A real-time fluorescence assay for CPSF73, the nuclease for pre-mRNA 3′-end processing

PEDRO A. GUTIERREZ, KIRK BAUGHMAN, YADONG SUN, and LIANG TONG

Department of Biological Sciences, Columbia University, New York, New York 10027, USA

ABSTRACT

CPSF73 is the endonuclease that catalyzes the cleavage reaction for 3′-end processing of mRNA precursors (pre-mRNAs) in two distinct machineries, a canonical machinery for the majority of pre-mRNAs and a U7 snRNP (U7 machinery) for replication-dependent histone pre-mRNAs in animal cells. CPSF73 also possesses 5′–3′ exonuclease activity in the U7 machinery, degrading the downstream cleavage product after the endonucleolytic cleavage. Recent studies show that CPSF73 is a potential target for developing anticancer, antimalarial, and antiprotozoal drugs, spurring interest in identifying new small-molecule inhibitors against this enzyme. CPSF73 nuclease activity has so far been demonstrated using a gel-based end-point assay, using radiolabeled or fluorescently labeled RNA substrates. By taking advantage of unique properties of the U7 machinery, we have developed a novel, real-time fluorescence assay for the nuclease activity of CPSF73. This assay is facile and high-throughput, and should also be helpful for the discovery of new CPSF73 inhibitors.

Keywords: U7 snRNP; endonuclease; exonuclease; histone pre-mRNA 3′-end processing

INTRODUCTION

Eukaryotic messenger RNA precursors (pre-mRNAs) undergo processing at the 3′ end as a part of their maturation, and the resulting mRNAs are exported to the cytoplasm for translation into proteins. For the majority of pre-mRNAs, 3′-end processing consists of cleavage of the RNA at a specific location near the 3′ end followed by the addition of a poly(A) tail, and a large cleavage and polyadenylation machinery with many protein subunits is required for this canonical processing (Zhao et al. 1999; Shi and Manley 2015; Sun et al. 2020a). On the other hand, replication-dependent histone pre-mRNAs in animal cells are cleaved at the 3′ end but not polyadenylated, and a distinct machinery, the U7 snRNP, is required for this processing (Dominski and Marzluff 2007; Romeo and Schumperli 2016).

Remarkably, most of the cleavage module is shared between the two machineries. In the canonical machinery, the mammalian cleavage factor (mCF) (Chan et al. 2014; Schonenmann et al. 2014), with CPSF73, CPSF100 and symplekin as its subunits, catalyzes the cleavage reaction, with CPSF73 as the endonuclease (Mandel et al. 2006). In the U7 snRNP, also referred to as the U7 machinery here, the histone pre-mRNA cleavage complex (HCC) is equivalent to mCF, with CPSF73, CPSF100, symplekin and CstF64 as its subunits, and CPSF73 is the endonuclease for the cleavage reaction as well (Dominski et al. 2005; Dominski 2010). CPSF73 also has processive 5′–3′ exonuclease activity in the U7 machinery and can degrade the downstream cleavage product (DCP) to mononucleotides (Dominski et al. 2005; Yang et al. 2020).

The U7 machinery contains the U7 snRNA (Fig. 1A), the heptamer Sm ring, the protein FLASH (Yang et al. 2009, 2011), and HCC. The complex of U7 snRNA with the Sm ring is referred to as the U7 Sm core here. The amino-terminal segment of FLASH, which is required for histone pre-mRNA 3′-end processing, forms a coiled-coil dimer (Aik et al. 2017). Histone pre-mRNAs contain a conserved stem–loop structure upstream of the cleavage site (Fig. 1A), which is recognized by the stem–loop binding protein (SLBP) (Tan et al. 2013; Zhang et al. 2014). The histone downstream element (HDE) is recognized by the U7 machinery through base-pairing with the 5′ region of the U7 snRNA (Schaufele et al. 1986; Mowry and Steitz 1987; Bond et al. 1991) (the HDE-U7 duplex, Fig. 1A).

We recently reconstituted the human U7 Sm core (Bucholc et al. 2020), and subsequently the human U7 machinery (Sun et al. 2020b). This fully reconstituted machinery is active for cleaving model histone pre-mRNA...
substrates (Sun et al. 2020b) as well as degrading the DCP (Yang et al. 2020; Sun et al. 2021). The cryo-EM structure of this machinery in complex with a model substrate and SLBP reveals that the HDE-U7 duplex is surrounded by the amino-terminal domain of symplekin, CPSF100 and CPSF73 (Fig. 1B; Sun et al. 2020b). The pre-mRNA substrate is captured in the active site of CPSF73, providing molecular insights into its activity in histone pre-mRNA 3′-end processing.

Human CPSF73 was recently identified as the target of the compound JTE-607 (Fig. 2A; Kakegawa et al. 2019; Ross et al. 2020). It binds in the active site of CPSF73 after hydrolysis to its acid analog and inhibits the growth of acute myeloid leukemia (AML) and Ewing’s sarcoma cell lines (Ross et al. 2020), establishing CPSF73 as a potential target for novel anticancer drug discovery. In addition, inhibitors of Plasmodium falciparum CPSF73, such as the oxaborole compound AN3661 (Fig. 2B), are potent inhibitors of the growth of this and other parasites (Jacobs et al. 2011a,b; Palencia et al. 2017; Sonoiki et al. 2017; Swale et al. 2019). The oxaborole moiety is directly coordinated to the two zinc ions in the active site of CPSF73 (Swale et al. 2019), while JTE-607 has no contacts with the zinc ions (Fig. 2C; Ross et al. 2020).

The nuclease activity of CPSF73 has so far been demonstrated using gel-based end-point assays. These assays are cumbersome, time consuming and low throughput, requiring hour-long incubation of the reconstituted machinery (or nuclear extract) with a radiolabeled or fluorescently labeled model RNA substrate, separation of the products from the substrate by gel electrophoresis, and finally imaging the gel to visualize the labeled substrate and product(s). Taking advantage of unique properties of the U7 machinery, we have developed a novel, real-time fluorescence assay for the nuclease activity of CPSF73. This assay is facile and high-throughput, and should also be helpful for the discovery of new CPSF73 inhibitors.

**FIGURE 1.** Fluorescently labeled RNAs for the real-time CPSF73 activity assay. (A) Sequences of the model histone pre-mRNA substrate (H2a*) with a 3′ FAM label and the U7 snRNA with a 5′ TAMRA label. Nucleotides in the HDE that are changed to form Watson–Crick base pairs with the U7 snRNA are shown in lower case. (B) Overall structure of the core of the U7 machinery (Sun et al. 2020b). The proteins are shown as cartoons as well as a transparent surface in gray. The H2a* pre-mRNA is in orange, and U7 snRNA in green. The 5′ end of the U7 snRNA and 3′ end of H2a* are in close proximity. The red arrowhead indicates the active site of CPSF73. Panel B produced with PyMOL (www.pymol.org).

**FIGURE 2.** Inhibitors of CPSF73. (A) Chemical structure of JTE-607. The ester group that is hydrolyzed to produce the active acid analog is indicated with the red arrow. (B) Chemical structure of AN3661. The oxaborole moiety is indicated with the red arrow. (C) The binding modes of JTE-607 acid analog and AN3661 clash with the active, open form of human CPSF73. The structure of JTE-607 in complex with the closed form of human CPSF73 (Ross et al. 2020) is superimposed with that of the open form (Sun et al. 2020b), using the metallo-β-lactamase domain as the reference. Similarly, the structure of AN3661 in complex with the closed form of C. hominis CPSF73 (Swale et al. 2019) is superimposed with that of human CPSF73 open form. Only the inhibitors from the two structures are shown.
RESULTS

A real-time fluorescence assay for CPSF73

This real-time fluorescence assay is based on two unique properties of the U7 machinery revealed by the biochemical and structural studies: (i) The 3′ end of the pre-mRNA substrate is located close to the 5′ end of the U7 snRNA because of the HDE-U7 duplex (Fig. 1A,B); and (ii) the endo and exonuclease activities of CPSF73 in this machinery lead to the degradation of the DCP, destroying the HDE-U7 duplex. Therefore, if the model histone pre-mRNA substrate carries a FAM label at the 3′ end (the last nucleotide of the HDE-U7 duplex) and the U7 snRNA carries a TAMRA label at the 5′ end (the first nucleotide of the HDE-U7 duplex) (Fig. 1A), fluorescence from FAM will be quenched by TAMRA in this duplex due to their proximity and the overlap of their emission/absorption spectra. During the processing reaction, the endo and exonuclease activities of CPSF73 lead to the release of FAM from the pre-mRNA substrate and the duplex, and the resulting fluorescence signal from FAM can be readily monitored in real time, and with high-throughput, using a plate reader. This assay is similar in principle to a real-time assay developed for other exonucleases (Sinturel et al. 2009).

Characterization of the real-time fluorescence assay for CPSF73

We reconstituted a U7 Sm core in complex with FLASH using a U7 snRNA with a 5′ TAMRA label, and incubated it with SLBP and a pre-mRNA with a 3′ FAM label to allow the formation of the HDE-U7 duplex. The pre-mRNA (H2a∗) contains a modified HDE (Sun et al. 2020b), which enables it to form 15 consecutive Watson-Crick base pairs with the U7 snRNA (Fig. 1A). H2a∗ was kept substoichiometric to the U7 snRNA to ensure all the pre-mRNAs are in the duplex to minimize background signals from FAM. The reaction was initiated by the addition of HCC.

After optimizing the parameters of the reaction and the plate reader, we identified conditions for the concentrations of U7 Sm core, H2a∗ and HCC that produced linear reaction progress curves for 30 min or longer at 30°C and were readily able to determine the reaction velocities (Fig. 3A). The reaction did not proceed without HCC, confirming that the release of FAM is dependent on the activity CPSF73. As shown earlier (Sun et al. 2020b), SLBP is not essential for the activity with this substrate, because of the enhanced base-pairing in the HDE-U7 duplex, although it did have a small stimulatory effect on the reaction. We supplemented additional FLASH in another reaction, in case not all the FLASH binding sites in the U7 Sm core were occupied, but the extra FLASH did not have a significant effect on the velocity.

We next determined the activity of CPSF73 as a function of H2a∗ concentration. The concentration of H2a∗ was varied between 0 and 40 nM, whereas the concentrations of U7 Sm core-FLASH complex and HCC were kept constant. The concentration of SLBP was varied together with that of H2a∗, being kept at a constant molar ratio (1.3:1) relative to the H2a∗ concentration. The reaction velocities showed a nearly linear increase as a function of H2a∗ concentration up to 40 nM (Fig. 3B), suggesting that 40 nM H2a∗ substrate is not sufficient to saturate the enzyme under the assay condition tested.

Assessing the inhibitory activities of JTE-607 and AN3661

To characterize the inhibitory activities of JTE-607 and AN3661, we varied their concentrations from 0.1 to 300 µM, while the H2a∗ concentration was fixed at 40 nM (Fig. 3B). The stock solutions of the compounds were at 100 mM in 100% DMSO, and therefore the highest
concentration of DMSO in the reactions was 0.3%. We carried out a control reaction in the presence of 0.3% DMSO but no compound, which showed no inhibitory activity. Both JTE-607 acid analog and AN3661 demonstrated inhibition of CPSF73 in this assay (Fig. 4A,B). The inhibition data were fitted to obtain apparent IC₅₀ values of 134 µM for JTE-607 and 3.6 µM for AN3661 under the conditions tested.

We also carried out gel-based end-point assays as an independent determination of the inhibitory potency of the two compounds against the nuclease activity of human CPSF73. The cleavage reaction was run at 30°C for 60 min in the presence of increasing concentrations of the inhibitor. A doubly fluorescently labeled H2a∗ substrate, with 5′ TAMRA and 3′ FAM labels, was used in these reactions, which allowed both the endo and exonuclease activities of CPSF73 to be monitored directly (Fig. 5A; Sun et al. 2021). While the 5′ product is stable, the 3′ product is degraded to mononucleotides by the exonuclease activity of CPSF73 (Yang et al. 2020; Sun et al. 2021). No intermediates are observed for the degradation of the 3′ product, consistent with a processive exonuclease activity for CPSF73. An IC₅₀ value was determined based on the amount of the 5′ product, with a TAMRA label, as a function of the inhibitor concentration (Fig. 5B,C). The IC₅₀ was 54 µM for JTE-607 acid analog and 0.85 µM for AN3661 based on these assays.

DISCUSSION

We have established a real-time fluorescence assay for the nuclease activities of CPSF73 based on unique properties of the U7 machinery. The assay is facile and high-throughput, as the plate reader format allows many different conditions (or inhibitors) to be assayed at the same time. A readout of 30 min is typically sufficient for the reactions, and shorter readout times (such as 10 min) may be adequate as well. These properties of the assay represent a substantial improvement over the current gel-based end-point assays. On the other hand, such a real-time assay is not possible for the canonical machinery, unless a method is found to set up the quenching pair of fluorophores for that RNA substrate. The cleavage module is equivalent between the two machineries, and the inhibitory activity of compounds against the canonical machinery is likely similar to that measured against the U7 machinery.

The FAM-labeled pre-mRNA substrate is kept sub-stoichiometric relative to the TAMRA-labeled U7 snRNA to minimize the FAM signal at the beginning of the reactions. In addition, excess FAM-labeled pre-mRNA could reduce the measured velocity of the reaction, because a FAM fluorophore that is released by the U7 machinery is replaced by a FAM-labeled RNA from the excess pool, which base-pairs to the U7 snRNA and becomes quenched, thereby producing no net change in fluorescence. Swapping the TAMRA and FAM fluorophores between the two RNAs would not solve this problem either. Much higher concentrations of both pre-mRNA and U7 Sm core would be needed to saturate CPSF73 and produce a Michaelis–Menten curve (Fig. 3B). For inhibition assays, it might be advantageous to purify a complex of U7 Sm core—FLASH with the H2a∗ substrate by gel filtration, and use that sample in all the reactions.

The FAM fluorescence signal is observed only after the DCP is degraded by the exonuclease activity of CPSF73. We did not observe any delay in the appearance of the FAM signal in our real-time assays, indicating that DCP degradation happens soon after the endonucleolytic cleavage. It is likely that the DCP does not dissociate from the U7 snRNA after the cleavage reaction, although the U7 machinery can degrade exogenously added DCP (Yang et al. 2020; Sun et al. 2021). The degradation of the DCP...
would allow the U7 machinery to be recycled to process the next pre-mRNA substrate.

The IC50 values determined from the real-time and endpoint activity assays for JTE-607 acid analog are comparable to each other, but they are much higher than the Kd value reported earlier (370 nM) (Ross et al. 2020), which could be explained by the following consideration. The Kd value was measured for JTE-607 acid analog against the catalytic module of CPSF73 (Ross et al. 2020), which exists in a closed form (Mandel et al. 2006), as also confirmed by the structure of JTE-607 in complex with the catalytic module (Ross et al. 2020). On the other hand, CPSF73 undergoes a large conformational change to an open form when it binds to the pre-mRNA substrate for catalysis (Sun et al. 2020b). The binding mode of JTE-607 is incompatible with the open form, for example it has extensive clashes with Phe241 in the open form (Fig. 2C). JTE-607 also has some steric clashes with this side chain in the closed form, and there is a conformational change for this residue in the JTE-607 complex that ameliorates these clashes. A much larger conformational change for this residue would be needed to accommodate the JTE-607 acid analog in the open form. The reported Kd value is a measure of the affinity of JTE-607 for the closed form of CPSF73, while the IC50 values measured here reflect the inhibitory potency of JTE-607 against the catalytic activity of CPSF73, which occurs in the open form. While other factors may also contribute to the observed differences in the inhibitory potencies, this analysis suggests that the open form of CPSF73 would be a distinct target for developing inhibitors against it. It may also be possible that JTE-607 inhibits CPSF73 partly by preventing the conformational transition from the closed to the open form.

The IC50 values determined from the real-time and endpoint activity assays for AN3661 are similar to each other as well. AN3661 also binds to the closed form of C. hominis CPSF73 (Swale et al. 2019). It clashes with Phe241 in the open form of human CPSF73 (Fig. 2C), although this clash is not as extensive as that for JTE-607 as it only involves the carboxylate group of AN3661. This group could reorient itself to avoid the contact, as it is connected to the rest of the compound through an ethylene linker. The IC50 of
this compound against wild-type *P. falciparum* is 31 nM (Sonoki et al. 2017), roughly 70-fold lower than the IC$_{50}$ measured here against human CPSF73. There is a conformational difference near Phe264 (equivalent to Phe241 in human CPSF73) in the closed form of *C. hominis* CPSF73 compared to human CPSF73, which preorganizes the binding site for AN3661 and could be the molecular basis for its higher activity against protozoal CPSF73.

Studies with the reconstituted U7 machinery show that U7 Sm core, FLASH, and HCC are sufficient for activity in vitro. Many other factors have been identified that may be important for histone pre-mRNA 3′-end processing in vivo, such as ZFP100 (Wagner and Marzluff 2006), nuclear cap binding complex (Narita et al. 2007), CF l,68 (Ruepp et al. 2010), and FUS/TLS (Racynska et al. 2015). This real-time assay should also allow any direct contributions of these factors to the cleavage reaction to be studied quantitatively in vitro.

**MATERIALS AND METHODS**

**Reconstitution of the U7 machinery**

U7 snRNA with 5′ TAMRA label and H2a∗ pre-mRNA with 3′ FAM label were purchased from IDT (Integrated DNA Technologies). The preparation of all the recombinant proteins and reconstitution of the U7 machinery followed protocols described earlier (Bucholc et al. 2020; Sun et al. 2020b, 2021). A U7 Sm core with TAMRA labeled U7 snRNA, in complex with FLASH (Aik et al. 2017), was reconstituted from its components. HCC and SLBP were expressed and purified from baculovirus-infected insect cells (Sun et al. 2020b; Zhang et al. 2020).

**Real-time fluorescence assays**

The buffer for all the reactions contained 15 mM HEPES (pH 8.0), 75 mM KCl, 15% (v/v) glycerol, and 20 mM EDTA. U7 Sm core-FLASH complex, HCC, H2a∗, SLBP, and inhibitor if present were mixed and incubated at 30°C for 60 min. The proteins were at 0.2 μM and the H2a∗ was at 0.1 μM concentration. The substrate RNA and products were separated using a 15% (w/v) denaturing polyacrylamide gel (1× TBE buffer, 8 M urea, 280 V, 20 min). The gel was imaged with the ChemiDoc MP imaging system (Bio-Rad) based on TAMRA fluorescence. The amount of the 5′ cleavage product was estimated from the gel image with the program ImageJ (Schneider et al. 2012).

**ACKNOWLEDGMENTS**

We thank Shi-Xian Deng for converting JTE-607 to its acid analog, funded by the Raymond and Beverley Sackler Center for Research at Convergence of Disciplines at Columbia University Medical Center. This research is supported by National Institutes of Health (NIH) grant R35GM118093 (to L.T.).

Author contributions: P.A.G. carried out the real-time assays. K.B. carried out gel-based assays and initial attempts with the real-time assay. Y.S. produced all the protein samples for the assays. L.T. designed the experiments and supervised the research. L.T. wrote the paper, and all authors commented on the paper.

Received March 25, 2021; accepted June 17, 2021.

**REFERENCES**

Aik WS, Lin MH, Tan D, Tripathy A, Marzluff WF, Dominski Z, Chou C-Y, Tong L. 2017. The hunt for the 3′-end of histone pre-mRNA: getting closer to the end. *Genes Dev* 31:1709–1722. doi:10.1101/gad.9.5.1709

Bucholc K, Aik WS, Yang X-C, Wang K, Zhou ZH, Dadlez M, Marzluff WF, Tong L, Dominski Z. 2020. Composition and processing activity of a semi-recombinant holo U7 snRNP. *Nucleic Acids Res* 48:1508–1530. doi:10.1093/nar/gkz1148

Chan SL, Huppersz I, Yao C, Weng L, Moresco JJ, YatesIII JR, Ule J, Manley JL, Shi Y. 2014. CPSF30 and Wdr33 directly bind to AAUAAA in mammalian mRNA 3′ processing. *Genes Dev* 28:2370–2380. doi:10.1101/gad.250993.114

Dominski Z. 2010. The hunt for the 3′ endonuclease. *Wiley Interdiscip Rev RNA* 1:325–340. doi:10.1002/wrna.33

Dominski Z, Marzlinke MF, 2007. Formation of the 3′ end of histone mRNA: getting closer to the end. *Gene* 373:373–390. doi:10.1016/j.gene.2007.04.021

Dominski Z, Yang X-C, Marzluff WF. 2005. The polyadenylation factor CPSF-73 is involved in histone-pre-mRNA processing. *Cell* 123:37–48. doi:10.1016/j.cell.2005.08.002

Jacobs RT, Nare B, Wrung SA, Orr MD, Chen D, Sligar JM, Jenks MX, Noe RA, Bowling TS, Mercer LT, et al. 2011a. SCYX-7158, an orally-active benzoxaborole for the treatment of stage 2 human
African trypanosomiasis. PLoS Negl Trop Dis 5: e1151. doi:10.1371/journal.pntd.0001151

Jacobs RT, Plattner JJ, Keenan M. 2011b. Boron-based drugs as antiprotozoals. Curr Opin Infect Dis 24: 586–592. doi:10.1097/QCO.0b013e32834e630e

Kakegawa J, Sakane N, Suzuki K, Yoshida T. 2019. JTE-607, a multiple cytokine production inhibitor, targets CPSF3 and inhibits pre-mRNA processing. Biochem Biophys Res Commun 518: 32–37. doi:10.1016/j.bbrc.2019.08.004

Mandel CR, Kaneko S, Zhang H, Gebauer D, Vethantham V, Palencia A, Bougdour A, Brenier-Pinchart MP, Touquet B, Bertini RL, Schaufele F, Gilmartin GM, Bannwarth W, Birnstiel ML. 1986. 'Cleoprotein and participates in 3′ end processing of animal nuclear RNA.

Barabino SM, Schümperli D. 2010. The 68 kDa subunit of mammalian histone mRNAs.

Yamaguchi Y, Handa H. 2007. NELF interacts with CBC and compensatory mutations suggest that base-pairing with a small nuclear RNA is required to form the 3′ end of H3 messenger RNA. Nature 476: 763–7650. doi:10.1038/nature05363

Wagner EJ, Marzluff WF. 2006. ZFP100, a component of the active U7 snRNP limiting for 3′ end processing of histone pre-mRNAs, is essential for 3′ exonucleolytic degradation of the downstream cleavage product. Mol Cell 21: 6702–6712. doi:10.1128/MCB.00391-06

Yang X-C, Burch BD, Yan Y, Marzluff WF, Dominski Z. 2009. FLASH, a proapoptotic protein involved in activation of caspase-8, is essential for 3′ end processing of histone pre-mRNAs. Mol Cell 36: 267–278. doi:10.1016/j.molcel.2009.08.016

Yang X-C, Xu B, Sabath I, Kunduru L, Burch BD, Marzluff WF, Dominski Z. 2011. FLASH is required for the endonucleolytic cleavage of histone pre-mRNAs but is dispensable for the 5′ exonucleolytic degradation of the downstream cleavage product. Mol Cell 31: 1492–1502. doi:10.1128/MCB.00979-10

Zhang J, Tan D, DeRose EF, Perera L, Dominski Z, Marzluff WF, Tong L, Tanaka Hall TM. 2014. Molecules for the regulation of histone mRNA stem-loop-binding protein by phosphorylation. Proc Natl Acad Sci 111: E2937–E2946. doi:10.1073/pnas.1406381111

Zhang Y, Sun Y, Shi Y, Walz T, Tong L. 2020. Structural insights into the human pre-mRNA 3′-end processing machinery. Mol Cell 77: 800–809. doi:10.1016/j.molcel.2019.11.005

Zhao J, Hyman L, Moore CL. 1999. Formation of mRNA 3′ ends in eukaryotes: mechanism, regulation, and interrelationships with other steps in mRNA synthesis. Microbiol Mol Biol Rev 63: 405–445. doi:10.1128/MMBR.63.2.405-445.1999