Volume 13 Number 19 1985 Nucleic Acids Research

Cloning of eukaryotic protein synthesis initiation factor genes: isolation and characterization of cDNA clones encoding factor eIF-4A

Peter J.Nielsen[1,3,4], Gary K.McMaster[2,4] and Hans Trachsel[1]

[1]Biocenter, University of Basel, CH-4056 Basel, and [2]Hoffmann-La Roche Inc., CH-4002 Basel, Switzerland

Received 25 July 1985; Revised and Accepted 9 September 1985

ABSTRACT

Monoclonal antibodies directed against rabbit reticulocyte protein synthesis initiation factor 4A (eIF-4A) were used to isolate mouse cDNA clones expressing eIF-4A protein sequences in E.coli. The identity of cDNA clones encoding eIF-4A sequences was confirmed by hybrid-selected translation and peptide mapping of the translation product. Analysis of the mRNA coding for eIF-4A from mouse liver and HeLa cells by Northern hybridization revealed two discrete mRNA species of approximately 2000 and 1600 nucleotides in length. The existence of two mRNAs in mouse and HeLa cells encoding eIF-4A was confirmed by cDNA sequencing.

INTRODUCTION

Eukaryotic protein synthesis initiation is a complex assembly process which involves binding of initiator Met-tRNA^{met} and mRNA to the 40S ribosomal subunit followed by the joining of the 60S ribosomal subunit (for reviews, see 1-5). The many steps in this pathway are catalyzed by at least ten initiation factors (on the order of twentyfive polypeptide chains) with a total molecular weight of over one million daltons. Most initiation factors are required for mRNA binding, but very little else is known at present about their function(s). The process of mRNA binding requires ATP and can be subdivided into recognition of the 5' end or cap of the mRNA by the 40S ribosomal subunit, scanning of the 5' untranslated sequences and recognition of an initiator AUG codon. Among the initiation factors required for mRNA binding is eIF-4A, a single polypeptide chain of 43 kilodaltons (6-8). This factor appears to be especially interesting since it was recently shown that it is both a subunit of a high molecular weight protein complex involved in cap recognition

(9,10) and is required as a single polypeptide chain for mRNA binding to ribosomes (11).

In an effort to elucidate the structure, function and possible regulation of this and other eukaryotic protein synthesis initiation factors, we have begun to clone their genes. Here, we describe the identification and characterization of cDNA clones encoding eukaryotic initiation factor 4A.

MATERIALS AND METHODS

Cell culture

HeLa cells were grown either in suspension culture in Joklik-modified minimal essential medium (GIBCO) containing 10% newborn calf serum or in monolayer culture in Eagles modified minimal essential medium (GIBCO) containing 10% new born calf serum. Cl9 mouse hybridoma cells (gift from Dr. T. Staehelin, Hoffmann-La Roche, CH-4002 Basel, Switzerland) were grown in monolayer culture as described above for HeLa monolayer culture.

Preparation of poly(A)$^+$RNA

All glassware was heated to 200°C for 2hr and all buffers were autoclaved before use. HeLa cells (about 10^8cells) were pelleted (200 xg, 10 min, 4°C) directly from suspension cultures (approximately $5x10^5$ cells/ml) or after detachment from monolayer culture dishes by incubation in PBS (20 mM potassium phosphate, pH 7, 150 mM NaCl) containing 1 mM EDTA for 5 min at 37°C. Cells were washed once in cold PBS, resuspended at approximately $4x10^6$ cells/ml in 10 mM Tris-HCl, pH 7.6, 10 mM NaCl, 1 mM EDTA on ice and immediately homogenized with 5 strokes in a Dounce homogenizer. The homogenate was centrifuged (2000xg, 10 min, 4°C) and sodium dodecyl sulfate (SDS) and LiCl added to the resulting supernatant to give final concentrations of 0,5 % and 0.1 M, respectively. Total cellular RNA was then deproteinized by 3 successive phenol extractions followed by ethanol precipitation (12). Total cytoplasmic RNA from Cl9 mouse hybridoma cells and Balb/c mouse liver was isolated as described by Schibler et al. (13). Poly(A)$^+$RNA was isolated from total RNA by two cycles of purification on oligo(dT)-cellulose as described (12).

cDNA synthesis

cDNA was synthesized according to the methods compiled by

Maniatis et al. (12) using purified enzymes from Anglian Biotechnology laboratory, except that denaturation of the mRNA-DNA hybrid following the first-strand synthesis was according to the boil-chill procedure of Buell et al. (14). Alternatively, cDNA was synthesized using the Reverse Transcriptase System from New England Nuclear (NEK-021). The ends of the cDNA were made flush using T4 DNA polymerase and blunt-end ligated to kinased EcoRI linkers in the presence of 15 % polyethylene glycol (PEG) (15). The cDNA was then digested with EcoRI restriction endonuclease and size-fractionated on a Biogel A-15M column (Bio-Rad). Fractions containing cDNA larger than approximately 200 base pairs were pooled and the cDNA precipitated with ethanol.

Construction of cDNA libraries in λgt11

Growth and maintenance of bacterial strains and phage DNA isolation was essentially as described (12). Lambda gt11 DNA (16) was digested with the restriction endonuclease EcoRI, dephosphorylated with calf intestine alkaline phosphatase (Boehringer) (12), phenol-extracted and ethanol-precipitated. cDNA with EcoRI linkers was then ligated at 14°C to dephosphorylated λgt11 DNA in the presence of 15 % PEG (15). The molar ratio of ends of cDNA to vector was approximately 2:1. The average size of the cDNA after EcoRI digestion was 1 kilobase (Kb). The resulting ligation products were packaged in vitro into phage particles (12) and either used directly for immuno-screening (see below) or amplified on plates (12). We obtained $5\text{-}20 \times 10^6$ recombinant phage per µg cDNA and approximately 95% of the phage contained inserts.

Immuno-screening

Immuno-screening of the λgt11 cDNA libraries was as described (17) with E.coli Y1090, except that the synthesis of the beta-galactosidase-hybrid protein was induced by laying a nitrocellulose filter on the plaque-containing bacterial lawn (13cm Ø plate) and spraying it with approximately 0.7 ml of 10 mM isopropylthiogalactoside (IPTG). After incubation of the plates for 1.5 hr at 37°C, the filters were removed and further incubated at room temperature first with 2.5 % bovine serum albumin (BSA) in TBS (20 mM Tris-HCl, pH 7.6, 150 mM NaCl) for 1

hr followed by hybridoma culture supernatant containing anti-eIF4A antibodies (10) diluted 1:3 in TBT (0.5% BSA, 0.1% Triton X-100 in TBS) for 4 hrs to overnight. The filters were washed for 1 hr in TBS, incubated with rabbit-anti mouse antibodies (DAKO, 1:1000 dilution in TBT) for 2 hrs, washed again for 1 hr in TBS followed by incubation with [^{125}I] protein A (10 mCi/mg protein) in TBT for 30 min. Filters were washed with TBS, dried and exposed to Kodak X-omat SO-282 film using intensifying screens.

Hybrid-selected translation

The cDNA inserts of recombinant λgtll clones were excised by EcoRI restriction endonuclease digestion and subcloned in the plasmids pUC8 or pUC9 (18). Plasmid preparations containing 0.1-0.5 µg of cDNA insert were linearized by digestion with the restriction endonuclease Bam Hl and spotted on untreated 0.5 cm^2 nylon membrane filters (Gene Screen, NEN). Filter-bound DNA was alkali-denatured, neutralized and washed as described (19). The dried filters were baked at 75°C for 1.5 hr. The filters were placed in a tube with sterile water, boiled for 1 min and chilled in an ice-water bath. The water was replaced by hybridization solution (50 mM piperazine-1,4-bis-2-ethane sulfonic acid (Pipes), pH 7.5, 0.5 M NaCl, 2 mM EDTA, 0.4% SDS, 30 % deionized formamide) and 1-4 mg total HeLa cellular RNA and the filters were incubated at 47°C for 8-16 hrs. The filters were then washed 10 times for 1 min in 1xSSC (30 mM Na-citrate, pH 7.5, 0.3 M NaCl) with 0.1 % SDS at 50°C followed by one wash for 1 min in 5 mM Tris-HCl, pH 7.5, 1 mM EDTA at 50°C. Hybridized RNA was eluted from the washed filters by two cycles of boiling for 1 min in 5 mM KCl, 10 µg/ml, yeast tRNA (Boehringer) followed by quick-freezing in liquid nitrogen. After thawing, the eluates were pooled and the RNA precipitated with ethanol at -20°C overnight. Hybrid-selected RNA was translated in a mRNA-dependent rabbit reticulocyte lysate as described (20). [^{35}S] methionine-labelled translation products were analyzed either directly by SDS-polyacrylamide gel electrophoresis (SDS-PAGE) (21) or after immunoprecipitation. Immunoprecipitation was carried out as described (22). For 15 µl translation mixture, 8-10 µg of monoclonal anti-elF-4A antibody and 20-30 µl of rabbit anti-mouse antibody (DAKO, 800 µl precipitates 1 mg mouse immunoglobulins)

in a total volume of 150 µl was used. After washing, the immunoprecipitate was dissolved in SDS-sample buffer, heated and applied to an acrylamide gel (21). Fixed gels were prepared for fluorography (23) and exposed to Kodak X-omat SO 282 film using intensifying screens at -70°C.

Peptide mapping

Partial proteolytic digestion of protein in excised gel pieces was performed according to Cleveland et al. (24). Following electrophoresis, gels were either silver-stained (25) or processed for fluorography as described above.

Northern blots

Poly(A)$^+$RNA (1 µg) was denatured by treatment with glyoxal (26), fractionated on agarose gels and transferred to Gene Screen filters (27). The filters were baked at 80°C for 1.5 hr, washed by boiling in 20 mM Tris-HCl, pH 7.8, for 5 min and then incubated at 45°C for 2-3 hrs in prehybridization buffer: 5xSSC, 0.2 % Ficoll, 0.2 % polyvinylpyrrolidone, 0.2 % BSA (Pentax fraction V), 0.05 % Na-pyrophosphate, 15 mM Na_2HPO_4, 0.1 % SDS, 50 % deionized formamide, 100 µg/ml denatured herring sperm DNA (SIGMA). The filters were then incubated at 45°C for 12-16 hrs with hybridization buffer (the same as prehybridization buffer plus 10 % dextran sulfate and [^{32}P]-labelled probe at 0.5-1.0 x 10^6 cpm/ml. For probes, cDNA-containing plasmids were labeled by nick-translation (12, specific activity approximately 5x10^7 cpm/µg). After hybridization, the filters were washed twice in 2xSSC at room temperature and once in 2xSSC with 0.1 % SDS at 45°C for 30 min. Single-stranded ribosomal RNA and DNA size standards were used to determine the length of the mRNA. For DNA standards, the plasmid pB40 (made by ligation of EcoRI-cut pBR322 with SV40) was digested with Hind III and the ends filled in with the Klenow fragment of E.coli DNA polymerase I in the presence of [^{32}]P-labeled deoxynucleotide triphosphates (12).

DNA sequencing

Cloned cDNAs were subcloned into M13mp8 or M13mp9 (28) either as entire EcoRI fragments or following digestion with the restriction endonucleases HindIII, PstI, HaeIII or Sau3A. Alternatively, cloned cDNAs were sonicated, end-repaired using T_4 DNA polymerase, size-fractionated on agarose gels and ligated

into the SmaI site of M13mp8 or M13mp9. Following transformation of E.coli JM 101, recombinant phage were identified (colorless plaques on plates containing the indicator 5-bromo-4-chlor-3-indolyl-β-D-galactoside and the inducer IPTG), single-stranded M13 DNA was isolated and sequenced using the dideoxy chain-termination method of Sanger et al. (29) as modified by Biggin et al. (30). Both strands of cloned cDNAs were completely sequenced.

RESULTS

Immuno-screening

Monoclonal antibodies directed against rabbit reticulocyte protein synthesis initiation factor 4A (eIF-4A) were used to isolate the corresponding cDNA sequences by expression in E. coli. Since several lines of evidence suggested that initiation factors from rabbit, mouse and man are structurally, functionally and immunologically very similar (31,32), we chose mouse hybridoma poly $(A)^+$RNA for cDNA synthesis. A mouse cDNA library representing approximately $2x10^6$ recombinants was constructed in the lamda expression vector λgt11 (16). Portions of the cDNA library were either immuno-screened directly as described in Materials and Methods or first amplified in E.coli Y1088 (17). We have isolated eight eIF-4A immunoreactive cDNA clones after screening approximately 1×10^6 recombinant phage with monoclonal anti-eIF-4A antibodies. cDNA insert sizes ranged from 120-2000 base pairs. One of the eIF-4A clones contains a cDNA insert of 1600 base pairs which is long enough to code for the entire 43 kilodaltons of eIF-4A. Consistent with this, the hybrid protein expressed by this clone is at least 40 kilodaltons longer than beta-galactosidase (not shown).

Hybrid-selected translation and peptide mapping

To verify by other than immunological criteria that cDNA inserts encoded eIF-4A gene sequences, hybrid-selected translation followed by peptide mapping of the translation products was performed. This is shown below for a cDNA encoding eIF-4A subcloned in the plasmid pUC8 (18). Fig. 1 shows that the cDNA selects from a crude mixture of mRNAs encoding many proteins (lane 1) an mRNA encoding a protein of the same molecular weight as eIF-4A (43 kilodaltons, lane 3). This protein can be

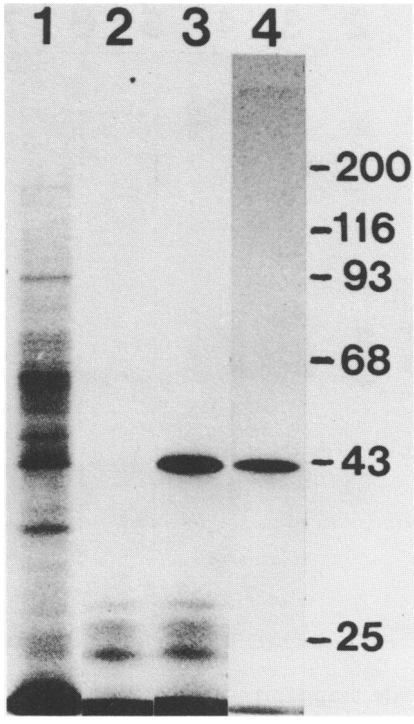


Figure 1. Hybrid-selected translation
The autoradiogram of a 12% acrylamide gel containing [^{35}S]-labeled translation products from a reticulocyte lysate is shown. The lysate was programmed with: lane 1, total HeLa cytoplasmic RNA; lane 2, hybrid-selected RNA hybridized to control plasmid DNA; lane 3, hybrid-selected RNA; lane 4, as for lane 3 except that the translation products were immunoprecipitated with the monoclonal anti-eIF-4A antibody as described in Materials and Methods. The position in the gel of the molecular weight markers myosin (200), beta-galactosidase (116), phosphorylase a (93), bovine serum albumin (68), ovalbumin (43) and chymotrypsin (25) are indicated in kilodaltons.

immunoprecipitated with monoclonal anti-eIF-4A antibody (lane 4). Plasmid DNA without insert does not select this mRNA (lane 2). Peptide mapping of the hybrid selection-derived 43 kilodalton polypeptide (Fig. 1, lane 3) and comparison of the resulting maps (Fig. 2, lanes 3,7) with maps generated from purified unlabelled eIF-4A (Fig. 2, lanes 1,2,6) and eIF-4A immunoprecipitated from total translation products (Fig. 2, lanes 4,8) confirmed that the cDNA insert contains eIF-4A sequences. The digestion patterns

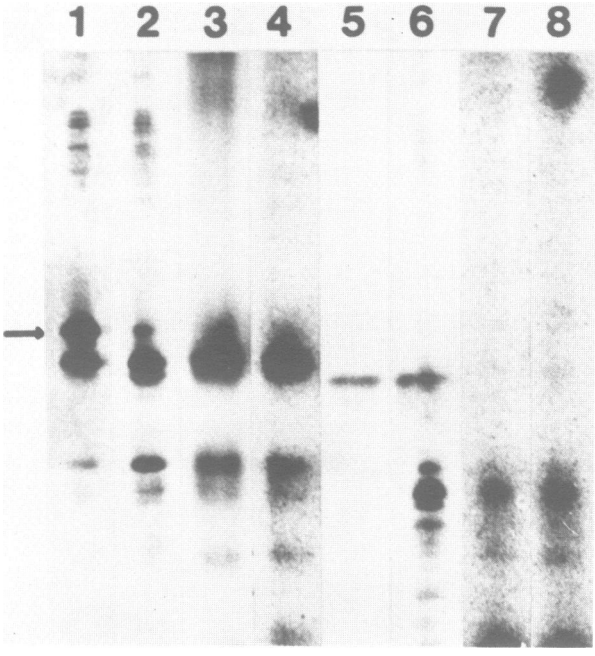


Figure 2. Peptide mapping
Bands containing purified unlabelled eIF-4A, [^{35}S]-labelled eIF-4A translated from hybrid-selected RNA and [^{35}S]-labelled eIF-4A immunoprecipitated from translation mixtures programmed with total HeLa RNA were excised from polyacrylamide gels, digested with S. aureus protease V8 and applied to a second polyacrylamide gel (24).
Lanes 1,2 and 6, unlabelled eIF-4A (1 µg), silver-stained; lanes 3 and 7, [^{35}S]-labelled eIF-4A translated from hybrid-selected RNA; lanes 4 and 8, [^{35}S]-labelled eIF-4A immunoprecipitated from translation mixtures programmed with total HeLa RNA. The amount of V8 protease varied between 1 ng (lane 1), 10 ng (lanes 2-4), and 100 ng (lanes 5-8). Lane 5 : 100 ng V8 protease, silver-stained. The position of undigested eIF-4A (arrow) was used to align the silver-stained lanes with the autoradiogram.

were nearly identical for all three preparations over a range of protease concentrations.

Northern blot analysis

A cDNA clone encoding eIF-4A sequences (see above) was used to identify mRNA encoding eIF-4A in poly $(A)^+$RNA from mouse liver and HeLa cells by the method of Northern blotting . Two bands of approximately 2000 and 1600 nucleotides in length were detected in both HeLa and mouse poly$(A)^+$ RNA (Fig. 3). The same result was

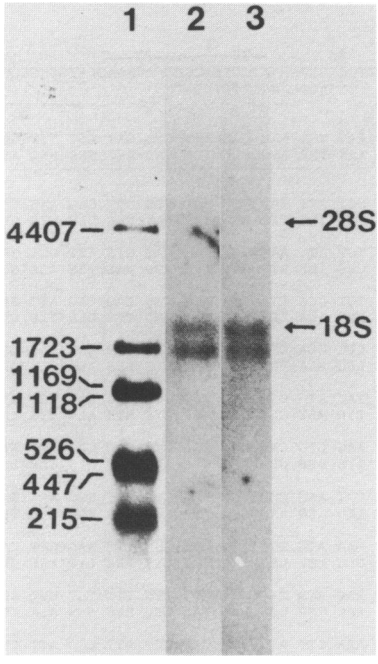
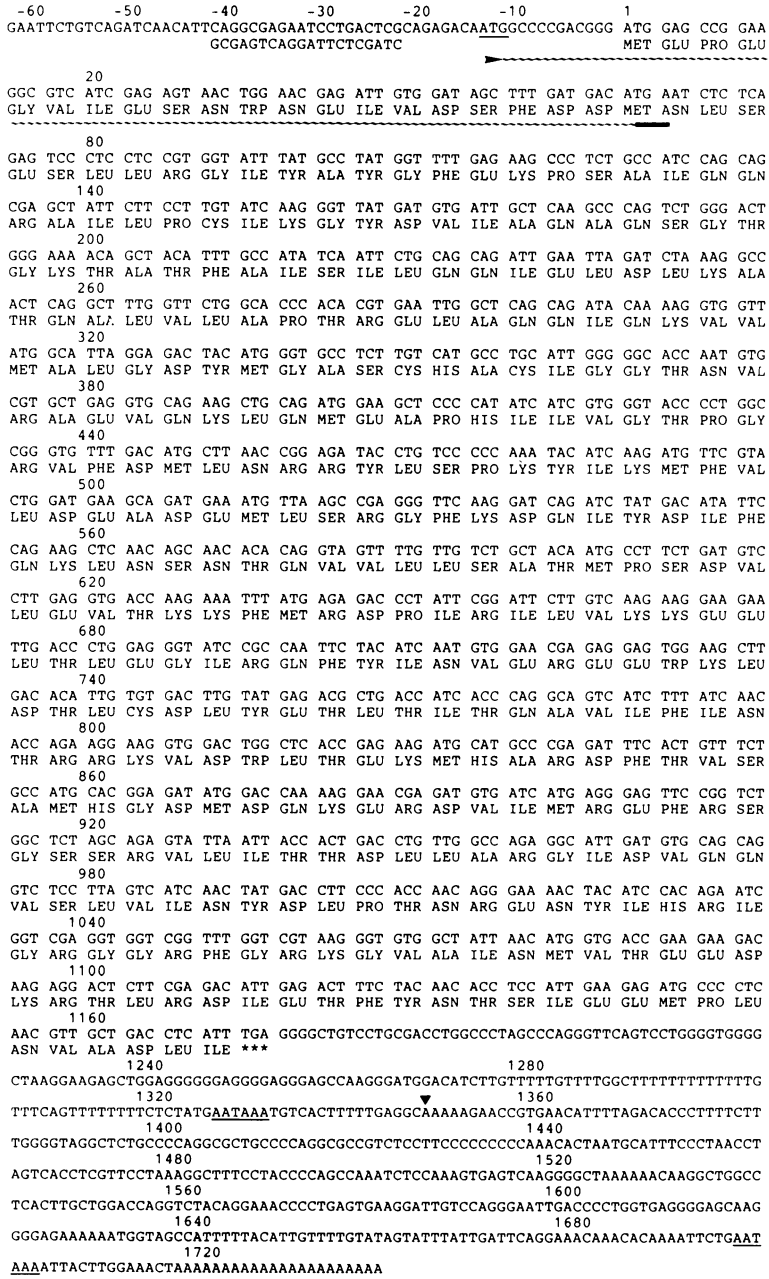


Figure 3. Northern blot
Cytoplasmic poly(A)-containing RNA from HeLa and mouse liver cells was denatured with glyoxal, fractionated on a 1.1% agarose gel, and transferred to Gene Screen. Plasmid DNA containing eIF-4A cDNA insert was labelled with [^{32}P] by nick-translation (12) and 25 ng (1.5×10^6 cpm) were hybridized to the blot in a total volume of 2 ml (see Materials and Methods). The resulting autoradiogram is shown. Lane 1, Hind III-digested pB40 fusion plasmid used as size markers (given in nucleotides); lane 2, 1 µg HeLa Poly(A)$^+$ cytoplasmic RNA, lane 3, 1 µg mouse liver poly(A)$^+$ cytoplasmic RNA. The positions of 18S and 28S ribosomal RNAs located by ethidium bromide staining of an adjacent lane are indicated.

obtained when total HeLa cytoplasmic RNA was used or when two other eIF-4A clones were used as hybridization probes (results not shown). In all cases the strength of hybridization was about equal for both bands.

Nucleotide sequence determination

cDNA clones encoding the short (about 1600 nucleotides) and long forms (about 2000 nucleotides) of eIF-4A mRNA isolated either by immuno-screening or by hybridization were sequenced as described in Materials and Methods. The nucleotide sequence of


-60 -50 -40 -30 -20 -10 1
GAATTCTGTCAGATCAACATTCAGGCGAGAATCCTGACTCGCAGAGACAATGGCCCCGACGGG ATG GAG CCG GAA
GCGAGTCAGGATTCTCGATC MET GLU PRO GLU
20
GGC GTC ATC GAG AGT AAC TGG AAC GAG ATT GTG GAT AGC TTT GAT GAC ATG AAT CTC TCA
GLY VAL ILE GLU SER ASN TRP ASN GLU ILE VAL ASP SER PHE ASP ASP MET ASN LEU SER
80
GAG TCC CTC CTC CGT GGT ATT TAT GCC TAT GGT TTT GAG AAG CCC TCT GCC ATC CAG CAG
GLU SER LEU LEU ARG GLY ILE TYR ALA TYR GLY PHE GLU LYS PRO SER ALA ILE GLN GLN
140
CGA GCT ATT CTT CCT TGT ATC AAG GGT TAT GAT GTG ATT GCT CAA GCC CAG TCT GGG ACT
ARG ALA ILE LEU PRO CYS ILE LYS GLY TYR ASP VAL ILE ALA GLN ALA GLN SER GLY THR
200
GGG AAA ACA GCT ACA TTT GCC ATA TCA ATT CTG CAG CAG ATT GAA TTA GAT CTA AAG GCC
GLY LYS THR ALA THR PHE ALA ILE SER ILE LEU GLN GLN ILE GLU LEU ASP LEU LYS ALA
260
ACT CAG GCT TTG GTT CTG GCA CCC ACA CGT GAA TTG GCT CAG CAG ATA CAA AAG GTG GTT
THR GLN ALA LEU VAL LEU ALA PRO THR ARG GLU LEU ALA GLN GLN ILE GLN LYS VAL VAL
320
ATG GCA TTA GGA GAC TAC ATG GGT GCC TCT TGT CAT GCC TGC ATT GGG GGC ACC AAT GTG
MET ALA LEU GLY ASP TYR MET GLY ALA SER CYS HIS ALA CYS ILE GLY GLY THR ASN VAL
380
CGT GCT GAG GTG CAG AAG CTG CAG ATG GAA GCT CCC CAT ATC ATC GTG GGT ACC CCT GGC
ARG ALA GLU VAL GLN LYS LEU GLN MET GLU ALA PRO HIS ILE ILE VAL GLY THR PRO GLY
440
CGG GTG TTT GAC ATG CTT AAC CGG AGA TAC CTG TCC CCC AAA TAC ATC AAG ATG TTC GTA
ARG VAL PHE ASP MET LEU ASN ARG ARG TYR LEU SER PRO LYS TYR ILE LYS MET PHE VAL
500
CTG GAT GAA GCA GAT GAA ATG TTA AGC CGA GGG TTC AAG GAT CAG ATC TAT GAC ATA TTC
LEU ASP GLU ALA ASP GLU MET LEU SER ARG GLY PHE LYS ASP GLN ILE TYR ASP ILE PHE
560
CAG AAG CTC AAC AGC AAC ACA CAG GTA GTT TTG TTG TCT GCT ACA ATG CCT TCT GAT GTC
GLN LYS LEU ASN SER ASN THR GLN VAL VAL LEU LEU SER ALA THR MET PRO SER ASP VAL
620
CTT GAG GTG ACC AAG AAA TTT ATG AGA GAC CCT ATT CGG ATT CTT GTC AAG AAG GAA GAA
LEU GLU VAL THR LYS LYS PHE MET ARG ASP PRO ILE ARG ILE LEU VAL LYS LYS GLU GLU
680
TTG ACC CTG GAG GGT ATC CGC CAA TTC TAC ATC AAT GTG GAA CGA GAG GAG TGG AAG CTT
LEU THR LEU GLU GLY ILE ARG GLN PHE TYR ILE ASN VAL GLU ARG GLU GLU TRP LYS LEU
740
GAC ACA TTG TGT GAC TTG TAT GAG ACG CTG ACC ATC ACC CAG GCA GTC ATC TTT ATC AAC
ASP THR LEU CYS ASP LEU TYR GLU THR LEU THR ILE THR GLN ALA VAL ILE PHE ILE ASN
800
ACC AGA AGG AAG GTG GAC TGG CTC ACC GAG AAG ATG CAT GCC CGA GAT TTC ACT GTT TCT
THR ARG ARG LYS VAL ASP TRP LEU THR GLU LYS MET HIS ALA ARG ASP PHE THR VAL SER
860
GCC ATG CAC GGA GAT ATG GAC CAA AAG GAA CGA GAT GTG ATC ATG AGG GAG TTC CGG TCT
ALA MET HIS GLY ASP MET ASP GLN LYS GLU ARG ASP VAL ILE MET ARG GLU PHE ARG SER
920
GGC TCT AGC AGA GTA TTA ATT ACC ACT GAC CTG TTG GCC AGA GGC ATT GAT GTG CAG CAG
GLY SER SER ARG VAL LEU ILE THR THR ASP LEU LEU ALA ARG GLY ILE ASP VAL GLN GLN
980
GTC TCC TTA GTC ATC AAC TAT GAC CTT CCC ACC AAC AGG GAA AAC TAC ATC CAC AGA ATC
VAL SER LEU VAL ILE ASN TYR ASP LEU PRO THR ASN ARG GLU ASN TYR ILE HIS ARG ILE
1040
GGT CGA GGT GGT CGG TTT GGT CGT AAG GGT GTG GCT ATT AAC ATG GTG ACC GAA GAA GAC
GLY ARG GLY GLY ARG PHE GLY ARG LYS GLY VAL ALA ILE ASN MET VAL THR GLU GLU ASP
1100
AAG AGG ACT CTT CGA GAC ATT GAG ACT TTC TAC AAC ACC TCC ATT GAA GAG ATG CCC CTC
LYS ARG THR LEU ARG ASP ILE GLU THR PHE TYR ASN THR SER ILE GLU GLU MET PRO LEU
1160
AAC GTT GCT GAC CTC ATT TGA GGGGCTGTCCTGCGACCTGGCCCTAGCCCAGGGTTCAGTCCTGGGGTGGGG
ASN VAL ALA ASP LEU ILE ***
1240 1280
CTAAGGAAGAGCTGGAGGGGGGAGGGGAGGGAGCCAAGGGATGGACATCTTGTTTTTGTTTTGGCTTTTTTTTTTTTG
1320 1360
TTTCAGTTTTTTTTCTCTATGAATAAATGTCACTTTTTGAGGCAAAAAGAACCGTGAACATTTTAGACACCCTTTTCTT
1400 1440
TGGGGTAGGCTCTGCCCCAGGCGCTGCCCCAGGCGCCGTCTCCTTCCCCCCCCCAAACACTAATGCATTTCCCTAACCT
1480 1520
AGTCACCTCGTTCCTAAAGGCTTTCCTACCCCAGCCAAATCTCCAAAGTGAGTCAAGGGGCTAAAAAACAAGGCTGGCC
1560 1600
TCACTTGCTGGACCAGGTCTACAGGAAACCCCTGAGTGAAGGATTGTCCAGGGAATTGACCCCTGGTGAGGGGAGCAAG
1640 1680
GGGAGAAAAAATGGTAGCCATTTTTACATTGTTTTGTATAGTATTTATTGATTCAGGAAACAAACACAAAATTCTGAAT
1720
AAAATTACTTGGAAACTAAAAAAAAAAAAAAAAAAAAA

the long form of eIF-4A mRNA (Fig. 4, upper line) begins with an EcoRI endonuclease restriction site and extends 1778 nucleotides to the poly(A) tail. The 5' end of the long form of eIF-4A was located by primer extension experiments (not shown) and found to be about 190 nucleotides upstream of the EcoRI site. The sequence contains a single open reading frame long enough to code for eIF-4A (Fig. 4, lower line). The start of translation is ambiguous due to two in-frame AUG start codons separated by 57 nucleotides (position +1 and +61). Comparison of the published molecular weights and amino acid composition of eIF-4A (8) with the proteins synthesized from the two possible AUG start codons does not permit the identification of the true start site. Several attempts to sequence the N-terminus of purified eIF-4A by manual Edman degradation were unsuccessful. In addition to the two in-frame AUGs, the mRNA for eIF-4A contains a third AUG which is 14 nucleotides upstream (position -14) and out-of-frame with respect to the long open reading frame. This AUG at position -14 is followed by a short open reading frame of 72 nucleotides which ends at the UGA codon (position +62) which overlaps with the second of the two in-frame AUG codons (position +61).

The nucleotide sequence of the short form of eIF-4 mRNA differs from the long form sequence only in the 5' and 3' untranslated regions (Fig. 4). The 175 nucleotide 3' untranslated region is identical to the first part of the 3' untranslated region of the long form except for the stretch of T's beginning at position 1290. Two cDNA clones corresponding to the long form of elF-4A mRNA contained 13 T's and one cDNA clone of the short mRNA form contained 14 T's. We have not yet determined whether

Figure 4. eIF-4A cDNA sequences
The nucleotide sequence of mouse cDNA corresponding to the long form of eIF-4A mRNA is shown. The nucleotide sequence of the 5' end of cDNA corresponding to the short form of eIF-4A mRNA differs from the long form sequence upstream of position -23 and is shown under the long form sequence. The polyadenylation site of the short form mRNA is marked by the large arrow head at position 1347. Possible polyadenylation signals are underlined (positions 1325 and 1696). The predicted amino acid sequence for eIF-4A is shown beginning at the first of the two in-frame ATG's. The upstream out-of-frame ATG at position -14 is underlined. The wavy line marks the open reading frame between the ATG at position -14 (►) and the TGA at position +62 (▬). For details, see text.

this difference is due to transcription from different elF-4A genes or is an artefact of cDNA synthesis or cloning. When cDNA sequences corresponding to the 3' untranslated region of the long form were used as probe in Northern hybridization experiments only the upper of the two eIF-4A specific mRNAs was labeled (not shown). The 5' end of the short form was located by primer extension experiments between position -40 and -50. The sequence between position -22 and +1 is identical for both the long and short forms and includes the out-of-frame AUG at position -14.

DISCUSSION

This report describes the isolation of cDNA clones encoding eukaryotic initiation factor 4A (eIF-4A) gene sequences. eIF-4A is a rather abundant protein representing 0.3-0.5% of total protein in HeLa cells (33). We obtained eight eIF-4A cDNA clones by immuno-screening of approximately 10^6 recombinant λgtll phage (carrying HeLa and mouse cDNA inserts) with a monoclonal anti-eIF-4A antibody. Hybrid-selected translation and identification of the translation product by immunoprecipitation and peptide mapping were used to confirm that the cloned cDNA sequences encode the initiation factor eIF-4A (Fig. 1 and 2). By Northern blot analysis we detected two discrete bands suggesting the existence of two mRNAs encoding eIF-4A in HeLa and mouse cells (Fig. 3). Nucleotide sequence determination of cDNA clones proved the presence of two mRNAs with identical coding sequences differing in their 5' and 3' untranslated regions (Fig. 4). We do not know whether these two mRNAs are encoded by two different genes although genomic DNA digested with any one of several different restriction enzymes and hybridized on Southern blots with DNA probes derived from the coding region of eIF-4A give multiple bands.

Of special interest is the initiation region of eIF-4A mRNAs since it contains three potential AUG start codons at positions -14, +1 and +61 (Fig. 4). This region is identical in the two mRNAs. By the criteria worked out by M. Kozak (34) the AUG at position -14 is an optimal start codon with respect to its flanking sequences. However, efficient initiation of translation at this codon would not allow translation of the long open reading frame encoding eIF-4A unless the ribosomes would re-

initiate at the third AUG (position +61) after reaching the UGA stop codon which overlaps the third AUG. Re-initiation at the third AUG codon would lead to the synthesis of an eIF-4A molecule consisting of 370 amino acids and having a molecular weight of 42,130 daltons. Alternatively, for reasons which are not apparent the AUG at position -14 could be an inefficient translation start codon and translation might begin at the AUG in position +1. This would lead to the synthesis of an eIF-4A molecule with 390 amino acids and a molecular weight of 44,435 daltons. It is even possible that both types of eIF-4A molecules are made. These speculations are particularly intriguing in light of previous observations which show that (a) part of the cytoplasmic eIF-4A is found associated with other polypeptides in a cap-binding factor complex (9,10), (b) the complex-bound eIF-4A differs slightly (as shown by peptide mapping) from free eIF-4A (10) and (c) the complex-bound and free eIF-4A have different functions in initiation of protein synthesis (11).

In vitro translation experiments with eIF-4A mRNA will help to answer the open questions concerning eIF-4A mRNA translation.

ACKNOWLEDGEMENTS

We thank C. Handschin for expert technical assistance, M. Hümbelin for monoclonal anti-eIF-A antibody and T. Staehelin for C19 hybridoma cells. This work was supported by the grant Nr. 3.117-0.81 from the Swiss National Science Foundation to H.T.

[3]To whom reprint requests should be sent

[4]Present address: Max Planck Institut für Immunologie Stübeweg 51, D-7800 Freiburg, FRG

REFERENCES

1. Safer, B. and Anderson, W.F. (1978) Critical Reviews in Biochemistry, 261-290, CRC Press, Ohio (Cleveland).
2. Revel, M. and Groner, Y. (1978) Ann. Rev. Biochem. 47, 1079-1126.
3. Kramer, G. and Hardesty, B. (1980) Cell Biology 4, 69-105.
4. Maitra, U., Stringer, E.A. and Chaudhuri, A. (1982) Ann. Rev. Biochem. 51, 869-900.
5. Perez-Bercoff, R., ed. (1982) Protein Biosynthesis in Eukaryotes, Plenum Press, New York and London.

6. Trachsel, H., Erni, B., Schreier, M.H. and Staehelin, T. (1977) J. Mol. Biol. 116, 755-767.
7. Benne, R., Luedi, M. and Hershey, J.W.B. (1977) J. Biol. Chem. 252, 5798-5803.
8. Grifo, J.A., Tahara, S.M., Leis, J.P., Morgan, M.A., Shatkin, A.J. and Merrick, W.C. (1982) J. Biol. Chem. 257, 5246-5252.
9. Grifo, J.A., Tahara, S.M., Morgan, M.A., Shatkin, A.J. and Merrick, W.C. (1983) J. Biol. Chem. 258, 5804-5810.
10. Edery, I., Hümbelin, M., Darveau, A., Lee, K.A.W., Milburn, S., Hershey, J.W.B., Trachsel, H. and Sonenberg, N. (1983) J. Biol. Chem. 258, 11398-11403.
11. Daniels-McQueen, S., Detjen, B.M., Grifo, J.A., Merrick, W.C. and Thach, R.E. (1983) J. Biol. Chem. 258, 7195-7199.
12. Maniatis, T., Fritsch, E.F., Sambrook, J. (1982) Molecular cloning; A laboratory manual, Cold Spring Harbor Laboratory.
13. Schibler, U., Tosi, M., Pittet, A.C., Fabiani, L. and Wellauer, P.K. (1980) J. Mol. Biol. 142. 93-116.
14. Buell, G.N., Wickens, M.P., Payvar, F. and Schimke, R.T. (1978) J. Biol. Chem. 253, 2471-2482.
15. Pheiffer, B.H. and Zimmermann, S.B. (1983) Nucl. Acid. Res. 11, 7853-7871.
16. Young, R.A. and Davis, R.W. (1983) Proc. Natl. Acad. Sci. USA 80, 1194-1198.
17. Young, R.A. and Davis, R.W. (1983) Science 222, 778-782.
18. Vieira, J. and Messing, J. (1982) Gene 19, 259-268.
19. Miller, J.S., Paterson, B.M., Ricciardi, R.P., Cohen, L. and Roberts, B.E. (1983) Meth. Enzymol. 101, 650-674.
20. Pelham, H.R.B. and Jackson, R.J. (1976) Eur. J. Biochem. 67, 247-256.
21. Anderson, C.W., Baum, P.R. and Gesteland, R.F. (1973) J. Virol. 12, 241-252.
22. Hughes, E.N. and August, J.T. (1981) J. Biol. Chem. 256, 664-671.
23. Chamberlain, J.P. (1979) Anal. Biochem. 98, 132-135.
24. Cleveland, D.W., Fischer, S.G., Kirschner, M.W. and Laemmli, U.K. (1977) J. Biol. Chem. 252, 1102-1106.
25. Oakley, B.R., Kirsch, D.R. and Morris, N.R. (1980) Anal. Biochem. 105, 361-363.
26. McMaster, G.K. and Carmichael, G.G. (1977) Proc. Natl. Acad. Sci. USA 74, 4835-4838.
27. Thomas, P.S. (1983) Meth. Enzymol. 100, 255-266.
28. Messing, J., and Vieira, J. (1982) Gene 19, 269-276.
29. Sanger, F., Nicklen, S., and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463-5467.
30. Biggin, M.d., Gibson, T.J. and Hong, G.F. (1983) Proc. Natl. Acad. Sci. USA 80, 3963-3965.
31. Trachsel, H., Erni, B., Schreier, M.H., Braun, L. and Staehelin, T. (1979) Biochim. Biophys. Acta 561, 484-490.
32. Meyer, L.J., Milburn, S.C. and Hershey, J.W.B. (1982) Biochem. 21, 4206-4212.
33. Duncan, R. and Hershey, J.W. B. (1983) J. Biol. Chem. 258, 7228-7235.
34. Kozak, M. (1981) Nucl. Acids Res. 9, 5233-5252.