Published in final edited form as:

Tetrahedron Lett. 2012 September 5; 53(36): 4843–4847. doi:10.1016/j.tetlet.2012.06.127.

5'-Terminal chemical capping of spliced leader RNAs

Karolina Piecyk^a, Richard E. Davis^b, and Marzena Jankowska-Anyszka^{*,a}

^aFaculty of Chemistry, University of Warsaw, 02-093, Warsaw, Poland ^bDepartment of Biochemistry and Molecular Genetics, University of Colorado School of Medicine, Aurora, CO 80045, USA

Abstract

Spliced leader (SL) RNA *trans*-splicing adds a 2,2,7-trimethylguanosine cap (TMG) and a 22-nucleotide sequence, the SL, to the 5' end of mRNAs. Both non-*trans*-spliced with a monomethylguanosine cap (MMG) and *trans*-spliced mRNAs co-exist in *trans*-splicing metazoan cells. Efficient translation of TMG-capped mRNAs in nematodes requires a defined core of nucleotides within the SL sequence. Here we present a chemical procedure for the preparation and purification of 5'-terminal capped MMG and TMG wild-type, and mutant 22 nt spliced leader RNAs (GGU/ACUUAUUACCCAAGUUUGAG) with or without a 3' biotin tag.

Keywords

cap analog; 5'-capped RNA; TMG-capped RNA; SL RNA; spliced leader

Eukaryotic mRNAs have a 5' cap that consists of a methylated guanosine linked by a 5'-5' triphosphate bridge (m⁷GpppG; MMG cap) (Figure 1A) to the first transcribed nucleotide of the mRNA strand. It is known that 70% of all mRNAs in nematodes such as *Caenorhabditis elegans* or *Ascaris suum* have an atypical trimethylated form of the mRNA cap (Figure 1B) with two methyl groups at the N2 position of 7-methylguanosine (m₃^{2,2,7}GpppG; TMG cap).¹ This unusual cap is added with a conserved spliced leader sequence (SL) during *trans*-splicing.² Recent findings^{3,4} indicated that the 22 nt nematode SL sequence significantly enhances translation of TMG-capped RNAs, having little effect on translation of MMG-capped messages. Efficient translation requires specific nucleotides and structural elements within a spliced leader sequence including a small stem-loop that contains a three base-pair stem (Figure 1C) immediately downstream of the cap.⁵ Recently, the first crystal structures

© 2012 Elsevier Ltd. All rights reserved.

Corresponding Author: Marzena Jankowska-Anyszka, Faculty of Chemistry, University of Warsaw, Poland, 02-093 Warsaw, Pasteura 1, marzena@chem.uw.edu.pl, phone (+48 22) 822 0211, fax (+48-22) 822 5996.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

of *Ascaris* eIF4E-3 (eukaryotic initiation factor eIF4E recognizes and binds cap structures during the initiation step of translation) in complex with an MMG and a TMG cap were published.⁶ These structures do not include the native substrate for eIF4E-3, the TMG-capped 22 nt spliced leader RNA and its stem loop, and therefore do not define the full interaction of the cap-binding protein with the capped RNA. In order to carry out mechanistic binding studies of the eIF4E-3 with the TMG-capped SL RNA sequence and to determine the structure of the RNA bound to eIF4E-3, large amounts (milligrams) of pure capped RNA are needed.

Numerous studies have been carried out to identify an efficient method to synthesize capped oligonucleotides. Two general methods have been explored: solid supported and “in solution” synthesis. Despite several limitations that are a consequence of the instability of 7-methylguanosine under acidic (depurination) and basic conditions (opening of the 7-methylguanosine imidazole ring), the synthesis of short 5'-capped oligonucleotides on solid supports have been described in the literature. The first synthesis, published in 2001 by Kadokura et al.,⁷ of m₃^{2,2,7}GpppAmUmA was carried out using an acid-labile phosphoramidate linker. The assembly of the cap was achieved by coupling of a 2',3'-phenylboronylated 5'-phosphoroimidazolide of N²,N², 7-trimethylguanosine with a 5'-diphosphorylated trinucleotide unit. The coupling efficiency and yield of isolated product for this approach were, however, very low. To improve the synthesis, the Sekine group⁸ developed a method for the preparation of a 13 nt oligomer (TCAGTCAGTCAGT) having a 5' terminal TMG cap with a pyrophosphate linkage using a silyl-type linker on polymer supports. However, to our knowledge, no biological studies have been reported using TMG capped mRNAs lacking one phosphate group. In addition to this method, Jemielity et al.⁹ described an alternative approach exploiting a disulfide linker and an imidazolide derivative of unprotected m⁷GDP as a capping agent. The reported yield of the capping reaction was approximately 50%. However, the product was not isolated and the efficacy of the synthesis was not determined. Sawai et al.,¹⁰ and Sekine and co-workers,¹¹ studied an “in solution” capping reaction of 5' phosphorylated oligonucleotides under aqueous or anhydrous conditions in the presence of double-charged ions (Mn²⁺, Zn²⁺). Their methods included the coupling of a nucleotide diphosphate imidazolide derivative with the 5'-monophosphorylated oligonucleotides. The longest 5' MMG-capped oligoribonucleotide prepared and isolated by HPLC was the 11 mer, pACACUUGCUUU. To show the effectiveness of the method for longer oligonucleotides, they examined the MMG capping reaction of a commercially available 72 chain length tRNAPhe with a 5'-terminal monophosphate, but the resulting product was not isolated. Recently, Thillier et al.¹² described a method for the synthesis of m⁷G-capped RNAs that combined a solid-phase chemical synthesis and an enzymatic methylation of the guanine moiety. The authors synthesized short RNAs (from 4 to 18 nt) on a solid support using phosphoramidite 2'-O-pivaloyloxymethyl chemistry, which were further phosphorylated at the terminal 5'-OH, activated by imidazole, and coupled with GDP. Nanomolar amounts of several, ion-exchange (IE) HPLC purified, GpppN-RNAs were methylated enzymatically at the N7 position using human (guanine-N7)-methyl transferase. This approach might be explored in the future to prepare TMG-capped RNAs, however it would require a totally new procedure that would involve an additional methyltransferase enzyme for the N2 methylation. Thus, while challenging, there remains a need for an efficient method for the synthesis of capped oligonucleotides.

Our goal was to synthesize (from nano to micromolar amounts) 22 nt wild type SL RNAs capped with m⁷G-, m₃^{2,2,7}G- and G- (as a control that should not interact with cap-binding proteins) in order to better understand the nature and the contribution of the TMG-cap and the sequence/structure of the SL to post-transcriptional gene expression in nematodes. We chose to develop an “in solution” method. The capping reaction was carried out by coupling

of a 5'-monophosphorylated oligonucleotide (TriLink BioTechnologies) with an activated nucleoside 5'-diphosphate (Scheme 1).¹³ GDP, m⁷GDP and m₃^{2,2,7}GDP were prepared as previously described¹⁴ and activated with imidazole in the presence of 2,2'-dithiopyridine and triphenylphosphine (see Supplementary Material). A large molar excess of the imidazole derivative (imGDP, im(m⁷GDP), im(m₃^{2,2,7}GDP) and metal ion (Zn²⁺ or Mn²⁺) with respect to the 5'-monophosphorylated oligonucleotide was used. We compared the coupling of imidazole derivatives of nucleoside 5'-diphosphates with the 5' monophosphorylated oligonucleotides under aqueous and anhydrous conditions. The reaction in aqueous conditions occurred with much higher efficiency than the similar reaction in anhydrous DMF. Following the reaction, the capped and uncapped oligonucleotides were readily precipitated from the other reactants thereby simplifying the purification procedure (see Supplementary Material). Coupling yields were dependent on the imidazole derivative used and correlated with the number of methyl groups present on the guanine moiety. The order of capping efficiency for the wild-type SL was imGDP > im(m⁷GDP) > im(m₃^{2,2,7}GDP) (Table 1).

We also prepared capped SL RNAs modified at the 3' end with biotin (Scheme 1). The coupling yields are shown in Table 1, and the HPLC purifications (see below) of all the biotinylated oligonucleotides were performed similarly to their non-biotinylated counterparts.¹³

Separation of the product from the reaction mixture, in particular the capped oligonucleotide from the uncapped example, can be problematic with increasing lengths of the RNA fragments. Purification of short, capped oligonucleotides has typically been carried out using ion-pair reversed-phase (IP-RP) HPLC based on the charge-to-charge interaction of phosphate internucleotide linkages with ion-pairing agents adsorbed onto the stationary phase, and hydrophobic interactions of the nucleobases with the reversed-phase adsorbent.¹⁵ Although ion-exchange HPLC has also recently been used,¹² it was not applicable in our case. The method can only be used if the coupled nucleotide is a non-methylated G. In the case of a 7-methylated G (m⁷G- or m₃^{2,2,7}G) the positive charge generated at the N7 position by the methyl substituent neutralizes the net negative charge of the two incorporated phosphate groups, and consequently the charge differences are not sufficient for good peak separation of capped from uncapped oligonucleotides. For that reason, ion-pair reversed-phase chromatography with inexpensive and volatile triethylammonium acetate (TEAA) as the ion pairing mobile phase was used in our research. This method was used to purify G-, m⁷G- and m₃^{2,2,7}G-capped spliced leader sequences with or without 3' biotin, and separation was also dependent on the number of methyl groups within the guanine moiety with the best peak separation being G > m⁷G > m₃^{2,2,7}G (Figure 2).

Previous studies³⁻⁵ demonstrated that specific nucleotides within the nematode spliced leader sequence are needed for efficient translation and are likely to contribute to a putative stem-loop structure. The mutations of the second and third (GT->AC) nucleotides of the spliced leader (Scheme 1) lead to significant reductions in translation and would disrupt potential base pairing within the SL stem-loop. Given the importance of these mutations in the biology of the SL sequence, we also prepared mutant 22 nt SL RNAs capped with G, m⁷G and m₃^{2,2,7}G, with or without 3' biotin. Surprisingly, this small modification in the oligonucleotide sequence led to substantial changes in the course of the reaction and purification. The coupling efficiency was lower in the mutant SL and separation of the substrate and product was much more difficult (Table 1, Figure 2). One possible explanation for the greater coupling efficiency for the wild-type (WT) SL may be related to the predicted stem-loop structure in the SL that might increase the availability or interaction of the 5' end of the oligonucleotide for the imidazole derivatives. The base sequence and the ion-pairing agent are known to play important roles in the retention behavior of oligonucleotides in

HPLC.¹⁶ Oligonucleotides as short as 10 bases with self-complementary sequences, possess decreased surface area for hydrophobic interactions of nucleobases with reversed phase adsorbents, and they are eluted faster than those without a self-complementary sequence in TEAA buffer system. We hypothesize that the SL sequence creating a stem-loop behaves more like a shorter oligonucleotide during the HPLC separation due to decreased surface area for hydrophobic interactions. The consequence of the orderly shape of WT SL together with the fact that the separation effectiveness decreases with increased oligonucleotide length, may explain the more difficult separation and purification of the mutant SL compared to the WT SL.

In conclusion, we have described the preparation of several capped (G-, m⁷G- and m₃^{2,2,7}G-) spliced leader RNAs (wild type and mutant SL, with or without biotin) in nano/micromolar amounts. To our knowledge, this is the first report of capping and purifying such long RNAs (22 nt). The effectiveness of the preparation of capped RNAs as a function of the imidazole derivatives (GDP, m⁷GDP and m₃^{2,2,7}GDP) and oligonucleotide sequence is described. In addition, the dependence of the SL sequence and the ratio of guanine methylation of the cap structure indicate features that need to be taken into account for both the coupling and purification.

The non-biotinylated WT and mutant SL RNAs are currently being used in NMR studies to characterize the structure of the SL that is necessary for translation.⁵ The biotinylated WT and mutant SL RNAs, are being used in surface plasmon resonance studies to compare the binding affinity of nematode eIF4E isoforms for the MMG and TMG-capped SL RNAs, and in pull-down assays with the nematode translation extracts to identify proteins that may be uniquely associated with TMG-SL mRNAs compared to MMG-capped mRNAs.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was partially supported by a grant N N301 096339 from the Ministry of Science and Higher Education, Poland and grant R0149558 from the National Institute of Health, U.S.A. (to R. E. D.). We thank Dr. Janusz Stepinski and Prof. Edward Darzynkiewicz from the Division of Biophysics, Institute of Experimental Physics, Warsaw University for sharing their time and advice over the course of this work.

References

1. (a) Thomas JD, Conrad RC, Blumenthal T. Cell. 1988; 54:533. [PubMed: 3401926] (b) Liou RF, Blumenthal T. Mol Cell Biol. 1990; 10:1764. [PubMed: 2157142]
2. (a) Van Doren K, Hirsh D. Mol Cell Biol. 1990; 10:1769. [PubMed: 1690851] (b) Maroney PA, Hannon GJ, Nilsen TW. Proc Natl Acad Sci U S A. 1990; 87:709. [PubMed: 2300557]
3. Lall S, Friedman CC, Jankowska-Anyszka M, Stepinski J, Darzynkiewicz E, Davis RE. J Biol Chem. 2004; 279:45573. [PubMed: 15322127]
4. Cheng G, Cohen L, Mikhli C, Jankowska-Anyszka M, Stepinski J, Darzynkiewicz E, Davis RE. Mol Biochem Parasitol. 2007; 153:95. [PubMed: 17391777]
5. Wallace A, Filbin ME, Veo B, McFarland C, Stepinski J, Jankowska-Anyszka M, Darzynkiewicz E, Davis RE. Mol Cell Biol. 2010; 30:1958. [PubMed: 20154140]
6. Liu W, Jankowska-Anyszka M, Piecyk K, Dickson L, Wallace A, Niedzwiecka A, Stepinski J, Stolarski R, Darzynkiewicz E, Kieft J, Zhao R, Jones DNM, Davis RE. Nucleic Acids Res. 2011; 39:8820. [PubMed: 21965542]
7. Kadokura M, Wada T, Seio K, Moriguchi T, Huber J, Luhrmann R, Sekine M. Tetrahedron Lett. 2001; 42:8853.

8. Ohkubo A, Sasaki K, Noma Y, Tsunoda H, Seio K, Sekine M. *Bioorg Med Chem.* 2009; 17:4819. [PubMed: 19493681]
9. Jemielity J, Heinonen P, Lonnberg H, Darzynkiewicz E. *Nucleosides Nucleotides Nucleic Acids.* 2005; 24:601. [PubMed: 16247996]
10. Sawai H, Wakai H, Nakamura-Ozaki A. *J Org Chem.* 1999:5836.
11. (a) Sekine M, Nishiyama S, Kamimura T, Osaki Y, Hata T. *Bull Chem Soc Jpn.* 1985; 58:850.(b) Sekine M, Iwase R, Hata T, Miura K. *J Chem Soc Perkin Trans.* 1989; 1:969.(c) Sekine M, Kadokura M, Satoh T, Seio K, Wada T, Fischer U, Sumpter V, Luhrmann R. *J Org Chem.* 1996; 61:4412. [PubMed: 11667346]
12. Thillier Y, Decroly E, Morvan F, Canard B, Vasseur J-J, Debart F. *RNA.* 2012; 18:856. [PubMed: 22334760]
13. **General procedure for the preparation of chemically capped WT and mutant Spliced Leader.** GGUUUAAUUACCCAAGUUUGAG+/-biotin (0.2-0.4 μ mol) or ACUUAUUUACCCAAGUUUGAG+/-biotin (0.2 μ mol) and $\text{MnCl}_2 \cdot 4\text{H}_2\text{O}$ (4 mg, 0.02 mmol) were stirred in 0.1 M *N*-ethylmorpholine buffer pH 6 (0.15 mL) with imidazolide 5'-diphosphate derivatives [imGDP, im(m^7 GDP) or im($m_3^{2,2,7}$ GDP)], for their preparation – see the Supplementary Material] (0.02 mmol) at room temperature for 36 h. The reaction was spun and the precipitate, containing only capped and uncapped oligonucleotide, was dissolved in an aqueous solution of EDTA (11.5 mg, 0.03 mmol) and subjected to HPLC purification. Capped oligonucleotides were separated and analyzed using a Knauer chromatograph equipped with a Supelcosil LC-18-T column (4.6 $\times$ 250 mm). The following mobile phases at flow rates of 1 mL/min were used: 0.1 M triethylammonium acetate (TEAA), pH7 (buffer A), and 25% CH_3CN in 0.1 M TEAA, pH 7 (buffer B). The 0.1 M TEAA buffer, pH 7, was prepared by mixing the appropriate amount of Et_3N and AcOH in H_2O and carefully adjusting to pH 7 using Et_3N or AcOH. The G- or m^7 G-capped spliced leaders were purified using a linear gradient of buffer B from 0 to 100% in buffer A in 20 min. $m_3^{2,2,7}$ G-capped SL was purified using a linear gradient of buffer B from 20% to 50% in buffer A in 30 min. Capped mutant SLs were purified at 50 $^\circ\text{C}$. UV detection was performed at 260 nm. Fractions containing pure product were repeatedly lyophilized and MS spectra of the obtained capped-RNAs were acquired using a Waters Micromass Q-TOF Premier spectrometer with a negative electrospray ionization source (see Table 1).
14. (a) Jankowska-Anyszka M, Piecyk K, Samonina-Kosicka. *J Org Biomol Chem.* 2011; 9:5564.(b) Piecyk K, Davis RE, Jankowska-Anyszka M. *Bioorg Med Chem.* 201210.1016/j.bmc.2012.05.078
15. Gilar M, Fountain KJ, Budman Y, Neue UD, Yardley KR, Rainville PD, Russell RJ II, Gebler JC. *J Chromatogr A.* 2002; 958:167. [PubMed: 12134814]
16. (a) Song R, Zhang W, Chen H, Ma H, Dong Y, Sheng G, Zhou Z, Fu J. *Eur J Mass Spectrom.* 2005; 11:83.(b) Kuznetsov SV, Shen Y, Benight AS, Ansari A. *Biophys J B.* 2001; 81:2864.

Figure 1.

Structure of the mRNA 5' end. (A) non-*trans*-spliced mRNAs have an MMG cap, which consists of a 7-methylguanosine connected *via* a 5',5'-triphosphate linkage to the first transcribed nucleotide of the RNA chain, (B) *trans*-spliced mRNA with a TMG cap, which consists of *N*²,*N*²-7-trimethylguanosine linked *via* a 5',5'-triphosphate linkage to the 22 nucleotide spliced leader (SL). (C) Predicted 5' stem-loop in *Ascaris* SL.⁵

Figure 2.

HPLC profiles of the separation of capped/uncapped WT and mutant SL RNAs.

Supercosil LC-18-T, buffer A: 0.1M TEAA, pH 7, buffer B: 15% CH₃CN in 0.1M TEAA, pH 7

^a linear gradient of buffer B from 0 to 100% in buffer A in 20 min

^b linear gradient of buffer B from 20% to 50% in buffer A in 30 min

^c HPLC analysis was performed at 50°C

Scheme 1.

General course of the synthesis of 5'-terminal capped spliced leader RNAs. (i) $\text{MnCl}_2 \cdot 4\text{H}_2\text{O}$, 0.1 M *N*-ethylmorpholine buffer, pH 6.

Table 1

Entry	5'-3' sequence	Molecular formula	Calc. (m/z) ^a	Found (m/z) ^a	Yield (%) ^b
1	p-GGUUAAUUACCCAAAGUUUGAG	C ₂₀₉ H ₂₅₈ N ₈₀ O ₁₅₇ P ₂₂	7084.2	-	-
2	m ⁷ GpppGGUUUAAUUACCCAAAGUUUGAG	C ₂₂₀ H ₂₇₄ N ₈₅ O ₁₆₇ P ₂₄	7524.4	7524.0	66
3	m ₃ ^{2,2',7} GpppGGUUUAAUUACCCAAAGUUUGAG	C ₂₂₂ H ₂₇₈ N ₈₅ O ₁₆₇ P ₂₄	7552.4	7552.5	37
4	pGGUUUAAUUACCCAAAGUUUGAG-biotin	C ₂₂₅ H ₂₈₇ O ₁₆₄ P ₂₃ N ₈₅ S ₁	7522.6	-	-
5	GpppGGUUUAAUUACCCAAAGUUUGAG-biotin	C ₂₃₅ H ₃₀₀ O ₁₇₄ P ₂₅ N ₈₈ S ₁	7947.8	7947.6	69
6	m ⁷ GpppGGUUUAAUUACCCAAAGUUUGAG-biotin	C ₂₃₆ H ₃₀₃ O ₁₇₄ P ₂₅ N ₈₈ S ₁	7962.9	7961.5	59
7	m ₃ ^{2,2',7} GpppGGUUUAAUUACCCAAAGUUUGAG-biotin	C ₂₃₈ H ₃₀₇ O ₁₇₄ P ₂₅ N ₈₈ S ₁	7990.9	7990.0	34
8	p-GACUUAAUUACCCAAAGUUUGAG	C ₂₀₉ H ₂₅₉ N ₈₁ O ₁₅₅ P ₂₂	7067.2	-	-
9	m ⁷ GpppGACUUAAUUACCCAAAGUUUGAG	C ₂₂₀ H ₂₇₅ N ₈₆ O ₁₆₅ P ₂₄	7507.4	7507.5	26
10	m ₃ ^{2,2',7} GpppGACUUAAUUACCCAAAGUUUGAG	C ₂₂₂ H ₂₇₉ N ₈₆ O ₁₆₅ P ₂₄	7535.5	7535.7	6
11	pGACUUAAUUACCCAAAGUUUGAG-biotin	C ₂₂₅ H ₂₈₉ O ₁₆₂ P ₂₃ N ₈₄ S ₁	7506.7	-	-
12	GpppGACUUAAUUACCCAAAGUUUGAG-biotin	C ₂₃₅ H ₃₀₂ O ₁₇₂ P ₂₅ N ₈₉ S ₁	7931.8	7931.8	40
13	m ⁷ GpppGACUUAAUUACCCAAAGUUUGAG-biotin	C ₂₃₆ H ₃₀₅ O ₁₇₂ P ₂₅ N ₈₉ S ₁	7946.9	7946.6	25
14	m ₃ ^{2,2',7} GpppGACUUAAUUACCCAAAGUUUGAG-biotin	C ₂₃₈ H ₃₀₉ O ₁₇₂ P ₂₅ N ₈₉ S ₁	7974.9	7974.6	4

^aTOF-MS characterization in negative mode^bPercentage yield of isolated capped-RNAs