The host Integrator complex acts in transcription-independent maturation of herpesvirus microRNA 3′ ends

Mingyi Xie,[1,2] Wei Zhang,[3] Mei-Di Shu,[1,2] Acer Xu,[1,2] Diana A. Lenis,[4] Daniel DiMaio,[2,3] and Joan A. Steitz[1,2]

[1]Howard Hughes Medical Institute, [2]Department of Molecular Biophysics and Biochemistry, Boyer Center for Molecular Medicine, Yale University School of Medicine, New Haven, Connecticut 06536, USA; [3]Department of Genetics, Yale University School of Medicine, New Haven, Connecticut 06510, USA; [4]Department of Biochemistry and Cell Biology, Stony Brook University, Stony Brook, New York 11794, USA

Herpesvirus saimiri **(HVS) is an oncogenic γ-herpesvirus that produces microRNAs (miRNAs) by cotranscription of precursor miRNA (pre-miRNA) hairpins immediately downstream from viral small nuclear RNAs (snRNA). The host cell Integrator complex, which recognizes the snRNA 3′ end processing signal (3′ box), generates the 5′ ends of HVS pre-miRNA hairpins. Here, we identify a novel 3′ box-like sequence (miRNA 3′ box) downstream from HVS pre-miRNAs that is essential for miRNA biogenesis. In vivo knockdown and rescue experiments confirmed that the 3′ end processing of HVS pre-miRNAs also depends on Integrator activity. Interaction between Integrator and HVS primary miRNA (pri-miRNA) substrates that contain only the miRNA 3′ box was confirmed by coimmunoprecipitation and an in situ proximity ligation assay (PLA) that we developed to localize specific transient RNA–protein interactions inside cells. Surprisingly, in contrast to snRNA 3′ end processing, HVS pre-miRNA 3′ end processing by Integrator can be uncoupled from transcription, enabling new approaches to study Integrator enzymology.**

[*Keywords*: Integrator; *Herpesvirus saimiri*; microRNA biogenesis; proximity ligation assay; snRNA 3′ end processing]

Supplemental material is available for this article.

Received June 5, 2015; revised version accepted July 2, 2015.

Integrator is a metazoan-specific multisubunit, multifunctional protein complex composed of 14 subunits named Int1–Int14 (Integrator subunits) (Baillat et al. 2005; Chen et al. 2012). Despite the large number of subunits, characteristic structural domains are largely absent, and there are no obvious paralogs of Integrator subunits across animal genomes, except for Int9 and Int11 (Baillat and Wagner 2015). Int9 and Int11 are paralogous to CPSF100 (cleavage and polyadenylation specificity factor 100-kD subunit) and CPSF73, respectively, factors essential for 3′ end cleavage of RNA polymerase II (Pol II) transcribed messenger RNAs (mRNAs) (Baillat et al. 2005; Dominski et al. 2005; Mandel et al. 2006). Therefore, initially, the Integrator complex was implicated in the 3′ end processing of small nuclear RNAs (snRNA), another class of abundant Pol II transcribed RNAs (Baillat et al. 2005). Recent studies of Integrator have extended its functions into a broader spectrum of Pol II transcription events, including transcription initiation, promoter-proximal pausing, and termination of protein-coding transcripts (Gardini et al. 2014; Stadelmayer et al. 2014; Skaar et al. 2015). In addition, Integrator is involved in microRNA (miRNA) biogenesis in *Herpesvirus saimiri* (HVS), a γ-herpesvirus that causes fatal T-cell leukemias and lymphomas in new world primates (Fickenscher and Fleckenstein 2001; Cazalla et al. 2011).

miRNAs are short (~22-nucleotide [nt]) RNA molecules that control gene expression at the post-transcriptional level. Since miRNAs regulate the majority of protein-coding genes in metazoan cells and contribute importantly to many essential biological processes, significant efforts have been devoted to delineating their biogenesis pathways (Ha and Kim 2014). In the canonical miRNA biogenesis pathway, hairpin structures in Pol II transcribed primary miRNAs (pri-miRNAs) are cleaved sequentially by the nuclear Microprocessor complex (Drosha/DGCR8) followed by cytoplasmic Dicer to produce mature miRNA duplexes (Hutvagner et al. 2001; Lee et al. 2003; Denli et al. 2004; Gregory et al. 2004). Nuclear–cytoplasmic transport of the precursor miRNA (pre-

Corresponding author: joan.steitz@yale.edu
Article is online at http://www.genesdev.org/cgi/doi/10.1101/gad.266973.115.

miRNA) intermediates is carried out by Exportin-5 (Yi et al. 2003; Lund et al. 2004).

Many alternative routes of miRNA biogenesis exist in animals and their viruses (Xie and Steitz 2014). In animals, there are miRNA biogenesis pathways that bypass either Drosha or Dicer cleavage. For example, some cellular pre-miRNAs are produced via splicing (mirtrons) or Pol II transcription initiation/termination (m^7G-capped pre-miRNA and other endogenous shRNAs), thereby bypassing the Drosha cleavage step (Okamura et al. 2007; Ruby et al. 2007; Babiarz et al. 2008; Xie et al. 2013). Dicer cleavage is replaced by AGO-2 cleavage in vertebrate miR-451 biogenesis (Cheloufi et al. 2010; Cifuentes et al. 2010). In animal viruses, several miRNA biogenesis pathways skip the Drosha cleavage step by using Pol III to produce pre-miRNAs directly or tRNA-pre-miRNA chimeric transcripts that are further processed by RNaseZ (Bogerd et al. 2010; Kincaid et al. 2012). HVS-infected monkey T cells express seven viral snRNAs called HSUR1–HSUR7 (*Herpesvirus saimiri* U RNAs); three HVS pre-miRNAs are cotranscribed immediately downstream from HSUR2, HSUR4, or HSUR5, giving rise to mature miRNAs named miR-HSUR2, miR-HSUR4, and miR-HSUR5, respectively (Cazalla et al. 2011). The Integrator complex was previously shown to cleave the chimeric pri-snRNA/miRNA transcript, thus generating the 5′ ends of the three HVS pre-miRNAs as well as the 3′ ends of all HSURs (Cazalla et al. 2011).

The 3′ end processing of snRNAs has been studied extensively beginning two decades before the discovery of Integrator. Two sequence elements are essential for snRNA 3′ end processing: (1) a snRNA-type promoter and (2) a 3′ end processing signal called the 3′ box, with the consensus 5′-GTTTN$_{0–3}$AAARNNAGA-3′ sequence found 9–19 nt downstream from the U1, U2, or U3 snRNA 3′ ends (Hernandez 1985; Yuo et al. 1985; Hernandez and Weiner 1986). Although the required sequence elements seemed to be clear, it has been surprisingly difficult to directly assay snRNA 3′ end processing in vitro (Uguen and Murphy 2003). snRNA 3′ end processing appears to occur cotranscriptionally, since substituting the snRNA-type promoter with an mRNA-type promoter severely reduces the efficiency of 3′ end formation (Hernandez and Weiner 1986; Hernandez and Lucito 1988). Accordingly, in HVS miRNA biogenesis, the snRNA promoter and 3′ box proved to be essential for both HSUR 3′ end processing and pre-miRNA 5′ end formation (Cazalla et al. 2011). However, the precise mechanism of HVS pre-miRNA 5′ end formation was unclear: The Integrator complex presumably cleaves upstream of the HSUR 3′ box, generating pre-miRNAs preceded by a 5′ leader that is further trimmed by Integrator itself or another exonuclease (Fig. 1A; Cazalla et al. 2011). Likewise, the 3′ end formation mechanism for the three HVS pre-miRNAs remained elusive.

Here, we identify a HVS miRNA-specific 3′ end processing signal, which we call the miRNA 3′ box (miR 3′ box). Mutational analyses reveal that the essential elements of this sequence are the central adenosine residues. The miR 3′ box therefore resembles the adenosine-rich snRNA 3′ box, which is required for 3′ end processing of the HSURs. Using a construct expressing a m^7G-capped HVS pre-miRNA, which circumvents Integrator-dependent 5′ end processing, we performed in vivo knockdown and rescue experiments to demonstrate a requirement for Integrator activity in pre-miRNA 3′ end processing. We confirmed a direct interaction between Integrator and HVS pri-miRNA substrates that contain only the miR 3′ box by coimmunoprecipitation and a novel in situ PLA (proximity ligation assay) that detects transient RNA–protein interactions inside cells. Most strikingly, unlike promoter-dependent snRNA 3′ end processing by Integrator, HVS pre-miRNAs can be generated from an mRNA-type promoter and successfully 3′ end-processed. We conclude that HVS has uniquely evolved sequences that hijack the host Integrator processing machinery to generate both the 5′ and 3′ ends of its viral pre-miRNAs.

Results

Identification of the miR 3′ box essential for HVS miRNA biogenesis

Multiple sequence alignment of pre-miR-HSUR2, pre-miR-HSUR4, and pre-miR-HSUR5 and downstream sequences from different strains of HVS revealed a conserved A-rich tract (5′-TNAAAANT-3′) (Fig. 1A; Supplemental Fig. S1A), indicating an important signal for 3′ end formation. We named it the miRNA 3′ box (miR 3′ box) to mirror the nomenclature of the snRNA 3′ box. Although we focused here on miR-HSUR4 because it is the most highly expressed viral miRNA in HVS-infected marmoset T cells, the presence of the miR 3′ box downstream from miR-HSUR2 and miR-HSUR5 suggests that they use the same signal and mechanism for 3′ end formation. In this study, we refer to the processing intermediate containing a HVS pre-miRNA hairpin and its downstream sequences as HVS pri-miRNA (Fig. 1B).

We first asked whether the 20-nt genomic sequence downstream from pre-miR-HSUR4 (nucleotides 193–212; nucleotide 1 is the first nucleotide of HSUR4), which contains the miR 3′ box, is necessary for miR-HSUR4 biogenesis. We transfected a plasmid containing the HVS strain A11 genomic sequence encoding HSUR4 and its downstream miR-HSUR4, driven by the cellular U1 snRNA promoter (referred to as pHSUR4 hereafter), into human embryonic kidney (HEK) 293T cells (Cazalla et al. 2011). HEK293T cells were used for ease of transfection and because they faithfully express HVS miRNAs upon transfection of HVS miRNA genes (Cazalla et al. 2011). Northern blots showed that transfection of miR 3′ box-containing plasmids generates pre-miR-HSUR4 and mature miR-HSUR4 RNAs that can be detected by probes complementary to miR-HSUR4-5p or miR-HSUR4-3p (mature miRNAs derived from the 5′ or 3′ arm of the pre-miR-HSUR4 hairpin) (Supplemental Fig. S1B, lanes 2,3). No specific RNA band longer than pre-miR-HSUR4 is detected, suggesting the absence of pri-miRNAs of a defined length. In contrast, the plasmid containing no HVS

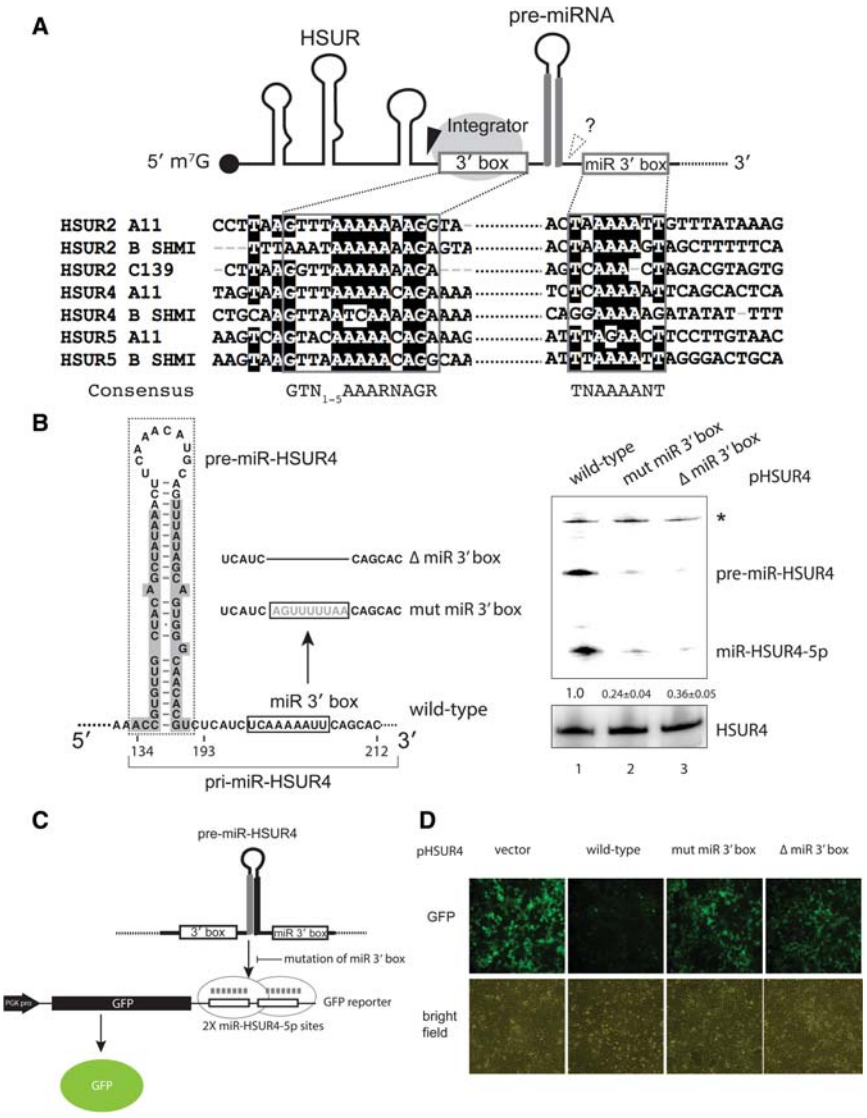

Figure 1. Identification of the miR 3′ box that is essential for HVS miRNA biogenesis. (*A*) Schematic of the HSUR pre-miRNA chimeric transcript. Gray bars represent mature miRNA-5p and miRNA-3p. The black triangle indicates cleavage by the Integrator complex (gray oval). The dashed-line triangle indicates the site of 3′ end formation investigated in this study. The multiple sequence alignment of HVS genomic sequences shows conserved nucleotides in black, while the 3′ box and miR 3′ box (identified in this study) are outlined by gray boxes; consensus sequences are shown *below* the alignment. The consensus sequence of the snRNA 3′ box is derived from HSUR1–7 as well as human snRNAs, including U1, U2, U4, U5, U7, U11, and U12. (*B*) pre-miR-HSUR4 from HVS strain A11 is boxed by a dashed line, with mature miR-HSUR4-5p and miR-HSUR4-3p shaded gray. The *right* panel shows a Northern blot probed for miR-HSUR4-5p (*above*) and HSUR4 (*below*). Total RNA was isolated from 293T cells transfected with pHSUR4 containing either the wild-type or mutated miR 3′ box, as shown in the schematic. Quantifications [mean ± standard deviation (SD)] show the relative average of pre-miR-HSUR4 and mature miR-HSUR4-5p levels derived from three independent experiments. The asterisk labels a nonspecific band. (*C*) A GFP reporter detected mutations affecting HVS miRNA biogenesis. When essential sequence elements are mutated, GFP expression is elevated because of the absence of miR-HSUR4-5p (gray), which targets sites in the GFP mRNA 3′ untranslated region. (*D*) 293T cells were cotransfected with the pGFP-miR-HSUR4-5p reporter and the miR-HSUR4 expression vectors shown in *B*. GFP fluorescence and bright-field images of the cells after 48 h are shown.

genomic sequence downstream from pre-miR-HSUR4 generates significantly less miR-HSUR4-5p and miR-HSUR4-3p (Supplemental Fig. S1B, lane 4). Together, these data argue that the 20-nt genomic sequence downstream from pre-miR-HSUR4, which contains the miR 3′ box, is necessary for miR-HSUR4 biogenesis.

The importance of the miR 3′ box sequence was probed by mutation to a complementary sequence (mut miR 3′ box) or deletion (Δ miR 3′ box) in pHSUR4 and transfection of the mutant plasmids into 293T cells (Fig. 1B). As anticipated, both mutants yielded significantly reduced levels (~30%) of pre-miR-HSUR4 and mature miR-HSUR4-5p (Fig. 1B, lanes 2,3). By cotransfecting a GFP reporter containing two sites complementary to miR-HSUR4-5p in its 3′ untranslated region (UTR) (Fig. 1C), we confirmed that miR-HSUR4-5p levels are reduced, as reflected by the elevated GFP signal (Fig. 1D). Together, these results establish that transfection of pHSUR4 into 293T cells produces functional miRNAs that repress protein expression and that the miR 3′ box is a pre-miRNA 3′ end formation signal.

To further characterize the miR 3′ box, we tested serially deleted sequences for their ability to generate HVS miRNAs (Fig. 2A). The results argue that the central adenosine residues (nucleotides 200–204) are essential, since miR-HSUR4-5p levels drop significantly only when the adenosines are deleted (Fig. 2B [lanes 5–8,13–16], C). Individual point mutations within the miR 3′ box confirmed that the two central adenosines (A202 and A203) are most critical to miR-HSUR4 biogenesis (Fig. 2D, lanes 5,6; Supplemental Fig. S1C). Whereas A202U and A203U in the double mutant (AA-UU) are not additive (Fig. 2D, cf. lanes 10,11 and lane 12), point mutation of A202 or A203 causes a reduction in miR-HSUR4-5p comparable with deletion of the whole miR 3′ box (Fig. 2D, cf. lanes 10,11 and lane 15). Note that none of the tested

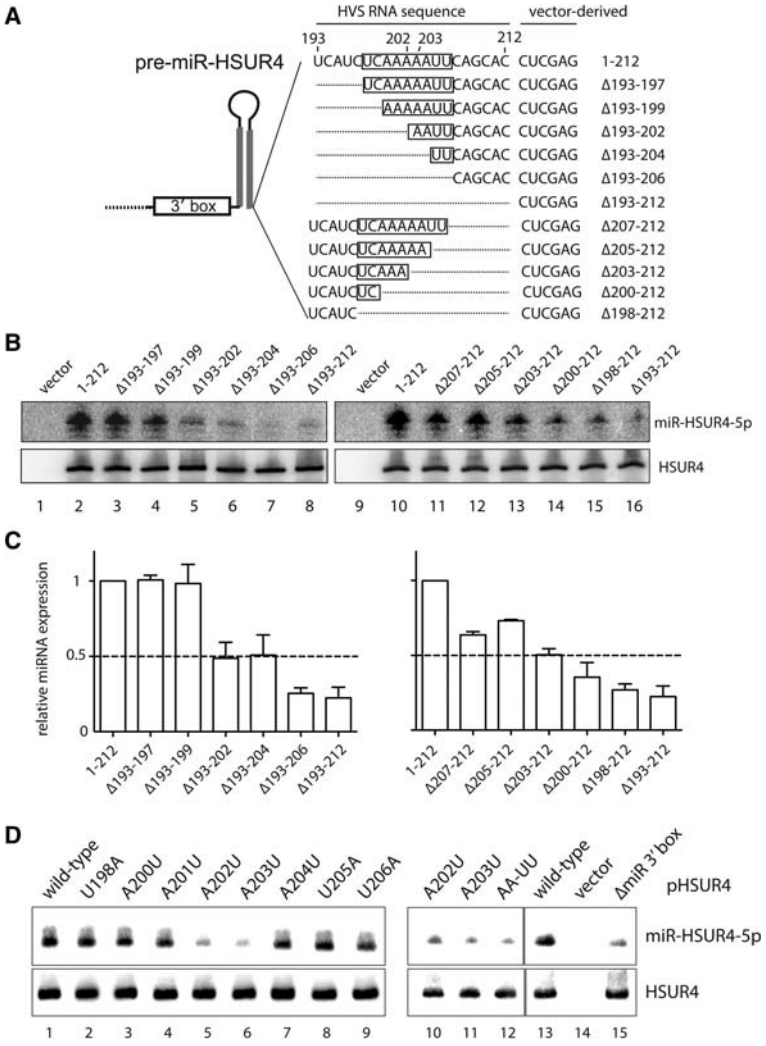

Figure 2. The central adenosine residues of the miR 3′ box are essential. (*A*) Serial deletions in the 20-nt HVS genomic sequence downstream from pre-miR-HSUR4 are illustrated, with the nucleotide numbers counting from the HSUR4 transcription start site. (*B*) The Northern blot analyzes expression of miR-HSUR4-5p and HSUR4 in total RNAs extracted from 293T cells transfected with the constructs illustrated in *A*. (*C*) Relative miR-HSUR4 levels in *B* are graphed. Error bars represent SD from three experiments. (*D*) Northern blot of total RNA extracted from 293T cells transfected with pHSUR4 wild-type or the indicated mutant constructs, probed for miR-HSUR4-5p and HSUR4.

mutations completely eliminates pre-miR-HSUR4 3′ end formation, reminiscent of the partial inhibition of snRNA 3′ end formation observed after mutating the snRNA 3′ box (Ach and Weiner 1987). We conclude that HVS evolved a unique adenosine-rich miR 3′ box that is essential for viral miRNA biogenesis.

The 3′ end of HVS pre-miRNA is not produced by exonuclease trimming

To elucidate the molecular mechanism of HVS pre-miRNA 3′ end formation, we first asked whether the HVS pre-miRNA 3′ end is created by the exosome complex, a major 3′-to-5′ exonuclease involved in RNA turnover in eukaryotic cells (Houseley et al. 2006). It was previously reported that the exosome trims 3′-tailed mirtrons (Flynt et al. 2010). We used siRNA to knock down Dis3 (also known as Rrp44), a core component of both the nuclear and cytoplasmic exosome complexes (Tomecki and Dziembowski 2010), followed by transfection of pHSUR4. Reduced miR-HSUR4-5p levels were not observed even though the Dis3 protein was successfully depleted (Supplemental Fig. S2A), arguing that the exosome complex does not participate in HVS pre-miRNA 3′ end formation.

To test the possibility that the HVS pre-miRNA might be generated by another 3′-to-5′ trimming activity, we inserted into pHSUR4 18 consecutive guanosine residues downstream from pre-miR-HSUR4 at two different locations 20 nt apart (G_{18} at nucleotide 200 or 220) (Supplemental Fig. S2B). Poly-G tracts were previously used to block exonuclease trimming, while poly-A tracts served as negative controls (Anderson and Parker 1998). When the mutant plasmids were transfected into 293T cells, insertion of either G_{18} or A_{18} at nucleotide 220 did not negatively impact miR-HSUR4-5p levels (Supplemental Fig. S2C, cf. lanes 5,6 and lane 1). In contrast, at nucleotide 200, G_{18} insertion severely reduced pre-miRNA and mature miRNA production, whereas insertion of A_{18} had no effect (Supplemental Fig. S2C, lanes 3,4). Because insertion of G_{18} farther downstream (nucleotide 220) did not block miR-HSUR4-5p biogenesis and because the miR 3′ box was disrupted by addition of G_{18} but not A_{18} at nucleotide 200, we conclude that the production of

the pre-miR-HSUR4 3′ end is likely to be exonuclease-independent.

Integrator catalytic activity is critical for HVS pre-miRNA 3′ end processing

Since the miR 3′ box resembles the adenosine-rich snRNA 3′ box, we asked whether it is recognized by the Integrator complex. We designed a m^{7}G-capped pre-miR-HSUR4 expression construct (pmiR-HSUR4) in which the 5′ end of pre-miR-HSUR4 corresponds to the Pol II transcription initiation site instead of an Integrator processing site, allowing specific examination of the role of Integrator in pre-miRNA 3′ end formation (Fig. 3A). The encoded RNA undergoes 3′ end formation to generate the pre-miR-HSUR4 intermediate and mature miR-HSUR4-3p but does not produce stable miR-HSUR4-5p due to the presence of the m^{7}G-cap (Supplemental Fig. S3, lane 3; Xie et al. 2013). As for pHSUR4, Northern blot analysis did not detect defined pri-miR-HSUR4 transcripts in total RNA extracted from pmiR-HSUR4 transfected cells (Supplemental Fig. S3, cf. lanes 2 and 3). When the abundance of the catalytic subunit of Integrator (Int11) was reduced by ~70% using siRNA, 293T cells transfected with pmiR-HSUR4 generated significantly reduced levels of both pre-miR-HSUR4 and mature miR-HSUR4-3p (Fig. 3B, lane 2). Coexpression of a wild-type but not a catalytically inactive (E203Q) Int11 rescued miRNA biogenesis (Fig. 3B, lanes 3,4). The E203Q mutation was previously shown not to affect Integrator complex assembly but to cause snRNA misprocessing (Baillat et al. 2005). Knockdown and rescue of Int11 showed similar effects on HSUR4 expression (encoded on a separate plasmid that does not generate miR-HSUR4) but not on Pol III transcribed EBER1 (Epstein-Barr virus-encoded RNA 1) (Fig. 3B). These data indicate that Integrator activity is essential for pre-miR-HSUR4 3′ end formation.

Other than processing the 3′ end of snRNAs, Integrator is also implicated in Pol II pause release and transcription elongation (Gardini et al. 2014; Yamamoto et al. 2014). To distinguish whether the reduced miRNA levels in the absence of Int11 are due to a processing or a transcriptional defect, we carried out RNase protection assays (RPAs) by annealing ^{32}P-body-labeled RNA probes complementary to pri-miR-HSUR4 with total RNA from transfected 293T cells and digesting with single-strand-specific RNases (Fig. 3C). Because RPA does not require the transcripts to have a uniform 3′ end, it is superior to Northern blot analysis for detecting pri-miR-HSUR4, which apparently contains a heterogeneous 3′ end (Supplemental Fig. S3). Knocking down Int11 led to reduced levels of both pre-miR-HSUR4 and mature miR-HSUR4-3p as well as increased amounts of pri-miR-HSUR4 (Fig. 3D, cf. lanes 5 and 6). These data argue that Int11 acts directly

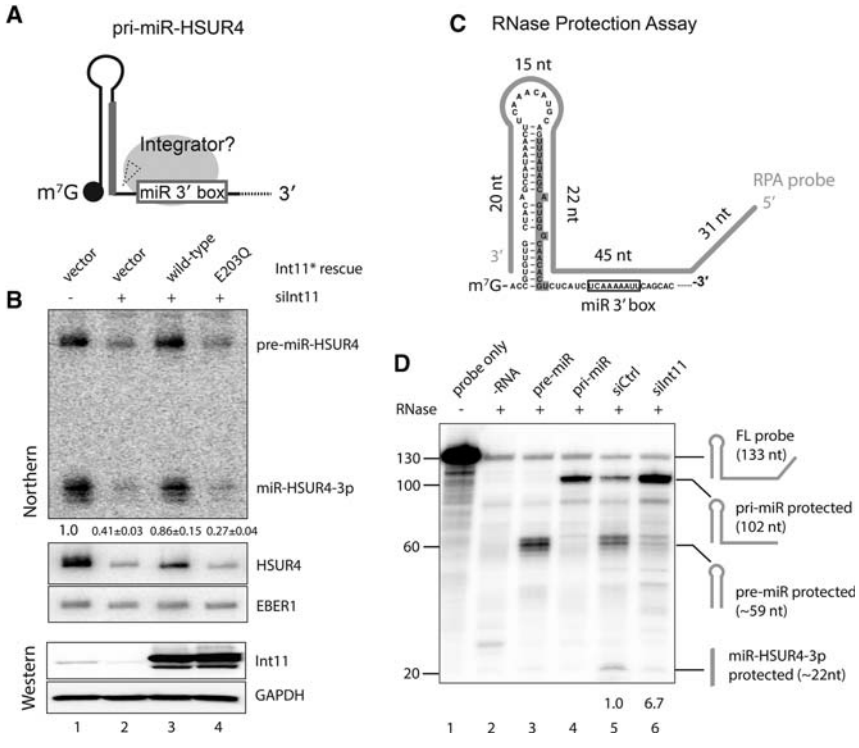

Figure 3. Integrator activity is required for HVS pre-miRNA 3′ end processing. (*A*) Schematic of pri-miR-HSUR4 transcribed from the pmiR-HSUR4 construct. The gray bar represents miR-HSUR4-3p. The dashed-line triangle indicates potential cleavage by the Integrator complex (gray oval). (*B*) Northern blot analyzing the levels of miR-HSUR4-3p, HSUR4, and EBER1 RNAs in 293T cells treated with a control siRNA (–) or a siRNA against Int11 (+) followed by cotransfection of plasmids expressing siRNA-resistant (*) Int11 wild-type or E203Q and the three RNAs. Western blots show knockdown of endogenous Int11 and the expression of siRNA-resistant Int11, with GAPDH as a loading control. Quantifications (mean ± SD) show the relative average of pre-miR-HSUR4-3p and mature miR-HSUR4-3p levels derived from three independent experiments. (*C*) The cartoon depicts the pri-miR-HSUR4 transcript, with miR-HSUR4-3p shaded gray, and the riboprobe used shown as a gray line. The lengths of each protected fragment are indicated. (*D*) The RNase protection assay detects pri-miR-HSUR4, pre-miR-HSUR4, and mature miR-HSUR4-3p simultaneously. Free probe is in lane *1*. RNase reactions were carried out in the presence of 5 µg of yeast total RNA with no target RNA (lane *2*), 50 pg of in vitro transcribed pre-miR-HSUR4 or pri-miR-HSUR4 marker RNAs (lanes *3*,*4*), or 5 µg of total RNA isolated from 293T cells pretreated with a nonspecific siRNA (siCtrl) or siInt11 followed by transfection of pmiR-HSUR4. Protected fragments are identified at the *right*. The ratio of pri-miRNA to the sum of pre-miR-HSUR4 and mature miR-HSUR4-3p is given for siCtrl and siInt11 reactions.

in 3′ end processing rather than by elevating the transcription level of pre-miR-HSUR4.

Integrator interacts with primary HVS miRNA substrate in vivo

Although our results indicate that Integrator activity is essential for HVS pre-miRNA 3′ end processing, the effect may be indirect. To test whether the catalytic subunit of Integrator complex associates with and processes HVS pri-miRNAs, we first performed RNA immunoprecipitation experiments. We knocked down endogenous Int11 by siRNA treatment of 293T cells and then transiently expressed both a Flag-tagged Int11, which carries an active site mutation (E203Q), and m^7G-capped pre-miR-HSUR4 followed by a wild-type or mut miR 3′ box (see Fig. 1B). The Int11 E203Q mutant assembles into a complete Integrator complex that may be retained longer on substrates (Baillat et al. 2005). After treating the cells with formaldehyde, we immunoprecipitated the Integrator complex from cell lysates using anti-Flag antibodies, eluted the cross-linked protein complexes

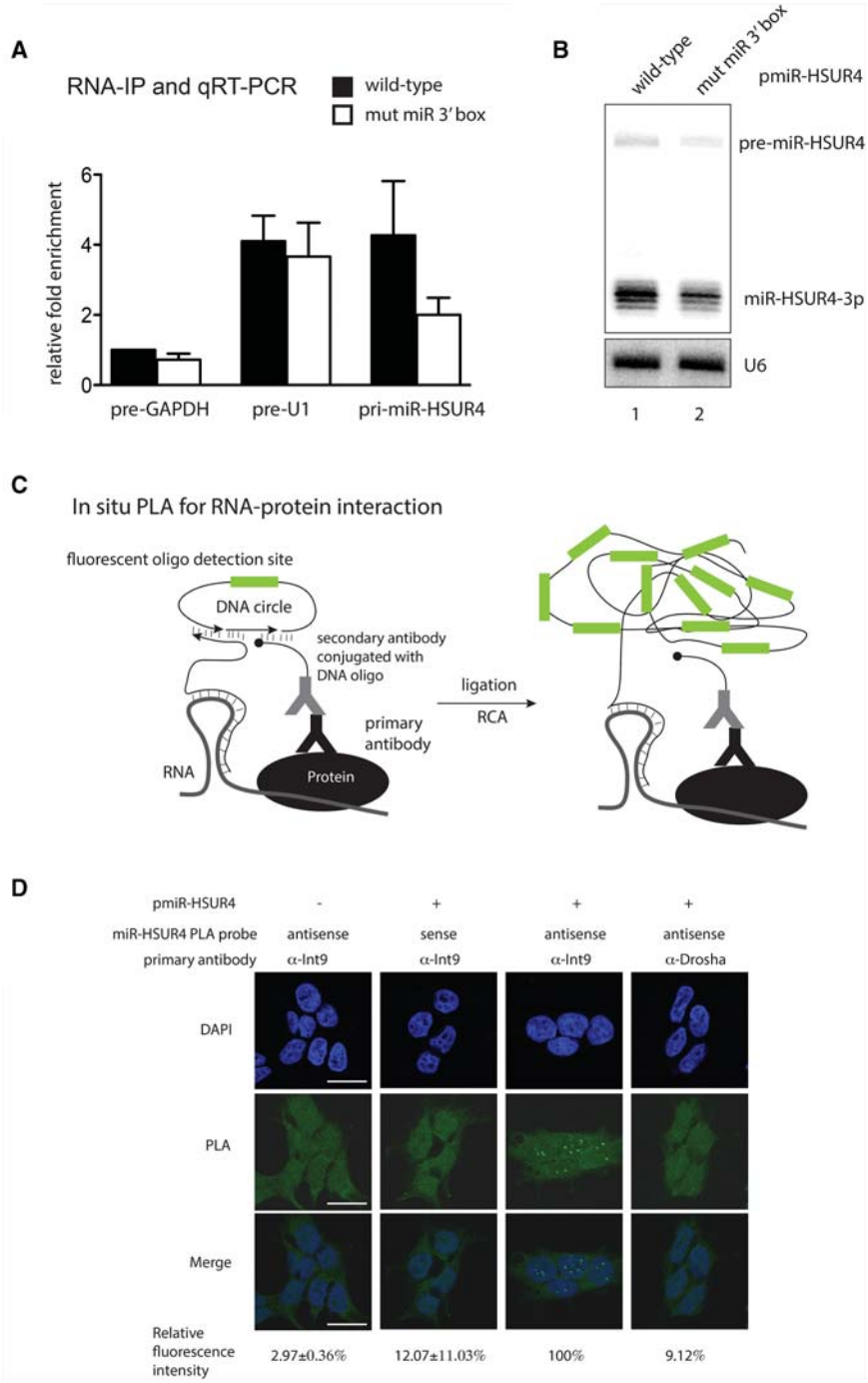

Figure 4. Coimmunoprecipitation and PLA evidence for association of the Integrator complex with the HVS pri-miRNA substrate. (*A*) 293T cells were treated with siInt11 and sequentially transfected with plasmids expressing siRNA-resistant Flag-Int11-E203Q and pri-miR-HSUR4 containing a wild-type or mutant miR 3′ box. Cell lysates were subjected to anti-Flag or control IgG selection. Precipitated proteins and bound RNAs were eluted with 3xFlag peptides. RT-qPCR quantified the relative enrichment (anti-Flag/IgG) of cellular pre-GAPDH, cellular pre-U1, and pri-miR-HSUR4 from cells transfected with pmiR-HSUR4 (shown in Fig. 3A) containing a wild-type (black) or mutated (white) miR 3′ box. Fold enrichment relative to pre-GAPDH enriched from wild-type pri-miR-HSUR4-expressing cells was calculated. Error bars represent SD from three experiments. (*B*) Northern blot analysis of miR-HSUR4-3p and endogenous U6 from 293T cells transfected with pmiR-HSUR4 containing a wild-type or mutated miR 3′ box. (*C*) Schematic of the in situ RNA–protein PLA procedure (see the Results and the Materials and Methods). (RCA) Rolling circle amplification. (*D*) 293T cells were transfected with empty vector or pmiR-HSUR4 for 48 h and assayed by PLA using DNA probes that were sense or antisense to the loop continuing into the 3p arm of the pre-miR-HSUR4 hairpin and anti-Int9 or anti-Drosha antibodies. Nuclei were stained with DAPI. Bar, 20 µm.

with 3xFlag peptides, reversed the cross-links, and performed RT-qPCR to detect associated RNAs (Supplemental Fig. S4A). Both Integrator-cleaved pre-U1 snRNA (endogenous) and pri-miR-HSUR4 were enriched in the precipitate compared with CPSF-cleaved GAPDH pre-mRNA (endogenous) (Fig. 4A). Furthermore, mutation of the miR 3′ box decreased the coimmunoprecipitation of Integrator and pri-miR-HSUR4 (Fig. 4A) as well as the yield of pre-miR-HSUR4 and mature miR-HSUR4-3p (Fig. 4B). These results argue that the miR 3′ box recruits Integrator to enable HVS miRNA biogenesis.

An in situ PLA visualizes the Integrator–HVS pri-miRNA interaction in vivo

To visualize the interaction between the Integrator complex and HVS pri-miRNA in cells, we developed an in situ PLA as a highly sensitive method to detect specific RNA–protein interactions in vivo. We adapted the classical PLA procedure, which is widely used to detect protein–protein interactions (Supplemental Fig. S4B; Soderberg et al. 2006). There, if two proteins are in close proximity (<40 nm), two primary antibodies and corresponding secondary antibodies conjugated to DNA oligonucleotides can bring together another pair of oligonucleotides that are circularized by DNA ligase. The ligated circle then serves as a template for rolling circle amplification (RCA) using one of the antibody-conjugated oligonucleotides as primer. This amplified DNA is subsequently readily detected by fluorescently labeled complementary oligonucleotides, indicating a protein–protein interaction.

For RNA–protein PLA, we replaced one pair of primary and secondary antibodies by a DNA oligonucleotide that both anneals to the RNA of interest and serves as the primer for RCA (Fig. 4C). Specifically, we transfected 293T cells with a plasmid expressing m^7G-capped pre-miR-HSUR4 and performed PLA using a primary antibody against Int9, which heterodimerizes with Int11, and an oligonucleotide that anneals to the pre-miR-HSUR4 loop region and the 3p arm. Green fluorescent signals were detected as dots in the cell nuclei only when pri-miR-HSUR4 was present and when the detection oligonucleotide was antisense to pri-miR-HSUR4 (Fig. 4D). Moreover, even though an anti-Drosha antibody detected abundant Drosha in the cell nuclei (Supplemental Fig. S4C), no strong positive RNA–protein PLA signals were observed, since miR-HSUR4 biogenesis is Drosha-independent (Fig. 4D; Cazalla et al. 2011). These data suggest that in situ PLA successfully captures the Integrator–pri-miR-HSUR4 interaction in cells. This new method could potentially be used to detect and localize other in vivo RNA–protein interactions with high sensitivity.

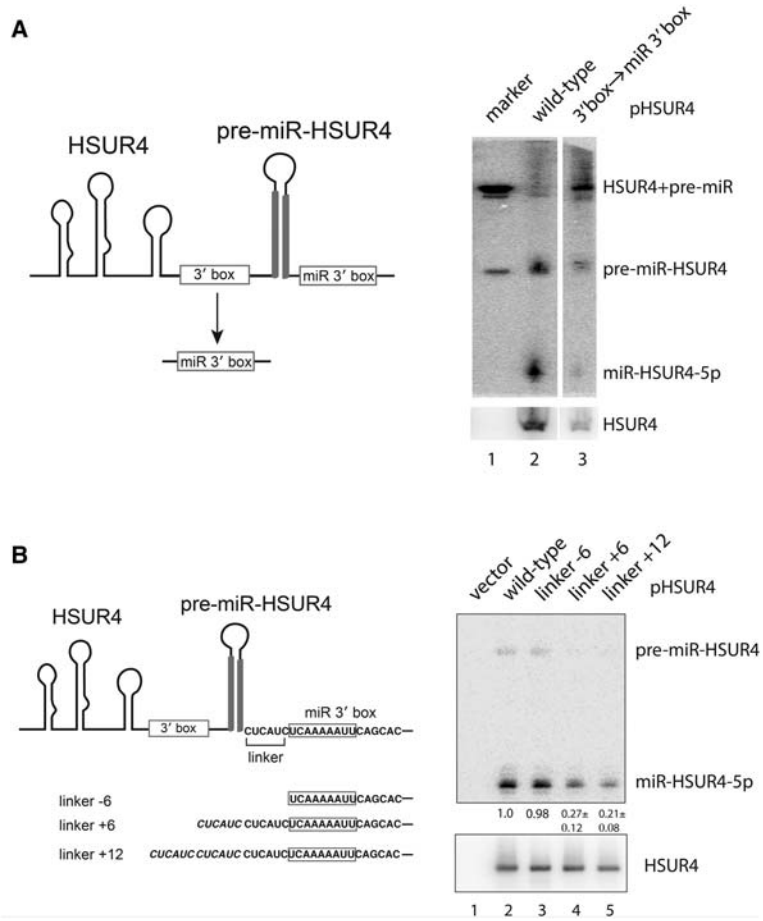

Figure 5. The miR 3′ box functions in proximity to a pre-miRNA stem–loop. Northern blots probed for miR-HSUR4-5p and HSUR4 analyzed total RNAs extracted from 293T cells transfected with an empty vector, pHSUR4 (wild type), pHSUR4 (3′ box → miR 3′ box) (*A*), or pHSUR4 carrying one of the linker deletion/insertion mutants depicted *below* (*B*). In pHSUR4 (3′ box → miR 3′ box), the HSUR4 3′ box (underlined) and its flanking sequences (5′-CTTAGTAAGTTTAAAAACAGAAAAAA-3′) were substituted by the miR 3′ box (underlined) and its flanking sequences (5′-TCATCTCTCAAAAATTCA-3′). In vitro transcribed markers are a 192-nt HSUR4 + pre-miR-HSUR4 chimeric transcript and a 59-nt pre-miR-HSUR4. Relative miR-HSUR4 expression levels (mean ± SD) were derived from two experiments.

Integrator-dependent HVS pre-miRNA 3′ end processing is distinct from snRNA 3′ end processing

Because both the snRNA 3′ box and the miR 3′ box direct Integrator to sites of 3′ end processing, we asked whether these signals are interchangeable. When the snRNA 3′ box is substituted by the miR 3′ box in pHSUR4, HSUR4 snRNA 3′ end processing is reduced, as levels of both HSUR4 and miR-HSUR4-5p decrease (Fig. 5A, cf. lanes 2 and 3). In contrast, 3′ end processing of pre-miR-HSUR4 is not affected, since a species the same size as the marker, which contains both HSUR4 and pre-miR-HSUR4, accumulates (Fig. 5A, cf. lanes 1 and 3). We reasoned that perhaps the miR 3′ box cannot replace the snRNA 3′ box because the miR 3′ box needs to be precisely positioned downstream from a pre-miRNA hairpin in order to be functional. Indeed, insertion of 6 or 12 nt between the miR 3′ box and pre-miR-HSUR4 severely reduces miR-HSUR4-5p production (Fig. 5B). These data suggest that although both 3′ end processing events that generate HSUR4 and pre-miR-HSUR4 are Integrator-dependent, they are distinct and functionally separable.

HVS pre-miRNA 3′ end processing is uncoupled from transcription

snRNA 3′ end processing requires a snRNA-type promoter and is believed to be coupled to transcription (Hernandez and Weiner 1986; Hernandez and Lucito 1988). We asked whether the requirement for a snRNA-type promoter likewise applies to pre-miR-HSUR4 3′ end processing. We compared pre-miR-HSUR4 generated from a construct containing its native HSUR4 (snRNA-type) promoter with a cytomegalovirus (CMV) immediate–early gene (mRNA-type) promoter in 293T cells (constructs c and d in Fig. 6A). Surprisingly, similar amounts of miR-HSUR4-3p were produced (Fig. 6B, lanes 3,4). Importantly, CMV promoter-driven pre-miR-HSUR4 3′ end processing was nonetheless Integrator-dependent (Supplemental Fig. S5A). In contrast, in cells transfected with constructs a and b (Fig. 6A), CMV promoter-driven HSUR4 snRNA and its downstream miR-HSUR4-3p were produced at much lower levels than from the native HSUR4 promoter, as observed previously (Fig. 6B, lanes 1,2; Cazalla et al. 2011).

These results suggest that pre-miR-HSUR4 3′ end processing by Integrator is uncoupled from transcription. To confirm this conclusion, we in vitro synthesized pri-miR-HSUR4 RNA substrates that contain a 5′ m^7G-cap and a 33-nt tail with a phosphorothioate backbone and 2′-*O*-methyl modifications near the 3′ end to counteract potential nonspecific exonuclease trimming (Supplemental Fig. S5B). After gel purification and transfection into 293T cells, the pri-miR-HSUR4 RNA with the wild-type miR 3′ box was successfully processed into mature miR-HSUR4-3p (Fig. 6C), demonstrating that processing

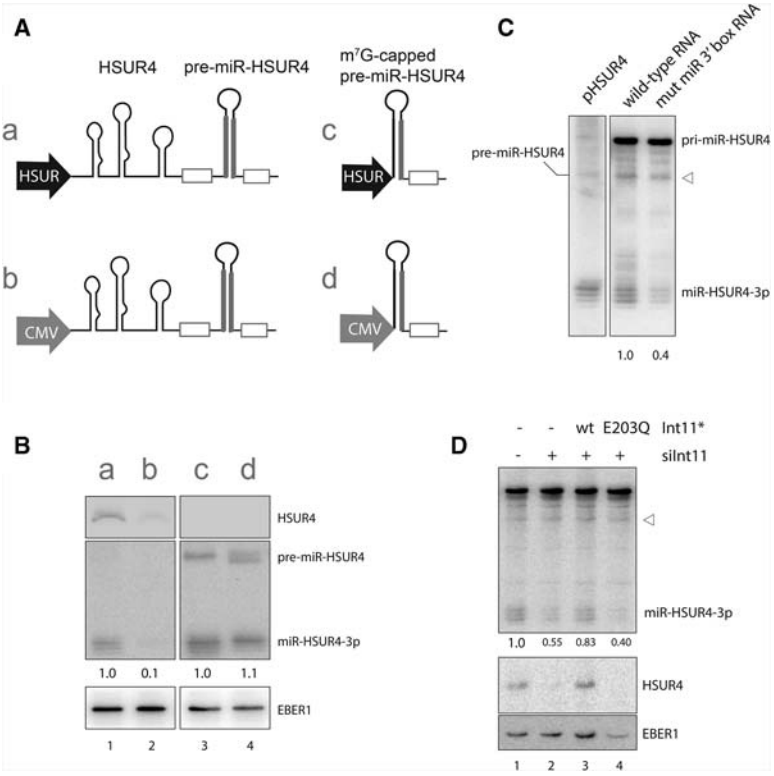

Figure 6. HVS pre-miRNA 3′ end processing is uncoupled from transcription. (*A*) Schematic of four different constructs that encode the HSUR4–pre-miR-HSUR4 chimeric transcript (constructs a and b) or pre-miR-HSUR4 alone (constructs c and d) driven by the native HSUR4 or CMV immediate–early gene promoter. Gray bars represent mature miR-HSUR4-5p and miR-HSUR4-3p. (*B*) 293T cells were cotransfected with one of the four different constructs depicted in *A* together with a plasmid expressing EBER1. Total RNAs were extracted and analyzed by Northern blot, probing for HSUR4, miR-HSUR4-3p, and EBER1. Relative levels of miR-HSUR4-3p are given. (*C*) Northern blot analyzing levels of miR-HSUR4-3p in total RNA extracted from 293T cells transfected with pHSUR4 or in vitro synthesized pri-miR-HSUR4 RNA with either a wild-type or mutated miR 3′ box. The arrowhead points to a degradation intermediate comigrating with pre-miR-HSUR4 (see Supplemental Fig. S5C). Relative levels of miR-HSUR4-3p are quantitated. The sequences of the in vitro synthesized pri-miR-HSUR4 RNAs are given in Supplemental Figure S5B. (*D*) Northern blot probed for miR-HSUR4-3p, HSUR4, and EBER1 to analyze total RNAs extracted from 293T cells or 293T cells stably expressing siRNA-resistant (*) Int11 (wild type or E203Q). The arrowhead points to a degradation intermediate comigrating with pre-miR-HSUR4, as in *C*. Relative levels of miR-HSUR4-3p are quantitated. Cells were sequentially transfected with siInt11 and in vitro synthesized pri-miR-HSUR4 RNA together with a HSUR4-expressing plasmid that does not generate downstream miR-HSUR4 and an EBER1-expressing plasmid.

can occur uncoupled from transcription. Similar to pri-miR-HSUR4 transcribed from plasmids (Figs. 1B, 3A), efficient processing of the transfected pri-miR-HSUR4 RNA depended on the presence of an intact miR 3′ box (Fig. 6C) and an active Integrator complex (Fig. 6D). Therefore, in contrast to snRNA 3′ end processing, HVS pre-miRNA 3′ end processing does not strictly require a snRNA-type promoter and can occur independently of transcription.

Discussion

An unusual Integrator cleavage in HVS miRNA biogenesis

We uncovered a novel mode of Integrator cleavage that occurs at the 3′ ends of HVS pre-miRNAs. The three HVS pre-miRNAs were previously discovered to be cotranscribed downstream from HSUR2, HSUR4, and HSUR5 and released from their upstream snRNAs by 5′ end cleavage requiring the Integrator complex (Cazalla et al. 2011); the resulting mature HVS miRNAs are involved in regulating cell cycle progression in HVS-infected cells (YE Guo, T Oei, and JA Steitz, in prep.). Fusing two noncoding RNAs into a chimeric primary transcript may be a strategy used by the virus because of its limited genome size. It also seems reasonable for the virus to hijack the host Integrator complex, which is recruited to all snRNA gene promoters (Egloff et al. 2007), to execute a second cleavage of the same substrate, thereby producing both the 5′ and 3′ ends of the HVS pre-miRNAs. Notably, Integrator-generated pre-miR-HUSR4 has noncanonical 5′ and 3′ overhangs compared with Microprocessor-generated pre-miRNAs, which contain 2-nt 3′ overhangs (Fig. 1B; Cazalla et al. 2011). Nonetheless, in vitro synthesized pre-miR-HSUR4 is efficiently cleaved by recombinant human Dicer (Cazalla et al. 2011), consistent with Dicer's versatile "counting rules" involving recognition of the 5′ end, the 3′ end, or the loop of the pre-miRNA (Park et al. 2011; Gu et al. 2012).

Similar to the 3′ box-directed processing of the HSUR snRNAs, pre-miRNA 3′ end processing is also directed by a conserved sequence element, the miR 3′ box (Figs. 1, 2). Whereas 5′ end formation of HVS pre-miRNAs was previously shown to be Integrator-dependent (Cazalla et al. 2011), use of a m^7G-capped pre-miRNA expression construct allowed us to unequivocally assign HVS pre-miRNA 3′ end processing to the catalytic activity of the Integrator complex as well (Fig. 3). Furthermore, we provide evidence that Integrator directly cleaves HVS pri-miRNA substrates, since in vivo interactions are detected by both coimmunoprecipitation and a novel RNA–protein PLA methodology (Fig. 4). Although both HSUR and HVS pre-miRNA 3′ end processing are Integrator-dependent, we were surprised to find that the latter is independent of snRNA-type promoter-directed transcription and can occur completely uncoupled from transcription (Fig. 6).

One difference between snRNA and HVS pre-miRNA 3′ end processing is apparent from comparing the snRNA 3′ box and miR 3′ box sequences. Although both are rich in adenosine, the miR 3′ box is shorter, missing the consensus GT and AGR signatures located at the 5′ and 3′ ends of the snRNA 3′ box, respectively (Fig. 1A). Perhaps the longer 3′ box evolved for snRNA processing because it is the only RNA sequence element required for Integrator cleavage. Although the 3′-terminal stem–loop of the *Drosophila* U7 snRNA is important for 3′ end processing with a GFP reporter attached (Ezzeddine et al. 2011), others reported that such a stem–loop is not always essential for Integrator-dependent cleavage of snRNAs (Hernandez 1985; Yuo et al. 1985). We likewise conclude that the HSUR4 terminal stem–loop is not essential for snRNA 3′ end processing, since removing the entire HSUR4 sequence does not affect miR-HSUR4 production, which requires 5′ cleavage by Integrator (Cazalla et al. 2011). On the other hand, the miR 3′ box functions in proximity to a pre-miRNA stem–loop. Positioning the miR 3′ box either downstream from the shorter HSUR4 terminal stem–loop (Fig. 5A) or >18 nt downstream from the pre-miRNA hairpin eliminated its ability to direct efficient Integrator cleavage (Fig. 5B). Probably there is not a defined pre-miRNA sequence required for the 3′ end processing, as the three HVS pre-miRNAs have diverse sequences (Supplemental Fig. S1A). We also swapped the 5p and 3p of pre-miR-HSUR4 and observed efficient production of miRNAs from that construct (Cazalla et al. 2011). Because our in vivo system uses the pre-miRNA and mature miRNA as indirect readouts of processing efficiency, we were unable to definitively measure the impact of the length of the stem–loop on 3′ end cleavage efficiency (data not shown). Likewise, although we propose that Integrator cleavage directly generates the HVS pre-miRNA 3′ ends, our in vivo processing system is inadequate to address the exact sites of cleavage. Future work will focus on devising an in vitro system to answer such questions.

An in vitro Integrator cleavage system for snRNA substrates has been difficult to develop, since the processing appears to be coupled to transcription (Uguen and Murphy 2003). By using a HVS pri-miRNA substrate whose cleavage is uncoupled from transcription, a direct in vitro assay for Integrator may be achievable. Such a system will undoubtedly provide new opportunities to dissect the roles of the various Integrator subunits. The composition of the Integrator complexes that process the 3′ ends of the HSUR snRNAs and the 3′ ends of HVS pre-miRNAs is likely to be the same. Indeed, we show that Int9 and Int11, believed to comprise the catalytic core of the snRNA-processing Integrator, both associate with the pri-miR-HSUR4 substrate (Fig. 4). It remains to be determined whether the RNA-binding subunit within the Integrator complex is the same for both snRNA and pre-miRNA processing and how two different modes of substrate specificity are conferred.

Noncanonical Integrator substrates on cellular RNAs?

Because viruses commonly acquire mechanisms from their hosts, it is possible that unusual HVS pri-miRNA-like substrates exist in uninfected cells (Tycowski et al.

2015). Moreover, because miR 3′ box processing occurs independently of snRNA promoter-driven transcription, such substrates would not be limited to snRNA precursors. Since the discovery of Integrator as the long-sought factor in snRNA processing (Baillat et al. 2005), new roles have been assigned to the Integrator complex: regulating RNA Pol II transcription promoter-proximal pausing as well as termination (Baillat and Wagner 2015). Chromatin association analyses have shown that Integrator subunits bind to not only snRNA genes but also numerous protein-coding genes (Gardini et al. 2014; Stadelmayer et al. 2014; Skaar et al. 2015). Intriguingly, when Stadelmayer et al. (2014) analyzed Integrator-responsive mRNA genes, 40% were found to contain 3′ box-like sequences close to the transcription termination site (<1.5 kb). Moreover, the presence of a 3′ box was linked to a Pol II processivity defect over termination sites in Int11-depleted cells (Stadelmayer et al. 2014). It would be interesting to examine whether there are stem–loop structures upstream of these 3′ box sequences. In the same study, Integrator was found to regulate transcription elongation at the HIV TAR element (Stadelmayer et al. 2014), which forms a pre-miRNA-like stem–loop structure (Ouellet et al. 2008). Integrator activity has also been implicated in the pausing of Pol II at the TAR element (Stadelmayer et al. 2014). Perhaps structure and sequence motifs like HVS pri-miRNAs exist in these cellular and viral RNAs and direct Integrator cleavage, which in turn regulates RNA Pol II pausing and termination.

The PLA for detecting transient RNA–protein interactions

PLA was initially designed to reveal in situ protein–protein interactions at single-molecule resolution and has been widely used on fixed cells and intact tissues (Fredriksson et al. 2002; Soderberg et al. 2006; Bellucci et al. 2014; Lipovsky et al. 2015). Recent studies have described PLA-based methods to examine specific RNA–protein interactions (Jung et al. 2013; Weibrecht et al. 2013); however, the procedure either requires converting RNA to cDNA or uses complicated peptide-modified multiplex RNA-imaging probes.

Here, we successfully developed a simpler and powerful PLA assay to detect RNA–protein interactions in situ. Even if contact is transient, robust signals are generated when the RNA and protein are within close proximity (30–40 nm); here, the RNA targeting probe contains an ~40-nt linker that is estimated to span ~13 nm, while each antibody is ~10 nm across. Compared with traditional methods for detecting RNA–protein association by combining fluorescence in situ hybridization and immunofluorescence, this new PLA protocol requires much closer proximity to capture an interaction, eliminates nonspecific colocalization, and provides data for statistical analyses.

Using this new PLA method, we captured specific interactions between Integrator and HVS pri-miRNA. The specificity is highlighted by the expected absence of PLA signals for Drosha and pri-miR-HSUR4 even though both are present in the nucleus (Fig. 4D; Supplemental Fig. S4C). Interestingly, although Int9 staining is distributed throughout the nucleus (Supplemental Fig. S4C), the PLA signals for Int9 and pri-miR-HSUR4 appeared as punctate dots (Fig. 4D), perhaps indicating the location of the RNA–protein complex at discrete nuclear sites. Alternatively, the punctate pattern may reflect abundantly amplified fluorescent oligonucleotide detection sites on successfully ligated probes or could be an artifact of overexpression of pri-miR-HSUR4 at specific intranuclear locations. Our success with PLA to confirm the transient interaction between Integrator and one of its substrates in cells adds to its record as a useful tool to study viral entry mechanisms of human papillomavirus (Lipovsky et al. 2015). It is likely that this method can be adapted to a broad range of applications to investigate the spatial and temporal RNA–protein interactions that occur not only between cellular molecules but also during host–pathogen interactions.

Materials and methods

Plasmids

For expression of HSUR4 together with downstream miR-HSUR4 from the U1 promoter in 293T cells, pU1-HSUR4 was used (Cazalla et al. 2011). To express pre-miR-HSUR4 as a m^7G-capped pre-miRNA, pU1-miR-HSUR4 was used (Xie et al. 2013). For simplicity, these two plasmids are referred to here as pHSUR4 and pmiR-HSUR4. Various mutations of the miR 3′ box, as indicated in the figures, were introduced into pHSUR4 and pmiR-HSUR4 by site-directed quick-change mutagenesis. p7.4-miR-HSUR4-3pmt was used to express HSUR4 without generating the downstream miRNA (Cazalla et al. 2011). To generate pCMV-miR-HSUR4 for the synthesis of m^7G-capped pre-miR-HSUR4 from a CMV promoter, the 3′ box and flanking sequences were removed from pCMV-HSUR4ΔsnRNA (Cazalla et al. 2011). For transient expression of siInt11-resistant Flag-Int11 (wild type or E203Q mutant), a Flag tag and the Int11-coding sequence were first cloned into pcDNA3 vector via the NotI and EcoRI sites. The siInt11 targeted sequence in Int11 cDNA was mutated from 5′-AGCACATCAAGGCCTTCG-3′ to 5′-AACATATT AAAGCTTTTG-3′, and the E203 codon was mutated from GAG to CAG to generate the pNFlag-Int11 siRNA-resistant and pNFlag-Int11 E203Q siRNA-resistant constructs, respectively. To generate 293T cells stably expressing siRNA-resistant Flag-Int11 or Flag-Int11 E203Q, the EF-1α (human elongation factor-1α) short promoter (EFS), 2xFlag tag, siRNA-resistant Int11-coding sequences, and BGH polyA site were cloned into the pTYF lentiviral vector between SalI and MluI sites. The plasmid pEBV RIJ expressed EBER1 (Rosa et al. 1981).

Multiple sequence alignment

One-hundred-eighty nucleotides downstream from HSUR2, HSUR4, and HSUR5 from the reported genomic sequences of HVS strains A11, B SHMI, and C139 were aligned with the BioEdit program using the ClustalW algorithm (Albrecht et al. 1992; Fickenscher et al. 1997; Hor et al. 2001).

Cell culture, plasmid, and RNA transfection

HEK293T and HeLa cells were maintained in DMEM containing heat-inactivated FBS, penicillin/streptomycin, and L-glutamine.

For 293T cells stably expressing Flag-Int11, the medium also contained 2 µg/mL puromycin. Plasmid transfections were carried out in six-well plates using Lipofectamine 2000 (Invitrogen) or TransIT-293 (Mirus Bio) according to the manufacturers' instructions. After 48 h, total RNA was harvested using Trizol reagent (Invitrogen) and analyzed by Northern blot of RNAs, as indicated in the figures. Relative miR-HSUR4 expression levels were quantitated using the average of pre-miR-HSUR4 and mature miR-HSUR4 signals normalized to transfection and loading control (HSUR4 or EBER1). For Int11 knockdown and rescue experiments, 30 nM nonspecific control siRNA (siCtrl, 5′-AAGCGAUACCUCGUGUGUGA-3′) or siRNA against Int11 (siInt11, 5′-UCGAAGGCCUUGAUGUGCU-3′) was transfected into 5×10^5 293T cells in six-well plates using Lipofectamine RNAiMAX according to the manufacturer's instructions. Twenty-four hours later, 2 µg of pNFlag-Int11 siRNA-resistant wild-type or E203Q was transfected using TransIT-293. After another 24 h, cells were transfected with siRNAs for the second time. Finally, 24 h after the second siRNA transfection, plasmids expressing miR-HSUR4 (1.5 µg), HSUR4 (0.4 µg), and EBER1 (0.1 µg) were cotransfected using TransIT293. Total RNA was harvested and analyzed 48 h later. For pri-miR-HSUR4 RNA transfection, 50 ng of in vitro synthesized RNA was transfected into 2×10^5 cells in 12-well plates using Lipofectamine RNAiMAX. Total RNA was harvested and analyzed after 24 h. For Int11 rescue experiments, 293T cells stably expressing siRNA-resistant Flag-Int11 wild type or E203Q were first transfected by siInt11 for 48 h followed by transfection of pri-miR-HSUR4 RNA for 24 h. Relative expression levels of miR-HSUR4-3p were normalized to transfected pri-miR-HSUR4 RNA (Fig. 6C) or EBER1 (Fig. 6D). Dis3 knockdown was performed as described (Xie et al. 2013).

RPA

RPAs were performed as described (Cazalla et al. 2011). Specifically, total RNA was extracted from 293T cells transfected with siCtrl or siInt11 followed by transfection of pmiR-HSUR4 and treated with DNase RQ1 (Promega) following the manufacturer's instructions. The in vitro T7 synthesized ^{32}P-body-labeled RPA probe annealed to nucleotides 135–236 of the HSUR4-miRNA chimeric transcript.

RNA immunoprecipitation and qRT–PCR

Total RNA was extracted from 293T cells prepared as described for the Int11 knockdown and rescue experiment, except that only Int11 E203Q was transfected. Cross-linking was performed in PBS with 1% formaldehyde for 10 min at room temperature; the reaction was quenched by addition of 250 µM glycine. After two washes with PBS, the cell pellet was resuspended in 1.2 mL of RIPA buffer (10 mM Tris-HCl at pH 7.5, 150 mM NaCl, 2.5 mM MgCl2, 1% Triton X-100, 0.1% SDS, 0.5% sodium deoxycholate, 1 mM DTT, 1× protease inhibitor [Calbiochem]), and cell extract was prepared by sonication. Five-hundred microliters of each lysate was loaded onto 20 µL of either anti-Flag or IgG antibody-coated beads and incubated for 2 h at 4°C with gentle agitation. The beads were then washed 12 times with high-salt wash buffer (50 mM Tris-HCl at pH 7.5, 500 mM NaCl, 1% NP40, 0.1% SDS, 1 mM EDTA, 0.5% sodium deoxycholate). Bound Integrator was eluted into 100 µL of PBS with 100 µg/mL 3xFlag peptides. The eluted material was then digested with 2 mg/mL proteinase K in a 300-µL reaction containing 50 mM Tris-HCl (pH 7.4), 5 mM EDTA, 300 mM NaCl, and 1.5% SDS for 1 h at 50°C. Cross-links were reversed by incubation for 45 min at 70°C.

For RT-qPCR, the RNA sample was first treated with RQ1 DNase (Promega), and cDNA was synthesized with SSIII RT and random primers (Invitrogen). Next, using a StepOnePlus instrument (Applied Biosystems), the cDNA was analyzed by RT-qPCR in technical triplicates using FastStart Universal SYBR Green Master (Rox) master mix (Roche) and primers for pre-GAPDH and pre-U1 (McCloskey et al. 2012) and for pri-miR-HSUR4 (forward, 5′-CCGTGTTGCTACAGCTATAAACT TC-3′; reverse, 5′-ATTACATCCTCTTCCTGTGTAATGTTTG AG-3′). The enrichment value for the RNA selected by anti-Flag antibodies was normalized to the control selection by IgG. Final enrichment values were normalized to the pre-GAPDH enrichment from the pmiR-HSUR4 transfected cell lysate. The experiment was performed three times.

PLA, immunofluorescence analysis, and antibodies

293T cells (5×10^5) were first transfected in a six-well plate with 2 µg of pmiR-HSUR4 for 24 h; 1×10^5 transfected cells were then seeded on poly-L-lysine-coated 12-mm coverslips in a 24-well plate for another 24 h. The cells were washed once with cold PBS and fixed in 4% formaldehyde in PBS for 30 min on ice. Permeabilization was performed with 1% saponin in PBS for 1 h at room temperature and washed three times with PBS. For PLA analyses, cells were blocked with 250 µL of 10 mM Tris-acetate (pH 7.5), 10 mM magnesium acetate, 50 mM potassium acetate, 250 mM NaCl, 0.25 µg/µL BSA, 0.05% Tween 20, and 20 µg/mL sheared salmon sperm DNA (sssDNA) for 1 h at 4°C. Next, 100 nM DNA oligonucleotide (sense, 5′-AC TTCAAACATGCAGTTTATAGCAGTGGGCAACACGTCTC AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAT ATGACAGAACTAGACACTCTT-3′; or antisense, 5′-GAGACG TGTTGCCCACTGCTATAAACTGCATGTTTGAAGTAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATATGAC AGAACTAGACACTCTT-3′) was added in fresh blocking buffer, heated for 3 min at 70°C, applied to the cells, and incubated for another hour at 37°C. Subsequently, the cells were washed once with 2× SSC and 0.1% Tween-20 and three times with PBS and blocked in PBST containing 1% BSA and 20 µg/mL sssDNA for 1 h at room temperature. The buffer was replaced by 1:1000 diluted primary antibody against Int9 (Abcam, ab70586) or Drosha (Cell Signaling, D28B1) in PBST blocking buffer. Subsequently, PLA was performed according to the manufacturer's instructions (Olink Bioscience), except that only an anti-rabbit minus probe was used. Quantification of the PLA signal was performed as previously described (Popa et al. 2015). For immunofluorescence, HeLa cells were prepared without transfection, and the secondary antibody used was 1:1000 diluted Alexa fluor 594 goat anti-rabbit (Invitrogen). Antibodies against Int11 (Abcam, ab75276), GAPDH (Cell Signaling, 14C10), and tubulin (Sigma) were used in Western blots.

In vitro synthesis of RNAs

RNA substrates and probes were transcribed in vitro by T7 RNA polymerase. Briefly, 100~200 ng of PCR-generated DNA template containing a T7 promoter sequence were incubated in a 10-µL reaction containing 40 mM Tris-HCl (pH 8.0); 25 mM NaCl; 8 mM $MgCl_2$; 2 mM Spermidine-$(HCl)_3$; 10 mM DTT; 1 mM ATP, CTP, GTP, and UTP; 20 U of RNase inhibitor (Roche); and 5 U of T7 RNA polymerase. To synthesize ^{32}P-labeled RNAs, 1 mM UTP was substituted with 100 µm UTP and 50 µM α-^{32}P UTP (800 Ci/mmol; Perkin-Elmer). Reactions were incubated overnight (or 1 h for ^{32}P-labeling) at 37°C. The reaction products were separated by polyacrylamide gel electrophoresis,

excised from the gel, eluted, ethanol-precipitated, and resuspended in double-distilled H_2O before use.

Acknowledgments

We are grateful to Evan Anderson, Sviatoslav Bagriantsev, Demian Cazalla, Xinguo Chen, Sam Gunderson, Nikolay Kolev, Nara Lee, and Shona Murphy for reagents and protocols; Kazimierz Tycowski, Nara Lee, Paulina Pawlica, and other members of the Steitz laoboratory for critical discussion; and Angela Miccinello for editorial assistance. This work was supported by grants from the National Institutes of Health (P01-CA016038 and K99-CA190886 to M.X., and F32-AI114132 to W.Z.) J.A.S. is an Investigator at the Howard Hughes Medical Institute. M.X. and J.A.S. conceived the project and were involved in all experiments. M.X., W.Z., and D. D. designed the RNA–protein PLA. W.Z. performed and analyzed the PLA. M.S., A.X., and D.A.L. helped with plasmid construction and Northern blotting. A.X. performed and analyzed the RNA immunoprecipitation. D.A.L. carried out pilot experiments for the RPA. M.X., W.Z., D.D., and J.A.S. wrote the paper.

References

Ach RA, Weiner AM. 1987. The highly conserved U small nuclear RNA 3′-end formation signal is quite tolerant to mutation. *Mol Cell Biol* **7:** 2070–2079.

Albrecht JC, Nicholas J, Biller D, Cameron KR, Biesinger B, Newman C, Wittmann S, Craxton MA, Coleman H, Fleckenstein B, et al. 1992. Primary structure of the *Herpesvirus saimiri* genome. *J Virol* **66:** 5047–5058.

Anderson JS, Parker RP. 1998. The 3′ to 5′ degradation of yeast mRNAs is a general mechanism for mRNA turnover that requires the SKI2 DEVH box protein and 3′ to 5′ exonucleases of the exosome complex. *EMBO J* **17:** 1497–1506.

Babiarz JE, Ruby JG, Wang Y, Bartel DP, Blelloch R. 2008. Mouse ES cells express endogenous shRNAs, siRNAs, and other Microprocessor-independent, Dicer-dependent small RNAs. *Genes Dev* **22:** 2773–2785.

Baillat D, Wagner EJ. 2015. Integrator: surprisingly diverse functions in gene expression. *Trends Biochem Sci* **40:** 257–264.

Baillat D, Hakimi MA, Naar AM, Shilatifard A, Cooch N, Shiekhattar R. 2005. Integrator, a multiprotein mediator of small nuclear RNA processing, associates with the C-terminal repeat of RNA polymerase II. *Cell* **123:** 265–276.

Bellucci A, Fiorentini C, Zaltieri M, Missale C, Spano P. 2014. The 'in situ' proximity ligation assay to probe protein-protein interactions in intact tissues. *Methods Mol Biol* **1174:** 397–405.

Bogerd HP, Karnowski HW, Cai X, Shin J, Pohlers M, Cullen BR. 2010. A mammalian herpesvirus uses noncanonical expression and processing mechanisms to generate viral MicroRNAs. *Mol Cell* **37:** 135–142.

Cazalla D, Xie M, Steitz JA. 2011. A primate herpesvirus uses the Integrator complex to generate viral microRNAs. *Mol Cell* **43:** 982–992.

Cheloufi S, Dos Santos CO, Chong MM, Hannon GJ. 2010. A dicer-independent miRNA biogenesis pathway that requires Ago catalysis. *Nature* **465:** 584–589.

Chen J, Ezzeddine N, Waltenspiel B, Albrecht TR, Warren WD, Marzluff WF, Wagner EJ. 2012. An RNAi screen identifies additional members of the *Drosophila* Integrator complex and a requirement for cyclin C/Cdk8 in snRNA 3′-end formation. *RNA* **18:** 2148–2156.

Cifuentes D, Xue H, Taylor DW, Patnode H, Mishima Y, Cheloufi S, Ma E, Mane S, Hannon GJ, Lawson ND, et al. 2010. A novel miRNA processing pathway independent of Dicer requires Argonaute2 catalytic activity. *Science* **328:** 1694–1698.

Denli AM, Tops BB, Plasterk RH, Ketting RF, Hannon GJ. 2004. Processing of primary microRNAs by the Microprocessor complex. *Nature* **432:** 231–235.

Dominski Z, Yang XC, Purdy M, Wagner EJ, Marzluff WF. 2005. A CPSF-73 homologue is required for cell cycle progression but not cell growth and interacts with a protein having features of CPSF-100. *Mol Cell Biol* **25:** 1489–1500.

Egloff S, O'Reilly D, Chapman RD, Taylor A, Tanzhaus K, Pitts L, Eick D, Murphy S. 2007. Serine-7 of the RNA polymerase II CTD is specifically required for snRNA gene expression. *Science* **318:** 1777–1779.

Ezzeddine N, Chen J, Waltenspiel B, Burch B, Albrecht T, Zhuo M, Warren WD, Marzluff WF, Wagner EJ. 2011. A subset of *Drosophila* integrator proteins is essential for efficient U7 snRNA and spliceosomal snRNA 3′-end formation. *Mol Cell Biol* **31:** 328–341.

Fickenscher H, Fleckenstein B. 2001. *Herpesvirus saimiri. Philos Trans R Soc Lond B Biol Sci* **356:** 545–567.

Fickenscher H, Bokel C, Knappe A, Biesinger B, Meinl E, Fleischer B, Fleckenstein B, Broker BM. 1997. Functional phenotype of transformed human αβ and γδ T cells determined by different subgroup C strains of *Herpesvirus saimiri*. *J Virol* **71:** 2252–2263.

Flynt AS, Greimann JC, Chung WJ, Lima CD, Lai EC. 2010. MicroRNA biogenesis via splicing and exosome-mediated trimming in *Drosophila*. *Mol Cell* **38:** 900–907.

Fredriksson S, Gullberg M, Jarvius J, Olsson C, Pietras K, Gustafsdottir SM, Ostman A, Landegren U. 2002. Protein detection using proximity-dependent DNA ligation assays. *Nat Biotechnol* **20:** 473–477.

Gardini A, Baillat D, Cesaroni M, Hu D, Marinis JM, Wagner EJ, Lazar MA, Shilatifard A, Shiekhattar R. 2014. Integrator regulates transcriptional initiation and pause release following activation. *Mol Cell* **56:** 128–139.

Gregory RI, Yan KP, Amuthan G, Chendrimada T, Doratotaj B, Cooch N, Shiekhattar R. 2004. The Microprocessor complex mediates the genesis of microRNAs. *Nature* **432:** 235–240.

Gu S, Jin L, Zhang Y, Huang Y, Zhang F, Valdmanis PN, Kay MA. 2012. The loop position of shRNAs and pre-miRNAs is critical for the accuracy of dicer processing in vivo. *Cell* **151:** 900–911.

Ha M, Kim VN. 2014. Regulation of microRNA biogenesis. *Nat Rev Mol Cell Biol* **15:** 509–524.

Hernandez N. 1985. Formation of the 3′ end of U1 snRNA is directed by a conserved sequence located downstream of the coding region. *EMBO J* **4:** 1827–1837.

Hernandez N, Lucito R. 1988. Elements required for transcription initiation of the human U2 snRNA gene coincide with elements required for snRNA 3′ end formation. *EMBO J* **7:** 3125–3134.

Hernandez N, Weiner AM. 1986. Formation of the 3′ end of U1 snRNA requires compatible snRNA promoter elements. *Cell* **47:** 249–258.

Hor S, Ensser A, Reiss C, Ballmer-Hofer K, Biesinger B. 2001. *Herpesvirus saimiri* protein StpB associates with cellular Src. *J Gen Virol* **82:** 339–344.

Houseley J, LaCava J, Tollervey D. 2006. RNA-quality control by the exosome. *Nat Rev Mol Cell Biol* **7:** 529–539.

Hutvagner G, McLachlan J, Pasquinelli AE, Balint E, Tuschl T, Zamore PD. 2001. A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA. *Science* **293:** 834–838.

Jung J, Lifland AW, Zurla C, Alonas EJ, Santangelo PJ. 2013. Quantifying RNA–protein interactions in situ using modified-MTRIPs and proximity ligation. *Nucleic Acids Res* **41:** e12.

Kincaid RP, Burke JM, Sullivan CS. 2012. RNA virus microRNA that mimics a B-cell oncomiR. *Proc Natl Acad Sci* **109:** 3077–3082.

Lee Y, Ahn C, Han J, Choi H, Kim J, Yim J, Lee J, Provost P, Radmark O, Kim S, et al. 2003. The nuclear RNase III Drosha initiates microRNA processing. *Nature* **425:** 415–419.

Lipovsky A, Zhang W, Iwasaki A, DiMaio D. 2015. Application of the proximity-dependent assay and fluorescence imaging approaches to study viral entry pathways. *Methods Mol Biol* **1270:** 437–451.

Lund E, Guttinger S, Calado A, Dahlberg JE, Kutay U. 2004. Nuclear export of microRNA precursors. *Science* **303:** 95–98.

Mandel CR, Kaneko S, Zhang H, Gebauer D, Vethantham V, Manley JL, Tong L. 2006. Polyadenylation factor CPSF-73 is the pre-mRNA 3′-end-processing endonuclease. *Nature* **444:** 953–956.

McCloskey A, Taniguchi I, Shinmyozu K, Ohno M. 2012. hnRNP C tetramer measures RNA length to classify RNA polymerase II transcripts for export. *Science* **335:** 1643–1646.

Okamura K, Hagen JW, Duan H, Tyler DM, Lai EC. 2007. The mirtron pathway generates microRNA-class regulatory RNAs in *Drosophila*. *Cell* **130:** 89–100.

Ouellet DL, Plante I, Landry P, Barat C, Janelle ME, Flamand L, Tremblay MJ, Provost P. 2008. Identification of functional microRNAs released through asymmetrical processing of HIV-1 TAR element. *Nucleic Acids Res* **36:** 2353–2365.

Park JE, Heo I, Tian Y, Simanshu DK, Chang H, Jee D, Patel DJ, Kim VN. 2011. Dicer recognizes the 5′ end of RNA for efficient and accurate processing. *Nature* **475:** 201–205.

Popa A, Zhang W, Harrison MS, Goodner K, Kazakov T, Goodwin EC, Lipovsky A, Burd CG, DiMaio D. 2015. Direct binding of retromer to human papillomavirus type 16 minor capsid protein L2 mediates endosome exit during viral infection. *PLoS Pathog* **11:** e1004699.

Rosa MD, Gottlieb E, Lerner MR, Steitz JA. 1981. Striking similarities are exhibited by two small Epstein-Barr virus-encoded ribonucleic acids and the adenovirus-associated ribonucleic acids VAI and VAII. *Mol Cell Biol* **1:** 785–796.

Ruby JG, Jan CH, Bartel DP. 2007. Intronic microRNA precursors that bypass Drosha processing. *Nature* **448:** 83–86.

Skaar JR, Ferris AL, Wu X, Saraf A, Khanna KK, Florens L, Washburn MP, Hughes SH, Pagano M. 2015. The Integrator complex controls the termination of transcription at diverse classes of gene targets. *Cell Res* **25:** 288–305.

Soderberg O, Gullberg M, Jarvius M, Ridderstrale K, Leuchowius KJ, Jarvius J, Wester K, Hydbring P, Bahram F, Larsson LG, et al. 2006. Direct observation of individual endogenous protein complexes in situ by proximity ligation. *Nat Methods* **3:** 995–1000.

Stadelmayer B, Micas G, Gamot A, Martin P, Malirat N, Koval S, Raffel R, Sobhian B, Severac D, Rialle S, et al. 2014. Integrator complex regulates NELF-mediated RNA polymerase II pause/release and processivity at coding genes. *Nat Commun* **5:** 5531.

Tomecki R, Dziembowski A. 2010. Novel endoribonucleases as central players in various pathways of eukaryotic RNA metabolism. *RNA* **16:** 1692–1724.

Tycowski KT, Guo YE, Lee N, Moss WN, Vallery TK, Xie M, Steitz JA. 2015. Viral noncoding RNAs: more surprises. *Genes Dev* **29:** 567–584.

Uguen P, Murphy S. 2003. The 3′ ends of human pre-snRNAs are produced by RNA polymerase II CTD-dependent RNA processing. *EMBO J* **22:** 4544–4554.

Weibrecht I, Lundin E, Kiflemariam S, Mignardi M, Grundberg I, Larsson C, Koos B, Nilsson M, Soderberg O. 2013. In situ detection of individual mRNA molecules and protein complexes or post-translational modifications using padlock probes combined with the in situ proximity ligation assay. *Nat Protoc* **8:** 355–372.

Xie M, Steitz JA. 2014. Versatile microRNA biogenesis in animals and their viruses. *RNA Biol* **11:** 673–681.

Xie M, Li M, Vilborg A, Lee N, Shu MD, Yartseva V, Sestan N, Steitz JA. 2013. Mammalian 5′-capped microRNA precursors that generate a single microRNA. *Cell* **155:** 1568–1580.

Yamamoto J, Hagiwara Y, Chiba K, Isobe T, Narita T, Handa H, Yamaguchi Y. 2014. DSIF and NELF interact with Integrator to specify the correct post-transcriptional fate of snRNA genes. *Nat Commun* **5:** 4263.

Yi R, Qin Y, Macara IG, Cullen BR. 2003. Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. *Genes Dev* **17:** 3011–3016.

Yuo CY, Ares M Jr, Weiner AM. 1985. Sequences required for 3′ end formation of human U2 small nuclear RNA. *Cell* **42:** 193–202.