SUPPLEMENTARY INFORMATION

A general method for rapid and cost efficient large-scale production of 5' capped RNA.

Anna-Lisa Fuchs[1], Ancilla Neu[1] and Remco Sprangers*

Max Planck Institute for Developmental Biology, Spemannstrasse 35, 72076 Tübingen, Germany

[1] Equal contribution.

* To whom correspondence should be addressed.

Email: remco.sprangers@tuebingen.mpg.de

```
gaaattaatacgactcactataggggaattgtgagcggataacaattcccctctagaaat
      T7 promotor #1                                    XbaI

aattttgatttaactttaagaaggagatatacc
                    RBS
```

START OF THE D1 PROTEIN

```
atgaaacatcaccatcaccatcaccccatgagcgattacgacatccccactactgagaat
 M  K  H  H  H  H  H  H  P  M  S  D  Y  D  I  P  T  T  E  N

ctttattttcagggcgccatggacgctaatgtcgtgtcttcttctaccatcgcaacctat
 L  Y  F  Q  G  A  M  D  A  N  V  V  S  S  S  T  I  A  T  Y
    TEV site      NcoI
attgacgctctggcaaaaaacgcctcggaactggaacaacgctcaaccgcgtatgaaatc
 I  D  A  L  A  K  N  A  S  E  L  E  Q  R  S  T  A  Y  E  I

aacaatgaactggaactggtgtttatcaaaccgccgctgattacgctgaccaacgtggtt
 N  N  E  L  E  L  V  F  I  K  P  P  L  I  T  L  T  N  V  V

aatatcagcaccattcaggaatcttttattcgtttcacggttaccaacaaagaaggcgtc
 N  I  S  T  I  Q  E  S  F  I  R  F  T  V  T  N  K  E  G  V

aaaatccgcacgaaaattccgctgagcaaagttcatggtctggatgtgaaaaacgttcaa
 K  I  R  T  K  I  P  L  S  K  V  H  G  L  D  V  K  N  V  Q

ctggtcgacgcaatcgataatattgtgtgggaaaagaaaagcctggttaccgaaaatcgt
 L  V  D  A  I  D  N  I  V  W  E  K  K  S  L  V  T  E  N  R

ctgcataaagaatgcctgctgcgtctgagcacggaagaacgccacatctttctggactat
 L  H  K  E  C  L  L  R  L  S  T  E  E  R  H  I  F  L  D  Y

aaaaaatacggcagctctatccgcctggaactggtgaacctgatccaggctaaaaccaaa
 K  K  Y  G  S  S  I  R  L  E  L  V  N  L  I  Q  A  K  T  K

aacttcacgatcgatttcaaactgaaatattttctgggcagtggtgctcaatccaaaagt
 N  F  T  I  D  F  K  L  K  Y  F  L  G  S  G  A  Q  S  K  S

tccctgctgcatgcgatcaaccacccgaaaagtcgtccgaatacctccctggaaattgaa
 S  L  L  H  A  I  N  H  P  K  S  R  P  N  T  S  L  E  I  E

ttcaccccgcgcgacaacgaaacggtgccgtacgatgaactgattaaagaactgaccacg
 F  T  P  R  D  N  E  T  V  P  Y  D  E  L  I  K  E  L  T  T

ctgtcacgtcatatctttatggcgtcgccggaaaacgttattctgagcccgccgatcaat
 L  S  R  H  I  F  M  A  S  P  E  N  V  I  L  S  P  P  I  N

gccccgattaaaaccttcatgctgccgaaacaggacattgttggcctggatctggaaaac
 A  P  I  K  T  F  M  L  P  K  Q  D  I  V  G  L  D  L  E  N

ctgtatgcggtcacgaaaaccgatggtattccgatcaccattcgcgtgacgtcgaatggc
 L  Y  A  V  T  K  T  D  G  I  P  I  T  I  R  V  T  S  N  G

ctgtattgctactttacccacctgggttatattatccgttacccggttaaacgcattatc
 L  Y  C  Y  F  T  H  L  G  Y  I  I  R  Y  P  V  K  R  I  I

gactccgaagtcgtggttttcggcgaagcggtcaaagataaaaattggaccgtgtatctg
 D  S  E  V  V  V  F  G  E  A  V  K  D  K  N  W  T  V  Y  L
```

```
atcaaactgattgaaccggtgaacgccatcaacgatcgtctggaagaatcaaaatacgtg
 I  K  L  I  E  P  V  N  A  I  N  D  R  L  E  E  S  K  Y  V

gaatcgaaactggttgacatctgtgatcgcatcgttttcaaaagcaaaaaatacgaaggt
 E  S  K  L  V  D  I  C  D  R  I  V  F  K  S  K  K  Y  E  G

ccgttcaccacgacctctgaagtcgtggatatgctgagtacctatctgccgaaacagccg
 P  F  T  T  T  S  E  V  V  D  M  L  S  T  Y  L  P  K  Q  P

gaaggcgtgatcctgttttacagcaaaggtccgaaatctaacatcgacttcaaaatcaaa
 E  G  V  I  L  F  Y  S  K  G  P  K  S  N  I  D  F  K  I  K

aaagaaaacaccatcgatcaaacggccaatgttgtctttcgttatatgtcatcggaaccg
 K  E  N  T  I  D  Q  T  A  N  V  V  F  R  Y  M  S  S  E  P

attatctttggcgaaagctctatcttcgtggaatacaaaaaattctcgaacgataaaggc
 I  I  F  G  E  S  S  I  F  V  E  Y  K  K  F  S  N  D  K  G

ttcccgaaagaatacggcagcggtaaaattgtcctgtataacggtgtgaattacctgaac
 F  P  K  E  Y  G  S  G  K  I  V  L  Y  N  G  V  N  Y  L  N

aatatctattgcctggaatacattaacacccataatgaagttggcattaaatctgtggtt
 N  I  Y  C  L  E  Y  I  N  T  H  N  E  V  G  I  K  S  V  V

gtcccgatcaaatttattgcagaattcctggtcaacggtgaaatcctgaaaccgcgtatt
 V  P  I  K  F  I  A  E  F  L  V  N  G  E  I  L  K  P  R  I

gacaaaaccatgaaatacatcaacagtgaagattactacggtaaccagcataacatcatc
 D  K  T  M  K  Y  I  N  S  E  D  Y  Y  G  N  Q  H  N  I  I

gtggaacacctgcgcgaccaatctatcaaaatcggcgatatcttcaacgaagacaaactg
 V  E  H  L  R  D  Q  S  I  K  I  G  D  I  F  N  E  D  K  L

agtgatgtcggtcaccagtatgcgaacaatgataaatttcgtctgaacccggaagtgtcc
 S  D  V  G  H  Q  Y  A  N  N  D  K  F  R  L  N  P  E  V  S

tacttcaccaataaacgtacgcgcggcccgctgggtatcctgtcaaattatgtcaaaacc
 Y  F  T  N  K  R  T  R  G  P  L  G  I  L  S  N  Y  V  K  T

ctgctgatttcaatgtactgttcgaaaacgtttctggatgacagcaacaaacgcaaagtt
 L  L  I  S  M  Y  C  S  K  T  F  L  D  D  S  N  K  R  K  V

ctggccattgactttggcaatggtgcagatctggaaaaatatttctacggcgaaatcgct
 L  A  I  D  F  G  N  G  A  D  L  E  K  Y  F  Y  G  E  I  A

ctgctggttgcgaccgatccggacgcggatgccattgcacgtggcaacgaacgctataac
 L  L  V  A  T  D  P  D  A  D  A  I  A  R  G  N  E  R  Y  N

aaactgaattctggtatcaaaaccaaatactacaaattcgactacatccaggaaaccatt
 K  L  N  S  G  I  K  T  K  Y  Y  K  F  D  Y  I  Q  E  T  I

cgtagtgatacgttcgtgagttccgttcgcgaagtcttttatttcggcaaattcaacatc
 R  S  D  T  F  V  S  S  V  R  E  V  F  Y  F  G  K  F  N  I

atcgattggcaattcgccatccattattctttccatccgcgtcactacgcaaccgtgatg
 I  D  W  Q  F  A  I  H  Y  S  F  H  P  R  H  Y  A  T  V  M

aacaatctgagtgaactgacggcttccggcggtaaagttctgattacgacgatggatggt
 N  N  L  S  E  L  T  A  S  G  G  K  V  L  I  T  T  M  D  G

gataaactgtccaaactgaccgataagaaaaccttcattatccacaaaaacctgccgtca
 D  K  L  S  K  L  T  D  K  K  T  F  I  I  H  K  N  L  P  S
```

```
tcggaaaactacatgtcagtggaaaaaatcgccgatgaccgcattgtggtttataacccg
 S  E  N  Y  M  S  V  E  K  I  A  D  D  R  I  V  V  Y  N  P

agcacgatgtctaccccgatgacggaatacatcattaagaaaaacgatatcgtccgtgtg
 S  T  M  S  T  P  M  T  E  Y  I  I  K  K  N  D  I  V  R  V

tttaatgaatacggtttcgttctggtcgacaacgttgattttgcaaccattatcgaacgc
 F  N  E  Y  G  F  V  L  V  D  N  V  D  F  A  T  I  I  E  R

agcaaaaaattcatcaatggcgcttccacgatggaagatcgtccgtcaacgcgcaacttt
 S  K  K  F  I  N  G  A  S  T  M  E  D  R  P  S  T  R  N  F

ttcgaactgaatcgcggtgcaattaaatgtgaaggtctggatgtggaagatctgctgtcc
 F  E  L  N  R  G  A  I  K  C  E  G  L  D  V  E  D  L  L  S

tattatgtcgtgtatgtgttctctaaacgctaa
 Y  Y  V  V  Y  V  F  S  K  R  -


gctagctaggatccgaattcgagctcggcgcgcctgcaggtcgacaagcttgcggccgca
         BamHI

taatgcttaagtcgaacagaaagtaatcgtattgtacacggccgcataatcgaaattaat
                                    DuetUP2 Primer
                                   DuetDOWN1 Primer
acgactcactatagggggaattgtgagcggataacaattccccatcttagtatattagtta
T7 promotor #2

agtataagaaggagatatacat
         RBS       NdeI
```

START OF THE D12 PROTEIN

```
atggatgaaatcgtcaaaaatatccgcgaaggcacgcacgtcctgctgccgttctatgaa
 M  D  E  I  V  K  N  I  R  E  G  T  H  V  L  L  P  F  Y  E

accctgccggaactgaatctgtcactgggcaaatctccgctgccgagtctggaatatggt
 T  L  P  E  L  N  L  S  L  G  K  S  P  L  P  S  L  E  Y  G

gcaaactactttctgcagatttctcgtgtgaacgatctgaatcgcatgccgaccgacatg
 A  N  Y  F  L  Q  I  S  R  V  N  D  L  N  R  M  P  T  D  M

ctgaaactgttcacgcatgatatcatgctgccggaaagcgatctggacaaagtctacgaa
 L  K  L  F  T  H  D  I  M  L  P  E  S  D  L  D  K  V  Y  E

atcctgaaaatcaactccgttaaatactacggccgttcaaccaaagcggatgccgtggtt
 I  L  K  I  N  S  V  K  Y  Y  G  R  S  T  K  A  D  A  V  V

gcagacctgtccgctcgcaataaactgtttaaacgtgaacgcgatgctattaaatcgaac
 A  D  L  S  A  R  N  K  L  F  K  R  E  R  D  A  I  K  S  N

aatcacctgaccgaaaacaacctgtacatcagcgattacaaaatgctgacgtttgacgtg
 N  H  L  T  E  N  N  L  Y  I  S  D  Y  K  M  L  T  F  D  V

ttccgtccgctgttcgatttcgttaacgaaaaatactgcatcatcaaactgccgaccctg
 F  R  P  L  F  D  F  V  N  E  K  Y  C  I  I  K  L  P  T  L

tttggccgtggtgtgattgatacgatgcgcatctactgcagcctgttcaaaaatgtccgc
 F  G  R  G  V  I  D  T  M  R  I  Y  C  S  L  F  K  N  V  R

ctgctgaaatgtgtgtcggatagctggctgaaagactctgcgattatggtggccagtgac
 L  L  K  C  V  S  D  S  W  L  K  D  S  A  I  M  V  A  S  D
```

```
gtttgtaagaaaaacctggacctgtttatgtcccatgtcaaatcagtgaccaaaagctct
 V  C  K  K  N  L  D  L  F  M  S  H  V  K  S  V  T  K  S  S

agttggaaagacgttaattcggtccaatttagcattctgaacaatccggttgatacggaa
 S  W  K  D  V  N  S  V  Q  F  S  I  L  N  N  P  V  D  T  E

ttcatcaacaaattcctggaattctctaaccgtgtttacgaagcactgtattacgtccac
 F  I  N  K  F  L  E  F  S  N  R  V  Y  E  A  L  Y  Y  V  H

agtctgctgtactcctcaatgacctcggactccaaatccatcgaaaataaacatcaacgc
 S  L  L  Y  S  S  M  T  S  D  S  K  S  I  E  N  K  H  Q  R

cgcctggtgaaactgctgctgtaa
 R  L  V  K  L  L  L  -


ctcgagcaccaccaccaccaccactgagatccggctgctaacaaagcccgaaaggaagct
 XhoI

gagttggctgctgccaccgctgagcaataactagcataaccccttggggcctctaaacgg
                  T7 Terminator

gtcttgaggggttttttgctgaaaggaggaactatatccggat
```

FIGURE S1. Vaccinia virus Capping Enzyme. Plasmid design, DNA and protein sequences. (Internal ID: 1323).

Figure S2

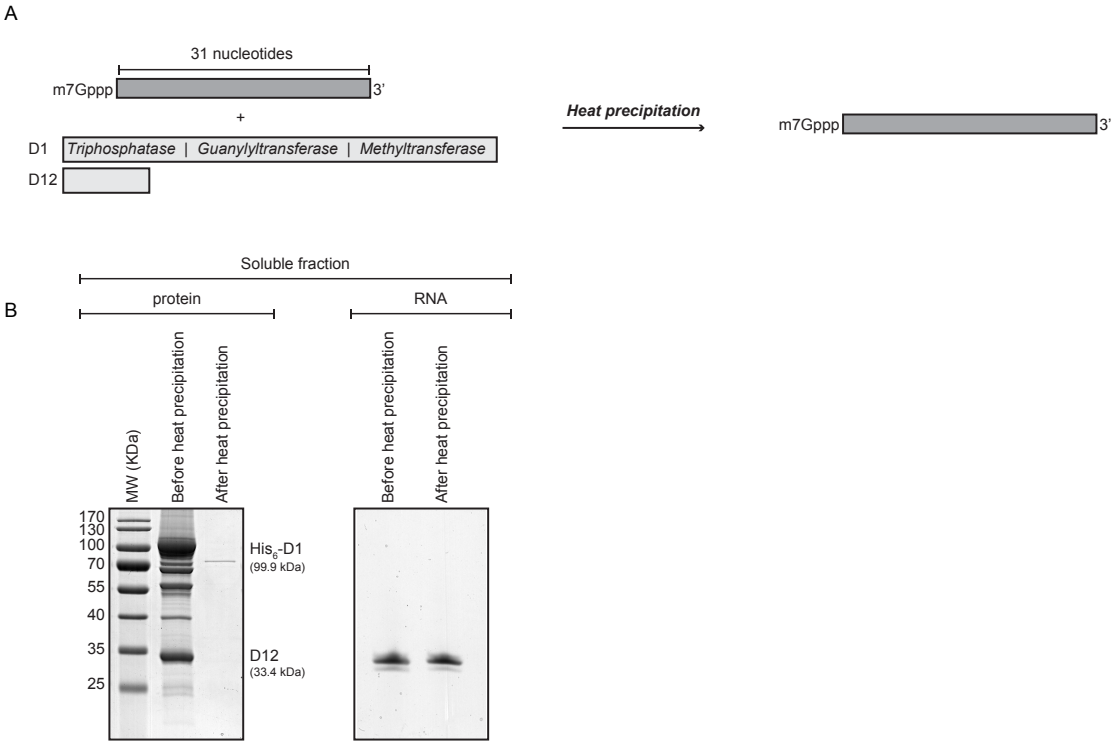

FIGURE S2: To separate the capped RNA from the capping enzyme the reaction mixture was incubated at 75°C for 10 minutes, followed by a 10 minute centrifugation step at 4500g. The capping enzyme was quantitatively removed from the mixture, whereas the RNA remained in solution.

Figure S3

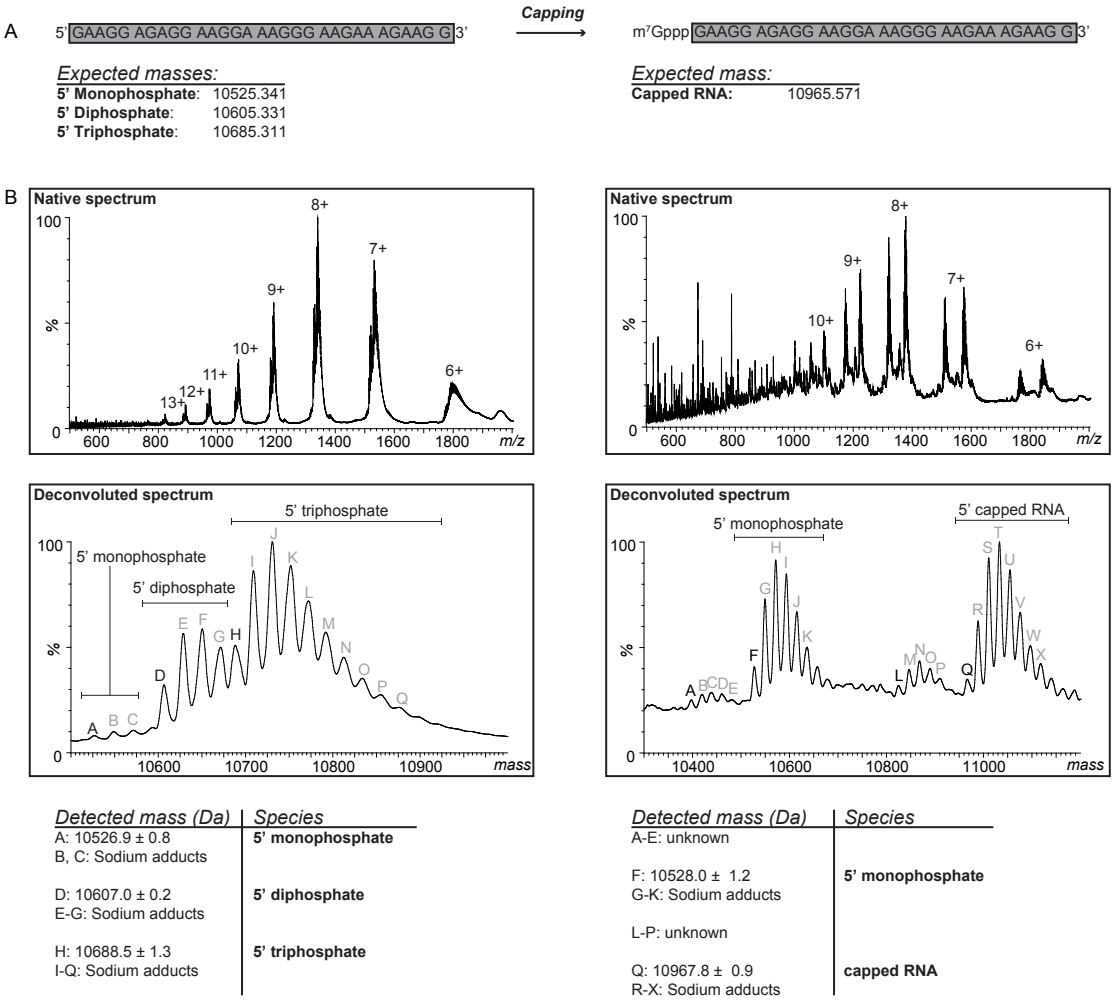

FIGURE S3. Native mass spectrometry analysis of the RNA capping reaction. *(A)* A mixture of 5' mono-, 5' di- and 5' triphosphate RNA was subjected to the vaccinia virus capping enzyme. The 3' end of the RNA contained a cyclic phosphate as a result of ribozyme cleavage. The RNA sequence and the expected masses of the substrates and product are indicated. *(B)* The experimental mass of the substrate confirmed the presence of mono-, di- and triphosphate at the 5' end of the substrate RNA. The experimental mass of the product shows a mixture of mono-phosphate RNA and of capped RNA. The analysis shows that both 5' di- and 5' triphosphate RNA are efficiently and properly capped by the vaccinia virus capping enzyme, whereas 5' monophosphate RNA appears not to be a substrate of the enzyme. The native mass spectrometry analysis was performed by AbLab (http://www.hecklab.com/ablab), a service unit within the Heck lab (http://www.hecklab.com) of the Utrecht University.

Figure S4

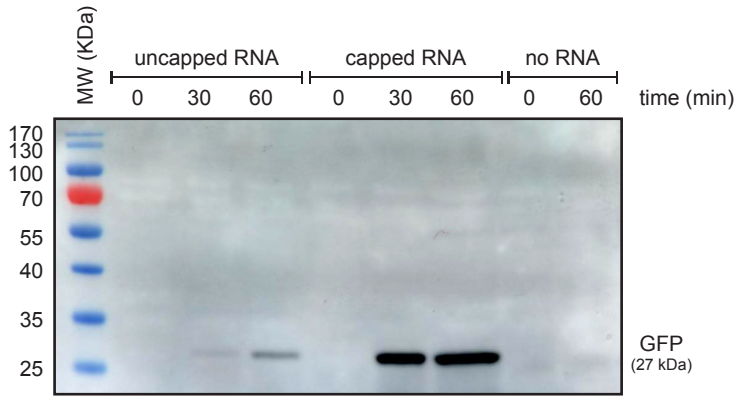

FIGURE S4. *In vitro* translation assays show efficient translation of a capped RNA that codes for the GFP protein.

Fractions of an *in vitro* translation reaction (see below) were loaded on an SDS PAGE gel and the transcribed GFP protein was visualized using a GFP antibody (see below). Samples were taken after 0, 30 or 60 minutes, as indicated on top of the gel. The left three lanes are samples that are taken from a translation reaction in the presence of RNA that has not been capped, in the middle three an RNA that was capped according to the protocols described here was used and in the right two lanes the reaction took place in the absence of RNA.

The DNA template coding for a GFP mRNA was prepared by polymerase chain reaction. PCR was carried out according to manufacturer's recommendations with Q5 polymerase (NEB) on a GFP carrying plasmid. In addition to the sequence specific part, the forward primer and reverse primer encompassed a Kozak sequence and a stop codon followed by 23 adenosine bases, respectively (Forward primer: CGTAATACGACTCACTATAGGCCGCCACCATGGTGAGCAAGGGCGAG, reverse primer: TTTTTTTTTTTTTTTTTTTTTTTTTTATTACTTGTACAGCTCGTCCATGCC). The resulting double stranded DNA was purified (Macherey-Nagel) and used as template in a transcription reaction at a final concentration of 8 ng/µl. The produced RNA was precipitated and dissolved in RNase free water and capped as described at a molar ratio of 1:100 at 37 °C for 1 hour. For uncapped RNA, the capping enzyme was omitted from the reaction. RNA was used without further purification for *in vitro* translation. Commercial wheat germ extract (Promega) was utilized following manufacturer's instruction including capped and uncapped RNA at 20µg/ml and 50 mM potassium acetate. Samples for SDS PAGE were taken at 0 min, 30 min and 60 min reaction time. Translated GFP protein was detected by western blotting using α-GFP (mouse) (Roche) and α-mouse-peroxidase (Sigma) antibodies as primary and secondary antibodies respectively. Chemiluminescence was developed in WesternBright substrate (Advansta) and detected on an Amersham Imager 600 (GE).

Figure S5

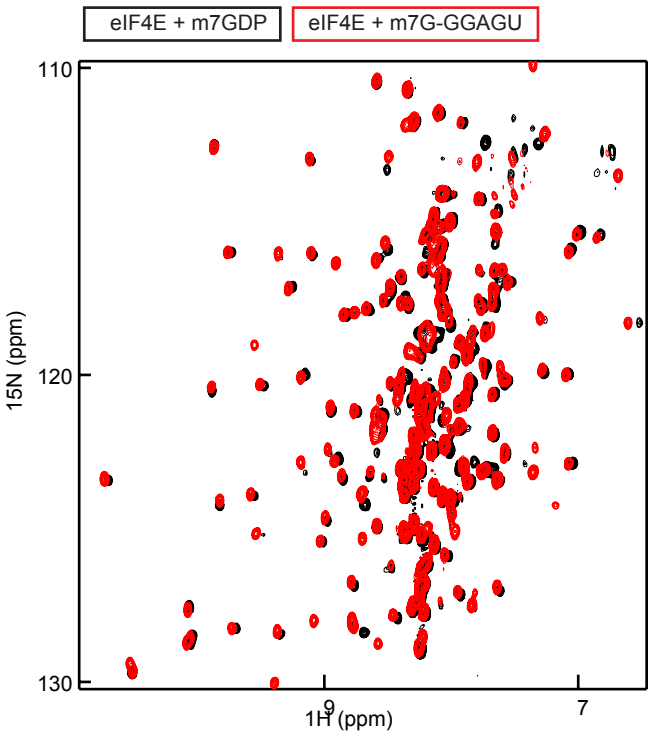

FIGURE S5. NMR spectra of *Dm* eIF4E in the presence of m^7GDP (black) and in the presence of capped RNA. The small chemical shift differences between the two spectra confirm that the RNA body does not interact extensively with the cap binding protein.

Table S1: RNA constructs used in this study.

Construct	RNA sequence	Anti-sense primer sequence
100GA14U	GGGAAGGAAGGGAAUAAAGAAG GGAAGAGGAAGGAGAGGAGGGA AGAAAGAAGAGGAGAGGAAGGA AGGGAAGAAAGAAGAGGGAAGA GGAAGGAGAGGA	TCCTCTCCTTCCTCTTCCCTCT TCTTTCTTCCCTTCCTTCCTCT CCTCTTCTTTCTTCCCTCCTCT CCTTCCTCTTCCCTTCTTTATT CCCTTCCTTCC**C**TATAGTGAGT CGTATTACG
80GA14U	GGGAAGGAAGGGAAUAAAGAAG GGAAGAGGAAGGAGAGGAGGGA AGAAAGAAGAGGAGAGGAAGGA AGGGAAGAAAGAAG	CTTCTTTCTTCCCTTCCTTCCT CTCCTCTTCTTTCTTCCCTCCT CTCCTTCCTCTTCCCTTCTTTA TTCCCTTCCTTCC**C**TATAGTGA GTCGTATTACG
60GA14U	GGGAAGGAAGGGAAUAAAGAAG GGAAGAGGAAGGAGAGGAGGGA AGAAAGAAGAGGAGAG	CTCTCCTCTTCTTTCTTCCCTC CTCTCCTTCCTCTTCCCTTCTT TATTCCCTTCCTTCC**C**TATAGT GAGTCGTATTACG
40GA14U	GGGAAGGAAGGGAAUAAAGAAG GGAAGAGGAAGGAGAGGA	TCCTCTCCTTCCTCTTCCCTTC TTTATTCCCTTCCTTCC**C**TATA GTGAGTCGTATTACG
15GA *	GGAGAAGAGAAGGAG	CTCCTTCTCTTCTC**C**TATAGTG AGTCGTATTA
15CU *	GGCCUCUUCGAAGCG	CGCTTCGAAGAGGC**C**TATAGT GAGTCGTATTA
Hp0 *	GCGGUUCGCCGCA	TGCGGCGAACCG**C**TATAGTGA GTCGTATTA
Hp2 *	GGGCGGUUCGCCGCA	TGCGGCGAACCGCC**C**TATAGT GAGTCGTATTA
Hp4 *	GGAAGCGGUUCGCCGCA	TGCGGCGAACCGCTTC**C**TATA GTGAGTCGTATTA
30GA2U	GTAAGGAGAGGAAGGAAGGGAA GAAAGAAG	CTTCTTTCTTCCCTTCCTTCCT CTCCTTA**C**TATAGTGAGTCGT ATTACG
30GA3U	GGTAGGAGAGGAAGGAAGGGAA GAAAGAAG	CTTCTTTCTTCCCTTCCTTCCT CTCCTAC**C**TATAGTGAGTCGT ATTACG
30GA5U	GGAGTGAGAGGAAGGAAGGGAA GAAAGAAG	CTTCTTTCTTCCCTTCCTTCCT CTCACTC**C**TATAGTGAGTCGT ATTACG
30GA15U	GGAGGAGAGGAAGGTAAGGGAA GAAAGAAG	CTTCTTTCTTCCCTTACCTTCC TCTCCTC**C**TATAGTGAGTCGT ATTACG
20AG	AGGAGGGAGAGGAAGGAAGG	CCTTCCTTCCTCTCCCTCC**T**AA TAGTGAGTCGTATTACG

Primers marked with * were methylated at the 2'-O position of the last two nucleotides. The sequence of the anti-sense primer duplexed with the promoter primer was underlined and the first transcribed nucleotide is marked in bold.

Table S2: T7 promoter sequences used in this study.

RNA promoter sequence	Promoter sequence
T7 promoter Φ6.5 (for start with G)	CGTAATACGACTCACTATAGG
T7 promoter Φ2.5 (for start with A)	CGTAATACGACTCACTATTAG