Identification of cDNA Clones for the Large Subunit of Eukaryotic Translation Initiation Factor 3

COMPARISON OF HOMOLOGUES FROM HUMAN, NICOTIANA TABACUM, CAENORHABDITIS ELEGANS, AND SACCHAROMYCES CEREVISIAE*

(Received for publication, March 19, 1996, and in revised form, December 12, 1996)

Keith R. Johnson‡, William C. Merrick§, Wendy L. Zoll§, and Yingxun Zhu¶

From the Department of Biology, University of Toledo, Toledo, Ohio 43606-3390 and the §Department of Biochemistry, School of Medicine, Case Western Reserve University, Cleveland, Ohio 44106-4935

Initiation of translation in eukaryotes is mediated by a set of initiation factors. Mammalian initiation factor 3 is composed of at least 8 subunits, with the largest being about 180 kDa in size. Here we report the cloning of the p180 subunit of human eukaryotic translation initiation factor (eIF) 3. The amino acid sequence deduced from the cDNA agrees with the sequences of CNBr fragments of eIF-3, confirming the identity of the clone. The 1382 amino acid open reading frame contains a high percentage of charged residues (48%) and an unusual repetitive domain near the carboxyl terminus composed of 25 repeats of 10 amino acids each. Data base searches identified related sequences found in members of the plant and fungal kingdoms as well as in other mammals and the nematode Caenorhabditis elegans. These sequences share significant identity with the human clone and probably represent the homologues of the p180 subunit in these organisms. This is the first report identifying the sequence of the large subunit of eIF-3.

Eukaryotic translation initiation factor (eIF) 3 is the largest of the protein synthesis initiation factors, with a size of about 650 kDa. eIF-3 purified from rabbit reticulocyte lysate consists of at least eight individual polypeptide chains (1, 2). Originally, eIF-3 was identified as a factor that binds to the 40 S ribosomal subunit and thereby prevents the association of the 40 and 60 S subunits with one another. This results in a pool of 40 S subunits, which are then able to participate in the initiation process. However, eIF-3 has also been implicated in a number of additional roles. One example is the association of eIF-3 with eIF-4F, where the interaction is sufficiently stable to separate the two initiation factors (3). eIF-3 interacts with a number of initiation factors in addition to eIF-4F. This suggests that eIF-3 may be the major protein which aligns the factors so that the mRNA is correctly positioned for initial binding to the 40 S subunit and the subsequent identification of the initiating AUG.

Of the many associations of eIF-3 with other translational components, most appear to be via the p180 subunit. The ability of eIF-3 to bind to and stabilize the ternary complex (eIF-2-GTP-Met-tRNAi) is dependent on the p180 subunit, since preparations that have been depleted of p180 fail to promote the formation of the ternary complex (4). The p180 subunit of eIF-3 has also been shown to interact with eIF-4B using "Far Western" blotting. As noted above, eIF-3 interacts with eIF-4F. The site of interaction between eIF-3/eIF-4F has been mapped to the middle region (amino acids 480–886) of the γ subunit of eIF-4F (5), but the binding site in eIF-3 has not yet been identified. The limited biochemical evidence indicates that the p180 subunit is important for all of these interactions and processes.

Many mammalian factors can replace their yeast counterparts in vivo (1), suggesting extensive homology between the factors. For eIF-3, the only detailed comparison that can be made is between mammalian and wheat germ eIF-3, and, in this instance, the extensive similarity seen between the translation factors appears to break down (1). The differences in p1 and apparent molecular weight have made it difficult to compare eIF-3 from different sources or infer functions of the subunits. At present, much of what is known about eukaryotic protein synthesis has been learned from fractionated, cell-free systems. With the availability of cloned genes and cDNAs, it has been possible to begin to manipulate the translation factors either by altering the levels of expression or by site-directed mutagenesis. To date, cDNA clones have been obtained for all the known initiation factors except eIF-3 and eIF-6. Just recently, cDNA clones have been identified for two subunits of the yeast Saccharomyces cerevisiae eIF-3 (6–9). As an initial effort to better understand the role and regulation of eIF-3, we have obtained full-length cDNA clones for the p180 subunit of human eIF-3. Data base searches revealed sequences from S. cerevisiae, the nematode Caenorhabditis elegans, and the plant Nicotiana tabacum with significant similarity to the human cDNA. The sequences likely represent the corresponding subunits of eIF-3 in these organisms.

MATERIALS AND METHODS

Reagents and Antibodies—Hybridoma 116, a cell fusion between a rat spleen cell and mouse myeloma Sp2/0-Ag14 (10), was isolated while searching for monoclonal antibodies (mAb) against cadherin-related...
proteins. The immunoglobulin secreted by this hybridoma is called mAb 116. Acrylamide and bisacrylamide were from Bio-Rad; molecular biology grade urea and DNA size markers were from Life Technologies, Inc. Size markers and other electrophoresis reagents were from Sigma. [α-32P]dCTP and [α-35S]dATP were purchased from Amersham Corp. Tran35S-label™ was from ICN Biochemicals. The human liver, human keratinocyte, and HL-60 cell (11) cDNA libraries in λgt11 were from Clontech (Palo Alto, CA). Fetal calf serum was purchased from HyClone Laboratories (Logan, UT). Other tissue culture reagents were from Sigma. Components of bacteriological media were purchased from Difco. Fluorescein-conjugated anti-rat IgG was from Organon-Teknika (Duram, NC).

Cell Culture, Preparation of Cell Extracts, and Immunofluorescence—Normal human keratinocytes were isolated and cultured as described (12). JAR choriocarcinoma cells (13) and A-431 cervical carcinoma cells (14) were cultured in Dulbecco’s modified Eagle’s medium supplemented with penicillin/streptomycin and 10% fetal calf serum. JAR choriocarcinoma cells (13) and A-431 cervical carcinoma cells (14) were cultured in Dulbecco’s modified Eagle’s medium supplemented with penicillin/streptomycin and 10% fetal calf serum.

For metabolic labeling, JAR cells were starved in methionine-free medium for 4 h and labeled for 2 h with Tran35S-label™ prior to extraction with Nonidet P-40. A preparation of mAb 116 was added to the cleared lysate, and immune complexes were collected with anti-rat-conjugated Sepharose™. The proteins present in the immunoprecipitate were resolved on a 7% SDS gel. The gel was processed for fluorography and exposed to Kodak X-Omat AR film at −70 °C. No distinct band corresponding to p180 appeared in the pulse labeled cells; instead, a diffuse set of proteins migrating between about 200 and 160 kDa were seen. Bands that specifically co-immunoprecipitated with mAb 116 are indicated with asterisks; other bands visible in the autoradiogram were present in control immunoprecipitations with irrelevant mAbs (data not shown). The estimated sizes of the co-immunoprecipitating bands are 116, 60, 47, 37, and 35 kDa. Lane 3, the proteins present in an Nonidet P-40 extract of JAR cells were resolved on a 7% SDS-polyacrylamide gel, transferred to nitrocellulose, and incubated with mAb 116. Incubation with alkaline-phosphatase-conjugated secondary antibody, followed by treatment with phosphatase substrates, revealed a strong signal at approximately 160 kDa. The other bands are likely to be breakdown products of the 160-kDa band.

Clone Isolation, Fusion Protein Purification, Antiserum Production, and Sequencing—Screening λgt11 libraries with mAb 116 was performed with slight modifications (15) to established procedures (16). Lysogens were lyed by repeated freeze/thaw, extracted with Nonidet P-40, and debris removed by centrifugation. Aliquots of clarified extract were passed through an anti-β-galactosidase column. Bound material was eluted with 50 mM diethylamine, dialyzed against phosphate-buffered saline, and injected into a rabbit using standard techniques (17). Screening with 32P-labeled probes was performed as described (18). Phage clones were precipitated from plate lysates with ammonium sulfate (19). Subcloning of inserts into pUC18/19 (20), exonuclease III deletions, and sequencing were performed as described (21). Sequence comparisons were made using the facilities at the National Center for Biotechnology Information and the European Molecular Biology Laboratory (GenBank release 92). Sequence alignments were made using the PILEUP and GAP programs of the University of Wisconsin Genetics Computer Group. Consensus sequences were identified using the

FIG. 1. Immunofluorescence localization of the antigen. A confluent monolayer of JAR (human choriocarcinoma) cells was fixed, permeabilized, and stained with mAb 116. The cells were then treated with fluorescein-conjugated anti-rat IgG and viewed with epifluorescence microscopy. The antigen is excluded from the periphery of the cells as well as the nuclei. A similar distribution of the antigen was also seen in other cell types, both normal and transformed (data not shown).

FIG. 2. Immunoblot and immunoprecipitation of the antigen. Lane 1 contained 14C-labeled molecular size markers. The relative molecular sizes are given to the left of lane 1. Lane 2, confluent monolayers of JAR cells were incubated for 2 h with Tran35S-label™ prior to extraction with Nonidet P-40. A preparation of mAb 116 was added to the cleared lysate, and immune complexes were collected with anti-rat-conjugated Sepharose™. The proteins present in the immunoprecipitate were resolved on a 7% SDS gel. The gel was processed for fluorography and exposed to Kodak X-Omat AR film at −70 °C. No distinct band corresponding to p180 appeared in the pulse labeled cells; instead, a diffuse set of proteins migrating between about 200 and 160 kDa were seen. Bands that specifically co-immunoprecipitated with mAb 116 are indicated with asterisks; other bands visible in the autoradiogram were present in control immunoprecipitations with irrelevant mAbs (data not shown). The estimated sizes of the co-immunoprecipitating bands are 116, 60, 47, 37, and 35 kDa. Lane 3, the proteins present in an Nonidet P-40 extract of JAR cells were resolved on a 7% SDS-polyacrylamide gel, transferred to nitrocellulose, and incubated with mAb 116. Incubation with alkaline-phosphatase-conjugated secondary antibody, followed by treatment with phosphatase substrates, revealed a strong signal at approximately 160 kDa. The other bands are likely to be breakdown products of the 160-kDa band.

FIG. 3. Summary of the cDNA clones. The central heavy line represents the 5.3-kb full-length sequence. There are three internal EcoRI sites whose positions are indicated by the letter E. The predicted positions of the first and stop codons are indicated by the letters M and X, respectively. Above the heavy line, two antibody-positive clones are shown (ab+). Below the heavy line, three of the clones isolated during the screening process are shown. In each clone the positions of the internal EcoRI sites are indicated with vertical lines. The EcoRI fragments that were completely sequenced on both strands are indicated with asterisks. The 5′ sequence up to the first EcoRI site was derived from a clone found in a human keratinocyte cDNA library; the rest of the sequence was derived from clones isolated from a HL-60 cell cDNA library. PROSITE data base (22) available on the EMBL server. 3 Molecular weight and isoelectric point calculations were performed using the ExPaSy molecular biology server of the Geneva University Hospital.

Internet address: http://www.ebi.ac.uk/searches/prosite_input.html.
Polyacrylamide Gel Electrophoresis, Immunoblotting, and Immunoprecipitation—Proteins were resolved on SDS-polyacrylamide gels (23). For immunodetection, proteins resolved on SDS gels were transferred to nitrocellulose (24). Blots were blocked with bovine serum albumin and then incubated sequentially with primary and alkaline-phosphatase-conjugated secondary antibodies. Positive bands were detected with nitro blue tetrazolium and 5-bromo-4-chloro-3-indolyl phosphate. Immunoprecipitations were performed as described (25). Protein molecular size markers and their designated molecular masses were: myosin, 205 kDa; β-galactosidase, 116 kDa; phosphorylase b, 97 kDa; bovine serum albumin, 68 kDa; ovalbumin, 45 kDa; and carbonic anhydrase, 30 kDa.

Purification of eIF-3—eIF-3 was purified from the 0.5M KCl salt fraction of the rat liver cytosol as described (26,27).

Cloning the Large Subunit of Human eIF-3

The predicted start (nucleotides 114–116) and stop (nucleotides 4260–4262) codons and two potential polyadenylation signals are capitalized and underlined. In the 3' flanking region, six ATTTA sequences are underlined. The sequence has been deposited in the data bases with accession no. U78311. Dots indicate identity between the mouse and human sequences. Dashes have been inserted in these sequences in order to maintain the best alignment between them. The mouse cDNA sequence was translated in several reading frames; positions where the reading frame was changed are indicated with an asterisk in the mouse sequence. Seven reading frame changes were introduced. In the human sequence, the three underlined regions correspond to the NH2-terminal amino acid sequences of peptides derived from rabbit eIF-3 (see Table I). The repetitive domains (residues 925–1172 in the human sequence) are in lowercase. 736 residues are aligned between the human and mouse sequences; of these, 693 (94%) are identical.

Fig. 4. Nucleotide and deduced amino acid sequences. A, the accumulated cDNA sequence of the large subunit of human eIF-3 is shown. The predicted start (nucleotides 114–116) and stop (nucleotides 4260–4262) codons and two potential polyadenylation signals are capitalized and underlined. In the 3' flanking region, six ATTTA sequences are underlined. The sequence has been deposited in the data bases with accession no. U78311. Dots indicate identity between the mouse and human sequences. Dashes have been inserted in these sequences in order to maintain the best alignment between them. The mouse cDNA sequence was translated in several reading frames; positions where the reading frame was changed are indicated with an asterisk in the mouse sequence. Seven reading frame changes were introduced. In the human sequence, the three underlined regions correspond to the NH2-terminal amino acid sequences of peptides derived from rabbit eIF-3 (see Table I). The repetitive domains (residues 925–1172 in the human sequence) are in lowercase. 736 residues are aligned between the human and mouse sequences; of these, 693 (94%) are identical.

5 Internet address: http://expasy.hcuge.ch/ch2d/pi_tool.html.
washed of rabbit reticulocyte polysomes as described previously (3). Briefly, the steps involved in purification were: batch chromatography on phosphocellulose (150–450 mM KCl, Whatman P-11 phosphocellulose), sucrose density-gradient centrifugation in 500 mM KCl, gradient elution from DEAE-cellulose, and gradient elution from phosphocellulose. eIF-3 activity was monitored using the hemoglobin synthesis assay and protein purity was monitored using SDS-gel electrophoresis of column fractions.

Amino Acid Sequencing of eIF-3—Preparations of eIF-3 were adjusted to 70% formic acid, and a 100-fold excess of CNBr was added to the solution. After an overnight incubation in the dark at room temperature, additional CNBr was added and the incubation continued for a total of 24 h. The sample was then diluted with 10 volumes of water and evaporated to dryness. Twice the sample was rehydrated and dried. Following the third lyophilization, the sample was dissolved in SDS sample buffer with heating at 90 °C for 10 min. Peptide fragments were then resolved by SDS-gel electrophoresis in an 18% gel. Peptide fragments were transferred to a polyvinylidene difluoride membrane (Immobilon™, Millipore Corp., Bedford, MA), stained with Coomassie Blue, and destained with 7% acetic acid, 5% methanol. After extensive destaining to remove both unbound Coomassie Blue and any Tris and glycine that had been cut out and subjected to protein sequencing. Amino acid sequencing was performed with an Applied Biosystems model 477A microsequencer with on-line phenylthiohydantoin analysis in the Molecular Biology Core Laboratory at Case Western Reserve University.

RESULTS

mAb 116 was isolated while attempting to produce antibodies against cadherin-related proteins. A protein fraction containing E-cadherin and its associated proteins was used as the antigen. The mAb was used for immunofluorescence of JAR epithelial cells (Fig. 1), where it recognized a cytoplasmic antigen. mAb 116 also recognized a cytoplasmic antigen in other normal and transformed cells (data not shown). In each case, there was a conspicuous lack of signal both in the nucleus and at the periphery of the cells. Immunoblots of proteins separated by SDS-PAGE showed that mAb 116 recognized a protein of about 160 kDa in cell extracts (Fig. 2, lane 3). The faster migrating bands are presumed degradation products of the 160-kDa band (see also Fig. 5). Immunoprecipitation using extracts of 35S-labeled cells revealed only a diffuse signal in the high molecular mass portion of the gel. However, several distinct bands whose migration on SDS gels suggested sizes of approximately 116, 60, 47, 37, and 35 kDa (Fig. 2, lane 2) were co-immunoprecipitated, suggesting that the 160-kDa antigen was part of a protein complex.

In order to further characterize the antigen, a human liver cDNA library was screened with mAb 116. Two positive clones were found in the approximately 2 x 10^6 plaques that were screened. Lysogens of both clones were prepared in Escherichia coli Y1089 (16). mAb 116 recognized fusion proteins of about 170 kDa in extracts of the lysogens; the synthesis of these proteins was induced with isopropyl-1-thio-b-D-galactopyranoside (data not shown). The b-galactosidase fusion protein was prepared from one of the clones by affinity chromatography on an anti-β-galactosidase column, and rabbit antiserum was raised against the fusion protein. This antiserum reacted with a single protein in immunoblots of extracts of tissue culture cells; this protein co-migrated with the antigen recognized by mAb 116 (data not shown). These data strongly suggested that the cDNA clones encoded the antigen recognized by mAb 116.

The inserts in the two antibody-positive clones were removed from purified bacteriophage Agt11 DNA and subcloned into pUC19. Sequencing the ends of the clones revealed that the clones overlapped (Fig. 3). Data base searches showed that the
two sequences were unique. Additional clones were isolated from hybridization screens of several cDNA libraries. The longest clone (4.3 kb, Fig. 3) was found in an HL-60 cell library. The 2.7- and 1.5-kb EcoRI fragments from this clone were completely sequenced on each strand, using a combination of exonuclease III deletions and restriction fragments. When the 5’ and 3’ ends of the insert in this clone were sequenced using Agt11 primers, it was seen that the 5’ and 3’ EcoRI sites were authentic cDNA sites; thus, the clone was not full-length on either end. An EcoRI fragment from an independent HL-60 clone extended to a polyadenylation consensus sequence (Fig. 3), but no clone was identified that extended further in the 5’ direction. A randomly primed human keratinocyte library was screened with a probe from the 5’ direction terminated soon after the 3’ end of the 2.7-kb EcoRI fragment. This screen resulted in the isolation of a number of clones that extended further 5’. The longest of these was chosen for sequencing.

The resulting composite cDNA sequence (Fig. 4A) is 5256 nucleotides long. The first ATG in the sequence is predicted to be the start codon. It is preceded by an in-frame stop codon and inaugurates a 1382-amino acid open reading frame (ORF). This ATG is preceded by a purine in the −3 position, so it has some of the characteristics of an optimal start codon (26). The 5’ untranslated region (UTR) in this clone is thus predicted to be 113 nucleotides in length. The predicted stop codon is quickly followed by additional stop codons in all three reading frames. The resulting 3’-UTR is about 1 kb in length. Six ATTTA sequences that can be associated with rapid turnover of mRNA (27, 28) are present in this region, as are two potential polyadenylation signals (AATAAA; Ref. 29). The clone that extended furthest in the 3’ direction terminated soon after the second potential polyadenylation signal. The deduced amino acid sequence (Fig. 4B) represents a protein of 166 kDa and pI of 6.4 that is highly charged; 664 residues (48%) are acidic (Asp and Glu) or basic (Lys and Arg). There is an unusual repetitive domain near its carboxyl terminus (amino acids 925–1172, shown in lowercase letters in Fig. 4B). This is addressed further under “Discussion.”

The complex identified by mAb 116 had a composition similar to that reported for mammalian eIF-3 (1). When a purified preparation of rabbit reticulocyte eIF-3 (see “Materials and Methods”) was probed with the antiserum raised against the β-galactosidase fusion protein, the serum recognized the p180 band in the eIF-3 preparation. This reticulocyte protein co-migrated with the band recognized by the antiserum in extracts of human cultured cells (Fig. 5).

In order to confirm that the cDNA encoded the p180 subunit of eIF-3, rabbit eIF-3 was subjected to CNBr fragmentation. The fragments were resolved by SDS-gel electrophoresis, and amino-terminal sequences were determined. Three of the fragments produced relevant amino acid sequences (Table I). Where positive identifications were made, 56 out of 56 matched the amino acid sequence deduced from the cDNA. Five residues were consistent with, but not proof of, the residues observed in the deduced amino acid sequence. If peptides 1 and 2 are placed in the deduced sequence, methionine precedes the first amino acid of each peptide, consistent with the specificity of CNBr cleavage. One inferred difference was observed when the third peptide was placed in the deduced amino acid sequence. In the human sequence, the residue just before lysine 824 is leucine, not methionine. It is possible that amino acid 823 in rabbit p180 is methionine, but this has not been determined. These data confirm the identity of the cDNA clone as encoding the large subunit of human initiation factor 3.

When the composite human cDNA sequence was compared to the data bases, a mouse sequence called centrosomin was identified (Ref. 30; accession nos. X17373 and X84651 for centrosomin A and B, respectively). As discussed below, centrosomin A and B probably represent partial cDNA clones for the p180 subunit of mouse eIF-3. In addition to centrosomin, 30 human, 2 mouse, and 1 rat expressed sequence tags (EST) were identiﬁed that were almost identical to portions of the sequence presented in Fig. 4A. Sequences from the nematode C. elegans, the dicotyledogenous plant N. tabacum (tobacco), and the yeast S. cerevisiae also showed short stretches of homology with the human sequence. All of these homologies are discussed further below.

When the protein sequence derived from the human cDNA was used to search the protein data bases, three sequences in addition to the deduced sequence of centrosomin were identiﬁed. One was an ORF found in a cosmid prepared from S. cerevisiae.6 No introns were predicted to interrupt this open reading frame. Another ORF was found in a cosmid prepared

---

**Table I**

| Peptide sequence | Residue no. in human p180 | Peptide SDS |
|------------------|--------------------------|-------------|
| SSVLKALEVIXPAHILQEG | 527–546 | 17,000 |
| ELKARQSVVEKLFQEE | 763–781 | 28,000 |
| KEREERERAERAKREEEYQ | 824–845 | 18,000 |

---

6 Q. J. M. Van der Aart, unpublished results (GenBank accession no. X76294, ORF YBR0734; Swiss-Prot accession no. P39249).
from *C. elegans*.\(^7\) In this case, the computer algorithm used to identify putative introns and exons identified 6 introns. Support for the position of one of these predicted introns is found in the partial sequence of a *C. elegans* cDNA clone.\(^8\) The deduced amino acid sequence of the *C. elegans* protein presented in Fig. 6 is derived from the conceptual splicing of the exons.

\(^7\) Z. Du, unpublished results (GenBank accession no. L23650; Swiss-Prot accession no. P34539).

\(^8\) Y. Kuhara, unpublished results (GenBank accession no. D36664).
The length of each sequence used for the comparisons is given in parentheses below the name. Each figure in the table is the percent identity (percent similarity in parentheses) that was obtained when the sequence in the row was compared to the sequence in the column. The abbreviations are: Tob, *N. tabacum*; Cel, *C. elegans*; and ND, not done. The data were generated using the GCG package.

**TABLE II**

| Sequence (no. of residues) vs. | Human (1382) | Tob (858) | Cel (1076) |
|-----------------------------|-------------|-----------|------------|
| Centrosomin (1076)          |            | 94%       | ND         |
| *N. tabacum* (958)          | 73% (96%)   | ND        | ND         |
| *C. elegans* (1076)         | 35% (59%)   | ND        | ND         |
| *S. cerevisiae* (964)       | 36% (56%)   | 33% (54%) | ND         |
|                             | 29% (54%)   | 30% (53%) | 28% (50%)  |

**Fig. 7.** The repetitive domain of human eIF-3 p180. Shown are 248 residues from the deduced sequence of the large subunit of human eIF-3 spanning amino acids 925–1172. The sequence has been divided into 25 repeats of about 10 amino acids each; deletions and/or insertions have been introduced into 7 of the repeats in order to maximize the alignment of identical residues. The consensus line (con) gives the amino acid residue(s) most likely to be found in each position in the repetitive domain.

immunoprecipitated with several other proteins, suggesting it resided in a protein complex. The composition of the complex was similar to that reported for eIF-3. This identification was confirmed by cross-reactivity of the anti-fusion protein serum with the p180 subunit of rabbit eIF-3 and by aligning the sequence of peptides derived from purified rabbit eIF-3 with the p180 subunit of rabbit eIF-3 and by aligning the sequence reported in Fig. 4A; they also support the reading frame changes introduced into the centrosomin sequence corresponding to amino acids 421 and 468, as well as the inclusion of glycine 921 (see Fig. 4B). The other changes introduced into the centrosomin sequence are not covered by any of the ESTs.

The polypeptide deduced from the centrosomin nucleotide sequence is almost identical to the human sequence, making it likely that centrosomin represents a partial cDNA clone for mouse p180. The high degree of sequence identity between human, rabbit, and (what is likely) mouse p180 is characteristic of the mammalian protein synthesis translation factors, which often show greater than 98% identity (31).

The repetitive domain of the deduced amino acid sequence of p180 is shown in detail in Fig. 7. Of the 248 residues included in the domain, 144 (58%) are charged. All 68 of the basic residues are arginine; 64 of the 76 acidic residues are aspartate. The sequence deduced from the centrosomin clone also spans this region (Fig. 4B). However, it was necessary to introduce gaps into the centrosomin sequence in order to maximize the alignment of identical residues, as well as to maintain the spacing patterns of those repeats that are not exactly 10 residues in length. This suggests there is variability in the number of repeats in vertebrates. The sequences from more distant organisms do not contain a repetitive domain (Fig. 6). The existence of a repeating element in the p180 subunit is reminiscent of the 10 copies of the pseudo repeat DRYR in eIF-4B (32). Since this portion of eIF-4B interacts with the p180 subunit of eIF-3, the repeated elements in both proteins may be responsible for their interaction. However, it should be noted that the repeated elements are lacking in both yeast eIF-4B (33, 34) and the large subunit of yeast eIF-3 (Fig. 6).

The biochemical data available for mammalian and plant eIF-3 can be compared to characteristics inferred from the cDNA clones. The predicted molecular mass values roughly correspond to determinations by SDS-gel electrophoresis, although in both cases the SDS estimates are larger (166 versus 180 kDa for mammalian and 111 versus 116 kDa for plant). Although not as well characterized, it appears the SDS molecular mass of the large subunit of yeast eIF-3 (130 kDa; Refs. 6, 7, and 9) also exceeds that deduced from the nucleotide sequence (110 kDa). These discrepancies may reflect the high percentage of charged residues in the proteins. The isoelectric points calculated for the human (6.4) and tobacco (9.4) proteins are consistent with those determined for the rabbit (6.7) and wheat (greater than 8) proteins. The biochemical data reveal heterogeneity in the sizes of the subunits comprising eIF-3 and suggest that eIF-3 of varying compositions may be found when the factor is isolated from other organisms. The ORFs found in the tobacco and nematode sequences are consistent with het-
erogeneity in the size of the large subunit of eIF-3 (Fig. 6).

Fig. 6 shows a comparison between all four deduced protein sequences as well as the consensus; Table II shows the results of pairwise comparisons of the sequences. Compared to other translation factors, the identity between the yeast and human eIF-3 subunits is somewhat lower (31). In all four sequences, the NH₂-terminal regions show more similarity than the COOH-terminal regions. The dissimilarities become more pronounced when the repetitive domain of the human sequence is reached.

The data in Figs. 4 and 6 together with the protein studies on rabbit eIF-3 suggest that the mammalian p180 subunits are unusually large when compared to the subunits of other organisms. It has been reported that yeast eIF-3 will replace HeLa eIF-3 in the methionyl-puromycin assay (6), again suggesting a general conservation of the pathway for initiation of protein synthesis (35). However, this is surprising given the size difference, relatively low sequence identity, and lack of a repetitive domain in the yeast large subunit. It may be that the function(s) of the repetitive domain found in the mammalian subunit is provided by a different subunit in other organisms. In this context, it is noteworthy that wheat germ eIF-3 is reported to contain more subunits than mammalian eIF-3 (1). However, yeast eIF-3 appears to contain the same number of subunits, with each being 10–20% smaller than its mammalian counterpart (7, 8). The availability of genes for subunits of eIF-3 combined with the power of yeast genetics make it likely that, in the near future, more will be known of the function of the high molecular weight complex eIF-3.

Acknowledgments—We thank Rebecca Castle for technical assistance and Dr. Margaret Wheelock for helpful discussion and assistance throughout the project.

Addendum—As this manuscript was under review, unidentified sequences of a human cDNA clone (accession no. D50929, submitted by N. Nomura) and a partial human genomic clone (accession no. U58047, submitted by J. K. Scholler) were placed in the data bases. These sequences correspond to the large subunit of eIF-3. In addition, the unidentified sequence of a mouse cDNA clone (accession no. U14172, submitted by R. Fisher) corresponding to the large subunit of eIF-3 has been placed in the data bases.

REFERENCES
1. Merrick, W. C. (1992) Microbiol. Rev. 56, 291–315
2. Hershey, J. W. B. (1991) Annu. Rev. Biochem. 60, 717–755
3. Grifo, J. A., Tahara, S. M., Morgan, M. A., Shatkin, A. J., and Merrick, W. C. (1983) J. Biol. Chem. 258, 5804–5810
4. Gupta, N. K., Roy, A. L., Nag, M. K., Kinzy, T. G., MacMillan, S., Hileman, R. E., Dever, T. E., Wu, S., Merrick, W. C., and Hershey, J. W. B. (1990) in Post-transcriptional Control of Gene Expression (McCarthy, J. E. G., and Tuite, M. F., eds) pp. 521–526, Springer-Verlag, Berlin
5. Lambhear, B. J., Kirchwege, R., Skern, T., and Ihnoud, R. E. (1995) J. Biol. Chem. 270, 21975–21983
6. Naranda, T., MacMillan, S. E., and Hershey, J. W. B. (1994) J. Biol. Chem. 269, 32286–32292
7. Danais, P., Wittmer, B., Altmann, M., and Trachsel, H. (1995) J. Biol. Chem. 270, 4288–4292
8. Evans, D. R. H., Rasmussen, C., Hanic-Joyce, P. J., Johnston, G. C., Singer, R. A., and Barnes, C. A. (1985) Mol. Cell. Biol. 15, 4525–4535
9. Garcia-Barrionuevo, M. T., Naranda, T., Vazquez de Aldana, C. R., Cuesta, R., Hinnebusch, A. G., Hershey, J. W. B., and Tamame, M. (1995) Genes Dev. 9, 1781–1796
10. Bennett, R. H. (1979) Methods Enzymol. 58, 345–359
11. Gallagher, R., Collins, S., Trujillo, J., McCredie, K., Ahearn, M., Tsai, S., Metzgar, R., Aulakh, G., Ting, R., Ruscetti, R., and Gallo, R. (1979) Blood 54, 713–733
12. Wheelock, M. J., and Jensen, P. J. (1992) J. Cell Biol. 117, 415–425
13. Pattillo, R. A., Ruckert, A., Hassa, R., Bernstein, R., and Delis, E. (1971) In Vitro 6, 398a–399a
14. Giard, D. J., Aaronson, S. A., Todaro, G. J., Arnstein, P., Kersey, J. H., Dosik, H., and Parks, W. P. (1973) J. Natl. Cancer Inst. 51, 1417–1423
15. Wheelock, M. J., Komuniecki, R., Duran, E., and Johnson, K. R. (1991) Mol. Biochem. Parasitol. 45, 9–18
16. Haytn, T. V., Young, R. A., and Davis, R. W. (1985) in DNA Cloning: A Practical Approach (Glover, D. M., ed) Vol. 1, pp. 49–78, IRL Press, Oxford
17. Harlow, E., and Lane, D. (1988) Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
18. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, 2nd Ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
19. Zia, M. R., Giordano, A., Aramendia, E., and Ferrone, S. (1988) Anal. Biochem. 171, 192–196
20. Yanisch-Perron, C., Vieira, J., and Messing, J. (1985) Gene (Amst.) 33, 103–119
21. Johnson, K. R. (1990) Anal. Biochem. 180, 170–174
22. Bairoch, A., Bucher, P., and Hofmann, K. (1996) Nucleic Acids Res. 24, 189–96
23. Laemmli, U. K. (1970) Nature 227, 680–685
24. Wheelock, M. J., Buck, C. A., Bechtol, K. B., and Damsky, C. H. (1987) J. Cell. Biochem. 34, 187–202
25. Johnson, K. R., Lewis, J. E., Li, D., Wahl, J., Peralta Soler, A., Knudsen, K. A., and Wheelock, M. J. (1993) Exp. Cell Res. 207, 252–260
26. Kozak, M. (1990) Annu. Rev. Cell Biol. 6, 197–225
27. Shaw, G., and Kamen, R. (1986) Cell 46, 659–667
28. Zubiaga, A. M., Belasco, J. G., and Greenberg, M. E. (1995) Mol. Cell. Biol. 15, 2219–2230
29. Proudfit, N. J., and Brownlee, G. G. (1976) Nature 263, 211–214
30. Joswig, G., Pettetz, C., and Werner, D. (1991) J. Cell Sci. 98, 37–43
31. Merrick, W. C., Dever, T. E., Kinzy, T. G., Conroy, S. C., Cavallius, J., and Owens, C. L. (1990) Biochem. Biophys. Acta 1050, 235–240
32. Milburn, S. C., Hershey, J. W. B., Davies, M. V., Kelleher, K., and Kaufman, R. (1990) EMBO J. 9, 2783–2790
33. Coppolecchia, R., Buser, P., Stutz, A., and Linder, P. (1993) EMBO J. 12, 4005–4011
34. Altmann, M., Wittmer, B., Methot, N., Sonenberg, N., and Trachsel, H. (1995) EMBO J. 14, 3820–3827
35. Hannig, E. M. (1995) BioEssays 17, 915–919
36. Zhu, Y., and Johnson, K. R. (1992) Mol. Cell. Biol. 3, 322a (abstr.)