RNase H-based analysis of synthetic mRNA 5’ cap incorporation

S. HONG CHAN, JOSEPH M. WHIPPLE, NAN DAI, THERESA M. KELLEY, KATHRYN WITHERS, GEORGE TZERTZINIS, IVAN R. CORRÊA JR., and G. BRETT ROBB

New England Biolabs, Ipswich, Massachusetts 01938, USA

ABSTRACT

Advances in mRNA synthesis and lipid nanoparticles technologies have helped make mRNA therapeutics and vaccines a reality. The 5’ cap structure is a crucial modification required to functionalize synthetic mRNA for efficient protein translation in vivo and evasion of cellular innate immune responses. The extent of 5’ cap incorporation is one of the critical quality attributes in mRNA manufacturing. RNA cap analysis involves multiple steps: generation of predefined short fragments from the 5’ end of the kilobase-long synthetic mRNA molecules using RNase H, a ribozyme or a DNAzyme, enrichment of the 5’ cleavage products, and LC-MS intact mass analysis. In this paper, we describe (1) a framework to design site-specific RNA cleavage using RNase H; (2) a method to fluorescently label the RNase H cleavage fragments for more accessible readout methods such as gel electrophoresis or high-throughput capillary electrophoresis; (3) a simplified method for post-RNase H purification using desthiobiotinylated oligonucleotides and streptavidin magnetic beads followed by elution using water. By providing a design framework for RNase H-based RNA 5’ cap analysis using less resource-intensive analytical methods, we hope to make RNA cap analysis more accessible to the scientific community.

Keywords: synthetic mRNA; 5’ cap; RNA capping; RNA cap analysis; RNase H; LC-MS

INTRODUCTION

mRNA synthesis and nano lipid particles technologies have advanced to a stage where mRNA therapeutics and vaccines have become a reality with a global public health impact. The relatively short development and manufacturing time have been the main attributes propelling the rapid deployment of effective mRNA vaccines in a global health emergency such as COVID-19.

In eukaryotic cells, the 5’ end of messenger RNA (mRNA) is characterized by a cap structure that plays key functional roles in many biological processes (Gonatopoulos-Pournatzis and Cowling 2014). Most relevant to mRNA as therapeutics or vaccines is its roles in efficient protein translation (Goodfellow and Roberts 2008; Maquat et al. 2010; Gonatopoulos-Pournatzis and Cowling 2014), evading innate immune surveillance (Kumar et al. 2014; Hyde and Diamond 2015; Devarkar et al. 2016), and regulation of 5’-mediated decay (Floor et al. 2012; Mugridge et al. 2018). Hence, a 5’ cap structure is crucial for the efficacy and safety of mRNA therapeutics (Ramaswamy et al. 2017) and vaccines (Pardi et al. 2018; VanBlargan et al. 2018).

The 5’ cap consists of an N7-methylated guanosine linked to the 5’ end of RNA through a 5’ to 5’ triphosphate group. In metazoans, the m7GpppN cap (Cap-0) is further modified by a cap-specific 2’-O-methyltransferase (MTase) to the m7GpppNm structure (Cap-1) (Fig. 1).

Currently, the cap structure can be incorporated into an in vitro transcript by two means: (1) incorporation of a chemically synthesized cap structure such as the anti-reverse cap analog (ARCA) or capped dinucleotides (Henderson et al. 2021) during transcription, or (2) post-transcriptional capping using a mRNA capping enzyme. These approaches have well-established methods that are readily adaptable and scalable (Fuchs et al. 2016; Vaidyanathan et al. 2018). It is, however, less trivial to verify the identity of the cap and evaluate the extent of capping on kilobase-long synthetic mRNA molecules. One of the most common methods for evaluating the efficiency of cap incorporation is a cell-based functional assay that measures the level of protein expression from the synthetic mRNA in vivo through immunometric or enzymatic reactions. These readouts can provide information about the efficacy of the synthetic mRNA translation, but do not directly reflect the extent of cap incorporation of a synthetic mRNA preparation.
To analyze by mobility-based methods such as gel or capillary electrophoresis. As with other hybridization-based approaches, efficient RNase H cleavage depends on factors such as overcoming the secondary structures of the analyte molecules and annealing of DNA–RNA chimera to the predefined cleavage site under given reaction conditions. These factors are often not well-defined, and pseudouridine and derivatives substitutions could change the base-pairing propensity of synthetic RNA. Hence, a framework for designing precise RNase H cleavage is desired.

Previously, it has been shown that by using a biotinylated DNA–RNA chimera oligo, the 5′ RNase H cleavage fragments could be readily isolated for qualitative and quantitative analyses using LC-MS (Beverly et al. 2016). However, there has been insufficient guidance on the DNA–RNA chimera design to achieve uniform RNase H cleavage, and the readout method is limited to resource-intensive LC-MS. Recently, Vlatkovic et al. (2022) showed that custom designed ribozymes can generate uniform 5′ fragments for cap analysis and that the 5′ cleavage fragments can be purified by silica-based spin columns. In this paper, by applying a DNA–RNA chimera selection framework that screens for RNase H specificity empirically, we showed that RNase H can generate highly uniform cleavage at predefined sites. We adopted and simplified an affinity-based method to enrich the 5′ cleavage fragment that requires lower RNA input.

In the cell, RNase H (RNase H1) is an endonuclease that removes the RNA primers from the Okazaki fragments of the replicating DNA and processes R-loops to modulate R-loop-mediated biological processes, such as gene expression, DNA replication, and DNA and histone modifications (Huang et al. 1994; Brocoli et al. 2004; Parajuli et al. 2017). It has been shown that Escherichia coli RNase H can be constrained to cleaving ssRNA at specific sites in vitro using a DNA–RNA chimera. However, E. coli RNase H may also cleave one or more nucleotides away from the 5′ or 3′ of the target site, giving rise to multiple cleavage products that differ by 1 nt (Lapham and Crothers 1996; Lapham et al. 1997; Yu et al. 1997). For the analysis of RNA 5′ capping, which is essentially the addition of one nucleotide at the 5′ triphosphate group, multiple cleavage products of a few nucleotides difference in length can make it impossible to analyze by mobility-based methods such as gel or capillary electrophoresis. As with other hybridization-based approaches, efficient RNase H cleavage depends on factors such as overcoming the secondary structures of the analyte RNA molecules and annealing of DNA–RNA chimera to the predefined cleavage site under given reaction conditions. These factors are often not well-defined, and pseudouridine and derivatives substitutions could change the base-pairing propensity of synthetic RNA. Hence, a framework for designing precise RNase H cleavage is desired.

Recently, Vlatkovic et al. (2022) showed that custom designed ribozymes can generate uniform 5′ fragments for cap analysis and that the 5′ cleavage fragments can be purified by silica-based spin columns. In this paper, by applying a DNA–RNA chimera selection framework that screens for RNase H specificity empirically, we showed that RNase H can generate highly uniform cleavage at predefined sites. We adopted and simplified an affinity-based method to enrich the 5′ cleavage fragment that requires lower RNA input. We further showed that the 3′ end of the RNase H cleavage fragments can be fluorescently labeled by the fill-in activity of the Klenow fragment. Because the fill-in activity is
dependent on the complementarity of the incoming fluorescently labeled dNTPs to the DNA–RNA chimera, the labeling step further constrains the size of the fluorescently labeled RNase H cleavage product to effectively eliminate nontarget cleavage products from mobility-based methods such as gel or capillary electrophoresis. Together with the DNAzyme- and ribozyme-based methods (Cairns et al. 2003; Vlatkovic et al. 2022), we hope that our methods (Fig. 2) can help make RNA cap analysis more accessible to the scientific community.

RESULTS AND DISCUSSION

DNA–RNA chimera-guided RNase H cleavage at preselected site

Escherichia coli RNase H has been shown to cleave ssRNA at predefined sites in vitro by hybridizing a DNA–RNA chimera to the target synthetic RNA, generating an RNA/DNA duplex akin to its physiological substrate. Two major cleavage sites, the adjacent phosphodiester bonds 5′ and 3′ to the ribonucleotide hybridized to the 5′ deoxynucleotide of the DNA–RNA chimera have been reported (Fig. 3A; Lapham and Crothers 1996; Lapham et al. 1997; Yu et al. 1997).

To try to understand whether RNase H can be restrained to make more uniform cuts, we first compared two commercially available RNase H enzymes (from E. coli and from Thermus thermophilus) for their efficiency and specificity in cleavage on a synthetic RNA oligonucleotide containing the 5′ sequence of a firefly luciferase FLuc transcript (synFLuc-AC). To maximize the availability of cleavage sites, we designed the DNA–RNA chimera (Targeting Oligos; TOs) to direct RNase H to cleave within a loop region of the predicted secondary structure (Fig. 3B; Supplemental Fig. 1). The substrate synFLuc-AC contains a 5′ FAM group so that the cleavage products can be analyzed using urea-PAGE electrophoresis and quantified using a flatbed laser scanner.

As shown in Figure 3C, using the same quantity of enzyme under identical reaction conditions (37°C for 1 h), E. coli and T. thermophilus (Tth) RNase H achieved similar levels of cleavage (98.3 ± 0.4% vs. 96.6 ± 0.2%, respectively). However, Tth RNase H generated more 24 nt cleavage product than the E. coli enzyme (87.6 ± 1.0% vs. 74.1 ± 3.3%, respectively). Hence, Tth RNase H was used in the following study.

FIGURE 2. A general scheme of RNase H-based RNA cap analysis. A DNA–RNA or DNA-2′-O-methyl-ribonucleotide chimera is designed to be complementary to part of the 5′ end of the target RNA molecule such that the chimera stays annealed to the cleavage fragment after RNase H cleavage. The chimera (called targeting oligo or TO in this paper) contains a 3′-deoxynucleotide group. After denaturation and annealing, RNase H cleaves at a predefined site within the RNA-TO duplex and generates a one-base recessive end at the 3′ end of the cleaved RNA. Because RNase H cleavage results in a 3′ hydroxyl group (24), this recessive 3′ end can be filled in with a fluorescently labeled deoxynucleotide using the Klenow fragment of DNA polymerase I. The fluorescently labeled 5′ cleavage fragment can be analyzed by denaturing PAGE directly without enrichment. The 5′ duplex cleavage fragment can be enriched using streptavidin magnetic beads. The enriched RNase H cleavage products can be analyzed by LC-MS or capillary electrophoresis (if filled in with a fluorescent deoxynucleotide).
Using the 5′ FAM-labeled synthetic RNA oligonucleotide synFLuc-AC as a surrogate, we next sampled several potential cleavage sites using different TOs. Cleavage reactions were done in triplicate and cleavage events were evaluated using urea PAGE and LC-MS intact mass analysis (see the schematic in Fig. 4A). Urea PAGE analysis showed that targeted RNase H cleavage efficiency was achieved with all TOs (Fig. 4B). LC-MS analysis showed that 86% of the fragments generated from the FLucTO-25 were products of cleavage at the –1 A|C site, 15% of the cleavage events took place at the –2 (A|A), and <1% at the –3 (U|A) site. Further extending the TO by 3 nt (FLucTO-28), the major cleavage site remained at the U|U site now at –2 position (98%). The –1 (U|U) site was cleaved at a low frequency (1%).

We next used a 3′ desthiobiotinylated version of FLucTO-25 (TO-1) to guide Tth RNase H to cleave the FLuc in vitro transcripts (1.7 kb). FLucTO-27 and FLucTO-28 that generated 87% to 98% of a specific cleavage product were not chosen because of the presence of uridine at the cleavage sites (see below for the effect of pseudouridine substitution in RNase H cleavage). After cleavage by Tth RNase H, the 5′ cleavage product was purified using the water elution method described below and analyzed by LC-MS intact mass analysis. Consistent with the findings with the RNA oligonucleotide surrogate, the median cleavage frequency at the –1 A|C site was 91%, with 8% at the –2 (A|A) site, and 2% at the +1 (C|U) site (Fig. 5B), from eight replicated experiments (Fig. 5). Therefore, highly uniform RNase H cleavage can be achieved by screening a series of TO on an RNA oligo surrogate containing the 5′ sequence of the target synthetic mRNA.

**Efficient enrichment of RNase H cleavage products**

It has been reported recently that ribozyme-cleaved 5′ cleavage fragments of in vitro transcript could be enriched for LC-MS analysis using silica-based spin columns. The spin column format, however, requires 20–60 µg RNA input (Vlatkovic et al. 2022). The use of RNase H and biotinylated targeting oligos allows for affinity enrichment (Beverly et al. 2016) that could potentially decrease the input RNA requirement. The enrichment method described previously by Beverly et al. (2016), however, requires elution of the 5′ cleavage fragment from streptavidin magnetic beads by heating in a methanol solution. We sought to develop an enrichment method that allows for a lower RNA input and eliminates the use of methanol. To this end, we adopted **FIGURE 3. Directed RNA cleavage with RNase H.** (A) Previously reported RNase H cleavage sites (17–19). (B) Synthetic RNA oligonucleotide containing the first 33 nt of an artificial FLuc transcript (synFLuc-AC) and targeting oligos FLucTO-25 designed based on A. Inverted triangles represent RNase H cleavage sites reported in the literature (A) and in this study (B). Deoxynucleotides in TOs are colored in blue; ribonucleotides are colored in green. The predicted loop region of the synthetic RNA oligonucleotides based on an RNA folding algorithm (RNAFold, University of Vienna) is indicated. (C) RNA cleavage using *E. coli* or *Thermus thermophilus* RNase H at 37°C. The cleavage efficiency of both enzymes was similar but Tth RNase H generates more uniform cuts than the *E. coli* enzyme.
the Beverly method with a few improvements. Using an artificial firefly luciferase (FLuc) in vitro transcript as a test case, the long 3' RNase H cleavage product and the uncleaved full-length transcript were first separated from the smaller 5' cleavage products and TO (TO-1; Fig. 6) by size selection and then affinity enrichment. Briefly, the RNase H treated sample (or the corresponding Klenow fill-in products; see below) were added to the NEBNext Sample Purification Magnetic Beads. The unbound fraction containing smaller RNAs (ub1, Fig. 6) was added to fresh beads for further enrichment. The subsequent unbound fraction (ub2, Fig. 6), which contained only a small fraction of the larger RNA molecules, was then subjected to affinity enrichment using streptavidin magnetic beads. Two elution methods were compared: biotin in aqueous solution versus water (Holmberg et al. 2005). The RNA-bound streptavidin magnetic beads were divided into two equal fractions. The first fraction was washed with a standard streptavidin-bead washing buffer (containing 1 M NaCl), and elution was carried out by incubating the clarified beads with a 0.1 M biotin solution at 37°C for 1 h. The second fraction of RNA-bound streptavidin magnetic beads was washed with a low salt buffer (containing 60 mM NaCl) and elution was carried out by incubating the clarified beads with nuclease-free water at 65°C for 5 min. Elution with either water or biotin solution recovered similar amounts of the TO-1 and the RNase H 5' cleavage products as analyzed by urea PAGE (Fig. 6) and LC-MS (Supplemental Fig. 2). For simplicity, a low salt wash buffer and elution with water was chosen to enrich desthiobiotin-tagged TO-RNase H cleavage product duplexes in subsequent experiments.

Upon further investigation, we found that although a small amount of higher molecular weight RNA was present when the size-selection step was not performed, the data quality of the LC-MS analysis was not affected (data not shown). Hence, the size-selection step can be considered optional.

We used the single step post-RNase H affinity enrichment method to process enzymatically capped Cap-1 CLuc or FLuc transcripts. The deconvoluted mass spectrums of the enriched 5' cleavage fragments are shown in Figure 7. Using this protocol, 5 pmol of the 1.8 kb CLuc, or 1.7 kb FLuc in vitro transcript (2.9 µg) was sufficient to generate high quality LC-MS data for quantitative analysis.
products using Klenow fragment was 8%, 91%, and 1% at (A|A), (A|C), and (C|U), respectively. Cleavage product using LC-MS. Median of cleavage frequencies of an artificial 5′-UTR and the corresponding targeting oligo TO-1. The targeting oligo contains six deoxynucleotides (blue) at the 5′ end followed by 19 ribonucleotides (green) and a desthiobiotin (DTB) group at the 3′ end. The size of the RNase H cleavage products is shown. (B) Frequency of cleavage events expressed as percentage of detected cleavage product using LC-MS. Median of cleavage frequencies was 8%, 91%, and 1% at (A/A), (A/C), and (C/U), respectively.

**Fluorescent labeling of RNase H 5′ cleavage products using Klenow fragment**

The length of TO allows the 5′ cleavage fragment to stay annealed to the TO as a duplex. From the TO selection, we found that the major cleavage product of FLucTO-25 (and its 3′ desthiobiotinylated derivative TO-1) produced a 3′ recessive end (Figs. 4B, 5). Because RNase H cleavage results in a hydroxyl group at the 3′ end (Lima and Crooke 1997), this 3′ recessive end is amenable to the fill-in activity of the Klenow fragment of *E. coli* DNA polymerase I (Huang and Szostak 1996; Sandin et al. 2009). We took advantage of this discovery to label the 3′ end of the 5′ cleavage product using a FAM-labeled dNTP complementary to the 5′ nucleotide of the TO. Because the fill-in activity is dependent on the complementarity of the incoming fluorescently labeled dNTPs to the DNA–RNA chimera, the labeling step further constrains the size of the fluorescently labeled RNase H cleavage product to effectively eliminate the nontarget cleavage products from the analysis. The filled-in 5′ fragments can then be analyzed using urea-PAGE or capillary electrophoresis (Greenough et al. 2016; Wulf et al. 2019) in addition to LC-MS.

As an example, a FLuc transcript capped using VCE was analyzed by treatment with RNase H followed by Klenow fill-in using FAM-labeled dCTP, as described in Materials and Methods (Fig. 8A). Figure 8B showed that Klenow fill-in enabled visualization and quantification of the extent of 5′ cap incorporation using urea PAGE. Compared to total RNA staining (SYBR Gold), the fluorescent signal offers a higher signal-to-noise ratio and eliminates the interference from the TO in data analysis. A small quantity of +1 nt cleavage product is visible in the FAM-signal image (Fig. 8B, right panel). LC-MS analysis indicated that the +1 nt is probably derived from RNA polymerase slippage at the GGG trinucleotide at the 5′ end of the FLuc transcript (data not shown) (Pleiss et al. 1998).

Urea-PAGE is a quick and easily implementable readout for 5′ RNA capping, but it does not resolve 5′ triphosphate (ppp-) from 5′ diphosphate (pp-), or m7G cap (Cap-0; m′Gppp-) from unmethylated G cap (Gppp-). Rather, 5′ triphosphate and 5′ diphosphate co-migrate as uncapped RNA, while m7G cap and unmethylated G cap co-migrate as capped RNA. An affinity-based denaturing PAGE method that can resolve m7G cap (Cap-0) from unmethylated G cap (Gppp-) and uncapped (pp- and pp-) RNA has been described (Matts et al. 2014). Capillary electrophoresis, on the other hand, resolves nucleic acids by size and charge. We showed that capillary electrophoresis can be adapted to resolve and quantify the four 5′ end structures related to enzymatic RNA capping on short RNA oligonucleotides (Supplemental Fig. 3; Wulf et al. 2019). To show that the capillary electrophoresis platform is applicable to cap analysis of long IVT RNA molecules, the FLuc transcript partially capped by VCE was processed by the RNase H/Klenow fill-in method and enriched as described above (to remove most of the free FAM-dCTP). Capillary electrophoresis of the processed samples allowed quantification of 5′ ppp-, pp-, Gppp-, and m′Gppp- (Fig. 8C).

Capillary electrophoresis can be routinely done in a 96-well format using an Applied Biosystems Genetic Analyzer (see Materials and Methods.) The high-throughput capability and resolution of enzymatic capping intermediates is of great interest for method development and quality control in platforms that utilize enzymatic capping for mRNA production.

To further validate the Klenow fill-in method, a partially capped FLuc preparation was subjected to the RNase H/Klenow fill-in processing and analyzed by urea-PAGE, capillary electrophoresis, and LC-MS (Fig. 8D). We found that cap analysis using capillary electrophoresis and LC-MS intact mass analysis produced highly comparable results. Although urea-PAGE cannot individually resolve m′Gppp- from Gppp-, or pp- from ppp-, the percentage of capped (m′Gppp- and Gppp-) and uncapped (pp- and ppp-) agreed well with the capillary electrophoresis and LC-MS results.

**The effect of uridine modifications on RNase H cleavage**

Uridine modifications (e.g., pseudouridine, 1-methylpseudoouridine and 5-methoxyuridine) have been shown to help
reduce cellular innate immune response and improve protein translation of synthetic mRNA (Karikó et al. 2008, 2012; Svitkin et al. 2017). Modified uridine has been widely used in mRNA vaccines, such as the FDA-approved SARS-CoV-2 vaccines manufactured by Pfizer/BioNTech and Moderna (Nance and Meier 2021). Pseudouridine is known to exhibit different base-pairing properties compared to uridine (Kierzek et al. 2014; Deb et al. 2019). Such differences can affect annealing of targeting oligos and/or RNase H cleavage.

To investigate the effect of pseudouridine (Ψ) on RNase H cleavage site preferences, a Cypridina luciferase (CLuc) transcript (1.8 kb) was transcribed in vitro in the presence of either UTP or ΨTP. Targeting oligo CLucTO-26 was used such that the major RNase H cleavage sites were centered around a uridine residue (Supplemental Fig. 5). Upon Tth RNase H cleavage, the 5′ cleavage product of the uridine and the Ψ-substituted CLuc transcripts were subjected to the enrichment and analysis as described above. Two major cleavage events were observed with the unsubstituted CLuc transcript: 73% of the products resulted from cleavage 5′ of the uridine (C|UC) with a length of 25 nt, while 27% resulted from cleavage 3′ of the uridine (CU|C) with a length of 26 nt (Fig. 9). When uridine residues were substituted with pseudouridine, cleavage events 5′ of the pseudouridine (C|ΨC) dropped to a median frequency of 34%, with 66% of cleavage taking place 3′ of the pseudouridine (CΨ|C). Hence, pseudouridine substitution can alter the cleavage specificity of RNase H through mechanisms possibly related to base-pairing with the TO and/or RNase H substrate binding.

Messenger RNA, as a single-stranded nucleic acid molecule, can adopt multiple stable conformations depending on formulation and reaction conditions. These stable conformers can affect RNase H, ribozyme or DNAzyme cleavage efficiency and specificity through factors such as substrate annealing and substrate-enzyme interactions. While these factors are still poorly understood, we showed that highly uniform RNase H cleavage can be achieved by systematic screening of Targeting Oligos (DNA–RNA chimera). The sequence complementarity-based method to fluorescently label the 3′ end of the 5′ RNase H cleavage fragments using...
the Klenow fragment further constrains the size of cleavage fragments for gel- or capillary electrophoresis-based fluorescence detection and quantitation. In addition, our simplified one-step affinity-based post-RNase H purification method allows for lower RNA input compared to the silica-based purification method (Vlatkovic et al. 2022).

Using the TO selection framework that screens for RNase H specificity empirically, we have achieved ≥90% cleavage specificity in at least one out of four or five TOs on a FLuc transcript (Figs. 4, 5), a CLuc transcript (Supplemental Fig. 5) and one unrelated transcript (data not shown). Sufficient level of cleavage and product recovery could be obtained using TOs ranging from 14 to 37 nt (data not shown). The fact that pseudouridine substitution could affect RNase H cleavage specificity showed that it is important to include modified nucleotides in the surrogate substrate to ensure the TO-screening results are applicable to transcripts that contain modifications. Finally, we believe that our empirical TO-screening framework and the downstream affinity enrichment method is a generalized approach to develop a highly effective RNA cap analysis method.

Our methods present additional tools for RNA cap analysis for the rapidly expanding applications of synthetic mRNA. Ultimately, the availability of multiple approaches to evaluate synthetic mRNA cap incorporation could help democratize synthetic mRNA research, accelerate technology advancement, and benefit basic research.

MATERIALS AND METHODS

In vitro transcription

In vitro transcription of firefly luciferase (FLuc) and Cypridina luciferase (CLuc) transcripts containing artificial 5′ UTRs were performed using PCR-amplified DNA templates and the HiScribe T7 High Yield RNA Synthesis Kit (New England Biolabs), according to manufacturer’s instructions. The RNA transcripts were purified using the Monarch RNA Cleanup Kit (New England Biolabs). RNA concentrations were determined using a Qubit RNA BR Assay Kit (Thermo Fisher Scientific).

RNA capping reaction

Purified in vitro transcripts or synthetic RNA oligonucleotides were capped using vaccinia RNA capping enzyme (Vaccinia Capping System, New England Biolabs) according to manufacturer’s instructions. In some cases, a lower concentration of vaccinia RNA capping enzyme was used to generate partially capped RNA to demonstrate the resolution of reaction intermediates by the analytical methods.

RNase H reactions

DNA–RNA chimeras (targeting oligo or TO) composed of 5′ deoxynucleotides and 3′ ribonucleotides with or without a 3′ TEG-des-thiobiotin group were designed as described in the main text and synthesized chemically (Integrated DNA Technologies or Bio-Synthesis, Inc). For the investigation of RNase H cleavage specificity, 0.5 μM of synthetic RNA oligonucleotides containing a 5′ FAM label were combined with 2.5 μM of the corresponding TOs in a 10 μL reaction containing 1 × RNase H reaction buffer (50 mM Tris-HCl, pH 8.3, 75 mM KCl, 3 mM MgCl2, 10 mM DTT). The mixtures were heated at 80°C for 30 sec and cooled.
down to 25°C at a rate of 0.1°C/sec. The mixtures were then incubated with *E. coli* RNase H (New England Biolabs) or *T. thermophilus* RNase H (Thermosable RNase H, New England Biolabs) at a final concentration of 5 U/µL at 37°C for 1 h. Reactions were quenched by addition of EDTA to a final concentration of 10 mM. The quenched RNase H reactions were then subjected to 15% urea-PAGE followed by fluorescent imaging using the Amersham Typhoon RGB laser scanner (Cytiva). Alternatively, quenched reactions were analyzed by LC-MS directly.

For cap analysis of in vitro transcripts, RNA samples (0.5 µM) were heated at 80°C for 30 sec in the presence of 2.5 µM of appropriate targeting oligos in a 10 µL reaction containing 1x RNase H reaction buffer or 1x RNA capping buffer (50 mM Tris-HCl, pH 8.0, 5 mM KCl, 1 mM MgCl₂, 1 mM DTT). After cooling down to 25°C at a rate of 0.1°C/sec, the reactions were subjected to RNase H cleavage by incubation with Thermosable RNase H (New England Biolabs) at a final concentration of 0.5 U/µL at 37°C for 1 h.

**Klenow fill-in and urea-PAGE analysis**

With the targeting oligos selection framework (Fig. 4), we were able to select a TO that generated a 3’ one-base recessed end after RNase H cleavage. Because RNase H cleavage results in a hydroxyl group at the 3’ end (Lima and Crooke 1997), the recessed 3’ end was filled in using a fluorescently labeled deoxynucleotide using the Klenow fragment. In this example, a FAM-labeled dCTP was incorporated into the 3’ end of the FLuc RNA H cleavage fragment and directly analyzed by urea PAGE (B) or capillary electrophoresis (C) after enrichment. (B) Polyacrylamide gel analysis of cleaved RNA fragments. The 1.7 kb FLuc transcript capped using the vaccinia RNA capping enzyme (VCE) was subjected to the RNase H/Klenow fill-in treatments. Reactions were analyzed directly by urea PAGE followed by laser scanning of total RNA stain using SYBR Gold (left panel) or fluorescent signal (right panel). The targeting oligo TO-1 was invisible when the gel was scanned using the FAM channel and did not interfere with quantitation of the 3’ cleavage products. (C) Resolution and quantification of capping and capping intermediates using capillary electrophoresis. The FLuc transcript was capped using a low concentration of VCE (10 nM) and subjected to the RNase H/Klenow fill-in reactions. After enrichment, the RNA was analyzed using capillary electrophoresis. In addition to substrate 3’ triphosphate (ppp-) and the product m7Gppp-capped forms, enzymatic intermediate products 3’ diphosphate (pp-) and the unmethyl-cap (Gppp-) can be resolved and quantified. (D) Synthetic mRNA cap analysis using gel electrophoresis, capillary electrophoresis, and LC-MS intact mass analysis produce comparable results. After RNase H/Klenow fill-in reactions and enrichment, an uncapped or partially capped FLuc transcript was analyzed using all three available methods. Capillary electrophoresis and LC-MS yielded comparable results in quantification of substrate (ppp-), product (m7Gppp-), and intermediate products (pp- and Gppp-). Urea-PAGE does not resolve pp- from ppp- or Gppp- from m7Gppp-. Considering pp- and pp- as uncapped and Gppp- and m7Gppp- as capped species, quantitation of fluorescently labeled RNase H cleavage products using urea-PAGE generates results comparable to CE or LC-MS, despite the lack of resolution for intermediate products.

For urea-PAGE analysis, 2 µL of the Klenow reactions were retrieved and added to 8 µL of 2x RNA Loading Dye (New England Biolabs). The mixtures were then analyzed by electrophoresis through a urea-15% polyacrylamide gel (Novex TBE-Urea gel 15%, Thermo Fisher Scientific). To acquire the fluorescent signals, the urea gels were scanned directly using the Amersham Typhoon RGB scanner (Cytiva) with the Cy2 channel. To acquire total RNA stain, the gels were stained using SYBR Gold (Thermo Fisher Scientific) and scanned using the Amersham Typhoon RGB scanner using the same settings. Data were analyzed using ImageQuant (Cytiva).

![FIGURE 8. Fluorescent labeling of RNase H 5’ cleavage products and analyses. (A) Schematic representation of targeting, cleavage, and labeling. The targeting oligo was designed to guide RNase H to generate a 1-nt 3’ recessive end, which can be filled in by a fluorescently labeled deoxynucleotide using the Klenow fragment. In this example, a FAM-labeled dCTP was incorporated into the 3’ end of the FLuc RNA H cleavage fragment and directly analyzed by urea PAGE (B) or capillary electrophoresis (C) after enrichment. (B) Polyacrylamide gel analysis of cleaved RNA fragments. The 1.7 kb FLuc transcript capped using the vaccinia RNA capping enzyme (VCE) was subjected to the RNase H/Klenow fill-in treatments. Reactions were analyzed directly by urea PAGE followed by laser scanning of total RNA stain using SYBR Gold (left panel) or fluorescent signal (right panel). The targeting oligo TO-1 was invisible when the gel was scanned using the FAM channel and did not interfere with quantitation of the 3’ cleavage products. (C) Resolution and quantification of capping and capping intermediates using capillary electrophoresis. The FLuc transcript was capped using a low concentration of VCE (10 nM) and subjected to the RNase H/Klenow fill-in reactions. After enrichment, the RNA was analyzed using capillary electrophoresis. In addition to substrate 3’ triphosphate (ppp-) and the product m7Gppp-capped forms, enzymatic intermediate products 3’ diphosphate (pp-) and the unmethyl-cap (Gppp-) can be resolved and quantified. (D) Synthetic mRNA cap analysis using gel electrophoresis, capillary electrophoresis, and LC-MS intact mass analysis produce comparable results. After RNase H/Klenow fill-in reactions and enrichment, an uncapped or partially capped FLuc transcript was analyzed using all three available methods. Capillary electrophoresis and LC-MS yielded comparable results in quantification of substrate (ppp-), product (m7Gppp-), and intermediate products (pp- and Gppp-). Urea-PAGE does not resolve pp- from ppp- or Gppp- from m7Gppp-. Considering pp- and pp- as uncapped and Gppp- and m7Gppp- as capped species, quantitation of fluorescently labeled RNase H cleavage products using urea-PAGE generates results comparable to CE or LC-MS, despite the lack of resolution for intermediate products.](image-url)
Capillary electrophoresis of FAM-labeled RNA was carried out using an Applied Biosystems 3730xl Genetic Analyzer (96 capillary array) using POP-7 polymer and GeneScan 120 LIZ dye Size Standard (Applied Biosystems) (Greenough et al. 2016; Wulf et al. 2019). Peak detection and quantification were performed using the Peak Scanner software v.1.0 (Thermo Fisher Scientific) and an in-house data analysis suite. Details of method validation and sample electropherograms can be found in Supplemental Figure 3.

Purification of RNase H cleavage products

After RNase H cleavage and the optional Klenow fill-in, the 5’ cleavage fragment/TO duplex was purified by size selection followed by affinity-purification (Fig. 2). Briefly, 45 µL of nuclease-free water was added to 5 µL of RNase H cleavage reactions. The mixture was then added to 100 µL of NEBNext Sample Purification Magnetic Beads (New England Biolabs) and incubated at room temperature for 5 min. The beads were then placed on a magnet for 2 min at room temperature. The clarified supernatant was retrieved and added to precleared beads derived from 100 µL of NEBNext Sample Purification Magnetic Beads. The beads were again placed on a magnet for 2 min at room temperature. The clarified supernatant was retrieved and added to precleared beads derived from 50 µL of Dynabeads MyOne Streptavidin C1 (Thermo Fisher Scientific). For the evaluation of biotin and water elution fractions, after washing with 200 µL of standard streptavidin beads wash buffer (5 mM Tris-HCl, pH 7.5, 1 M NaCl, 0.5 mM EDTA), the slurry was divided into two equal fractions. The first fraction was washed three more times with 100 µL of the standard streptavidin beads wash buffer, and the bound RNA was eluted by incubating the clarified beads in 10 µL of 0.1 M biotin solution at 37°C for 1 h. The other slurry fraction was washed three more times with 100 µL of a low salt wash buffer (5 mM Tris, pH 7.5, 0.5 mM EDTA, 60 mM NaCl). The bound RNA was eluted by incubating the clarified beads in 15 µL of nuclease-free water at 65°C for 5 min (Holmberg et al. 2005). The eluted RNA was filtered through a 0.22 µm Ultrafree-MC centrifugal filter device (hydrophilic PVDF, 0.5 mL) (Millipore Sigma) prior to capillary electrophoresis or LC-MS analysis. Although higher molecular weight RNA was present when the size-selection step was omitted, the data quality of the LC-MS analysis was not affected (data not shown). Hence, the size-selection step was considered optional. For cap analysis, the streptavidin magnetic beads with bound 5’ cleavage fragment/TO duplexes were washed four times with 100 µL low salt wash buffer. The RNA was eluted by incubating the washed beads in 15 µL of nuclease-free water at 65°C for 5 min and filtered through a 0.22 µm Ultrafree-MC centrifugal filter device (hydrophilic PVDF, 0.5 mL) (Millipore Sigma) prior to capillary electrophoresis or LC-MS analysis.

LC-MS intact mass analysis

Intact oligonucleotide mass analyses were done using in-house facilities or by external contractor Novatia LLC. For in-house analyses, intact oligonucleotide mass analyses were performed by liquid chromatography-mass spectrometry (LC-MS) on a Vanquish Horizon UHPLC system equipped with a diode array detector (Thermo Scientific) and a Q-Exactive Plus Orbitrap Mass Spectrometer operating under negative electrospray ionization mode (–ESI) (Thermo Scientific). UHPLC was performed using a DNA Pac RP Column (2.1 x 50 mm, 4 µm; Thermo Scientific) at 70°C and 0.3 mL/min flow rate with a gradient mobile phase consisting of a hexafluorosilaneopropyl (HFIP) and N,N-diisopropylethylamine (DIEA) aqueous buffer and methanol. UV detection was performed at 260 nm. Intact mass analysis was performed under full scan mode at a resolution of 70,000 (FWHM) at m/z 200. ESI-MS raw data were deconvoluted using Prommass HR (Novatia, LLC). The relative abundance of each deconvoluted mass peak was used to calculate the percentage of the substrate (ppp—), intermediates (pp—, Gppp—) and final products (m7Gppp— or m7GpppNm—) of enzymatic RNA capping. Methods and results for validation of deconvoluted mass peak intensities for RNA quantitation can be found in Supplemental Figure 4.

DATA DEPOSITION

Additional data are available upon request.

SUPPLEMENTAL MATERIAL

Supplemental material is available for this article.
REFERENCES

Bevery M, Dell P, Parmar P, Houghton L. 2016. Label-free analysis of mRNA capping efficiency using RNase H probes and LC-MS. Anal Bioanal Chem 408: 5021–5030. doi:10.1007/s00216-016-9605-x
Broccoli S, Rallu F, Sanscartier P, Cerritelli SM, Crouch RJ, Drolet M. 2004. Effects of RNA polymerase modifications on transcription-induced negative supercoiling and associated R-loop formation. Mol Microbiol 52: 1769–1779. doi:10.1111/j.1365-2958.2004.04092.x
Cairns MJ, King A, Sun LQ. 2003. Optimisation of the 10–23 DNAzyme-substrate pairing interactions enhanced RNA cleavage activity at purine-cytosine target sites. Nucleic Acids Res 31: 2883–2889. doi:10.1093/nar/gkg378
Deb J, Popenda Ł, Samorska J, Małgowska M, Lahir A, Gdaniecz K, Kierzek R. 2019. Computational and NMR studies of RNA duplexes with an internal pseudouridine-adenosine base pair. Sci Rep 9: 16278. doi:10.1038/s41598-019-52637-0
Devarkar SC, Wang C, Miller MT, Ramanathan A, Jiang F, Khan AG, Patel SS, Marcotrigiano J. 2016. Structural basis for m7G recognition and 2′-O-methyl discrimination in capped RNAs by the innate immune receptor RIG-I. Proc Natl Acad Sci 113: 596–601. doi:10.1073/pnas.1515152113
Floor SN, Borja MS, Gross JD. 2012. Interdomain dynamics and coactivation of the mRNA degrading enzyme Dcp2 are mediated by a gatekeeper tryptophan. Proc Natl Acad Sci 109: 2872–2877. doi:10.1073/pnas.1113620109
Fuchs A-L, Neu A, Sprangrs R. 2016. A general method for rapid and cost-efficient large-scale production of S′ capped RNA. RNA 22: 1454–1466. doi:10.1261/rna.056614.116
Gonatopoulos-Pournatzis T, Cowling VHH. 2014. Cap-binding complex (CBC). Biochem J 457: 231–242. doi:10.1042/BJ20131214
Goodfellow IG, Roberts LO. 2008. Eukaryotic initiation factor 4E. Int J Biochem Cell Biol 40: 2675–2680. doi:10.1016/j.biocel.2007.10.023
Greenough L, Schemerhorn KM, Mazzola L, Bybee J, Rivizigoz D, Cantin E, Slatko BE, Gardner AF. 2016. Adapting capillary gel electrophoresis as a sensitive, high-throughput method to accelerate characterization of nucleic acid metabolic enzymes. Nucleic Acids Res 44: e15. doi:10.1093/nar/gkv899
Henderson JM, Ujita A, Hill E, Yousif-Rosales S, Smith C, Ko N, McReynolds T, Cabral CR, Escamilia-Powers JR, Houston ME. 2021. Cap 1 messenger RNA synthesis with co-transcriptional CleanCap® analog by in vitro transcription. Curr Protoc 1: e39. doi:10.1002/cpct.1.39
Holmberg A, Blomstergren A, Nord O, Lukacs M, Lundeberg J, Uhlen M. 2005. The biotin-streptavidin interaction can be reversibly broken using water at elevated temperatures. Electrophoresis 26: 501–510. doi:10.1002/elps.200410070
Huang Z, Szostak JW. 1996. A simple method for 3′-labeling of RNA. Nucleic Acids Res 24: 4360–4361. doi:10.1093/nar/24.21.4360
Huang J, Kim Y, Turchi JJ, Bambara RA. 1994. Structure-specific cleavage-age of the RNA primer from Okazaki fragments by calf thymus RNase H1. J Biol Chem 269: 25922–25927. doi:10.1074/jbc.2592184334-6
Hyde JL, Diamond MS. 2015. Innate immune restriction and antigen-ism of viral RNA lacking 2′-O methylation. Virology 479–480: 66–74. doi:10.1016/j.virol.2015.01.019
Kariko K, Muramatsu H, Welch FA, Ludwig J, Kato H, Akira S, Weissman D. 2008. Incorporation of pseudouridine into mRNA yields superior nonimmunogenic vector with increased transla-tional capacity and biological stability. Mol Ther 16: 1833–1840. doi:10.1038/mt.2008.200
Kariko K, Muramatsu H, Keller JM, Weissman D. 2012. Increased erythropoiesis in mice injected with submicrogram quantities of pseudouridine-containing mRNA encoding erythropoietin. Mol Ther 20: 948–953. doi:10.1038/mt.2012.7
Kierzek E, Malgowska M, Lisowiec J, Turner DH, Gdaniecz K, Kierzek R. 2014. The contribution of pseudouridine to stabilities and structure of RNAs. Nucleic Acids Res 42: 3492–3501. doi:10.1093/nar/gkt1330
Kumar P, Sweeney TR, Skabbkin MA, Skabkina O V, Hellen CUT, Pestova T V. 2014. Inhibition of translation by IFIT family members is determined by their ability to interact selectively with the 5′-terminal regions of cap0-, cap1- and 5′ppp- mRNAs. Nucleic Acids Res 42: 3228–3245. doi:10.1093/nar/gkt1321
Lapham J, Crothers DM. 1996. RNase H cleavage for processing of in vitro transcribed RNA for NMR studies and RNA ligation. RNA 2: 289–296.
Lapham J, Yu YT, Di SM, Steitz JA, Crothers DM. 1997. The position of site-directed cleavage of RNA using RNase H and 2′-O-methyl oligonucleotides is dependent on the enzyme source. RNA 3: 950.
Lima WF, Crooke ST. 1997. Cleavage of single strand RNA adjacent to RNA-DNA duplex regions by Escherichia coli RNase H1. J Biol Chem 272: 7253–72516. doi:10.1074/jbc.272.44.27513
Maquat LE, Tan W-Y, Isken O. 2010. The pioneer round of translation: features and functions. Cell 142: 368–374. doi:10.1016/j.cell.2010.07.022
Matsuda JS, Sytnikova Y, Chinn GW, Igloi GL, Lau NC. 2014. Small RNA library construction from minute biological samples. Methods Mol Biol 1093: 123–136. doi:10.1007/978-1-62703-694-8_10
Mugridge JS, Tibble RW, Ziemiak M, Jemielity J, Gross JD. 2018. Structure of the activated Edc1-Dcp1-Dcp2-Edc3 mRNA degrading complex with substrate analog poised for catalysis. Nat Commun 9: 1152. doi:10.1038/s41467-018-03536-x
Nance KD, Meier JL. 2021. Modifications in an emergency: the role of N1-methylpseduouridine in COVID-19 vaccines. ACS Cent Sci 7: 748–756. doi:10.1021/acscentsci.1c00197
Parajuli S, Teasley DC, Murali B, Jackson J, Vindigni A, Stewart SA. 2017. Human ribonuclease H1 resolves R-loops and thereby en-ables progression of the DNA replication fork. J Biol Chem 292: 15216–15224. doi:10.1074/jbc.M117.787473
Pardi N, Hogan MJ, Porter FW, Weissman D. 2018. mRNA vaccines: a new era in vaccinology. Nat Rev Drug Discov 17: 261–279. doi:10.1038/nrd.2017.243
Pleiss JA, Derrill MK, Uhlenbeck OC. 1998. T7 RNA polymerase produces S′ end heterogeneity during in vitro transcription from certain templates. RNA 4: 1313–1317. doi:10.1017/S13558389000106X

COMPETING INTEREST STATEMENT

S.H.C., J.M.W., N.D., G.T., I.R.C., Jr., and G.B.R. are current employees of New England Biolabs that markets molecular biology reagents. The affiliation does not affect the authors’ impartiality, adherence to journal standards and policies, or availability of data.

ACKNOWLEDGMENTS

We would like to thank Daniel Kneller for critical review of the manuscript. We are also grateful to the NEB sequencing core for the analysis of capillary electrophoresis samples as well as all the facilities personnel for daily support of operation. This work is supported by New England Biolabs, Inc.

Received March 21, 2022; accepted May 27, 2022.

REFERENCES
Ramanathan A, Robb GB, Chan SH. 2016. mRNA capping: biological functions and applications. *Nucleic Acids Res* **44**: 7511–7526. doi:10.1093/nar/gkw551

Ramaswamy S, Tonnu N, Tachikawa K, Limphong P, Vega JB, Karmali PP, Chivukula P, Verma IM. 2017. Systemic delivery of factor IX messenger RNA for protein replacement therapy. *Proc Natl Acad Sci* **114**: E1941–E1950. doi:10.1073/pnas.1619653114

Sandin P, Stengel G, Ljungdahl T, Börjesson K, Macao B, Wilhelmsson LM. 2009. Highly efficient incorporation of the fluorescent nucleotide analogs tC and tC\(^{\text{D}}\) by Klenow fragment. *Nucleic Acids Res* **37**: 3924–3933. doi:10.1093/nar/gkp266

Svitkin YV, Cheng YM, Chakraborty T, Presnyak V, John M, Sonenberg N. 2017. N1-methyl-pseudouridine in mRNA enhances translation through eIF2\(\alpha\)-dependent and independent mechanisms by increasing ribosome density. *Nucleic Acids Res* **45**: 6023–6036. doi:10.1093/nar/gkx135

Vaidyanathan S, Azizan KT, Ashiqul Haque AKM, Henderson JM, Hendel A, Shore S, Antony JS, Hogrefe RI, Kormann MSD, Porteus MH, et al. 2018. Uridine depletion and chemical modification increase Cas9 mRNA activity and reduce immunogenicity without HPLC purification. *Mol Ther Nucleic Acids* **12**: 530–542. doi:10.1016/j.omtn.2018.06.010

VanBlargan LA, Himansu S, Foreman BM, Ebel GD, Pierson TC, Diamond MS. 2018. An mRNA vaccine protects mice against multiple tick-transmitted flavivirus infections. *Cell Rep* **25**: e3. doi:10.1016/j.celrep.2018.11.082

Vlatkovic I, Ludwig J, Boros G, Szabó GT, Reichert J, Buff M, Baiersdörfer M, Reinholz J, Mahiny AJ, Şahin U, et al. 2022. Ribozyme assays to quantify the capping efficiency of in vitro-transcribed mRNA. *Pharmaceutics* **14**: 328. doi:10.3390/pharmaceutics14020328

Wulf MG, Buswell J, Chan SH, Dai N, Marks K, Martin ER, Tzertzinis G, Whipple JM, Corrêa IR, Schildkraut I. 2019. The yeast scavenger decapping enzyme DcpS and its application for in vitro RNA recapping. *Sci Rep* **9**: 8594. doi:10.1038/s41598-019-45083-5

Yu YT, Di SM, Steitz JA. 1997. A new method for detecting sites of 2′-O-methylation in RNA molecules. *RNA* **3**: 324–331.