arabidopsis. Curr Biol. 2012; 22:695–700. [PubMed: 22464191]

38. Zhao Y, et al. The Arabidopsis nucleotidyl transferase HESO1 uridylates unmethylated small RNAs to trigger their degradation. Curr Biol. 2012; 22:689–694. [PubMed: 22464194]

39. Iliopoulos D, Hirsch HA, Struhl K. An epigenetic switch involving NF-kappaB, Lin28, Let-7 MicroRNA, and IL6 links inflammation to cell transformation. Cell. 2009; 139:693–706. [PubMed: 19878981]

40. Viswanathan SR, et al. Lin28 promotes transformation and is associated with advanced human malignancies. Nat Genet. 2009; 41:843–848. [PubMed: 19483683]

41. Urbach A, et al. Lin28 sustains early renal progenitors and induces Wilms tumor. Genes Dev. 2014; 28:971–982. [PubMed: 24732380]

42. Guo Y, et al. Identification and characterization of lin-28 homolog B (LIN28B) in human hepatocellular carcinoma. Gene. 2006; 384:51–61. [PubMed: 16971064]

43. Wang YC, et al. Lin-28B expression promotes transformation and invasion in human hepatocellular carcinoma. Carcinogenesis. 2010; 31:1516–1522. [PubMed: 20525879]

44. West JA, et al. A role for Lin28 in primordial germ-cell development and germ-cell malignancy. Nature. 2009; 460:909–913. [PubMed: 19578360]

45. Fu X, et al. miR-26a enhances miRNA biogenesis by targeting Lin28B and Zcchc11 to suppress tumor growth and metastasis. Oncogene. 2013; doi: 10.1038/onc.2013.385

46. Astuti D, et al. Germline mutations in DIS3L2 cause the Perlman syndrome of overgrowth and Wilms tumor susceptibility. Nat Genet. 2012; 44:277–284. [PubMed: 22306653]

47. 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. 2012; 18:1875–1885. [PubMed: 22898984]

48. Nam Y, et al. Molecular basis for interaction of let-7 microRNAs with Lin28. Cell. 2011; 147:1080–1091. [PubMed: 22078496]

49. Kim B, et al. TUT7 controls the fate of precursor microRNAs by using three different uridylation mechanisms. EMBO J. 2015; 34:1801–1815. [PubMed: 25979828]

50. Yeom KH, et al. Single-molecule approach to immunoprecipitated protein complexes: insights into miRNA uridylation. EMBO Rep. 2011; 12:690–696. [PubMed: 21637296]

51. Nakel K, Bonneau F, Eckmann CR, Conti E. Structural basis for the activation of the C. elegans noncanonical cytoplasmic poly(A)-polymerase GLD-2 by GLD-3. Proc Natl Acad Sci USA. 2015; 112:8614–8619. [PubMed: 26124149]

52. Hamill S, Wolin SL, Reinisch KM. Structure and function of the polymerase core of TRAMP, a RNA surveillance complex. Proc Natl Acad Sci USA. 2010; 107:15045–15050. [PubMed: 20696927]

53. Bai Y, Srivastava SK, Chang JH, Manley JL, Tong L. Structural basis for dimerization and activity of human PAPD1, a noncanonical poly(A) polymerase. Mol Cell. 2011; 41:311–320. [PubMed: 21292163]

54. Lapkouski M, Hällberg BM. Structure of mitochondrial poly(A) RNA polymerase reveals the structural basis for dimerization, ATP selectivity and the SPAX4 disease phenotype. Nucleic Acids Res. 2015; 43:9065–9075. [PubMed: 26319014]

55. Yates LA, et al. Structural plasticity of Cid1 provides a basis for its distributive RNA terminal uridylyl transferase activity. Nucleic Acids Res. 2015; 43:2968–2979. [PubMed: 25712096]

56. Wang L, et al. LIN28 Zinc Knuckle Domain Is Required and Sufficient to Induce let-7 Oligouridylation. Cell Rep. 2017; 18:2664–2675. [PubMed: 28297670]

57. Adams PD, et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr D Biol Crystallogr. 2010; 66:213–221. [PubMed: 20124702]

58. 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. [PubMed: 18094468]

59. Kabsch W. XDS. Acta Crystallogr D Biol Crystallogr. 2010; 66:125–132. [PubMed: 20124692]

60. Vonrhein C, et al. Data processing and analysis with the autoPROC toolbox. Acta Crystallogr D Biol Crystallogr. 2011; 67:293–302. [PubMed: 21460447]

61. McCoy AJ, et al. Phaser crystallographic software. J Appl Crystallogr. 2007; 40:658–674. [PubMed: 19461840]

62. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr. 2004; 60:2126–2132. [PubMed: 15572765]

63. Murshudov GN, Vagin AA, Dodson EJ. Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr D Biol Crystallogr. 1997; 53:240–255. [PubMed: 15299926]

64. Davis IW, et al. MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. Nucleic Acids Res. 2007; 35:W375–83. [PubMed: 17452350]
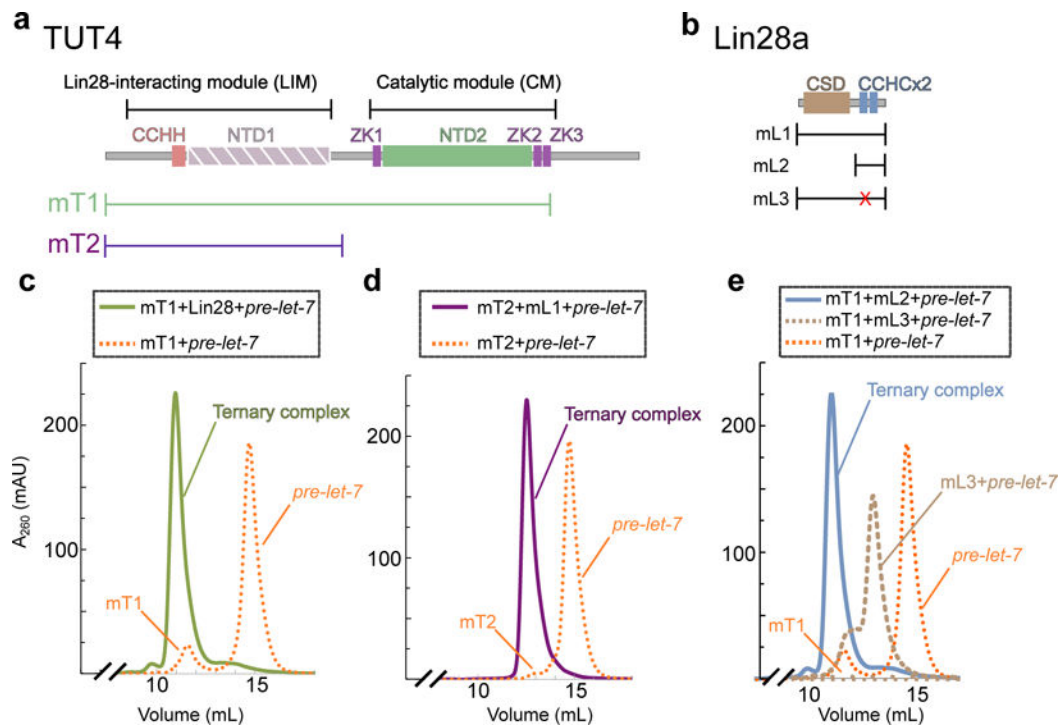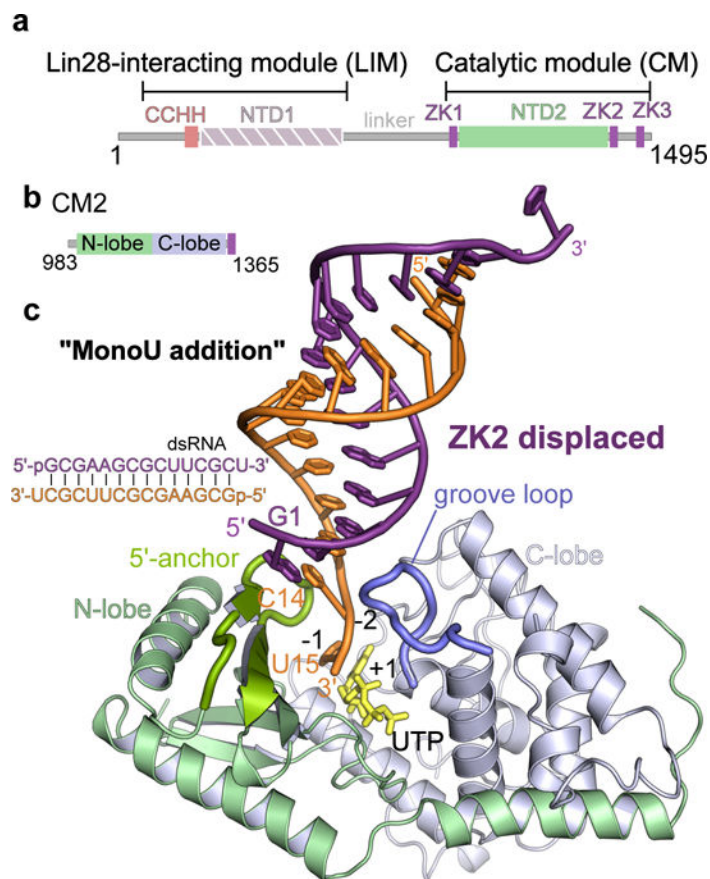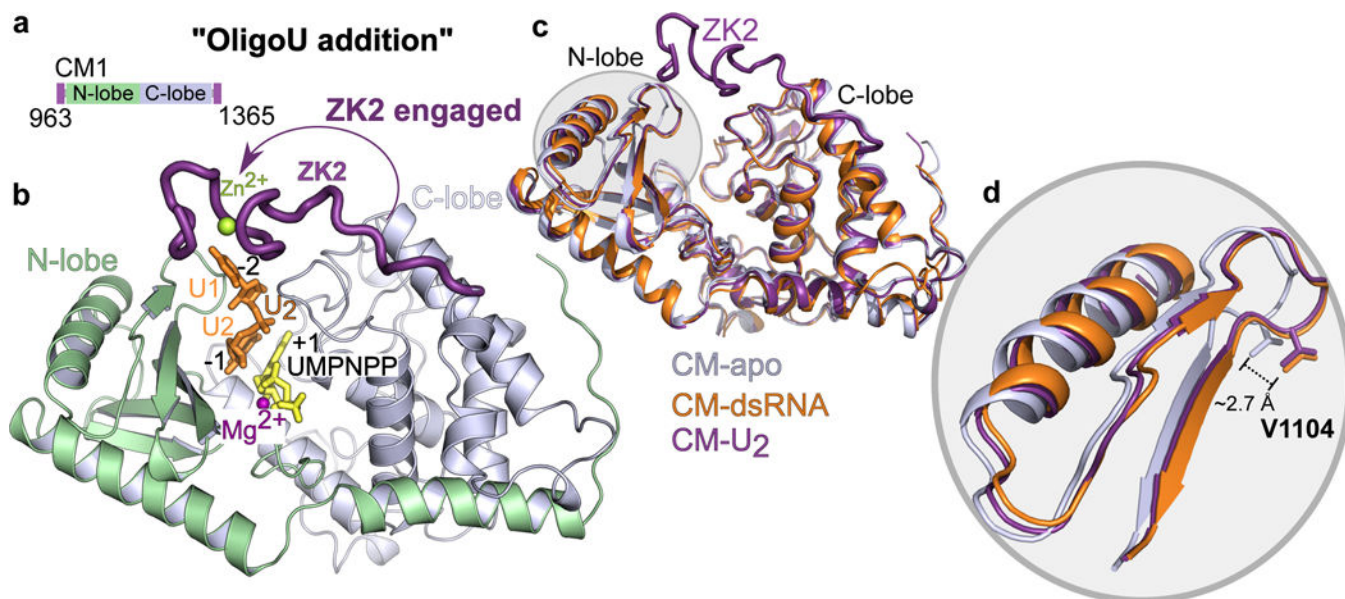
Figure 1.

Assembly of the Lin28/pre-let-7/TUT4 ternary complex. (**a**) Domain layout of mouse TUT4 (mTUT4). The Lin28-interating module (LIM) and catalytic module (CM) are indicated. The LIM is composed of the CCHH zinc finger (pink) and inactive NTD1 (light purple, white hash marks). The CM contains zinc knuckle domains ZK1-ZK3 (purple) and the active NTD2 (green). The mTUT4 truncation constructs analyzed in this study are labeled (mT1-mT2). We found that construct mT1 (truncation of C-terminal domain up to ZK3) has the same activity as full-length mTUT4, but is easier to purify and more stable. We therefore used mT1 as the backbone for other mutant constructs used in this study. (**b**) Domain layout of mouse Lin28 with cold shock domain (CSD, brown) and CCHC zinc knuckles (CCHCx2, blue) labeled. The Lin28 mutant constructs used for experiments presented in panels **c–e** are labeled. Point mutants of the Lin28 CCHC ZK1 (C139A, C142A) are indicated (red X). (**c**) Gel filtration (GF) binding assay of mTUT4 (mT1) and pre-let-7g with (green solid curve) and without Lin28 (orange dashed curve). (**d**) GF binding assay of the LIM of mTUT4 (mT2) and pre-let-7g with (purple solid curve) and without (orange dashed curve) Lin28. (**e**) GF binding assay of mTUT4 (mT1) and pre-let-7g with (blue solid curve) and without (orange dashed curve) mL2 (CCHCx2 only). Also shown is GF binding assay with mL3 (C139A, C142A mutant). Chromatograms in panels **c–e** are a plot of the absorbance (260 nm) vs. elution volume. Gel filtration assay curves are representative of three technical replicates.
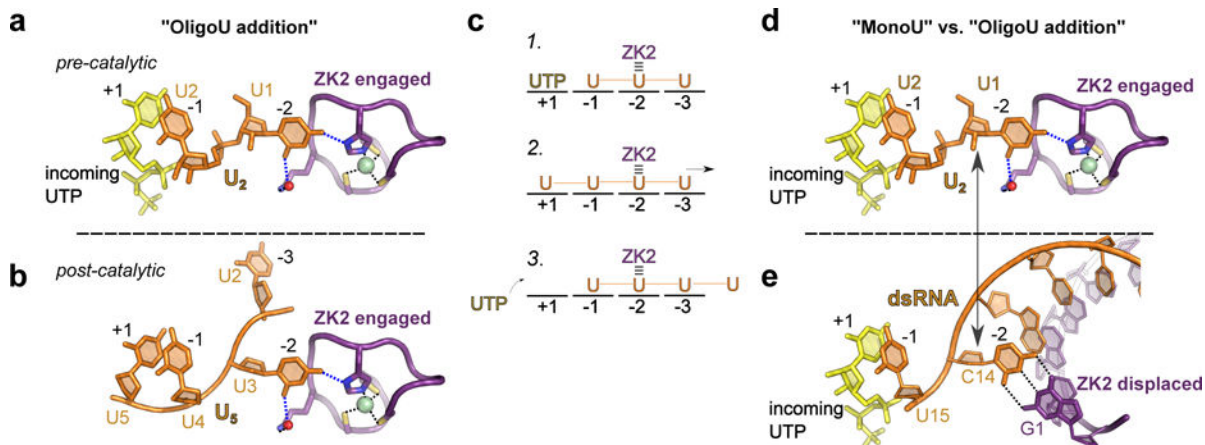
Figure 2.
Structure of hTUT7 CM in the "monoU" state. (**a**) Domain schematic of human TUT7 with Lin28-interacting module (LIM) and catalytic module (CM) indicated. The CCHH zinc finger (pink), inactive NTD1 (light purple, white hash marks), disordered linker, three CCHC zinc knuckles (ZK1-3, purple), and the catalytic NTD2 (green) are labeled. (**b**) Schematic of the truncated CM2 construct used for crystallization of CM-dsRNA, with the N-lobe (light green) and C-lobe (gray) labeled. (**c**) Structure of CM-dsRNA representing the "monoU addition" state. The N-lobe and C-lobe are labeled as in panel **b**. The RNA duplex (purple and orange strands) rests against the 5′-anchor (dark green) and groove loop (blue). The 1-nt U overhang (U15, orange strand) sits in the −1 position and is stacked with the incoming UTP nucleotide (yellow sticks) in the +1 position. During monoU addition, ZK2 is displaced by the 5′-strand (purple) of RNA duplex.
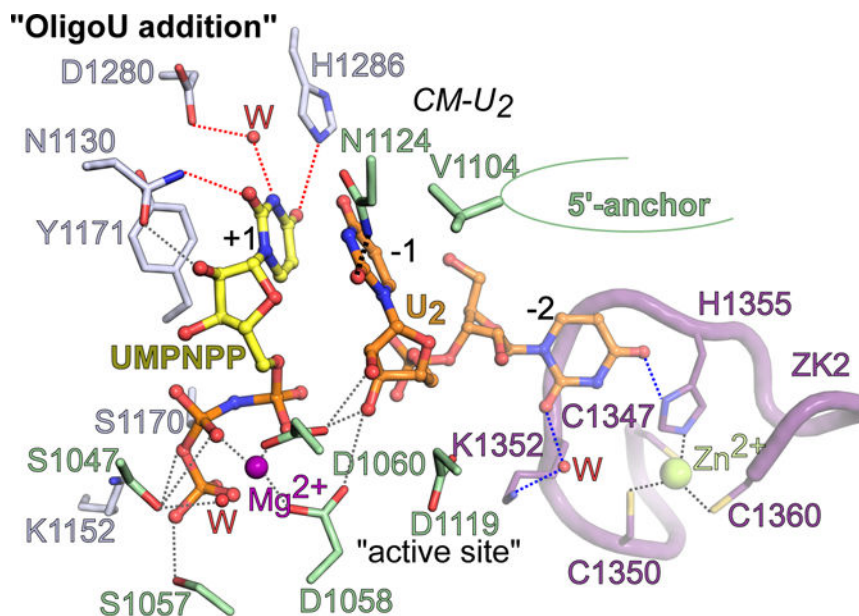
Figure 3.
Structure of hTUT7 CM in the "oligoU" state. (**a**) Schematic of the truncated CM1 construct used for crystallization of CM-U$_2$, with the N-lobe (light green) and C-lobe (gray) labeled. (**b**) Structure of CM-U$_2$ trapped in the "oligoU state". Here, ZK2 (purple ribbon) coordinates Zn^{2+} (sphere, lime green) and is engaged with U1 in the –2 position. Incoming nucleotide (UMPNPP, yellow stick) sits in the +1 position and is poised for the next uridylation cycle. (**c**) Superposition of a single chain of CM-apo (light blue cartoon) vs. CM-U$_2$ (purple cartoon) and CM-dsRNA (orange cartoon). The movement of the N-lobe in the presence of RNA substrates is highlighted (gray shade). (**d**) Close-up view of the N-lobe movement induced upon substrate binding. The degree of movement is illustrated by the displacement of Val1104 (sticks) by ~2.7 Å going from "open" to "closed".
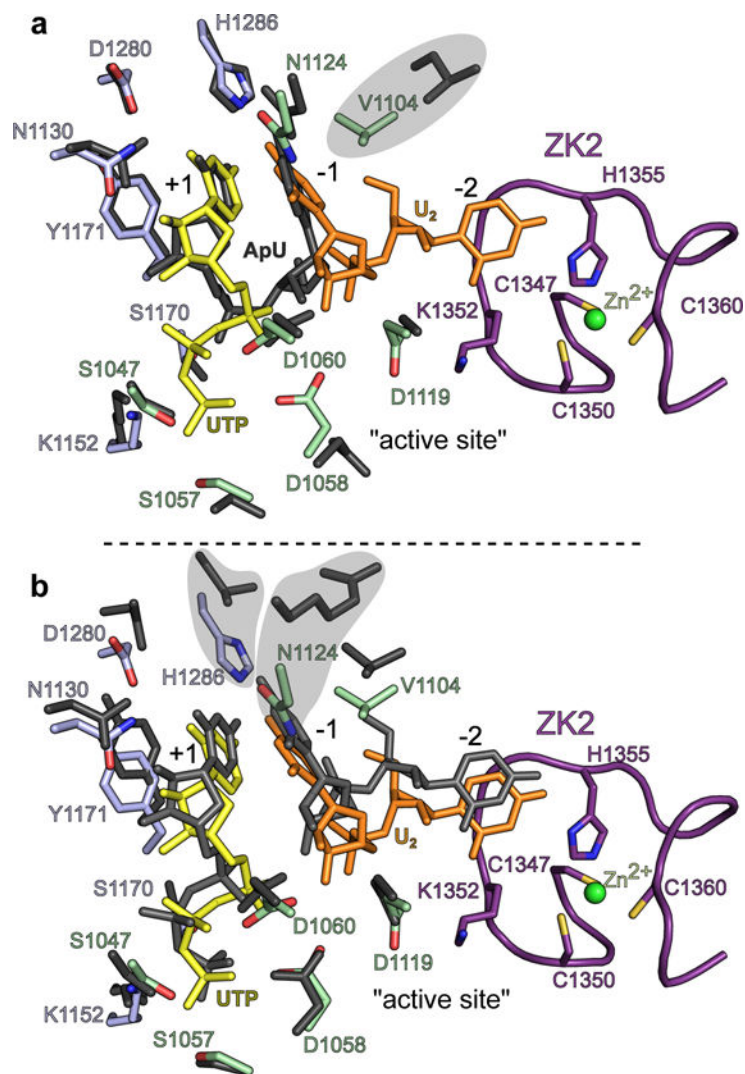
Figure 4.
Role of ZK2 in "Lin28-dependent oligoU" vs. "monoU addition". (**a**) The pre-catalytic state captured in the CM-U$_2$ structure. Shown are incoming nucleotide (UMPNPP, yellow stick), U$_2$ substrate (orange stick), and ZK2 (purple ribbon and stick). ZK2 is engaged with U1 in the −2 position in a U-specific manner (blue dashed lines). (**b**) The post-catalytic state captured in the CM-U$_5$ structure aligned with CM-U$_2$ (panel a). Shown is the U$_5$ substrate (orange sticks) and ZK2. ZK2 is engaged with the nucleotide in the −2 position in both the pre- and post-catalytic states. (**c**) Reaction scheme of the oligoU addition cycle from pre- (*1.*) to post-catalytic states (*2.*), highlighting the role of ZK2 in promoting oligoU addition (*3.*). (**d**) Identical to panel **a**. (**e**) Comparison of the pre-catalytic state during "monoU addition" (CM-dsRNA) to the pre-catalytic state in CM-U$_2$ (panel d). Instead of ZK2 engaged with the −2 position (CM-U$_2$, panel d), the −2 position (C14) is stabilized by base pair with G1 (double arrow). ZK2 is displaced by the 5′-strand of dsRNA.
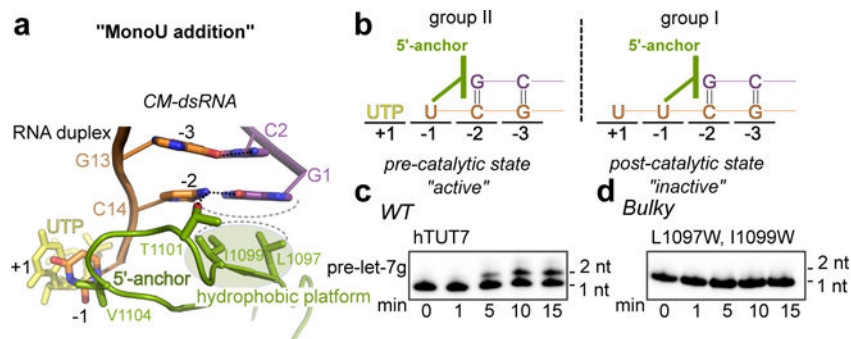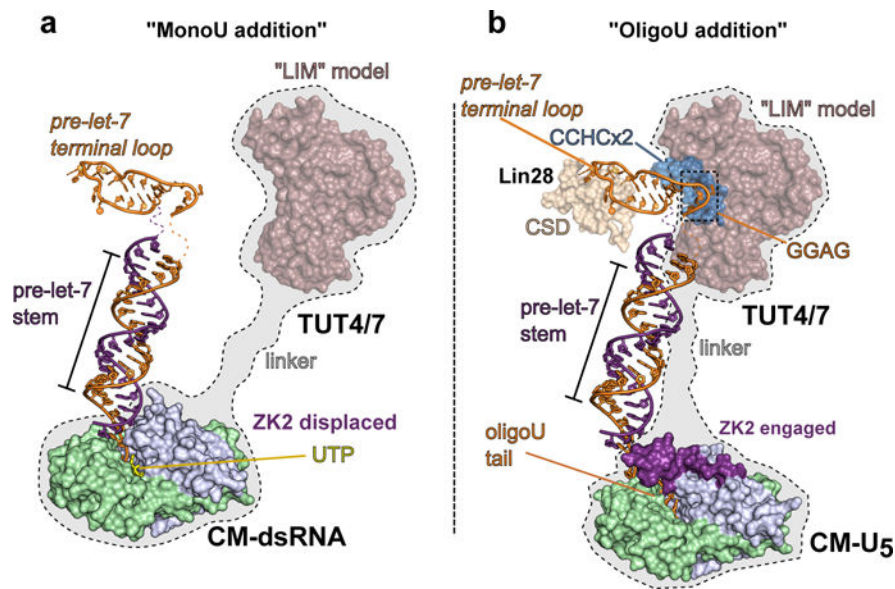
Figure 5.
Active site interactions in CM-U$_2$. Incoming nucleotide (UMPNPP, yellow stick) in the +1 position, U$_2$ substrate RNA (orange stick), and CM substrate interacting residues (N-lobe residues; green sticks, C-lobe; gray sticks, ZK2; purple sticks) are shown. General H-bond interactions (gray dashed lines) are displayed, as are U-specific interactions with the +1 position (red dashed lines), −1 position (black dashed line), and the −2 position (blue dashed lines).

Figure 6.
Active site comparison of TUT7 CM with CID1 and TbTUT4. (**a**) Superposition of CM-U_2 with *S. pombe* CID1 (PDB ID 4NKU) in complex with ApU product mimic. The active site residues for CM-U_2 are labeled. Conserved residues in CID1 are shown as dark gray sticks. Non-equivalent residues are highlighted with gray shade. (**b**) Superposition of CM-U_2 with TbTUT4 (PDB ID 5KAL) in complex with UTP and UpU. The active site residues for CM-U_2 are labeled. Conserved residues in TbTUT4 are shown as dark gray sticks. Non-equivalent residues are highlighted with gray shade.

Figure 7.
Group II pre-let-7 substrate recognition. (**a**) Group II miRNAs are measured by the 5′-anchor and groove loop. From the 5′-anchor: Val1104, Thr1101, and the hydrophobic platform (Leu1097, Ile1099) make direct contact with the dsRNA (green sticks). The C14:G1 base pair of dsRNA is stabilized by the hydrophobic platform. (**b**) Schematic of group II vs. group I pre-let-7 substrate recognition during monoU addition. (**c**) MonoU addition assay time course conducted with full-length hTUT7. (**d**) Bulky substitutions (L1097W, I1099W) within the hydrophobic platform severely disturb monoU addition. Gels shown are representative of three technical replicates. Uncropped gel images are shown in Supplementary Data Set 1.

Figure 8.
Model of the TUT4(7) activity switch. (**a**) A model of monoU addition. Group II miRNAs are monouridylated by the CM through recognition of the 1 nt overhang end structure. The CM-dsRNA is displayed as a molecular surface cartoon. The LIM is a homology model built from the CM-dsRNA structure (light pink, transparent surface). The pre-let-7 terminal loop is rendered from the structure of the Lin28/pre-let-7g complex (orange cartoon) (PDB ID 3TS2). (**b**) Model of Lin28-dependent oligoU addition. By way of its CCHCx2 domain (blue surface), Lin28 recruits the LIM of TUT4(7) to pre-let-7. The stable ternary complex promotes processive oligoU addition by the CM aided by direct engagement of ZK2.

Table 1

Data collection and refinement statistics.

	SeMet	CM-apo (5W0B)	CM-U$_2$ (5W0N)	CM-U$_5$ (5W0M)	CM-dsRNA (5W0O)
Data Collection					
Space Group	P6$_1$	P6$_1$	P6$_1$	P6$_1$	P2$_1$2$_1$2$_1$
Cell dimensions					
a, b, c (Å)	135, 135, 179	141, 141, 174	136, 136, 182	135, 135, 179	67, 81, 182
α,β,γ (°)	90, 90, 120	90, 90, 120	90, 90, 120	90, 90, 120	90, 90, 90
Resolution (Å)	178–3.0 (3.03–3.0)[a]	100–2.6 (2.62–2.6)	99–2.5 (2.51–2.5)	179–2.3 (2.31–2.3)	90–2.5 (2.51–2.5)
Rmerge (%)	9.4 (111)	4.6 (59.7)	6.3 (63.7)	4.9 (55.6)	4.9 (73.4)
I/σ(I)	22.8 (2.2)	32.8 (2.1)	15.5 (2.1)	22.7 (2.1)	19.3 (2.1)
Completeness (%)	100 (99.7)	100 (98.4)	99.6 (99.9)	99.9 (99.6)	99.5 (100)
Redundancy	10.4 (10.7)	9.3 (4.9)	4.4 (4.5)	5.6 (4.2)	4.7 (4.7)
Refinement					
Resolution (Å)		100–2.6	99–2.5	179–2.3	91–2.5
No. Reflections		59474	65534	82606	35149
R_{work}/R_{free}		0.232/0.255	0.187/0.215	0.188/0.229	0.205/0.256
No. atoms					
Macromolecules		8361	8936	9094	5977
Ligands		39	211	26	58
Water		22	218	205	50
B factors					
Macromolecules		86.3	63.2	59.9	78.9
Ligands		103.7	53.8	61.4	62.5
Water		59.9	51.0	52.3	61.8
R.m.s. deviations					
Bond lengths (Å)		0.003	0.003	0.005	0.006
Bond angles (°)		0.61	0.56	0.68	0.84

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