Cap analogs with a hydrophobic photocleavable tag enable facile purification of fully capped mRNA with various cap structures.

Masahito Inagaki [1+], Naoko Abe[1+], Zhenmin Li[1], Yuko Nakashima[1,2], Susit Acharyya[1], Kazuya Ogawa[1], Daisuke Kawaguchi[1], Haruka Hiraoka[1], Ayaka Banno[1], Zheyu Meng[1], Mizuki Tada[1], Tatsuma Ishida[1], Pingxue Lyu[1], Kengo Kokubo[1], Hirotaka Murase[1], Fumitaka Hashiya[2], Yasuaki Kimura[1], Satoshi Uchida[5,6], Hiroshi Abe[1,2,3,4]*

1 Department of Chemistry, Graduate School of Science, Nagoya University, Furo-cho, Chikusa-ku, Nagoya, Aichi 464-8602, Japan
2 Research Center for Materials Science, Nagoya University, Furo-cho, Chikusa-ku, Nagoya, Aichi 464-8602, Japan
3 CREST, Japan Science and Technology Agency, 7, Gobancho, Chiyoda-ku, Tokyo 102-0076, Japan
4 Institute for Glyco-core Research (iGCORE), Nagoya University, Furo-cho, Chikusa-ku, Nagoya, Aichi 464-8601, Japan
5 Graduate School of Medical Science, Kyoto Prefectural University of Medicine, 1-5 Shimogamohangi-cho, Sakyo-ku, Kyoto 606-0823, Japan
6 Innovation Center of NanoMedicine (iCONM), Kawasaki Institute of Industrial Promotion, 3-25-14 Tonomachi, Kawasaki-ku, Kawasaki 210-0821, Japan
*Corresponding author, h-abe@chem.nagoya-u.ac.jp
[+] *These authors contributed equally to this work*

ABSTRACT

Removing immunogenic uncapped mRNA from in vitro transcribed mRNA is critical in mRNA research and clinical applications. Commonly used capping methods provide a maximum capping efficiency of around 80-90% for widely used Cap-0- and Cap-1-type mRNAs. However, uncapped and capped mRNA possesses almost identical physicochemical properties, posing challenges to their physical separation. Herein, we developed hydrophobic photocaged tag-modified cap analogs, which separated capped mRNA from uncapped mRNA by reversed-phase HPLC. Subsequent photo-irradiation recovers footprint-free native capped mRNA. This approach provided 100% capping efficiency even in Cap-2-type mRNA with versatility applicable to 650 nt and 4,247 nt mRNA. The Cap-2-type mRNA showed up to 3 to 4-fold higher translational activity in cultured cells and animals than mRNA prepared by the standard capping method. Notably, the purification process simultaneously removed immunogenic double-stranded mRNA, another major contaminant of in vitro transcribed mRNA, drastically reducing mRNA immunogenicity in cultured cells.

IINTRODUCTION

Decades of research into mRNA therapeutics[1, 2] have culminated in two approved mRNA vaccines for the COVID-19 pandemic[3, 4]. The 5' cap plays a critical role in mRNA therapeutics, wherein N^7-methylguanosine (m^7G) is attached to the 5' end of mRNA *via* a 5'-5' triphosphate bridge (Fig. 1a)[5, 6]. The interaction between the 5' cap and the poly-A chain at the 3' end of mRNA enhances the translational activity of mRNA[5, 6]. The other cap structure roles include splicing, protection to nuclease, and nuclear transport[5, 6]. Structurally, the 2'-hydroxyl groups at the 5' end of mRNA are methylated in the cap structure[5, 6, 7, 8]. The cap structures of the first two nucleotides possessing zero, one, and two 2' *O*- methyl (OMe) groups are called Cap-0, Cap-1, and Cap-2, respectively (Fig. 1a). The methylation in the Cap-1 structure is a marker of 'self' mRNA, making Cap-1 mRNA with less immunogenic than Cap-0[9, 10, 11], while the Cap-2 functionalities have yet to be fully discovered. In mammals, when the first transcribed base of capped mRNA is A, most of its N^6 position is methylated, forming m^6A[5, 6]. The presence of m^6A increases resistance to an mRNA decapping enzyme, Dcp2[12].

In chemical or semi-chemical methods of capped RNA synthesis[13, 14, 15, 16], the length of the RNA strand is limited to about 150 bases. For synthesizing longer capped mRNA with high efficiency, RNA polymerase of phage origin is mainly used, especially T7 RNA polymerase (T7 RNAP). The current capping methods include enzymatic and co-transcriptional capping[5]. In the enzymatic capping, transcribed mRNA is treated with a capping enzyme[5, 17], often followed by methyltransferase treatment to convert the Cap-0 structure to the Cap-1[5]. In co-transcriptional capping, a synthetic cap analog, *i.e.*, m^7G-containing nucleotide, is added to the *in vitro* transcription (IVT) reaction[18, 19]. The first generation dinucleotide cap analog $m^7G(5')ppp(5')G$ had the problem of the chain-elongation starting from the 3' hydroxyl group of the m^7G. The unintended elongation results in 30–50% of the reversely capped RNA inactive for translation[19]. This issue was solved by an anti-reverse cap analog [ARCA, $m_2^{7,3'\text{-}O}G(5')ppp(5')G$] methylated in the 3'-hydroxyl group of m^7G, ensuring the chain-elongation exclusively from the desired side[20, 21]. Recently, new cap analogs consisting of trinucleotides and tetranucleotides have been developed[22, 23, 24, 25, 26] to directly introduce the Cap-1 or Cap-2 structure into mRNA. Despite such progress, co-transcriptional capping still has a challenge in obtaining capped mRNA with 100% efficiency because nucleotides such as GTP compete with cap analogs for the initiation step of the transcription reaction.

The uncapped byproduct has a triphosphate (ppp) at the 5' end, a motif found in viral mRNA, triggering an immune response via innate immune receptors such as retinoic acid-inducible gene-I (RIG-1) and melanoma differentiation-associated protein 5 (MDA5)[11, 27]. Moradian *et al.* have reported that a small amount of 5' ppp-RNA byproduct from a standard co-transcriptional capping method induces substantial immune responses[28], wherein the capping method provides around 80 – 90 % capping efficiency[29]. These studies highlight the necessity to maximize capping efficiency or

minimize uncapped ppp-RNA contaminants. The enzymatic treatment is a standard method to remove uncapped mRNA byproducts. For example, the 5'-terminal triphosphate of uncapped RNA will be converted to the monophosphate by RNA 5' polyphosphatase, followed by degrading monophosphorylated RNA using XRN I, a 5'→3' exoribonuclease requiring 5′ monophosphate [23, 25, 26, 30]. However, considering mRNA susceptibility to non-specific degradation, the number of purification processes should be minimized. In addition, repeated enzymatic treatment significantly increases the production cost, hampering commercial mass preparation of mRNA. These issues motivate us to develop a non-enzymatic methodology for diminishing the uncapped mRNA immunogenicity.

Physical separation of capped mRNA from uncapped mRNA would become a simple, robust, cost-effective purification strategy. However, capped and uncapped mRNA have almost the same physicochemical properties, posing challenges for their separation. In this study, we designed hydrophobic cap analogs, namely PureCap analogs, for purifying capped mRNA. The PureCap cap analogs are integrated into mRNA 5'-terminal during IVT reaction, and its hydrophobicity enables the purification of the capped mRNA using reversed-phase high-performance liquid chromatography (RP-HPLC) (Figs. 1b, 2). Using a photodegradable 2-nitrobenzyl derivative as a hydrophobic tag enables tag removal under mild conditions by light irradiation[31]. This method relies on RP-HPLC, a standard purification process for nucleic acid medicine, rather than enzymatic treatment. Especially, purification of mRNA by RP-HPLC is an established method for eliminating double-stranded RNA (dsRNA) contaminants produced during the IVT, which increases the immunogenicity of mRNA and inhibits its translation[32, 33, 34, 35]. Thus, the PureCap method can isolate capped mRNA from the two major impurities, uncapped ppp-RNA and dsRNA, simultaneously providing a practical advantage in mRNA production.

Using the PureCap technology, we succeeded in a footprint-free purification of the capped mRNA in a native form with 100% efficiency for the first time. The PureCap mRNA showed undetectable immunogenicity even without enzymatic treatment. Achieving 100% efficiency in various cap species, our approach provides an unbiased platform to study their structure-activity relationship by excluding the influence of different capping efficiency observed among cap species. Notably, we prepared fully capped Cap-2 mRNA, which showed up to 3 to 4-fold higher translational activity in cultured cells and animals than mRNA prepared by standard capping methods.

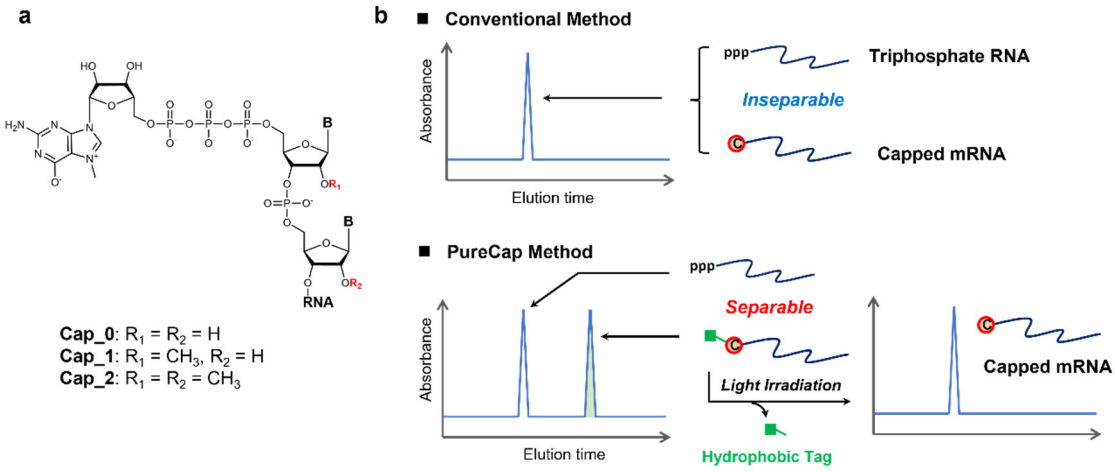


Figure 1 Hydrophobic cap analogs for the purification of capped mRNA using RP-HPLC (the PureCap method). (**a**) Structure of the 5' cap in eukaryotic mRNA. (**b**) Capped mRNA can be isolated from uncapped impurities using RP-HPLC by the hydrophobicity of a tag attached to the cap moiety. After the isolation of the capped mRNA, light irradiation can remove the tag.

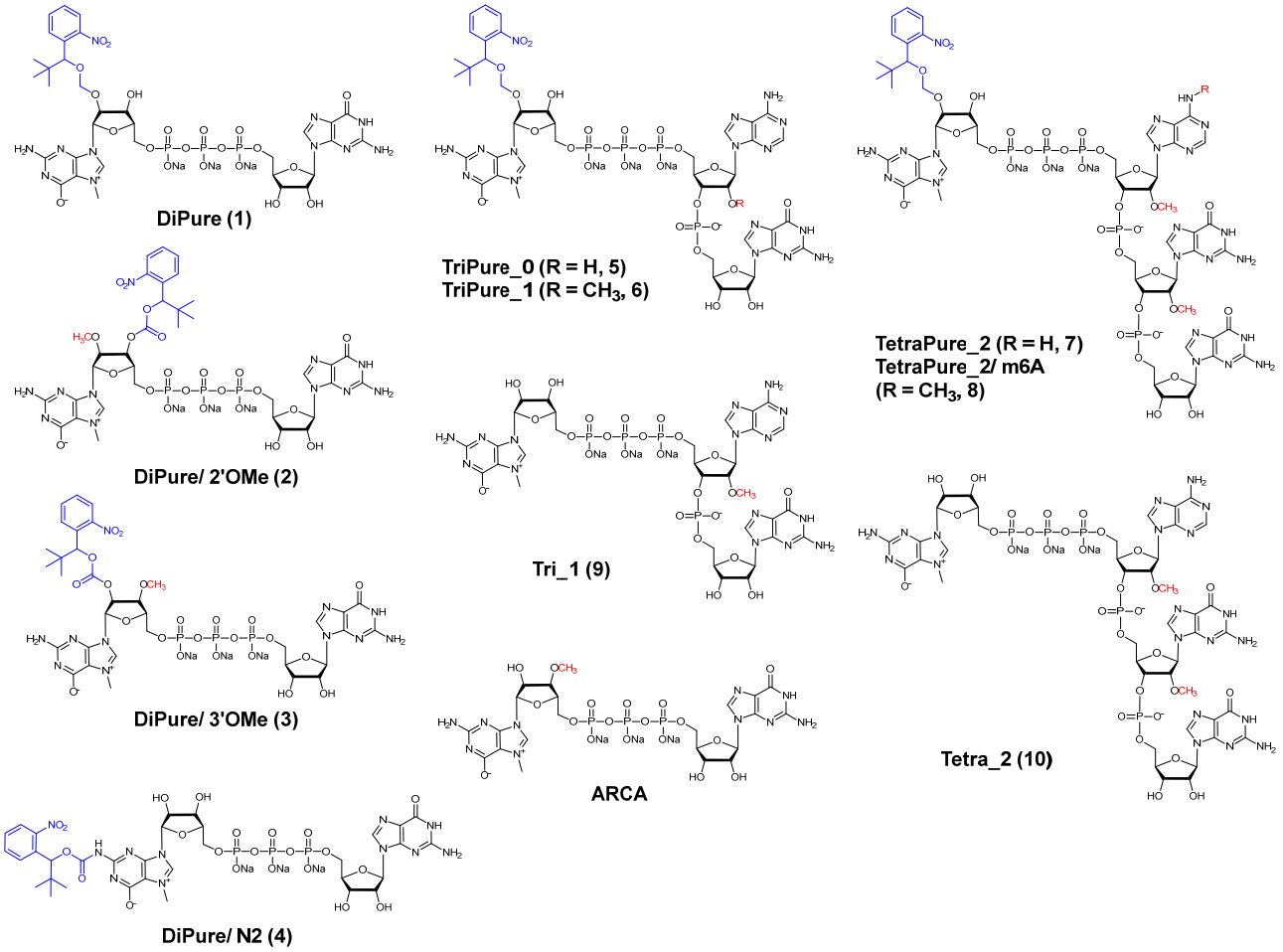
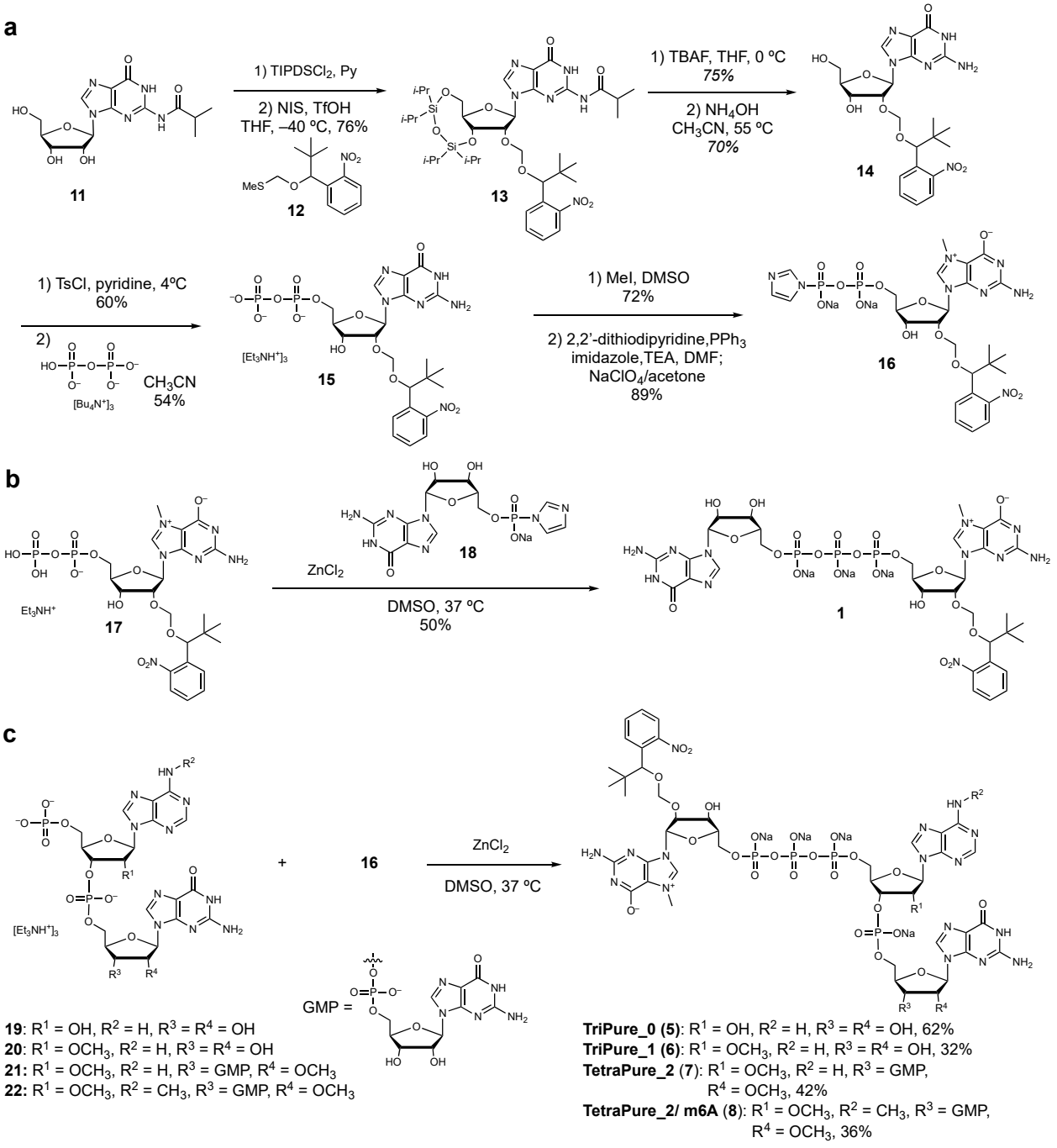


Figure 2. Structure of cap analogs modified with a photodegradable hydrophobic tag "the PureCaps" and untagged control cap analogs synthesized/ evaluated in this study.

RESULTS

Design and synthesis of dinucleotide PureCap analogs



Scheme 1. Synthesis of PureCap analogs. (**a**) Synthesis of 7-methyl 2'-*O*-Nb-guanosine 5' diphosphate imidazolide (**16**). (**b**) Synthesis of dinucleotide PureCap analog **DiPure** (**1**). (**c**) Synthesis of tri/tetranucleotide PureCap analogs; **TriPure_0** (**5**), **TriPure_1** (**6**), **TetraPure_2** (**7**) and **TetraPure_2/ m6A** (**8**).

To begin with, we designed and synthesized four new dinucleotide PureCap cap analogs modified with a hydrophobic tag (Fig. 2, Scheme 1). The structure of a hydrophobic tag contains a *tert*-butyl (*t*Bu) group in a 2-nitobenzyl (Nb) photocaging molecule[31], where *t*Bu moiety works to enhance its hydrophobicity and chemical stability[36]. In three of the four designed cap analogs, the Nb tag was introduced into either the 2' or 3' hydroxyl group of m^7G. **DiPure** (**1**) is an analog that links the Nb group *via* an acetal group at the 2' position and has a free 3' hydroxyl group since methylation of the 2'-hydroxyl has been reported to inhibit chain elongation from m^7G as well as that of the 3'-hydroxyl[37]. The other two analogs have *O*-methyl (OMe) and *O*-Nb modifications at the 2'/3' positions of m^7G. Analogs with OMe modifications on either 2' or 3' were named **DiPure/2'OMe** (**2**) or **DiPure/3'OMe** (**3**), respectively. Using the ONb and OMe modifications, all three analogs have anti-reverse activity. Being deprotected after introduction into mRNA, three types of purely capped mRNA can be prepared that differ only in the presence or absence of the OMe modification. As a result, we planned to make a precise comparison in the translation activity with the structure without OMe modification, which was impossible before. The remaining compound, **DiPure/N2** (**4**), is a cap analog modified with the Nb at the exocyclic amino group of m^7G.

The key step in synthesizing cap analogs is synthesizing the corresponding diphosphate by phosphorylation of guanosine derivatives. For the synthesis of cap analogs known as the ARCA derivatives, protocols consisting of monophosphorylation by the Yoshikawa method using phosphoryl chloride, activation of monophosphate thorough phosphorimidazolide formation, and diphosphate formation by reaction with alkylammonium phosphate were applied[18]. However, this method is complicated because it is based on the stepwise introduction of phosphate groups and requires multiple purifications with aqueous solvents, such as ion exchange and reversed-phase chromatography. The introduction of a lipophilic moiety, such as in the PureCap analogs, is expected to decrease the solubility of the intermediates in aqueous solvents, making purification particularly difficult. Therefore, we applied a one-pot synthesis of the diphosphate directly from the nucleoside to avoid the solubility problem and simplify and shorten the synthetic processes. In the first approach, phosphoryl chloride is applied to a guanosine derivative, and the diphosphate is obtained by adding an alkylammonium phosphate salt to the resulting phosphorodichloridate intermediate[38, 39, 40]. The second approach is to obtain the diphosphate directly from 5'-tosylated guanosine by reacting with tetrabutylammonium pyrophosphate salt[41, 42, 43, 44]. The resulting diphosphate is converted to the

desired dinucleotide PureCap analog by methylation at the N^7 position and subsequent condensation reaction with guanosine monophosphate imidazolide in the presence of zinc chloride.

2'-*O*-Nb-modified guanosine **14** as a phosphorylation precursor was synthesized from N^2-isobutyrylguanosine **11** in four steps (Scheme 1a). First, N^2-isobutyrylguanosine **11** was reacted with 1,3-dichloro-1,1,3,3-tetraisopropyldisiloxane ($TIPDSCl_2$) in pyridine to afford the 3',5'-*O*-protected compound in a quantitative yield[45]. The key Nb group was introduced by reaction with nitrobenzyl alcohol methylthioacetal **12** in the presence of *N*-iodosuccinimide (NIS) and TfOH in THF at -40 ºC to afford the corresponding product **13** in 76% yield. The silyl protecting group was subsequently removed with tetrabutylammonium fluoride to afford the unprotected compound in 75% yield, which was then treated with 28% aqueous ammonia at 55 ºC to remove the isobutyryl group to afford 2'-*O*-Nb modified guanosine **14** in 70% yield. In the first approach, **14** was reacted with phosphoryl chloride in trimethyl phosphate, and then tetrabutylammonium monophosphate salt was applied to afford the corresponding diphosphate **15** in one pot after hydrolysis (Supporting Scheme S3). Although **15** could be synthesized by this method in a short step, the low yield of the diphosphate formation (13%) was a problem. Therefore, as a second approach, direct diphosphorylation via 5'-*O*-tosylated form **14** was investigated. First, 2'-*O*-Nb guanosine (**14**) was reacted with tosyl chloride in pyridine to obtain the 5'-*O*-tosylated form in 60% yield, which was then reacted with tetrabutylammonium diphosphate salt to give the diphosphate **15** in 54% yield. The resulting diphosphate was reacted with methyl iodide in DMSO to afford the 7-methylguanosine diphosphate derivative in 72% yield. This compound was then converted to the corresponding phosphorimidazolide **16** by treatment with imidazole, 2,2'-dithiodipyridine, and triphenylphosphine in 89% yield. Finally, compound **17** was reacted with guanosine monophosphate phosphoroimidazolide **18** in the presence of zinc chloride to afford PureCap analog **DiPure** (**1**) in 50% yield (Scheme 1b). 2'-*O*-methylated analog **DiPure/2'OMe** (**2**), 3'-*O*-methylated analog **DiPure/3'OMe** (**3**), and N^2-Nb-protected analog **DiPure/N2** (**4**) were synthesized similarly, as shown in the Supporting Schemes S6 to S8.

Design and synthesis of tri/tetranucleotide PureCap analogs

The cap structure produced by the co-transcription of dinucleotide cap analogs is Cap-0. However, in 2009, Ishikawa *et al.* reported that mRNAs with the Cap-1 structures could be prepared using trinucleotide cap analogs[22]. Sikorski *et al.* recently synthesized a series of trinucleotide cap analogs and compared their translational activity of mRNA containing Cap-1 with that of mRNA containing Cap-0, reporting a higher translation activity of the former in some mammalian cultured cells[23]. More recently, Drazkowska *et al.* evaluated the translational activity of Cap-2-containing mRNAs prepared using tetranucleotide cap analogs[26].

We synthesized tri- or tetranucleotide PureCap analogs in this study to produce quantitatively capped mRNAs with Cap-1 and Cap-2 structures (Figs. 1 and 2, Scheme 1c). We selected A and m^6A bases as the base of the sites to which m^7G links in the tri- or tetranucleotides; according to the report that the highest translational activity with Cap-1 mRNA was observed when the base in the cap structure was A or m^6A[23].

Sikorski *et al.* reported the synthesis of trinucleotide cap analogs by the reaction of 5'-monophosphorylated dinucleotides synthesized by a solid phase method with N^7-methylguanosine diphosphate imidazolide in the presence of zinc chloride[23]. Senthilvelan *et al.* also reported the synthesis of trinucleotide cap analogs using 5'-monophosphorylated dinucleotides synthesized by a liquid phase method[25]. The 5'-monophosphorylated dinucleotides (**19, 20**) and 5'-monophosphorylated trinucleotides (**21, 22**) synthesized by these synthetic methods were reacted with N^7-methyl-2'-*O*-Nb-guanosine diphosphate imidazolide **16** in the presence of zinc chloride to synthesize PureCap analogs in trinucleotide (**5**, **6**) and tetranucleotide (**7**, **8**) forms. (Scheme 1c).

Preparation and purification of mRNA using dinucleotide PureCap analogs

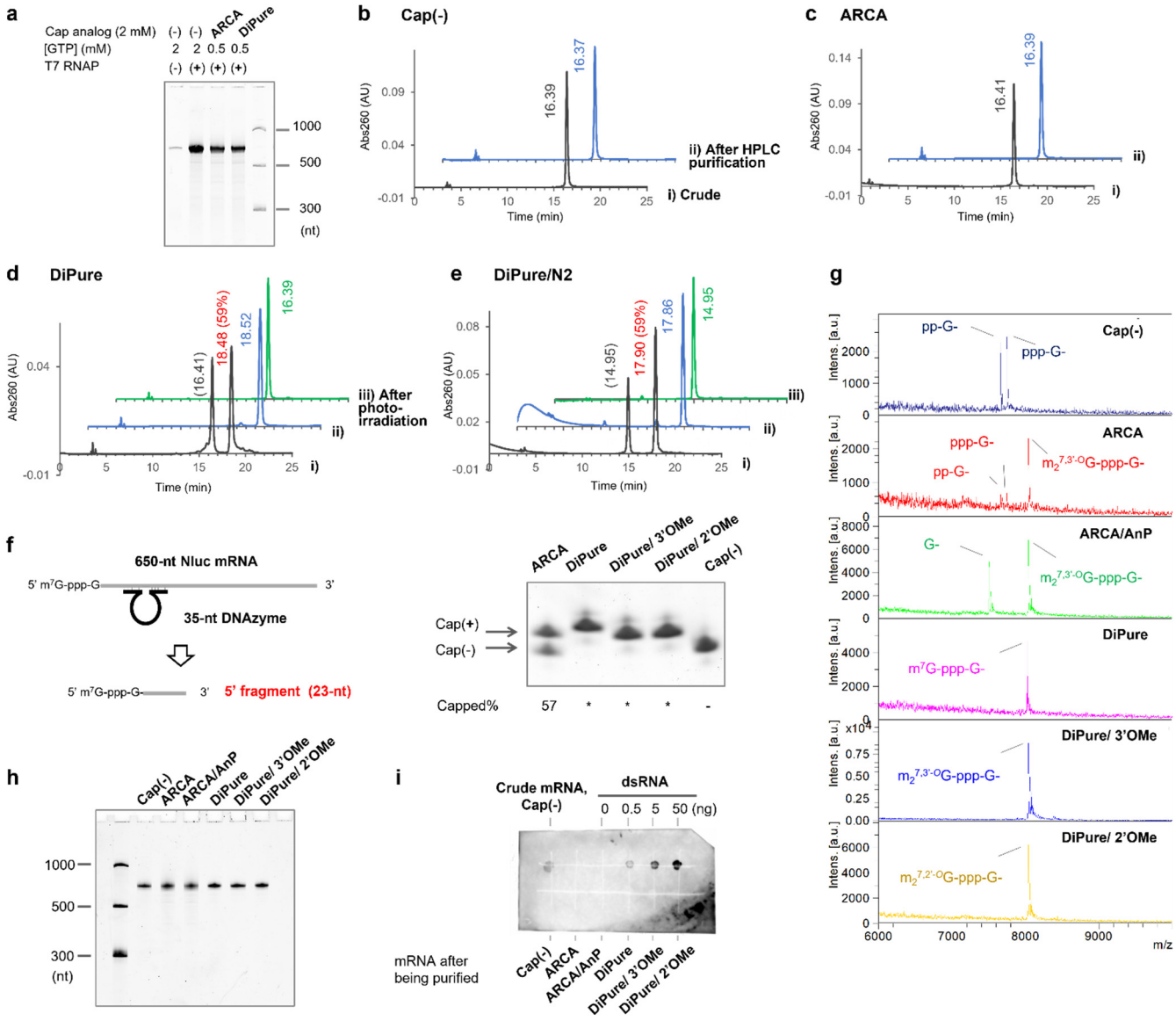


Figure 3. IVT and purification of mRNA by RP-HPLC using dinucleotide PureCap analogs. Nluc mRNA (650-nt) was transcribed from its DNA template using T7 RNAP in the absence or presence of a cap analog. (**a**) dPAGE analysis of the transcription reaction. (**b**–**e**) RP-HPLC analysis of the RNA transcript. The transcript was analyzed as a crude mixture (**i**, black line) or after being purified by preparative HPLC (**ii**, blue line). Purified RNA was analyzed after deprotection by irradiating 365-nm light (**iii**, green line). Elution time (min) of the peak was noted nearby. A ratio (%) of capped mRNA calculated based on the peak area was listed in parentheses after the elution time. The elution time in the experiment shown in **e** was different from others (**b**–**d**) because HPLC conditions were not completely identical. (**f, g**) Analysis of 5' terminus of purified mRNAs after cleavage by DNAzyme 10-23. Cleaved 5' RNA fragments were analyzed by dPAGE. The capped ratio was

calculated based on the band intensities of the capped and uncapped RNA fragments. An asterisk (*) means that uncapped fragments were judged to be undetectable. (**g**) MALDI-TOF MS analysis of the 5' RNA fragments. For these observed peaks, the calculated and measured MS agreed well, which were shown in Extended Data Table 1. (**h, i**) Purity confirmation of mRNAs by dPAGE (**h**) and dsRNA removal check by a dot blot assay using an anti-dsRNA antibody (**i**). 25 ng (**h**) or 500 ng RNAs (**i**) were analyzed.

First, we examined whether the synthesized dinucleotide cap analog **DiPure** could be used to isolate and purify capped mRNA by RP-HPLC as planned. Using 676-bp double-stranded DNA containing the T7 promoter and encoding NanoLuc luciferase (Nluc) as a template, mRNA of 650 bases was transcriptionally synthesized using T7 RNAP[46]. **DiPure** was added to this reaction, and the resulting RNA transcripts were analyzed by RP-HPLC (Fig. 3). Control samples were prepared without the cap analog and with the commercially available ARCA [$m_2^{7,3'\text{-}O}$G(5')ppp(5')G]. First, as is common in capped mRNA preparation reactions using ARCA to increase the introduction of the cap analog, the GTP concentration was reduced to 0.5 mM, and NTPs other than GTP and cap analog concentrations were set to 2 mM for the transcription reaction. (Fig. 3a). The resulting RNA transcripts were analyzed by RP-HPLC, and RNA was eluted at 16.4 min in case no cap analog was added or with ARCA (Figs. 3b, 3c). On the other hand, with the **DiPure** cap analog, a peak at 18.5 min was observed in addition to the peak at 16.4 min (Fig. 3d). We judged that the elution time of capped RNA containing **DiPure** changed from 16.4 min to 18.5 min due to its hydrophobicity (Fig. 3d). To confirm this, the peak at 18.5 min was collected, and 365-nm light was irradiated to the RNA. The elution time of the RNA changed from 18.5 min to 16.4 min due to the removal of the hydrophobic protecting group. To confirm the structure of the 5' end of the RNA, 650-nt Nluc mRNA purified by RP-HPLC was cleaved at 23 bases from the 5' end by DNAzyme 10-23[47]. The product was subjected to denaturing polyacrylamide gel electrophoresis (dPAGE) and mass analysis (Figs. 3f, 3g, Extended Data Table 1). The results showed that the mRNA prepared with ARCA had a capping rate of 57% and was a mixture of capped and uncapped mRNA, whereas the mRNA purified by RP-HPLC with **DiPure** contained no uncapped mRNA. For all mRNAs, including uncapped mRNA, minor bands were observed above and below the main band of the cleavage product, which we believe is due to the 5' end heterogeneity of the transcript produced by T7 RNAP (Fig. 3f)[48]. These results indicate that capped mRNA can be quantitatively prepared by RP-HPLC purification using the **DiPure** cap analog.

DiPure/N2, a cap analog with a protecting group at the N^2 position of m^7G, was similarly added to the IVT reaction, and the generated RNA was analyzed by RP- HPLC. As a result, it was found that capped mRNA could be isolated and purified for mRNA containing this analog, using the hydrophobicity of the cap analog (Fig. 3e).

When T7 RNAP transcribes RNA, it is known that highly immunogenic dsRNA is produced as a byproduct[32, 33, 34, 35]. When this dsRNA is included in the mRNA as an impurity, the translation activity of the mRNA is significantly impaired due to the triggering of cellular immune response[23, 32, 34]. Purification by RP-HPLC or removal using cellulose is known for removing dsRNA impurity[23, 32, 34]. In our preliminary study, we also noticed that purifying capped RNA by RP-HPLC increased the translational activity in HeLa cells by about 50–100-fold. Detection experiments using anti-dsRNA antibody showed that this increase in activity was due to the removal of dsRNA by RP-HPLC purification. Using RNA transcripts with no cap analog added, we examined the elution time of the dsRNA impurity by RP-HPLC and found that it eluted about 2–3 minutes later than the main product, 650-nt mRNA, under our analytical conditions (Extended Data Fig. 1). This position coincides with the elution position of mRNA containing PureCap analogs. Therefore, mRNA containing PureCap had to be purified once by HPLC preparative purification of capped mRNA using its hydrophobicity, followed by deprotection by light irradiation. Then a second HPLC purification was added for complete dsRNA removal. Control mRNA samples were also purified by HPLC twice to match the number of cycles in the HPLC purification process.

Capped Nluc mRNAs using **DiPure/3'OMe** and **DiPure/2'OMe** were also prepared (Extended Data Fig. 2). In this case, we examined the transcription reaction conditions. We found that if the GTP concentration was not reduced in the reaction and kept at 2 mM, the same as other NTPs, the capped mRNA yield after RP-HPLC isolation increased, though the incorporation rate of the PureCap analog decreased. Therefore, in subsequent experiments, the GTP concentration was not reduced, and the desired pure-capped mRNA was isolated from the resulting RNA transcript mixture by RP-HPLC.

The purity of the resulting capped mRNAs was confirmed to be nearly identical by dPAGE (Fig. 3h). A detection assay for dsRNA contamination using an anti-dsRNA antibody confirmed that dsRNA was similarly removed from the capped mRNAs (Fig. 3i). These highly purified capped mRNAs were used for subsequent immunogenicity and translation activity evaluations.

Preparation and purification of mRNA using tri/tetranucleotide PureCap analogs

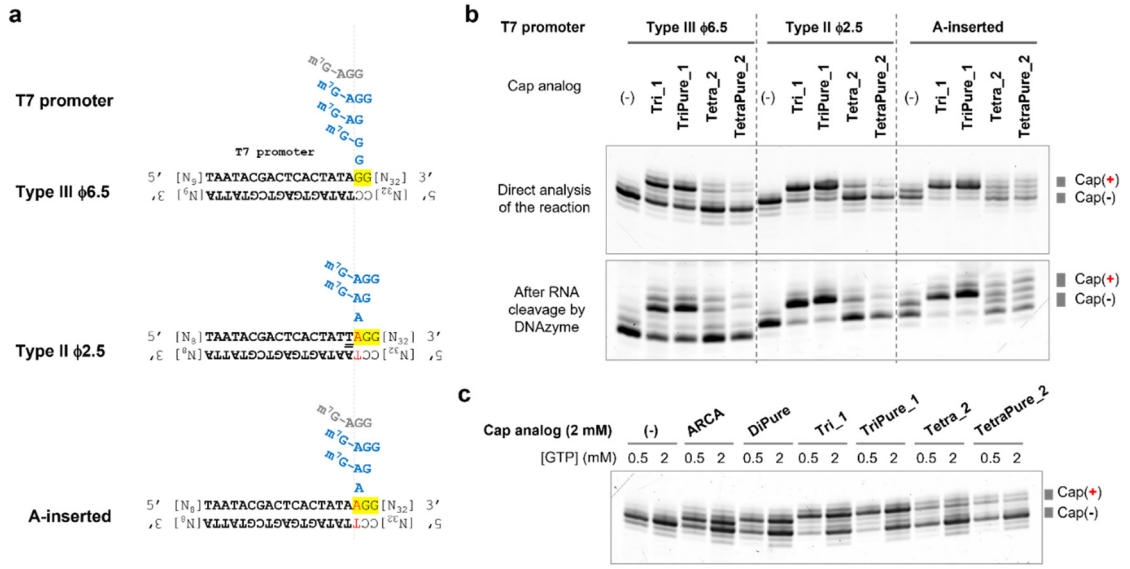


Figure 4. Tri/tetranucleotide PureCap analogs are incorporated by T7 RNAP in a model system. (**a**) Three types of promoter sequences were tested using 60-bp dsDNA. The estimated position of cap analogs when they were incorporated into RNA was depicted in the figure. Major or minor species were shown in blue or gray color, respectively. (**b**,**c**) dPAGE analysis of the reaction. An aliquot was taken from the reaction mixture and analyzed (upper panel in **b** and **c**). Or they were analyzed after being cleaved by DNAzyme 10-23 to align the heterogeneous 3' ends of the RNA transcripts (lower panel in **b**). (**b**) Effect of the promoter sequence on the incorporation of the cap analogs. The concentrations of NTPs and cap analogs were 2 mM. (**c**) Analysis of IVT on the template with Type III ϕ6.5 promoter. NTP other than GTP and cap analog concentrations were set at 2 mM, and GTP concentrations were set at either 0.5 or 2 mM.

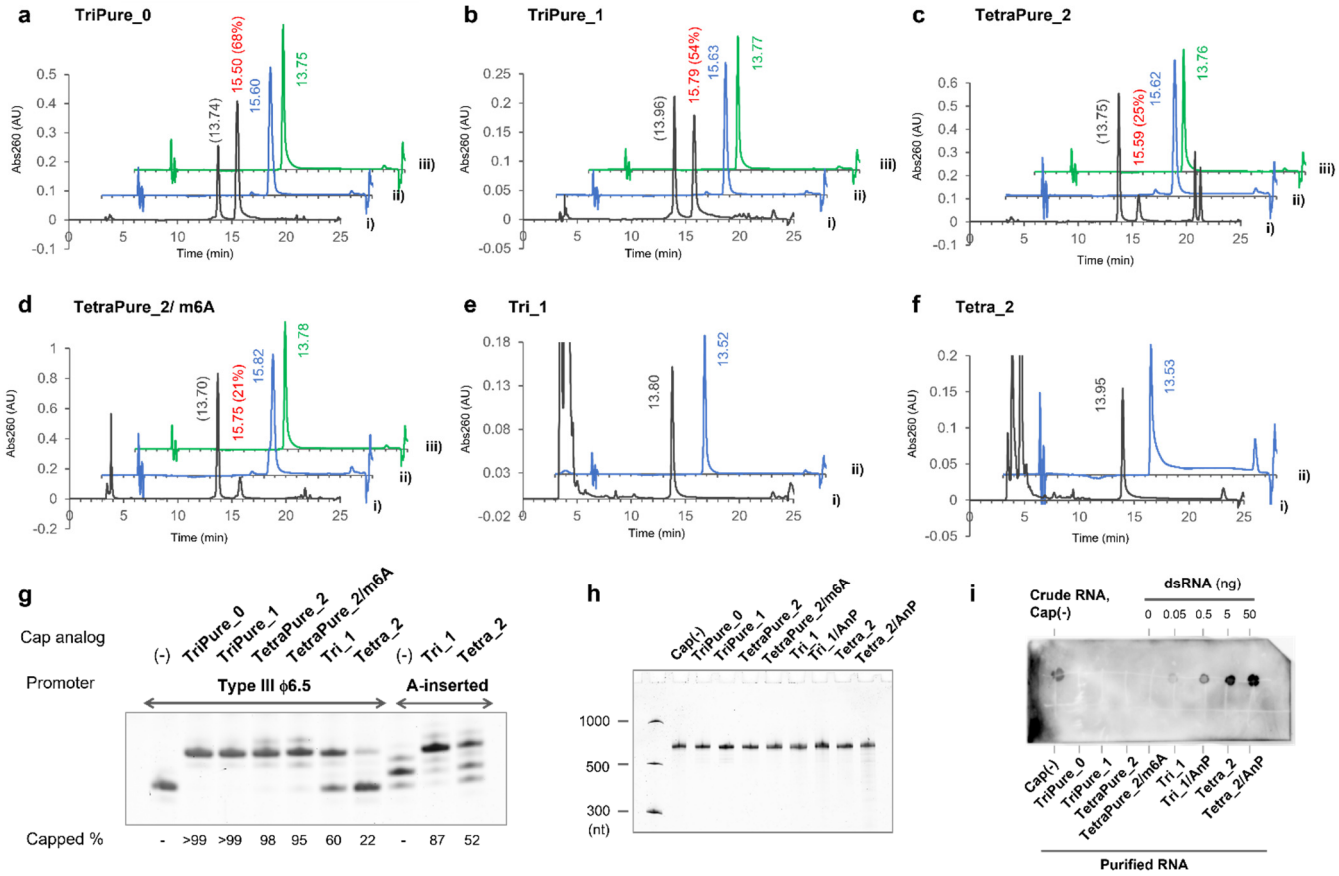
First, a trinucleotide PureCap analog **TriPure_1** and a tetranucleotide PureCap analog **TetraPure_2** were tested to determine whether they could be incorporated in the IVT using a short RNA transcription system on 60-bp dsDNA templates (Fig. 4). As a control experiment, the efficiency of incorporation of untagged cap analogs was evaluated to know the effect of the modification by comparison. We tested three different promoter sequences: the first was a sequence called Type III ϕ6.5, known as the most versatile and powerful promoter sequence; the second was Type II ϕ2.5, initiated from ATP, anticipating the previously reported 5' end high uniformity[49, 50]. And third, a modified T7 promoter with a corresponding base A inserted at the +1 position of the ϕ6.5 promoter, referred to as the "A-inserted" promoter in the figure, was tested that is reported to improve the trinucleotide cap analog incorporation in IVT[24]. First, IVT reactions were performed with all NTPs and cap analog at 2 mM concentration, and transcripts were analyzed by dPAGE

(Figure 4). To resolve the 3' end heterogeneity of the purified RNA transcripts and to facilitate confirmation of the product, the RNA was digested by DNAzyme at ten bases upstream from the 3' end of the RNA[47]. They were analyzed by dPAGE in the same manner. The results showed that the incorporation efficiency of the trinucleotide cap analog was higher than that of the tetranucleotide for all promoters used and that tag introduction to the trinucleotide did not affect the incorporation efficiency. Indeed, with the A-inserted promoter, the generation of the uncapped transcript was lower, and the selectivity of the transcription initiation from the trinucleotide cap was higher[24]. For the tetranucleotide cap analogs, unlike the trinucleotide counterpart, a reduction in the incorporation efficiency due to the introduction of the tag was observed. Next, we used the Type III ϕ6.5 promoter to determine whether decreasing the GTP concentration increases the amount of capped RNA obtained (Fig. 4c)[26]. The results showed that lowering the concentration of GTP, which competes with the cap analog in the initiation phase, increased the capping ratio, but decreased the amount of capped RNA produced. When using the PureCap analog, the amount of uncapped RNA was not an issue because RP-HPLC could separate the capped RNA. From these results, we chose to keep the GTP concentration the same as the other NTPs in the transcription reaction with PureCap analogs to prepare the mRNA.

Nluc mRNA was then prepared using these PureCap analogs to study the effects of Cap-0, Cap-1, and Cap-2 structures on translational activity (Fig. 5). Cap-0 and Cap-1 structures were introduced into mRNA using the corresponding trinucleotide PureCaps, **TriPure_0**, **TriPure_1**, respectively (Fig. 1). Cap-2 structures were introduced using the tetranucleotide PureCaps **TetraPure_2** or **TetraPure_2/m6A**. Preliminary experiments were performed to compare the amount of capped RNA produced by the two promoters shown in Fig. 4, Type III ϕ6.5 and A-inserted. As a result, we found that the Type III ϕ6.5 promoter made the highest amount of the capped mRNA and that it was appropriate to use this promoter to prepare mRNA (Extended Data Fig 3). Nluc mRNA was transcribed with 2 mM concentrations of NTPs and cap analog from DNA containing the Type III ϕ6.5 promoter. As shown in Figure 5, similar to the case with **DiPure** caps, mRNAs with PureCap analogs showed different elution times than uncapped mRNAs, and RP-HPLC could isolate capped mRNAs. Consistent with the previous results producing short RNA transcripts, the incorporation efficiency of tetranucleotide PureCaps was lower than those of trinucleotide PureCaps: 68% and 54% for **TriPure_0** and **TriPure_1**, and 25% and 21% for **TetaPure_2**, **TetraPure_2 /m6A,** respectively. To check the identity of the 5' end of the HPLC-purified mRNAs, they were cleaved using DNAzyme 10-23 at 23 bases from the 5' end and analyzed by dPAGE and MALDI-TOF MS (Fig. 5, Extended Data Fig. 4). The capped mRNA produced using PureCap analogs showed almost no band derived from uncapped RNA and observed molecular weight of the 5' end fragments were consistent with theoretical ones. Control mRNAs were prepared using two cap analogs without tag, a trinucleotide cap analog **Tri_1** and a tetranucleotide cap analog **Tetra_2** from templates containing

two different promoters, Type III ϕ6.5 and A-inserted one, and the differences in their capping efficiencies were examined (Fig 5, Extended Data Fig. 4). As shown in Fig. 5, both cap analogs were incorporated into mRNA at higher ratios when the template containing A-inserted promoter was used. Therefore, the control mRNAs were transcribed from the template containing the A-inserted promoter and utilized for subsequent biological activity measurements.

Figure 5. NanoLuc luciferase mRNA (650-nt) preparation using tri/tetranucleotide PureCap analogs.



(**a**-**f**) RP-HPLC analysis of IVT-RNAs. PureCap analogs **TriPure_0** (**a**), **TriPure_1** (**b**), **TetraPure_2** (**c**) and **TeraPure_2/m6A** (**d**) were added to the reaction. Control mRNA using cap analogs **Tri_1** (**e**) or **Tetra_2** (**f**) was transcribed from a DNA template containing the A-inserted promoter. The transcript was analyzed as a crude mixture (**i**, black line) or after being purified by preparative HPLC (**ii**, blue line). Purified RNA was analyzed after deprotection by irradiating 365-nm light (**iii**, green line). Elution time (min) of the peak was noted nearby. A ratio (%) of capped mRNA calculated based on the peak area was listed in parentheses after the elution time. (**g**) Analysis of 5' terminus of purified mRNAs after cleavage by DNAzyme 10-23. Cleaved 5' RNA fragments were analyzed by dPAGE. Capped ratio (%) of mRNA was calculated based on the band intensities of the capped and uncapped RNA fragments. (**h**, **i**) Purity confirmation of mRNAs by

dPAGE (**h**) and dsRNA removal check by a dot blot assay using an anti-dsRNA antibody (**i**). 25 ng (**h**) or 500 ng RNAs (**i**) were analyzed.

Immunogenicity of mRNAs prepared using PureCap analogs

Mammalian cells have an immune system as a defense mechanism against infection by foreign organisms such as viruses[11, 27]. Externally introduced RNA is recognized by pattern recognition receptors such as Toll-like receptor 3 (TLR3), TLR7, TLR8, and RIG-I, triggering an immune response in host cells[27]. If the prepared mRNA contains uncapped 5' ppp-RNA and/or dsRNA impurity, the defense mechanism against antiviral infection will be activated[11, 27]. When they are sensed by RIG-1, a cellular-type immune response occurs, activating protein kinase R and inhibiting intracellular translation. 5' ppp-RNA is also recognized by the antiviral proteins Interferon-induced protein with tetratricopeptide repeats 1 (IFIT1) and IFIT5[51]. Therefore, removing as much uncapped mRNA as possible when preparing capped mRNAs is essential, which may cause unwanted immune responses.

Since highly purified, quantitatively capped 650-nt Nluc mRNAs were prepared by the PureCap method described above, the immunostimulatory properties of these mRNAs were evaluated using NF-κB reporter (Luc)-HEK293 cell line[52]. Control mRNAs and mRNAs prepared by the PureCap method were introduced into the cells by lipofection, cultured, and lysed. The activation level of the NF-κB pathway was measured as the expression of the firefly luciferase reporter gene. In this system, after activating pro-inflammatory cytokines or agonists of lymphokine receptors, endogenous NF-κB transcription factors bind to the DNA response elements, inducing the luciferase reporter gene transcription. As shown in Fig 6, the expression level of the luciferase gene by mRNAs prepared with the PureCap analogs was almost the same as that of the mock-transfected sample (**No RNA**). In contrast, when mRNA prepared using the control cap analogs, ARCA or Tri_1 was transfected, and 2.7- and 2.2-fold higher luciferase expression was observed than in the mock-transfected sample. The expression decreased when the mRNA was dephosphorylated using Antarctic phosphatase, which indicates that their immunogenicity was attributed mainly to the uncapped 5' ppp-RNA byproduct[9, 11, 53]. We confirmed that RP-HPLC purification significantly reduced the amount of luciferase expression, which would reflect the removal of dsRNA impurities (Fig. 6a, **ARCA, crude** vs. **ARCA**)[32].

These experiments demonstrate the effect of removing 5'-triphosphate from uncapped mRNA, which is included as an impurity, on lowering the immunostimulatory effect of prepared mRNA. At the same time, they show the effectiveness of the PureCap method, which can remove both 5'-ppp-RNA and dsRNA impurities in the RP-HPLC purification process.

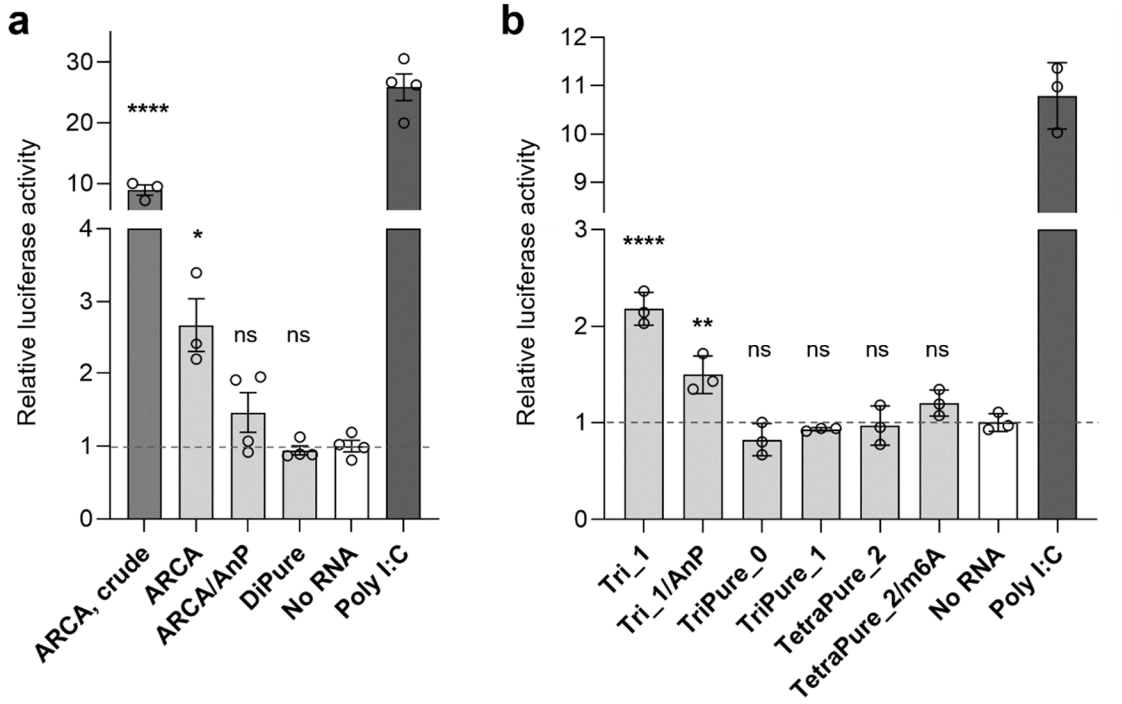


Figure 6. Evaluation of the immunostimulatory effect of capped mRNAs using NF-κB reporter (Luc)–HEK293 cells. 650-nt Nluc mRNA prepared/ purified using dinucleotide cap analogs (**a**) or tri/tetranucleotide cap analogs (**b**) was transfected into the cells using Lipofectamine MessengerMAX reagent. Poly I: C was used as the positive control of this assay. The cap analog used for mRNA preparation is indicated in the graph. AnP means the mRNA was further treated with Antarctic phosphatase to remove its 5' phosphate. ARCA-capped RNA before the HPLC purification was also analyzed (**ARCA, crude**). The expression level was normalized to the mock-transfected control sample (**No RNA**). Data are expressed as the mean ± standard error (n = 3 or 4). The statistically significant differences of Nluc mRNA-transfected samples from the mock-transfected sample in One-way ANOVA test followed by Dunnett test were marked as follows: ns, $p > 0.05$; *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$.

Translation activity measurement of Nluc mRNAs prepared by the PureCap method

Nluc mRNAs were transfected into human cervical carcinoma-derived HeLa cells, a representative of non-immune cells, and mouse immature dendritic cell-derived JAWS II cells, a representative of immune cells, to examine the effect of different cap structures on the translational activity of mRNA (Fig. 7). As shown in Figure 7a, in cultured HeLa cells mRNA prepared with **DiPure** showed 2–3 times higher translation activity than mRNA prepared with **ARCA**. Dephosphorylation of the uncapped mRNA in ARCA-mRNA hardly changed the translational activity (**ARCA** vs. **ARCA/AnP**). Similar results were confirmed using JAWS II cells, where DiPure-mRNA showed approximately 3-fold more potent activity than ARCA-mRNA (Fig. 7b). These differences in

translational activity between DiPure-mRNA and ARCA-mRNA will be attributed mainly to the difference in the capping efficiency. DiPure-mRNA was quantitative, although ARCA-mRNA was about 60% (Fig. 3f). From these results, we confirmed that the preparation of quantitatively capped mRNA using PureCap analogs contributes to improving mRNA translational activity.

To investigate the effect of the methyl group introduced into the 2' or 3' hydroxyl of m^7G on the translational activity, mRNAs prepared with three PureCap analogs, **DiPure**, **DiPure/2'OMe**, and **DiPure/3'OMe** were compared. In cultured HeLa and JAWS II cells, *O*-methylation at the 2' position slightly reduced its translational activity. Still, no apparent change in the activity was observed when one methyl group was introduced into the cap structure (Extended Data Fig 5a)[54]. These experimental results are consistent with *in silico* docking simulations showing that the methyl group introduced at the 3'-*O* position of m^7G of dinucleotide cap analog would not interfere with the interaction of the cap-binding protein V39 of cap-specific mRNA 2'-*O*-methyltransferase or the eukaryotic translation initiation factor eIF4E[21]. In the end, we conclude that **DiPure** is the best of the three analogs because it is the easiest to synthesize, has the best incorporation by T7 RNAP, and exhibits high mRNA translation activity.

The translation activity of mRNA prepared using **DiPure/N2**, in which the Nb modification was introduced into the exocyclic amino group of m^7G, was measured in cultured HeLa cells (Extended Data Fig. 6). The most important finding of this experiment was that the presence of Nb protecting group in the cap significantly suppressed mRNA translation, showing only 1.6% of the activity after deprotection. This observation is consistent with a recent report by Klocker *et al.* that the translation of mRNA containing a cap structure with a photocleavage group on the exocyclic amino group of m^7G could be triggered by light irradiation in cultured cells[30]. They attributed the translation inhibition effect of the photocleavage group to the inhibition of binding to eIF4E[30]. Unlike the case above with **DiPure/N2**, it should be noted that the mRNA prepared with **DiPure** showed almost the same translation activity before and after light irradiation, indicating that the Nb group at the 2' *O*-position does not impair the translational activity.

The translational activity of Nluc mRNAs with Cap-0, Cap-1, and Cap-2 cap structures was compared using cultured HeLa and JAWS II cells (Fig. 7). In these experiments, unlike previous reports[23], there was no apparent difference in activity between mRNAs with Cap-0 and Cap-1 in both HeLa and JAWS II cells. On the other hand, the translational activity of Cap-2-mRNAs introduced using the tetranucleotide PureCap analogs **TetraPure_2** and **TetraPure_2/m6A** were higher than that of mRNAs with Cap-1 and Cap-0[26]. The enhanced translational activity of Cap-2 mRNA was more pronounced in JAWS II cells than in HeLa cells and was up to 4.2-fold more active than Cap-1-mRNA prepared with **Tri-1** cap analog. The results of our experiment, in which the translational activity of mRNA containing the Cap-2 structure is higher than that having the Cap-1 structure in JAWS II cells, differ from the results of a study reported by Drazkowska *et al.*[26]. Still,

the reason for this difference is unclear at this point. For control mRNAs prepared from a DNA template containing the A-inserted promoter, the translation activity of Cap-1-mRNA prepared using the **Tri_1** trinucleotide cap, which had a high capping efficiency of 90%[24], was comparable to that of mRNA prepared using the PureCap analog **TriPure_1**, both in HeLa and JAWS II cells. The activity of Cap-2 mRNA prepared with the tetranucleotide cap **Tetra_2**, in which only about 50% of the cap was introduced, was only about 1/3 of that of Cap-2 mRNA prepared with **TetraPure_2**. This result reflects the low capping efficiency of tetranucleotide cap analogs and indicates the PureCap method's usefulness in introducing Cap-2 structures using tetranucleotide cap analogs with low incorporation efficiency. As for Cap-2 mRNA containing m^6A prepared using **TetraPure_2/m6A,** a slight increase in translational activity was observed in JAWS II cells.

We also compared the translational activity of mRNAs with this Cap-0, 1, and 2 structures in a cell-free translation system based on HeLa cells[55]. In this experiment, Cap-2-mRNAs also showed higher translation activity than Cap-0/1-mRNA, especially the Cap-2 with m^6A (Extended Data Fig. 5b).

In our experiments, since the Cap-2 cap structure was found to increase the translational activity of its mRNA compared to the Cap-1 cap structure in cultured cell lines, these translational activities were subsequently compared in animal experiments using mice. mRNA was encapsulated into ionizable lipid-based lipid nanoparticles (LNPs) for systemic administration. *In vivo* luminescence imaging revealed that **TetraPure_2** and **TetraPure_2/m6A** tended to show enhanced Nluc expression compared to the Cap-1-containing mRNA prepared with the PureCap analog **TriPure_1**, or control compound **Tri_1** (Fig. 8a). LNPs provided strong Nluc expression in the liver. At the same time, the expression in the spleen may also contribute to the luminescence from the left mouse side. This observation led us to evaluate the Nluc expression level more quantitatively using the liver and spleen homogenates. As observed in the *in vivo* imaging, Cap-2-containing mRNA showed approximately 3-fold higher translation in the liver and spleen with statistical significance compared to the Cap-1-containing mRNA in Nluc quantification from the homogenates (Fig. 8b, c). To the best of our knowledge, this result demonstrated the advantage of Cap2-mRNA over Cap1-mRNA in increasing protein expression efficiency in animals for the first time.

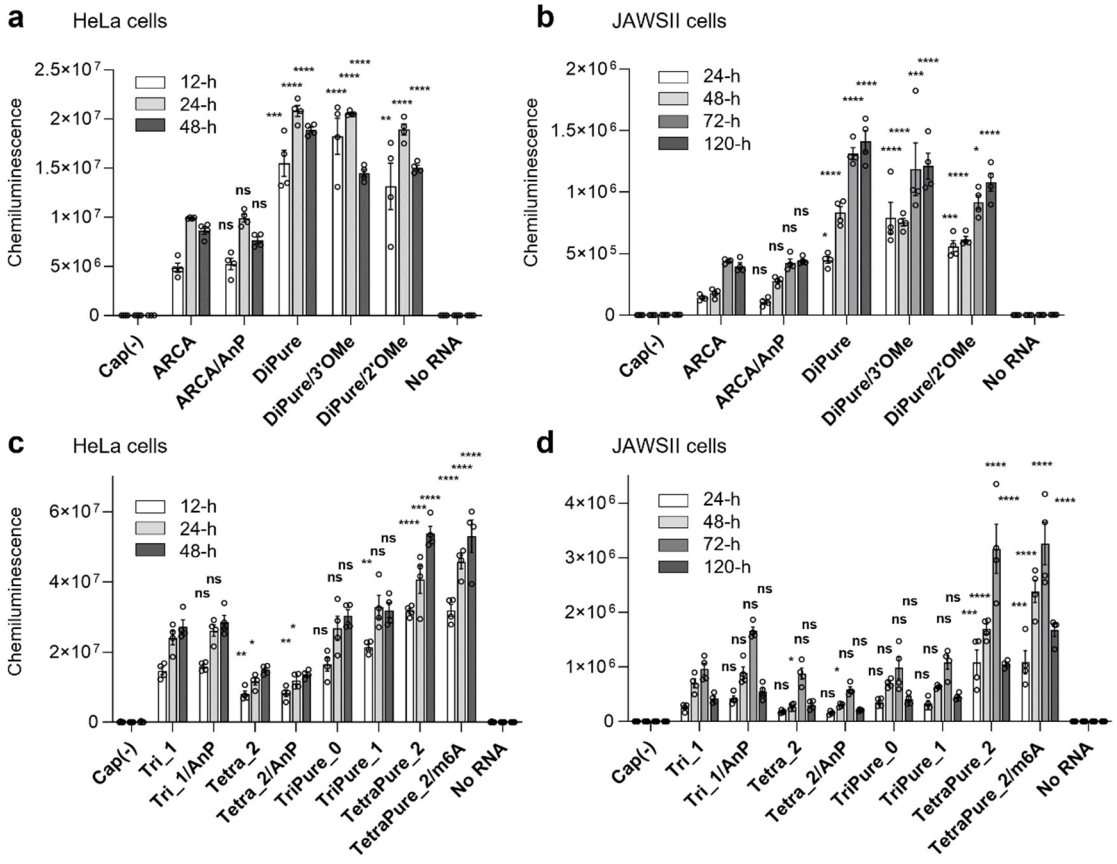


Figure 7. Translation activities of Nluc mRNAs prepared using PureCap analogs in cultured mammalian cell lines. Capped mRNAs prepared using dinucleotide (**a, b**) or tri/tetranucleotide (**c, d**) cap analogs were introduced into HeLa cells (**a, c**) or JAWS II cells (**b, d**). Cells were incubated for the described time and then lysed, and Nluc expression was measured. The cap analog used for mRNA preparation is indicated in the graph. AnP means the mRNA was further treated with Antarctic phosphatase to remove its 5' phosphate. Data are expressed as the mean ± standard error ($n = 4$). The statistically significant differences for the capped RNA samples from "**ARCA**" (**a**, **b**) or "**Tri_1**" (**c**, **d**) in One-way ANOVA followed by Dunnett's test were marked as follows. ns, > 0.05; *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$.

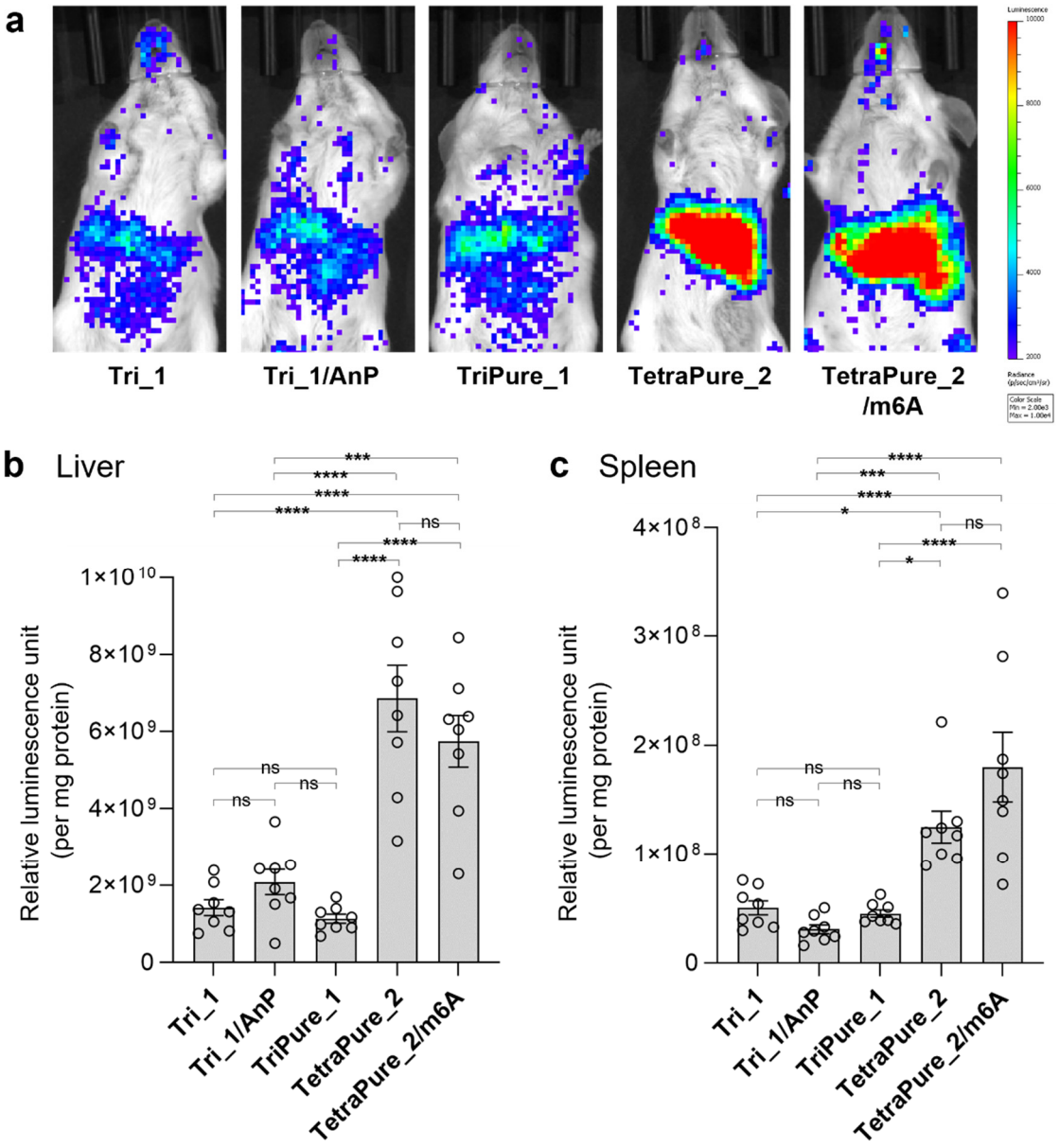


Figure 8. Translation activity of 650-nt Nluc mRNA in mice. The expression level of Nluc was evaluated four hours after intravenous injection of mRNA lipid nanoparticles (LNPs) into mice. The cap analog used for mRNA preparation is indicated in the Figure. AnP means the mRNA was further treated with Antarctic phosphatase to remove its 5' phosphate. (**a**) *In vivo* images of mice. (**b,c**) Quantifying Nluc expression levels using homogenates of the liver (**b**) and spleen (**c**). Data are expressed as the mean ± standard error (n = 8). The statistically significant differences in One-way ANOVA followed by Tukey's test were marked as follows: ns, $p > 0.05$ (not significant); *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$, ****, $p < 0.0001$.

Preparation/purification of longer mRNA by the PureCap method

We attempted to purify capped mRNAs with different longer sequences using the PureCap method. We found that the separation decreased as the length of the mRNA increased. Figure 9 shows the results of RP-HPLC purification of the capped mRNA when the **DiPure** cap analog was added to the transcription reaction of mRNA encoding the spike protein of SARS-CoV2[3], which consists of 4,247-nt. Notably, the proportion of cap structure in the whole mRNA is 6.5-fold less in the spike mRNA compared to Nluc mRNA (one cap in 4,247-nt in the spike mRNA and one cap in 650-nt in Nluc mRNA), posing more challenges to separating capped mRNA from uncapped mRNA. Even when targeting this sequence, which is 6.5 times longer than Nluc mRNA, the peak of capped mRNA was almost completely separable from that of uncapped mRNA due to the hydrophobicity of the Nb protecting group. Capped mRNA was isolated by repeating this system twice. After isolation, the capped mRNA was digested by DNAzyme. We found that the capped mRNA with this length can also be separable from uncapped mRNA using the hydrophobicity of the PureCap analog (Fig. 9c)

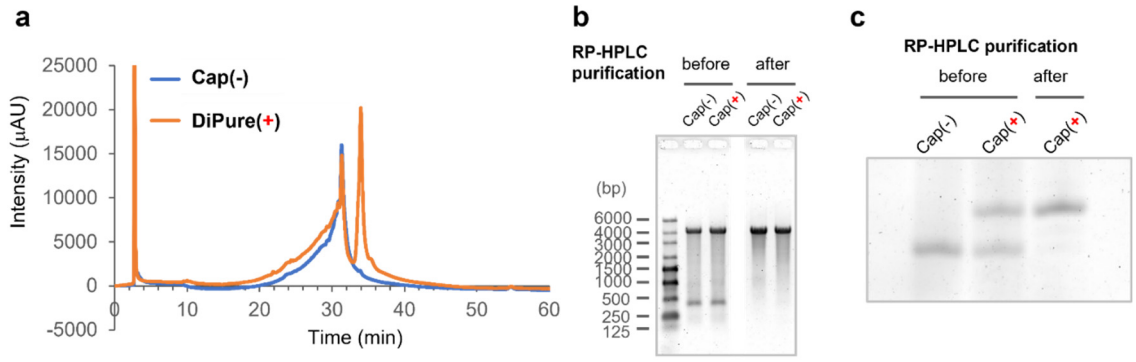


Figure 9. Preparation and purification of capped 4,247-nt mRNA encoding spike protein of SARS-CoV-2 using **DiPure** cap analog. (**a**) RP-HPLC analysis of IVT-RNAs with or without the cap analog. (**b**) Denaturing agarose gel electrophoresis of transcribed RNAs before and after RP-HPLC purification (**c**) Analysis of 5' terminus of mRNA after RP-HPLC purification. After being cleaved by DNAzyme 10-23, it was analyzed by15% dPAGE.

DISCUSSION & CONCLUSION

Various cap analogs with different structural features have been developed to prepare capped mRNAs by IVT[18]. First to mention are the dinucleotide cap analogs ARCAs[20, 21, 37, 54, 56], which were designed to control the direction of introduction of the dinucleotide cap analog m^7G(5')ppp(5')G. It was reported that methylation of the 2' or 3' hydroxyl group completely controlled the direction of strand elongation, resulting in an approximately 2–3-fold increase in mRNA translation activity in cultured cells[54, 56]. Following the successful development of ARCA, many modified analogs for

dinucleotide caps have been developed. In particular, many examples of modification of the phosphate moieties[57, 58]. Extensions of triphosphate linkage to tetra- and pentaphosphates were reported, and improved translation activity of capped mRNA prepared using these analogs was reported[37]. The development of analogs in which bridging or non-bridging oxygen atoms were replaced by S, Se, CH_2, and BH_3 was also reported[57, 58]. Synthesis of cap analogs in which a non-phosphate linkage was introduced into the phosphate moiety by the click reaction has been reported[59, 60]. These modifications in the linkage often increase the translational activity of the mRNA, mainly because they increase its resistance to degradation by the scavenger mRNA-decapping enzyme DcpS. Synthesis of functional cap analogs that can create labeled translatable mRNA by introducing fluorescent groups or biotin into the 2' and 3' hydroxyl groups of m^7G was also reported[61, 62]. The introduction of a benzyl group at the N^2 position of m^7G or the substitution of a 7-methyl group with an arylmethyl group was reported to enhance the translational activity of mRNA in human cultured cells[54, 63, 64]. Some cap analogs have been developed for on/off control of the mRNA translation[30, 65, 66, 67]. For example, Ogasawara *et al.* reported the synthesis of cap analogs that contained photo-responsible groups such as phyenylazo or styryl group or at the N^2 or C^8 position of m^7G and that their cis/trans conformation enabled photo-switchable translational control[65, 66, 67]. As the biological significance of the Cap-1 and Cap-2 structures becomes clearer, tri- and tetranucleotide cap analogs have also been developed to introduce them directly into mRNA[22, 23, 24, 25, 26].

Compared to the development of novel cap structures mentioned above, much less effort has been devoted to improving capping efficiency. The standard co-transcriptional capping method using trinucleotide cap analogs yields a substantial amount of contaminant 5' ppp-RNA, inducing strong innate immunity[28]. Meanwhile, capping efficiency in the method was reported to be 80–90%[29], requiring a further increase in capping efficiency. In addition, T7 RNAP produces an immunostimulatory dsRNA byproduct in mRNA preparation[32, 33]. We undertook the present study to remove these byproducts for the therapeutic use of mRNA. The novelty of this study is designing cap analogs for increasing capping efficiency, while previous studies derivatized numerous cap analogs for improving cap functionalities as mentioned above; the PureCap method we are developing here is unprecedented in that it removes 5' ppp-RNA by chromatography which can remove dsRNA at the same time[32]. Indeed, we succeeded in completely suppressing innate immune responses without enzymatic treatment, such as Antarctic phosphatase (Fig. 6). Further notably, the hydrophobic tag introduced for purification can be removed simply by light irradiation, leaving no unnatural modification in the capped mRNA.

PureCap technology also provides a reliable platform to study structure-activity the relationship of cap analogs. In co-transcriptional capping, different cap structures resulted in different capping efficiency (Fig. 5g), causing bias for their comparison. Indeed, our PureCap method achieved 100%

capping efficiency in various capping structures, including Cap-0, Cap-1, and Cap-2, with or without modification in 2' or 3' hydroxyl of m^7G and N^6 position of the first transcribed A (m^6A). Especially, 100% capping efficiency was obtained in preparing Cap-2 mRNA, which shows low capping efficiency in conventional methods (Fig. 5g). Cap-2 mRNA thus prepared showed up to 3 to 4-fold higher translational activity in cultured cells and animals than mRNA prepared by standard capping methods without inducing innate immune responses (Figs. 7,8). Intriguingly, Cap-2 mRNA improved protein translational efficiency by approximately 3-fold in the liver and spleen, two major target organs of mRNA therapeutics. Numerous studies and clinical developments target the liver for treating single gene disorders providing therapeutic proteins systemically[68, 69], and the spleen for cancer vaccines[70], immunotolerance vaccines[71], and chimeric antigen receptor T cell potentiation[72]. The Cap-2 mRNA technology is capable of potentiating these therapeutics.

One issue to be resolved in the future is the low incorporation efficiency of the tetranucleotide cap analogs to give the Cap-2 structure in mRNA, which was about 25% of the total. Currently, wild-type T7 RNA polymerase is used, but in the future, we would like to improve this, for example, with modification of the polymerase.

In summary, we synthesized quantitatively capped mRNAs by the PureCap method using synthesized hydrophobic cap analogs. To the best of our knowledge, our system provides a footprint-free purification method of capped mRNA with 100% capping efficiency for the first time. Practically, this technology requires only RP-HPLC for purifying capped mRNA, avoiding repeated enzymatic treatment commonly performed for reducing mRNA immunogenicity. Achieving 100% efficiency in various cap species, our approach provides an unbiased platform to study their structure-activity relationship by excluding the influence of different capping efficiency observed among cap species. In this study, we prepared fully capped Cap-2 mRNA, which shows low capping efficiency in conventional methods. Ultimately, Cap-2 mRNA thus prepared showed up to 3 to 4-fold higher translational activity in cultured cells and animals than mRNA prepared by standard capping methods without inducing innate immune responses. Worth notably, Cap-2 mRNA improved protein translational efficiency in the liver and spleen, two major target organs of mRNA therapeutics, providing an excellent tool for future mRNA-based medicines.

REFERENCES

1. Wolff JA, Malone RW, Williams P, Chong W, Acsadi G, Jani A*, et al.* Direct Gene Transfer into Mouse Muscle in Vivo. *Science* 1990, **247**(4949)**:** 1465-1468.

2. Sahin U, Kariko K, Tureci O. mRNA-based therapeutics - developing a new class of drugs. *Nat Rev Drug Discov* 2014, **13**(10)**:** 759-780.

3. Polack FP, Thomas SJ, Kitchin N, Absalon J, Gurtman A, Lockhart S*, et al.* Safety and Efficacy of the BNT162b2 mRNA Covid-19 Vaccine. *New Engl J Med* 2020, **383**(27)**:** 2603-2615.

4. Baden LR, El Sahly HM, Essink B, Kotloff K, Frey S, Novak R*, et al.* Efficacy and Safety of the mRNA-1273 SARS-CoV-2 Vaccine. *New Engl J Med* 2021, **384**(5)**:** 403-416.

5. Ramanathan A, Robb GB, Chan SH. mRNA capping: biological functions and applications. *Nucleic Acids Res* 2016, **44**(16)**:** 7511-7526.

6. Pelletier J, Schmeing TM, Sonenberg N. The multifaceted eukaryotic cap structure. *WIREs RNA* 2021, **12**(2)**:** e1636.

7. Galloway A, Atrih A, Grzela R, Darzynkiewicz E, Ferguson MAJ, Cowling VH. CAP-MAP: cap analysis protocol with minimal analyte processing, a rapid and sensitive approach to analysing mRNA cap structures. *Open Biol* 2020, **10**(2)**:** 190306.

8. Abbas YM, Laudenbach BT, Martinez-Montero S, Cencic R, Habjan M, Pichlmair A*, et al.* Structure of human IFIT1 with capped RNA reveals adaptable mRNA binding and mechanisms for sensing N1 and N2 ribose 2 '-O methylations. *Proc Natl Acad Sci USA* 2017, **114**(11)**:** E2106-E2115.

9. Devarkar SC, Wang C, Miller MT, Ramanathan A, Jiang FG, Khan AG*, et al.* Structural basis for m7G recognition and 2 '-O-methyl discrimination in capped RNAs by the innate immune receptor RIG-I. *Proc Natl Acad Sci USA* 2016, **113**(3)**:** 596-601.

10. Zust R, Cervantes-Barragan L, Habjan M, Maier R, Neuman BW, Ziebuhr J*, et al.* Ribose 2 '-O-methylation provides a molecular signature for the distinction of self and non-self mRNA dependent on the RNA sensor Mda5. *Nat Immunol* 2011, **12**(2)**:** 137-143.

11. Thompson MG, Sacco MT, Horner SM. How RNA modifications regulate the antiviral response. *Immunol Rev* 2021, **304**(1)**:** 169-180.

12. Mauer J, Luo XB, Blanjoie A, Jiao XF, Grozhik AV, Patil DP*, et al.* Reversible methylation of m(6)A(m) in the 5 ' cap controls mRNA stability. *Nature* 2017, **541**(7637)**:** 371-375.

13. Ohkubo A, Kondo Y, Suzuki M, Kobayashi H, Kanamori T, Masaki Y*, et al.* Chemical Synthesis of U1 snRNA Derivatives. *Org Lett* 2013, **15**(17)**:** 4386-4389.

14. Abe N, Imaeda A, Inagaki M, Li ZM, Kawaguchi D, Onda K*, et al.* Complete Chemical Synthesis of Minimal Messenger RNA by Efficient Chemical Capping Reaction. *ACS Chem Biol* 2022, **17**(6)**:** 1308-1314.

15. Thillier Y, Decroly E, Morvan F, Canard B, Vasseur JJ, Debart F. Synthesis of 5 ' cap-0 and cap-1 RNAs using solid-phase chemistry coupled with enzymatic methylation by human (guanine-N-7)-methyl transferase. *RNA* 2012, **18**(4)**:** 856-868.

16. Muthmann N, Guez T, Vasseur JJ, Jaffrey SR, Debart F, Rentmeister A. Combining Chemical Synthesis and Enzymatic Methylation to Access Short RNAs with Various 5 ' Caps. *ChemBioChem* 2019, **20**(13)**:** 1693-1700.

17. Ensinger MJ, Martin SA, Paoletti E, Moss B. Modification of the 5'-terminus of mRNA by soluble guanylyl and methyl transferases from vaccinia virus. *Proc Natl Acad Sci USA* 1975, **72**(7)**:** 2525-2529.

18. Shanmugasundaram M, Senthilvelan A, Kore AR. Recent Advances in Modified Cap Analogs: Synthesis, Biochemical Properties, and mRNA Based Vaccines. *Chem Rec* 2022, **22**(8)**:** e202200005.

19. Grudzien-Nogalska E, Stepinski J, Jemielity J, Zuberek J, Stolarski R, Rhoads RE*, et al.* Synthesis of anti-reverse cap analogs (ARCAs) and their applications in mRNA translation and stability. *Methods Enzymol* 2007, **431:** 203-227.

20. Stepinski J, Waddell C, Stolarski R, Darzynkiewicz E, Rhoads RE. Synthesis and properties of mRNAs containing the novel "anti-reverse" cap analogs 7-methyl(3 '-O-methyl)GpppG and 7-methyl(3 '-deoxy)GpppG. *RNA* 2001, **7**(10)**:** 1486-1495.

21. Peng ZH, Sharma V, Singleton SF, Gershon PD. Synthesis and application of a chain-terminating dinucleotide mRNA cap analog. *Org Lett* 2002, **4**(2)**:** 161-164.

22. Ishikawa M, Murai R, Hagiwara H, Hoshino T, Suyama K. Preparation of eukaryotic mRNA having differently methylated adenosine at the 5′-terminus and the effect of the methyl group in translation. *Nucleic Acids Symp Ser* 2009, **53**(1)**:** 129-130.

23. Sikorski PJ, Warminski M, Kubacka D, Ratajczak T, Nowis D, Kowalska J*, et al.* The identity and methylation status of the first transcribed nucleotide in eukaryotic mRNA 5 ' cap modulates protein expression in living cells. *Nucleic Acids Res* 2020, **48**(4)**:** 1607-1626.

24. Henderson JM, Ujita A, Hill E, Yousif-Rosales S, Smith C, Ko N*, et al.* Cap 1 Messenger RNA Synthesis with Co-transcriptional CleanCap(®) Analog by In Vitro Transcription. *Curr Protoc* 2021, **1**(2)**:** e39.

25. Senthilvelan A, Vonderfecht T, Shanmugasundaram M, Pal I, Potter J, Kore AR. Trinucleotide Cap Analogue Bearing a Locked Nucleic Acid Moiety: Synthesis, mRNA Modification, and Translation for Therapeutic Applications. *Org Lett* 2021, **23**(11)**:** 4133-4136

26. Drazkowska K, Tomecki R, Warminski M, Baran N, Cysewski D, Depaix A*, et al.* 2′-O-Methylation of the second transcribed nucleotide within the mRNA 5′ cap impacts the protein production level in a cell-specific manner and contributes to RNA immune evasion. *Nucleic Acids Res* 2022**:** https://doi.org/10.1093/nar/gkac1722.

27. Mann CDC, Hornung V. Molecular mechanisms of nonself nucleic acid recognition by the innate immune system. *Eur J Immunol* 2021, **51**(8)**:** 1897-1910.
28. Moradian H, Roch T, Anthofer L, Lendlein A, Gossen M. Chemical modification of uridine modulates mRNA-mediated proinflammatory and antiviral response in primary human macrophages. *Mol Ther Nucl Acids* 2022, **27:** 854-869.
29. Vlatkovic I, Ludwig J, Boros G, Szabo GT, Reichert J, Buff M*, et al.* Ribozyme Assays to Quantify the Capping Efficiency of In Vitro-Transcribed mRNA. *Pharmaceutics* 2022, **14**(2)**:** 328.
30. Klocker N, Weissenboeck FP, van Dulmen M, Spacek P, Huwel S, Rentmeister A. Photocaged 5 ' cap analogues for optical control of mRNA translation in cells. *Nat Chem* 2022, **14**(8)**:** 905-913.
31. Klan P, Solomek T, Bochet CG, Blanc A, Givens R, Rubina M*, et al.* Photoremovable Protecting Groups in Chemistry and Biology: Reaction Mechanisms and Efficacy. *Chem Rev* 2013, **113**(1)**:** 119-191.
32. Kariko K, Muramatsu H, Ludwig J, Weissman D. Generating the optimal mRNA for therapy: HPLC purification eliminates immune activation and improves translation of nucleoside-modified, protein-encoding mRNA. *Nucleic Acids Res* 2011, **39**(21)**:** e142.
33. Mu X, Greenwald E, Ahmad S, Hur S. An origin of the immunogenicity of in vitro transcribed RNA. *Nucleic Acids Res* 2018, **46**(10)**:** 5239-5249.
34. Baiersdorfer M, Boros G, Muramatsu H, Mahiny A, Vlatkovic I, Sahin U*, et al.* A Facile Method for the Removal of dsRNA Contaminant from In Vitro-Transcribed mRNA. *Mol Ther Nucl Acids* 2019, **15:** 26-35.
35. Wu MZ, Asahara H, Tzertzinis G, Roy B. Synthesis of low immunogenicity RNA with high-temperature in vitro transcription. *RNA*
36. Litosh VA, Wu WD, Stupi BP, Wang JC, Morris SE, Hersh MN*, et al.* Improved nucleotide selectivity and termination of 3'-OH unblocked reversible terminators by molecular tuning of 2-nitrobenzyl alkylated HOMedU triphosphates. *Nucleic Acids Res* 2011, **39**(6)**:** e39.
37. Jemielity J, Fowler T, Zuberek J, Stepinski J, Lewdorowicz M, Niedzwiecka A*, et al.* Novel "anti-reverse" cap analogs with superior translational properties. *RNA* 2003, **9**(9)**:** 1108-1122.
38. Besada P, Shin DH, Costanzi S, Ko H, Mathe C, Gagneron J*, et al.* Structure-activity relationships of uridine 5 '-diphosphate analogues at the human P2Y(6) receptor. *J Med Chem* 2006, **49**(18)**:** 5532-5543.
39. El-Tayeb A, Qi AD, Muller CE. Synthesis and structure-activity relationships of uracil nucleotide derivatives and analogues as agonists at human P2Y(2), P2Y(4), and P2Y(6) receptors. *J Med Chem* 2006, **49**(24)**:** 7076-7087.

40. Toti KS, Jain S, Ciancetta A, Balasubramanian R, Chakraborty S, Surujdin R*, et al.* Pyrimidine nucleotides containing a (S)-methanocarba ring as P2Y(6) receptor agonists. *MedChemComm* 2017, **8**(10)**:** 1897-1908.

41. Davisson VJ, Davis DR, Dixit VM, Poulter CD. Synthesis of nucleotide 5'-diphosphates from 5'-O-tosyl nucleosides. *J Org Chem* 1987, **52**(9)**:** 1794-1801.

42. Liang FT, Jain N, Hutchens T, Shock DD, Beard WA, Wilson SH*, et al.* alpha,beta-Methylene-2 '-deoxynucleoside 5 '-Triphosphates as Noncleavable Substrates for DNA Polymerases: Isolation, Characterization, and Stability Studies of Novel 2 '-Deoxycyclonucleosides, 3,5 '-Cyclo-dG, and 2,5 '-Cyclo-dT. *J Med Chem* 2008, **51**(20)**:** 6460-6470.

43. Wu W, Bergstrom DE, Jo Davisson V. Chemoenzymatic preparation of nucleoside triphosphates. *Curr Protoc Nucleic Acid Chem* 2004, **Chapter 13:** Unit 13.12.

44. Korhonen HJ, Bolt HL, Hodgson DRW. A procedure for the preparation and isolation of nucleoside-5 '-diphosphates. *Beilstein J Org Chem* 2015, **11:** 469-472.

45. Featherston AL, Kwon Y, Pompeo MM, Engl OD, Leahy DK, Miller SJ. Catalytic asymmetric and stereodivergent oligonucleotide synthesis. *Science* 2021, **371**(6530)**:** 702-707.

46. Hall MP, Unch J, Binkowski BF, Valley MP, Butler BL, Wood MG*, et al.* Engineered Luciferase Reporter from a Deep Sea Shrimp Utilizing a Novel Imidazopyrazinone Substrate. *ACS Chem Biol* 2012, **7**(11)**:** 1848-1857.

47. Santoro SW, Joyce GF. A general purpose RNA-cleaving DNA enzyme. *Proc Natl Acad Sci USA* 1997, **94**(9)**:** 4262-4266.

48. Pleiss JA, Derrick ML, Uhlenbeck OC. T7 RNA polymerase produces 5 ' end heterogeneity during in vitro transcription from certain templates. *RNA* 1998, **4**(10)**:** 1313-1317.

49. Coleman TM, Wang GC, Huang FQ. Superior 5 ' homogeneity of RNA from ATP-initiated transcription under the T7 phi 2.5 promoter. *Nucleic Acids Res* 2004, **32**(1)**:** e14.

50. Dunn JJ, Studier FW. Complete nucleotide sequence of bacteriophage T7 DNA and the locations of T7 genetic elements. *J Mol Biol* 1983, **166**(4)**:** 477-535.

51. Diamond MS, Farzan M. The broad-spectrum antiviral functions of IFIT and IFITM proteins. *Nat Rev Immunol* 2013, **13**(1)**:** 46-57.

52. van Dulmen M, Muthmann N, Rentmeister A. Chemo-Enzymatic Modification of the 5 ' Cap Maintains Translation and Increases Immunogenic Properties of mRNA. *Angew Chem Int Ed* 2021, **60**(24)**:** 13280-13286.

53. Ren XM, Linehan MM, Iwasaki A, Pyle AM. RIG-I Selectively Discriminates against 5 '-Monophosphate RNA. *Cell Rep* 2019, **26**(8)**:** 2019-2027.

54. Kocmik I, Piecyk K, Rudzinska M, Niedzwiecka A, Darzynkiewicz E, Grzela R*, et al.* Modified ARCA analogs providing enhanced translational properties of capped mRNAs. *Cell Cycle* 2018, **17**(13)**:** 1624-1636.

55. Kobayashi T, Machida K, Imataka H. Human Cell Extract-Derived Cell-Free Systems for Virus Synthesis. *Methods Mol Biol* 2014, **1118:** 149-156.
56. Kore AR, Shanmugasundaram M, Charles I, Vlassov AV, Barta TJ. Locked Nucleic Acid (LNA)-Modified Dinucleotide mRNA Cap Analogue: Synthesis, Enzymatic Incorporation, and Utilization. *J Am Chem Soc* 2009, **131**(18)**:** 6364-6365.
57. Rydzik AM, Warminski M, Sikorski PJ, Baranowski MR, Walczak S, Kowalska J*, et al.* mRNA cap analogues substituted in the tetraphosphate chain with CX2: identification of O-to-CCl2 as the first bridging modification that confers resistance to decapping without impairing translation. *Nucleic Acids Res* 2017, **45**(15)**:** 8661-8675.
58. Warminski M, Sikorski PJ, Kowalska J, Jemielity J. Applications of Phosphate Modification and Labeling to Study (m) RNA Caps. *Topics Curr Chem* 2017, **375**(1)**:** 16.
59. Walczak S, Nowicka A, Kubacka D, Fac K, Wanat P, Mroczek S*, et al.* A novel route for preparing 5 ' cap mimics and capped RNAs: phosphate-modified cap analogues obtained via click chemistry. *Chem Sci* 2017, **8**(1)**:** 260-267.
60. Walczak S, Sikorski PJ, Kasprzyk R, Kowalska J, Jemielity J. Exploring the potential of phosphotriazole 5 ' mRNA cap analogues as efficient translation initiators. *Org Biomol Chem* 2018, **16**(36)**:** 6741-6748.
61. Ziemniak M, Szabelski M, Lukaszewicz M, Nowicka A, Darzynkiewicz E, Rhoads RE*, et al.* Synthesis and evaluation of fluorescent cap analogues for mRNA labelling. *RSC Adv* 2013, **3**(43)**:** 20943-20958.
62. Jemielity J, Lukaszewicz M, Kowalska J, Czarnecki J, Zuberek J, Darzynkiewicz E. Synthesis of biotin labelled cap analogue - incorporable into mRNA transcripts and promoting cap-dependent translation. *Org Biomol Chem* 2012, **10**(43)**:** 8570-8574.
63. Wojcik R, Baranowski MR, Markiewicz L, Kubacka D, Bednarczyk M, Baran N*, et al.* Novel N7-Arylmethyl Substituted Dinucleotide mRNA 5 ' cap Analogs: Synthesis and Evaluation as Modulators of Translation. *Pharmaceutics* 2021, **13**(11)**:** 1941.
64. Grudzien E, Stepinski J, Jankowska-Anyszka M, Stolarski R, Darzynkiewicz E, Rhoads RE. Novel cap analogs for in vitro synthesis of mRNAs with high translational efficiency. *RNA* 2004, **10**(9)**:** 1479-1487.
65. Ogasawara S. Control of Cellular Function by Reversible Photoregulation of Translation. *ChemBioChem* 2014, **15**(18)**:** 2652-2655.
66. Ogasawara S. Duration Control of Protein Expression in Vivo by Light-Mediated Reversible Activation of Translation. *ACS Chem Biol* 2017, **12**(2)**:** 351-356.
67. Ogasawara S, Maeda M. Photoresponsive 5 '-cap for the reversible photoregulation of gene expression. *Bioorg Med Chem Lett* 2011, **21**(18)**:** 5457-5459.

68. Stadler CR, Bahr-Mahmud H, Celik L, Hebich B, Roth AS, Roth RP*, et al.* Elimination of large tumors in mice by mRNA-encoded bispecific antibodies. *Nat Med* 2017, **23**(7)**:** 815-817.
69. Kose N, Fox JM, Sapparapu G, Bombardi R, Tennekoon RN, de Silva AD*, et al.* A lipid-encapsulated mRNA encoding a potently neutralizing human monoclonal antibody protects against chikungunya infection. *Sci Immunol* 2019, **4**(35)**:** eaaw6647.
70. Sahin U, Oehm P, Derhovanessian E, Jabulowsky RA, Vormehr M, Gold M*, et al.* An RNA vaccine drives immunity in checkpoint-inhibitor-treated melanoma. *Nature* 2020, **585**(7823)**:** 107-112.
71. Krienke C, Kolb L, Diken E, Streuber M, Kirchhoff S, Bukur T*, et al.* A noninflammatory mRNA vaccine for treatment of experimental autoimmune encephalomyelitis. *Science* 2021, **371**(6525)**:** 146-153.
72. Reinhard K, Rengstl B, Oehm P, Michel K, Billmeier A, Hayduk N*, et al.* An RNA vaccine drives expansion and efficacy of claudin-CAR-T cells against solid tumors. *Science* 2020, **367**(6476)**:** 446-453.

METHODS

Preparation of Nluc mRNA by IVT and its purification by RP-HPLC. The sequence of 650-nt Nluc mRNA is described in the Supporting Information file. Template DNA for IVT was prepared by PCR reaction using pNL1.1 TK vector (Promega) as a template. The sequence of the PCR primers was as follows: Forward (Type III ϕ6.5 promoter-containing), 5' CCCGGATCCTAATACGACTCACTATAGGCGCATATTAAGGTGACGCGT 3'; Reverse, 5' $(T)_{30}$-CTAGAATTACGCCAGAATGCG 3'. To prepare DNA containing different T7 promoters as Type II ϕ2.5 or the one named "A-inserted," a different forward primer was used as described below: For Type II ϕ2.5: 5' CCCGGATCCTAATACGACTCACTATTAGGCGCATATTAAGGTGACGCGT 3': For "A-inserted"; 5' CCCGGATCCTAATACGACTCACTATAAGGCGCATATTAAGGTGACGCGT 3'.The PCR mixture was consisted of 0.5 μM primers, 1 ng/μL pNL1.1TK vector, 0.2 mM dNTPs, 1.5 mM $MgSO_4$, 1×PCR Buffer for KOD -Plus- Neo, 0.02 units/μL KOD -Plus- Neo (Toyobo). The mixture was subjected to the thermal cycling reaction as follows: 94 °C, 2 min → (98 °C, 10 s → 55 °C, 30 s →68 °C, 90 s) × 25 cycles → 68 °C, 5 min. The reaction was analyzed by agarose gel electrophoresis, and dsDNA was recovered by alcohol precipitation after TE-saturated phenol/ chloroform (1:1) extraction. Typical transcription reaction was carried out in a reaction mixture containing 15 ng/μL DNA (PCR product), 2 mM NTPs, 2 mM cap analog, 40 mM Tris-HCl (pH 8.0), 8 mM $MgCl_2$, 2 mM spermidine, 5 mM DTT, 0.002 U/μL inorganic (yeast) pyrophosphatase (New England Biolabs), 9.4 ng/μL T7 RNA polymerase. For the control cap analogs, ARCA was purchased from Jena Bioscience, while Tri_1 and Tetra_2 were synthesized in-house. After the mixture was incubated at 37 ºC for 2 hours, DNase I (Takara bio) was added to the mixture at a final concentration of 0.1 units/μL, and it was further incubated at 37 ºC for 15 min. After the mixture was extracted with TE-saturated phenol/ chloroform and chloroform, RNA was recovered by alcohol precipitation. Progress of the IVT reaction or mRNA purity after being purified was confirmed by denaturing PAGE analysis [5(w/v) % acrylamide (ratio of acrylamide and bis-acrylamide was 19:1), 7.5 M urea, 1× Tris-borate-EDTA (TBE) buffer (89 mM Tris, 89 mM boric acid, 2 mM EDTA, pH 8.3)]. The gel was stained with SYBR Green II and visualized on the ChemiDoc Touch MP imaging system (Bio-Rad). Nluc mRNA was analyzed and purified by RP-HPLC using YMC TriartBio C4 column (250 × 4.6 mm I.D., S-5 μm, 12 nm) on Hitachi Chromaster or LaChrom HPLC system with Solution_A (50 mM triethylammonium acetate (TEAA, pH 7.0) containing 5% acetonitrile) and Solution_B (acetonitrile) at a flow rate of 1 mL/min. The content of Solution_B was raised from 0 to 20% over 20 min. The column temperature was maintained at 50 °C. After purification of mRNA by RP-HPLC, RNA was recovered by alcohol-precipitation from the eluate in the presence of NaOAc (pH 5.2) and 2-propanol. RNA concentration was determined by measuring absorbance at 260 nm on NanoDrop 2000 spectrophotometer (Thermo). Nb-protection of the capped mRNAs prepared using

PureCap analogs was removed by 365-nm light irradiation to the mRNA. mRNA solution was transferred into a well of 96-well multi-well plate at a volume of 50 μL/well. 365-nm light was irradiated to the RNA at 4 mW/cm^2 for 10 min at room temperature using MAX-350 compact xenon light source (Asahi Spectra). After light irradiation, the progress of the deprotection reaction was confirmed by the change in the elution time in RP-HPLC analysis. The deprotected capped mRNA was purified by RP-HPLC under the same conditions as above to remove residual dsRNA contamination. RP-HPLC purified control mRNAs were dephosphorylated using Antarctic phosphatase (AnP, New England Biolabs) in a reaction mixture containing 250 ng/μL Nluc mRNA, 50 mM Bis-Tris-propane-HCl, 1 mM $MgCl_2$, 0.1 mM $ZnCl_2$ (pH 6), 0.25 U/μL AnP. After being incubated at 37 °C for 30 min, the mixture was extracted with TE-saturated phenol/ chloroform (1:1) and chloroform. RNA was then recovered by alcohol precipitation.

Dot blot analysis to detect dsRNA. mRNA samples were spotted onto the Amersham Hybond-N$^+$ membrane and air-dried. After being blocked in TBT-T [50 mM Tris-HCl (pH 7.4), 150 mM NaCl, 0.05%(v/v) Tween-20] containing 5% (w/v) skim milk (Wako) for 1 h at room temperature, the blot was incubated with anti-dsRNA clone rJ2 (Sigma-Aldrich) for 1 h at room temperature, which 1,000-fold diluted with TBS-T containing 0.5% skim milk. After being washed with TBS-T, the blot was further incubated with anti-mouse IgG-HRP (Sigma-Aldrich) for 1 h at room temperature, diluted by 5,000-fold with TBS-T containing 0.5% skim milk. After washing with TBS-T, the blot was incubated with SuperSignal West Femto Maximum Sensitivity Substrate (Thermo) and visualized on a ChemiDoc Touch MP imaging system (Bio-Rad). siRNA Ladder Marker (Takara) was used as a positive control of dsRNA.

RNA cleavage using a DNAzyme 10-23 to analyze the 5' end of mRNA. DNAzyme 10-23 was designed according to a previous report[47]. The sequences used in this study are listed as follows (catalytic domain underlined): for Nluc mRNA, 5' TTCGAGGCCA<u>GGCTAGCTACAACGA</u>ACGCGTCACC 3'; for 34/35-nt short RNA, 5' TTGTAGTCCA<u>GGCTAGCTACAACGA</u>CGGATATATCTCCT 3'; for spike protein mRNA, 5' TCTGTGGGGA<u>GGCTAGCTACAACGA</u>CAGAAGAATACTAG 3'. Target RNA (0.5 μM) and DNA (1 μM) were incubated in a buffer composed of 50 mM Tris-HCl (pH 8.0) and 50 mM $MgCl_2$ at 37 C for 1 h. The mixture was then alcohol precipitated. 1.57 μg RNA for 650-nt Nluc mRNA was digested in a 15 μL reaction and analyzed by 15(w/v)% denaturing PAGE containing 7.5 M urea and 1 × TBE. The gel was stained with SYBR Gold and visualized on a ChemiDoc Touch MP imager. As for MALDI-TOF MS analysis of the reaction, the reaction mixture (10 μL) was twice alcohol-precipitated using ammonium acetate to remove sodium ions. It was measured in a positive mode

with an UltrafleXtreme MALDI-TOF/ TOF mass spectrometer (Bruker Daltonics) using 3-hydroxypicolinic acid as a matrix.

Short RNA transcription system using 60-bp DNA as a template. IVT templates containing three different T7 RNA promoter sequences were prepared by the extension reaction of 26-nt DNA oligo annealed to the 60-nt oligo using DNA polymerase. The sequences of the oligos used in this experiment were listed below (5'→3'direction; underlined promoter sequence; N_m, 2'-*O*-methyl-modified).Type III ϕ6.5 containing (60-nt), 5' CCCGGATCCTAATACGACTCACTATAGGGATCCGAAGGAGATATATCCGATGGACTACAA 3'; Type II ϕ2.5 containing (60-nt), 5' CCGGATCCTAATACGACTCACTATTAGGGATCCGAAGGAGATATATCCGATGGACTACAA 3'; A-inserted containing (60-nt), 5' CCGGATCCTAATACGACTCACTATAAGGGATCCGAAGGAGATATATCCGATGGACTACAA 3'; Reverse strand for all (26-nt): 5' U_m_U_m_GTAGTCCATCGGATATATCTCCTT 3'. 60-bp dsDNA IVT templates were prepared as follows. A reaction mixture composed of 2 μM DNA oligos (60-nt and 26-nt), 0.2 mM dNTPs, 1×PrimeSTAR HS Buffer, 0.025 U/μL PrimeSTAR HS DNA polymerase was incubated at 95 °C for 2 min, 55 °C for 1 min and 72 °C for 60 min. It was extracted with TE-saturated phenol/chloroform (1:1) chloroform, and the DNA was alcohol-precipitated. Transcription of 34/ 35-nt RNA was carried out in a reaction mixture containing 15 ng/μL 60-bp dsDNA, 2 mM NTPs, 2 mM PureCap analog, 40 mM Tris-HCl (pH 8.0), 8 mM $MgCl_2$, 2 mM spermidine, 5 mM DTT, 9.4 ng/μL T7 RNA polymerase. After incubation at 37 °C for 2 hours, a portion (0.5 μL) was taken from the mixture and analyzed by 15% denaturing PAGE containing 7.5 M urea. The gel was visualized by SYBR Green II staining.

Immunogenicity measurement of mRNAs. NF-κB reporter (Luc)-HEK293 cells (BPS Bioscience) were grown in Dulbecco's modified Eagle's medium (DMEM; Wako) supplemented with 10% fetal bovine serum (FBS; Invitrogen) and 100 μg/ml of Hygromycin B at 37 °C under 5% CO_2 atmosphere. One day before transfection, the cells were seeded in a 48-well cell culture plate at 1.0×10^5 cells/well. Just before transfection, the medium was replaced with Opti-MEM I Reduced Serum Medium (240 μL/well, Thermo). mRNA (100 ng/well) was mixed with Lipofectamine MessengerMAX (0.15 μL/well) in Opti-MEM I medium (10 μL/well) and added to cells. Three hours after the transfection, DMEM containing 20% FBS was added to the cells (250 μL per well). Twenty-four hours after transfection, the cells were lysed using Glo Lysis Buffer (50 μL/well, Promega), and the luciferase expression was measured using ONE-Glo Luciferase Assay System (Promega). The chemiluminescence was measured on the TriStar5 plate reader (Berthold). The

luciferase expression was normalized by total protein concentration in the lysate measured using Pierce BCA Protein Assay Kit (Thermo).

Translation activity measurement of Nluc mRNA using cultured mammalian cells. HeLa cells (Riken Cell Bank) were grown in DMEM (Wako) supplemented with 10% FBS (Invitrogen) at 37 °C under a 5% CO_2 atmosphere. JAWS II cells (ATCC) were grown in MEM α, nucleoside (Gibco) supplemented with 10% FBS (Invitrogen) and 5 ng/mL murine GM-CSF (PeproTech) at 37 °C under a 5% CO_2 atmosphere. One day before transfection, the cells were seeded in a 96-well cell culture plate, typically at 5×10^3 cells/well for HeLa cells or 1×10^4 cells/well for JAWS II cells. Just before transfection, the medium was replaced with Opti-MEM I medium (40 μL/well, Thermo). mRNA (10 ng/well) mixed with Lipofectamine MessengerMAX (0.15 μL/well) in Opti-MEM I medium (10 μL/well) was added to the cells. Three hours after the transfection, the growing medium was added to the cells (100 μL per well). At indicated time points, the cells were lysed, and the luciferase expression was measured using Nano Glo Luciferase Assay System (Promega).

Translation activity measurement of Nluc mRNA using HeLa cell-based cell-free translation system. Translation mixture containing 1.67 ng/μL Nluc mRNA (650-nt) was incubated at 32 °C, and the expression of Nluc was measured after 0.5-, 1-, 2-, 4- and 6 hours using Nano Glo Luciferase Assay System (Promega). The reaction was composed by 4.35 μL of HeLa S3 cell extract, 3.25 μL of Mixure-2 [12.7 mM DTT, 5.42 mM magnesium acetate, 85.4 mM potassium acetate, 282 mM HEPES-KOH (pH 7.5)], 0.9 μL of Mixture-3 [12.6 mM ATP, 1.22 mM GTP, 200 mM creatine phosphate, 0.6 mg/mL creatine kinase, 0.08–1.2 mM 20-amino acids mixture], 0.5 μL of 30 ng/μL mRNA. The HeLa S3 cell extract was prepared based on a previous report[55], details described in the Supporting Information file.

Translation activity measurement of Nluc mRNA using mice. For lipid nanoparticle (LNP) preparation, 595 μg/mL of mRNA was dissolved in citrate buffer (pH 3.0) and lipid mixture from D-Lin-MC3-DMA (MedChemExpress), 1,2-distearoyl-sn-glycero-3-phosphocholine (DSPC, Wako), cholesterol (Sigma), and 1,2-dimyristoyl-rac-glycero-3-methoxypolyethylene glycol-2000 (DMG-PEG2000, Avanti) was dissolved in ethanol at D-Lin-MC3-DMA/DSPC/Sigma/DMG-PEG2000 molar ratio of 50/10/38.5/1.5 and total ethanol concentration of 35 mM. The citrate buffer and ethanol solution were mixed at a volume ratio of 2 to obtain a nitrogen/phosphate ratio of 5 using microfluidics (NanoAssemblr® Spark, Precision Nanosystems). The solvent was exchanged for PBS by diluting the solution 40-fold and concentrating LNP using Amicon Ultra-15-30K centrifugal units (Merck Millipore). 2 μg of mRNA encapsulated in LNP was injected into Balb/c mouse (female, 7-week-old, Charles River Laboratories Japan, Inc). For in vivo imaging, 5 μg/mouse of furimazine

(Chem Shuttle) was injected from the mouse tail vein. Immediately after the injection, the mouse was observed using IVIS Spectrum SP-BFM-T1 (Perkin Elmer). The liver and spleen were excised at 4 h after the LNP injection for quantification using tissue homogenates. The organs were homogenized with Passive lysis buffer (Promega), followed by luciferase assay using Lumat3 LB9508 luminometer (Berthold Technologies) and Nano-Glo® Luciferase Assay System (Promega). The animal experiments were conducted under the approval of the animal care and use committees at Kyoto Prefectural University of Medicine (Kyoto, Japan) and the Innovation Center of NanoMedicine (iCONM, Kawasaki, Japan).

Preparation and purification of spike protein mRNA of SARS-CoV-2. Reaction conditions in the transcription of capped 4,247-nt mRNA encoding spike protein of SARS-CoV-2 using **DiPure** cap analog were the same as that of 650-nt Nluc mRNA. IVT template was prepared by a PCR reaction using poly-dT_{80}-containing reverse PCR primer to introduce a poly-A_{80} tail to mRNA. mRNA sequence is described in the Supporting Information file. Transcribed mRNA was analyzed and purified by RP-HPLC using Nacalai Cosmosil RNA RP-1 column (200 × 4.6 mm I.D.) on Shimadzu Prominence HPLC system with Solution_A (100 mM TEAA (pH 7.0) containing 5% acetonitrile) and Solution_B (100 mM TEAA (pH 7.0) containing 50% acetonitrile) at a flow rate of 1 mL/min. The content of Solution_B was raised from 10 to 25% over 60 min. The column temperature was maintained at 50 °C.

ACKNOWLEDGEMENTS

This work was supported by Japan Science and Technology Agency (JST) [CREST, JPMJCR18S1 to H. A.) and Japan Agency for Medical Research and Development (AMED) [LEAP, JP21gm0010008 to H. A.]. We thank the Chemical Instrumentation Facility of the Research Center for Materials Science, Nagoya University, for allowing us to conduct the analyses.

AUTHOR CONTRIBUTIONS

MI, NA, HM, YK, and HA designed the cap analogs. MI, ZL, SA, KO, AB, ZM, MT, TI, PL and KK performed the chemical synthesis. NA, YN, DK, HH, FH and SU designed and performed biochemical experiments, including mRNA preparation and cell experiments. SU designed and performed animal experiments. MI, NA, YK, SU and HA wrote the manuscript.

ADDITIONAL INFORMATION

Extended data is available for this paper (Extended Data Table 1 and Figures 1 to 6).

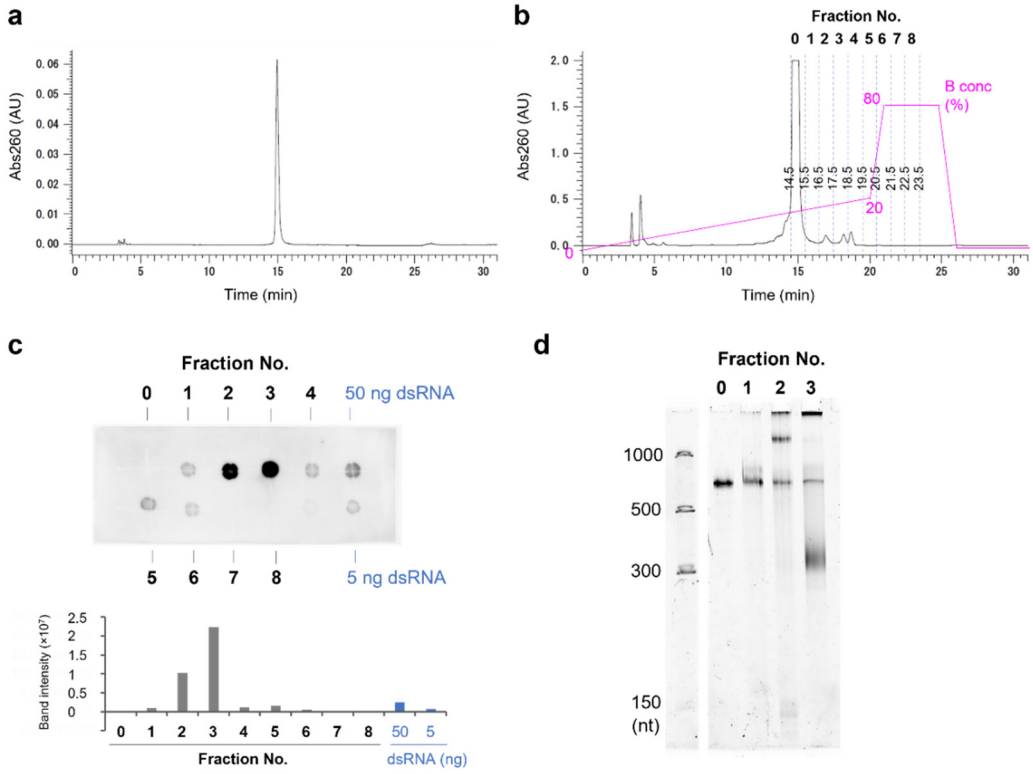
Supplementary Information is also available, which includes Supporting Schemes, experimental details for organic synthesis and analytical data for the synthesized compounds.

Competing financial interests

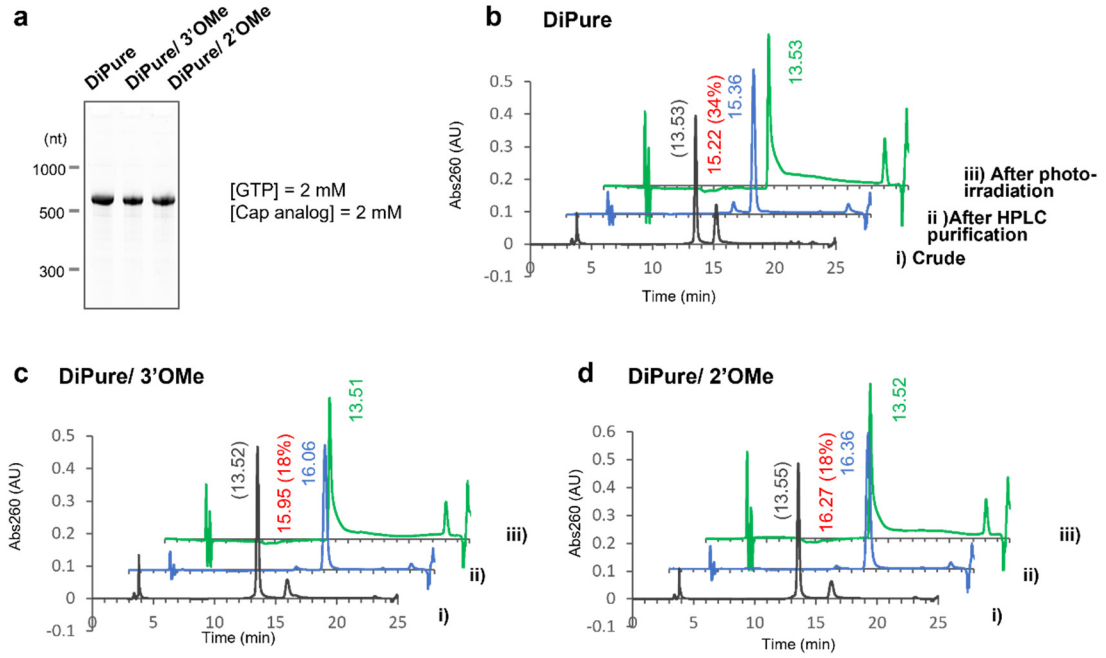
HA and SU are cofounder of Crafton Biotechnology. The company focus on the development of mRNA therapeutics. A international patent application covering part of this work has been filed by Nagoya university.

Extended Data Table 1. MALDI TOF MS data of 5' fragments originated from capped mRNAs after cleavage by DNAzyme 10-23. Calculated and observed masses were shown. The sequence shown as **RNA** is 5' -GCGCAUAUUAAGGUGACGCGUG-p (2',3' cyclic).

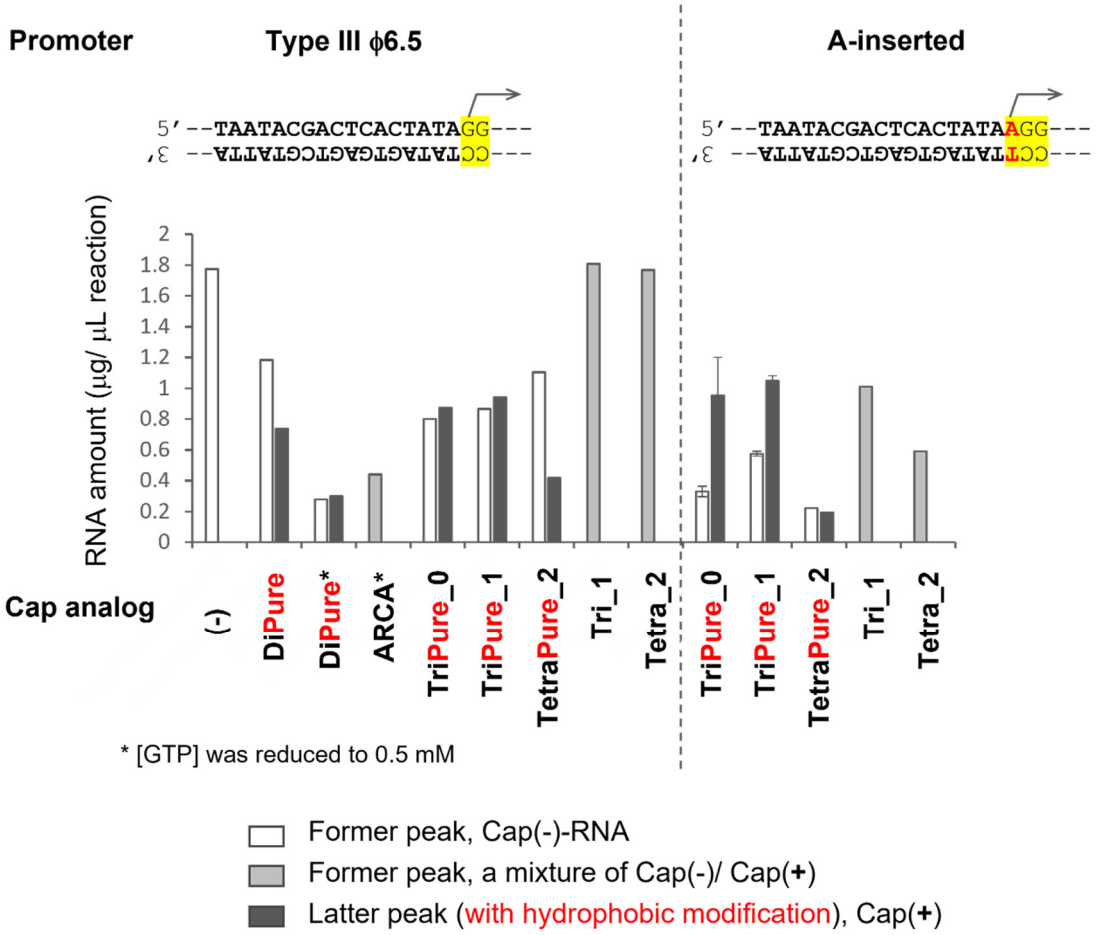
Sequence of RNA fragment	$[M+H]^+$ calcd	$[M+H]^+$ observed					
		Cap(-)	**ARCA**	**ARCA/AnP**	**DiPure**	**DiPure/3'OMe**	**DiPure/2'OMe**
G-RNA	7505.5			7506.0 (+0.5)			
pp-G-RNA	7665.4	7666.9 (+1.5)	7663.9 (-1.5)				
ppp-G-RNA	7745.4	7746.9 (+1.5)	7743.9 (-1.5)				
m^7G-ppp-G-RNA	8024.7				8025.2 (+0.5)		
$m_2^{7,3'\text{-O}}$G-ppp-G-RNA	8038.7		8038.7 (0)	8039.1 (+0.4)		8038.6 (-0.1)	
$m_2^{7,2'\text{-O}}$G-ppp-G-RNA	8038.7						8038.4 (-0.3)



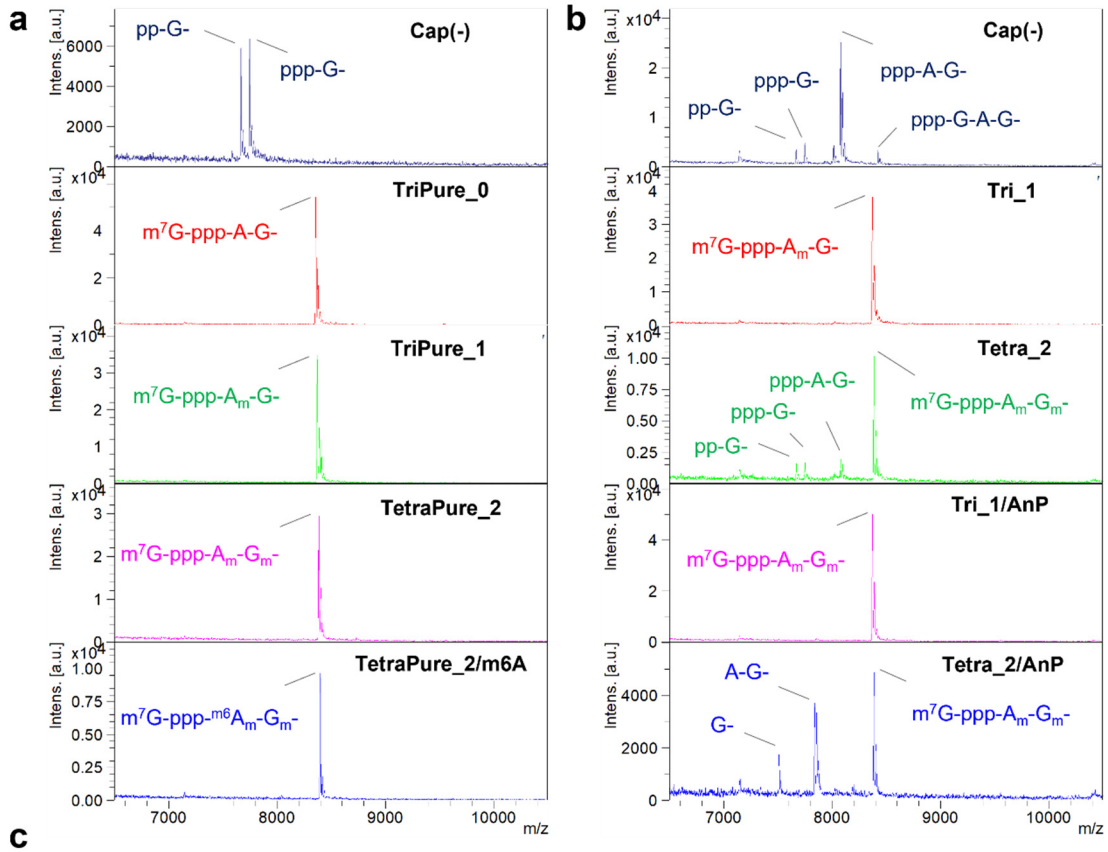
Extended Data Figure 1. Removal of double-stranded RNA (dsRNA) impurity from transcribed NanoLuc luciferase mRNA (650-nt) using reversed-phase HPLC. (**a, b**) Analytical (**a**) or preparative reversed-phase HPLC profiles of the crude RNA, using YMC Triart Bio C4 column (4.6 × 250 mm) using Solution_A [50 mM TEAA (pH 7), 5% acetonitrile] and Solution_B (acetonitrile). Solution_B content was increased from 0% to 20% over 20 minutes at a flow rate of 1 mL. The column temperature was maintained at 50 °C. (**b**) Eluted RNA was fractionated 1 min at a time, from 14.5 min to 23.5 min. (**c**) dsRNA detection by a dot blot assay using anti-dsRNA clone rJ2 antibody. An actual photograph of the blot and a graph of the signal quantification results were shown. The RNA fractions shown in (**b**) were concentrated, and 50 ng RNA was dotted. (**d**) 5% dPAGE analysis of fractionated RNAs (50 ng each). The gel was visualized by SYBR Green II staining.



Extended Data Figure 2. IVT using DiPure cap analogs. NanoLuc luciferase mRNA (650-nt) was transcribed by T7 RNAP in the presence of DiPure cap analogs. (**a**) dPAGE analysis of the transcription reaction. (**b-d**) Reverse-phase HPLC analysis of the RNA transcript. The transcript was analyzed as a crude mixture (**i**, black line) or after being purified by preparative HPLC (**ii**, blue line). Purified RNA was analyzed after deprotection by irradiating 365-nm light (**iii**, green line). Elution time (min) of the peak was noted nearby. A ratio (%) of capped mRNA calculated based on the peak area was listed in parentheses after the elution time.



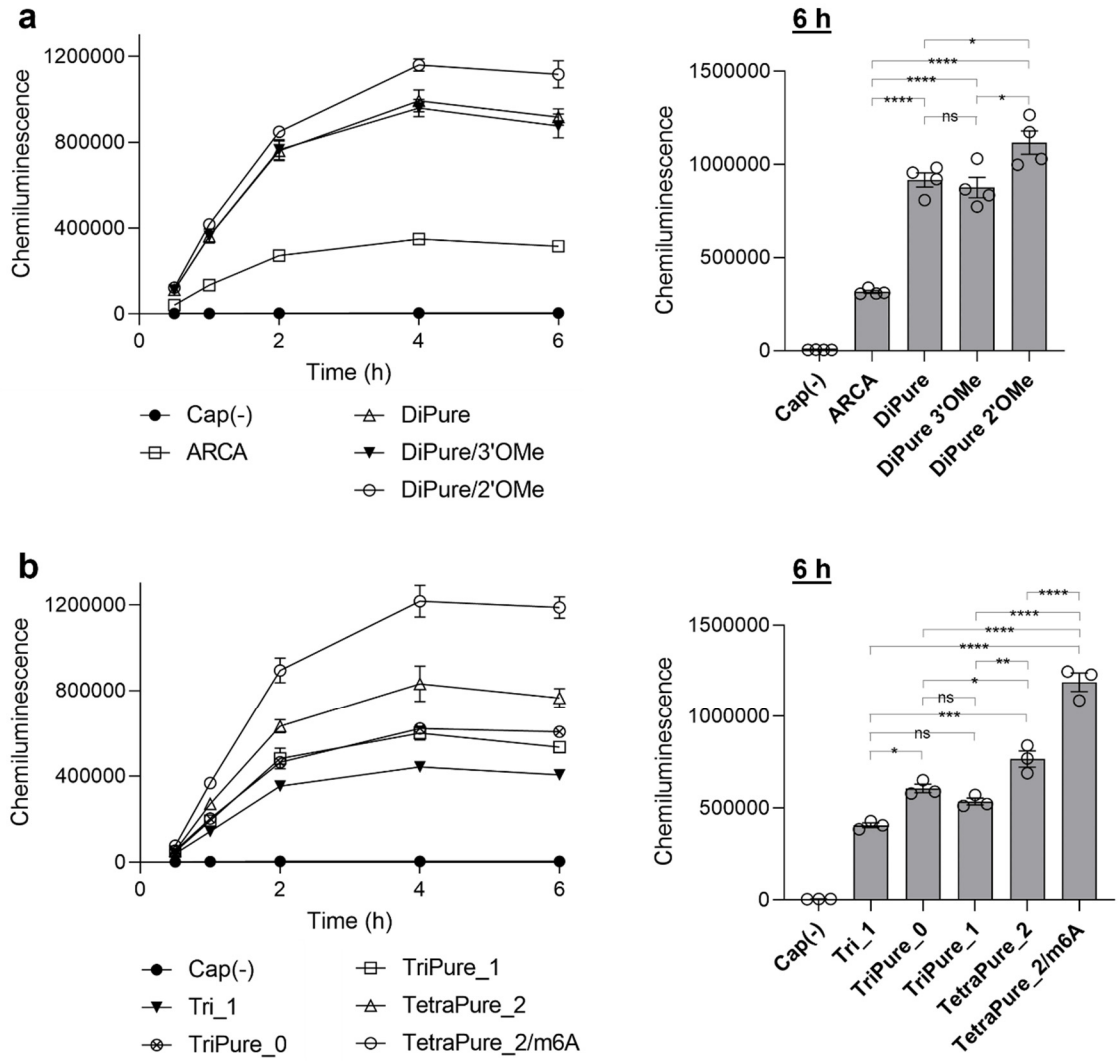
Extended Data Figure 3. RP-HPLC quantification of the capped/uncapped transcripts in the IVT reaction using the PureCap analogs. 650-nt Nluc mRNA was transcribed from a DNA template containing canonical Type III ϕ6.5 promoter or a modified promoter ("A-inserted"). The cap analog added to the reaction was indicated in the graph.



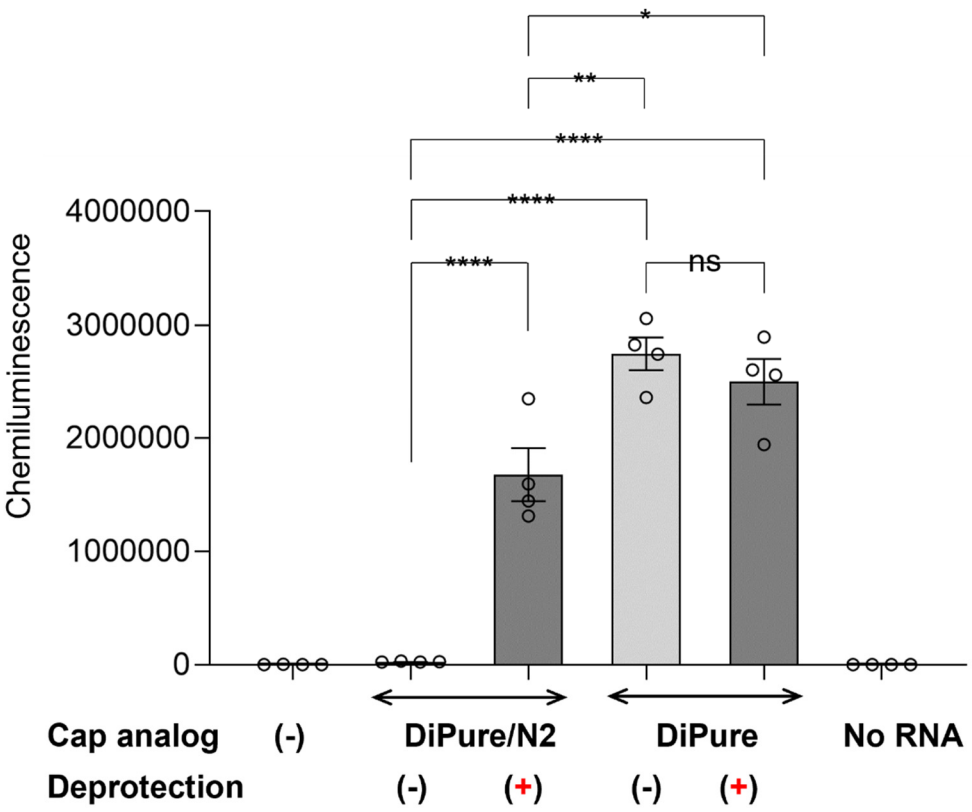
c

Sequence of RNA fragment	[M+H]+ calcd	Promoter sequence									
		Type III phi6.5 (panel_a)					A-inserted (panel_b)				
		Cap(-)	TriPure_0	TriPure_1	TetraPure_2	TetraPure_2/m6A	Cap(-)	Tri_1	Tetra_2	Tri_1/AnP	Tetra_2/AnP
G-RNA	7505.5										7510.2 (+4.7)
pp-G-RNA	7665.4	7666.5 (+1.1)					7667.9 (+2.5)		7669.5 (+4.1)		
ppp-G-RNA	7745.4	7746.2 (+0.8)					7748.1 (+2.7)		7749.9 (+4.5)		
A-G-RNA	7834.7										7839.3 (+4.6)
pp-A-G-RNA	7994.7										
ppp-A-G-RNA	8074.7						8077.3 (+2.7)		8078.4 (+3.8)		
ppp-G-A-G-RNA	8419.9						8422.4 (+2.5)				
m7G-ppp-A-G-RNA	8353.9		8356.6 (+2.7)								
m7G-ppp-Am-G-RNA	8367.9			8370.2 (+2.3)				8371.4 (+3.5)		8371.6 (+3.7)	
m7G-ppp-Am-Gm-RNA	8381.9				8384.8 (+2.9)				8385.8 (+3.9)		8386.6 (+4.7)
m7G-ppp-m6Am-Gm-RNA	8395.9					8398.0 (+2.1)					

Extended Data Figure 4. MALDI-TOF MS analysis of 5' terminus of capped mRNAs prepared using tri/tetranucleotide cap analogs. The names of the cap analogs are shown in the figure; AnP means that the RNA was dephosphorylated using Antarctic phosphatase. Capped mRNAs prepared from a DNA template containing Type III ϕ6.5 promoter (**a**) or a modified T7 promoter named "A-inserted" (**b**) were analyzed. The 5' fragment was generated by cleavage with DNAzyme. (**c**) Calculated and observed masses shown in panels (**a**) and (**b**) are listed. The sequence shown as **RNA** is 5'-GCGCAUAUUAAGGUGACGCGUG-p (2',3' cyclic). All prominent peaks were accompanied by a peak with a molecular weight of +18, probably due to hydrolysis of the 2' and 3' cyclic monophosphate initially present at the 3' end of the RNA fragment.



Extended Data Figure 5. Nluc mRNAs with different cap structures were compared in a HeLa-based cell-free translation system. Capped mRNAs prepared using dinucleotide (**a**) or tri/tetranucleotide cap analogs (**b**) were compared. Time-course of the expression profiles (left) and the extracted data at 6-h (right) were shown. Data are expressed as the mean ± standard error (n = 3 or 4). One-way ANOVA test between capped mRNAs followed by Tukey's test were marked as follows: ns, $p > 0.05$ (not significant); *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$, ****, $p < 0.0001$.



Extended Data Figure 6. Effect of the hydrophobic tag on the translation activity of 650-nt Nluc mRNAs in cultured HeLa cells. mRNAs (20 ng per well in a 96-well plate) were transfected into HeLa cells (seeded at 1×10^4 cells one day before transfection) using Lipofectamine MessengerMAX reagent (0.3 μL per well). After six hours of incubation, the cells were lysed, and the expression of Nluc protein was evaluated. mRNAs were transcribed in the presence of **DiPure/N2** or **DiPure** cap analog, and then the capped RNA was isolated by RP-HPLC. The activity was compared before and after removing a hydrophobic tag from the cap moiety by 365-nm light irradiation. Data are expressed as the mean ± standard error ($n = 4$). One-way ANOVA test between capped mRNAs followed by Tukey's test were marked as follows: ns, $p > 0.05$ (not significant); *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$, ****, $p < 0.0001$.