THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 279, No. 44, Issue of October 29, pp. 45573–45585, 2004

Printed in U.S.A.

Contribution of Trans-splicing, 5′-Leader Length, Cap-Poly(A) Synergism, and Initiation Factors to Nematode Translation in an *Ascaris suum* Embryo Cell-free System*[S]♦

Received for publication, July 6, 2004, and in revised form, August 18, 2004
Published, JBC Papers in Press, August 20, 2004, DOI 10.1074/jbc.M407475200

Sabbi Lall‡§, Cassandra C. Friedman‡, Marzena Jankowska-Anyszka¶, Janusz Stepinski‖, Edward Darzynkiewicz‖, and Richard E. Davis‡

From the ‡Department of Biology, City University of New York Graduate Center, College of Staten Island, Staten Island, New York 10314 and the ¶Department of Chemistry and ‖Department of Biophysics, University of Warsaw, 02-093 Warsaw, Poland

Trans-splicing introduces a common 5′ 22-nucleotide sequence with an *N*-2,2,7-trimethylguanosine cap ($m_3^{2,2,7}$GpppG or TMG-cap) to more than 70% of transcripts in the nematodes *Caenorhabditis elegans* and *Ascaris suum*. Using an *Ascaris* embryo cell-free translation system, we found that the TMG-cap and spliced leader sequence synergistically collaborate to promote efficient translation, whereas addition of either a TMG-cap or spliced leader sequence alone decreased reporter activity. We cloned an *A. suum* embryo *eIF4E* homolog and demonstrate that this recombinant protein can bind m^7G- and TMG-capped mRNAs in cross-linking assays and that binding is enhanced by eIF4G. Both the cap structure and the spliced leader (SL) sequence affect levels of *A. suum* eIF4E cross-linking to mRNA. Furthermore, the differential binding of eIF4E to a TMG-cap and to trans-spliced and non-trans-spliced RNAs is commensurate with the translational activity of reporter RNAs observed in the cell-free extract. Together, these binding data and translation assays with competitor cap analogs suggest that *A. suum* eIF4E-3 activity may be sufficient to mediate translation of both trans-spliced and non-trans-spliced mRNAs. Bioinformatic analyses demonstrate the SL sequence tends to trans-splice close to the start codon in a diversity of nematodes. This evolutionary conservation is functionally reflected in the optimal SL to AUG distance for reporter mRNA translation in the cell-free system. Therefore, trans-splicing of the SL1 leader sequence may serve at least two functions in nematodes, generation of an optimal 5′-untranslated region length and a specific sequence context (SL1) for optimal translation of trimethylguanosine capped transcripts.

Spliced leader (SL)[1] trans-splicing is present in diverse metazoan phyla including the chordates, cnidarians, platyhelminths, and nematodes (reviewed in Ref. 1). Spliced leader addition forms the mature 5′ end of the mRNA providing an unusual cap structure and common sequence to the mRNA. In the case of *Caenorhabditis elegans*, 70% of transcripts have one of two 22-base sequences, spliced leader 1 or spliced leader 2 (Ref. 2 and reviewed in Ref. 3). Because the spliced leader donor RNA exon has a 2,2,7-trimethylguanosine cap, rather than the usual monomethylguanosine cap, the product of the trans-splicing reaction is a TMG-capped mRNA (4–7). Trans-splicing may therefore affect translation initiation.

The major function of SL2 trans-splicing in *C. elegans* is to contribute to the resolution of polycistronic transcripts to generate mature mRNAs (8–10). Trans-splicing also plays this role in some other organisms, including the kinetoplastid protozoans, and perhaps platyhelminths (11, 12). However, it is not clear that resolution of polycistronic pre-mRNAs is the sole role for spliced leader trans-splicing, particularly in platyhelminths (see 13).[2] The predominant form of trans-splicing in *C. elegans* and other nematodes is SL1 trans-splicing (57% of *C. elegans* and at least 70% of *Ascaris suum* transcripts) (2, 14). However, the major role(s) of SL1 trans-splicing has yet to be defined in nematodes, and the role of trans-splicing in other organisms such as *Ciona intestinalis* and *Hydra vulgaris* also remains unknown (15, 16).

Since trans-splicing has a profound effect on the 5′-UTR of mRNAs, it has long been postulated that SL1 and TMG-cap addition might influence translational efficiency of mRNAs (see Ref. 17). Indeed a survey of SL-containing transcripts showed that trans-splicing tends to occur close to the start codon of mRNAs, providing circumstantial evidence for a relationship between trans-splicing and translation (3). While the nematodes are a major focus of both genomic and developmental biology research, *in vitro* translation studies have not been conducted in *C. elegans* because of the lack of a competent cell-free system. Embryos of the parasitic nematode *Ascaris* are highly suitable for preparation of several types of biochemically competent extracts that have been used to examine multiple molecular processes in nematodes (18–22). *Ascaris* translation extracts were used to show that 80–90% of transcripts are

* This work was supported by National Institutes of Health Grants AI49558, AI043300, and AI039714 and City University of New York-College of Staten Island startup funds (to R. E. D.) and by Grants PBZ-KBN-059/T09/10 and KBN 3 P04A 021 25 from the Polish Committee for Scientific Research to (E. D.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

♦ This article was selected as a Paper of the Week.

[S] The on-line version of this article (available at http://www.jbc.org) contains supplemental Figs. 1 and 2 and Tables 1–3.

§ Current address: Dept. of Biology, New York University, New York, NY 10012.

** To whom correspondence should be addressed. Current address: UCHSC at Fitzsimons, Depts. of Biochemistry and Molecular Genetics and Pediatrics, Mail Stop 8101, P. O. Box 6511, 12801 East 17th Ave., Aurora, CO 80045. Tel.: 303-724-3226; Fax: 303-724-3215; E-mail: richard.davis@uchsc.edu.

[1] The abbreviations used are: SL, spliced leader; $m_3^{2,2,7}$GpppG (or TMG), 2,2,7-trimethylguanosine; m^7GpppG (or m^7G), monomethylguanosine; UTR, untranslated region; PBS, phosphate-buffered saline; DTT, dithiothreitol; PMSF, phenylmethylsulfonyl fluoride; RACE, rapid amplification of cDNA ends; EST, expressed sequence tag; SAH, *S*-adenosylhomocysteine.

[2] R. E. Davis, unpublished data.

trans-spliced (14). This same study found that the TMG-cap and SL1 sequence collaborate to increase translational efficiency in a micrococcal nuclease treated *Ascaris* extract.

In recent years, development and characterization of *in vitro* eukaryotic translation systems have used cap dependence and synergy between the mRNA cap and poly(A)-tail as an important characteristic for validation of cell-free systems (Ref. 23 and reviewed in Ref. 24). Cap dependence and cap/poly(A) synergy are *in vivo* properties of translation, and some *in vitro* systems do not recapitulate these characteristics. Indeed only a few systems exhibiting cap/poly(A)-tail synergy exist, none of them allowing translation studies in a nematode cell-free system (25–29). The effect of trans-splicing on protein synthesis in nematodes in a system known to reflect *in vivo* conditions has not been carried out. In fact, little has been done to functionally characterize many attributes of nematode translation, including 5′-UTR length and start codon choice, factors known to be important in mammalian and yeast translation systems (30–33).

Translation initiation in yeast and mammals involves the binding of a multiprotein complex, eIF4F to the 5′ end of the mRNA (reviewed in Refs. 34 and 35). eIF4F consists of eIF4E, eIF4G, and eIF4A. Whereas eIF4E recognizes the 5′ mRNA cap structure, eIF4G acts as a coordinating scaffold, connecting eIF4E, eIF4A, and the poly(A)-binding protein (36–38). The interaction of eIF4G with eIF4E has been shown to enhance binding of the latter to the mRNA cap (39, 40). eIF4G is involved in recruiting the S 40 preinitiation complex to the mRNA, and the interaction of eIF4G with the poly(A)-binding protein may be the basis of the synergistic effect of a cap and poly(A)-tail upon translational efficiency (reviewed in Ref. 34). Of these events, only the binding of eIF4E proteins to methylated nucleoside triphosphates (as a proxy for the mRNA cap structure) has been examined in nematodes. Whereas many organisms have one to two homologs of eIF4E, *C. elegans* has five (named ife-1 through 5, Ref. 41) that differ in their binding specificity for mono- and trimethyl-cap analogs. Of these proteins, *C. elegans* IFE-3 and IFE-4 most closely resemble the two human eIF4E proteins in primary structure, and both bind to an m^7GTP-Sepharose column. IFE-1, -2, and -5 are capable of binding both m^7GTP- and $m_3^{2,2,7}$GTP-Sepharose matrices (41). Such binding experiments, along with structural modeling, were used to predict and define residues that might enable IFE-1, -2, and -5 binding to $m_3^{2,2,7}$GTP-Sepharose (42). These binding studies and *in vivo* functional experiments (43) were used to support the hypothesis that *C. elegans* IFE-3 and -4 might be responsible for m^7GpppG-based translation (with IFE-4 playing a non-essential role). IFE-1, -2, and -5 were predicted to play a role in $m_3^{2,2,7}$GpppG-cap based translation. However, conservation of these TMG-binding proteins across nematodes, differential binding of IFEs to a capped RNA (*i.e.* the *in vivo* substrate) rather than a cap analog, and their contribution to the translation of trans-spliced transcripts has not been examined.

In the current study, we describe and use an *Ascaris* cell-free system to examine the effect of trans-splicing on translation. We show that the *Ascaris in vitro* system is mRNA cap-dependent and retains translational synergy between the cap structure (both m^7G and TMG) and the poly(A)-tail, an effect that we have observed *in vivo* in separate studies.[3] We find that in our *Ascaris* extract, non-trans-spliced transcripts are most efficiently translated. Whereas either the TMG-cap or SL sequence alone have a detrimental effect on protein synthesis, together they promote translation of a reporter mRNA. We then show that *Ascaris* eIF4E-3 is capable of binding both m^7G- and TMG-capped RNA. The cross-linking efficiency of *Ascaris* eIF4E-3 to mRNA may explain the translational activity of trans-spliced *versus* non-trans-spliced transcripts in the *in vitro* system. We have also used a bioinformatic analysis to show that the spliced leader (SL1) tends to splice close to the translation start codon across nematode species. This correlates in part with the most efficient SL to start codon distance for translation in our cell-free system and may therefore reflect evolutionary selection for optimal expression. The *Ascaris* system should be valuable for further characterizing and understanding nematode translation and the role of trans-splicing in post-transcriptional gene expression. The parasite cell-free extract may also prove useful for drug testing and identification of targets for rational drug design.

EXPERIMENTAL PROCEDURES

A. suum Extract Preparation—Embryos were collected from the proximal third (~4 cm) of *A. suum* uteri (Carolina Biological Supply). After rinsing in cold 1× PBS (140 mM NaCl, 2.7 mM KCl, 8 mM Na_2HPO_4, 2 mM $KHPO_4$ (pH 7.5)), the uterine tissue was dissolved in 0.5 N NaOH. Embryos were washed extensively in cold H_2O, twice in cold 1× PBS (pH 2.0), and stored in 5 volumes of PBS (pH 2.0) at 4 °C until required. Development was initiated by agitating embryos at 110 rpm, 30 °C in 1× PBS (pH 2.0). At day 5 (32–64 cell stage, determined by 4′,6-diamidino-2-phenylindole staining), embryos were spun for 5 min at 750 × *g*, and the eggshells were removed by a 90-min 0.4 N KOH, 1.4% sodium hypochlorite treatment at 30 °C. After four washes in 10–20 volumes of cold 1× PBS (pH 7.4), the embryo pellet was resuspended in 0.4 volume of buffer A (10 mM KCl, 1.5 mM $MgCl_2$, 20 mM Tris (pH 7.9), 1 mM DTT, 1× Roche Applied Science EDTA-free complete protease inhibitors, 0.1 mM PMSF, 2.5 μg/ml leupeptin, and 2 μg/ml pepstatin). The embryo suspension was homogenized within 5 min with 8–10 passes in a 15 ml Wheaton Dura-Grind® stainless steel dounce (Millville, NJ) and the homogenate centrifuged at 27,000 × *g*, 4 °C, 20 min. The intermediate layer (soluble non-lipid layer) was dialyzed twice against 1000–2000 volumes of ice-cold buffer (50 mM KCl, 20 mM Tris-HCl (pH 7.9), 0.2 mM EDTA, 1 mM DTT, 0.5 mM PMSF, 20% glycerol). The extract was then centrifuged (21,000 × *g*, 4 °C, 10 min), frozen in liquid nitrogen, and stored at −80 °C until use.

In Vitro Transcription—PCR-generated templates for *in vitro* transcription were synthesized using the primers in supplemental Table I (which introduce a T7 promoter), from either pRL-null for *Renilla* luciferase or pGL3-Basic for firefly luciferase (Promega). The *in vitro* transcription reactions (T7 Megascript, Ambion) typically contained 1 μg of template, 10 mM ATP, 10 mM CTP, 10 mM UTP, 12 mM cap analogue (ApppG, m^7GpppG, or $m_3^{2,2,7}$GpppG, Ref. 44), 1.5 mM GTP, 1× Reaction buffer, 1× T7 polymerase mix and was carried out at 37 °C for 4 h. Uncapped transcription reactions lacked cap analog. After DNase I treatment, the mRNA transcript was purified, qualified by formaldehyde denaturing gel electrophoresis, and quantified using UV spectroscopy and the Ribogreen reagent (Molecular Probes). Analysis of cap orientation on transcripts was carried out as described (45) and indicated that 90% of the m^7GpppG and 80% of the $m_3^{2,2,7}$GpppG caps were added in the correct orientation. These analyses demonstrated that differences in the translation efficiencies observed were not a function of differences in cap orientation on the transcripts.

Uniformly labeled transcripts for UV cross-linking were synthesized with the Promega T7 Riboprobe kit according to the manufacturer's instructions using 0.5 μg of template in a 20-μl reaction containing 50 μCi of [α-^{32}P]GTP (3000 Ci/mmol, PerkinElmer Life Sciences), a 0.5 mM concentration each of CTP, UTP, and ATP, 0.2 mM GTP, and 1.6 mM cap analog. Uncapped transcripts were similarly generated without cap analog. The transcripts were gel-purified from 5% denaturing PAGE gels. Cap-labeled RNAs were prepared from uncapped RNA substrates using [α-^{32}P]GTP (PerkinElmer Life Sciences) and recombinant vaccinia RNA guanylyltransferase and (guanine-N^7)-methyltransferase (generously provided by Stewart Shuman) or human guanylyltransferase (generously provided by Aaron Shatkin) (46). RNA sequences used for translation and cross-linking are provided in supplemental Table II.

In Vitro Translation and Reporter Quantification—Ten-microliter translation reactions contained 5–12.5 μg/ml reporter RNA, 50% *Ascaris* embryo extract (~125–150 μg of extract protein), 10 mM Tris-HCl

[3] L. S. Cohen, C. Mihkli, M. Jankowska-Anyszka, J. Stepinski, E. Darzynkiewicz, and R. E. Davis, manuscript in preparation.

(pH 7.9), 125 mM KOAc, 25 mM KCl, 2.6 mM $Mg(OAc)_2$, 0.5 mM GTP, 0.5 mM ATP, 10 mM creatine phosphate, 50 μg/ml creatine kinase, 2.5 mM DTT, 0.1 mM EDTA, 50 μM Promega amino acid mix, 0.25 mM PMSF, and 10% glycerol. After 30–60 min at 30 °C, the translation reaction was diluted in the appropriate lysis buffer (Promega) and frozen on dry ice. Typically, 0.1% of the reaction was read in a Sirius luminometer v.2.2 (Berthold Detection Systems). Dual assay experiments were performed by mixing 62.5 ng of test mRNA (*Renilla* luciferase) with 62.5 ng m^7G-firefly luciferase. After 1 h, translation was assayed using the Promega dual luciferase assay kit. Statistical analysis was carried out using Prism 4.0.

Cloning and Sequence Analysis of eIF4E Genes—*A. suum eIF4E-3* (GenBank™ Accession number AY599493) was cloned from 32–64 cell cDNA (synthesized as described below) (12, 47–50) by 5′ RACE using a degenerate primer (supplemental Table I) based on conserved residues identified in nematode eIF4E sequences. RACE products were directionally cloned into the pAMP vector (Invitrogen), as described by the manufacturer. 3′ RACE was then used to obtain the entire open reading frame of *A. suum eIF4E-3* (primers in supplemental Table I). *C. elegans ife-1* and *ife-3* coding regions were obtained from mixed stage oligo(dT) or random primed *C. elegans* cDNA using the primers in supplemental Table I. Phylogenetic analysis was performed using neighbor-joining and Bayesian methods (see overviews in Refs. 51–53).

Recombinant Protein Production and Nucleotide Binding Studies—The eIF4E open reading frames were cloned into the NdeI/BamHI sites of pET16b and transformed into Rosetta(DE3) bacteria (Novagen). Protein production was induced overnight (25 °C) using 0.4 mM isopropyl β-D-thiogalactopyranoside. The bacterial pellet was frozen at −80 °C, then sonicated in lysis buffer (20 mM HEPES-KOH (pH 7.5), 100 mM KCl, 2 mM EDTA, 1 mM DTT, 10% glycerol, 0.4× Roche Applied Science EDTA-free protease inhibitors), and the fusion protein was purified on a Ni^{2+}-nitrilotriacetic acid-agarose column (Qiagen). The bound protein was eluted using an imidizole gradient (10–500 mM), dialyzed against buffer B (50 mM KCl, 20 mM Tris (pH 7.5), 0.2 mM EDTA, 1 mM DTT, 0.5 mM PMSF, 20% glycerol) and stored at −80 °C.

For cap analog binding studies, recombinant protein was added to m^7GTP-Sepharose or $m_3^{2,2,7}$GTP-Sepharose 4B beads (41) and incubated for 45 min at 4 °C on a rotating platform. The beads were collected, the supernatant removed (unbound fraction), and the beads washed three times in buffer B (above). Protein bound to the beads and in the unbound fraction were analyzed by 12.5% SDS-PAGE to determine protein binding efficiency to the matrices. Proteins were detected by Coomassie Blue staining or SYPRO Ruby labeling (Molecular Probes).

UV cross-linking was performed by incubating 500 ng of recombinant protein with 250,000 dpm of uniformly labeled RNA (or 100,000 dpm of cap-labeled RNA) for 10 min at 30 °C in 20 μl of buffer (100 mM KOAc, 2 mM $Mg(OAc)_2$, 20 mM HEPES-KOH (pH 7.5), 2 mM DTT, 3% glycerol). Cross-linking was then performed in a Hoefer UV Cross-linker, at a distance of 4 cm (120,000 $\mu J/cm^2$) for 30 min on ice. The cross-linked mRNA was then treated with 20 μg of RNase A (United States Biochemical, 37 °C for 30 min) and analyzed by SDS-PAGE on a 4–15% Tris-HCl gel (Bio-Rad), and the gel was exposed to Kodak BioMax film overnight at −80 °C. Cross-linked proteins were also imaged and analyzed using the STORM 860 and ImageQuant programs (Amersham Biosciences).

Search for Spliced Leaders in Ascaris Embryos—To assess the presence of SL sequences on *Ascaris* embryo mRNAs, an oligonucleotide-capped cDNA library was constructed from 32–64 cell stage *A. suum* embryo RNA (12, 47–50). In brief, total RNA was dephosphorylated with alkaline phosphatase (Roche Applied Science), decapped using tobacco acid pyrophosphatase (Epicenter Technologies), and an RNA oligonucleotide (see supplemental Table I) ligated to the 5′ end of RNAs using T4 RNA ligase (New England Biolabs) under the conditions described by the manufacturers. First strand cDNA was synthesized using an adaptor oligo(dT) primer and Superscript II reverse transcriptase (Invitrogen) at 42 °C for 30 min, followed by addition of more Superscript enzyme and a further incubation at 50 °C for 30 min. Reverse transcription-PCR products were generated using primers complementary to the RNA oligonucleotide and the 3′ adaptor sequence. PCR products ~500–1000 and 1000–3000 bp were gel-purified and directionally cloned using ligation-independent cloning with the pAMP system (Invitrogen). The 5′ ends (only ~30 nucleotides) of ~1500 randomly selected clones were directly sequenced from bacterial lysates using ^{32}P-end-labeled primers and cycle sequencing.

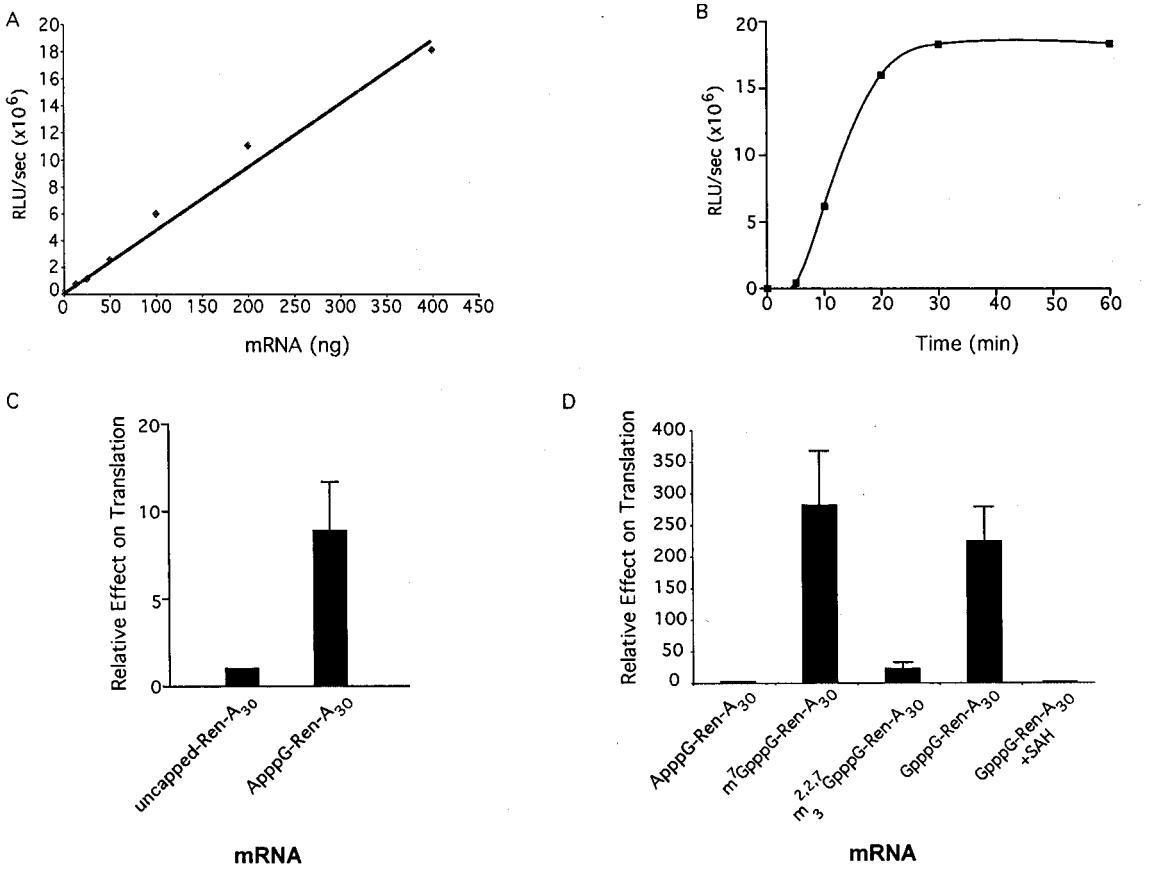
Bioinformatic Analysis of SL to AUG Distance—The position of the spliced leader sequence relative to the start codon in trans-spliced mRNAs was ascertained using cDNA and EST sequences from available developmental stages in current databases. We took the following steps to ensure that only cDNA sequences that truly contain a spliced leader were assessed. Nematode cDNA sequences that contain an entire spliced leader sequence at their 5′ end were identified using a BLASTN search of the NCBI nr data base. Additional sequences were found using BLASTN searches of EST data bases at NCBI and the NEMBLAST server (nema.cap.ed.ac.uk/ncbi_blast.html). Libraries based on PCR amplification with an SL containing primer were not used (to avoid sequences where an SL had been fortuitously introduced by PCR primers). The predicted start codon was either that specified in the data base entry or the most 5′ AUG of the longest open reading frame in the sequence (in the case of ESTs).

RESULTS

The A. suum in Vitro Translation Extract Is Cap-dependent and Shows Synergy between the Cap and Poly(A)-tail—*A. suum* embryonic extracts have previously been used to study translation *in vitro* (14). However, these experiments were carried out following micrococcal nuclease treatment of endogenous mRNAs and several characteristics of the cell-free system, including potential cap dependence and cap/poly(A)-tail synergism, were not described. To further develop and characterize a nematode *in vitro* system, we modified the extract preparation (see "Experimental Procedures"), the micrococcal nuclease treatment was excluded and demonstrated that the extracts were cap-dependent and exhibited cap/poly(A)-tail synergism (see below). Our preparation yielded extracts with increased protein concentration and allowed evaluation of reporter translation under conditions where exogenously added transcripts compete with endogenous mRNAs for translation complexes.

There was a linear relationship between reporter transcript concentration up to at least 400 ng (*m^7GpppG-Ren-A_{85}*) and translation as measured by *Renilla* activity (Fig. 1*A*). This linear message dependence extended up to 1 μg of RNA for the firefly luciferase reporter (data not shown). We assayed 50–125 ng of transcript in subsequent experiments to evaluate translation at the lower end of the linear range of the dose-response relationship. Time courses of reporter translation showed there is a 5-min lag followed by *Renilla* luciferase accumulation with time, leveling off at around 30 min (Fig. 1*B*). The leveling off was observed with different transcripts and seems to be due to diminishing extract activity rather than mRNA instability insofar as ~70% of uniformly labeled full-length mRNAs persisted until at least 60 min (data not shown).

The cap dependence of translation was assessed by comparing reporter translation from uncapped and ApppG-, GpppG-, and m^7GpppG-capped transcripts. Uncapped transcript translation was barely detectable above background (Fig. 1*C*). ApppG-capped transcripts demonstrated a low but reproducible level of translation in the presence (Fig. 1*C*) or absence of a poly(A)-tail (*ApppG-Ren*, Fig. 1*E*). Substitution of the ApppG-cap on a reporter RNA with a GpppG-cap led to a 200-fold increase in reporter detection (*GpppG-Ren-A_{30}*, Fig. 1*D*). This initially suggested that translation in nematodes might not be dependent on N^7-methylation of the first guanine of GpppG. However, addition of 0.5 mM *S*-adenosylhomocysteine (SAH), a methylation inhibitor, to the translation reactions virtually eliminated translation of the GpppG-capped RNA (Fig. 1*D*) but not m^7GpppG-capped RNA (data not shown). These data suggested that efficient translation of a GpppG-capped transcript might be dependent on N^7-methylation of the cap to m^7GpppG in the translation extract. This was confirmed using transcripts containing a radiolabeled cap (Gp*ppG-RNA). Following 30–60 min of translation in the cell-free system, the methylation status of the mRNA cap was assessed by resolving the RNA caps derived from P1 or T1 digestion of the RNA by thin-layer chromatography on polyethyleneimine-cellulose developed with 0.45 M ammonium sulfate (46). The Gp*ppG-cap-labeled transcripts were efficiently N^7-methylated and well



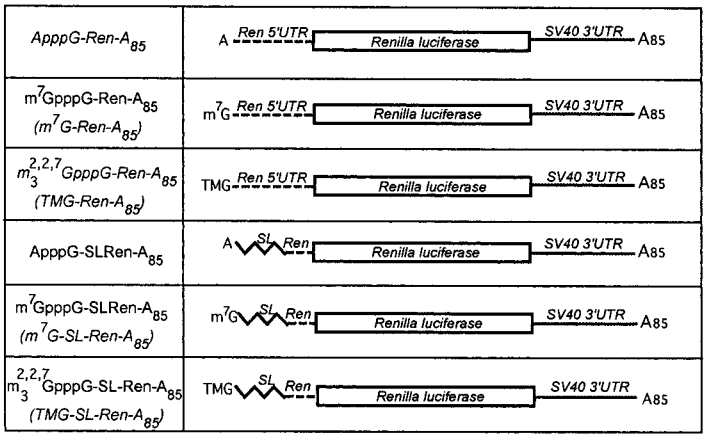
E

mRNA sample	*Renilla* Luciferase Activity (RLU/sec)	Effect of methylated cap on translation[a]	Effect of poly(A)-tail on translation[b]		Effect of cap and poly(A)-tail on translation[c]	Cap and tail synergistic effect[d]
ApppG-Ren	2400	1	1		1	
m^{7}GpppG-Ren	638,400	266		1		
ApppG-Ren-A$_{85}$	226,100	1	94			
m^{7}GpppG-Ren-A$_{85}$	41,670,200	184		65	17362	48
ApppG-SL-Ren	1,000	1	1		1	
m$_3^{2,2,7}$GpppG-SL-Ren	114,500	115		1		
ApppG-SL-Ren-A$_{85}$	92,000	1	92			
m$_3^{2,2,7}$GpppG-SL-Ren-A$_{85}$	17,979,000	195		157	17979	87

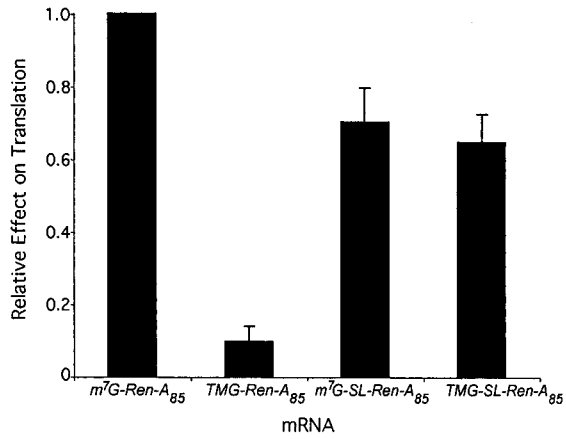
FIG. 1. **Characterization of the *A. suum in vitro* translation system.** *A*, *Renilla* luciferase mRNA had a linear dose-response relationship for the given concentration range when incubated in translation extract at 30 °C for 1 h. RLU/sec, relative light units/second of 0.1% of the translation reaction. *B*, a time course of reporter accumulation from 125 ng of *m^{7}GpppG-Ren-A$_{85}$* transcript indicated that reporter accumulation occured until at least 30 min after mRNA addition. *C*, ApppG-capped transcript translational efficiency relative to an uncapped mRNA (ppp-RNA). *D*, translational efficiency in the *A. suum* extract is methylation dependent, as shown by reporter activity of m^{7}G- or TMG-capped transcripts relative to an ApppG-capped *Renilla* luciferase mRNA. A GpppG-capped transcript is translated in the extract, except in the presence of 0.5 mM SAH (a methylation inhibitor) (see "Results"). SAH itself has no significant effect on the translation of m^{7}GpppG capped RNA (data not shown). *E*, effect of transcript cap and poly(A)-tract on translation in *Ascaris* extracts. Translational efficiency in the *A. suum* extract was methylated cap-dependent, as shown by reporter activity of m^{7}GpppG- or m$_3^{2,2,7}$GpppG-capped transcripts relative to an ApppG-capped *Renilla* luciferase mRNA. Translational efficiency was also poly(A)-dependent, as shown by increased *Renilla* luciferase production when an A$_{85}$ tail was added. The cap and poly(A)-tail had a synergistic effect on translation in *A. suum* extract, shown by the greater than additive effect on translation of endowing a transcript with an m^{7}GpppG- or TMG-cap as well as a tail. Numbers in *E* are from a representative experiment with two replicates that was repeated on at least five occasions (combined data shown in *C* and *D*). All translation experiments illustrated were carried out in duplicate at least twice using independently prepared extracts and RNA transcripts. Footnotes in *E*: a, activity of given reporter/activity of ApppG-capped mRNA with same tail; b, activity of given reporter/activity of capped mRNA without tail; c, activity of given reporter/activity of ApppG-capped and untailed mRNA; d, synergistic effect of cap and tail ($c/(a+b)$) for given transcript.

translated in the absence of the methylation inhibitor, SAH, but not methylated and poorly translated in the presence of SAH (Fig. 1*D*, data not shown). Furthermore, comparison of the translation of an ApppG- compared with m^{7}GpppG-capped transcript showed that the latter reporter had drastically more activity (266-fold, Fig. 1, *D* and *E*). In addition, no significant methylation or demethylation of guanine was observed in the extracts during translation of an m^{7}Gp*ppG-capped reporter.

A



B



C

mRNA sample	*Renilla* Luciferase (RLU/sec)	Effect of methylated cap on translation[a]	Effect of spliced leader on translation[b]		Effect of methylated cap and SL on translation[c]	Cap and SL synergistic effect[d]
ApppG-Ren-A85	226,100	1	1		1	
m7GpppG-Ren-A85	41,670,200	184		1		
ApppG-SL-Ren-A85	92,900	1	0.4			
m7GpppG-SL-Ren-A85	17,612,400	190		0.4	78	0.4
ApppG-Ren-A85	226,100	1	1		1	
m3 2,2,7GpppG-Ren-A85	3,404,300	15		1		
ApppG-SL-Ren-A85	92,900	1	0.4			
m3 2,2,7GpppG-SL-Ren-A85	17,979,100	194		5.3	80	5.1

FIG. 2. **The SL1 sequence synergistically enhances translation of TMG-capped transcripts but negatively affects translation of m^{7}G-capped transcripts.** *A*, transcripts used to study the effect of cap/SL1 combinations in *A. suum* translation extracts. The *left column* indicates names and abbreviated names (in *parentheses*) used to refer to transcripts. *B*, while a TMG cap alone reduced translation of the reporter mRNA (relative to *m^{7}G-Ren-A$_{85}$*), the SL1 sequence significantly enhanced the translation levels of TMG-capped RNAs. However, the SL1 sequence diminished relative translational efficiency of an m^{7}G-capped reporter transcript. *Error bars* reflect variation in five experiments using independently prepared extracts and RNA transcripts. *C*, interaction of the TMG-cap with the spliced leader sequence was synergistic, as shown in a representative experiment involving two sets of replicates. Footnotes in *C*: a, activity of given reporter/Activity of ApppG-capped reporter with same 5′-UTR; b, activity of given reporter/activity of methylated, capped reporter without SL1; c, activity of given reporter/Activity of uncapped reporter without SL1; d, synergistic effect of cap and SL sequence ($c/(a+b)$) for given transcript.

In all subsequent experiments, ApppG-capped RNA was used as the baseline control representing m^{7}GpppG cap-independent translation.

We then tested the effect of a hypermethylated (TMG, m$_3^{2,2,7}$GpppG) cap on translation. TMG-caps dramatically increase the efficiency of reporter translation, up to 195-fold relative to ApppG-capped mRNA (in the presence of SL1 and an A$_{85}$ tail, Fig. 1*E*). Together this shows that the cell-free system is extremely methylated cap-dependent and that both the m^{7}G- and TMG-cap promote reporter production.

Poly(A)-tail dependence was assessed by comparing the translational efficiency of m^{7}G- and TMG-capped *Renilla* luciferase transcripts with and without a poly(A)-tract (A$_{30}$ and A$_{85}$). The effect of the poly(A)-tail on translation was found to be similar for both m^{7}G- and TMG-capped transcripts, with an increase in reporter activity upon addition of an A$_{85}$ tract (Fig. 1*E*). Addition of an A$_{85}$ tail to the transcript led to a 65- and 157-fold increase in reporter production from m^{7}G- and TMG-capped transcripts, respectively. Transcripts with a poly(A)-tract of 85 nucleotides demonstrated a higher translation efficiency in the extract (~2.5-fold) than transcripts with a shorter 30 nucleotide poly(A)-tract.

We then tested how translation is affected by addition of both a cap and poly(A)-tail. The presence of the m^{7}GpppG-cap with an A$_{85}$ tail increased *Renilla* luciferase reporter translation by up to 17,362-fold, compared with an ApppG-capped transcript without a tail (*ApppG-Ren*, Fig. 1*E*). This is much greater than the predicted additive effect based on reporter transcripts carrying an m^{7}GpppG-cap or A$_{85}$ tail alone. We therefore conclude that this is a synergistic effect (Fig. 1*E*) based on the equation: effect on translation of both a cap and tail/(effect of cap alone + effect of tail alone). By this equation, the m^{7}G-cap and poly(A)-tail have a synergistic effect of 48. The TMG-cap and poly(A)-tail also drastically promoted translation when compared with *ApppG-SLRen* (17,979-fold increase). The TMG-cap and poly(A)-tail had a synergistic effect of 87 on reporter production. The same effects were seen with multiple preparations of mRNA as well as in independently prepared extracts. Therefore both the m^{7}G- and TMG-cap synergistically enhance translation when combined with a poly(A)-tail in the *A. suum* embryonic extract.

The Effect of Cap Structure and the Spliced Leader Sequence on Translation in the A. suum Cell-free System—Since the discovery of trans-splicing, it has been speculated that spliced leader addition might have a profound effect on mRNA translation. Therefore, we set out to test the hypothesis that trans-splicing affects translational efficiency in *A. suum*. To test the translational efficiency of mRNAs with different 5′-leaders in

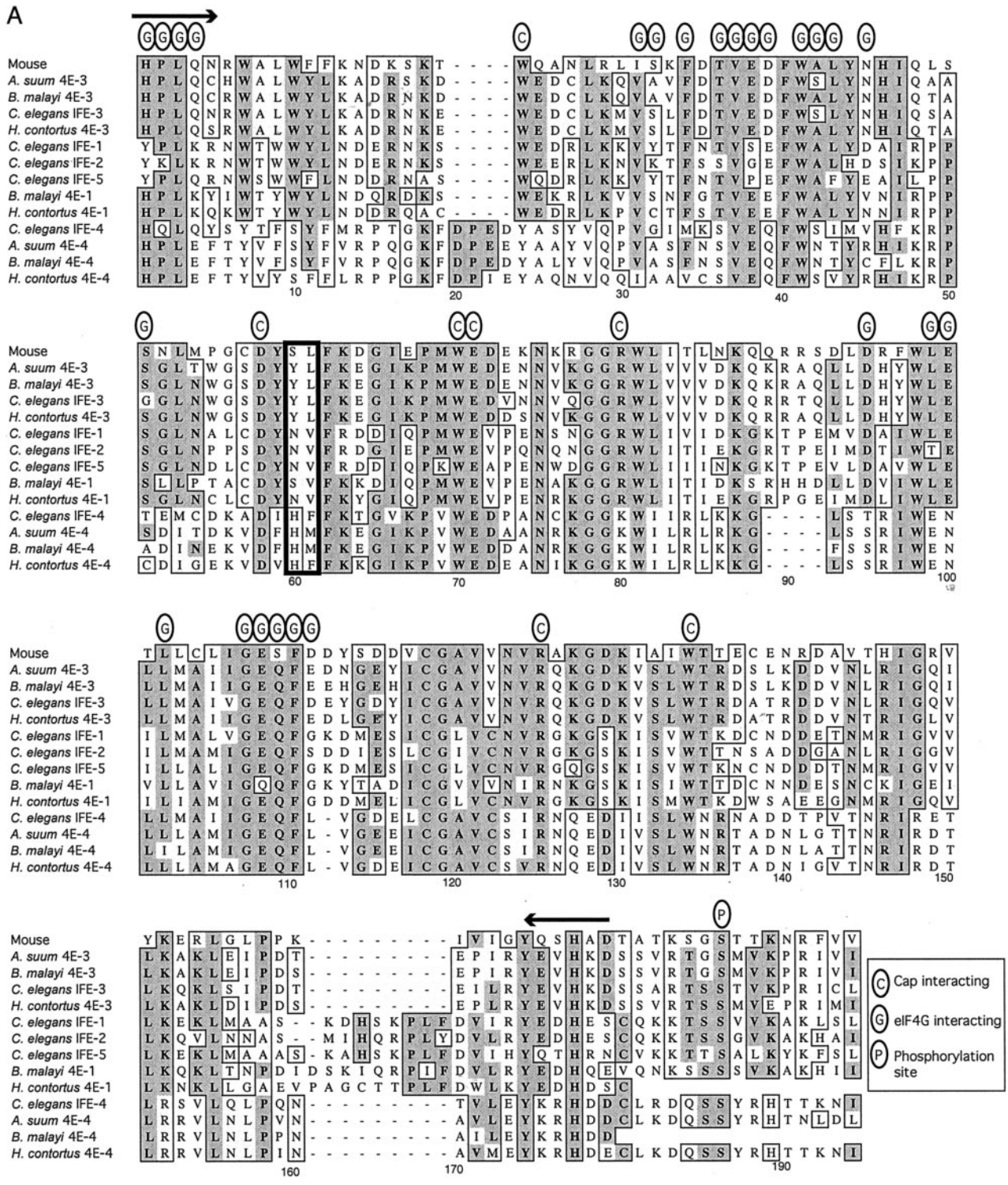


FIG. 3. **Multiple sequence alignment and phylogenetic analyses of the nematode eIF4E proteins reveal that *A. suum* eIF4E-3 is most closely related to *C. elegans* IFE-3.** *A*, alignment of the *A. suum* eIF4E-3 with other nematode homologs. Only the conserved region of eIF4E (~*Ascaris* amino acid 50–230) are illustrated in the alignment. *Circled letters* indicate residues previously implicated in specific functions, *G* = eIF4G interacting, *C* = cap binding, and *p* = phosphorylation site as described (34, 40). *Boxed residues* at 60–61 are thought to be involved in TMG-cap binding specificity as described for *C. elegans* eIF4E-5 (42). *Arrows* indicate the region used in phylogenetic analyses. *B*, phylogenetic analysis using the Neighbor-joining method (and strongly supported by Bayesian analysis of the corresponding nucleotide sequence, data not shown) supported the hypothesis that at the sequence level *A. suum* eIF4E is most closely related to *C. elegans* IFE-3. *Starred nodes* are supported with a posterior probability of >0.83 in Bayesian likelihood analysis, indicating strong support at these positions using multiple phylogenetic analysis tools. Species are *Brugia malayi*, *Hemonchus contortus*, *C. elegans*, and *A. suum*.

our system, we synthesized the transcripts represented in Fig. 2A. *In vitro* translation of *TMG-Ren-A*$_{85}$ transcripts was considerably less efficient than the equivalent m^{7}G-capped transcript (Fig. 2, *B* and *C*). Addition of the SL1 sequence increased translational efficiency of TMG-capped transcripts almost 7-fold to 64% of m^{7}G-capped mRNA. Interestingly, the SL1 sequence had the opposite effect on the mean efficiency of m^{7}G-based translation, reducing reporter protein accumulation to 73%. Therefore, the SL1 sequence specifically enhances expression from TMG-capped transcripts but leads to a de-

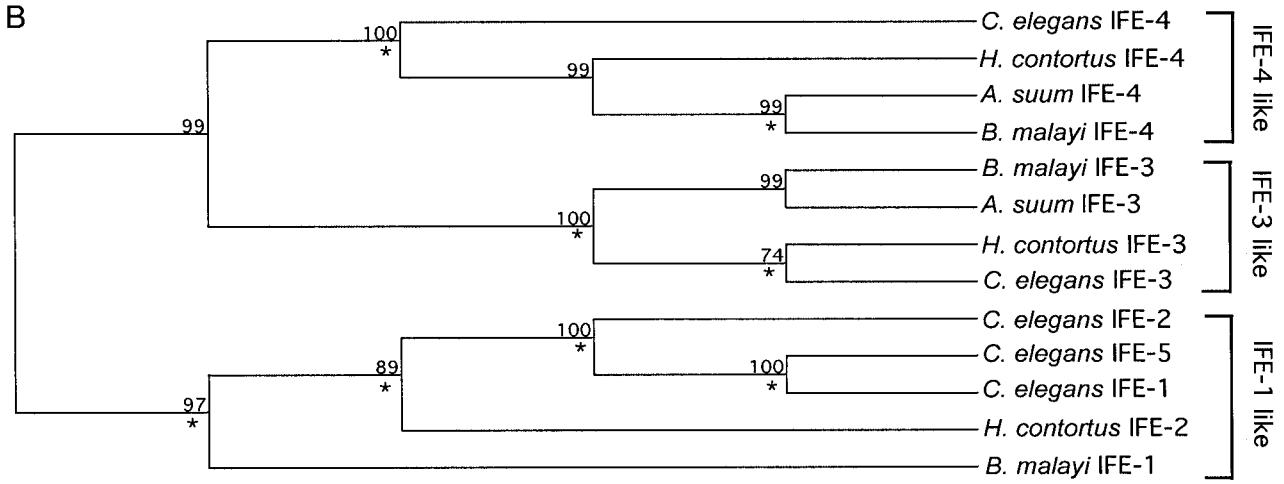


FIG. 3—*continued*

crease in translation of m^7G-capped mRNAs. By following decay of uniformly radiolabeled reporters, we observed that the effects of the $m_3^{2,2,7}$GpppG-cap and SL1 sequence on reporter production were not a function of differential transcript stability (data not shown).

We next tested whether the TMG-cap and SL1 sequence had an additive or synergistic effect on translation. We compared the translational efficiency of $m_3^{2,2,7}$GpppG- and m^7GpppG-capped transcripts with and without the SL sequence with that of ApppG-capped mRNAs. From these analyses, we observed that the SL1 sequence has a synergistic effect of 5.1 on TMG-cap based translation and 0.4 on an m^7G-capped reporter production (Fig. 2*C*). The positive effect of the SL sequence on translation of TMG-capped transcripts is independent of mRNA identity as it is consistently seen in the context of at least two different luciferase reporters, as well as with both an *Ascaris gene 12* 3′-UTR and the SV40 3′-UTR (data not shown). In addition, the synergism between cap and tail was greatest with the TMG-SL transcripts.

A. suum eIF4E-3 Binds Both Monomethylated and Trimethylated Capped RNA and Cap Analogs—To examine the basis of cap recognition on trans-spliced *versus* non-trans-spliced transcripts in nematode embryos, we cloned an *A. suum eIF4E* cDNA using a combination of degenerate PCR and RACE from 32–64 cell stage oligonucleotide-capped RNA. All of the independent clones characterized corresponded to a single *eIF4E* sequence. Amino acid alignments and phylogenetic analysis revealed that this *A. suum* eIF4E is most closely related to *C. elegans* IFE-3 (Fig. 3, supplemental Table III), and we have named it eIF4E-3. Since *A. suum* eIF4E-3 appears homologous to the eIF4E with a well defined role in translation in most eukaryotes and is the only ife isoform proven to be present in 32–64 cell embryos (also see below), we focused our attention on its functional characterization.

The ability of recombinant *A. suum* eIF4E-3 to bind m^7GTP-Sepharose 4B was examined and compared with two previously characterized eIF4Es from *C. elegans*, IFE-1 and IFE-3. *C. elegans* IFE-1 was previously shown to bind both immobilized $m_3^{2,2,7}$GTP- and m^7GTP-Sepharose 4B, whereas IFE-3 (analogous to the typical eukaryotic eIF4E) primarily only bound the latter. Recombinant *C. elegans* IFE-1 and IFE-3 and *A. suum* eIF4E-3 all bound to m^7GTP-Sepharose (Fig. 4*A*). Furthermore, the addition of 200 μM m^7GTP cap competitor decreased the amount of protein bound to m^7GTP-Sepharose and increased the amount of protein in the unbound fraction. In contrast, addition of up to 5 mM GTP to the binding buffer did not affect *A. suum* eIF4E-3 binding to the m^7GTP-Sepharose (data not shown). As expected, $m_3^{2,2,7}$GTP competitor increased the amount of *C. elegans* IFE-1 in the unbound fraction, decreasing the amount of protein bound to the m^7GTP-Sepharose 4B (Fig. 4*A*). $m_3^{2,2,7}$GTP competitor did not affect *C. elegans* IFE-3 levels in the flow-through or on the matrix, indicating that hypermethylated cap analog cannot compete with binding of this protein to the m^7GTP-Sepharose matrix (Fig. 4*A*). However, 200 μM $m_3^{2,2,7}$GTP competitor decreased the amount of *A. suum* eIF4E-3 bound to the matrix. These results demonstrated that the $m_3^{2,2,7}$GTP cap analog could effectively compete with *A. suum* eIF4E-3 binding to m^7GTP-Sepharose 4B suggesting that the protein might be capable of binding trimethylated caps. To further characterize binding in light of this result, we incubated recombinant *A. suum* eIF4E-3 with $m_3^{2,2,7}$GTP-Sepharose. We found that whereas *C. elegans* IFE-3 did not significantly bind $m_3^{2,2,7}$GTP-Sepharose, both IFE-1 and *A. suum* eIF4E-3 did bind to this matrix (Fig. 4*B*). Furthermore, *A. suum* eIF4E-3 binding to $m_3^{2,2,7}$GTP-Sepharose was competed by the addition of either m^7GTP or $m_3^{2,2,7}$GTP (Fig. 3*C*). These data indicated that *Ascaris* eIF4E-3 was capable of binding both m^7GTP or $m_3^{2,2,7}$GTP cap analogs.

Having shown that *A. suum* eIF4E-3 bound both m^7GTP- and $m_3^{2,2,7}$GTP-Sepharose, we wanted to examine its cap-binding properties in the context of an RNA and check whether such data could shed light on the translation efficiencies of different transcripts we observed in the cell-free system. We used UV cross-linking to examine the binding of *A. suum* eIF4E-3 to a uniformly labeled, capped 50-nucleotide transcript (*Ren50*) (cap-labeled RNA was also used: all data not shown). We first evaluated the cross-linking of *A. suum* eIF4E-3 to *m^7G-Ren50*, *TMG-Ren50*, uncapped-*Ren50*, *m^7G-SL-Ren50*, *TMG-SL-Ren50*, and uncapped-*SL-Ren50* substrates. We found that *A. suum* eIF4E-3 did not cross-link to uncapped substrates but cross-linked to all capped substrates (Fig. 4*D*). Cross-linking to *m^7G-Ren50* was more efficient than to *TMG-Ren50*. Interestingly, addition of the SL1 sequence to the 5′ end of m^7G-Ren50 reduced cross-linking (Fig. 4*D*). This result was seen with analysis of both a cap-labeled (*bottom panel*, Fig. 4*D*) and a uniformly labeled mRNA substrate. These data correlated with our translation assays where SL1 reduces translation from an m^7G-capped transcript. Furthermore, the presence of the SL1 sequence increased cross-linking of *A. suum* eIF4E-3 to TMG-capped substrates, once again correlating with the combined effect of the TMG cap and SL1 on cell-free translation (Fig. 2). Quantitation reveals that cross-linking of *A. suum* eIF4E-3 is greatest to *m^7G-Ren50* (arbitrarily 100%), followed by *m^7G-SL-*

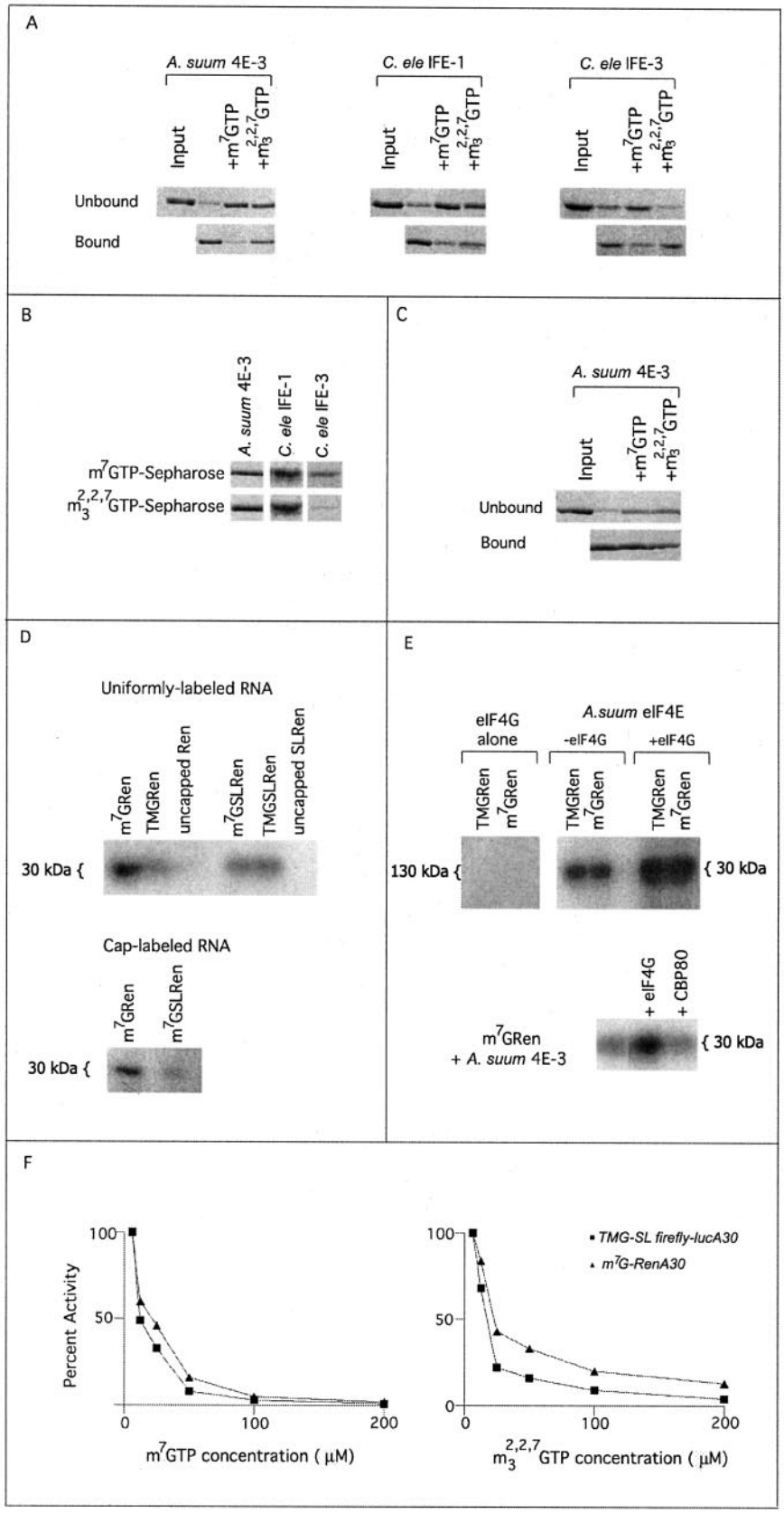


FIG. 4. ***A. suum* eIF4E binds both m^{7}G- and TMG-capped mRNA in a manner that is enhanced by eIF4G and influenced by the presence of SL1.** *A*, cap-analog binding and competition experiments demonstrate *A. suum* eIF4E binds to m^{7}GTP-Sepharose 4B and that binding is reduced in the presence of either 200 μM m^{7}GTP or m$_3^{2,2,7}$GTP. The binding and competition data for *Ascaris* eIF4E to m^{7}GTP-Sepharose 4B resemble those previously described for *C. elegans* eIFE-1 but not eIFE-3 (43). *B*, binding of recombinant eIF4E proteins to m^{7}GTP-Sepharose and m$_3^{2,2,7}$GTP-Sepharose 4B revealed that *A. suum* eIF4E-3 binds both cap analogs and thus has cap analog binding properties resembling *C. elegans* IFE-1, not IFE-3. *C*, either m^{7}GTP or m$_3^{2,2,7}$GTP can act as a competitor for *Ascaris* eIF4E-3 binding to m$_3^{2,2,7}$GTP-Sepharose 4B. Overall, the data in *A–C* demonstrate that *Ascaris* eIF4E-3 can bind both m^{7}GTP and m$_3^{2,2,7}$GTP cap analogs, and this binding can be competed against with either m^{7}GTP or m$_3^{2,2,7}$GTP. *D*, *A. suum* eIF4E cross-linking to mRNA is affected by cap identity and 5′-UTR sequence. *A. suum* eIF4E cross-linked more strongly to uniformly labeled m^{7}G- than TMG-capped *Renilla* luciferase RNA. Addition of the SL1 sequence to a TMG-capped transcript increased *A. suum* eIF4E cross-linking, while the presence of SL1 decreased the interaction with an m^{7}G-capped mRNA. The *bottom panel* demonstrates cross-linking to cap-labeled RNA substrates where the ^{32}P is located in the first phosphate of the triphosphate linkage (m^{7}G^{32}pppG-RNA). eIF4E-3 cross-linking to uniformly or cap-labeled m^{7}G-capped transcripts decreases upon addition of SL1 sequence. *E*, UV cross-linking of *A. suum* eIF4E to TMG- or m^{7}G-capped RNA is increased in the presence of eIF4G. In addition to CBP80, bovine serum albumin and His-tag-purified β-galactosidase did not increase cross-linking. Sepharose binding and cross-linking experiments were carried out multiple times with independently prepared transcripts and protein preparations. *F*, increasing levels of m^{7}GTP and m$_3^{2,2,7}$GTP competitors similarly inhibited translation of both m^{7}GpppG-*Renilla* luciferase A_{85} and TMG-SL firefly luciferase A_{85} reporter transcripts in the *A. suum* cell-free system. Equal amounts of the two reporters were added to the translation system and translation of both reporters simultaneously carried out in the presence of the indicated inhibitor. Following 30 min of translation, the translated proteins were assayed using the dual luciferase assay. Data illustrated represent 0.1% of reporter activity assayed.

Ren50 (75%), *TMG-SL-Ren50* (60%), and *TMG-Ren50* (54%). Additional experiments labeling nucleotides further 3′ in the RNA subsrate demonstrated that 90% of the eIF4E cross-linking is to the guanine at either +1 or +2 (Gppp^{+1}G*$\mathbf{p}^{+2}$G-RNA) with labeling of the protein derived from the phosphate (indicated by the boldface p) between the +1 and +2 guanines. Importantly, this nucleotide is labeled to a similar extent in the different cross-linking substrates. These and other data indicate that the different cross-linking efficiencies of eIF4E observed are a function of the cap and SL sequence and not due to

differential labeling of substrates (data not shown).

Cap binding of *A. suum* eIF4E-3 was further examined using cross-linking to determine whether recombinant *C. elegans* eIF4G (GenBank™ NM_063328) influences the nature and affinity of cap binding. eIF4G enhanced UV cross-linking (~3–4-fold) of *A. suum* eIF4E-3 to substrates with either a TMG- or an m^7GpppG-cap structure (Fig. 4*E*).

Ascaris eIF4E-3 May Translate Both m^7G- and TMG-capped mRNAs—The ability of *A. suum* eIF4E-3 to bind both m^7G- and TMG-capped substrates led us to hypothesize that *Ascaris* embryo translation of trans-spliced and non-trans-spliced mRNAs might be carried out by *A. suum* eIF4E-3 and thus involve fewer eIF4E homologs than predicted for *C. elegans*. Preliminary translation experiments directed toward immunodepletion or m^7GTP-Sepharose depletion of eIF4E from the extracts followed by reconstitution with recombinant *A. suum* eIF4E-3 were not successful. As an alternative approach, we therefore examined the effect of m^7GTP and $m_3^{2,2,7}$GTP inhibition on translation of m^7GpppG-firefly luciferase and TMG-*Renilla* luciferase mRNAs simultaneously introduced into the *A. suum* translation extract. We predicted that if *A. suum* had eIF4E homologs with different cap specificities (as in *C. elegans*), translation of the m^7G- and TMG-capped substrates would be differentially affected by the competitor nucleotides. As shown in Fig. 4*F*, m^7GTP inhibited the translation of both m^7G- and TMG-capped mRNAs in the extract with a similar inhibition curve. This is to be expected from the previous analyses of the *C. elegans* IFE proteins, which all bound m^7GTP-Sepharose. However, $m_3^{2,2,7}$GTP also inhibited translation of both m^7G- and TMG-capped transcripts. Previous $m_3^{2,2,7}$GTP-Sepharose binding studies using *C. elegans* IFE-3 and -4 suggest these proteins would not be inhibited by $m_3^{2,2,7}$GTP (43). Together with the binding and cross-linking studies, our evidence therefore indicates that a single protein, *A. suum* eIF4E-3, might be fully capable of translating both trans-spliced and non-trans-spliced transcripts.

A. suum and Brugia malayi May Have Fewer eIF4E Isoforms than C. elegans—As *C. elegans* has five ife isoforms, we searched the *Ascaris* EST data base (~40,000 ESTs from a variety of different stages and tissues) for other *A. suum* eIF4E cDNAs. Only one other eIF4E was identified and is present in a gut-derived cDNA library (*e.g.* GenBank™ BQ380672). Both multiple sequence alignment and phylogenetic analyses show that *A. suum* eIF4E-4 has closest identity to *C. elegans* IFE-4 at the amino acid level (Fig. 3, supplemental Table III). The whole genome sequence (5-fold coverage) of the filarial nematode (www.tigr.org/tdb/e2k1/bma1) contains three eIF4E corresponding by phylogenetic analysis to a *C. elegans ife-3*, *ife-4*, and one form of the *ife-1*, *-2*, and *-5* group (Fig. 3). A similar set of *ife* genes is found in another nematode, *H. contortus*. Thus, we have identified *eIF4E* genes resembling IFE-3 and -4 in all of the nematodes examined but have been unable to identify multiple copies of genes related to *ife-1*, *-2*, and *-5* in nematodes outside of the *Caenorhabditis* group. This suggests that fewer eIF4E isoforms might be present in nematodes other than *C. elegans*.

The SL Sequence Trans-splices Close to the Start Codon of Transcripts—It has previously been observed that the SL1 sequence tends to splice close to the start codon (3), a phenomenon that will affect the sequence and length of the 5′-UTR of trans-spliced transcripts. As previous analysis had not examined SL1 and SL2 sequences separately, we also addressed whether both SL1 and SL2 would be found close to the start codon. We identified a set of *C. elegans* cDNAs in the data bases that contain an SL1 or SL2 sequence and noted the SL to AUG distance (as described under "Experimental Procedures"). We found that both SL1 and SL2 do indeed tend to splice close to the start codon, with 49% of transcripts containing a spliced leader sequence within 10 nucleotides of the initiator AUG (Fig. 5*A*). Even within the set of cDNAs where the SL1 to AUG distance is less than 10 nucleotides, there is a tendency for the spliced leader to lie close to the start codon (Fig. 5*B*). When we extended this analysis to *Ascaris* and other nematode species, we found that the SL1 sequence trans-splices close to the AUG in nematodes other than *C. elegans*. The sample size is somewhat small as many nematode ESTs do not contain the entire spliced leader sequence. However, in species that phylogenetically group with *A. suum* (Clade III nematodes), at least 45% of cDNAs contain an SL1 sequence within 10 nucleotides of the start codon (Fig. 5*C* and see supplemental Fig. 1).

Our cross-species analysis focused on the SL1 sequence. The presence and importance of alternative spliced leaders, such as SL2, in other nematode groups is currently unclear, having been identified in only a few nematodes (8, 54–56). To determine whether a second spliced leader was easily identifiable on mRNAs derived from 32–64 cell stage *A. suum* embryos, we generated an oligonucleotide-capped cDNA library to maximize representation of full-length clones. Analysis of the extreme terminal 5′ sequence of ~1500 randomly selected clones (corresponding to 200 unique cDNAs) indicated that 70% of *A. suum* transcripts are SL1 spliced, a number that is similar to, although slightly less than, previous analysis of trans-splicing levels in *A. suum* (14). Furthermore, despite sequencing the extreme 5′ ends of at least 200 unique cDNAs, we did not find a second spliced leader sequence. This suggests SL1 may be the only spliced leader sequence in *A. suum*.

The SL1 Sequence Has an Optimal Distance for Translation from the Start Codon—To test the hypothesis that the position of SL1 trans-splicing affects translation, we synthesized the transcripts illustrated in Fig. 6*A* containing an increasing number of random bases in the 5′-UTR (Ns, where N = A, U, G, or C; derived from degenerate oligonucleotide primers between the spliced leader and the start AUG) (see supplemental Table II). It has previously been shown that the three nucleotides upstream of the start codon, particularly the −3 position, affect mammalian translation (31). To eliminate the effect of randomizing the three nucleotides at the −3 to −1 positions, we initially fixed these as ACC (nucleotides that are found at these positions when nematode transcripts are surveyed, as well as the Kozak consensus for these positions). To facilitate normalization of our translation assays, an m^7GpppG-capped firefly luciferase mRNA was included as an internal control with the *Renilla* luciferase test transcript. The translation of both mRNAs was measured using the dual luciferase assay system and luminescence from the latter reporter normalized relative to the former. When the distance between the SL1 sequence and the start codon was increased by three nucleotides (sequence ACC), translational efficiency increased (up to 7.5-fold, Fig. 6*B*, *panel 1*, *first two bars*). When seven additional nucleotides were added to the ACC (increasing SL to AUG distance to 10 bases), there was a drop in translational efficiency (Fig. 6*B*). As SL to AUG distance was further increased, reporter activity remained relatively low. This may explain the tendency of SL1 to splice close to the start codon described above (Fig. 5, *A–C*).

Placing three nucleotides between the SL and start codon increased reporter activity. This phenomenon either reflected optimal spacing or was due to the specific identity of the intervening bases. Indeed, in a survey of 49 *A. suum* cDNAs, 55% contained an A at the −3 position, 26% a G, and 8% had a U (analysis at www.changbioscience.com/). The bias toward an A at the −3 position was also evident in trans-spliced transcripts

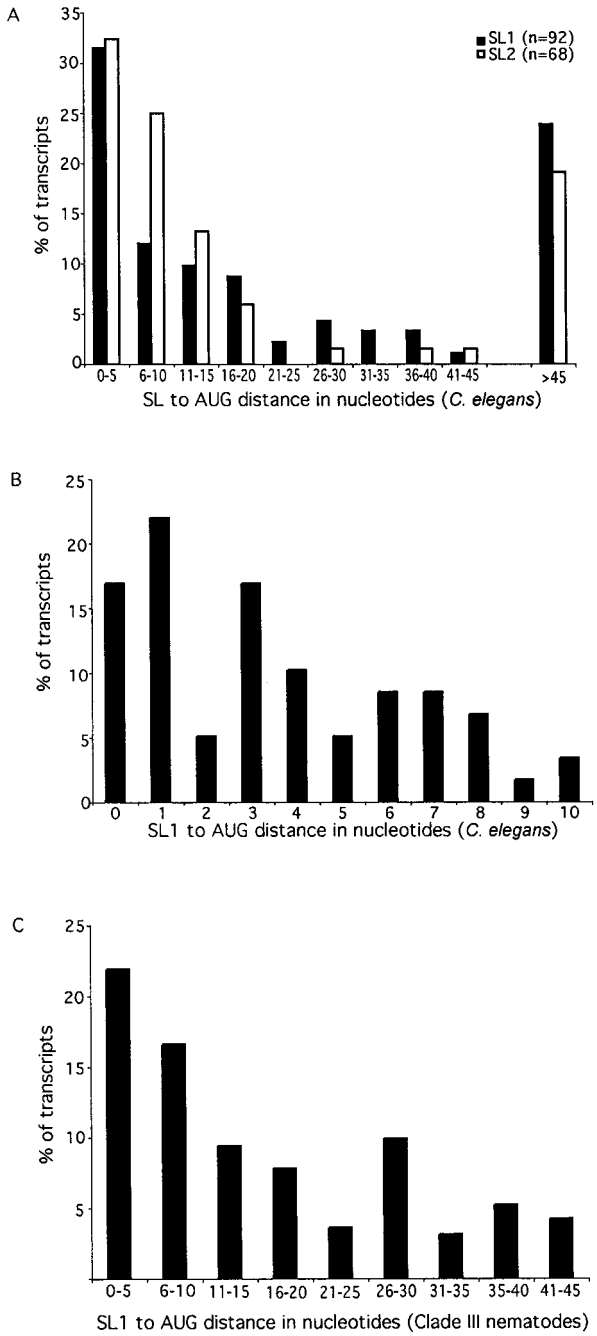


FIG. 5. **SL sequences tend to trans-splice close to the start codon in nematode transcripts.** *A*, a survey of SL1 and SL2 trans-spliced *C. elegans* transcripts revealed that trans-splicing tends to occur close to the start AUG. *B*, examination of 59 *C. elegans* cDNAs with less than 10 nucleotides separating the SL1 sequence and start codon reveals that greater than 70% of the transcripts have the SL1 sequence within four nucleotides of the predicted start AUG. *C*, the SL1 sequence also tends to trans-splice close to the start codon in 182 mRNAs from nematodes including *Ascaris* and five species that are phylogenetically related to *A. suum* (so-called clade III nematodes including *Dirofilaria immitis*, *Onchocerca volvulus*, *Litomosoides sigmodontis*, *B. malayi*, and *Toxocara canis*). Data for each nematode included in the compilation are illustrated in supplementary Fig. 1.

(61% had an A as the −3 nucleotide) examined from databases. These data suggest that there is a bias in *A. suum* toward having a purine, and in particular an A, at the −3 position relative to the start codon. We therefore compared placing an A, G, or a pyrimidine (U) at the −3 position relative to the start codon on translation while randomizing the identity of the nucleotides at −2 and −1 (constructs in Fig. 6*C*). We found that an A at the −3 position had the greatest translational efficiency in the cell-free system (arbitrarily 100%), with G less efficient, and U the least efficient (Fig. 6*C*). Transcripts containing an SL1 sequence immediately adjacent to the start codon have a GAG at the −3 to −1 positions. However, the effect of having an A rather than a G at the −3 position does not fully account for increased reporter activity seen when ACC is inserted between the SL1 sequence and the start AUG (an ~2-fold increase in Fig. 6*C*, which does not account for the up to 7.5-fold increase in the *first two bars* in Fig. 6*B*, *panel 1*). This suggests that the increase between *TMG-SL-Ren-A*$_{85}$ and *TMG-SL-ACC-Ren-*$_{85}$ is in part due to having an A rather than a G at the −3 position but also reflects an optimal distance for SL1 from the start codon.

To test whether this effect was due to the SL1 sequence itself or the increasing length of the 5′-UTR in these experiments, we replaced TMG-SL1 with m^{7}G-N$_{21}$. The optimal 5′-UTR length for SL1 containing mRNAs is distinct from that observed for *m^{7}G-ACC-Ren*, where the optimal total length is 31 nucleotides (Fig. 6*B*, *panel 2*; see also supplemental Fig. 2). Thus the drop in reporter activity with increasing SL1 to AUG distance was not due to increasing 5′-UTR length (or the presence of randomized nucleotide sequence) but was related to either the SL1 sequence or TMG-cap. Note also that the optimal TMG-cap to start codon distance is greater when SL1 is replaced with random nucleotides (Fig. 6*B*, *panel 3*). Consequently, changes in translation efficiency were specifically due to varying distance between the SL sequence and the start codon and were not simply a function of 5′-UTR length or cap identity. The tendency of SL to trans-splice close to the start codon may therefore reflect evolutionary selection for optimal expression (see "Discussion").

DISCUSSION

Insights into Nematode Translation and the Role of the Spliced Leader Sequence—We have described a nematode *in vitro* translation system that displays attributes of *in vivo* translation. Translation in the *A. suum* cell-free extract is cap-dependent, increases with poly(A)-tail length, and shows a synergistic interaction between the cap and poly(A)-tail. These and other features of the cell-free system reflect our *in vivo* translation data obtained following biolistic introduction of the same reporter transcripts in *Ascaris* embryos.[3] Start codon context and 5′-UTR length were shown to contribute to translation efficiency. In particular, an adenine at the −3 position in the transcripts analyzed leads to higher translation than either G or U at this position. Aspects of translation defined in other metazoa are conserved in nematodes including the contribution of start codon context and 5′-UTR length and cap and poly(A)-tail synergism on translation efficiency. The synergy between the TMG-cap and poly(A)-tail indicates that the same interactions occur in trans-spliced transcripts. Notably, the synergism between the TMG-SL and poly(A)-tail was greater than that observed for the m^{7}G-cap and poly(A)-tail. Cross-linking of eIF4E to m^{7}G- and TMG-capped RNAs was also enhanced by the presence of eIF4G, further demonstrating that interactions between translation initiation factors are conserved in nematodes.

While the presence of either the spliced leader or a trimethylguanosine cap alone to a test transcript leads to a significant drop in translation efficiency, together these features synergistically enhance translation. Previous studies also observed a collaborative effect between the TMG-cap and SL sequence and a detrimental effect of either the TMG cap or SL sequence alone on an endogenous trans-spliced mRNA (14). Our experiments have extended these studies using an optimal 5′-UTR length and by examining multiple reporters to eliminate sequence specific effects on translation by the body or 3′-UTR of the

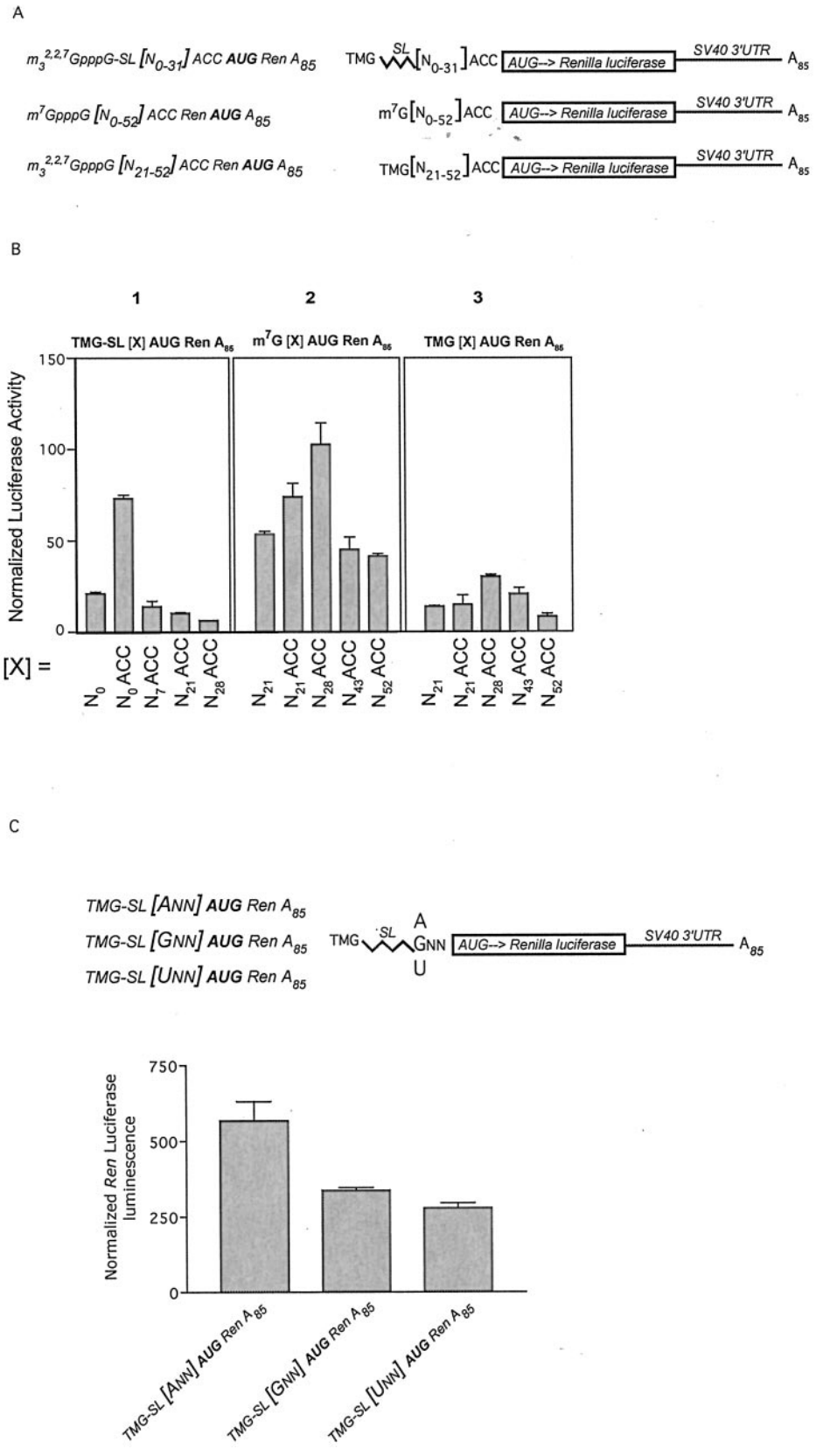


FIG. 6. **Analysis of SL1 sequence to start codon distance indicates there is an optimal SL1 to start codon distance for translation of reporter transcripts.** *A*, transcripts used to analyze the effect of SL1 to AUG distance on *in vitro* translation in *A. suum* embryonic extracts. In a series of test transcripts, SL1 sequence was separated from the AUG by an increasing number of random bases derived from degenerate oligonucleotides synthesized with equal proportions of each base at a position (see "Results" and supplemental Table II). The effect of 5′-UTR length and cap identity was also tested, using a series of constructs with an increasing number of random bases between the cap and the start codon (see supplemental Fig. 2). (B) The placement of "ACC" between the SL1 sequence and the start AUG enhanced translation (*two left bars*, *panel 1*), an effect that diminished in transcripts as the SL1 to AUG distance was increased (*left group* of *bars*). *Renilla* luciferase activity was normalized to an internal control (m^7GpppG-capped firefly luciferase reporter). The drop in translational efficiency with distance was not simply due to increasing 5′-UTR length or the TMG cap to AUG distance, as translation was enhanced with 5′-UTR length (up to a certain distance) in the absence of SL1 sequence (*panels 2* and *3*). Although the relative differences between reporter activity vary from experiment to experiment, the illustrated trend was consistently seen. *C*, the effect of the −3 residue on translation efficiency. Constructs used to test the effect of −3 residue identity relative to the start codon are illustrated and the translation efficiencies of the three constructs shown. The presence of an A at the −3 position enhanced reporter production over a G or U in the nematode *in vitro* translation extract. All data shown ($n = 2$) are representative of results from experiments repeated at least six times with independently synthesized mRNAs and prepared translation extracts.

transcript. In contrast to Maroney *et al.* (14), we find that non-trans-spliced transcripts are more efficiently translated than trans-spliced mRNAs, an effect consistent with our *in vivo* observations.[3] We attribute the differences observed to 1) extract preparation, 2) potential contributions of the endogenous *Ascaris* mRNA sequences to translation efficiency in the earlier study, and 3) differences in the 5′-UTR sequences. In addition, our translation assays measured luciferase activity derived through addition of an exogenous mRNA that must compete for translation with all endogenous mRNAs in the extract preparation. In contrast, previous studies removed endogenous RNAs by micrococcal nuclease treatment and measured

[^{35}S]methionine incorporation into protein.

We have further found that the optimal 5′-UTR length for trans-spliced test transcripts is ~25 nucleotides (including the SL) and ~31 nucleotides for non-trans-spliced test transcripts. Our analyses (see Fig. 5 and supplemental Fig. 1) indicate that transcripts with a non-optimal SL to AUG distance are definitely present *in vivo*. This is not unexpected as factors affecting translation of each mRNA are complex. Indeed, mRNAs are unlikely to be translated at similar levels but their expression finely tuned to produce the required levels of protein by a variety of intrinsic and extrinsic factors.

The SL1 sequence is remarkably conserved across nematode species. This sequence conservation may reflect several roles for the SL1 sequence. For example, the SL sequence has been shown to be a promoter element for SL RNA transcription (20). Although lethality caused by a deletion of the SL1 donor genes can be rescued by introduction of SL sequences that are substantially changed or even extended in length (57, 58), more directed and subtle assays show that mutations in the SL1 sequence can significantly alter the translatability of a trans-spliced mRNA (14). These SL1 mutations presumably affect the ability of the spliced leader to synergistically collaborate with the TMG-cap to promote translation. Overall, these and our data suggest that the SL1 sequence itself influences translation.

Bioinformatics data indicate that the SL sequence tends to trans-splice close to the start codon in a diversity of nematode species and that this evolutionary conservation is functionally reflected in general with the optimal SL to AUG distance for reporter mRNA translation in the cell-free system. Our data also show a correlation between translational efficiency of TMG-SL1 containing mRNAs and cross-linking levels to *A. suum* eIF4E-3. This suggests that the TMG-cap and SL1 together might influence interactions between the transcript and the translation initiation complex, perhaps reflecting specific requirements for translation initiation of trans-spliced transcripts. This model is consistent with recent evidence that the spliced leader, as well as the so-called cap 4 methylation status, can affect association of transcripts with polysomes in the kinetoplastid *Leishmania tarentolae* (59).

The SL1 sequence has a specific effect on optimal 5′-UTR length when compared with random sequence. Thus, addition of the spliced leader sequence may also provide a relatively short, uniform 5′-UTR that is free of small upstream open reading frames and optimal for ribosomal scanning. This model might predict that trans-spliced nematode primary transcripts may not be optimal for translation. We have attempted to examine this question by identifying the site of transcription initiation for several trans-spliced *A. suum* genes to define and examine the region upstream of the AUG in primary transcripts. This has proven difficult as multiple techniques fail to consistently identify a discrete transcription start site for two genes (*v-ATPase* and the ribosomal gene *L23a*, data not shown), although these assays can define a transcription start site for a non-trans-spliced gene (glutathione *S*-transferase). Notably, transcription initiation sites for trans-spliced *C. elegans* genes also have not been definitively characterized. Our inability to discretely define transcription initiation sites could be consistent with the hypothesis that transcription initiation may be heterogeneous for such genes (17). Trans-splicing might then function to trim these heterogeneous primary transcripts, generating uniform 5′-UTRs that are more optimal for translation.

Overall our data suggest that the translation machinery in nematodes may have evolved to translate trans-spliced mRNAs with 5′-UTRs of a particular length containing both the spliced leader and its TMG-cap. Therefore, trans-splicing of the SL1 sequence may serve at least two functions in nematodes, generation of an optimal 5′-UTR length as well as the sequence context for translation of trimethylguanosine capped transcripts.

Translation Initiation and the Evolution of the eIF4E Genes in Trans-splicing Organisms—We have cloned *eIF4E-3* from *A. suum* and our cross-linking studies indicate that its cap-binding characteristics may explain differential activity of TMG-*versus* m^{7}GpppG-capped reporters in the extract, as well as the effect of the SL1 sequence on translation. *A. suum* eIF4E-3 can bind both immobilized m^{7}GTP- and m$_3^{2,2,7}$GTP-Sepharose, as well as m^{7}G and TMG-capped mRNAs. Our results further indicate that while *A. suum* eIF4E-3 is most closely related to *C. elegans* IFE-3 in sequence, its binding properties resemble those of *C. elegans* IFE-1, -2, or -5. Interestingly, *A. suum* eIF4E-3 binds TMG-cap even though residues previously postulated as being critical to *C. elegans* IFE-3 m^{7}GTP-selective cap binding are conserved in the two proteins. In particular, mutational analysis of *C. elegans* IFE-5 (^{64}NV^{65} → YL; see equivalent residues 60–61 *boxed* in Fig. 4) increased IFE-5 m^{7}GTP binding, indicating that these residues play a role in cap-binding selectivity (42). *C. elegans* eIF4E isoforms with YL residues at this position were reported to be less efficient at binding m$_3^{2,2,7}$GTP. *C. elegans* IFE-3, which only binds m^{7}GTP, has the YL residues. However, *A. suum* eIF4E-3 has residues YL at the equivalent position but binds both m^{7}GTP- and m$_3^{2,2,7}$GTP-Sepharose and cross-links to the corresponding RNAs. This indicates the YL residues may not be as important for selective binding to m^{7}GTP in the *Ascaris* eIF4E-3, and additional residues that have yet to be defined may contribute to cap-binding specificity.

Regions of *H. sapiens* eIF4E are relatively unstructured until nucleotide binding (60, 61). In addition, both eIF4E and eIF4G undergo conformational changes on interaction with each other and on nucleotide binding (40). Steric flexibility and conformational changes might mean that although *A. suum* eIF4E-3 has a similar sequence to *C. elegans* IFE-3, it can still accommodate a trimethylated cap nucleotide. The structure of the *A. suum* eIF4E-3 binding pocket containing a nucleotide is therefore predicted to differ in its dimensions from that defined for the human protein and predicted for *C. elegans* IFE-3 (42). Indeed preliminary structural homology modeling indicates that the dimensions of the *A. suum* eIF4E-3 might be larger than those of its closest sequence homolog IFE-3, a prediction that could explain why the *A. suum* protein can accommodate the TMG-cap while *C. elegans* IFE-3 cannot.[2]

Our characterization of *A. suum* IFE-3 suggests that additional eIF4E homologs such as those observed in *C. elegans* (*e.g. ife-1*, *-2*, and *-5*) might not be necessary for TMG-cap binding of transcripts in *Ascaris* embryos. Analysis of the sequenced genomes of two other trans-splicing metazoa, *C. intestinalis* and *Schistosoma mansoni*, suggests that they do not have a distinct *ife-1*, *-2*, and *-5* class homolog based on molecular phylogenetic analyses. Therefore, current data indicate there might not be an ancestral *eIF4E* homolog specifically shared by trans-splicing organisms, as proposed previously (16). In addition, it has become clear that many non-trans-splicing organisms, including *H. sapiens*, have multiple *eIF4E* isoforms. For example, spliced leader trans-splicing has not been detected in *Drosophila melanogaster*, yet this organism has 6 *eIF4E*-like genes (62).

Trans-splicing and Polycistronic Transcripts—While operons have not yet been shown to exist in *A. suum*, the apparent absence of an SL2 sequence suggests that if polycistronic transcripts do arise they may be resolved by SL1 trans-splicing. Indeed SL1 has been shown to resolve polycistronic transcripts in *C. elegans* (63). The genome sequence of *B. malayi* is currently available (www.tigr.org/tdb/e2k1/bma1/) and will pro-

vide information on whether additional spliced leaders are present in this nematode and whether gene organization suggests that polycistronic transcripts are likely to be present. However, the origins of SL trans-splicing and its ancestral role (in resolving polycistronic transcripts *versus* promoting translation) within the nematodes will likely only be revealed if basal species (those potentially closer to the common ancestor of nematodes) are examined. In addition, functional analysis of trans-splicing in flatworms (*S. mansoni*), *H. vulgaris*, and *C. intestinalis* will shed light on whether trans-splicing arose in a common ancestor or has independently evolved in multiple lineages. Such issues will eventually be resolved by identifying the full phylogenetic extent of trans-splicing, coupled with functional analysis in multiple species.

Acknowledgments—We acknowledge Leah Cohen, Claudette Mikhli, David Ndegwa, Yubin Zhang, Scott Hodgson, Eugenia Ribeiro, and Angelica Parra for contributions to this work and excellent technical assistance. We also thank Dr. Timothy Nilsen's laboratory for advice on preparation of extracts and for providing trans-spliced and non-trans-spliced genes and their sequences. We thank Paul Copeland and members of the Davis laboratory for their comments on the manuscript and Frank Burbrink for help with phylogenetic analysis.

REFERENCES

1. Nilsen, T. W. (2001) *Trends Genet.* **17,** 678–680
2. Zorio, D. A., Cheng, N. N., Blumenthal, T., and Spieth, J. (1994) *Nature* **372,** 270–272
3. Blumenthal, T., and Steward, K. (1997) in *C. elegans II* (Riddle, D. L., Blumenthal, T., Meyers, B. J., and Priess, J. R., eds) pp. 117–145, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
4. Maroney, P. A., Hannon, G. J., and Nilsen, T. W. (1990) *Proc. Natl. Acad. Sci. U. S. A.* **87,** 709–713
5. Thomas, J., Conrad, R. C., and Blumenthal, T. (1988) *Cell* **54,** 533–539
6. Liou, R.-F., and Blumenthal, T. (1990) *Mol. Cell. Biol.* **10,** 1764–1768
7. Van Doren, K., and Hirsh, D. (1990) *Mol. Cell. Biol.* **10,** 1769–1772
8. Huang, X. Y., and Hirsh, D. (1989) *Proc. Natl. Acad. Sci. U. S. A.* **86,** 8640–8644
9. Spieth, J., Brooke, G., Kuersten, S., Lea, K., and Blumenthal, T. (1993) *Cell* **73,** 5121–5132
10. Blumenthal, T., and Gleason, K. S. (2003) *Nat. Rev. Genet* **4,** 110–118
11. Ullu, E., Tschudi, C., and Günzl, A. (1996) in *Molecular Biology of Parasitic Protozoa* (Smith, D. F., and Parsons, M., eds) pp. 115–133, Oxford University Press, Oxford, UK
12. Davis, R. E., and Hodgson, S. (1997) *Mol. Biochem. Parasitol.* **89,** 25–39
13. Davis, R. E., Hardwick, C., Tavernier, P., Hodgson, S., and Singh, H. (1995) *J. Biol. Chem.* **270,** 21813–21819
14. Maroney, P. A., Denker, J. A., Darzynkiewicz, E., Laneve, R., and Nilsen, T. W. (1995) *RNA* (*N. Y.*) **1,** 715–723
15. Vandenberghe, A. E., Meedel, T. H., and Hastings, K. E. M. (2001) *Genes Dev.* **15,** 294–303
16. Stover, N. A., and Steele, R. E. (2001) *Proc. Natl. Acad. Sci. U. S. A.* **98,** 5693–5698
17. Davis, R. E. (1996) *Parasitol. Today* **12,** 33–40
18. Maroney, P. A., Hannon, G. J., and Nilsen, T. W. (1989) *Mol. Biochem. Parasitol.* **35,** 277–284
19. Hannon, G. J., Maroney, P. A., Denker, J. A., and Nilsen, T. W. (1990) *Cell* **61,** 1247–1255
20. Hannon, G. J., Maroney, P. A., Ayers, D. G., Shambaugh, J. D., and Nilsen, T. W. (1990) *EMBO J.* **9,** 1915–1921
21. Denker, J. A., Zuckerman, D. M., Maroney, P. A., and Nilsen, T. W. (2002) *Nature* **417,** 667–670
22. Denker, J. A., Maroney, P. A., Yu, Y. T., Kanost, R. A., and Nilsen, T. W. (1996) *RNA* (*N. Y.*) **2,** 746–755
23. Gallie, D. R. (1991) *Genes Dev.* **5,** 2108–2116
24. Sachs, A. B., Sarnow, P., and Hentze, M. W. (1997) *Cell* **89,** 831–838
25. Gebauer, F., Corona, D. F. V., Preiss, T., Becker, P. B., and Hentze, M. W. (1999) *EMBO J.* **21,** 6146–6154
26. Svitkin, Y. V., and Sonenberg, N. (2004) *Methods Mol. Biol.* **257,** 155–170
27. Michel, Y. M., Poncet, D., Piron, M., Kean, K. M., and Borman, A. M. (2000) *J. Biol. Chem.* **275,** 32268–32276
28. Thoma, C., Ostareck-Lederer, A., and Hentze, M. W. (2004) *Methods Mol. Biol.* **257,** 171–180
29. Iizuka, N., Najita, L., Franzusoff, A., and Sarnow, P. (1994) *Mol. Cell. Biol.* **14,** 7322–7330
30. Kozak, M. (1978) *Cell* **15,** 1109–1123
31. Kozak, M. (1981) *Nucleic Acids Res.* **12,** 857–872
32. Slusher, L. B., Gillman, E. C., Martin, N. C., and Hopper, A. K. (1991) *Proc. Natl. Acad. Sci. U. S. A.* **88,** 9789–9793
33. Peri, S., and Pandey, A. (2001) *Trends Genet.* **17,** 685–687
34. Gingras, A.-C., Raught, B., and Sonenberg, N. (1999) *Annu. Rev. Biochem.* **68,** 913–963
35. Von Der Haar, T., Gross, J. D., Wagner, G., and McCarthy, J. E. (2004) *Nat. Struct. Mol. Biol.* **11,** 503–511
36. Lamphear, B. J., Kirchweger, R., Skern, T., and Rhoads, R. E. (1995) *J. Biol. Chem.* **270,** 21975–21983
37. Mader, S., Lee, H., Pause, A., and Sonenberg, N. (1995) *Mol. Cell. Biol.* **15,** 4990–4997
38. Imataka, H., Gradi, A., and Sonenberg, N. (1998) *EMBO J.* **17,** 7480–7489
39. Haghighat, A., and Sonenberg, N. (1997) *J. Biol. Chem.* **272,** 21677–21680
40. Gross, J. D., Moerke, N. J., von der Haar, T., Lugovsky, A. A., Sachs, A. B., McCarthy, J. E. G., and Wagner, G. (2003) *Cell* **115,** 739–750
41. Jankowska-Anyszka, M., Lamphear, B. J., Aamodt, E. J., Harrington, T., Darzynkiewicz, E., Stolarski, R., and Rhoads, R. E. (1998) *J. Biol. Chem.* **273,** 10538–10542
42. Miyoshi, H., Dwyer, D. S., Keiper, B. D., Jankowska-Anyszka, M., Darzynkiewicz, E., and Rhoads, R. E. (2002) *EMBO J.* **17,** 4680–4690
43. Keiper, B. D., Lamphear, B. J., Deshpande, A. M., Jankowska-Anyszka, M., Aamodt, E. J., Blumenthal, T., and Rhoads, R. E. (2000) *J. Biol. Chem.* **275,** 10590–10596
44. Darzynkiewicz, E., Stepinski, J., Ekiel, I., Jin, Y., Haber, D., Sijuwade, T., and Tahara, S. M. (1988) *Nucleic Acids Res.* **16,** 8953–8962
45. Pasquinelli, A. E., Dahlberg, J. E., and Lund, E. (1995) *RNA* (*N. Y.*) **1,** 957–967
46. Cohen, L. S., Mikhli, C., Friedmann, C., Jankowska-Anyszka, M., Stepinski, J., Darzynkiewicz, E., and Davis, R. E. (2004) *RNA* (*N.Y.*) **10,** 1609–1624
47. Fromont-Racine, M., Bertrand, E., Pictet, R., and Grange, T. (1993) *Nucleic Acids Res.* **21,** 1683–1684
48. Sallie, R. (1993) *PCR Methods Applications* **3,** 54–56
49. Marayuma, K., and Sugano, S. (1994) *Gene* (*Amst.*) **138,** 171–174
50. Volloch, V., Schweitzer, B., and Rits, S. (1994) *Nucleic Acids Res.* **22,** 2507–2511
51. Swofford, D. L., Olsen, G. J., Waddell, P. J., and Hillis, D. M. (1996) in *Molecular Systematics* (Hillis, D. M., Moritz, C., and Mable, B. K., eds) Sinauer, Sunderland, MA
52. Nei, M., and Kumar, S. (2000) *Molecular Evaluation and Phylogenetics*, Oxford University Press, New York
53. Huelsenbeck, J. P., Ronquist, F., Nielsen, R., and Bollback, J. P. (2001) *Science* **294,** 2310–2314
54. Lee, K. Z., and Sommer, R. J. (2003) *Mol. Biol. Evol.* **20,** 2097–2103
55. Redmond, D. L., and Knox, D. P. (2001) *Mol. Biochem. Parasitol.* **117,** 107–110
56. Evans, D., Zorio, D., MacMorris, M., Winter, C. E., Lea, K., and Blumenthal, T. (1997) *Proc. Natl. Acad. Sci. U. S. A.* **94,** 9751–9756
57. Xie, H., and Hirsh, D. (1998) *Proc. Natl. Acad. Sci. U. S. A.* **95,** 4235–4240
58. Ferguson, K. C., and Rothman, J. H. (1999) *Mol. Cell. Biol.* **19,** 1892–1900
59. Zeiner, G. M., Sturm, N. R., and Campbell, D. A. (2003) *J. Biol. Chem.* **278,** 38269–38275
60. Tomoo, K., Shen, X., Okabe, K., Nozoe, Y., Fukuhara, S., Morino, S., Ishida, T., Taniguchi, T., Hasegawa, H., Terashima, A., Sasaki, M., Katsuya, Y., Kitamura, K., Miyoshi, H., Ishikawa, M., and Miura, K. (2002) *Biochem. J.* **362,** 539–544
61. Tomoo, K., Shen, X., Okabe, K., Nozoe, Y., Fukuhara, S., Morino, S., Sasaki, M., Taniguchi, T., Miyagawa, H., Kitamura, K., Miura, K., and Ishida, T. (2003) *J. Mol. Biol.* **328,** 365–383
62. Lasko, P. (2000) *J. Cell Biol.* **150,** F51–F56
63. Hengartner, M. O., and Horvitz, H. R. (1994) *Cell* **76,** 665–676

Supplementary Table 1. Primers.

	Primers
T7-SL-Renilla luciferase	**5′ Primer =** ATCGAAATTAATACGACTCACTATAGGTTTAATTAATTACCCAAGTTTGAGGGCTAGCCACCATGACTT **3′ Primer =** $T_{0,30,85}$ CCAAACTCATCAATGTATCTTA
T7-SL-N-Renilla luciferase **T7-SL-N-ACC-Renilla luciferase**	**5′ Primer =** ATCGAAATTAATACGACTCACTATAGGTTTAATTAATTACCCAAGTTTGAGN$_n$ATGACTTCGAAAGTTTATG **5′ Primer =** ATCGAAATTAATACGACTCACTATAGGTTTAATTAATTACCCAAGTTTGAGN$_n$ACCATGACTTCGAAAGTTTATG
T7-Renilla luciferase	**5′ Primer =** ATCGAAATTAATACGACTCACTATAGGTCATTCCGGTACTGTTGGTATGCTAGCCACCATGACTTC **3′ Primer =** $T_{0,30,85}$ CCAAACTCATCAATGTATCTTA
T7-N-Renilla luciferase **T7-N-ACC-Renilla luciferase**	**5′ Primer =** ATCGAAATTAATACGACTCACTATAGN$_n$ATGACTTCGAAAGTTTATG **5′ Primer =** ATCGAAATTAATACGACTCACTATAGN$_n$ACCATGACTTCGAAAGTTTATG
T7-Firefly luciferase	**5′ Primer =** ATCGAAATTAATACGACTCACTATAGGGCATTCCGGTACTGTTGGTAAAGCCACCATGGAAGA **3′ Primer =** $T_{0,30,85}$ TGAATGCAATTGTTGTTGTT
T7-Ren50	**5′ Primer =** TTAATACGACTCACTATAGGTCATTCCGGTACTGTTGGTATGCTAGCCACCATGACTTC **5′ Primer =** AATACACCGCGCTACTG
T7 SL-Ren50	**5′ Primer =** ATCGAAATTAATACGACTCACTATAGGTTTAATTAATTACCCAAGTTTGAGGGCTAGCCACCATGACTTC **3′ Primer =** AATACACCGCGCTACTG
Degenerate primer for eIF4E	**3′ Primer =** RAAYTGYTCICCIAYIADIGCCAT
A.suum eIF4E-3 3′RACE	**5′ Primer =** TCAATTCCATATGACGATGGCTGACGG **3′ Primer =** AGCGAGGATCCGTCGACTCTAGATTTTTTTTTTTTT
C. elegans ife-1	**5′ Primer =** GCGAATTCCATATGTCAGATTCTGAAATAGCATTTG **3′ Primer =** CGCGGATCCTTAGACGGCGGATTTCTCGG
C. elegans ife-3	**5′ Primer =** GCGAATTCCTCGAGATGAGCACATCCGTAGCGGAAA **3′ Primer =** CGCATTCTCGAGTTAAGGAGTTGGGGTGGCTGGAGA
RNA for oligo-capped cDNA library construction	**5′ Primer =** AUAGGACUAACUACACUGAGGAUCCAAA

Supplementary Table 2. RNA Sequences of 5' UTRs and Cross-linking RNAs.

5' UTRs

Nontrans-spliced test RNA	
Uncapped =	pppGGTCATTCCGGTACTGTTGGTATGCTAGCCACC**ATG**..Renilla..ORF→Stop
ApppG capped =	ApppG……..……**ATG**..Renilla..ORF→Stop
GpppG capped =	GpppG……..……**ATG**..Renilla..ORF→Stop
m^7G -capped =	m^7GpppG……..……**ATG**..Renilla..ORF→Stop
TMG-capped =	$m_3^{2,2,7}$GpppG……..……**ATG**..Renilla..ORF→Stop
Trans-spliced test RNA RNA	
Uncapped =	pppGGTTTAATTACCCAAGTTTGAGGGCTAGCCACC**ATG**..Renilla..ORF→Stop
ApppG capped =	ApppG……..……..**ATG**..Renilla..ORF→Stop
m^7G -capped =	m^7GpppG……..……..**ATG**..Renilla..ORF→Stop
TMG-capped =	$m_3^{2,2,7}$GpppG……..……..**ATG**..Renilla..ORF→Stop

5' UTRs with random nucleotides

Without fixed –1 to –3	*Cap dinucleotide-N_n-**ATG**..Renilla..ORF→Stop
	Cap dinucleotide-SL-N_n-**ATG**..Renilla..ORF→Stop
With ACC fixed at –1 to –3	Cap dinucleotide-N_n-ACC-**ATG**..Renilla..ORF→Stop
	Cap dinucleotide-SL-N_n-ACC-**ATG**..Renilla..ORF→Stop

Cross-linking 50-mer RNAs

Ren50 uncapped	pppGGTCATTCCGGTACTGTTGGTATGCTAGCCACCATGACTTCGAAAGTTTA
SL-Ren50 uncapped	pppGGTTTAATTACCCAAGTTTGAGGGCTAGCCACCATGACTTCGAAAGTTTA
m^7GpppG-Ren50	m^7GpppGGTCATTCCGGTACTGTTGGTATGCTAGCCACCATGACTTCGAAAGTTTA
$m_3^{2,2,7}$GpppG-SL-Ren50	$m_3^{2,2,7}$GpppGGTTTAATTACCCAAGTTTGAGGGCTAGCCACCATGACTTCGAAAGTTTA

GGTCATTCCGGTACTGTTGGTATGCTAGCCACC**ATG** = 5' UTR derived from Luciferase (unshaded) and Renilla (shaded)
ATG…….Renilla ORF……..Stop = AUG and renilla open reading frame
GGTTTAATTACCCAAGTTTGAG = spliced leader sequence
*Cap dinucleotide = m^7GpppG or $m_3^{2,2,7}$GpppG

Supplementary Table 3. Sequence Identity of Nematode eIF4Es.

	% identity to given *C. elegans* protein				
Nematode cDNA	IFE-1	IFE-2	IFE-3	IFE-4	IFE-5
A. suum eIF4E-3	47	44	**83**	34	43
A. suum eIF4E-4	29	31	36	**71**	29
B. malayi eIF4E-1	**60**	**54**	48	30	**52**
B. malayi eIF4E-3	49	45	**84**	34	45
B. malayi eIF4E-4	27	29	35	**69**	27
H. contortus eIF4E-1	**68**	**62**	47	30	**60**
H. contortus eIF4E-3	49	46	**88**	35	44
H. contortus eIF4E-4	26	28	35	**67**	26

Supplementary Figure Legends

Supplementary Figure 1. The SL1 sequence splices close to the start codon in many nematode species. Data for nematodes in the same group as *A. suum* (i.e. all but *H. contortus* which lies in the same clade as *C. elegans*), were combined to give Figure 5A. Species are *Ascaris suum*, *Dirofilaria immitis*, *Onchocerca volvulus*, *Litomosoides sigmodontis*, *Brugia malayi*, *Toxocara canis*, and *Haemonchus contortus*

Supplementary Figure 2. The effect of increasing 5′UTR length on m7G capped transcripts in *A. suum* extract. Transcripts containing the given number of random nucleotides were translated in the *A. suum extract*. Reporter activity was expressed relative to the activity of a transcript with the 5′UTR: m7G-N_{21}ACC. Translation initially increases with 5′UTR length, then drops off.

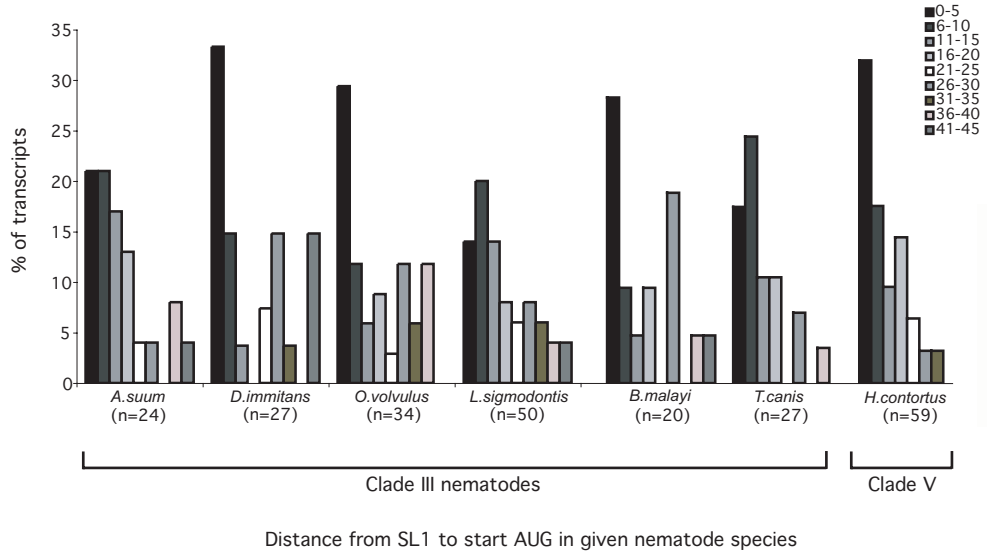
Supplementary Table Legends

Supplementary Table 1. Primers.

Supplementary Table 2. RNA sequences of 5′ UTRs and cross-linking RNAs.

Supplementary Table 3. Percentage identity of discussed nematode eIF4E proteins compared to *C. elegans* IFE proteins. Identities were calculated based on the conserved region of the proteins illustrated that were used for phylogenetic analyses (between arrows in Figure 3A). Numbers in bold highlight highest level of identity.

Supplementary Figure 1



Supplementary Figure 2