

# Minding your caps and tails – considerations for functional mRNA synthesis

Breton Hornblower, Ph.D., G. Brett Robb, Ph.D. and George Tzertzinis, Ph.D., New England Biolabs, Inc.

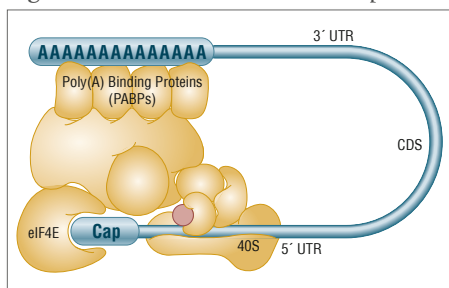
Applications of synthetic mRNA have grown and become considerably diversified in recent years. Examples include the generation of pluripotent stem cells (1-3), vaccines and therapeutics (4), and CRISPR/Cas9 genome editing applications (5-7). The basic requirements for a functional mRNA – a 7-methylguanylate cap at the 5' end and a poly(A) tail at the 3' end – must be added in order to obtain efficient translation by eukaryotic cells. Additional considerations can include the incorporation of modified bases, modified cap structures and polyadenylation strategies. Strategies for *in vitro* synthesis of mRNA may also vary according to the desired scale of synthesis. This article discusses options for selection of reagents and the extent to which they influence synthesized mRNA functionality.

Before translation in eukaryotic organisms, nascent mRNA (pre-mRNA) receives two significant modifications in addition to splicing. During synthesis, a 7-methylguanylate structure, also known as a “cap”, is added to the 5' end of the pre-mRNA, via 5' → 5' triphosphate linkage. This cap protects the mature mRNA from degradation, and also serves a role in nuclear export and translation initiation. The second modification is the addition of approximately 200 adenylate nucleotides (a poly(A) tail) to the 3' end of pre-mRNA by *E. coli* Poly(A) Polymerase. Polyadenylation is coupled to transcription termination, export of mRNA from the nucleus, and, like the cap, formation of the translation initiation complex. The mature mRNA forms a circular structure by bridging the cap to the poly(A) tail via the cap-binding protein eIF4E (eukaryotic initiation factor 4E) and the poly-(A) binding protein, both of which interact with eIF4G (eukaryotic initiation factor 4G). (Figure 1, (8))

RNA can be efficiently synthesized *in vitro* with prokaryotic phage polymerases, such as T7, T3 and SP6. The cap and poly(A) tail structures characteristic of mature mRNA can be added during or after the synthesis by enzymatic reactions with capping enzymes and Poly(A) Polymerase, respectively.

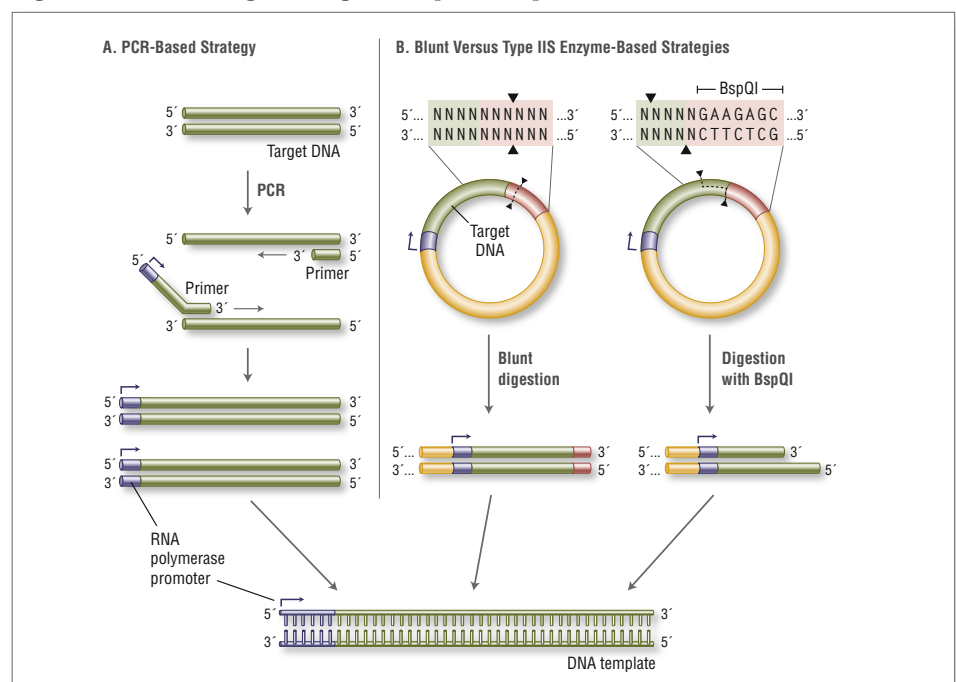
There are several factors to consider when planning for *in vitro* mRNA synthesis that will influence the ease of experimental setup and yield of the final mRNA product. These are discussed in the following sections.

Figure 1. Translation initiation complex.



The mature RNA forms a circular structure connected by protein complexes that bind the cap structure and poly(A) tail.

Figure 2. Methods for generating transcription templates



(A) PCR can be used to amplify target DNA prior to transcription. A promoter can be introduced via the upstream primer.

(B) When using plasmid DNA as a template, linearize with an enzyme that produces blunt or 5'-overhanging ends. Using a type IIS restriction enzyme (e.g., BspQI) allows RNA synthesis with no additional 3'-nucleotide sequence from the restriction site.

## DNA template

The DNA template provides the sequence to be transcribed downstream of an RNA polymerase promoter. There are two strategies for generating transcription templates: PCR amplification and linearization of plasmid with a restriction enzyme (Figure 2). Which one to choose will depend on the downstream application. In general, if multiple sequences are to be made and transcribed in parallel, PCR amplification is recommended as it generates many templates quickly. On the other hand, if large amounts of one or a few templates are required, plasmid DNA is recommended, because of the relative ease of producing large quantities of high quality, fully characterized plasmids.

PCR allows conversion of any DNA fragment to a transcription template by appending the

T7 promoter to the forward primer (Fig 2A). Additionally, poly(d)T-tailed reverse primers can be used in PCR to generate transcription templates with A-tails. This obviates the need for a separate polyadenylation step following transcription. Repeated amplifications should, however, be avoided to prevent PCR-generated point mutations. Amplification using PCR enzymes with the highest possible fidelity, such as Q5<sup>®</sup> High-Fidelity DNA Polymerase (NEB # M0491), reduces the likelihood of introducing such mutations.

The quality of the PCR reaction can be assessed by running a small amount on an agarose gel, and DNA should be purified before *in vitro* transcription using a spin column or magnetic beads (e.g., AMPure<sup>®</sup> beads). Multiple PCR reactions can be purified and combined to generate a

DNA stock solution that can be stored at  $-20^{\circ}\text{C}$ , and used as needed for *in vitro* transcription.

Plasmid templates are convenient if the template sequence already exists in a eukaryotic expression vector also containing the T7 promoter (e.g., pcDNA vector series). These templates include 5'- and 3'-untranslated regions (UTR), which are important for the expression characteristics of the mRNA.

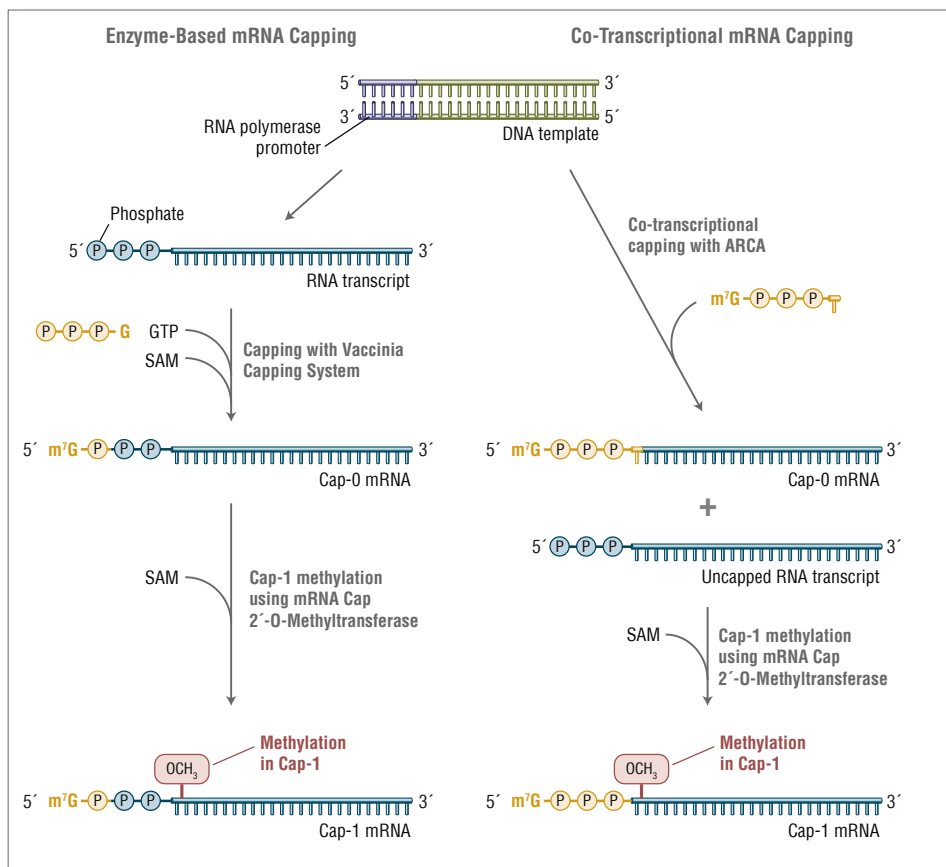
Plasmid DNA should be purified and linearized downstream of the desired sequence, preferably with a restriction enzyme that leaves blunt or 5' overhangs at the 3' end of the template. These are favorable for proper run-off transcription by T7 RNA Polymerase, while 3' overhangs may result in unwanted transcription products. To avoid adding extra nucleotides from the restriction site to the RNA sequence, a Type IIS restriction enzyme can be used (e.g., BspQI, NEB #R0712), which positions the recognition sequence outside of the transcribed sequence (Fig-

ure 2B). The plasmid DNA should be completely digested with the restriction enzyme, followed by purification using a spin column or phenol extraction/ethanol precipitation. Although linearization of plasmid involves multiple steps, the process is easier to scale for the generation of large amounts of template for multiple transcription reactions.

### *In vitro* transcription

There are two options for the *in vitro* transcription reaction depending on the capping strategy chosen: standard synthesis with enzyme-based capping following the transcription reaction (post-transcriptional capping) or incorporation of a cap analog during transcription (co-transcriptional capping) (Fig. 3). Method selection will depend on the scale of mRNA synthesis required and number of templates to be transcribed.

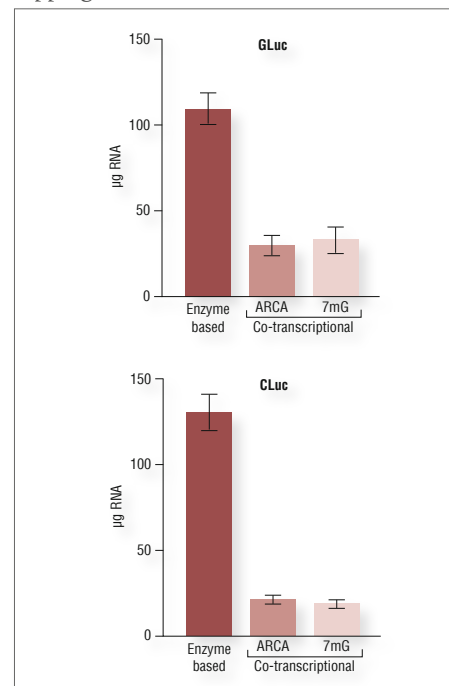
Figure 3. *In vitro* transcription options based upon capping strategy



Enzyme-based capping (left) is performed after *in vitro* transcription using 5'-triphosphate RNA, GTP, and S-adenosyl-methionine (SAM). Cap 0 mRNA can be converted to cap 1 mRNA using mRNA cap 2'-O-methyltransferase (MTase) and SAM in a subsequent or concurrent reaction. The methyl group transferred by the MTase to the 2'-O of the first nucleotide of the transcript is indicated in red. Conversion of ~100% of 5'-triphosphorylated transcripts to capped mRNA is routinely achievable using enzyme-based capping.

Co-transcriptional capping (right) uses an mRNA cap analog (e.g., ARCA; anti-reverse cap analog), shown in yellow, in the transcription reaction. The cap analog is incorporated as the first nucleotide of the transcript. ARCA contains an additional 3'-O-methyl group on the 7-methylguanosine to ensure incorporation in the correct orientation. The 3'-O-methyl modification does not occur in natural mRNA caps. Compared to reactions not containing cap analog, transcription yields are lower. ARCA-capped mRNA can be converted to cap 1 mRNA using mRNA cap 2'-O-MTase and SAM in a subsequent reaction.

Figure 4. RNA yields from transcriptional capping reactions



Reactions were set up according to recommended conditions for two templates: *Glossinia luciferase* (GLuc) and *Cypridina luciferase* (CLuc). The RNA was quantified spectrophotometrically after purification with spin columns.

### Transcription for enzyme-based capping (post-transcriptional capping)

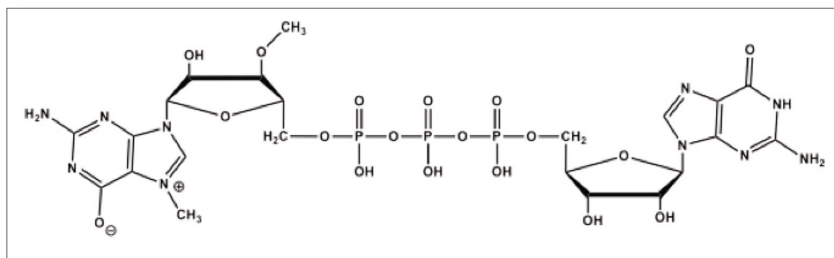
Standard RNA synthesis reactions produce the highest yield of RNA transcript (typically  $\geq 100$  µg per 20 µl in a 1 hr reaction using the HiScribe Quick T7 High Yield RNA Synthesis Kit, NEB #E2050S). Transcription reactions are highly scalable, and can be performed using an all-inclusive kit (e.g., HiScribe kits), or individual reagents. More information on the HiScribe kits can be found on page 7.

Following transcription, the RNA is treated with DNase I to remove the DNA template, and purified using an appropriate column, kit or magnetic beads, prior to capping. This method produces high yields of RNA with 5'-triphosphate termini that must be converted to cap structures. In the absence of template-encoded poly(A) tails, transcripts produced using this method bear 3' termini that also must be polyadenylated in a separate enzymatic step, as described below in "Post-transcriptional capping and Cap-1 methylation".

### Transcription with co-transcriptional capping

With co-transcriptional capping, a cap analog is introduced into the transcription reaction, along with the four standard nucleotide triphosphates, in an optimized ratio of cap analog to GTP 4:1. This allows initiation of the transcript with the cap structure in a large proportion of the synthe-

Figure 5. Structure of the anti-reverse cap analog, ARCA



The 3' position of the 7-methylated G is blocked by a methyl group.

sized RNA molecules. This approach produces a mixture of transcripts, of which ~80% are capped, and the remainder have 5'-triphosphate ends. Decreased overall yield of RNA products results from the lower concentration of GTP in the reaction (Fig. 4).

There are several cap analogs used in co-transcriptional RNA capping. The most common are the standard 7-methyl guanosine (m7G) cap analog and anti-reverse cap analog (ARCA), also known as 3' O-me 7-meGpppG cap analog (Fig. 5). ARCA is methylated at the 3' position of the m7G, preventing RNA elongation by phosphodiester bond formation at this position. Thus, transcripts synthesized using ARCA contain 5'-m7G cap structures in the correct orientation, with the 7-methylated G as the terminal residue. In contrast, the m7G cap analog can be incorporated in either the correct or the reverse orientation.

HiScribe T7 ARCA mRNA Synthesis Kits (NEB# E2060 and E2065) contain reagents, including an optimized mix of ARCA and NTPs for streamlined reaction setup for synthesis of co-transcriptionally capped RNAs.

### Transcription with complete substitution with modified nucleotides

RNA synthesis can be carried out with a mixture of modified nucleotides in place of the regular mixture of A, G, C and U triphosphates. For expression applications, the modified nucleotides of choice are the naturally occurring 5'-methylcytidine and/or pseudouridine in the place of C and U, respectively. These have been demonstrated to confer desirable properties to the mRNA, such as higher expression levels and avoidance of unwanted side effects in the key applications of protein replacement and stem-cell differentiation (1). It is important to note that nucleotide choice can influence the overall yield of mRNA synthesis reactions.

Fully substituted RNA synthesis can be achieved using the HiScribe T7 High-Yield RNA Synthesis Kit (NEB# E2040) in conjunction with NTPs with the desired modification. Transcripts made with complete replacement of one or more nucleotides may be post-transcriptionally capped (see next section), or may be co-transcriptionally capped by including ARCA or another cap ana-

log, as described previously.

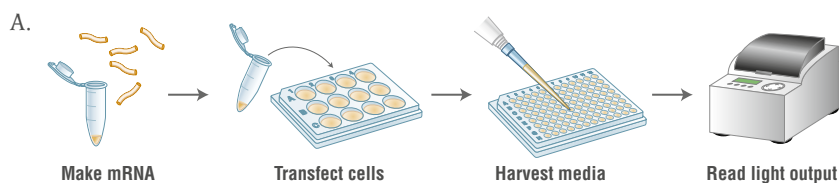
If partial replacement of nucleotides is desired, the HiScribe T7 ARCA mRNA Synthesis Kits (NEB# E2060 and E2065), may be used with added modified NTPs, to produce co-transcriptionally capped mRNAs, as described above. Alternatively, the HiScribe T7 Quick RNA Synthesis Kit may be used to produce transcripts for

post-transcriptional capping (see below).

### Post-transcriptional capping and Cap-1 methylation

Post-transcriptional capping is often performed using the mRNA capping system from *Vaccinia* virus. This enzyme complex converts the 5'-triphosphate ends of *in vitro* transcripts to the m7G-cap structures. The Vaccinia Capping System (NEB #M2080) is comprised of three enzymatic activities (RNA triphosphatase, guanylyl-transferase, guanine N7-methyltransferase) that are necessary for the formation of the complete Cap-0 structure, m7Gppp5' N, using GTP and the methyl donor S-adenosylmethionine. As an added option, the inclusion of the mRNA Cap 2' O-Methyltransferase (NEB #M0366) in the same reaction results in formation of the Cap-1 structure, which is a natural modification in many eukaryotic mRNAs. This enzyme-based capping approach results in the highest proportion of capped message, and it is easily scalable. The resulting capped RNA can be further modified by poly(A) addition before final purification.

### Analysis of capped RNA function in transfected mammalian cells



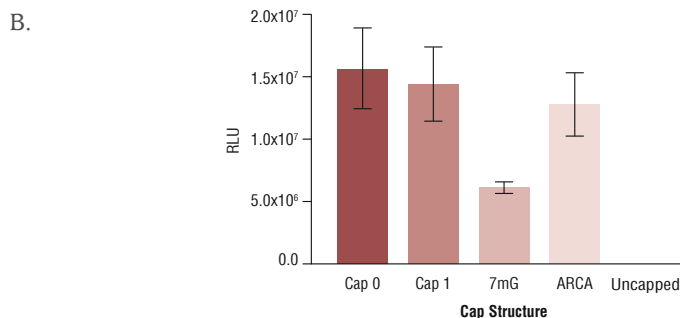
Schematic representation of reporter mRNA transfection workflow.

The effect of capping can be studied by delivering the mRNA to cultured mammalian cells and monitoring its translation. Using RNA encoding secreted luciferases (e.g., *Cypridina* luciferase, CLuc) the translation can be monitored by assaying its activity in the cell culture medium (Fig. A).

CLuc mRNA was synthesized and capped post-transcriptionally (Cap 0 or Cap 1) or co-transcriptionally (as described above) using standard (7mG) or anti-reverse cap analog (ARCA). For consistency, the mRNAs were prepared from templates encoding poly-A tails of the same length.

After capping, the mRNA was purified using magnetic beads and quantified before transfection into U2OS cells using the TransIT<sup>®</sup> mRNA transfection reagent following the manufacturer's protocol. CLuc activity was measured 16 hrs after transfection using the BioLux<sup>®</sup> Cypridina Luciferase Assay Kit (NEB #E3309).

Virtually no luciferase reporter activity was observed in conditions where uncapped RNA was transfected (Fig. B). In contrast, robust activity was detected from cells transfected with RNA capped using the methods described above. As anticipated, lower activity was observed from cells transfected with mRNA capped using the 7mG cap analog as compared to ARCA-capped mRNA.



Expression of *Cypridina* luciferase (CLuc) after capping using different methods. High activity from all capped RNAs is observed.

## A-tailing using *E. coli* Poly(A) Polymerase

The poly(A) tail confers stability to the mRNA and enhances translation efficiency. The poly(A) tail can be encoded in the DNA template by using an appropriately tailed PCR primer, or it can be added to the PCR-amplified template by enzymatic treatment of RNA with *E. coli* Poly(A) Polymerase (NEB #M0276). The lengths of the added tails can be adjusted by titrating the Poly(A) Polymerase in the reaction (Fig. 6).

The importance of the A-tail is demonstrated by transfection of untailed vs. tailed mRNA. When luciferase activity from cells transfected with equimolar amounts of tailed or untailed mRNAs were compared, a significant enhancement of translation efficiency was evident (Fig. 6). Increasing the length of poly(A) tails did not markedly further enhance reporter activity.

HiScribe T7 ARCA mRNA Synthesis Kit (with tailing) (NEB# E2060) includes *E. coli* Poly(A) Polymerase, and enables a streamlined workflow

for the enzymatic tailing of co-transcriptionally capped RNA. For mRNA synthesis from templates with encoded poly(A) tails, the HiScribe T7 ARCA mRNA Synthesis Kit (NEB# E2065) provides an optimized formulation for co-transcriptionally capped transcripts.

## Summary

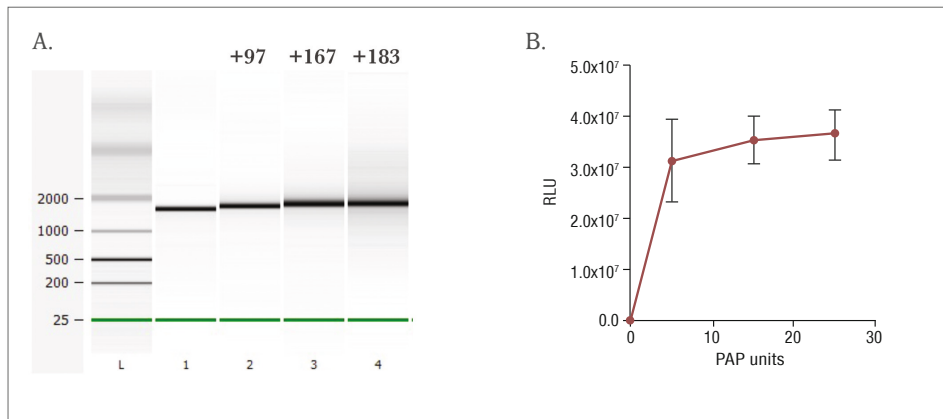
In summary, when choosing the right workflow for your functional mRNA synthesis needs, you must balance your experimental requirements for the mRNA (e.g., internal modified nucleotides) with scalability (i.e., ease-of-reaction setup vs. yield of final product).

In general, co-transcriptional capping of mRNA with template encoded poly(A) tails or post-transcriptional addition of poly(A) tail is recommended for most applications. This approach, using the HiScribe T7 ARCA mRNA Synthesis Kits (NEB# E2060 and E2065), enables the quick and streamlined production of one or many transcripts with typical yields of ~20 µg per reaction, totaling ~400 µg per kit.

Post-transcriptional mRNA capping with Vaccinia Capping System is well suited to larger scale synthesis of one or a few mRNAs, and is readily scalable to produce gram-scale quantities and beyond. Reagents for *in vitro* synthesis of mRNA are available in kit form and as separate components to enable research and large-scale production.

Products available from NEB for each step of the functional mRNA synthesis workflow, from template construction to tailing, are shown to the left.

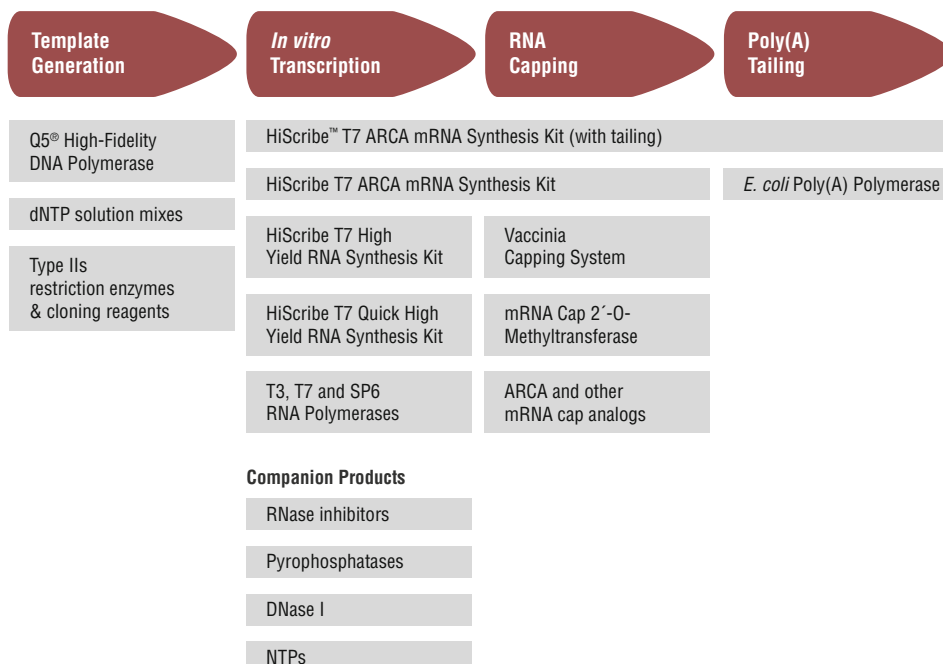
Figure 6. Analysis of capped and polyadenylated RNA



A. Agilent Bioanalyzer<sup>®</sup> analysis of capped and polyadenylated RNA. Longer tails are produced by increasing the enzyme concentration in the reaction. Calculated A-tail lengths are indicated over each lane. Lanes: L: size marker, 1: No poly-A tail, 2: 5 units, 3: 15 units, 4: 25 units of *E. coli* Poly(A) Polymerase per 10 µg CLuc RNA in a 50 µL reaction.

B. Effect of enzymatic A-tailing on the luciferase reporter activity of CLuc mRNA

Products from NEB are available for each step of the RNA Synthesis Product Workflow



## References:

- Warren, L., et al. (2010) *Cell Stem Cell*, 7, 618-630.
- Angel, M. and Yanik, M.F. (2010) *PLoS One*, 5:e11756.
- Yakubov, E., et al. (2010) *Biochem. Biophys. Res. Commun.* 394, 189.
- Geall, A.J., et al. (2012) *Proc. Natl. Acad. Sci. USA*, 109, 14604-14609.
- Ma, Y., et al. (2014) *PLoS One*, 9:e89413.
- Ota, S., et al. (2014) *Genes Cells*, 19, 555-564.
- Andrew R. B., et al. (2013) *Cell Rep.* 4, 220-228.
- Wells, S.E., et al. (1998) *Molecular Cell* 2, 135-140.

www.neb-online.de

NEW ENGLAND  
BioLabs<sup>GmbH</sup>

New England Biolabs GmbH, Brünningstr. 50, Geb. B852, 65926 Frankfurt/Main, Germany

Tel: +49/(0)69/305-23140 Toll Free: (Germany) 0800/246-5227 Toll Free: (Austria) 00800/246-52277 Fax: +49/(0)69/305-23149 e-mail: info.de@neb.com