

Cloning and Characterization of 4EHP, a Novel Mammalian eIF4E-related Cap-binding Protein*

(Received for publication, August 21, 1997, and in revised form, February 18, 1998)

Eran Rom,^a Hyung Chan Kim,^b Anne-Claude Gingras,^c Joseph Marcotrigiano,^{d,e} Daniel Favre,^f Henrik Olsen,^g Stephen K. Burley,^{d,h,i} and Nahum Sonenberg^j

From the Department of Biochemistry, McGill University, Montreal, Quebec, Canada H3G1Y6, the ^dLaboratories of Molecular Biophysics, The Rockefeller University, New York, New York 10021, and the ^hHoward Hughes Medical Institute and ^gHuman Genome Science Inc., Rockville, Maryland 20850

All eukaryotic mRNAs (except organellar) are capped at their 5' end. The cap structure (m⁷GpppN, where N is any nucleotide) is extremely important for the processing and translation of mRNA. Several cap-binding proteins that facilitate these processes have been characterized. Here we describe a novel human cytoplasmic protein that is 30% identical and 60% similar to the human translation initiation factor 4E (eIF4E). We demonstrate that this protein, named 4E Homologous Protein (4EHP), binds specifically to capped RNA in an ATP- and divalent ion-independent manner. The three-dimensional structure of 4EHP, as predicted by homology modeling, closely resembles that of eIF4E and site-directed mutagenesis analysis of 4EHP strongly suggests that it shares with eIF4E a common mechanism for cap binding. A putative function for 4EHP is discussed.

The cap structure is added cotranscriptionally to all RNA polymerase II transcripts (1, 2). The cap consists of a guanosine methylated at position 7 and linked via a triphosphate bridge to the 5' end of the mRNA (formulated as m⁷GpppN, where N is any nucleotide). The cap plays an important role in gene expression as it stabilizes mRNA (3–6), allows the RNA to be efficiently processed in the nucleus (7–13), and promotes transport to the cytoplasm (14, 15). In the cytoplasm, the cap facilitates ribosome binding to the mRNA (1, 16).

The eukaryotic translation initiation factor eIF4E is the best characterized cap-binding protein (17, 18). A large body of evidence shows that eIF4E mediates cap function during translation initiation (18, 19). eIF4E is the cap-binding subunit of eIF4F, which also contains eIF4G and eIF4A, an RNA helicase.

* This work was supported in part by a grant from the Medical Research Council of Canada (to N. S.). 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.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EBI Data Bank with accession number(s) AF047695.

^a Recipient of a postdoctoral long term fellowship from the Human Frontier Science Program.

^b Recipient of a studentship from the Fonds pour la Formation de Chercheurs et l'Aide à la Recherche.

^c Recipient of a 1967 Centennial Studentship from the Natural Sciences and Engineering Research Council of Canada, Canada.

^d Supported by National Institutes of Health Training Grant GM07982.

^e Recipient of Postdoctoral Fellowship 823A-46646 from the Swiss National Science Foundation.

^f Investigator of the Howard Hughes Medical Institute.

^g Distinguished scientist of the Medical Research Council of Canada and a Howard Hughes Medical Institute International Scholar. To whom correspondence should be addressed. Tel.: 514-398-7274; Fax: 514-398-1287; E-mail: sonenberg@medcor.mcgill.ca.

(20). Following cap binding, eIF4F, in conjunction with eIF4B, unwinds the mRNA secondary structure, permitting recruitment of the ribosome to the translation start site (reviewed in Ref. 18). eIF4E is present in limiting amounts relative to other initiation factors in cells (with the possible exception of rabbit reticulocytes) (21–23). Regulation of its activity via transcription (24), post-translational modifications (25–28), and inhibition of eIF4F assembly (29–32) affects cell proliferation (33), apoptosis (34), and morphogenesis (35).

Recently, proteins involved in cap nuclear functions have been identified and characterized. A nuclear cap-binding complex composed of two subunits, CBP20 and CBP80, facilitates spliceosome assembly (8). Interestingly, the same complex has been implicated in the cap-dependent export of U small nuclear RNA and mRNA (15, 36). eIF4E is also found in the nucleus (37), but the functional significance of this finding remains unclear. Here, we describe the isolation and biochemical properties of a novel eIF4E-like cap-binding protein and discuss its potential biological function.

EXPERIMENTAL PROCEDURES

Materials—Restriction enzymes were obtained from New England Biolabs. T7 RNA polymerase and RNasin were purchased from Promega. T7 DNA polymerase sequencing kit, GpppG, and m⁷GpppG were obtained from Amersham Pharmacia Biotech. Hybond-N nylon membrane and chemiluminescence system were purchased from Amersham Pharmacia Biotech. [α -³²P]GTP, [α -³²P]dCTP, and [α -³²P]UTP (3000 Ci/mmol) were obtained from NEN Life Science Products. The silver staining kit was from Sigma. An antibody to a peptide (YSRRTPGRPTSSQSYE, underlined in Fig. 1) of 4EHP was raised in rabbit and affinity-purified using the AminoLink Plus immobilization kit from Pierce. Oligonucleotides and peptide were prepared at the Sheldon Biotechnology Center, McGill University. All other reagents were reagent grade or better.

Isolation of 4EHP cDNA—An expressed sequence tag (EST)¹ from Human Genome Science Inc. (number 217861) was used to screen a human follicular cell Auni-ZAP-XR (Stratagene) cDNA library (kindly provided by Dr. Karl Heinz Scheit). One million plaques were screened in duplicate, and all procedures to screen, purify, and excise the cDNA inserts from positive clones were performed as described in the Stratagene ZAP Express™ instruction manual. The sequence of the largest cDNA (designated as p4EHP) was obtained by using the T7 DNA sequencing kit from Amersham Pharmacia Biotech. The sequence was deposited in GenBank™, accession number AF047695.

Northern Analysis—Total RNA was isolated from HeLa cells using Trizol (Life Technology, Inc.) and purified twice in batch mode with oligo(dT) beads (Amersham Pharmacia Biotech). Poly(A)⁺ RNA was fractionated through a 1.5% agarose/formaldehyde gel and blotted onto Hybond-N membrane (Amersham Pharmacia Biotech). The blot was probed with random-primed (38) 4EHP cDNA in 50% formamide hybridization buffer at 42 °C and washed with 0.1× SSC, 0.1% SDS at

¹ The abbreviations used are: EST, expressed sequence tag; ORF, open reading frame; bp, base pair(s); PAGE, polyacrylamide gel electrophoresis; HA, hemagglutinin.

65 °C. Exposure to Kodak X-Omat™ AR film was for 24 h.

In Vitro Transcription—p4EHP (5 µg) was linearized with *Xho*I. The template was transcribed with T3 RNA polymerase (Promega) under standard conditions. eIF4E RNA was prepared by transcribing the *Sal*I-linearized pKS-eIF4E plasmid (29) with T7 RNA polymerase.

Overexpression in *Escherichia coli* and Purification of Recombinant 4EHP—p4EHP was used as a template in a polymerase chain reaction. The forward oligonucleotide ACGTGAATTTCGGAACAACAAGTT, the reverse oligonucleotide ACGTGAATTCTCATGGCACATTCACCCG, and the *Pfu* polymerase (Stratagene) were used to amplify the 4EHP open reading frame (ORF). The resulting 730-base pair (bp) fragment was digested with *Eco*RI and cloned into the *Eco*RI site of BlueScript KS(+) (Stratagene) to generate the plasmid p4EHPORF, which was then sequenced and found to completely match the corresponding sequence in the original cDNA clone. This fragment was then subcloned in the *Eco*RI site of pARΔ[59/69] (39) to create a carboxyl-terminal fusion with the Flag-HMK sequence encoded by the vector; the resulting plasmid was designated pAR4EHP. BL21(DE3) bacterial cells (Novagen) were transformed with pAR4EHP, and selected colonies were grown in M9 + glucose to $A_{600} = 0.6$ at which time 1 mM isopropyl-1-thio-β-D-galactopyranoside was added. Cells induced for 4 h were pelleted, resuspended in 5% culture volume of 50 mM Tris-HCl (pH 7.4), 150 mM NaCl, 1 mM phenylmethylsulfonyl fluoride, and sonicated twice on ice for 20 s (1-min interval between sonications) at 60% output using a Vibra-cell sonicator (Sonics & Materials). Cleared lysate was processed on a Flag column according to the manufacturer's instructions (Eastman Kodak Co.). A recombinant histidine-tagged 4EHP protein (His-4EHP) was generated as follows: the 4EHP open reading frame cloned in BlueScript KS(+) was excised with *Bam*HI and *Hind*III and inserted into pQE30 (Qiagen). M15(pREP4) bacterial cells were transformed with the resulting plasmid, grown, and induced as described above, except that the protein lysate was prepared in 20 mM phosphate buffer (pH 7.4), 300 mM NaCl, 1 mM phenylmethylsulfonyl fluoride. Cleared lysate was processed on a nickel column according to the manufacturer's instructions (Qiagen).

The human eIF4E cDNA was isolated from human placenta total RNA via polymerase chain reaction and subcloned into pBluescript KS(+). The coding sequence was inserted into pET8c. BL21(DE3) cells were transformed with pET8c-eIF4E, and protein expression was induced by addition of 0.5 mM isopropyl-1-thio-β-D-galactopyranoside for 4 h at 37 °C. Recombinant proteins were purified by m⁷GDP-agarose resin as described (40). A silver staining procedure (Sigma silver staining kit) was applied to purified His-4EHP and eIF4E to determine their concentrations, as compared with a bovine serum albumin standard (Pierce).

Western Blotting—Proteins were resolved by SDS-PAGE and transferred to nitrocellulose membrane. After blocking for 2 h (5% skim milk in Tris-buffered saline containing 0.1% Tween 20), the membranes were incubated with either anti-eIF4E antibody 5853 (26) or affinity-purified anti-4EHP peptide antibody for 2 h in Tris-buffered saline containing 0.1% Tween 20. ¹²⁵I-Protein A or anti-rabbit Ig-peroxidase coupled (followed by ECL reaction) were used for detection.

Construction of Mutants—Oligonucleotide-directed point mutations were introduced into the 4EHP cDNA using the uracil incorporation method of site-directed mutagenesis (41). Mutations were confirmed by sequencing.

Cross-linking to Capped RNA—Luciferase RNA supplied by Promega (catalog number L4561) was capped with [α -³²P]GTP using vaccinia virus guanylyltransferase, as described (42). For UV cross-linking, cap-labeled RNA (2×10^4 cpm) was incubated with purified Flag-HMK4EHP or His-4EHP proteins and processed for analysis by SDS-PAGE according to Ref. 42.

Electrophoretic Mobility Retardation Analysis—pBlueScript KS(-), digested with *Bam*HI, was used as a template by T7 RNA polymerase. RNA was synthesized in the presence of 0.5 mM m⁷GpppG (methylated probe) or GpppG (nonmethylated probe) and internally labeled with [α -³²P]UTP (50 µCi, 3000 Ci/mmol). Probes were spun through a G-25 Sephadex column, phenol-extracted, and ethanol-precipitated. The dried pellets were dissolved in water at 2×10^4 cpm/µl. Recombinant eIF4E (0.5 µg) or His-4EHP (0.5 µg; wild-type or mutant) proteins were incubated with either a methylated or nonmethylated probe (2×10^4 cpm) as described in Ref. 43.

Cell Lines and Indirect Immunofluorescence Microscopy—The tetracycline-repressible expression plasmid pBPSTR1 (44) was used to create cell lines stably overexpressing an HA epitope-tagged 4EHP (HA-4EHP). Three hours before transfection, 1 µg/ml tetracycline was added to logarithmically growing NIH 3T3 cells (obtained from the American Type Culture Collection, Rockville, MD). Cells were transfected with

pBPSTR1 containing an HA-tagged 4EHP cDNA using LipofectAMINE™ reagent following manufacturer's (Life Technologies, Inc.) guidelines. Three hours post-transfection, the cells were washed and placed in fresh Dulbecco's modified Eagle's medium containing 10% fetal bovine serum. Puromycin (1.5 µg/ml) and tetracycline (1 µg/ml) were added at 30 h post-transfection and for an additional 10 days. Resistant colonies were tested for HA-4EHP expression by Western analysis using the monoclonal anti-HA antibody HA.11 (Babco) at a dilution of 1:1000. A control cell line was derived from transfection with pBPSTR1 alone.

Cells stably expressing HA-tagged 4EHP proteins, or mock-transfected cells, were processed for immunofluorescence as described (37), using HA.11 antibody at a dilution of 1:1000.

Homology Modeling of 4EHP—Several models of 4EHP were generated using the MODELER program (45), with the alignment shown in Fig. 2. Z scores were calculated for the overall structure and as a function of each amino acid by PROSAIL (46). The models with the lowest Z scores were visually inspected for any serious errors. Structures were then displayed and analyzed using the InsightII program (Biosym Technologies, San Diego, CA).

RESULTS

Cloning and Features of 4EHP cDNA

A human EST derived from a fetal brain library was found to encode an ORF with significant homology to the eukaryotic translation initiation factor eIF4E. This cDNA is 816 bp long, of which 594 bp encode a potential 198 amino acid polypeptide with 34% identity to the human eIF4E. The rest of the clone is composed of 193 bp of 3'-untranslated region containing a polyadenylation signal followed by a stretch of 29 adenosine residues. As the cloned EST lacked a 5'-untranslated region and an initiator ATG, a human follicular cell cDNA library was screened utilizing the EST DNA as a probe. Screening of 1×10^6 plaques yielded three independent clones with overlapping sequences. The longest, designated as 4EHP (for eIF4E Homologous Protein), was 989 base pairs in length and contained a 245-amino acid ORF encoding a protein with a calculated mass of 28 kDa and pI of 8.9 (Figs. 1, A and B). 4EHP is 30% identical and 60% similar to eIF4E at the amino acid level (Fig. 2). A BLAST search (47) detected mouse, *Caenorhabditis elegans*, and plant (several species) 4EHP homologues. The predicted 4EHP polypeptide contains no characterized protein motifs. As the isolated 5'-untranslated region lacks an upstream stop codon, we analyzed whether 4EHP represents a full-length clone. Thus, poly(A)⁺ RNA isolated from HeLa cells was probed with the 4EHP cDNA. The RNA prepared using the cDNA as template migrates at the same position as the cellular RNA, indicating that p4EHP harbors a cDNA corresponding to the entire 4EHP transcript (Fig. 1C). The detected HeLa mRNA is not the eIF4E mRNA, as demonstrated by including eIF4E RNA as a control. Various rat tissues, probed with anti-4EHP antibodies, demonstrated the presence of a ubiquitously expressed polypeptide doublet that migrates at 30–32 kDa (data not shown). Translation of 4EHP *in vitro* using rabbit reticulocyte lysate yielded the same polypeptide doublet (data not shown). To determine the concentration of 4EHP relative to eIF4E, a Western blot analysis, using purified recombinant eIF4E or 4EHP as standards (Fig. 3, lanes 1–4), was performed in extracts from HeLa and 293 cells (lanes 5–8 and 9–12, respectively). There is roughly 5–10 times more eIF4E than 4EHP in both cell types. For example, in HeLa cells (lanes 5–8), there is approximately 50 ng/mg 4EHP and 450 ng/mg eIF4E. 4EHP protein was detected in all rat tissues analyzed, indicating that it is ubiquitously expressed (not shown).

4EHP Is A Cap-binding Protein

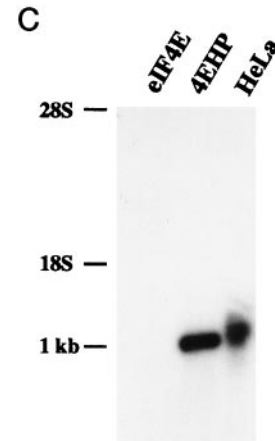
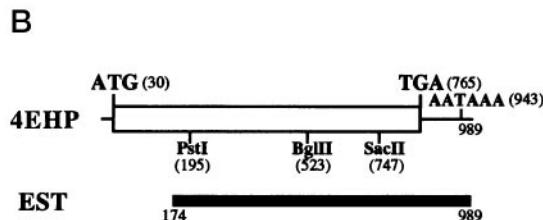
Cross-linking to Capped RNA—To demonstrate the ability of 4EHP to specifically recognize and interact with capped RNA, we used a photochemical cross-linking assay (42). Purified Flag-HMK-tagged 4EHP was UV-irradiated in the presence of

FIG. 1. **The 4EHP cDNA.** A, nucleotide sequence and the deduced amino acid sequence (single letter code) from the full-length cDNA clone. Noncoding regions are in *lowercase letters*, and the polyadenylation signal is *underlined*. B, schematic presentation of 4EHP cDNA. The coding region is shown as an *open box* and noncoding regions as *lines*. The positions of the first ATG, the stop codon, and the polyadenylation signal are indicated, as well as the location of three unique restriction sites. The region spanned by the EST described in the text is shown below as a *closed box*. C, Northern blot analysis using 20 μ g of HeLa poly(A)⁺ RNA and the described cDNA as probe. As controls, 20 ng of eIF4E and 10 ng of 4EHP RNAs transcribed *in vitro* were included in the analysis. Size markers are indicated on the left.

gaggcagtgggcagcagcgccggcgagagg 29

A

M	N	N	K	F	D	A	L	K	D	D	S	G	D	15	
ATG	AAC	AAC	AAG	TTC	GAC	GCT	TTG	AAA	GAT	GAT	GAC	AGT	GGG	GAC	74
H	D	Q	N	E	E	N	S	T	Q	K	D	G	E	K	30
CAT	GAT	CAG	AAT	GAA	GAA	AAC	AGC	ACA	CAG	AAA	GAT	GGT	GAG	AAG	119
E	K	T	E	R	D	K	N	Q	S	S	S	K	R	K	45
GAA	AAA	ACG	GAA	CGA	GAC	AAG	AAT	CAG	AGC	AGT	AGC	AAG	AGA	AAG	164
A	V	V	P	G	P	A	E	H	P	L	Q	Y	N	Y	60
GCT	GTT	GTC	CCT	GGA	CCG	GCA	GAG	CAT	CCC	CTG	CAG	TAC	AAC	TAC	209
T	F	W	Y	S	R	R	T	P	G	R	P	T	S	S	75
ACT	TTT	TGG	TAC	TCC	AGG	AGA	ACC	CCC	GGC	CGT	CCC	ACG	AGC	TCA	254
Q	S	Y	E	Q	N	I	K	Q	I	G	T	F	A	S	90
CAG	AGC	TAT	GAA	CAG	AAT	ATC	AAA	CAG	ATT	GGC	ACC	TTT	GCC	TCT	299
V	E	Q	F	W	R	F	Y	S	H	M	V	R	P	G	105
GTG	GAG	CAG	TTC	TGG	AGG	TTT	TAT	AGC	CAC	ATG	GTA	CGT	CCT	GGG	344
D	L	T	G	H	S	D	F	H	L	F	K	E	G	I	120
GAC	CTG	ACA	GGC	CAC	AGT	GAC	TTC	CAT	CTC	TTC	AAA	GAA	GGG	ATT	389
K	P	M	W	E	D	D	N	A	K	N	G	G	K	W	135
AAA	CCC	ATG	TGG	GAG	GAT	GAT	GCA	AAT	AAA	AAT	GGT	GGC	AAG	TGG	434
I	I	R	L	R	K	G	L	A	S	R	C	W	E	N	150
ATT	ATT	CGG	CTG	CGG	AAG	GGC	TTG	GCC	TCC	CGT	TGC	TGG	GAG	AAT	479
L	I	L	A	M	L	G	E	Q	F	M	V	G	E	E	165
CTC	ATT	TTG	GCC	ATG	CTG	GGG	GAA	CAG	TTC	ATG	GTT	GGG	GAG	GAG	524
I	C	G	A	V	V	S	V	R	F	Q	E	D	I	I	180
ATC	TGT	GGG	GCT	GTG	GTG	TCT	GTC	CGC	TTT	CAG	GAA	GAC	ATT	ATT	569
S	I	W	N	K	T	A	S	D	Q	A	T	T	A	R	195
TCA	ATA	TGG	AAT	AAG	ACT	GCC	AGT	GAC	CAA	GCA	ACC	ACA	GCC	CGA	614
I	R	D	T	L	R	R	V	L	N	L	P	P	N	T	210
ATC	CGG	GAC	ACA	CTT	CGG	CGA	GTG	CTT	AAC	CTA	CCT	CCC	AAC	ACC	659
I	M	E	Y	K	T	H	T	D	S	I	K	M	P	G	225
ATT	ATG	GAA	TAC	AAA	ACT	CAC	ACC	GAC	AGC	ATC	AAA	ATG	CCA	GGC	704
R	L	G	P	Q	R	L	L	F	Q	N	L	W	K	P	240
AGG	CTG	GGC	CCC	CAA	AGG	CTC	CTT	TTT	CAA	AAC	CTC	TGG	AAG	CCG	749
R	L	N	V	P	*										245
CGG	TTG	AAT	GTG	CCA	TGA	ccctctccctctctggatggcaccatcattgaagc									802
tggtgctcatcggagtctcttctgttctgttggcgtgctacctggaagatccttctgtcctg															861
gacaagagggaatttgggaagagcattttatgttttaagaacaggctgacacgcagcagct															920
acaacaacagctgagatcacttataaatgggtgctaactaaaaaaaaaaaaaaaaaaaaa															979
aaaaaaaaaa															989



³²P-cap-labeled luciferase RNA. Cross-linked products were then subjected to SDS-PAGE. 4EHP cross-linking was not observed without UV irradiation and was inefficient when a nonmethylated capped RNA was used (Fig. 4, compare lanes 1 and 2 with lane 3). The cross-linking of 4EHP was only slightly affected by the presence of excess of GpppG (lane 4), but was abolished by excess of m⁷GpppG (lane 5). The extent of cross-linking was not affected by ATP or Mg²⁺ (data not shown). The requirements for 4EHP cross-linking are, thus, similar to those reported for eIF4E (48).

Gel Mobility Shift of Capped RNA—To further demonstrate the specific binding of 4EHP to the cap, two radioactively labeled 58 nucleotide RNAs capped with either m⁷GpppG or with GpppG were incubated with wild-type 4EHP, and the resulting complexes analyzed by nondenaturing PAGE (Fig. 5; yeast RNA was included in the reactions to minimize nonspecific interactions between the protein and the RNA probe).

eIF4E and a mutant 4EHP, WED124–126FAA, in which tryptophan, glutamic acid, and aspartic acid at positions 124–126 were replaced by phenylalanine, alanine, and alanine, respectively (this mutant protein is unable to bind to the cap structure), were included as controls. eIF4E, as shown previously (49), formed a complex with the methylated capped RNA (Fig. 5, lane 2) and to a much lesser extent with the nonmethylated capped RNA (lane 3). Likewise, wild-type 4EHP formed a complex only with the methylated probe (lanes 4 and 5). As expected, no complex was observed with the mutant 4EHP (lanes 6 and 7). In accordance with these results, 4EHP, but not the mutant protein, also bound to a m⁷GDP-coupled resin (data not shown).

Amino Acid Residues Required for Cap Binding

Recently, the three-dimensional structure of murine eIF4E, bound to m⁷GDP, was determined by x-ray crystallography

In the crystal structure of eIF4E, Asp¹⁰⁴, which is absolutely conserved in all species, does not contact the cap structure. A change at the corresponding position (Asp¹²⁶ to Ala) in 4EHP has only a modest effect on cross-linking (*lane 5*). eIF4E, from all species so far examined, contains eight tryptophans. Of these, six are conserved in 4EHP, whereas the two remaining are replaced by tyrosine residues in 4EHP (see Fig. 2). We individually replaced the tryptophans in 4EHP that are conserved between 4EHP and eIF4E (Trp⁶³, Trp⁹⁵, Trp¹³⁵, and Trp¹⁴⁸; see also Fig. 2) with alanine and analyzed the cap binding activity of these mutants by UV cross-linking. With the exception of Trp⁹⁵ (which possessed 40% of wild-type activity; compare *lane 1* with *lane 8*), mutation of any of the tryptophan residues (63, 135, and 148) abolished cap binding (compare *lane 1* with *lanes 6, 10, and 12*). These findings substantiate the results of the homology modeling, which place Trp⁶³, Trp¹³⁵, and Trp¹⁴⁸ in the hydrophobic core. Trp⁹⁵, on the other hand, is predicted to lie on the convex dorsal surface of 4EHP. We also substituted the putative cap-binding tryptophans (Trp¹²⁴ and Trp¹⁸³) with phenylalanine, which also contains an aromatic ring to determine whether the benzyl ring of phenylalanine could stack with the m⁷G ring as well as the indol group of tryptophan. Substitution of Trp¹⁸³ with phenylalanine eliminated cap binding (*lane 15*), while substitution of Trp¹²⁴ with phenylalanine reduced cross-linking to 13% of control (*lane 3*). A similar effect was observed when the corresponding amino acids were substituted in yeast eIF4E (51). This suggests that the nature of the aromatic ring of the carboxyl-distal aromatic amino acid is more important than the amino-proximal amino acid. Individual Trp to Phe changes of tryptophans located in the hydrophobic core (Trp⁶³, Trp¹³⁵, and Trp¹⁴⁸; *lanes 7, 11, and 13*, respectively) had predictably less severe effects on cap binding.

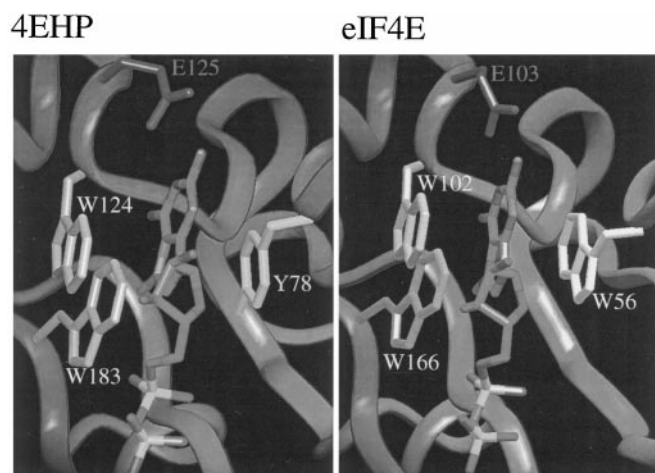
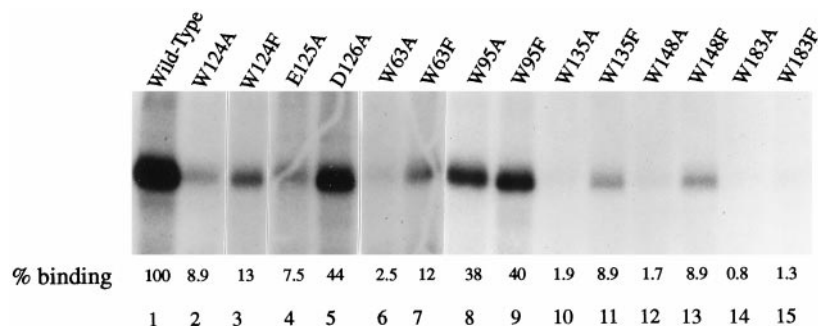


FIG. 6. Computer modeling of the structure of 4EHP cap binding site as compared with the crystal structure of eIF4E. Amino acids that form direct contact with the alkylated base are indicated.

FIG. 7. Effect of mutations on cap binding. Recombinant His-4EHP proteins mutated at the indicated residues were analyzed by cross-linking to capped RNA as described in the legend to Fig. 3. The extent of cross-linking was estimated using the Fuji Bas 2000 phosphorimager and is indicated as a percentage of wild-type.



Subcellular Localization of 4EHP

Antisera raised against 4EHP did not detect endogenous 4EHP by indirect immunofluorescence (data not shown). We, therefore, analyzed the subcellular localization of 4EHP in an NIH 3T3 cell line stably expressing an HA epitope-tagged version of the protein (HA-4EHP). As determined by a quantitative Western analysis, the level of HA-4EHP was twice that of the endogenous 4EHP (data not shown). A monoclonal anti-HA antibody stained HA-4EHP in the cytoplasm, while the fluorescent signal in the nucleus was comparable with that seen in mock-transfected cells (Fig. 8). As controls, we stained both cell lines with antibodies directed against eIF4A (a cytoplasmic translation initiation factor) or Sm (a component of small nuclear RNPs in the nucleus). eIF4A and Sm were found exclusively in the cytoplasmic and nuclear compartments, respectively (data not shown). To corroborate this result we analyzed the biochemical localization of endogenous 4EHP in HeLa cells. Cytoplasmic and nuclear fractions were prepared and analyzed by Western blotting using anti-4EHP antibodies. In accord with the results of the immunofluorescence study, 4EHP was found exclusively in the cytoplasm (data not shown).

DISCUSSION

Biochemical characterization, using cap-analog resin, UV cross-linking, and gel mobility shift assays, demonstrated that, like eIF4E, 4EHP specifically recognizes the cap structure. Recently, the three-dimensional structure of eIF4E, bound to m⁷GDP, was determined by x-ray crystallography (50). The alkylated base is held in position by π - π stacking interactions between the indol group of two tryptophans (Trp⁵⁶ and Trp¹⁰²). Cap binding is enhanced by additional hydrogen bonds with the backbone amino group of Trp¹⁰² and the side chain of Glu¹⁰³ and by a van der Waals contact between the methyl group of the guanine ring and the side chain of a third tryptophan (Trp¹⁶⁶). The ribose makes an additional van der Waals contact with Trp⁵⁶. The two m⁷GDP phosphate groups are hydrogen-bonded to Arg¹⁵⁷ and Lys¹⁶². Alignment of eIF4E and 4EHP shows that all but two of these residues are conserved (Fig. 2).

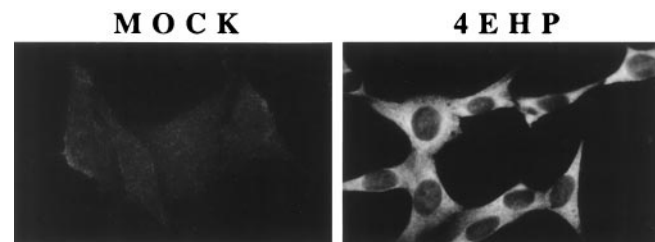


FIG. 8. Subcellular localization of HA-4EHP by indirect immunofluorescence. NIH 3T3 cells that stably overexpress an HA-4EHP and mock-transfected cells were grown in chamber slides. Cells at 50% confluence were fixed with paraformaldehyde and incubated with a monoclonal anti-HA antibody, followed by incubation with a rhodamine-conjugated rabbit anti-mouse IgG. Photographs were taken with a rhodamine filter at a $\times 1000$ magnification.

The two residues that differ are Tyr⁷⁸ and Ile¹⁷⁹, which correspond to Trp⁵⁶ and Lys¹⁶² in murine eIF4E. Tyrosine 78 is a conservative change and is expected to fulfill the same role as that of Trp⁵⁶ in cap binding. However, the Lys to Ile substitution is not conservative, suggesting that a positively charged residue at this position is not essential for cap binding. Not surprisingly, sequence alignments and computer modeling predict that the overall structure of 4EHP, including the cap binding cleft, is very similar to that of eIF4E (Fig. 6). Cap binding analysis of 4EHP mutants of Trp¹²⁴, Glu¹²⁵, and Trp¹⁸³ (residues predicted to form contacts with the cap) demonstrated that these amino acids are absolutely required for cap binding. Thus, we conclude that the mechanism of cap recognition is similar for eIF4E and 4EHP. The structure of eIF4E and that predicted for 4EHP differs markedly from that of the vaccinia protein VP39, which also binds specifically to the cap structure (52, 53). However, VP39 also sandwiches the cap structure between two aromatic side chains (phenylalanine and tyrosine) (53). It seems, therefore, that eIF4E, 4EHP, and VP39 bind the cap in a similar fashion.

Given the putative structural similarity with eIF4E and its cytoplasmic localization, it is possible that 4EHP plays a role in translation. We investigated possible interactions of 4EHP with candidate translation factors or inhibitors of translation, such as eIF4G, the 4E-BPs, and the poly(A)-binding protein, by various methods, without success. However, we cannot exclude the possibility that 4EHP interacts weakly with one of these proteins. Cap recognition is also required for several enzymatic activities, such as those performed by 1) 2'-O-methyltransferases to form m⁷G(5')pppN_mpN (cap I) and m⁷G(5')pppN_m-pN_m (cap II) (54–56); 2) m⁷G(5')pppN-pyrophosphatases that remove the cap from capped-RNAs (6, 57–59); and 3) Sm-dependent methyltransferase to form trimethylguanosine cap structures found mainly in small nuclear RNAs (54). 4EHP is most likely not involved in functions 1 and 2 because: (a) although cap I and cap II structures are not found in plants (60), we have retrieved ESTs corresponding to 4EHP cDNA from several plant species (*Brassica Napus*, *Arabidopsis thaliana*, *Zea mays*). (b) Recombinant 4EHP failed to exhibit methyltransferase or decapping activities.² It is possible, however, that 4EHP is involved in recognition of the trimethylguanosine of small nuclear RNAs for import into the nucleus. A 4EHP homologue does not exist in yeast. Although it is clear that in higher eukaryotes the small nuclear RNAs exit the nucleus and become hypermethylated in the cytoplasm (61), it is not known whether this process occurs in yeast.

Acknowledgments—We thank Chantal Binda for excellent technical assistance and Sharon Okouneff for her help in analyzing the subcellular location of 4EHP. We are grateful to Dr. Karl Heinz Scheit for the human follicular cell cDNA library and to Alessandra Gradi, Greg Cosentino, Christine DeMaria, Brian Raught, and Kianoush Khaleghpour for critically reading the manuscript. We are also grateful to A. Sali and R. Sanchez for their help with homology modeling of 4EHP using the structure of eIF4E.

REFERENCES

- Shatkin, A. J. (1976) *Cell* **9**, 645–653
- Salditt-Georgieff, M., Harpold, M., Chen-Kiang, S., and Darnell, J. E. (1980) *Cell* **19**, 69–78
- Furuichi, Y., LaFiandra, A., and Shatkin, A. J. (1977) *Nature* **266**, 235–239
- Shimotohno, K., Kodama, Y., Hashimoto, J., and Miura, K. I. (1977) *Proc. Natl. Acad. Sci. U. S. A.* **74**, 2734–2738
- Murthy, K. G. K., Park, P., and Manley, J. L. (1991) *Nucleic Acids Res.* **19**, 2685–2692
- Beelman, C. A., Stevens, A., Caponigro, G., LaGrandeur, T. E., Hatfield, L., Fortner, D. M., and Parker, R. (1996) *Nature* **382**, 642–646
- Hart, R. P., McDevitt, M. A., and Nevins, J. R. (1985) *Cell* **43**, 677–683
- Izaurrealde, E., Lewis, J., Mcguigan, C., Jankowska, M., Darzynkiewicz, E., and Mattaj, I. W. (1994) *Cell* **78**, 657–668
- Inoue, K., Ohno, M., Sakamoto, H., and Shimura, Y. (1989) *Genes Dev.* **3**, 1472–1479
- Ohno, M., Sakamoto, H., and Shimura, Y. (1987) *Proc. Natl. Acad. Sci. U. S. A.* **84**, 5187–5191
- Kraimer, A. R., Maniatis, T., Ruskin, B., and Green, M. R. (1984) *Cell* **36**, 993–1005
- Konarska, M. M., Padgett, R. A., and Sharp, P. A. (1984) *Cell* **38**, 731–736
- Edery, I., and Sonenberg, N. (1985) *Proc. Natl. Acad. Sci. U. S. A.* **82**, 7590–7594
- Lewis, J. D., Gunderson, S. I., and Mattaj, I. W. (1995) *J. Cell Sci.* **19**, 13–19
- Izaurrealde, E., Lewis, J., Gamberi, C., Jarmolowski, A., McGuigan, C., and Mattaj, I. W. (1995) *Nature* **376**, 709–712
- Shatkin, A. J. (1985) *Cell* **40**, 223–224
- Sonenberg, N., Morgan, M. A., Merrick, W. C., and Shatkin, A. J. (1978) *Proc. Natl. Acad. Sci. U. S. A.* **75**, 4843–4847
- Sonenberg, N. (1996) *mRNA 5' Cap Binding Protein eIF4E and Control of Cell Growth: Translational Control* (Hershey, J. W. B., Mathews, M. B., and Sonenberg, N., eds) pp. 245–269, Cold Spring Harbor Press, Cold Spring Harbor, NY
- Rhoads, R. E. (1988) *Trends Biochem. Sci.* **13**, 52–56
- Rozen, F., Edery, I., Meerovitch, K., Dever, T. E., Merrick, W. C., and Sonenberg, N. (1990) *Mol. Cell. Biol.* **10**, 1134–1144
- Hiremath, L. S., Webb, N. R., and Rhoads, R. E. (1985) *J. Biol. Chem.* **260**, 7843–7849
- Duncan, R., Milburn, S. C., and Hershey, J. W. (1987) *J. Biol. Chem.* **262**, 380–388
- Rau, M., Ohlmann, T., Morley, S. J., and Pain, V. M. (1996) *J. Biol. Chem.* **271**, 8983–8990
- Jones, R. M., Branda, J., Johnston, K. A., Polymenis, M., Gadd, M., Rustgi, A., Callanan, L., and Schmidt, E. V. (1996) *Mol. Cell. Biol.* **16**, 4754–4764
- Joshi-Barve, S., Rychlik, W., and Rhoads, R. E. (1990) *J. Biol. Chem.* **265**, 2979–2983
- Frederickson, R. M., Montine, K. S., and Sonenberg, N. (1991) *Mol. Cell. Biol.* **11**, 2896–2900
- Rinker-Schaeffer, C. W., Austin, V., Zimmer, S., and Rhoads, R. E. (1992) *J. Biol. Chem.* **267**, 10659–10664
- Morley, S. J., and Traugh, J. A. (1993) *Biochimie (Paris)* **75**, 985–989
- Pause, A., Belsham, G., Gingras, A.-C., Donzé, O., Lin, T.-A., Lawrence, J. C., Jr., and Sonenberg, N. (1994) *Nature* **371**, 762–767
- Lin, T.-A., Kong, X., Haystead, T. A. J., Pause, A., Belsham, G., Sonenberg, N., and Lawrence, J. C., Jr. (1994) *Science* **266**, 653–656
- Graves, L. E., Bornfeldt, K. E., Argast, G. M., Krebs, E. G., Kong, X., Lin, T.-A., and Lawrence, J. C. (1995) *Proc. Natl. Acad. Sci. U. S. A.* **92**, 7222–7226
- Haystead, T. A. J., Haystead, C. M. M., Hu, C., Lin, T.-A., and Lawrence, J. C., Jr. (1994) *J. Biol. Chem.* **269**, 23185–23191
- Smith, M. R., Jaramillo, M., Liu, Y., Dever, T. E., Merrick, W. C., Kung, H., and Sonenberg, N. (1990) *New Biol.* **2**, 648–654
- Polunovsky, V. A., Rosenwald, I. B., Tan, A. T., White, J., Chiang, L., Sonenberg, N., and Bitterman, P. B. (1996) *Mol. Cell. Biol.* **16**, 6573–6581
- Klein, P. S., and Melton, D. A. (1994) *Science* **265**, 803–806
- Visa, N., Izaurrealde, E., Ferreira, J., Daneholt, B., and Mattaj, I. W. (1996) *J. Cell Biol.* **133**, 5–14
- Lejbkowitz, F., Goyer, C., Darveau, A., Néron, S., Lemieux, R., and Sonenberg, N. (1992) *Proc. Natl. Acad. Sci. U. S. A.* **89**, 9612–9616
- Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) in *Molecular Cloning: A Laboratory Manual* (Nolan, C., ed) Vol. 2, 2nd Ed., pp. 10.13–10.17, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
- Blonar, M. A., and Rutter, W. J. (1992) *Science* **256**, 1014–1018
- Edery, I., Altmann, M., and Sonenberg, N. (1988) *Gene (Amst.)* **74**, 517–525
- Kunkel, T. A. (1985) *Proc. Natl. Acad. Sci. U. S. A.* **82**, 488–492
- Pelletier, J., and Sonenberg, N. (1985) *Mol. Cell. Biol.* **5**, 3222–3230
- Izaurrealde, E., Stepinski, J., Darzynkiewicz, E., and Mattaj, I. W. (1992) *J. Cell Biol.* **118**, 1287–95
- Paulus, W., Baur, I., Boyce, F. M., Breakfield, X. O., and Reeves, S. A. (1996) *J. Virol.* **70**, 62–67
- Sali, A., and Blundell, T. L. (1993) *J. Mol. Biol.* **234**, 779–815
- Sippel, M. J. (1993) *Proteins Struct. Funct. Genet.* **17**, 355–362
- Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. P. (1990) *J. Mol. Biol.* **215**, 403–410
- Sonenberg, N. (1981) *Nucleic Acids Res.* **9**, 1643–1656
- Edery, I., Chu, L. L., Sonenberg, N., and Pelletier, J. (1995) *Mol. Cell. Biol.* **15**, 3363–3371
- Marcotrigiano, J., Gingras, A.-C., Sonenberg, N., and Burley, S. (1997) *Cell* **89**, 951–961
- Altmann, M., Edery, I., Trachsel, H., and Sonenberg, N. (1988) *J. Biol. Chem.* **263**, 17229–17232
- Hodel, A. E., Gershon, P. D., Shi, X., and Quicho, F. A. (1996) *Cell* **85**, 247–256
- Hodel, A. E., Gershon, P. D., Shi, X. N., Wang, S. M., and Quicho, F. A. (1997) *Nat. Struct. Biol.* **4**, 350–354
- Reddy, R., Singh, R., and Shimba, S. (1992) *Pharmacol. Ther.* **54**, 249–267
- Langberg, S. R., and Moss, B. (1981) *J. Biol. Chem.* **256**, 10054–10060
- Shi, X., Yao, P., Jose, T., and Gershon, P. D. (1996) *RNA (N. Y.)* **2**, 88–101
- Nuss, D. L., Furuichi, Y., Koch, G., and Shatkin, A. J. (1975) *Cell* **6**, 21–27
- Nuss, D. L., and Furuichi, Y. (1977) *J. Biol. Chem.* **252**, 2815–2821
- Lavers, G. C. (1977) *Mol. Biol. Rep.* **3**, 413–420
- Banerjee, A. K. (1980) *Microbiol. Rev.* **44**, 175–205
- Mattaj, I. W. (1986) *Cell* **46**, 905–911

² E. Rom, unpublished observations.