Efficient Synthesis of $^{13}C,^{15}N$-Labeled RNA Containing the Cap Structure m^7GpppA

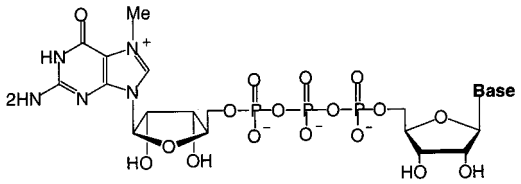
Hiroshi Matsuo,[†,‡] Tomohisa Moriguchi,[§] Toshimitsu Takagi,[†] Takahiro Kusakabe,[†,⊥] Stephen Buratowski,[†] Mitsuo Sekine,[§] Yoshimasa Kyogoku,[‡,∥] and Gerhard Wagner*[,†]

Contribution from the Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, 240 Longwood Avenue, Boston, Massachusetts 02115, Institute for Protein Research, Osaka University, Suita, Osaka 565-0871, Japan, and Department of Life Science, Tokyo Institute of Technology, Nagatsuta, Midoriku, Yokohama 226-8501, Japan

Received July 29, 1999

Abstract: Studying the mechanism by which the cap structure performs its numerous biological roles in mRNA metabolism by NMR spectroscopy requires the large-scale production of isotopically labeled capped RNA. We present an efficient method for production of short RNA containing the cap structure m^7GpppA, by combining chemical synthesis with enzymatic synthesis using a DNA primase. This method was employed to synthesize three capped RNA molecules, $m^7Gppp^{[13C,15N]}A$, $m^7Gppp^{[13C,15N]}ACC$, and $m^7GpppA^{[13C,15N]}C^{[13C,15N]}C$, that contain selective $^{13}C,^{15}N$-labeled adenosine or cytosine nucleotides. This selective labeling technique enabled us to obtain the chemical shift assignments of these oligonucleotides in complex with the eukaryotic translation initiation factor 4E (eIF4E). Furthermore, we were able to observed intermolecular NOE interactions between the RNA and protein.

Eukaryotic messenger RNA contains the cap structure, m^7-GpppN- (where N is any nucleotide), in which 7-methylguanosine is linked to the 5′ end of RNA via a 5′−5′ triphosphate bridge:[1,2]



Cap plays critical roles in many aspects of messenger RNA metabolism, including processing, nuclear transport, translation, and the protection of mRNA from untimely degradation.[3,4] To elucidate the molecular details of the interactions formed between capped RNA and proteins, structural studies of such complexes have recently been actively pursued. The vaccinia virus protein VP39[5−7] and eukaryotic translation initiation factor 4E (eIF4E),[8,9] however, are the only proteins whose structures have been solved in complexes with cap analogues. While there is no similarity in their primary or secondary structures, the interactions responsible for the cap binding of VP39 and eIF4E are quite similar as in both cases the 7-methyl guanosine ring is sandwiched between two aromatic rings. Furthermore, the complex of VP39 with a capped mRNA hexamer ($m^7GpppG(A)_5$) revealed sequence-independent interactions between the RNA hexamer and the cap binding protein.[7] The eIF4E/cap analogue complex is currently the only cap−protein complex whose structure has been determined by NMR.[9] Detailed studies of cap interactions using NMR spectroscopy have been hampered by the lack of suitable methods for producing large amounts of isotopically labeled capped mRNA.

Here we describe an efficient method for the large-scale production of short RNA containing the cap structure m^7GpppA by combining chemical synthesis with enzymatic synthesis using a DNA primase. Previously, methods have been described to produce capped RNA *in vitro* by using bacteriophage RNA polymerases (SP6, T3, or T7).[10−13] These techniques, however, are inefficient and require large amounts of cap analogue. Furthermore, the strict preference of bacteriophage RNA polymerase for guanosine at the RNA 5′-terminal position has caused previous structural studies of capped RNA molecules to be restricted to $m^7GpppG(N)_x$.[5−7]

Gene 4 primase, encoded by T7 phage, uses DNA templates to produce short RNA primers beginning with the sequence pppAC.[14] In the method described here, gene 4 primase was

* To whom corresponding should be addressed. Telephone: (617) 432-3213. Fax: (617) 432-4383. E-mail: wagner@wagner.med.harvard.edu.
† Harvard Medical School.
‡ Osaka University.
§ Tokyo Institute of Technology.
∥ Present address: Fukui Institute of Technology, 3-6-1, Fukui 910-8505, Japan
⊥ Present address: Kyusyu University, Fukuoka 812-8581, Japan.

(1) Shatkin, A. J. *Cell* **1976**, *9*, 645−653.
(2) Banerjee, A. K. *Microbiol. Rev.* **1980**, *44*, 175−205.
(3) Levis, J. D.; Izaurralde, E. *Eur. J. Biochem.* **1997**, *247*, 461−469.
(4) Sonenberg, N. *Translational Control*; Cold Spring Harbor Laboratory Press: New York, 1996.
(5) Hodel, A. E.; Gershon, P. D.; Shi, X.; Quiocho, F. A. *Cell* **1996**, *85*, 247−256.
(6) Hodel, A. E.; Gershon, P. D.; Shi, X.; Wang, S. M.; Quiocho, F. A. *Nat. Struct. Biol.* **1997**, *4*, 350−354.
(7) Hodel, A. E.; Gershon, P. D.; Quiocho, F. A. *Mol. Cell* **1998**, *1*, 443−447.
(8) Marcotrigiano, J.; Gingras, A.-C.; Sonenberg, N.; Burley, S. K. *Cell* **1997**, *89*, 951−961.
(9) Matsuo, H.; Li, H.; McGuire, A. M.; Fletcher, C. M.; Gringras, A.-C.; Sonenberg, N.; Wagner, G. *Nat. Struct. Biol.* **1997**, *4*, 717−724.
(10) Green, M. R.; Maniatis, T.; Melton, D. A. *Cell* **1983**, *32*, 681−694.
(11) Konarska, M. M.; Padgett, R. A.; Sharp, P. A. *Cell* **1984**, *38*, 731−736.
(12) Pelletier, J.; Sonenberg, N. *Cell* **1985**, *40*, 515−526.
(13) Nielsen, D. A.; Shapiro, D. J. *Nucleic Acids Res.* **1986**, *14*, 5936.

10.1021/ja9926820 CCC: $19.00
Published on Web 03/07/2000

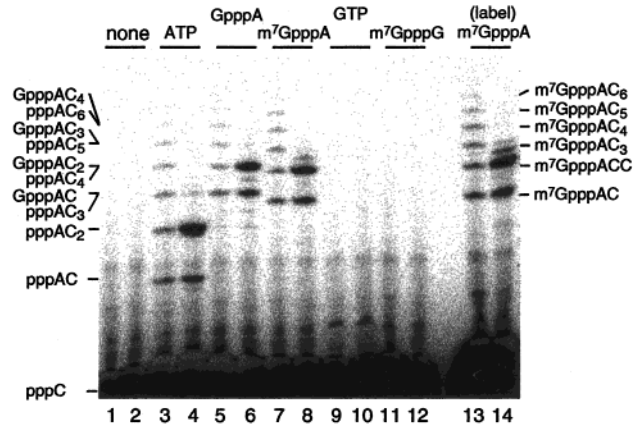


Figure 1. GpppA and m^{7}GpppA can be incorporated as the first nucleotide of primers synthesized by T7 primase. The oligonucleotide synthesis reactions with gene 4 primase were performed as described in the text, with 1 mM [α-^{32}P]-CTP, 1 mM of the indicated nucleotide, and with either 1 nM 5′-(G)$_{19}$GTCATTCTTGGACCT-3′ (odd-numbered lanes) or 1 nM 5′-CCCCCGGTC(T)$_{25}$-3′ (even-numbered lanes). The reaction products were resolved by electrophoresis on a 36% polyacrylamide gel containing 3 M urea and were subsequently subjected to autoradiography.

used to produce selectively labeled capped oligonucleotides with the sequence m^{7}GpppACC. Three selectively labeled capped RNA molecules were synthesized by using this method, including m^{7}Gppp$^{[13C,15N]}$A, m^{7}Gppp$^{[13C,15N]}$ACC, and m^7-GpppA$^{[13C,15N]}$C$^{[13C,15N]}$C. This technique enabled us to obtain the chemical shift assignments and intermolecular NOE interactions of these oligonucleotides in complexes with the eukaryotic translation initiation factor 4E (eIF4E).

Experimental Section

Cap Analogue Incorporation Using Gene 4 T7 Primase. The ability of gene 4 T7 primase to incorporate different cap analogues and nucleotides into the first position of primers was tested as described below. All reactions (10 μL) contained 40 mM Tris-HCl (pH 7.5), 10 mM $MgCl_2$, 10 mM dithiothreitol, 50 μg/mL bovine serum albumin, 50 mM potassium glutamate, 1 μM zinc chloride, 2 mM dTTP, 1 mM [α-^{32}P]-CTP (500−1000 cpm/pmol) (NEN/DuPond), 1 nM DNA template [either 5′-(G)$_{19}$GTCATTCTTGGACCT-3′ to produce RNA of variable lengths[14] or the template (5′-CCCCCGGTC(T)$_{25}$-3′) for trinucleotide RNA], and 1 μM (hexamer) gene 4 primase. The T7 gene 4 protein was purified as described previously,[15] and the oligonucleotide templates were chemically synthesized (Integrated DNA Technologies, Inc.). These reactions were performed in the absence or presence of either 1 mM ATP, GTP, or cap analogue dinucleotides (GpppA, m^7-GpppA, and m^{7}GpppG (New England Biolabs Inc.)). After incubation at 37 °C for 60 min, synthesis was stopped by the addition of 5 μL of 98% formamide, 10 mM EDTA (pH 8.0), 0.1% xylene cyanol FF, and 0.1% bromophenol blue. After being denatured by heating (at 95 °C for 5 min), the reaction mixtures were separated on a 36% polyacrylamide gel containing 3 M urea (Figure 1).

Chemical Synthesis of m^{7}Gppp$^{[13C,15N]}$A. Synthesis of m^{7}GpppA was performed by using a bifunctional phosphorylating reagent, *O*-8-(5-chloroquinolyl) *S*-phenyl phosphorothioate as shown in Scheme 1.[16]

General Procedures. Pyridine was distilled after refluxing over *p*-toluenesulfonyl chloride for several hours, redistilled from CaH_2, and then stored over molecular sieves 4A. ^{1}H NMR spectra were obtained at 270 MHz on a JEOL-EX-270 spectrometer in D_2O and referenced with sodium 3-(trimethylsilyl)propionesulfonate (DSS) as an external standard. ^{31}P NMR spectra were obtained at 109.25 MHz on a JEOL-EX-270 spectrometer using 85% H_3PO_4 as an external reference. Thin-layer chromatography was performed on precoated glass plates of Kieselgel 60 F_{254} (Merck, No. 5715). Anion-exchange chromatography was performed with a DEAE-Sephadex A-25 (HCO_3^- form) column by using a linear gradient of 5−50% 1 M NH_4HCO_3 at 4 °C.

***P*^{1}-*O*-8-(5-Chloroquinolyl) *P*^{2}-7-Methylguanosine-5′-*O*-yl Pyrophosphate.** Pyridinium guanosine-5′-monophosphate (221 mg, 0.5 mmol) was dissolved in *N*-methyl pyrrolidinone (NMP)−hexamethylphosphoric triamine (HMPA) (3:1,v/v, 12 mL). Methyl iodide (125 μL, 2.0 mmol) was added to this solution and stirred at room temperature overnight. The reaction mixture that contained compound **1** was concentrated to a small volume and then coevaporated three times with dry pyridine. The residue was dissolved in dry pyridine (0.5 mL), after which silver nitrate (170 mg, 1.0 mmol) and cyclohexylammonium *O*-8-(5-chloroquinolyl) *S*-phenyl phosphorothiate (225 mg, 0.5 mmol) were added at 0 °C. This mixture was stirred at 0 °C for 4 h and subsequently stirred at room temperature for 1 h. 4-Chlorobenzenethiol (1 g) in $CHCl_3$ (5 mL) and a small amount of H_2O were added to the reaction mixture at room temperature. The H_2O layer was washed four times with $CHCl_3$ and filtered by Celite. The filtrate was then concentrated to a small volume under reduced pressure and purified by anion-exchange chromatography to give the desired product (compound **2**, 104 mg, 35%). ^{1}H NMR (CD_3OD) δ 9.55 (1H, s, 8-H of G), 8.88−7.59 (5H, m, quinoline), 5.93 (1H, br, 1′-H), 4.58 (2H, m, 2′-H and 3′-H), 4.41 (1H, m, 4′-H), 4.25 (2H, m, 5′-H and 5″-H), 3.98 (3H, s, 7-CH_3 of G); ^{13}C NMR (CD_3OD) δ 158.10, 155.71, 152.05, 151.51, 150.16, 143.34, 139.56, 130.08, 129.03, 128.72, 126.62, 124.43, 120.58, 109.57, 92.64 (1′-C), 86.30 (d, J = 8.5 Hz, 4′-C), 77.53 (3′-C), 70.99 (2′-C), 65.80 (d, J = 6.1 Hz, 5′-C), 37.47 (7-CH_3); ^{31}P NMR (CD_3OD) δ −10.24 (d, J = 21.4 Hz, P^2), −15.66 (d, J = 22.1 Hz, P^1).

***P*^{1}-7-Methylguanosine-5′-*O*-yl *P*^{3}-Adenosine-5′-*O*-yl Triphosphate *P*^{1}-*O*-8-(5-Chloroquinolyl).** *P*^{2}-7-methylguanosine-5′-*O*-yl pyrophosphate (104 mg, 0.17 mmol) and 5′-AMP (74 mg, 0.17 mmol) were dissolved in NMP−HMPA (3:1, v/v, 4 mL) and dried by coevaporating three times with dry pyridine. $CuCl_2$, dissolved in NMP (2 mL) (68 mg, 0.50 mmol in NMP solution), was added to this solution. After stirring at room temperature for 36 h, the reaction was stopped upon addition of water. The precipitate was filtered by Celite, and the filtrate was washed three times with $CHCl_3$. The aqueous layer was then filtered by Celite, and the filtrate concentrated to a small volume under reduced pressure. The residue was purified by anion-exchange chromatography to give the desired product (compound **3**, 33 mg, 38 μmol). ^{1}H NMR (D_2O) δ 8.38 (1H, 8-H of A), 8.13 (1H, 2-H of A), 5.98 (1H, 1′-H of A), 4.61 (1H, 2′-H of A), 4.46 (1H, 3′-H of A), 4.25 (2H, 5′-H and 5″-H of A); ^{13}C NMR (D_2O) δ 155.55 (2-C of A), 142.16 (8-C of A), 89.09 (1′-C of A), 86.61 (4′-C of A), 77.22 (2′-C of A), 73.16 (3′-C of A), 68.14 (5′-C of A); ^{31}P NMR (D_2O) δ −10.88 (d, J = 18.2 Hz, P^1, P^3), −22.56 (t, J = 17.1 Hz, P^2).

The eIF4E/m^{7}Gppp$^{[13C,15N]}$A complex was produced by eluting yeast eIF4E from a m^{7}GDP-agarose affinity column with m^{7}Gppp$^{[13C,15N]}$A.[9,17] Fractions containing the complex were concentrated to 1 mM to be used for NMR experiments.

Enzymatic Synthesis of m^{7}Gppp$^{[13C,15N]}$ACC and m^{7}GpppA$^{[13C,15N]}$-C$^{[13C,15N]}$C. The m^{7}Gppp$^{[13C,15N]}$A, prepared by using the procedure outlined above, is ready for enzymatic extension by gene 4 DNA primase of bacteriophage T7 as shown in Scheme 2. The gene 4 T7 primase was used to extend m^{7}GpppA to produce the isotopically labeled RNA trinucleotides m^{7}Gppp$^{[13C,15N]}$ACC and m^{7}GpppA$^{[13C,15N]}$C$^{[13C,15N]}$C, in sufficient quantity for NMR spectroscopy. The oligonucleotide template described above (Figure 1) was used for these large-scale reactions. Reactions (20 × 100 μL) were performed in 40 mM Tris-Cl (pH 7.5), 10 mM $MgCl_2$, 10 mM dithiothreitol, 50 μg/mL bovine serum albumin, 50 mM potassium glutamate, 1 μM $ZnCl_2$, 10 mM dTTP, 10 mM CTP, 1 mM m^{7}GpppA, 1 μM template, and 500 nM gene 4 primase. Uniformly ^{13}C,^{15}N-labeled m^{7}Gppp$^{[13C,15N]}$A or CTP (ISOTEC INC) were used to obtain selectively labeled m^{7}GpppACC. After overnight incubation at 37 °C, the reaction mixture was applied to an ion-exchange

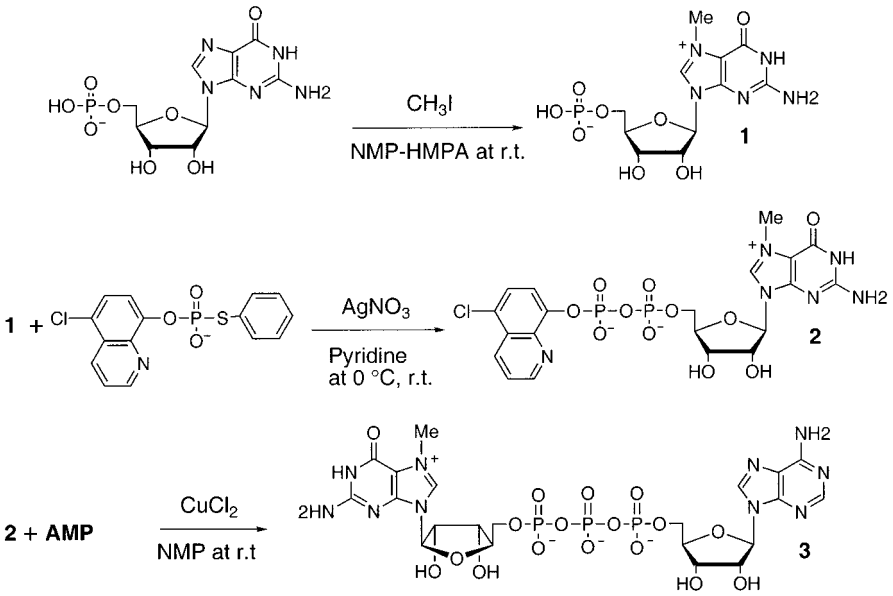
(14) Kusakabe, T.; Richardson, C. C. *J. Biol. Chem.* **1997**, *272*, 5943−5951.

(15) Mendelman, L. V.; Richardson, C. C. *J. Biol. Chem.* **1991**, *266*, 23240−23250.

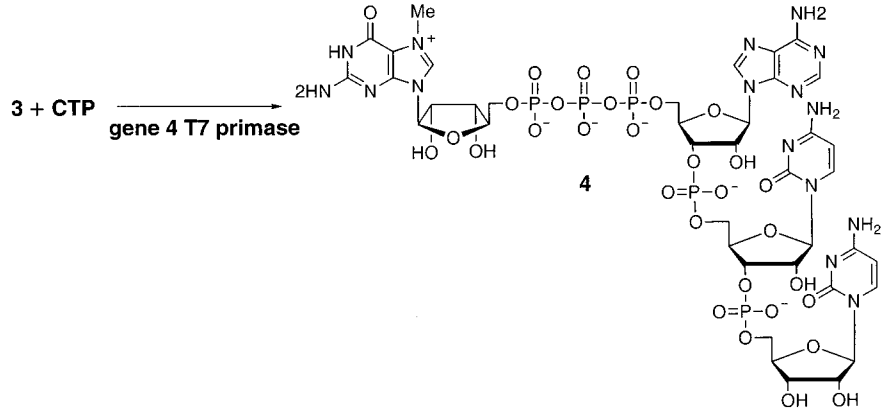
(16) Fukuoka, K.; Suda, F.; Suzuki, R.; Ishikawa, M.; Takaku, H.; Hata, T. *Nucleosides Nucleotides* **1994**, *13*, 1557−1567.

(17) Edery, I.; Altmann, M.; Sonenberg, N. *Gene* **1988**, *74*, 517−525.

Scheme 1. Chemical Synthesis of m^{7}Gppp$^{[13C,15N]}$A



Scheme 2. Procedure for Extending m^{7}Gppp$^{[13C,15N]}$A by Using Gene4 Primase to Produce the m^{7}Gppp$^{[13C,15N]}$ACC



column (Dionex Nucleopac Column PA-100) and eluted by applying a linear salt gradient ranging from 0 to 0.35 M NaCl for 30 min. Two micromoles of m^{7}GpppA was used to produce 1.5 μmol of m^{7}GpppACC. Fractions containing m^{7}GpppACC were collected and concentrated to dryness or used to produce the eIF4E/ m^{7}GpppACC complex by eluting yeast eIF4E protein from a m^{7}GDP-agarose affinity column.[9,17] Fractions containing the complex were concentrated to either 0.2 or 1 mM to be used for NMR experiments.

Results

GpppA and m^{7}GpppA supported oligonucleotide synthesis by gene4 primase almost as effectively as ATP (Figure 1, lanes 3−8). When the template for trinucleotide RNA (5′-CCCCCG-GTC(T)$_{25}$-3′) was used (Figure 1, lanes 4, 6, and 8), pppACC, GpppACC, or m^{7}GpppACC was the predominant product. Gene 4 protein has a substrate specificity for adenosine at the first position of the primer,[14] causing neither GTP nor m^{7}GpppG to be incorporated (Figure 1, lanes 9−12) and excluding the possibility that some of the products in lanes 5−8 have a "reverse" cap structure (ApppG and Apppm^{7}G).

As revealed in the electrophoresis gel shown in Figure 1 (lane 14), the extension of our synthetically produced m^{7}Gppp$^{[13C,15N]}$A by gene 4 primase was comparable to that of unlabeled m^{7}GpppA (New England Biolabs, Inc.) (lane8). Furthermore, the template for trinucleotide RNA (5′-CCCCCGGTC(T)$_{25}$-3′) predominantly produced m^{7}GpppACC. A large-scale reaction (20 × 100 μL reactions) produced enough m^{7}GpppACC for the eIF4E/m^{7}GpppACC sample used in the NMR experiments. The assigned [^{13}C,^{1}H]-HMQC (heteronuclear multiquantum coherence) spectrum of m^{7}Gppp$^{[13C,15N]}$A is provided in Figure 2. Assignments were obtained by using a DQF-COSY (double quantum filter correlated spectroscopy) spectrum (data not shown).[18] The presence of the triphosphate bridge connecting the 5′ end of m^{7}G to the 5′ end of $^{[13C,15N]}$A was confirmed from phosphorus 1D NMR spectroscopy.

Figure 3 illustrates the application of this method to characterize interactions between capped RNA and the translation initiation factor eIF4E carried by yeast *Saccharomyces cerevisiae*, a 213-residue cap binding protein that controls eukaryotic translation.[4] The eIF4E structure has been solved in its complexed form with the cap analogue m^{7}Gpp by both crystallography [8] and NMR spectroscopy.[9] To investigate the interactions between longer fragments of capped RNA and eIF4E, complexes were produced containing m^{7}Gppp$^{[13C,15N]}$A or m^{7}GpppA$^{[13C,15N]}$C$^{[13C,15N]}$C with ^{13}C,^{15}N-labeled adenosine or cytosine, respectively. Figure 3a shows 2D slices at selected ^{13}C frequencies from a ^{13}C-dispersed NOESY spectrum of m^{7}Gppp$^{[13C,15N]}$A in complex with eIF4E. This spectrum reveals not only intra- and internucleotide NOEs but also contains intermolecular NOEs between the adenosine and eIF4E. There

(18) Piantini, U.; Sørensen, O.; Ernst, R. R. *J. Am. Chem. Soc.* **1982**, *104*, 6800−6801.

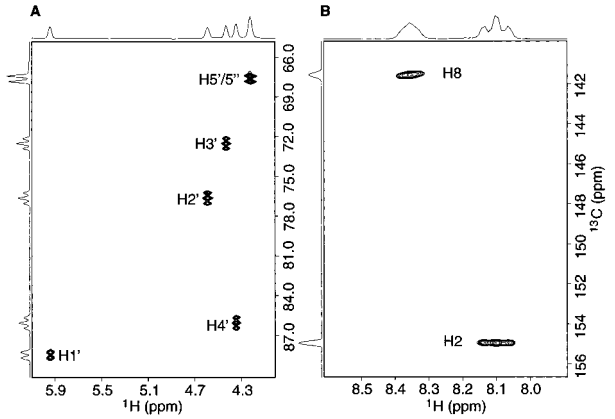


Figure 2. Selected regions of a [^{1}H,^{13}C]-HMQC spectrum of m^{7}GpppA displaying the adenosine sugar (A) and base (B) protons. This spectrum was recorded on a Varian Unity Plus 400 at 283 K in 100% D_2O and processed with FELIX software packages (Molecular Simulations). The spectral widths in the ^{13}C and ^{1}H dimensions are 13076 and 4000 Hz, respectively. Carbon−carbon and proton−proton couplings are observed for the signals displayed in A and B, respectively.

are numerous intermolecular NOE interactions between eIF4E and the observable adenosine base protons, 2H and 8H, and a sugar proton, H1′. Two-dimensional slices at selected ^{13}C frequencies from a ^{13}C-dispersed NOESY spectrum of m^{7}GpppA$^{[13C,15N]}$C$^{[13C,15N]}$C in complex with eIF4E reveal intra- and internucleotide NOEs but lack intermolecular NOEs (Figure 3b).

Furthermore, a ^{13}C-dispersed NOESY spectrum of m^{7}GpppACC in complex with eIF4E displayed line shape broadening and chemical shift changes of the adenosine signals upon binding. The sample used for this initial experiment was too dilute, however, to detect intermolecular NOEs, and therefore this study requires further investigation.

Discussion

The chemical synthesis of m^{7}GpppA enabled us to observe intermolecular NOEs between adenosine and eIF4E (Figure 3a). The m^{7}GpppA/eIF4E complex is only stable in a detergent micelle, causing its final molecular weight to be approximately 40 kDa.[9] NOE interactions of such a large system are often weak and difficult to detect. However, by using the cap analogues produced by the method described here, we were able to detect numerous intermolecular interactions between the protein and capped RNA.

We were able to observe NOE interactions between the base protons of the first transcribed nucleotide attached to cap and eIF4E. Since the first transcribed nucleotide is restricted to either guanosine or adenosine in yeast system, sequence specific interactions between capped mRNA and cap binding proteins can be inferred. m^{7}GpppA was previously found to interact with eIF4E residues 199−206, situated in helix 6 and the following loop, from chemical shift perturbation studies of eIF4E.[9] In contrast to the interactions we observed in the m^{7}GpppA/eIF4E complex, the interactions observed in the crystal complex between VP39 and the capped RNA hexamer (m^{7}GpppG(A)$_5$) were nonsequence specific, between the sugar and phosphorus backbone atoms of the RNA and the protein.[7] It would be interesting to establish if nonsequence specific interactions are formed between eIF4E and the subsequent nucleotides as there is no sequence preference for these positions.

To investigate such interactions, m^{7}GpppAC$^{[13C,15N]}$C was synthesized with gene4 primase and used to produce the m^{7}GpppAC$^{[13C,15N]}$C/eIF4E complex for NMR experiments. The ^{13}C-edited NOESY spectrum lacked intermolecular NOE interactions between the cytosine nucleotides of the cap analogue and eIF4E (Figure 3b). Since eIF4E functions in the eIF4F complex with eIF4A and eIF4G, it is conceivable that these other proteins are required to form such interactions with the transcribed RNA.

Since the commonly used bacteriophage polymerases (T7, SP6, and T3) have a strict preference for guanosine nucleotides at the first position of their transcripts, the 5′ end of the capped RNAs produced by using these polymerases are restricted to m^{7}GpppG. We have described a method that can produce sufficient quantities of capped RNA containing m^{7}GpppA at the 5′ end for structural studies by X-ray crystallography or NMR spectroscopy. Furthermore, in contrast to primase, all of the transcription systems of bacteriophage RNA polymerase

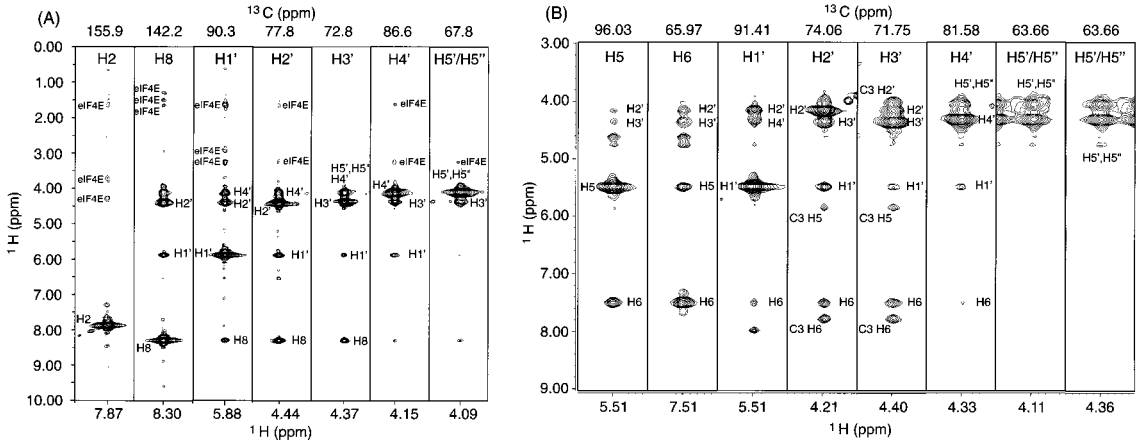


Figure 3. 2D strips at the ^{13}C resonance frequencies of (A) adenosine and (B) the second cytosine from 3D ^{13}C-dispersed NOESY-HSQC spectra of unlabeled eIF4E with m^{7}GpppA and m^{7}GpppAC$^{[13C,15N]}$C, respectively. Both spectra were recorded with a NOESY mixing time of 100 ms and in 100% D_2O. The frequency (in ppm) of the carbon indirect and proton direct dimensions are indicated above and below each strip, respectively. Processing and analysis were performed on a Silicon Graphics work stations with FELIX software packages (Molecular Simulations) and XEASY,[23] respectively. (A) The ^{13}C-dispersed NOESY-HSQC spectrum of m^{7}GpppA /eIF4E was recorded on a Bruker AVANCE 800 at 293 K. The spectral widths in the ^{13}C indirect, ^{1}H indirect, and ^{1}H direct dimensions are 8048, 11201, and 11160 Hz, respectively. Assignments are provided to the intra- and internucleotide cross-peaks. Intermolecular cross-peaks to the eIF4E are indicated. (B) The ^{13}C-dispersed NOESY-HSQC spectrum of m^{7}GpppAC$^{[13C,15N]}$C/eIF4E was recorded on a Varian Unity 500 at 298 K. The spectral widths in the ^{13}C indirect, ^{1}H indirect, and ^{1}H direct dimensions are 4651, 6300, and 6300 Hz, respectively. Intra- and internucleotide NOE cross-peaks are labeled.

require GTP for chain elongation, and consequently, not all of the transcripts produced by using polymerase are capped. The production of m^{7}GpppA(C)$_x$ by means of primase does not require additional ATP, causing all oligonucleotide to be capped. A further disadvantage of using polymerases is that dinucleotide cap analogues frequently get misincorporated in undesirable orientations such that the methylguanosine nucleotide is adjacent to the RNA sequence, forming a "reverse" cap structure (Gpppm^{7}GN$_x$ and mGpppm^{7}GN$_x$).[19] More recently, it has been published that T7 RNA polymerase makes mistakes at the 5′ end of in vitro-transcribed RNA.[20] These mistakes may affect the efficiency of the incorporation of cap analogues into the first position. Due to the very strict substrate specificity of gene 4 protein for adenosine nucleotide at the first position,[14] (Figure 1, lanes 9−12) the reverse cap structures ApppG and Apppm^{7}G are not produced using primase (Figure 1).

Although gene 4 primase cannot be used to produce long oligonucleotides, it is ideally suited for producing the short capped oligonucleotides to investigate interactions between capped RNA and cap binding proteins. Furthermore, general modifications of mammalian mRNA occur at either the cap itself or the first or second nucleotide directly following the cap, as in the 2′-methylation of the ribose in cap1 (m^{7}GpppN$_1$mN$_2$...) and cap2 (m^{7}GpppN$_1$mN$_2$m...)[2]. Consequently, capped oligonucleotides produced by the method described here can also serve to study the recognition of variant cap structures by proteins and to investigate internal ribose methylation.

Alternatively, capped RNA can be produced by using purely chemical synthesis, as has been reported for m$_3^{2,2,7}$GpppAmpUmpA[21] and m^{7}GpppGUAUUAAUA.[22] Although these methods are efficient, they are more difficult to perform, and the large-scale production of the isotopically labeled molecules required for NMR spectroscopy by such methods is costly.

In summary, we have developed an efficient method to prepare large amounts of short capped RNA containing m^{7}GpppA at the 5′ end by combining the chemical synthesis of the cap part with enzymatic synthesis using gene 4 DNA primase. This method is well suited for the production of isotopically labeled capped RNA and greatly facilitates structural studies by NMR spectroscopy of important mechanisms involving RNA metabolism.

Acknowledgment. This work was supported by a Grant from the "Research for the future" program of the Japan Society for Promotion of Science (JSPS-RFTF97100301) (M.S.), Grants-in-Aid for Scientific Research on Priority Areas (10179102) (H.M.), and the NIH (Grant CA68262 to G.W.). T.T. is a senior postdoctoral fellow of the American Cancer Society (Massachusetts Division). We thank Dr. Kylie Walters for proofreading the manuscript. We thank Dr. Koichi Uegaki for contributing to this project.

JA9926820

(19) Pasquinelli, A. M.; Dahlberg, J. E.; Lund, E. *RNA* **1995**, *1*, 957−967.

(20) Helm, M.; Brule, H.; Giege, R.; Florentz, C. *RNA* **1999**, *5*, 618−621.

(21) Sekine, M.; Kadokura, M.; Satoh, T.; Seio, K.; Wada, T. *J. Org. Chem.* **1996**, *61*, 4412−4422.

(22) Iwase, R.; Maeda, M.; Fujiwara, T.; Sekine, M.; Hata, T.; Miura, K. *Nucleic Acids Res.* **1992**, *20*, 1643−1648.

(23) Bartels, C.; Xia, T.-H.; Billeter, M.; Güntert, P.; Wüthrich, K. *J. Biomol. NMR* **1995**, *5*, 1−10.