



**HAL**  
open science

## Access to High-Purity 7m G-cap RNA in Substantial Quantities by a Convenient All-Chemical Solid-Phase Method

Mathieu Noël, Theo Guez, Yann Thillier, Jean-jacques Vasseur, Françoise Debart

► **To cite this version:**

Mathieu Noël, Theo Guez, Yann Thillier, Jean-jacques Vasseur, Françoise Debart. Access to High-Purity 7m G-cap RNA in Substantial Quantities by a Convenient All-Chemical Solid-Phase Method. *ChemBioChem*, 2023, 24 (2), pp.756-765. 10.1002/cbic.202300544 . hal-04235335

**HAL Id: hal-04235335**

**<https://cnrs.hal.science/hal-04235335>**

Submitted on 10 Oct 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

DOI: 10.1002/cbic.202300544

Research Article

Received: 02.August 2023

Revised: 24.August 2023

Accepted Article published: 04.September 2023

## Access to High-Purity <sup>7m</sup>G-cap RNA in Substantial Quantities by a Convenient All-Chemical Solid-Phase Method

Mathieu Noël,<sup>[a]</sup> Theo Guez,<sup>[a]</sup> Dr. Yann Thillier,<sup>[a,b]</sup> Dr. Jean-Jacques Vasseur<sup>0000-0002-4379-6139,</sup><sup>[a]</sup> Dr. Françoise Debart<sup>0000-0003-3422-3926\*</sup><sup>[a]</sup>

[a] <orgName/>Institut des Biomolécules Max Mousseron  
<orgDiv/>UMR 5247 CNRS-University of Montpellier-ENSCM, Equipe  
ChemBioNAC  
<orgDiv/>Pôle Chimie Balard Recherche  
<street/>1919 Route de Mende, <postCode/>34293 <city/>Montpellier Cedex 5  
(<country/>France)  
E-mail: francoise.debart@umontpellier.fr

[b] Present address: <orgName/>Chemgenes  
<street/>900 Middlesex Turnpike, <city/>Billerica, <postCode/>MA 01821  
(<country/>USA)

VIP

<spi> This article is part of the SCF-ChemBio: Chemical Biology Tour de France Special Collection. Please see our <url href="https://chemistry-europe.onlinelibrary.wiley.com/topic/vi-categories-14397633/special-collections/14397633">homepage</url> for more articles in the collection.

< pict> Supporting information for this article is available on the WWW under <url>http://dx.doi.org/10.1002/cbic.202300544</url>

A practical and valuable all-chemical method has been developed to synthesize 5'-<sup>7m</sup>Gppp RNA on solid-support without damaging the <sup>7m</sup>G-cap or the RNA during RNA deprotection and release from the support under mild basic conditions. Substantial quantities of high-purity 5'-<sup>7m</sup>Gppp RNA are thus affordable as useful research tools.

Given the importance of mRNA with 5'-cap, easy access to RNA substrates with different <sup>7m</sup>G caps, of high quality and in large quantities is essential to elucidate the roles of RNA and the regulation of underlying processes. In addition to existing synthetic routes to 5'-cap RNA based on enzymatic, chemical or chemo-enzymatic methods, we present here an all-chemical method for synthetic RNA capping. The novelty of this study lies in the fact that the capping reaction is performed on solid-support after automated RNA assembly using commercial 2'-*O*-propionyloxymethyl ribonucleoside phosphoramidites, which enable final RNA deprotection under mild conditions while preserving both <sup>7m</sup>G-cap and RNA integrity. The capping reaction is efficiently carried out between a 5'-phosphoroimidazolide RNA anchored on the support and <sup>7m</sup>GDP in DMF in the presence of zinc chloride. Substantial amounts of <sup>7m</sup>G-cap RNA (from 1 to 28 nucleotides in length and of any sequence with or without internal methylations) containing various cap structures (<sup>7m</sup>GpppA, <sup>7m</sup>GpppA<sub>m</sub>, <sup>7m</sup>Gppp<sup>m6</sup>A, <sup>7m</sup>Gppp<sup>m6</sup>A<sub>m</sub>, <sup>7m</sup>GpppG, <sup>7m</sup>GpppG<sub>m</sub>) were obtained with high purity after IEX-HPLC purification. This capping method using solid-phase chemistry is convenient to perform and provides access to valuable RNA substrates as useful research tools to unravel specific issues regarding cap-related processes.

cap structure

nucleic acids

ribonucleosides

RNA

solid-phase synthesis

@DebartFrancoise; @VasseurJeanJacq; @ChemBioNAC; @IBMM<?\_>Balard;

@INC<?\_>CNRS

## Introduction

The cap is a co-transcriptional modification at the 5'-end of eukaryotic and viral messenger RNA (mRNA).<sup>[1]</sup> The chemical structure of the cap consists of an *N*<sup>7</sup>-methylguanosine (<sup>7</sup>mG) connected to the first transcribed nucleotide (N1) at the 5'-end of the RNA by a 5'-5' triphosphate bridge (Figure<sup>1</sup>).<sup>[2]</sup> This so-called cap-0 (<sup>7</sup>mGpppN1) can undergo further methylations. The most common one occurs at the 2'-hydroxyl position of the ribose of the first and sometimes of the second nucleotide, resulting in cap-1 (<sup>7</sup>mGpppN1<sub>m</sub>) and cap-2 (<sup>7</sup>mGpppN1<sub>m</sub>N2<sub>m</sub>), respectively. These RNA structural modifications are crucial for efficient mRNA translation, protection against premature degradation by 5'-exonucleases and recognition by the innate immune system. Additional methylation at the *N*<sup>6</sup> position of adenosine has been detected in some cap structures if the first encoded nucleotide adjacent to the <sup>7</sup>mGppp cap is 2'-*O*-methyladenosine (*A*<sub>m</sub>), resulting in *N*<sup>6</sup>,2'-*O*-dimethyladenosine (<sup>m6</sup>*A*<sub>m</sub>) (Figure<sup>1</sup>). This double-methylated nucleotide <sup>m6</sup>*A*<sub>m</sub> present in the cap of ~30% of all cellular mRNAs has been shown to be a dynamic and reversible epitranscriptomic modification that confers stability to mRNA in mammalian cells.<sup>[3]</sup> Since these findings in 2017, the functions of *N*<sup>6</sup>-methylation have been widely investigated.<sup>[4]</sup>

Besides cellular events, RNA viruses have evolved strategies to synthesize a cap that mimics the cellular cap; thus, viral RNA can escape detection by the sensors of the host innate immune system. Consequently, many viral capping enzymes are unique in nature and have no counterparts in infected cells. These proteins are therefore attractive targets for antiviral drug design, as an antiviral drug can be expected to have good selectivity.<sup>[5]</sup> To accelerate drug design, a thorough understanding of these viral capping enzymes at the molecular and atomic level is required, and several RNA virus enzymes have already been characterized for this purpose.<sup>[6]</sup> The RNA capping machinery has been shown to be remarkably diverse structurally and functionally, reinforcing the interest of these enzymes for drug-design. However, the characterization and use of the viral RNA capping system as an antiviral target has lagged far

behind, impeding accurate characterization of crystallographic binary complexes and structure-based drug design.

Indeed, the availability of 5'-capped RNA substrates of any length, sequence, and structure in large amounts and high purity remains the main bottleneck for unraveling the roles of RNA and the regulation of underlying processes. Nevertheless, different routes for the synthesis of 5'-capped mRNAs based on enzymatic, chemical or chemo-enzymatic methods have been developed with varying difficulties and limitations.<sup>[7]</sup>

A few years ago, several chemical breakthroughs in our group helped to overcome this bottleneck. First, in 2008, we developed an original method for solid-phase synthesis of RNA oligonucleotides using only base-labile groups for the protection of nucleobases, phosphates, and 2'-hydroxyl functions.<sup>[8]</sup> The innovation is based on the use of the pivaloyloxymethyl (PivOM) group and other acyloxymethyl groups at 2'OH,<sup>[9]</sup> which enable concomitant deprotection and cleavage of the oligonucleotide from the solid support by mild aqueous ammonia treatment (3<sup>h</sup> at room temperature), without degrading the RNA. This method represented a major step forward because, in addition to facilitating RNA preparation, the deprotection conditions were totally compatible and applicable to solid-support synthesis of RNA with a 5'-triphosphate (TP) moiety<sup>[10]</sup> or a *N*<sup>7</sup>-unmethylated cap structure (Gppp).<sup>[11]</sup> Over the past decade, we have been faced with a high demand for synthetic 5'-cap RNAs from worldwide collaborators aimed at elucidating the mechanism, function, and structure of the capping system associated with methyltransferases in eukaryotes and viruses. These useful research tools have proven essential to unravel specific issues regarding cap-related processes.<sup>[3,12]</sup>

Unlike enzymatic preparation of capped mRNA by *in vitro* transcription of a DNA template, chemical synthesis offers the possibility to reliably introduce modified nucleotides at specific positions of the RNA sequence, in particular 2'-*O*-methyl ribonucleotides or *N*<sup>6</sup>-

methyl adenosine. Apart from our method combining chemical synthesis and enzymatic *N*<sup>7</sup>-methylation,<sup>[11]</sup> a few examples of capped RNA prepared by solid-phase synthesis have been reported in the literature.<sup>[13]</sup> Recently, Abe et al. reported a complete chemical process for cap RNA synthesis.<sup>[14]</sup> In this process, 5'-phosphorylated RNA are elongated by solid-phase synthesis, deprotected, and removed from the solid support before being coupled in solution with activated *N*<sup>7</sup>-guanosine diphosphate (Im-<sup>7</sup>mGDP) in DMSO. Several chemically modified 107-nt RNAs with cap-0, cap-1, cap-2, and cap-4 were prepared in small amounts (100 picomoles) using phosphoramidite chemistry and after capping, their translation activity was evaluated in HeLa cells. However, the insolubility of the different compounds in organic solvents, the salt dependence of the coupling efficiencies and the fact that the uncapped RNAs were inseparable from the capped RNAs are obstacles to the systematic use of this method. To overcome this drawback, the same authors have developed PureCap technology, which involves introducing hydrophobic photocaged tag-modified cap analogs into the mRNA 5'-end during the *in vitro* transcription.<sup>[15]</sup> This hydrophobicity enables the separation of capped mRNA from uncapped mRNA by RP-HPLC purification. Subsequent photoirradiation removes the hydrophobic tag. Thus, a cap-2 mRNA of 4247-nt was successfully prepared following this versatile approach.

With the aim to improve the efficiency and rapidity of <sup>7</sup>mG-cap RNA preparation, we propose an all-chemical method for the large-scale preparation of short <sup>7</sup>mG-cap RNA (up to 28 nucleotides long) and various derivatives where RNA elongation, its first phosphorylation, and coupling of *N*<sup>7</sup>-methyl guanosine diphosphate (<sup>7</sup>mGDP) are performed on solid support. This method could address the limitations reported above with the Abe method.<sup>[14]</sup>

## Results and Discussion

The all-chemical synthesis of 5'-<sup>7</sup>mG-cap RNA using a solid-phase process faces technical limitations due to the instability of *N*<sup>7</sup>-methylguanosine under acidic (glycosidic

bond cleavage) or basic (G imidazole ring opening) conditions. This fragility remains the main obstacle, especially during the final deprotection of the 7<sup>m</sup>G-cap RNA under standard 28% aqueous ammonia treatment. Therefore, to preserve the integrity of the 7<sup>m</sup>G-cap structure, we sought a milder basic medium to remove the RNA protecting groups and release the 7<sup>m</sup>G-cap RNA from the solid support. Anhydrous amine reagents used neat or dissolved in an organic solvent have been described as an alternative method for cleavage and deprotection of synthetic oligonucleotides.<sup>[16]</sup> Indeed, in a previous study, we found that a butylamine/THF solution (1:1, v/v) could remove the phenoxyacetyl group from adenosine (A) and cleave the succinyl linker between short RNA 2'-5' tetramers and the support at room temperature for 1 h while retaining 2'-*O*-PivOM groups.<sup>[17]</sup> Thus, milder basic conditions for RNA deprotection imply the use of 2'-*O* protecting groups that are easier to remove than PivOM under basic conditions. The stability of various acyloxymethyl groups (acetyloxymethyl, propionyloxymethyl, butyryloxymethyl, and isobutyryloxymethyl) in oligouridylates was then evaluated over a set of deprotection conditions based on different anhydrous alkylamines in organic solvents (THF, toluene).<sup>[18]</sup> From this evaluation, the propionyloxymethyl (PrOM) group was preferred for its complete removal upon treatment with anhydrous butylamine solution in THF (2:1, v/v) for 5 h at 30°C. Based on these preliminary data, we determined that an all-chemical synthesis of 7<sup>m</sup>G-cap RNA on solid support would be possible.

Therefore, we followed the strategy developed previously to introduce the G-cap at the 5'-end of the RNA.<sup>[11]</sup> After RNA assembly on an automated synthesizer using 2'-*O*-PrOM ribonucleoside phosphoramidites,<sup>[18]</sup> the 5'-hydroxyl group of the 5'-terminal nucleoside of the RNA sequences, still anchored to the solid support, was converted into an *H*-phosphonate derivative with diphenylphosphite (Scheme 1).<sup>[10]</sup> Then, oxidation and activation under strict anhydrous conditions with bromotrichloromethane (BrCl<sub>3</sub>C) in the presence of imidazole and *N*, *O*-bis-trimethylsilylacetamide yielded the 5'-phosphoroimidazolide RNA.

Imidazolide phosphate was reacted with the tri-*n*-butylammonium salt of <sup>7m</sup>GDP in DMF in the presence of ZnCl<sub>2</sub> for 18<sup>h</sup>. Finally, the cyanoethyl groups were eliminated from the phosphates by a short DBU treatment (3<sup>min</sup>), followed by removal of the *N*-acyl and 2'-*O*-PrOM groups, and cleavage of the succinyl linker to the solid-support with a solution of anhydrous butylamine in THF (2<sup>:1</sup>, v/v) for 5<sup>h</sup> at 30<sup>°C</sup>. In this deprotection protocol, the <sup>7m</sup>G-cap RNA that were cleaved from the support were retained in the glass beads. After washing the beads with THF, a buffered aqueous solution was used to collect the <sup>7m</sup>G-cap RNA in the solution. Both organic and aqueous solutions were analyzed by IEX-HPLC and, depending on the RNA sequence length, crude <sup>7m</sup>G-cap RNA material could be partially found in the organic butylamine/THF solution, which was observed especially for short sequences (up to 6-mer). Following this method, several <sup>7m</sup>G-cap RNA (from 3 to 13 nucleotides-long) were isolated with high purity after HPLC purification and without degradation of the <sup>7m</sup>G-cap. Nevertheless, the method showed several weaknesses such as incomplete removal of PrOM groups from RNA sequences longer than 10 nucleotides. MALDI-TOF MS analysis showed additional peaks with higher mass corresponding to 2 or 3 remaining PrOM groups randomly distributed in the sequence resulting in a mixture of RNA species. Thus, the butylamine/THF ratio (from 2<sup>:1</sup> to 4<sup>:1</sup> v/v) and the deprotection time (from 2<sup>h</sup> to 4<sup>h</sup>) were adjusted according to the length of the RNA sequence (related to the number of PrOM groups (3 to 13) to be removed). Moreover, the crude RNA material (short oligomers <10 nucleotides) was not efficiently retained in the beads after cleavage from the support, resulting in loss of material in the organic solution.

With the objective of strengthening <sup>7m</sup>G-cap RNA synthesis, we aimed to improve the conditions for RNA deprotection by using a 7<sup>N</sup> ammonia solution in methanol, instead of a butylamine/THF solution. Similarly, this treatment allowed the complete removal of the *N*-acyl and 2'-*O*-PrOM protecting groups and cleavage of RNA from CPG beads in 3<sup>h</sup> at 40<sup>°C</sup> without degrading the <sup>7m</sup>G-cap. Large amounts of <sup>7m</sup>G-cap RNA crude material (from



1 to 28 nucleotides in length and of various sequences with or without internal methylations) containing diverse caps ( $^7\text{mGpppA}$ ,  $^7\text{mGpppA}_m$ ,  $^7\text{mGppp}^{\text{m6}}\text{A}$ ,  $^7\text{mGppp}^{\text{m6}}\text{A}_m$ ,  $^7\text{mGpppG}$ ,  $^7\text{mGpppG}_m$ ), were collected in methanolic ammonia solution that was simply evaporated under reduced pressure prior to IEX-HPLC analysis and purification (Table<sup>1</sup>, Supporting information Figures<sup>S3-S6</sup> for some examples). All  $^7\text{mG}$ -cap RNA **1--35** and various cap structures **36--41** were characterized by mass spectrometry MALDI-TOF (Supporting information Figures<sup>S7-S47</sup>) and were quantified by UV-spectrometry at 260 nm. It is noteworthy that to prepare RNA with  $^7\text{mGppp}^{\text{m6}}\text{A}$  **16--17** and  $^7\text{mGppp}^{\text{m6}}\text{A}_m$  **29--35**, we needed to prepare the corresponding *N*-acyl 2'-*O*-PrOM and 2'-OMe<sup>m6</sup>A phosphoramidites using a convenient procedure of selective one-step *N*6-methylation of the 2'-*O*-PrOM and 2'-OMe phosphoramidite building blocks.<sup>[19]</sup> In addition, our capping method provided easy access to tens to hundreds of nanomoles of several natural cap structures and analogues **36--41** that could be useful for the 5-functionalization of long RNA.

Finally, a more cost-effective synthesis of  $^7\text{mGpppRNA}$  was achieved by using the tri-*n*-butylammonium salt of *N*7-methyl-guanosine diphosphate made in-house instead of purchasing the expensive commercially available  $^7\text{mGDP}$  sodium salt. First, large amounts of GDP sodium salt were converted to GDP tri-*n*-butylammonium salt on which methylation was performed with methyl iodide (6<sup>eq.</sup>) in DMF for 2 h at room temperature in 63% yield (Scheme<sup>2</sup>).<sup>[20]</sup>

## Conclusions

Although the recent COVID-19 pandemic **has sparked** **correct wording** renewed interest in mRNA, which can be used for therapeutic applications or as a vaccine, many studies involving mRNA had already been underway for years and required chemically synthesized RNA in large quantities and of high purity. To this end, to prepare synthetic RNA with various 5'-cap structures dedicated to elucidating the roles of

RNA and the regulation of underlying processes, we have been working for over a decade on automated RNA synthesis and capping reaction using solid-phase chemistry. In this study, we have presented an optimized all-chemical method for synthesizing 5'-<sup>7m</sup>Gppp RNA using solid-support chemistry without damaging the <sup>7m</sup>G-cap or the RNA during RNA deprotection and release from the support. To do so, 2'-*O*-PrOM ribonucleoside building blocks protected by fast-labile groups were used for RNA assembly, and mild basic conditions were found to release free <sup>7m</sup>Gppp RNA efficiently and safely. This improved synthetic methodology allows simple and convenient access to a wide range of tens to hundreds of nanomoles of 5'-cap RNA substrates, of high purity (>95%), of any sequence with natural or modified nucleotides, of any length (up to 28 nucleotides). The availability of such substrates will facilitate further research in the fields of viral methyltransferase mechanisms, crystallography, and even the design of selective antiviral drugs.

## Experimental Section

### General methods

Triethylamine was distilled from calcium hydride, other anhydrous reagents were purchased from Aldrich without further purification. Oligonucleotide synthesis reagents, LCAA-CPG solid support, 2'-*O*-PrOM and 2'-*O*-Me ribonucleoside phosphoramidites were purchased from ChemGenes. Guanosine diphosphate sodium salt was purchased from Carbosynth.

Analytical and semi-preparative high-performance anion-exchange and reverse-phase chromatographies were performed on a ThermoScientific Ultimate 3000 system equipped with an LPG3400RS pump or HPG3200BX pump, a DAD 3000 detector and a WPS-3000TBRS autosampler, column oven TCC-3000SD. IEX-HPLC was performed on DNAPac PA200 columns (4x250<sup>mm</sup>, 9x250<sup>mm</sup>, or 22x250<sup>mm</sup>, Thermo Scientific). The following HPLC solvent systems were used: 5% CH<sub>3</sub>CN in 25<sup>mM</sup> Tris-HCl buffer pH<sup>8</sup> (eluent A) and 5% CH<sub>3</sub>CN containing 400<sup>mM</sup> of

NaClO<sub>4</sub> in 25 mM Tris-HCl buffer pH 8 (eluent B). Flow rates were 1 mL/min, 4 mL/min, or 9 mL/min for analytical, semi-preparative and preparative purposes, respectively.

Reverse-phase C<sub>18</sub> HPLC analysis was conducted on a Nucleodur column (C<sub>18</sub> 3 μm, 4.6×75 mm, Macherey Nagel) column. The following HPLC solvent systems were used: 1% CH<sub>3</sub>CN in 50 mM TEAAc (eluent C) and 80% CH<sub>3</sub>CN in 50 mM TEAAc (eluent D). The flow rate was 1 mL/min. UV detection was performed at 260 nm. Chromeleon software was used.

MALDI-TOF mass spectra were recorded on an AXIMA Assurance (Shimadzu), using a saturated solution of 2,4,6-trihydroxyacetophenone in a solution of acetonitrile/ammonium citrate (0.1 M) (1:1; v/v). The samples were mixed with the matrix in a 1:1 ratio (v/v) crystallized on 384-well stainless steel plate and analyzed.

UV quantification of RNAs was performed on a VWR UV-1600 PC UV/visible spectrophotometer by measuring absorbance at 260 nm and using quartz cuvettes.

RMN analyses were performed on a Bruker 400 MHz and spectra were edited with MestReNova.

### **Preparation of <sup>7</sup>mGDP bis(tributylammonium) salt**

To a solution of guanosine diphosphate bis(tributylammonium) salt (550 mg, 0.62 mmol, 1.0 eq) prepared from commercial guanosine diphosphate sodium salt<sup>[11]</sup> in anhydrous DMF (5 mL) was added methyl iodide (251 μL, 4.06 mmol, 6.0 eq). The reaction mixture was stirred at room temperature for 2 h. The solution turned yellow and a gum or a precipitate appeared. The reaction was monitored by reverse-phase C<sub>18</sub> HPLC with a 10 min isocratic elution (eluent C) then a 10 min linear gradient of 10% eluent D in eluent C. The reaction was quenched by adding water (20 mL) in case of gum, or the precipitate was filtered off (<sup>7</sup>mGDP bis(tributylammonium) is not soluble in anhydrous DMF). The aqueous layer was washed 2 times with diethyl ether, and concentrated under vacuum. The pale yellow crude material was purified by flash column chromatography on a reverse-

phase silica gel column C<sub>18</sub> (120<sup>μ</sup>g, 40<sup>μ</sup>m) with an isocratic elution of 1% CH<sub>3</sub>CN in 30<sup>mM</sup> TEAAc buffer during 8<sup>CV</sup> then a 0–5% linear gradient of 5% CH<sub>3</sub>CN in 30<sup>mM</sup> TEAAc buffer for 5<sup>CV</sup>. The fractions containing the pure compound were pooled and concentrated under vacuum. The residue was taken up in water, passed through a Dowex-50<sup>W</sup> X8(H) resin column<sup>[11]</sup> and the resulting solution was lyophilized from water to obtain a white powder (324<sup>mg</sup>, 0.39<sup>mmol</sup>, 63%). The product was characterized by <sup>31</sup>P NMR and <sup>1</sup>H NMR.

### Automated synthesis of RNA oligonucleotides

RNA synthesis was performed on an ABI 394 synthesizer (Applied Biosystems) from long chain alkylamine controlled-pore glass (LCAA-CPG) solid support with a pore size of 1000<sup>Å</sup> derivatized through the succinyl linker with 5'-*O*-dimethoxytrityl-2'-*O*-acetyl-(uridine or *N*4-acetylcytidine or *N*2-isopropylphenoxyacetyl guanosine or *N*6-phenoxyacetyl adenosine) (Link Technologies) or with biotin 3'-LCAA-CPG 1000<sup>Å</sup> (Chemgenes). RNA sequences were assembled on a 1<sup>μ</sup>mol scale in Twist oligonucleotide synthesis columns (Glen Research) using various building blocks: 5'-*O*-DMTr-2'-*O*-propionyloxymethyl-3'-*O*-(*O*-cyanoethyl-*N,N*-diisopropylphosphoramidite)-[uridine, *N*4-acetylcytidine, *N*2-isopropylphenoxyacetyl guanosine or *N*6-phenoxyacetyl adenosine], 5'-*O*-DMTr-2'-*O*-methyl-3'-*O*-(*O*-cyanoethyl-*N,N*-diisopropylphosphoramidite)-[*N*6-phenoxyacetyl adenosine or *N*2-isopropylphenoxyacetyl guanosine] (Chemgenes). The 5' terminal adenosine can be unmodified A, or methylated in 2-OH (A<sub>m</sub>), or in *N*6 position (m<sup>6</sup>A) or in both positions (m<sup>6</sup>A<sub>m</sub>). The 5'-*O*-DMTr-2'-*O*-Me-3'-*O*-(*O*-cyanoethyl-*N,N*-diisopropylphosphoramidite)-A<sup>Pac</sup> (Chemgenes) was used to introduce A<sub>m</sub> at the 5'-end of RNA or internal positions. For synthesis of m<sup>6</sup>A-containing RNAs or m<sup>6</sup>A<sub>m</sub>-containing RNAs, the preparation of m<sup>6</sup>A and m<sup>6</sup>A<sub>m</sub> phosphoramidite building blocks was performed by a selective one-step methylation<sup>[19]</sup> of the commercially available 2'-*O*-PrOM-A<sup>Pac</sup>-CE phosphoramidite or 2'-*O*-Me-A<sup>Pac</sup>-CE phosphoramidite, respectively. Synthesis and data on 2'-*O*-PrOM-m<sup>6</sup>A<sup>Pac</sup>-CE phosphoramidite are presented in the supporting information.

### Synthesis of solid-supported 5'-O-phosphorimidazolide RNA

After assembly, the CPG beads were dried under a stream of argon. A solution (2 mL) of 1 M diphenyl phosphite (0.4 mL) in dry pyridine (1.6 mL) with activated 4 Å molecular sieves was passed manually (with a plastic syringe) through the Twist column and left to stand for 30 min at 40°C. The CPG was then washed with acetonitrile and a 0.1 M solution of triethylammonium bicarbonate (TEAB, pH 7.5) was applied to the column and left to react for 45 min at 40°C. After several washings and drying under argon, an oxidation solution containing imidazole (150 mg) in acetonitrile (0.8 mL), bromotrichloromethane (0.8 mL), triethylamine (0.1 mL) and *N,O*-bis-trimethylsilylacetamide (0.4 mL) with activated 3 Å molecular sieves, were added under argon and left to react for 2 h at 40°C. The solution was removed and the support was washed with 3 × 4 mL of anhydrous CH<sub>3</sub>CN and dried under argon.

### Capping reaction with <sup>7m</sup>GDP bis(tributylammonium) salt

In a dry 4-mL screw-capped glass vial, <sup>7m</sup>GDP bis(tributylammonium) salt (45 mg, 54 μmol), zinc chloride (60 mg, 0.44 mmol), and activated 4 Å molecular sieves were mixed in anhydrous DMF (0.5 mL). This solution was passed manually through the Twist column and was left for 17 hours at 40°C. The CPG was washed with dry DMF (1 × 4 mL), then a 0.1 M solution of EDTA disodium salt (3 × 4 mL), water (3 × 4 mL), anhydrous acetonitrile (3 × 4 mL), and then dried under argon.

### Deprotection and release of <sup>7m</sup>GpppRNAs from solid support

The solid-supported RNAs were deprotected and cleaved from the support as follows. First, a 1 M solution of diazobicyclo-(5,4,0)-undec-7-ene (DBU) in anhydrous CH<sub>3</sub>CN was applied through the column for 3 min. Then beads were transferred in a screw-capped glass vial and a 7 N methanolic ammonia solution (3 mL) was added. After 3 hours at 40°C, the mixture was filtered and the filtrate was recovered into a round-bottom flask 1 (25 mL), the beads were washed twice

with anhydrous CH<sub>3</sub>CN. In a second round-bottom flask 2 (25<sup>^</sup>mL), the beads were washed with a 0.1<sup>^</sup>M solution of TEAAc (pH<sup>^</sup>7) then with water twice. Isopropylamine (50<sup>^</sup>μL) was added in each round-bottom flask,<sup>[8]</sup> and the solutions were evaporated and co-evaporated with water under reduced pressure with a water bath at 30<sup>^</sup>C maximum. The residues were dissolved in water, transferred in a 2<sup>^</sup>mL-Eppendorf<sup>™</sup> tube and lyophilized.

### **Analysis and purification of <sup>7</sup>mGppp-RNAs 1--41**

The crude <sup>7</sup>mGppp-RNAs 1--41 were analyzed by MALDI-TOF spectrometry and anion-exchange HPLC using a 0--15% or 5--30% linear gradient of buffer B in buffer A. The crude mixture was then purified by semi-preparative or preparative automatic IEX-HPLC using a 0--15% or 5--30% linear gradient of buffer B in buffer A at 25<sup>^</sup>C or 75<sup>^</sup>C. The pure fractions of <sup>7</sup>mGppp-RNAs were pooled in a 100<sup>^</sup>mL round-bottom flask and concentrated to reduced pressure at 30<sup>^</sup>C. The residue was dissolved in 10<sup>^</sup>mL of 100<sup>^</sup>mM TEAAc buffer (pH<sup>^</sup>7) and loaded on a Sep-Pack® classic cartridge (Waters). Desalting was performed with 10<sup>^</sup>mL of 100<sup>^</sup>mM TEAAc buffer then oligonucleotides were eluted with 10<sup>^</sup>mL of 80% CH<sub>3</sub>CN in water in a 25<sup>^</sup>mL round-bottom flask and lyophilized. The residue was dissolved in 1.5<sup>^</sup>mL of water and transferred to a 2<sup>^</sup>mL-Eppendorf<sup>™</sup> tube, lyophilized and stored at <M->20<sup>^</sup>C.

### **Acknowledgements**

The research was made possible by funding from the <cgs>European Union Seventh Framework Programme</cgs> (<cgn>FP7/2007-2013</cgn>) under <cgs>SILVER grant</cgs> agreement no.<cgn>260644</cgn>, from the <cgs>Agence Nationale pour la Recherche (ANR)</cgs> under SARS-RNA-SPA grant agreements no.<cgn>ANR-12-BSV3-0007-02</cgn> and under <cgs>DecRisp grant</cgs> agreement no.<cgn>ANR-19-CE11-0017-02</cgn>. Y.T. thanks the <cgs>University of Montpellier</cgs> for the financial support of his PhD work. M.N., J.J.V., F.D. are researchers from CNRS.

### **Conflict of Interests**

The authors declare no conflict of interest.

## Data Availability Statement

The data that support the findings of this study are available in the supplementary material of this article.

<lit1><lit\_a><jnl>A.<sup>^</sup>J. Shatkin, *Cell* **1976**, *9*, 645--653</jnl>; <lit\_b><jnl>E. Decroly, F. Ferron, J. Lescar, B. Canard, *Nat. Rev. Microbiol.* **2011**, *10*, 51--65</jnl>.

<lit2><jnl>Y. Furuichi, A.<sup>^</sup>J. Shatkin, *Adv. Virus Res.* **2000**, *55*, 135--184</jnl>.

<lit3><jnl>J. Mauer, X.<sup>^</sup>B. Luo, A. Blanjoie, X.<sup>^</sup>F. Jiao, A.<sup>^</sup>V. Grozhik, D.<sup>^</sup>P. Patil, B. Linder, B.<sup>^</sup>F. Pickering, J.<sup>^</sup>J. Vasseur, Q.<sup>^</sup>Y. Chen, S.<sup>^</sup>S. Gross, O. Elemento, F. Debart, M. Kiledjian, S.<sup>^</sup>R. Jaffrey, *Nature* **2017**, *541*, 371--375</jnl>.

<lit4><jnl>X. Jiang, B. Liu, Z. Nie, L. Duan, Q. Xiong, Z. Jin, C. Yang, Y. Chen, *Signal Transduct. Target Ther.* **2021**, *6*, 74</jnl>.

<lit5><jnl>E. Decroly, B. Canard, *Curr. Opin. Virol.* **2017**, *24*, 87--96</jnl>.

<lit6><jnl>R. Kasprzyk, J. Jemielity, *ChemBioChem* **2021**, *22*, 3236--3253</jnl>.

<lit7><jnl>F. Muttach, N. Muthmann, A. Rentmeister, *Beilstein J. Org. Chem.* **2017**, *13*, 2819--2832</jnl>.

<lit8><jnl>T. Lavergne, J.-R. Bertrand, J.-J. Vasseur, F. Debart, *Chem. Eur. J.* **2008**, *14*, 9135--9138</jnl>.

<lit9><jnl>A.<sup>^</sup>R. Martin, T. Lavergne, J.-J. Vasseur, F. Debart, *Bioorg. Med. Chem. Lett.* **2009**, *19*, 4046--4049</jnl>.

<lit10><jnl>I. Zlatev, T. Lavergne, F. Debart, J.-J. Vasseur, M. Manoharan, F. Morvan, *Org. Lett.* **2010**, *12*, 2190--2193</jnl>.

<lit11><jnl>Y. Thillier, E. Decroly, F. Morvan, B. Canard, J.-J. Vasseur, F. Debart, *RNA* **2012**, *18*, 856--868</jnl>.

- <lit12><lit\_a><jnl>G.<sup>^</sup>C. Paesen, A. Collet, C. Sallamand, F. Debart, J.-J. Vasseur, B. Canard, E. Decroly, J.<sup>^</sup>M. Grimes, *Nat. Commun.* **2015**, *6*, 8749</jnl>;
- <lit\_b><jnl>W. Aouadi, A. Blanjoie, J.<sup>^</sup>J. Vasseur, F. Debart, B. Canard, E. Decroly, *J. Virol.* **2017**, *91*, e02217--02216</jnl>; <lit\_c><jnl>A. Tamarkin-Ben-Harush, J.<sup>^</sup>J. Vasseur, F. Debart, I. Ulitsky, R. Dikstein, *eLife* **2017**, *6*, e21907</jnl>; <lit\_d><jnl>B. Martin, B. Coutard, T. Guez, G.<sup>^</sup>C. Paesen, B. Canard, F. Debart, J.-J. Vasseur, J.<sup>^</sup>M. Grimes, E. Decroly, *Nucleic Acids Res.* **2018**, *46*, 7902--7912</jnl>; <lit\_e><jnl>K. Boulias, D. Toczydlowska-Socha, B.<sup>^</sup>R. Hawley, N. Liberman-Isakov, K. Takashima, S. Zaccara, T. Guez, J.-J. Vasseur, F. Debart, L. Aravind, S.<sup>^</sup>R. Jaffrey, E. Lieberman<sup>^</sup>Greer, *Mol. Cell* **2019**, *75*, 631--643.e638</jnl>; <lit\_f><jnl>J. Mauer, M. Sindelar, V. Despic, T. Guez, B.<sup>^</sup>R. Hawley, J.-J. Vasseur, A. Rentmeister, S.<sup>^</sup>S. Gross, L. Pellizzoni, F. Debart, H. Goodarzi, S.<sup>^</sup>R. Jaffrey, *Nat. Chem. Biol.* **2019**, *340--347*</jnl>; <lit\_g><jnl>C. Valle, B. Martin, F. Debart, J.-J. Vasseur, I. Imbert, B. Canard, B. Coutard, E. Decroly, *J. Virol.* **2020**, *94*, e00520--00520</jnl>.
- <lit13><lit\_a><jnl>R. Iwase, M. Maeda, T. Fujiwara, M. Sekine, T. Hata, K. Miura, *Nucleic Acids Res.* **1992**, *20*, 1643--1648</jnl>; <lit\_b><jnl>M. Kadokura, T. Wada, K. Seio, T. Moriguchi, J. Huber, R. Luhrmann, M. Sekine, *Tetrahedron Lett.* **2001**, *42*, 8853--8856</jnl>; <lit\_c><jnl>I.<sup>^</sup>I. Koukhareva, A.<sup>^</sup>V. Lebedev, *Nucleosides Nucleotides Nucleic Acids* **2004**, *23*, 1667--1680</jnl>; <lit\_d><jnl>J. Jemielity, P. Heinonen, H. Lonngberg, E. Darzynkiewicz, *Nucleosides Nucleotides Nucleic Acids* **2005**, *24*, 601--605</jnl>; <lit\_e><jnl>S. Nagata, T. Hamasaki, K. Uetake, H. Masuda, K. Takagaki, N. Oka, T. Wada, T. Ohgi, J. Yano, *Nucleic Acids Res.* **2010**, *38*, 7845--7857</jnl>; <lit\_f><jnl>A. Ohkubo, Y. Kondo, M. Suzuki, H. Kobayashi, T. Kanamori, Y. Masaki, K. Seio, K. Nagai, M. Sekine, *Org. Lett.* **2013**, *15*, 4386--4389</jnl>; <lit\_g><jnl>J. Leiter, D. Reichert, A. Rentmeister, R. Micura, *ChemBioChem* **2020**, *21*, 265--271</jnl>.



<lit14><jnl>N. Abe, A. Imaeda, M. Inagaki, Z. Li, D. Kawaguchi, K. Onda, Y. Nakashima, S. Uchida, F. Hashiya, Y. Kimura, H. Abe, *ACS Chem. Biol.* **2022**, *17*, 1308--1314</jnl>.

<lit15><jnl>M. Inagaki, N. Abe, Z. Li, Y. Nakashima, S. Acharyya, K. Ogawa, D. Kawaguchi, H. Hiraoka, A. Banno, Z. Meng, M. Tada, T. Ishida, P. Lyu, K. Kokubo, H. Murase, F. Hashiya, Y. Kimura, S. Uchida, H. Abe, *Nat. Commun.* **2023**, *14*, 2657</jnl>.

<lit16><other>T. Kempe, **1998**</other>.

<lit17><jnl>Y. Thillier, S. Stevens, C. Moy, J. Taylor, J.-J. Vasseur, L. Beigelman, F. Debart, *Bioorg. Med. Chem.* **2013**, *21*, 5461--5469</jnl>.

<lit18><jnl>A. Biscans, M. Bos, A. Martin, N. Ader, G. Sczakiel, J.-J. Vasseur, C. Dupouy, F. Debart, *ChemBioChem* **2014**, *15*, 2674--2679</jnl>.

<lit19><jnl>S. Kruse, S. Zhong, Z. Bodi, J. Button, M. Alcocer, C. Hayes, R. Fray, *Sci. Rep.* **2011**, *1*, 126</jnl>.

<lit20><jnl>K. Piecyk, M. Lukaszewicz, E. Darzynkiewicz, M. Jankowska-Anyszka, *RNA* **2014**, *20*, 1539--1547</jnl>.

Figure<sup>1</sup> Chemical structure of various caps at the 5' end of RNA and potential methylation sites in the cap structure.

Scheme<sup>1</sup> On-column synthesis of <sup>7m</sup>G-cap RNA anchored to solid-support then deprotection and release of <sup>7m</sup>G-cap RNA.

Scheme<sup>2</sup> Preparation of large quantities of <sup>7m</sup>GDP tri-*n*-butylammonium salt from commercially available and affordable GDP sodium salt after a counterion exchange.

Table<sup>1</sup> List and data for synthesized <sup>7m</sup>G-cap RNA **1--41**.<w=3>

O N	L e	RNA Sequences (5' → 3')	Synt hesis	m/z th.	m/z exp.	Yield

	<u>length</u>		<u>scale</u> ( $\mu$ mol)	( <u>M</u> - <u>&gt;1H</u> )	<u>(neg)</u>	<u>nmoles</u>
<b>Cap-0 RNA</b>						
1	4-mer	<sup>7m</sup> GpppGAUC	1	1740.98	1740.97	88
2	4-mer	<sup>7m</sup> GpppAUAU	1	1726.98	1727.55	83
3	6-mer	<sup>7m</sup> GpppAUAUUA	1	2362.35	2362.44	72
4	7-mer	<sup>7m</sup> GpppAGUUGUU-biotin	4	3270.12	3270.31	171
5	9-mer	<sup>7m</sup> GpppGGGACAAAA	3	3443.07	3443.77	171
6	9-mer	<sup>7m</sup> GpppGGGACAAGU	3	3436.03	3436.39	140
7	10-mer	<sup>7m</sup> GpppAGUUGUUAGU	1	3681.11	3680.64	66
8	11-mer	<sup>7m</sup> GpppAAUAUAAUACC	1	3960.34	3960.74	92

	m er					
9	1 2- m er	<sup>7m</sup> GpppGAAUAUAAUACC	1	4305. 55	4305. 39	143
10	1 3- m er	<sup>7m</sup> GpppAGAAUAUAAUACC	1	4634. 52	4634. 76	156
11	1 3- m er	<sup>7m</sup> GpppGAUGAAGAUUAAG	2	4730. 81	4730. 86	216
12	1 5- m er	<sup>7m</sup> GpppAACACUUCUCAUCUU	1	5134. 41	5133. 13	69
13	1 7- m er	<sup>7m</sup> GpppAGUGGACUAACCACCAU	1	5910. 07	5909. 03	33
14	2 1- m er	<sup>7m</sup> GpppACGCGAAAAAUGCGUACAA C	1	7267. 97	7269. 08	30
15	2 1- m er	<sup>7m</sup> GpppACACUUGCUIIUUGACACAAC U	1	7094. 24	7095. 59	31

16	2 1- mer		$^{7m}GpppACACUUGC$ $^{m6}ACU$ $^{m6}ACU$	1	7136. 20	7136. 90	20
17	2 1- mer		$^{7m}Gppp^{m6}ACACUUGC$ $^{m6}ACU$ $^{m6}ACU$	1	7136. 20	7136. 64	10
&cl in eb; 18	& cl in eb; 2 8- mer	& cl in eb;	&clineb; $^{7m}GpppGGACAGAUCGCCUGG$ AGACGCCAUCCAC	&cli neb; 1	&clin eb;94 92.67	&clin eb;94 92.41	&cline b;99
&clineb;Cap-1 RNA							
19	8- mer		$^{7m}GpppA_mGAAAUAG$	1	3112. 87	3112. 14	170
20	9- mer		$^{7m}GpppG_mGGACAAAA$	3	3457. 10	3457. 23	82
21	1 0- mer		$^{7m}GpppA_mCGCGAAAAA$	1	3746. 01	3746. 11	56
22	1 0-		$^{7m}GpppA_mGAUUAUAG$	1	3725.	3725.	137

	m er			21	40	
23	1 1- m er	<sup>7m</sup> GpppA <sub>m</sub> CGCGAAAAAA	1	4075. 49	4075. 76	57
24	1 2- m er	<sup>7m</sup> GpppA <sub>m</sub> GAUGC UAAUAG	1	4375. 60	4375. 26	216
25	1 3- m er	<sup>7m</sup> GpppA <sub>m</sub> ACACUUCUCAUC	1	4537. 68	4537. 84	89
26	1 4- m er	<sup>7m</sup> GpppA <sub>m</sub> GAUGC UAUAAUAG	1	5010. 97	5011. 44	187
27	2 1- m er	<sup>7m</sup> GpppA <sub>m</sub> CACUUGC UUUUGACACAA CU	2	7110. 21	7110. 62	72
28	2 1- m er	<sup>7m</sup> GpppA <sub>m</sub> CACUUGC UUUUGA <sub>m</sub> CA <sub>m</sub> CA A <sub>m</sub> CU	1	7153. 11	7154. 66	15
29	5- m er	<sup>7m</sup> Gppp <sup>m6</sup> A <sub>m</sub> GUCC	1	2075. 22	2075. 04	126

30	5- m er		${}^7\text{mGppp}^{\text{m6}}\text{A}_\text{m}\text{UACU}$	1	2060. 21	2060. 35	86
31	7- m er		${}^7\text{mGppp}^{\text{m6}}\text{A}_\text{m}\text{GUCCCG}$	1	2725. 62	2725. 84	84
32	7- m er		${}^7\text{mGppp}^{\text{m6}}\text{A}_\text{m}\text{UACUUA}$	1	2695. 59	2695. 61	79
33	2 1- m er		${}^7\text{mGppp}^{\text{m6}}\text{A}_\text{m}\text{CACUUGC UUUUGACACA}$ ACU	1	7125. 24	7124. 37	26
34	2 1- m er		${}^7\text{mGppp}^{\text{m6}}\text{A}_\text{m}\text{CACUUGC UUUUG}^{\text{m6}}\text{A}_\text{m}\text{C}^{\text{m6}}$ $\text{A}_\text{m}\text{CA}^{\text{m6}}\text{A}_\text{m}\text{CU}$	1	7209. 72	7211. 20	19
&c lin eb; 35	& cl in e b; 2 1- m er	& cl in e b;	&c lineb; ${}^7\text{mGppp}^{\text{m6}}\text{A}_\text{m}\text{CACUUGC UUUUG}$ ${}^{\text{m6}}\text{A}_\text{m}\text{CA}_\text{m}\text{CAA}_\text{m}\text{CU}$	&cli neb; 1	&clin eb;71 81.61	&clin eb;71 81.23	&c line b;19
&c lineb;Cap structures							
36			${}^7\text{mGpppA}$	1	785.4 3	785.9 4	164

37		${}^7\text{mGpppA}_m$	1	799.4 6	800.1 3	285
38		${}^7\text{mGppp}^{\text{m6}}\text{A}$	1	799.4 6	799.4 1	26
39		${}^7\text{mGppp}^{\text{m6}}\text{A}_m$	1	813.4 8	813.4 7	34
40		${}^7\text{mGpppA}_m\text{-biotin}$	1	1369. 06	1369. 05	122
41		${}^7\text{mGppp}^{\text{m6}}\text{A}_m\text{-biotin}$	1	1383. 09	1383. 82	86

[a] MALDI-TOF characterization in negative mode [ $\text{M}^+ \leftarrow \text{M} \rightarrow 2\text{H}$ ]. [b] Isolated amount of  ${}^7\text{mG}$ -cap RNA after purification by IEX-HPLC, determined by UV measurement at 260<sup>nm</sup>.